FINITE SAMPLE TESTS FOR ARCH EFFECTS AND VARIANCE CHANGE-POINTS IN LINEAR REGRESSIONS

Jean-Marie Dufour [a], Lynda Khalaf [b], Jean-Thomas Bernard [c] and Ian Genest [d]

[a] CIRANO, CIREQ, and Département de sciences économiques, Université de Montréal,
C.P. 6128, succursale Centre Ville, Montréal, Québec H3C 3J7

[b] CIREQ and GREEN, Université Laval. Pavillon J.-A.-De Sève, Québec, Canada

[c] GREEN, Université Laval, Pavillon J.-A.-De Sève, Québec, Canada, G1K 7P4

[d] Hydro-Québec and GREEN, Université Laval, Pavillon J.-A.-De Sève, Québec, Canada, G1K 7P4

Key words: heteroskedasticity; homoskedasticity; regression; Monte Carlo test; exact test; specification test; ARCH; GARCH; structural stability.

1. Introduction

A wide range of tests for heteroskedasticity have been proposed in econometrics and statistics.[1] Although a few exact tests are available (e.g. Goldfeld-Quandt's F-test, its extensions and Szroeter's procedures),[2] common heteroskedasticity tests are asymptotic which may not control size in finite samples. So a number of recent studies have tried to improve the reliability of these tests using Edgeworth, Bartlett, jackknife and bootstrap methods; for references, see Dufour et al. (2001). Yet the latter remain approximate, while Szroeter's exact tests require computing the distributions of general quadratic forms in normal variables and are seldom used.

In this paper, we describe a general solution to the problem of controlling the size of homoskedasticity tests in linear regressions. We exploit the technique of Monte Carlo (MC) tests [Dwass (1957), Barnard (1963), Dufour and Kiviet (1996, 1998)] to obtain provably exact randomized analogues of the tests considered. This simulation-based procedure yields an exact test when the distribution of the test statistic is *pivotal* under the null hypothesis: all we need is the possibility of simulating the relevant test statistic under the null hypothesis. In particular, this covers many cases where the finite-sample distribution of the test statistic is intractable or involves parameters which are unidentified under the null hypothesis (see Andrews and Ploberger (1995), Hansen (1996) and Andrews (2001)). Further the method allows one to consider any error distribution that can be simulated.

This paper makes five contributions. First, we show that all the standard homoskedasticity test statistics considered are pivotal in finite samples, hence allowing the construction of finite-sample MC versions of these. Second, we extend the tests for which a finite-sample theory has been supplied for Gaussian distributions (e.g. the Goldfeld-Quandt and Szroeter tests), to allow for non-Gaussian distributions. Third, our results cover ARCH-type alternatives. In the ARCH-M case, the MC procedure circumvents the unidentified nuisance parameter problem. Fourth, we define a number of new test statistics which include combined Breusch-Pagan-Godfrey and Goldfeld-Quandt tests against variance breaks at unknown points, based on the minimum or the product of individual p-values. Finally, we report a simulation study on ARCH and change point tests.

2. Framework

We consider the linear model:

$$y_t = x_t'\beta + u_t , \tag{2.1}$$

$$u_t = \sigma_t \varepsilon_t , \; t = 1, \ldots, T, \tag{2.2}$$

where $x_t = (x_{t1}, x_{t2}, \ldots, x_{tk})'$, $X \equiv [x_1, \ldots, x_T]'$ is a full-column rank $T \times k$ matrix, $\beta = (\beta_1, \ldots, \beta_k)'$ is a $k \times 1$ vector of unknown coefficients, $\sigma_1, \ldots, \sigma_T$ are

[1] In linear regressions, popular procedures include the Goldfeld-Quandt F-test [Goldfeld and Quandt (1965)], the Breusch-Pagan-Godfrey LM test [Godfrey (1978), Breusch and Pagan (1979), Koenker (1981)], White's test [White (1980)]; see Dufour, Khalaf, Bernard and Genest (2001). In time series and finance, modeling variances is viewed as an important aspect of data analysis, leading to the popular ARCH-type models; see Engle (1982), Lee and King (1993), Bera and Ra (1995) and Hong and Shehadeh (1999).

[2] See Szroeter (1978), Harrison and McCabe (1979), Harrison (1980, 1981, 1982), King (1981) and Evans and King (1985).

(possibly random) scale parameters, and

$$\varepsilon \text{ is a random vector with a completely specified continuous distribution conditional on } X\,, \tag{2.3}$$

where $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_T)'$. We wish to test

$$H_0 : \sigma_t^2 = \sigma^2,\ t = 1, \ldots, T,\ \text{for some } \sigma, \tag{2.4}$$

against the alternative $H_A : \sigma_t^2 \neq \sigma_s^2$, for at least one t and s. In most cases of practical interest, one would further restrict the distribution of $\boldsymbol{\varepsilon}$, for example by assuming

$$\varepsilon_1, \ldots, \varepsilon_T \text{ are i.i.d. according to } F_0\,, \tag{2.5}$$

where F_0 is given, so that $u_1, \ldots, u_T$ are i.i.d. with distribution $\mathsf{P}[u_t \leq v] = F_0(v/\sigma)$ under H_0; for example:

$$\varepsilon_1, \ldots, \varepsilon_T \overset{i.i.d.}{\sim} N[0, 1]\,. \tag{2.6}$$

However, normality is not needed for several of our results. We focus on the following alternatives (H_A):

H_1 : GARCH and ARCH-M;

H_2 : σ_t^2 increases monotonically with an exogenous variable;

H_3 : $\sigma_t^2$is constant within and differs across p subsamples.

H_3 includes discrete changes in variance at some (specified) point in time; we also propose exact tests for unknown break points. Let $\widehat{\beta} = (X'X)^{-1}X'y$ and

$$\widehat{\sigma}^2 = \widehat{u}'\widehat{u}/T\,,\ \ \widehat{u} = (\widehat{u}_1, \ldots, \widehat{u}_T)' = y - X\widehat{\beta}\,. \tag{2.7}$$

3. Test statistics

3.1. Tests based on auxiliary regressions

3.1.1. Standard auxiliary regression tests

Consider the auxiliary regressions:

$$\begin{aligned}
\widehat{u}_t^2 &= z_t'\alpha + w_t,\ t = 1, \ldots, T, && (3.8)\\
\widehat{u}_t^2 - \widehat{\sigma}^2 &= z_t'\alpha + w_t,\ t = 1, \ldots, T, && (3.9)\\
|\widehat{u}_t| &= z_t'\alpha + w_t,\ t = 1, \ldots, T, && (3.10)
\end{aligned}$$

where $z_t = (1, z_{t2}, \ldots, z_{tm})'$ is a vector of m fixed regressors on which σ_t may depend, $\alpha = (\alpha_1, \ldots, \alpha_m)'$ and $w_t,\ t = 1, \ldots, T$, are treated as error terms. The Breusch-Pagan-Godfrey (BPG) LM criterion [Breusch and Pagan (1979), Godfrey (1978)] may be obtained as the explained sum of squares (ESS) from (3.8) divided by $2\widehat{\sigma}^4$. The Koenker (K) test statistic [Koenker (1981)] is T times the uncentered R^2 from (3.9). White's (W) test statistic is T times the uncentered R^2 from (3.8) using for z_t the $r \times 1$ the non redundant variables in $x_t \otimes x_t$. The F statistic to test $\alpha_2 = \ldots = \alpha_m = 0$ in (3.10) yields Glejser's test (G) [Glejser (1969)].

3.1.2. Auxiliary regression tests against an unknown variance breakpoint

Tests against discrete breaks in variance at some specified date τ may be applied in the above framework by defining z_t as a dummy variable of the form $z_t = z_t(\tau)$,

$$z_t(\tau) = \begin{cases} 0, & \text{if } t \leq \tau\,, \\ 1, & \text{if } t > \tau\,. \end{cases} \tag{3.11}$$

If τ is unspecified, a different test statistic may be computed for each one of the possible break-dates $\tau = 1, \ldots, T-1$. Here we provide a procedure to combine inference based on the resulting multiple tests. Let BPG_τ be the BPG statistic obtained on using $z_t = z_t(\tau)$, where $\tau = 1, \ldots, T-1$. If used as a single test, the BPG_τ statistic is significant at level α when $BPG_\tau \geq \chi^2_\alpha(1)$, or (equivalently) when $\mathsf{G}_{\chi_1}(BPG_\tau) \leq \alpha$, where $\mathsf{G}_{\chi_1}(BPG_\tau)$ is the asymptotic p-value associated with BPG_τ. We propose here two combination methods. The first rejects H_0 when at least one of the p-values for $\tau \in J$ is sufficiently small, where J is some appropriate subset of the time interval $\{1, 2, \ldots, T-1\}$, such as $J = [\tau_1, \tau_2]$ where $1 \leq \tau_1 < \tau_2 \leq T-1$. More precisely, we reject H_0 at level α when

$$pv_{\min}(BPG;\, J) \leq p_0(\alpha;\, J)\,, \tag{3.12}$$

where $pv_{\min}(BPG;\, J) \equiv \min\{\mathsf{G}_{\chi_1}(BPG_\tau) : \tau \in J\}$ or, equivalently, when

$$F_{\min}(BPG;\, J) \geq F_{\min}(\alpha;\, J)\,, \tag{3.13}$$

where $F_{\min}(BPG;\, J) \equiv 1 - \min\{\mathsf{G}_{\chi_1}(BPG_\tau) : \tau \in J\}$; $p_0(\alpha;\, J)$ is the largest point such that $\mathsf{P}[pv_{\min}(BPG;\, J) \leq p_0(\alpha;\, J)] \leq \alpha$ under H_0, and $F_{\min}(\alpha;\, J) = 1 - p_0(\alpha;\, J)$. This method was suggested by Tippett (1931) and Wilkinson (1951) in the case of independent statistics; of course the variables BPG_τ, $\tau \in J$, are not independent.

The second method consists in rejecting H_0 when the product of the p-values $pv_\times(BPG;\, J) \equiv \prod_{\tau \in J} \mathsf{G}_\chi(BPG_\tau)$ is small, or when

$$F_\times(BPG;\, J) \geq \overline{F}_\times(J;\, \alpha)\,, \tag{3.14}$$

where $F_\times(BPG;\, J) \equiv 1 - pv_\times(BPG;\, J)$; $\overline{F}_\times(J; \alpha)$

is the largest point such that $\mathsf{P}[F_{\times}(BPG;\,J) \geq \overline{F}_{\times}(J;\alpha)] \leq \alpha$ under H_0. This method was originally suggested by Fisher (1932) and Pearson (1933) for independent test statistics.[3] We also propose a modified version of $F_{\times}(BPG;\,J)$ based on a subset of the p-values. The latter is a variant of $F_{\times}(BPG;\,J)$ based on the m smallest p-values:

$$F_{\times}(BPG;\,\widehat{J}_{(4)}) = 1 - \prod_{\tau \in \widehat{J}_{(m)}} \mathsf{G}_{\chi_1}(BPG_{\tau})\;, \quad (3.15)$$

where $\widehat{J}_{(m)}$ is the set of the m smallest p-values in the series $\{\mathsf{G}_{\chi_1}(BPG_{\tau}) : \tau = 1,\, 2,\, \ldots,\, T-1\}$. The maximal number of p-values retained (m in this case) may be chosen to reflect (prior) knowledge on potential breaks; see Christiano (1992).

3.2. ARCH-type tests

Engle (1982)'s ARCH test statistic based on (2.1) where $\varepsilon_t|_{t-1} \sim N(0,1)$, is $E = TR^2$, where R^2 is the coefficient of determination in the regression of $\widehat{u}_t^2$ on a constant and $\widehat{u}_{t-i}^2$ $(i = 1,\, \ldots\,,\, q)$. Under standard conditions $E \overset{asy}{\backsim} \chi^2(q)$. Lee (1991) has shown that the same test is appropriate against GARCH(p, q) alternatives

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \theta_i \sigma_{t-i}^2 + \sum_{i=1}^{q} \alpha_i\, \sigma_{t-i}^2 \varepsilon_{t-i}^2\;, \quad (3.16)$$

with $H_0 : \alpha_1 = \cdots = \alpha_q = \theta_1 = \cdots = \theta_p = 0$. Lee and King (1993) proposed an alternative (asymptotically standard normal) GARCH test statistic which exploits the one-sided nature of H_A:

$$LK = \frac{\left\{T_q \sum\limits_{t=q+1}^{T} \left[\frac{\widehat{u}_t^2}{\widehat{\sigma}^2} - 1\right] S_q\right\} / \left\{\sum\limits_{t=q+1}^{T} \left[\frac{\widehat{u}_t^2}{\widehat{\sigma}^2} - 1\right]^2\right\}^{1/2}}{\left\{T_q \sum\limits_{t=q+1} S_q^2 - \left(\sum\limits_{t=q+1}^{T} S_q\right)^2\right\}^{1/2}}\,. \quad (3.17)$$

where $T_q = T - q$ and $S_q = \sum_{i=1}^{q} \widehat{u}_{(t-i)}^2$. In the context of

$$y_t = x_t'\beta + \sigma_t\phi + u_t\;,\; t = 1,\, \ldots\,,\, T\,, \quad (3.18)$$

the LM ARCH-M for given ϕ is [see Bera and Ra (1995)]:

$$LM(\phi) = \widehat{\gamma}'V\left[V'V - \frac{\phi^2}{2+\phi^2}V_X\right]^{-1} V'\widehat{\gamma}/(2+\phi^2)\;, \quad (3.19)$$

where $V_X = V'X(X'X)^{-1}X'V$, $\widehat{\gamma}$ is a $T \times 1$ vector with elements $\widehat{\gamma}_t = [(\widehat{u}_t/\widehat{\sigma})^2 - 1] + \phi\widehat{u}_t/\widehat{\sigma}$ and V is a $T \times (q+1)$ matrix whose t-th row is $V_t = (1\,,\; \widehat{u}_{t-1}^2\,,\, \ldots\,,\; \widehat{u}_{t-q}^2)$. In this case, under H_0, ϕ is unidentified. Bera and Ra (1995) also show that the Davies sup-LM test applied to this problem leads to more reliable inference.

[3] For further discussion of methods for combining tests, see Dufour and Torrès (1998, 2000), and Dufour and Khalaf (2002*a*).

3.3. Tests based on grouping

An alternative class of tests assumes that observations can be ordered according to non-decreasing variance. Let $\widehat{u}_{(t)}$, $t = 1,\, \ldots\,,\, T$, denote reordered residuals.

3.3.1. Goldfeld-Quandt tests against an unknown variance breakpoint

The most familiar test in this class is the Goldfeld and Quandt (1965, GQ) test which involves separating the ordered sample into three subsets and computing OLS regressions on the first and last subsets. Let T_i, $i = 1,\, 2,\, 3$, denote the number of observations in each of these subsets ($T = T_1 + T_2 + T_3$). The test statistic, which follows a $F(T_3 - k, T_1 - k)$ distribution under H_0, is

$$GQ(T_1, T_3, k) = \frac{S_3/(T_3 - k)}{S_1/(T_1 - k)} \quad (3.20)$$

where S_1 and S_3 are the sum of squared residuals from the first T_1 and the last T_3 observations. Setting $\mathsf{G}_{F(T_3-k,T_1-k)}(x) = \mathsf{P}[F(T_3 - k, T_1 - k) \geq x]$, we denote $pv[GQ;\, T_1, T_3, k] = \mathsf{G}_{F(T_3-k,T_1-k)}[GQ(T_1, T_3, k)]$ the p-value associated with $GQ(T_1, T_3, k)$. To account for an unknown breakdate, we propose statistics of the form:

$$F_{\min}(GQ;\, K) \equiv 1 - \min_{(T_1,T_3)\in K} \{pv[GQ;\, T_1, T_3, k]\},\quad (3.21)$$

$$F_{\times}(GQ;\, K) \equiv 1 - \prod_{(T_1,T_3)\in K} pv[GQ;\, T_1, T_3, k]\,,\quad (3.22)$$

where K is any appropriate non-empty subset of

$$K(k,\, T) = \quad \{(T_1, T_3) \in \mathbb{Z}^2 : k+1 \leq T_1 \leq T-k-1 \text{ and } k+1 \leq T_3 \leq T - T_1\}\,,$$

the set of the possible subsample sizes compatible with the definition of the GQ. Reasonable choices could be $K = S_1(T,\, T_2,\, L_0,\, U_0)$ with

$$S_1(.) \equiv \{(T_1, T_3) : L_0 \leq T_1 \leq U_0, T_3 = T - T_1 - T_2 \geq 0\}\,, \quad (3.23)$$

where T_2 is the number of central observations while L_0 and U_0 are minimal and maximal sizes for the subsamples $(0 \leq T_2 \leq T-2k-2,\; L_0 \geq k+1,\; U_0 \leq T-T_2-k-1)$, or $K = S_2(T,\, L_0,\, U_0)$ with

$$S_2(.) = \{(T_1, T_3) : L_0 \leq T_1 = T_3 \leq U_0\} \quad (3.24)$$

where $L_0 \geq k+1$ and $U_0 \leq I[T/2]$; $I[x]$ is the largest integer less than or equal to x. According to (3.23), $\{GQ(T_1, T_3, k) : (T_1, T_3) \in K\}$ defines a set of GQ statistics, such that the number T_2 of central observations is kept constant; with (3.24), $\{GQ(T_1, T_3, k) : (T_1, T_3) \in K\}$ leads to GQ statistics such that $T_1 = T_3$. We also consider

$$F_{\times}(GQ;\, \widehat{K}_{(4)}) \equiv 1 - \prod_{(T_1,T_3)\in \widehat{K}_{(4)}} pv[GQ;\, T_1, T_3, k]$$

where $\widehat{K}_{(4)}$ selects the four smallest p-values from the set $\{pv[GQ;\, T_1, T_3, k] : (T_1, T_3) \in K\}$.

3.3.2. Szroeter-type tests

From Szroeter (1978) class of tests, we consider:

$$SKH = \sum_{t=1}^{T} 2\left[1 - \cos\left(\frac{\pi t}{t+1}\right)\right] \widehat{u}_{(t)}^2 / S_z \ , \tag{3.25}$$

$$S_N = 6T/(T^2-1)\left(\sum_{t=1}^{T} t\,\widehat{u}_{(t)}^2/S_z - \frac{(T+1)/2}{2}\right) \ , \tag{3.26}$$

$$S_F(T_1,\, T-T_1-T_2) = \sum_{t=T_1+T_2+1}^{T} \widehat{u}_{(t)}^2/S_z \ . \tag{3.27}$$

where $S_z = \sum_{t=1}^{T_1} \widehat{u}_{(t)}^2$. Under the null hypothesis, S_N follows a $N(0,1)$ distribution asymptotically. Exact critical points for SKH [under (2.6)] may be obtained using the Imhof method. King (1981) provided bounds for use with SKH; Harrison (1980, 1981, 1982) however showed there is a high probability that these bounds be inconclusive.

4. Finite-sample theory

In Dufour et al. (2001), we show that all the statistics described above have null distributions which are free of nuisance parameters. Here we show how this fact can be used to perform a finite-sample MC test of homoskedasticity using any one of these statistics.

We can proceed as follows based on any statistic, say $S_0 = S(y,\, X)$, whose null distribution (given X) is continuous and free of nuisance parameters. Let $G(x)$ be the survival function associated with S_0 under H_0 and consider a right-tailed continuous procedure. Generate N i.i.d. replications of the error vector ε according to (2.3), which leads to N realizations of the test statistic $S_1, \ldots, S_N$. The associated MC critical region is

$$\widehat{p}_N(S_0) \le \alpha \ , \tag{4.28}$$

where $\widehat{p}_N(x) = (N\widehat{G}_N(x) + 1)/(N+1)$ and $\widehat{G}_N(x)$ is the number of simulated statistics $\geq x$ and

$$P_{H_0}\left[\widehat{p}_N(S_0) \le \alpha\right] = \frac{I\left[\alpha(N+1)\right]}{N+1} \ . \tag{4.29}$$

Table 1. Tests against ARCH and GARCH heteroskedasticity:
normal error distribution and D1 design

	$T = 100$; 99 MC reps.			
	E		LK	
$(\phi, \alpha_1, \theta_1)$	ASY	MC	ASY	MC
$(0,0,0)$	3.63	5.18	4.72	5.22
$(0,.1,0)$	11.83	13.61	17.01	**17.22**
$(0,.5,0)$	64.18	66.43	71.93	**71.54**
$(0,.9,0)$	84.38	85.82	88.99	**88.97**
$(0,1,0)$	86.40	87.50	90.60	**90.24**
$(0,.1,.5)$	12.54	14.39	17.89	**18.35**
$(0,.25,.65)$	38.36	40.74	46.93	**47.29**
$(0,.40,.50)$	57.44	59.65	65.94	**65.69**
$(0,.15,.85)$	28.97	31.04	37.11	**37.93**
$(0,.05,.95)$	18.17	20.15	25.92	**26.28**

If $\alpha(N+1)$ is an integer, $\mathsf{P}_{H_0}\left[\widehat{p}_N(S_0) \le \alpha\right] = \alpha$; see Dufour and Kiviet (1998). Thus (4.28) has the same size as the critical region $G(S_0) \le \alpha$, irrespective of N. For applications, see Dufour and Kiviet (1996, 1998), Kiviet and Dufour (1997), Dufour and Khalaf (2001*a*, 2001*b*, 2002*a*, 2002*b*).

5. Simulation experiments [4]

The first experiment was based on (3.16), (3.18), with $q = 1$, $T = 100$, $k = I[T^{1/2}] + 1$, $\beta = (1,\, 1,\, \ldots,\, 1)'$ and $\alpha_0 = 1$. Four experiments were conducted with parameters set as follows: (i) $\phi = \alpha_1 = \theta_1 = 0$; (ii) $\phi = \theta_1 = 0$, $\alpha_1 = .1,\ .5,\ .9,\ 1$; (iii) $\phi = 0$, $(\alpha_1, \theta_1) = (.1, .5)$, $(.25, .65)$, $(.4, .5)$, $(.15, .85)$, $(.05, .95)$; (iv) $\theta_1 = 0$, $\phi = -2$, $\alpha_1 = 0,\ .1,\ .9$. Following Godfrey (1996), the error distributions considered are: $N(0,1)$, $t(5)$ and $Cauchy$ [denoted ε_t^n, $\varepsilon_t^{t(5)}$, ε_t^c]. The regressors were generated as i.i.d. according to a $U(0,10)$ distribution [D1 design]. In the case of (iv), we also considered an alternative regressor set, obtained by drawing (independently) from a Cauchy distribution (centered and re-scaled conformably with the previous design) [D2 design]. D1 and D2 include a constant.

Tables 1 and 2 report rejection percentages for a nominal level of 5%; 10000 replications were considered for (i) - (ii) and 1000 for (iv).[5] The most notable observation is that the Engle test is undersized; the size of the Lee-King test is better than that of the Engle test but is still below the nominal level. Although undersize problems

[4] For more complete results and a study of various other alternatives, see Dufour et al. (2001).

[5] In these tables and in later ones, the figures associated with best performing exact procedures in terms of power are set in **bold**.

Table 2. Size and Power of MC ARCH-M tests: various error distributions, D2 design; $\phi = -2$

$T = 100$; 99 MC reps						
		E		LK		BR
ε_t	α_1	ASY	MC	ASY	MC	MC
ε_t^n	0	3.9	5.3	5.4	5.1	6.2
$\varepsilon_t^{t(5)}$		2.8	4.8	3.4	5.0	4.5
ε_t^c		1.8	5.2	2.3	5.2	5.3
ε_t^n	.1	17.4	19.8	24.8	23.3	**25.0**
$\varepsilon_t^{t(5)}$		53.2	71.1	61.4	72.6	**79.0**
ε_t^c		66.1	81.7	71.3	83.9	**86.7**
ε_t^n	.9	61.7	64.2	68.5	65.9	**82.1**
$\varepsilon_t^{t(5)}$		45.8	62.3	54.2	63.5	**73.1**
ε_t^c		71.6	85.2	76.5	87.5	**90.0**

are evident under D1 and normal errors, more serious size distortions are observed with non normal errors.

MC tests yield noticeable effective power gains, even with uniform designs and normal errors; improvements in power are quite substantial in non-normal contexts. The Lee-King MC test is always more powerful than the Engle test. For ARCH-M alternatives, there is a substantial power gain from using the sup-LM MC test.

We next considered model (2.1) with: $T = 50$ and $k = 6$ and $\sigma_t^2 = \sigma_1$, if $t \leq \tau_0$, and $\sigma_t^2 = \sigma_1 + \delta$, if $t > \tau_0$, where $\delta \geq 0$ and τ_0 is the break time (assumed unknown). The regressors were generated as $U(0, 10)$ and the regression coefficients were set to one. Furthermore $\alpha_0 = 1$, and δ and τ_0 were set so that: $(\sigma_1 + \delta)/\sigma_1 = 1,\ 4,\ 16$, and $\tau_0/T = .3,\ .5\ ,\ .7$. We applied the MC versions of the standard tests GQ and BPG (using artificial regressions on $z_t = t,\ 1 \leq t \leq T$), S_F, SKH, and S_N tests, as well as the proposed combined tests $F_{\min}(GQ;\ K)$, $F_{\min}(BPG;\ J)$, $F_{\times}(BPG;\ \widehat{J}_{(m)})$, $F_{\times}(GQ; \widehat{K}_{(m)})$ with $m = 4$. For each one of the combined tests, we considered two possible "windows" $(J,\ K)$. The first one is a relatively uninformative "wide" window: $J^A = \{1, \ldots, T-1\}$, $K^A = S_1(T, T_2, k+1, T-T_2-k-1)$, with $T_2 = [T/5]$. The second set of windows were based on a predetermined interval around the true break-date, namely we considered: $J^S = \{L_0, L_0+1, \ldots, U_0\}$, $K^S = S_1(T, T_2, \tau_0^L(k), \tau_0^L(k))$, where $T_2 = [T/5]$, $\tau_0^L(k) = \max\{k+1, \tau_0 - I[T/5]\}$, $\tau_0^U(k) = \min\{T-k-T_2, \tau_0 + I[T/5]\}$. This yields the statistics $F_{\times}(BPG;\ \widehat{J}^i_{(4)})$, $F_{\min}(BPG;\ J^i)$, $F_{\times}(GQ; \widehat{K}^i_{(4)})$, $F_{\min}(GQ;\ K^i)$, $i = A, S$. The results are reported in Table 3.

The MC versions of all the tests achieve perfect size control. A remarkable finding here is the good performance of the Szroeter-type MC tests, which outperform commonly used tests such as the BPG and the GQ tests. For $\tau_0/T = 0.3$, the S_F test has the best power. The combined criteria perform well, and in several cases exhibit the best performance. Among these tests, product-type combined criteria perform better than min-type. The combined GQ criteria clearly dominate the standard GQ; the same holds for the BPG-based tests, if the search window is not uninformative. Power increases quite notably, where we consider the sup-tests maximized over the more informative window.

Table 3. Break in variance at unknown points

$T = 50$	H_0	$\tau_0/T = .3$		$\tau_0/T = .7$	
σ_2/σ_1	1	2	4	2	4
$GQ\,(F)$	5.5	35.0	59.9	57.1	99.3
BPG (MC)	5.8	33.7	59.1	73.2	99.6
S_F (MC)	7.2	47.0	70.3	73.3	99.4
S_N (MC)	6.2	45.5	70.8	80.4	99.7
SKH (MC)	6.8	46.7	71.9	81.1	99.8
MC Tests, maximized over the whole sample					
$F_{\times}(BPG;\ \widehat{J}^S_{(4)})$	5.3	16.1	26.1	72.5	99.5
$F_{\min}(BPG;\ \widehat{J}^S_{(4)})$	5.5	12.7	18.0	62.6	98.7
$F_{\times}(GQ;\ \widehat{J}^S_{(4)})$	5.6	56.8	98.6	73.3	99.7
$F_{\min}(GQ;\ \widehat{J}^S_{(4)})$	6.0	50.0	98.2	67.2	99.4
MC Tests maximized over a sub-sample					
$F_{\times}(BPG;\ \widehat{J}^S_{(4)})$	6.0	38.0	79.8	81.7	99.9
$F_{\min}(BPG;\ \widehat{J}^S_{(4)})$	5.8	37.1	77.8	79.4	99.9
$F_{\times}(GQ;\ \widehat{J}^S_{(4)})$	5.4	60.7	98.0	74.9	99.7
$F_{\min}(GQ;\ \widehat{J}^S_{(4)})$	6.2	53.1	98.3	75.0	99.5

6. Conclusion

We have described how finite-sample homoskedasticity tests can be obtained for a regression model with a specified error distribution. The latter exploit the MC test procedure which yields simulation-based exact randomized p-values irrespective of the number of replications used. The tests considered include tests for GARCH-type heteroskedasticity and sup-type tests against breaks in variance at unknown points. On observing that all test criteria are pivotal, the problem of "robustness to estimation effects" emphasized in Godfrey (1996) becomes irrelevant from our viewpoint. The general approach we used to obtain exact tests is not limited to normal errors. In particular, the method proposed allows one to consider possibly heavy-tailed error distributions, for which standard asymptotic theory would not apply.

References

Andrews, D. W. K. (2001), 'Testing when a parameter is on the boundary of the maintained hypothesis', *Econometrica* **69**, 683–733.

Andrews, D. W. K. and Ploberger, W. (1995), 'Admissibility of the likelihood ratio test when a parameter is present only under the alternative', *The Annals of Statistics* **23**, 1383–1414.

Barnard, G. A. (1963), 'Comment on 'The spectral analysis of point processes' by M. S. Bartlett', *Journal of the Royal Statistical Society, Series B* **25**, 294.

Bera, A. K. and Ra, S. (1995), 'A test for the presence of conditional heteroscedasticity within ARCH-M framework', *Econometric Reviews* **14**, 473–485.

Breusch, T. S. and Pagan, A. R. (1979), 'A simple test for heteroscedasticity and random coefficient variation', *Econometrica* **47**, 1287–1294.

Christiano, L. J. (1992), 'Searching for a break in GNP', *Journal of Business and Economic Statistics* **10**, 237–249.

Dufour, J.-M. and Khalaf, L. (2001*a*), Finite sample tests in seemingly unrelated regressions, *in* D. E. A. Giles, ed., 'Computer-Aided Econometrics', Marcel Dekker, New York. Forthcoming.

Dufour, J.-M. and Khalaf, L. (2001*b*), Monte Carlo test methods in econometrics, *in* B. Baltagi, ed., 'Companion to Theoretical Econometrics', Blackwell Companions to Contemporary Economics, Basil Blackwell, Oxford, U.K., chapter 23, pp. 494–519.

Dufour, J.-M. and Khalaf, L. (2002*a*), 'Exact tests for contemporaneous correlation of disturbances in seemingly unrelated regressions', *Journal of Econometrics* **106**(1), 143–170.

Dufour, J.-M. and Khalaf, L. (2002*b*), 'Simulation based finite and large sample tests in multivariate regressions', *Journal of Econometrics* **forthcoming**.

Dufour, J.-M., Khalaf, L., Bernard, J.-T. and Genest, I. (2001), Simulation-based finite-sample tests for heteroskedasticity and ARCH effects, Technical report, C.R.D.E., Université de Montréal, and GREEN, Université Laval. 48 pages.

Dufour, J.-M. and Kiviet, J. F. (1996), 'Exact tests for structural change in first-order dynamic models', *Journal of Econometrics* **70**, 39–68.

Dufour, J.-M. and Kiviet, J. F. (1998), 'Exact inference methods for first-order autoregressive distributed lag models', *Econometrica* **66**, 79–104.

Dufour, J.-M. and Torrès, O. (1998), Union-intersection and sample-split methods in econometrics with applications to SURE and MA models, *in* D. E. A. Giles and A. Ullah, eds, 'Handbook of Applied Economic Statistics', Marcel Dekker, New York, pp. 465–505.

Dufour, J.-M. and Torrès, O. (2000), 'Markovian processes, two-sided autoregressions and exact inference for stationary and nonstationary autoregressive processes', *Journal of Econometrics* **99**, 255–289.

Dwass, M. (1957), 'Modified randomization tests for nonparametric hypotheses', *Annals of Mathematical Statistics* **28**, 181–187.

Engle, R. F. (1982), 'Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation', *Econometrica* **50**, 987–1007.

Evans, M. A. and King, M. L. (1985), 'Critical value approximations for tests of linear regression disturbances', *Australian Journal of Statistics* pp. 68–83. 27.

Fisher, R. A. (1932), *Statistical Methods for Research Workers*, Oliver and Boyd, Edinburgh.

Glejser, H. (1969), 'A new test for heteroscedasticity', *Journal of the American Statistical Association* **64**, 316–323.

Godfrey, L. G. (1978), 'Testing for multiplicative heteroscedasticity,', *Journal of Econometrics* **8**, 227–236.

Godfrey, L. G. (1996), 'Some results on the Glejser and Koenker tests of heteroscedasticity', *Journal of Econometrics* **72**, 275–299.

Goldfeld, S. M. and Quandt, R. (1965), 'Some tests for heteroscedasticity', *Journal of the American Statistical Association* **60**, 539–547.

Hansen, B. E. (1996), 'Inference when a nuisance parameter is not identified under the null hypothesis', *Econometrica* **64**, 413–430.

Harrison, M. J. (1980), 'The small sample performance of the Szroeter bounds test for heteroscedasticity and a simple test for use when Szroeter's test is inconclusive', *Oxford Bulletin of Economics and Statistics* **42**, 235–250.

Harrison, M. J. (1981), 'A comparison of the bounds, beta-approximate, and exact variants of two tests for heteroscedasticity based on ordinary least squares residuals', *The Economic and Social Review* **12**, 235–252.

Harrison, M. J. (1982), 'Table of critical values for a beta approximation to szroeter's statistic for testing heteroscedasticity', *Oxford Bulletin of Economics and Statistics* **44**, 159–166.

Harrison, M. J. and McCabe, B. P. (1979), 'A test for heteroscedasticity based on ordinary least squares residuals', *Journal of the American Statistical Association* **74**, 494–499.

Hong, Y. and Shehadeh, R. D. (1999), 'A new test for ARCH effects and its finite-sample performance', *Journal of Business and Economic Statistics* **17**, 91–108.

King, M. L. (1981), 'A note on Szroeter's bound test', *Oxford Bulletin of Economics and Statistics* pp. 315–326. 43.

Kiviet, J. F. and Dufour, J.-M. (1997), 'Exact tests in single equation autoregressive distributed lag models', *Journal of Econometrics* **80**, 325–353.

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

Lee, J. H. H. (1991), 'A Lagrange multiplier test for GARCH models', *Economics Letters* **37**, 265–271.

Lee, J. H. and King, M. L. (1993), 'A locally most mean powerful based score test for arch and garch regression disturbances', *Journal of Business and Economic Statistics* **11**, 17–27. Correction 12 (1994), 139.

Pearson, K. (1933), 'On a method of determining whether a sample of size n supposed to have been drawn from a parent population', *Biometrika* **25**, 379–410.

Szroeter, J. (1978), 'A class of parametric tests for heteroscedasticity in linear econometric models', *Econometrica* **46**, 1311–1327.

Tippett, L. H. (1931), *The Methods of Statistics*, Williams and Norgate, London.

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

Wilkinson, B. (1951), 'A statistical consideration in psychological research', *Psychology Bulletin* **48**, 156–158.