Modified Variance Ratio Test for Autocorrelation in the Presence of Heteroskedasticity

Sohail Chand[*] and Nuzhat Aftab[**]

Abstract

Given that autocorrelation tests do not perform well in the presence of heteroskedasticity and in variance-break cases, we present three modified weighted variance ratio tests of autocorrelation. The numerical results show that the proposed tests perform better for small samples. They provide a better approximation of asymptotic distributions and are more powerful when the lag length is mis-specified. The study also applies these tests to data on the daily returns of two companies listed on the Pakistan Stock Exchange.

Keywords: Regression, variance break, wild bootstrap.

JEL classification: C40.

1. Introduction

Testing for autocorrelation in data is an important step in building a linear regression model. While the literature provides many tests for this purpose, the Breusch–Godfrey (BG) Lagrange multiplier (LM) test suggested by Breusch (1978) and Godfrey (1978) is the most popular among practitioners. The objective of these tests of misspecification is to obtain correct inferences about autocorrelation. Ignoring autocorrelation in data can lead to wrong conclusions when testing the significance of the regression parameters of central interest.

Generally, the BG-LM test is used without considering the presence of heteroskedasticity. Several studies investigate the performance of autocorrelation tests in the presence of conditional heteroskedasticity, including Wooldridge (1991), Guo and Phillips (2001) and Mantalos and Shukur (2005), among others. More recently, Jeong and Kang (2012) and Hyun et al. (2010) have used simulations to show that the BG-LM test can generate substantial size distortions in the presence of unconditional heteroskedasticity and misleading results in the case of a variance break.

[*]Associate Professor, College of Statistical and Actuarial Sciences, University of the Punjab, Pakistan.
[**] PhD Scholar, College of Statistical and Actuarial Sciences, University of the Punjab, Pakistan.

Shim et al. (2006) propose a modified LM test based on feasible generalized least squares (FGLS) and a revised variance ratio (VR) test for autocorrelation as an alternative to the BG LM test in the presence of variance breaks. Importantly, the FGLS-based LM test is not robust to the type of heteroskedasticity. Moreover, if the heteroskedasticity is not due to a variance break or if no prior information on the breakpoints is available, then the FGLS-based test gives invalid results.

Jeong and Kang (2012) argue that the VR test proposed by Shim et al. (2006), while robust to the type of heteroskedasticity, does not work well for finite samples. They suggest a modified VR test and wild bootstrap version of different VR tests in the presence of unconditional heteroskedasticity. Mun et al. (2014) propose a modified LM test for autocorrelation when there is a break in variance. Their simulation results show that the proposed test has good size and power properties when applied to small samples and is robust to variance-break heteroskedasticity.

This study focuses on the low power of these tests under a mis-specified lag length, especially for small samples. We propose three new VR tests by modifying the weights used in their standard versions. The simulation results show that these modified tests perform reasonably well for both static and dynamic regression models compared to existing tests. Moreover, they are robust to the presence of heteroskedasticity, have good size properties and remain more powerful even in variance-break cases.

The rest of the paper is organized as follows. Section 2 defines the model and VR tests used. The proposed modified VR tests are presented in Section 3. The Monte Carlo simulation results are given in Sections 4 and 5. Section 6 applies these tests to real-life contexts and Section 7 concludes the study.

2. Model and VR Test

We consider the following linear regression model:

$$y_i = X_i^t \beta + u_i, i = 1, 2, \dots, n \tag{1}$$

where y_i is the response variable, X_i is the ith row of the matrix X of order ($n \times (k+1)$), n is the sample size and k is the number of parameters. β is a vector of parameters of order (($k+1$) × 1) and $E(u_i) = 0$ for all i. We are interested in whether u_i is serially uncorrelated. The BG-LM test for autocorrelation is reliable when the disturbance term u_i has a constant variance, that is,

homoskedastic errors, but can be misleading in the case of heteroskedastic errors. To consider this scenario, we define the error term u_i in model (1) as

$$u_i = \varepsilon_i \eta_i \qquad (2)$$

where ε_i is a heteroskedasticity factor used to create different types of heteroskedasticity, η_i is white noise with a zero mean and unit variance, and $E(\eta_i \eta_j) = 0$ for $i \neq j$. The factor ε_i will result in multiplicative heteroskedasticity if it is a function of a pre-determined variable. In model (2), if we define the factor ε_i in the form of a categorical variable, then it will produce multiple variance-break cases as follows:

$$\varepsilon_i = \sigma_1^2 1[i \leq \tau_1 n] + \sigma_2^2 1[\tau_1 n < i \leq \tau_2 n] + \ldots + \sigma_{s+1}^2 1[i > \tau_s n] \qquad (3)$$

where $1[.]$ is the indicator function and $\tau \in (0, 1)$ is the break timing. Shim et al. (2006) and Hyun et al. (2010) study the simple variance break by considering a single variance breakpoint. In this case, the variance of u_i, that is, σ_i^2, is σ_1^2 for $i \leq \tau_1 n$ and σ_2^2 for $i > \tau_1 n$. Similarly, when we have s breaks, there will be $(s + 1)$ variances in ε_i.

The VR test is designed to identify the presence of serial correlation in a time series by investigating whether the variance of the q-period return is exactly q times larger than the variance of the one-period return. While the literature provides different versions of the VR test, in its simplest form the test compares the variance of the q-period return and q times the variance of the one-period return, that is:

$$VR(q) = \frac{\hat{\sigma}(q)^2}{q\hat{\sigma}(1)^2} - 1 \qquad (4)$$

where $q > 1$, $\hat{\sigma}(q)^2 = [\frac{n}{q(n-q+1)(n-q)}] \sum_{i=q}^{n} u_i(q)^2$, $\hat{\sigma}(1)^2 = [\frac{1}{n-k}] \sum_{i=1}^{n} u_i^2$, $u_i(q) = \sum_{j=0}^{q-1} \hat{u}_{i-j}$ and $\hat{u}_i = y_i - X_i \hat{\beta}$, where $\hat{\beta}$ is the estimator of β. If the errors are serially uncorrelated, then the q-period return variance will be precisely q times the variance of the single-period return. In this case, the value of VR(q) will be near 0.

Lo and MacKinlay (1988) show that the VR(q) given in model (4) can be written as

$$VR(q) = 2\left(1 - \frac{1}{q}\right)\rho(1) + 2\left(1 - \frac{2}{q}\right)\rho(2) + \cdots + 2\rho(q - 1) \qquad (5)$$

$$VR(q) \approx 2 \sum_{h=1}^{q-1} \left(1 - \frac{h}{q}\right) \rho(h) \tag{6}$$

where q is the lag length and $\rho(h) = \sum_{i=h+1}^{n} \hat{u}_i \hat{u}_{i-h} / \sum_{i=1}^{n} \hat{u}_i^2$ is the autocorrelation coefficient estimator at lag h. They argue that the VR(q) test has the following limiting distribution under the null hypothesis:

$$\sqrt{nq} \, VR(q) \overset{a}{\sim} N\left(0, \frac{2(2q-1)(q-1)}{3q}\right). \tag{7}$$

Model (5) describes the computation of variance ratios for an aggregate value q. We can see that these are a linear combination of the first $(q-1)$ autocorrelation coefficient estimators with mathematically declining weights. When $q = 2$, the VR(q) value will be approximately the first-order autocorrelation coefficient.

Jeong and Kang (2012) argue that the standard VR(q) test given in models (4) and (6) may be misleading in the presence of negative autocorrelation among the errors. For example, for $q = 3$, the VR(q) approaches 0 if $\Sigma \hat{u}_i \hat{u}_{i-2} = \Sigma \hat{u}_i \hat{u}_{i-1} = \Sigma \hat{u}_{i-2} \hat{u}_{i-1} = 0$, which implies that the errors have no serial correlation. When we have serially correlated errors, the VR(q) statistic will still approach 0, but in this case, $\Sigma \hat{u}_i \hat{u}_{i-2} = -\Sigma \hat{u}_i \hat{u}_{i-1} - \Sigma \hat{u}_{i-2} \hat{u}_{i-1} \neq 0$.

Pagan (1984, 1986) shows that, while the presence of negative autocorrelation might not lead the value of VR(q) towards 0, it can affect the power of the test. To resolve this, Jeong and Kang (2012) suggest using the quadratic version of the VR(q) test, QVR(q), which is defined as:

$$QVR(q) = \sum_{h=1}^{q-1} \left(1 - \frac{h}{q}\right) \rho(h)^2 \tag{8}$$

Here, Jeong and Kang (2012) employ the bootstrap distribution, indicating that the asymptotic distribution for it was not derived.

3. Modified VR Tests

This section proposes three new tests: WVR, WQVR and WAVR. The first two are modified versions of the VR and QVR tests. The third test suggests using absolute autocorrelation, given that (as discussed in Section 2) the presence of negative autocorrelation adversely affects the performance of VR tests. Unlike the QVR test, this uses the original magnitude of

autocorrelation while addressing the problem of negative autocorrelation. The new tests are defined as follows:

$$WVR(q) = p \sum_{h=1}^{q-1} W_h \rho(h) \tag{9}$$

$$WQVR(q) = p \sum_{h=1}^{q-1} W_h \rho(h)^2 \tag{10}$$

$$WAVR(q) = p \sum_{h=1}^{q-1} W_h |\rho(h)| \tag{11}$$

where $p = q - 1$, $W_h = \alpha(1 - \alpha)^{h-1}$ and $0 < \alpha < 1$. The rationale for using the W_h term is that it enables one to simultaneously control the lag length and assign geometrically decaying weights to the autocorrelations. The parameter α is user-defined but based on empirical evidence, we propose that $\alpha = 0.7$. For these suggested tests, we do not attempt to derive their asymptotic distribution since we employ the wild bootstrap procedure for these statistics.

4. Numerical Results

This section provides numerical results for the empirical size and power of the suggested VR tests, including a comparison with the existing tests. We consider two types of heteroskedasticity: multiplicative heteroskedasticity and the variance-break case. In both cases, we use static l and dynamic regression models. In comparing our results with those reported in the literature, we follow the simulation design used by Jeong and Kang (2012).

The static regression model is defined as:

$$y_i = \alpha_0 + \alpha_1 X_i + u_i, i = 1, 2, \dots, n \tag{12}$$

where the predictor $X_i \sim N(0, 1)$ and the values of α_0 and α_1 are set to 1.

The dynamic regression model is given as:

$$y_i = \beta_0 + \beta_1 y_{i-1} + \beta_2 X_i + \beta_3 X_{i-1} + u_i, i = 1, 2, \dots, n \tag{13}$$

The value of β_1 is set to 0.5 to avoid the complexity of nonstationarity, while the remaining parameters are set to 1. To remove the effect of initial values, we have burned out the first 100 simulated observations.

To compute empirical size, we define the serially independent errors of the regression model of multiplicative heteroskedasticity as follows:

$$u_i = \varepsilon_i \eta_i \tag{14}$$

where $\varepsilon_i \sim N(0,1)$, sorted by absolute values, is used to create the heteroskedasticity and $\eta \sim U(-0.5, 0.5)$ is the random error term. Thus, $\text{var}(u_i)$ is increasing in i. To compare power, we consider the following seven data-generating processes (DGPs) for the error term:

$$AR(1)+: \quad u_i = 0.7u_{i-1} + \varepsilon_i \eta_i \tag{15}$$

$$AR(1)-: \quad u_i = -0.7u_{i-1} + \varepsilon_i \eta_i \tag{16}$$

$$AR(2): \quad u_i = -0.5u_{i-1} + 0.4u_{i-2} + \varepsilon_i \eta_i \tag{17}$$

$$AR(4): \quad u_i = 0.8u_{i-1} + 0.0u_{i-2} + 0.2u_{i-3} + 0.06u_{i-4} + \varepsilon_i \eta_i \tag{18}$$

$$MA(1)+: \quad u_i = \varepsilon_i \eta_i + 0.7\varepsilon_{i-1}\eta_{i-1} \tag{19}$$

$$MA(1)-: \quad u_i = \varepsilon_i \eta_i - 0.7\varepsilon_{i-1}\eta_{i-1} \tag{20}$$

$$MA(2): \quad u_i = \varepsilon_i \eta_i - 0.5\varepsilon_{i-1}\eta_{i-1} + 0.4\varepsilon_{i-2}\eta_{i-2} \tag{21}$$

The following wild-bootstrap algorithm is used to compute size and power. The only difference between the two is that, for empirical size, the sample is simulated under the null hypothesis, while for power, the simulation is under the alternative hypothesis. In both cases, the model used under the null hypothesis is fitted to the simulated sample. T generally denotes any of the tests, that is, VR, QVR, WVR, WQVR and WAVR.

1. Compute the OLS residuals, $\hat{u}_i = y_i - X_i \hat{\beta}$.

2. Calculate the test (say, T) using the OLS residuals obtained in step 1.

3. Bootstrap y_i (say, y_i^*) using $y_i^* = X_i\hat{\beta} + \{\frac{\hat{u}_i V_i}{\sqrt{(1-h_i)}}\}$, where $h_i = X_i^t(X_t X)^{-1}X_i$ and V_i is a random vector such that $E(V_i) = 0$ and $E(V_1^2) = 1$. Here, V_i is taken from the Rademacher distribution as $\Pr(V_i = 1) = \Pr(V_i = -1) = 0.5$.

4. Repeat steps 1–3 B times (the number of bootstrap samples) to obtain the test for each bootstrap sample, that is, $T_1^*, T_2^*, \ldots, T_B^*$.

5. Compute the $100(1 - \alpha)\%$ percentile (say, $T_{1-\alpha}^*$) of the bootstrap distribution obtained in step 4.

6. Repeat steps 1–5 N times. The empirical size and power are computed as $\#(T > T_{1-\alpha}^*)/N$.

Kim (2006) and Jeong and Kang (2012) show that, using the bootstrap method, one can construct an empirical (bootstrap) distribution for the VR test that is robust to the form of heteroskedasticity. They argue that a bootstrap distribution gives more accurate critical values than an asymptotic normal distribution. We consider three choices of sample size – n = 30, 50 and 100 – and four different lag lengths of the autoregressive serial correlation, p = 1, 2, 4, 12. The nominal size of the test is fixed at 5 percent and $q = p + 1$. The number of Monte Carlo runs is set to 1,000, while 500 bootstrap samples are used.

The results for the empirical size of the proposed tests for static and dynamic models are given in Tables 1 and 2, respectively. The static and dynamic models given in (12) and (13) are considered null models. The results given in Table 1 show that, when the null model is static, the size of the proposed test is a very close approximation to the corresponding nominal size. While the standard VR and QVR tests also yield a good size approximation, they deviate from the empirical size the larger the chosen lag. The WVR, WQVR and WAVR tests give better size results and remain stable at varying lag values. Thus, the approximation to the asymptotic distribution of the proposed tests is not affected by the mis-specified lag length.

Table 1: Size comparison for static model with multiplicative heteroskedasticity

Test	T = 30				T = 50				T = 100			
	p = 1	p = 2	p = 4	p = 12	p = 1	p = 2	p = 4	p = 12	p = 1	p = 2	p = 4	p = 12
VR	0.044	0.053	0.062	0.044	0.045	0.043	0.055	0.062	0.055	0.054	0.049	0.074
QVR	0.055	0.050	0.061	0.055	0.043	0.053	0.052	0.050	0.055	0.059	0.056	0.056
WVR	0.053	0.052	0.052	0.048	0.054	0.049	0.057	0.056	0.055	0.044	0.058	0.042
WQVR	0.050	0.048	0.053	0.056	0.049	0.047	0.043	0.046	0.051	0.040	0.056	0.055
WAVR	0.05	0.052	0.054	0.062	0.049	0.039	0.043	0.048	0.051	0.044	0.058	0.051

Note: Nominal size = 5%.
Source: Authors' calculations

Table 2: Size comparison for dynamic model with multiplicative heteroskedasticity

Test	T = 30				T = 50				T = 100			
	p = 1	p = 2	p = 4	p = 12	p = 1	p = 2	p = 4	p = 12	p = 1	p = 2	p = 4	p = 12
VR	0.062	0.055	0.058	0.048	0.048	0.054	0.054	0.053	0.045	0.050	0.055	0.046
QVR	0.043	0.052	0.047	0.048	0.050	0.044	0.054	0.055	0.036	0.049	0.055	0.058
WVR	0.053	0.047	0.054	0.054	0.052	0.044	0.047	0.052	0.050	0.045	0.059	0.055
WQVR	0.040	0.034	0.051	0.051	0.056	0.051	0.038	0.045	0.043	0.054	0.048	0.051
WAVR	0.040	0.036	0.047	0.053	0.056	0.049	0.039	0.047	0.043	0.054	0.048	0.053

Source: Authors' calculations

Table 2 gives the results for the dynamic model defined in (13). In this case, the proposed WVR test shows a closer size approximation than the other tests at different lag values. The numerical results show that the size of the modified VR tests is comparable with that of the classical VR tests as the lag and sample size increases.

The results of empirical power against various alternatives are given in Tables 3 and 4. The sample is generated under seven different models: AR(1)+, AR(1)−, AR(2), AR(4), MA(1)+, MA(1)− and MA(2), as defined in equations (15) to (21). The results show that the proposed tests generally have a higher empirical power than the existing tests, outperforming the latter for all considered choices of p. For all considered alternative DGPs, we find that the power of the VR and QVR tests decreases as the lag increases.

Table 3: Empirical power for static model of multiplicative heteroskedasticity

	Test	T = 30				T = 50				T = 100			
		p = 1	p = 2	p = 4	p = 12	p = 1	p = 2	p = 4	p = 12	p = 1	p = 2	p = 4	p = 12
	VR	0.926	0.927	0.826	0.608	0.992	0.988	0.959	0.785	1	1	1	0.951
	QVR	0.835	0.85	0.805	0.716	0.986	0.976	0.963	0.939	0.999	1	1	1
AR(1)+	WVR	0.825	0.894	0.906	0.906	0.99	0.986	0.981	0.982	1	1	1	0.999
	WQVR	0.825	0.827	0.844	0.838	0.975	0.98	0.974	0.974	1	1	0.999	0.999
	WAVR	0.825	0.827	0.822	0.822	0.975	0.976	0.97	0.969	1	1	0.999	0.999
	VR	0.665	0.368	0.234	0.168	0.851	0.591	0.378	0.164	0.99	0.891	0.71	0.289
	QVR	0.665	0.663	0.607	0.511	0.851	0.844	0.801	0.722	0.99	0.983	0.974	0.923
AR(1)-	WVR	0.916	0.935	0.921	0.913	0.991	0.997	0.991	0.989	1	1	1	1
	WQVR	0.916	0.925	0.917	0.904	0.991	0.994	0.989	0.985	1	1	1	1
	WAVR	0.916	0.907	0.91	0.9	0.991	0.994	0.987	0.985	1	1	1	1
	VR	0.902	0.297	0.219	0.233	0.986	0.6	0.386	0.217	0.999	0.911	0.709	0.296
	QVR	0.902	0.946	0.946	0.93	0.986	0.988	0.991	0.99	0.999	1	1	1
AR(2)	WVR	0.937	0.955	0.954	0.955	0.982	0.987	0.992	0.996	1	1	1	1
	WQVR	0.937	0.898	0.891	0.889	0.982	0.963	0.971	0.961	1	0.999	0.997	0.996
	WAVR	0.937	0.862	0.879	0.865	0.982	0.954	0.967	0.956	1	1	0.997	0.994
	VR	0.925	0.919	0.786	0.456	0.995	0.996	0.958	0.677	1	1	0.988	0.884
	QVR	0.925	0.915	0.890	0.831	0.995	0.996	0.992	0.974	1	1	1	1
AR(4)	WVR	0.960	0.950	0.943	0.949	0.999	0.995	0.993	0.996	1	1	1	1
	WQVR	0.987	0.907	0.905	0.921	0.997	0.995	0.993	0.993	1	1	1	1
	WAVR	0.987	0.900	0.902	0.908	0.997	0.993	0.990	0.992	1	1	1	1
	VR	0.794	0.562	0.404	0.234	0.928	0.756	0.527	0.283	0.998	0.923	0.715	0.414
	QVR	0.575	0.597	0.513	0.352	0.835	0.874	0.823	0.677	0.987	0.994	0.988	0.943
MA(1)+	WVR	0.756	0.858	0.871	0.876	0.939	0.970	0.985	0.979	0.999	1	1	1
	WQVR	0.535	0.543	0.561	0.551	0.84	0.797	0.826	0.826	0.999	0.973	0.975	0.968
	WAVR	0.535	0.539	0.555	0.55	0.84	0.769	0.795	0.807	0.999	0.967	0.965	0.955
	VR	0.458	0.411	0.312	0.215	0.722	0.707	0.584	0.279	0.934	0.941	0.9	0.484
	QVR	0.458	0.451	0.378	0.253	0.722	0.69	0.639	0.472	0.934	0.927	0.898	0.767
MA(1)-	WVR	0.658	0.552	0.526	0.531	0.909	0.779	0.751	0.759	0.996	0.94	0.938	0.945
	WQVR	0.658	0.662	0.662	0.652	0.909	0.873	0.848	0.884	0.996	0.981	0.987	0.99
	WAVR	0.658	0.653	0.664	0.645	0.909	0.852	0.835	0.867	0.996	0.97	0.98	0.987
	VR	0.705	0.356	0.157	0.104	0.871	0.530	0.250	0.098	0.977	0.823	0.441	0.113
	QVR	0.705	0.699	0.657	0.547	0.871	0.870	0.875	0.788	0.977	0.992	0.988	0.969
MA(2)	WVR	0.668	0.725	0.699	0.693	0.865	0.881	0.884	0.868	0.972	0.987	0.988	0.979
	WQVR	0.668	0.695	0.664	0.670	0.865	0.849	0.860	0.859	0.972	0.962	0.976	0.974
	WAVR	0.668	0.669	0.648	0.640	0.865	0.830	0.841	0.838	0.972	0.949	0.969	0.969

Source: Authors' calculations

Table 4: Empirical power for dynamic model of multiplicative heteroskedasticity

	Test	T = 30				T = 50				T = 100			
		p = 1	p = 2	p = 4	p = 12	p = 1	p = 2	p = 4	p = 12	p = 1	p = 2	p = 4	p = 12
	VR	0.841	0.830	0.700	0.514	0.981	0.979	0.919	0.698	0.999	1	0.997	0.919
	QVR	0.647	0.621	0.563	0.507	0.926	0.924	0.899	0.821	0.999	1	0.999	0.989
AR(1)+	WVR	0.849	0.806	0.803	0.815	0.982	0.959	0.965	0.973	1	0.999	1	1
	WQVR	0.618	0.640	0.641	0.636	0.933	0.932	0.931	0.934	1	0.997	1	1
	WAVR	0.618	0.609	0.628	0.618	0.933	0.918	0.917	0.913	1	0.996	0.999	1
	VR	0.848	0.292	0.168	0.132	0.990	0.680	0.434	0.184	1	0.996	0.825	0.378
	QVR	0.848	0.870	0.815	0.734	0.990	0.980	0.985	0.942	1	1	1	0.998
AR(1)-	WVR	0.865	0.872	0.866	0.869	0.984	0.983	0.985	0.981	1	1	1	1
	WQVR	0.865	0.840	0.823	0.833	0.984	0.980	0.981	0.979	1	1	1	1
	WAVR	0.865	0.801	0.795	0.803	0.984	0.980	0.979	0.972	1	1	1	1
	VR	0.833	0.058	0.053	0.067	0.973	0.264	0.097	0.06	0.999	0.738	0.379	0.119
	QVR	0.833	0.909	0.912	0.898	0.973	0.986	0.989	0.987	0.999	1	1	1
AR(2)	WVR	0.861	0.928	0.92	0.933	0.971	0.99	0.991	0.993	1	1	1	1
	WQVR	0.861	0.759	0.777	0.763	0.971	0.947	0.947	0.947	1	0.996	0.998	0.997
	WAVR	0.861	0.69	0.739	0.727	0.971	0.932	0.94	0.937	1	0.996	0.999	0.998
	VR	0.924	0.912	0.732	0.438	0.996	0.994	0.939	0.614	1	1	0.997	
	QVR	0.776	0.742	0.678	0.626	0.983	0.997	0.971	0.923	1	1	1	
AR(4)	WVR	0.927	0.906	0.874	0.893	0.993	0.988	0.984	0.991	1	0.999	1	1
	WQVR	0.764	0.798	0.748	0.791	0.977	0.979	0.974	0.982	1	0.999	1	1
	WAVR	0.764	0.771	0.737	0.768	0.977	0.973	0.968	0.976	1	0.999	1	1
	VR	0.771	0.489	0.337	0.219	0.9	0.687	0.467	0.265	0.996	0.989	0.641	0.373
	QVR	0.369	0.392	0.342	0.281	0.746	0.799	0.744	0.58	0.98	0.992	0.979	0.916
MA(1)+	WVR	0.969	0.814	0.812	0.813	0.926	0.975	0.968	0.962	0.996	0.997	1	1
	WQVR	0.383	0.372	0.398	0.397	0.752	0.74	0.728	0.744	0.977	0.955	0.956	0.965
	WAVR	0.383	0.364	0.397	0.39	0.752	0.717	0.716	0.715	0.977	0.943	0.943	0.949
	VR	0.501	0.484	0.336	0.221	0.816	0.848	0.727	0.397	0.988	0.992	0.987	0.788
	QVR	0.501	0.444	0.403	0.297	0.816	0.831	0.75	0.569	0.988	0.994	0.984	0.914
MA(1)-	WVR	0.470	0.441	0.434	0.400	0.828	0.678	0.689	0.679	0.983	0.922	0.930	0.928
	WQVR	0.470	0.503	0.521	0.505	0.828	0.798	0.818	0.794	0.983	0.961	0.979	0.984
	WAVR	0.470	0.504	0.521	0.515	0.828	0.789	0.809	0.785	0.983	0.951	0.974	0.974
	VR	0.589	0.235	0.085	0.084	0.821	0.438	0.167	0.069	0.985	0.730	0.34	0.078
	QVR	0.589	0.571	0.567	0.438	0.821	0.826	0.833	0.703	0.985	0.973	0.979	0.952
MA(2)	WVR	0.604	0.657	0.607	0.585	0.830	0.872	0.846	0.873	0.986	0.985	0.981	0.975
	WQVR	0.604	0.590	0.548	0.558	0.830	0.822	0.810	0.826	0.986	0.959	0.97	0.951
	WAVR	0.604	0.541	0.521	0.536	0.830	0.801	0.776	0.797	0.986	0.946	0.953	0.941

Source: Authors' calculations

Table 4 gives the empirical power results for the dynamic model of multiplicative heteroskedasticity. These are generally similar to the results for the static model. The power of the WVR test is higher, especially for small samples. In the dynamic model, as the lag length increases, the power

of the VR and QVR tests decreases even when the lag length is properly specified. This decrease is more evident in the case of negative autocorrelation, for example, for MA(1)− and AR(1)−. The VR test becomes almost powerless for larger lags in the MA(2) and AR(2) models. Again, the proposed tests indicate good power except in some cases, for instance, MA(1)− for small samples.

5. Variance-Break Case

This section considers variance-break heteroskedasticity. Following Jeong and Kang (2012), the heteroskedastic factor ε_i is defined as a categorical variable as follows:

$$\varepsilon_i^2 = \sigma_1^2 \mathbb{1}[i \leq \tau_1 n] + \sigma_2^2 \mathbb{1}[\tau_1 n < i \leq \tau_2 n] + \sigma_3^2 \mathbb{1}[i > \tau_2 n]. \tag{24}$$

We have specified $\sigma_1^2 = 1$, $\sigma_2^2 = 4$, $\sigma_3^2 = 9$, $\tau_1 = 0.3$ and $\tau_2 = 0.7$ in this study. Except for the change in the type of heteroskedasticity, the static and dynamic DGPs used are similar to those employed in the previous section.

Table 5 gives the empirical size of the considered VR tests for the static model, using a two-point variance break. The results show that the proposed tests are comparable in terms of empirical size and that the size of the VR test is adversely affected by the presence of a break in variance, especially when the sample is small, that is, for $n = 30$ and $n = 50$. The results for the dynamic model with a two-point variance break, given in Table 6, are not much different from the results for the static model. The proposed tests yield a good size approximation for both models in the variance-break case.

Table 5: Empirical size under static model with variance break

	T = 30				T = 50				T = 100			
Test	p = 1	p = 2	p = 4	p = 12	p = 1	p = 2	p = 4	p = 12	p = 1	p = 2	p = 4	p = 12
VR	0.048	0.063	0.059	0.059	0.054	0.059	0.056	0.06	0.056	0.041	0.041	0.048
QVR	0.057	0.056	0.049	0.052	0.06	0.049	0.060	0.048	0.047	0.051	0.044	0.053
WVR	0.055	0.054	0.06	0.054	0.058	0.058	0.060	0.048	0.056	0.046	0.047	0.057
WQVR	0.048	0.051	0.055	0.049	0.058	0.047	0.055	0.057	0.050	0.053	0.052	0.052
WAVR	0.048	0.051	0.052	0.055	0.058	0.044	0.052	0.059	0.050	0.053	0.049	0.052

Source: Authors' calculations

Table 6: Empirical size under dynamic model with variance break

Test	T = 30				T = 50				T = 100			
	p = 1	p = 2	p = 4	p = 12	p = 1	p = 2	p = 4	p = 12	p = 1	p = 2	p = 4	p = 12
VR	0.060	0.059	0.057	0.06	0.065	0.049	0.056	0.054	0.049	0.045	0.051	0.051
QVR	0.042	0.046	0.048	0.051	0.05	0.048	0.054	0.047	0.038	0.044	0.045	0.055
WVR	0.042	0.044	0.057	0.034	0.05	0.048	0.052	0.047	0.04	0.048	0.04	0.046
WQVR	0.042	0.035	0.04	0.045	0.05	0.043	0.049	0.038	0.038	0.041	0.033	0.04
WAVR	0.042	0.033	0.039	0.043	0.05	0.04	0.045	0.041	0.038	0.041	0.034	0.038

Source: Authors' calculations

The empirical power results of the variance-break case for the five tests are given in Tables 7 and 8 and lead to the same conclusion as for the multiplicative heteroskedasticity case. The tests' good performance is, therefore, unaffected by the change in type of heteroskedasticity.

Table 7: Empirical power under static model with variance-break heteroskedasticity

	Test	T = 30				T = 50				T = 100			
		p = 1	p = 2	p = 4	p = 12	p = 1	p = 2	p = 4	p = 12	p = 1	p = 2	p = 4	p = 12
AR(1)+	VR	0.992	0.896	0.818	0.582	0.994	0.987	0.955	0.803	1	1	1	1
	QVR	0.825	0.794	0.744	0.685	0.973	0.971	0.956	0.936	1	1	1	1
	WVR	0.922	0.626	0.601	0.63	0.994	0.982	0.985	0.985	1	1	1	1
	WQVR	0.825	0.799	0.779	0.786	0.973	0.975	0.97	0.973	1	1	1	1
	WAVR	0.825	0.777	0.761	0.779	0.973	0.962	0.965	0.968	1	1	1	1
AR(1)-	VR	0.859	0.406	0.271	0.25	0.99	0.764	0.542	0.315	1	0.995	0.943	0.611
	QVR	0.859	0.882	0.844	0.771	0.99	0.987	0.98	0.96	1	1	1	1
	WVR	0.859	0.894	0.898	0.876	0.99	0.987	0.984	0.986	1	1	1	1
	WQVR	0.859	0.876	0.889	0.879	0.99	0.987	0.987	0.985	1	1	1	1
	WAVR	0.859	0.857	0.875	0.864	0.99	0.987	0.984	0.98	1	1	1	1
AR(4)	VR	0.955	0.93	0.834	0.515	0.998	0.998	0.949	0.671	1	1	1	0.87
	QVR	0.877	0.876	0.852	0.771	0.993	0.992	0.98	0.982	1	1	1	1
	WVR	0.877	0.736	0.721	0.722	0.993	0.962	0.961	0.96	1	1	1	1
	WQVR	0.877	0.878	0.879	0.877	0.993	0.987	0.985	0.987	1	1	1	1
	WAVR	0.877	0.873	0.856	0.859	0.993	0.986	0.983	0.986	1	1	1	1
MA(1)+	VR	0.701	0.507	0.359	0.235	0.923	0.701	0.493	0.294	1	0.953	0.727	0.441
	QVR	0.473	0.491	0.427	0.356	0.83	0.855	0.776	0.597	0.997	0.998	0.996	0.959
	WVR	0.701	0.835	0.835	0.856	0.923	0.973	0.981	0.97	1	1	1	1
	WQVR	0.476	0.504	0.503	0.528	0.83	0.794	0.792	0.798	0.997	0.982	0.985	0.984
	WAVR	0.473	0.502	0.503	0.524	0.83	0.768	0.778	0.775	0.997	0.979	0.978	0.976
MA(1)-	VR	0.581	0.56	0.461	0.365	0.827	0.607	0.444	0.24	0.994	0.885	0.678	0.962
	QVR	0.581	0.552	0.490	0.36	0.827	0.874	0.791	0.602	0.994	0.998	0.997	0.961
	WVR	0.581	0.463	0.481	0.474	0.827	0.891	0.885	0.879	0.994	0.998	0.998	0.949
	WQVR	0.581	0.593	0.609	0.59	0.827	0.789	0.812	0.787	0.994	0.989	0.985	0.984
	WAVR	0.581	0.578	0.614	0.583	0.827	0.762	0.789	0.772	0.994	0.985	0.978	0.979
MA(2)	VR	0.61	0.289	0.109	0.111	0.842	0.504	0.201	0.085	0.98	0.823	0.462	0.123
	QVR	0.61	0.642	0.594	0.491	0.842	0.851	0.824	0.726	0.98	0.991	0.988	0.967
	WVR	0.61	0.673	0.651	0.624	0.842	0.87	0.854	0.857	0.98	0.992	0.987	0.986
	WQVR	0.61	0.63	0.632	0.612	0.842	0.819	0.838	0.793	0.98	0.975	0.987	0.974
	WAVR	0.61	0.597	0.6	0.593	0.842	0.803	0.798	0.767	0.98	0.968	0.974	0.966
AR(2)	VR	0.85	0.145	0.11	0.143	0.977	0.482	0.249	0.175	1	0.908	0.692	0.277
	QVR	0.85	0.887	0.887	0.858	0.977	0.99	0.981	0.975	1	1	1	1
	WVR	0.85	0.902	0.907	0.912	0.977	0.994	0.986	0.986	1	1	1	1
	WQVR	0.85	0.812	0.827	0.832	0.977	0.962	0.956	0.958	1	1	1	0.998
	WAVR	0.85	0.777	0.801	0.796	0.977	0.955	0.946	0.952	1	1	1	0.998

Source: Authors' calculations

Table 8: Empirical power under dynamic model with variance-break heteroskedasticity

	Test	T = 30				T = 50				T = 100			
		p = 1	p = 2	p = 4	p = 12	p = 1	p = 2	p = 4	p = 12	p = 1	p = 2	p = 4	p = 12
AR(1)+	VR	0.825	0.779	0.685	0.456	0.974	0.954	0.906	0.633	1	0.999	0.996	0.889
	QVR	0.572	0.546	0.501	0.409	0.897	0.88	0.854	0.742	0.998	0.997	0.998	0.983
	WVR	0.828	0.769	0.772	0.773	0.974	0.931	0.941	0.944	0.998	0.997	0.998	0.999
	WQVR	0.572	0.55	0.548	0.562	0.897	0.878	0.887	0.861	0.998	0.997	0.998	0.999
	WAVR	0.572	0.539	0.529	0.536	0.897	0.866	0.872	0.854	0.998	0.996	0.998	0.999
AR(1)-	VR	0.777	0.172	0.114	0.098	0.97	0.453	0.231	0.095	1	0.881	0.602	0.164
	QVR	0.777	0.759	0.737	0.629	0.97	0.97	0.961	0.924	1	1	1	1
	WVR	0.777	0.793	0.807	0.811	0.97	0.974	0.973	0.981	1	1	1	1
	WQVR	0.777	0.707	0.728	0.723	0.97	0.953	0.962	0.966	1	1	1	1
	WAVR	0.777	0.662	0.702	0.673	0.97	0.937	0.953	0.948	1	1	1	1
AR(4)	VR	0.894	0.885	0.711	0.400	0.991	0.989	0.907	0.525	1	1	0.992	0.767
	QVR	0.698	0.691	0.635	0.548	0.964	0.963	0.932	0.853	1	1	1	1
	WVR	0.698	0.851	0.846	0.854	0.964	0.978	0.972	0.848	1	1	1	1
	WQVR	0.698	0.704	0.701	0.692	0.964	0.959	0.951	0.931	1	1	1	1
	WAVR	0.698	0.687	0.691	0.676	0.964	0.952	0.941	0.925	1	0.999	0.999	1
MA(1)+	VR	0.651	0.451	0.306	0.203	0.907	0.591	0.383	0.209	0.996	0.906	0.614	0.296
	QVR	0.348	0.345	0.314	0.263	0.700	0.724	0.691	0.483	0.983	0.993	0.986	0.93
	WVR	0.651	0.774	0.805	0.779	0.907	0.947	0.968	0.972	0.996	1	1	1
	WQVR	0.348	0.349	0.365	0.364	0.700	0.642	0.65	0.641	0.983	0.957	0.96	0.963
	WAVR	0.348	0.351	0.365	0.357	0.700	0.619	0.628	0.621	0.983	0.952	0.943	0.945
MA(1)-	VR	0.42	0.353	0.237	0.164	0.742	0.73	0.599	0.302	0.987	0.995	0.982	0.7
	QVR	0.42	0.386	0.322	0.253	0.742	0.716	0.633	0.426	0.987	0.992	0.976	0.888
	WVR	0.42	0.358	0.361	0.360	0.742	0.593	0.592	0.574	0.987	0.927	0.91	0.91
	WQVR	0.42	0.405	0.439	0.415	0.742	0.72	0.723	0.709	0.987	0.972	0.972	0.966
	WAVR	0.42	0.412	0.433	0.415	0.742	0.706	0.702	0.691	0.987	0.964	0.957	0.957
MA(2)	VR	0.51	0.177	0.059	0.062	0.773	0.338	0.092	0.081	0.977	0.653	0.205	0.047
	QVR	0.51	0.526	0.505	0.382	0.773	0.786	0.778	0.743	0.977	0.984	0.989	0.954
	WVR	0.51	0.551	0.561	0.551	0.773	0.812	0.801	0.784	0.977	0.984	0.985	0.975
	WQVR	0.51	0.483	0.475	0.478	0.773	0.743	0.759	0.751	0.977	0.957	0.964	0.954
	WAVR	0.51	0.446	0.445	0.443	0.773	0.709	0.717	0.696	0.977	0.942	0.951	0.938
AR(2)	VR	0.766	0.04	0.026	0.057	0.947	0.073	0.033	0.039	1	0.259	0.059	0.02
	QVR	0.766	0.85	0.831	0.826	0.947	0.984	0.988	0.971	1	1	1	1
	WVR	0.766	0.856	0.861	0.882	0.947	0.986	0.992	0.983	1	1	1	1
	WQVR	0.766	0.61	0.561	0.64	0.947	0.829	0.844	0.845	1	0.965	0.97	0.974
	WAVR	0.766	0.511	0.508	0.573	0.947	0.792	0.837	0.839	1	0.974	0.98	0.986

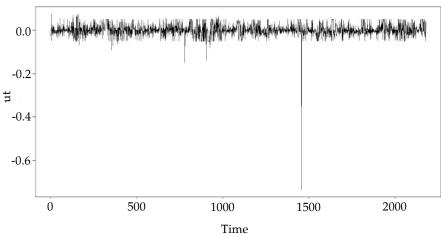*Source:* Authors' calculations

6. Real-Life Applications

This section applies the classical and proposed VR tests to daily returns data for two companies listed on the Pakistan Stock Exchange and

to the Microsoft Corporation. As Yasmin (2009) points out, heteroskedasticity is very common in financial data series. The data used here represents the daily stock returns of Dawood Hercules (DAWH) and Kohinoor Textile Mills Limited (KTML) for the period 1 July 2004 to 28 February 2014 as well as the daily stock returns of Microsoft for the period May 2000 to May 2017.

We test the series for the presence of conditional heteroskedasticity, using Engle's (1982) LM test. As Figure 1 shows, the daily returns series for DAWH does not exhibit heteroskedasticity (LM = 0.0717, p-value = 1). The KTML and Microsoft series, however, both indicate the presence of heteroskedasticity (LM = 479.9, p-value = 0.000; and LM = 261.27, p-value = 0.000, respectively).

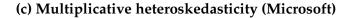**Figure 1: Daily returns data for DAWH, KTML and Microsoft**
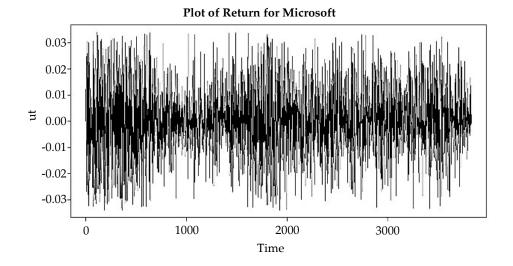
(a) Homoskedasticity (DAWH)

Plot of Return for Microsoft

(b) Variance break (KTML)

Plot of Return for Microsoft

(c) Multiplicative heteroskedasticity (Microsoft)

Plot of Return for Microsoft

The results of the VR tests are given in Table 9. The DAWH return series does not exhibit any conditional heteroskedasticity. All the tests are significant, indicating the presence of autocorrelation in the data even at a 1 percent level of significance for all considered choices of p. In the variance-break case – for the KTML returns – the VR and QVR tests are not significant for any lag value, even at a 10 percent level of significance. The modified QVR and AVR tests show very strong levels of rejection.

When the heteroskedasticity is multiplicative, the VR and WVR tests yield insignificant results at all lag values, while the QVR, WQVR and WAVR tests are significant at lags 1, 2 and 4 at the 5 percent level. The behavior of the proposed tests is consistent with the change in lag value, although the critical region for QVR changes with the lag value. The proposed tests yield significant results for all values of p at the 10 and 5 percent levels of significance and present fairly strong evidence against the null hypothesis.

Table 9: VR tests for daily returns data

p	VR	p value	QVR	p value	WVR	p value	WAVR	p value	WQVR	p value
DAWH										
1	0.096***	0.0000	0.0046***	0.0000	0.0672***	0.0000	0.0672***	0.0000	0.0064***	0.0000
2	0.1653***	0.0000	0.0072***	0.0000	0.1108***	0.0000	0.0018***	0.0000	0.0116***	0.0000
4	0.2447***	0.0000	0.0095***	0.0010	0.2267***	0.0000	0.2267***	0.0000	0.0233***	0.0000
12	0.4199***	0.0000	0.0132***	0.0050	0.6808***	0.0020	0.6808***	0.0010	0.0699***	0.0010
KTML										
1	-0.0345	0.9502	0.0006*	0.0938	-0.0242	0.9502	0.0242*	0.0938	0.0006*	0.0938
2	-0.0451	0.9222	0.0008	0.1634	-0.0489	0.945	0.0477*	0.0818	0.0017*	0.0646
4	-0.0294	0.7268	0.0013	0.2046	-0.0906	0.9296	0.1026*	0.0768	0.0035**	0.0561
12	-0.0994	0.871	0.0038	0.1548	-0.2725	0.9286	0.3085*	0.0802	0.0107**	0.0578
Microsoft										
1	-0.0365	0.9884	0.0007**	0.0254	-0.0255	0.9884	0.0255**	0.0254	0.0009**	0.0254
2	-0.053	0.9846	0.0009**	0.0422	-0.0540	0.9870	0.0540**	0.0408	0.0019**	0.0320
4	-0.078	0.9896	0.0012*	0.0676	-0.1117	0.9926	0.1119**	0.0342	0.0038**	0.0288
12	-0.1475	0.9920	0.0034**	0.021	-0.3376	0.9902	0.3385**	0.0338	0.0115**	0.0262

Note: * = significant at 10% level, ** = significant at 5% level, *** = significant at 1% level.

7. Conclusion

The performance of autocorrelation tests is adversely affected if the errors do not have a constant variance. In this context, we propose three alternative VR tests that yield very good size properties for both small and large samples and are not heavily dependent on the specified lag. These tests retain their power across static and dynamic models regardless of whether the heteroskedasticity is multiplicative or if there is a variance break.

References

Breusch, T. S. (1978). Testing for autocorrelation in dynamic linear models. *Australian Economic Papers*, *17*(31), 334–355.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, *50*(4), 987–1007.

Godfrey, L. G. (1978). Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables. *Econometrica*, *46*(6), 1293–1301.

Guo, B., & Phillips, P. C. (2001). *Testing for autocorrelation and unit roots in the presence of conditional heteroskedasticity of unknown form* (Economics Working Paper No. 540). Santa Cruz, CA: UC Santa Cruz.

Hyun, J.-Y., Mun, H. H., Kim, T.-H., & Jeong, J. (2010). The effect of a variance shift on the Breusch–Godfrey's LM test. *Applied Economics Letters*, *17*(4), 399–404.

Jeong, J., & Kang, B. (2012). Wild-bootstrapped variance-ratio test for autocorrelation in the presence of heteroskedasticity. *Journal of Applied Statistics*, *39*(7), 1531–1542.

Kim, J. H. (2006). Wild bootstrapping variance ratio tests. *Economics Letters*, *92*(1), 38–43.

Lo, A. W., & MacKinlay, A. C. (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of Financial Studies*, *1*(1), 41–66.

Mantalos, P., & Shukur, G. (2005). The effect of the GARCH (1, 1) on autocorrelation tests in dynamic systems of equations. *Applied Economics*, *37*(16), 1907–1913.

Mun, H.-H., Shim, E.-Y., & Kim, T.-H. (2014). A robust test for autocorrelation in the presence of a structural break in variance. *Journal of Statistical Computation and Simulation*, *84*(7), 1552–1562.

Pagan, A. (1984). Econometric issues in the analysis of regressions with generated regressors. *International Economic Review*, *25*(1), 221–247.

Pagan, A. (1986). Two stage and related estimators and their applications. *Review of Economic Studies*, 53(4), 517–538.

Shim, E., Moon, H., & Kim, T. (2006). *Autocorrelation tests robust to a structural break in variance*. Unpublished manuscript, Yonsei University, Seoul.

Wooldridge, J. M. (1991). On the application of robust, regression-based diagnostics to models of conditional means and conditional variances. *Journal of Econometrics*, 47(1), 5–46.

Yasmin, B. (2009). Trade liberalization and the lead role of human capital and job attributes in wage determination: The case of Pakistan's labor market. *Lahore Journal of Economics*, 14(1), 1–37.