Econometric Theory, **25**, 2009, 1734–1753. Printed in the United States of America.
doi:10.1017/S0266466609990314

HETEROSKEDASTICITY-ROBUST TESTING FOR A FRACTIONAL UNIT ROOT

HSEIN KEW
Monash University

DAVID HARRIS
University of Melbourne

This paper shows how fractional unit root tests originally derived under stationarity can be made robust to heteroskedasticity. This is done by using existing tests nested in a regression framework and then implementing these tests using White's heteroskedasticity consistent standard errors (White, 1980). We show this approach is effective both asymptotically and in finite samples. We also provide some evidence on the asymptotic local power of different implementations of the tests, under both homoskedasticity and heteroskedasticity.

1. INTRODUCTION

This paper considers testing for a unit root against a fractionally integrated alternative in the presence of heteroskedasticity. Specifically we consider the parametric time series model

$$\alpha(L)\Delta^{d_0} y_t = \varepsilon_t, \tag{1}$$

for $t = 1, \ldots, T$, where ε_t is a zero mean serially uncorrelated disturbance and $\alpha(L)$ is a stable lag polynomial of fixed order p and

$$\Delta^d = \sum_{j=0}^{\infty} \pi_j(d) L^j, \qquad \pi_j(d) = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)}. \tag{2}$$

Presample values are assumed to be zero, and so this is a "Type 2" model of fractional integration (Marinucci and Robinson, 1999). We aim to test

$$H_0 : d_0 = 1, \qquad H_1 : d_0 < 1. \tag{3}$$

This paper is a development of the second chapter of the first author's Ph.D. thesis. We are grateful to the co-editor Robert Taylor and two referees for very helpful comments on an earlier version of this paper and to participants at the Conference in Honour of Paul Newbold at the University of Nottingham and in seminars at the University of Liverpool, the University of Cambridge, Monash University, and the Singapore Econometric Study Group at Singapore Management University for helpful comments. Address correspondence to David Harris, Department of Economics, University of Melbourne, Victoria 3010, Australia. e-mail: harrisd@unimelb.edu.au.

 © 2009 Cambridge University Press 0266-4666/09 $15.00

https://doi.org/10.1017/S0266466609990314 Published online by Cambridge University Press

In homoskedastic settings this problem has been extensively analyzed, first by Robinson (1991, 1994), who suggested a Lagrange multiplier (LM) approach that was further developed by Agiakloglou and Newbold (1994), Tanaka (1999), Breitung and Hassler (2002), and Demetrescu, Kuzin, and Hassler (2008) (henceforth DKH). Lobato and Velasco (2007) (henceforth LV) recently introduced a Wald-type test, and Dolado, Gonzalo, and Mayoral (2002) (henceforth DGM) and Lobato and Velasco (2006) considered an alternative "fractional Dickey–Fuller" approach based on an extension of an analogy to unit root testing.

It has recently been recognized that unconditional heteroskedasticity is a feature of many economic time series. For example, Watson (1999) found evidence of variance shifts in U.S. long-term interest rates, and McConnell and Perez-Quiros (2000) found similar behavior in U.S. gross domestic product (GDP), as did van Dijk, Osborn, and Sensier (2002) in various G7 business cycle variables. The literature on testing for unit roots against $I(0)$ alternatives has begun to account for this issue; see, for example, Hamori and Tokihisa (1997), Kim, Leybourne, and Newbold (2002), Cavaliere (2004), Cavaliere and Taylor (2007, 2008a, 2008b), and Beare (2007). However in the literature on fractional unit root testing, only Robinson (1991) and DKH have allowed for conditional heteroskedasticity in ε_t, with the remainder of the literature assuming that ε_t is independent and identically distributed (i.i.d.). Robinson (1991) and DKH proposed LM tests that allow for conditional heteroskedasticity by using heteroskedasticity-consistent variance estimators, as in White (1980). Whereas conditional heteroskedasticity means that the conditional variance $\mathrm{E}(\varepsilon_t^2|\mathcal{F}_{t-1})$ is not constant, both Robinson (1991) and DKH assumed that the unconditional variance $\mathrm{E}(\varepsilon_t^2)$ is constant over time, so that unconditional heteroskedasticity is ruled out. We remove this latter restriction in this paper.

The paper proceeds as follows. In Section 2 we provide a brief overview of existing fractional unit root tests by nesting them all within a regression framework and showing that each test in the literature arises as a certain choice of weights in defining a regressor in the testing equation. This overview is useful in understanding the relations between the various tests. Also in Section 2 we define a class of fractional unit root tests within the regression framework making use of the White (1980) heteroskedasticity-robust standard error. This class of tests includes the test of DKH as a special case and also provides heteroskedasticity-robust versions of the Wald-type tests of DGM and LV. The main theoretical contribution of the paper is in Section 3, where we show that these tests are asymptotically valid in the presence of both conditional and unconditional heteroskedasticity, whereas tests based on ordinary least squares (OLS) standard errors suffer from asymptotic size distortions. The asymptotic distributions of the tests under local alternatives are also derived, and some implications for local power are discussed. Section 4 provides Monte Carlo evidence to demonstrate the effectiveness of the heteroskedasticity correction in this context. Section 5 concludes. Mathematical proofs are given in the Appendix of Kew and Harris (2008).

https://doi.org/10.1017/S0266466609990314 Published online by Cambridge University Press

2. FRACTIONAL UNIT ROOT TESTS

A class of regression-based tests for (3) can be defined using the equation[1]

$$\Delta y_t = \phi \hat{z}_{t-1}(\hat{d}) + \sum_{j=1}^{p} \varphi_j \Delta y_{t-j} + e_t, \tag{4}$$

where

$$\hat{z}_{t-1}(\hat{d}) = \hat{\beta}(L;\hat{d})\Delta y_t \tag{5}$$

and

$$\hat{\beta}(L;\hat{d}) = \sum_{j=1}^{t-1} \hat{\beta}_j(\hat{d})L^j, \tag{6}$$

for some potentially sample-dependent constants $\hat{d}$ and $\hat{\beta}_1(\hat{d}), \ldots, \hat{\beta}_{T-1}(\hat{d})$.[2] Various existing tests in the literature correspond to differing choices of these constants, which we explain subsequently.

The test statistic for (3) is the t-statistic constructed from $\hat{\phi}$, the OLS estimator in equation (4). Because the alternative in (3) is lower tailed, the null hypothesis is rejected for small values of the test statistic. The t-statistic can be constructed using either OLS or White's standard errors (White, 1980); that is,

$$t_\phi^{OLS} = \frac{\hat{\phi}}{\hat{\omega}_{\hat{\phi}}^{OLS}}, \qquad t_\phi = \frac{\hat{\phi}}{\hat{\omega}_{\hat{\phi}}},$$

where $\hat{\omega}_{\hat{\phi}}^{OLS}$ and $\hat{\omega}_{\hat{\phi}}$ are, respectively, the OLS and White's standard errors for $\hat{\phi}$,

$$\hat{\omega}_{\hat{\phi}}^{OLS} = \frac{\hat{\sigma}}{\sqrt{\sum_{t=1}^{T} \hat{z}_{t-1}^*(\hat{d})^2}}, \qquad \hat{\omega}_{\hat{\phi}} = \frac{\sqrt{\sum_{t=1}^{T} \hat{e}_t^2 \hat{z}_{t-1}^*(\hat{d})^2}}{\sum_{t=1}^{T} \hat{z}_{t-1}^*(\hat{d})^2},$$

and $\hat{z}_{t-1}^*(\hat{d})$ are OLS residuals from a regression of $\hat{z}_{t-1}(\hat{d})$ on $\Delta y_{t-1}, \ldots, \Delta y_{t-p}$, whereas $\hat{e}_t$ are the OLS residuals from (4) and $\hat{\sigma}^2 = T^{-1}\sum_{t=1}^{T} \hat{e}_t^2$. White's results were initially for exogenous regressors, but Nicholls and Pagan (1983) and more recently Phillips and Xu (2006) have shown that the approach is valid for autoregressive (AR) models. Although (4) is not an autoregression, we now show that this approach gives heteroskedasticity-robust inference for (3).

We focus on the hypotheses in (3) because these have been the subject of recent research. However it is straightforward to adapt our analysis to an alternative of the form $H_1' : d_0 > 1$ by rejecting the null hypothesis for large values of the t-statistic in (4). A two-sided alternative $H_1'' : d_0 \neq 1$ is similarly handled. Also we could consider a null hypothesis of the form $H_0^\delta : d_0 = \delta$ by replacing Δy_t in (4) and (5) by $\Delta^\delta y_t$.

https://doi.org/10.1017/S0266466609990314 Published online by Cambridge University Press

To discuss various specifications of $\hat{\beta}(L;\hat{d})$ it is convenient to define the lag operator $\lambda(L;d)$ to be the truncated version of $(1-\Delta^{d-1})/(d-1)$

$$\lambda(L;d)=\sum_{j=1}^{t-1}\lambda_j(d)L^j, \tag{7}$$

where $\lambda_1(d)=1$ and $\lambda_j(d)=\prod_{i=2}^{j}(i-d)/i$ for $j=2,3,\ldots$. Note that $\lambda_j(1)=j^{-1}$. We also define a lag polynomial $\hat{\alpha}(L;d)$ for any d to be the estimated AR lag polynomial from a regression of $\Delta^d y_t$ on $\Delta^d y_{t-1},\ldots,\Delta^d y_{t-p}$.

Robinson (1991) initially provided an LM test of $H_0: d_0=1$ against the two-sided alternative $H_1'': d_0\neq 1$ where $p=0$. This resulted in a choice of $\hat{\beta}_j(\hat{d})=j^{-1}=\lambda_j(1)$, and his test can be seen to be asymptotically equivalent to the t-test obtained from (4) with the lagged differences of Δy_t omitted. Although Robinson assumed throughout that ε_t is strictly stationary and ergodic, he gave a version of his test under the assumption of conditional homoskedasticity and another allowing for conditional heteroskedasticity. Although his tests were not expressed as regression-based tests, these two versions of the test are asymptotically equivalent to using OLS and White's standard errors (White, 1980), respectively, to compute the t-statistic from (4). Robinson (1994) showed that short-run dynamics, not necessarily of AR(p) form, can be allowed in his test.

Robinson's LM test has been written as a regression-based test (nested in equation (4)) by Agiakloglou and Newbold (1994)[3] and Breitung and Hassler (2002).[4] In the AR(p) case, this amounts to choosing $\hat{\beta}(L;\hat{d})=\hat{\alpha}(L;1)\lambda(L;1)$.[5] DKH proposed a one-step variation using $\hat{\beta}(L;\hat{d})=\lambda(L;1)$. DKH assumed stationarity of ε_t but allowed for conditional heteroskedasticity using White's standard errors. DKH also discussed the choice of p when the AR(p) model is an approximation to a more general linear process.

LV suggested a Wald-type approach by choosing $\hat{\beta}(L;\hat{d})=\hat{\alpha}(L;\hat{d})\lambda(L;\hat{d})$, where $\hat{d}$ is some preliminary consistent estimator of d_0. This is a Wald-type approach in the sense that an unrestricted estimator $\hat{d}$ is used to define $\hat{\beta}(L;\hat{d})$, as opposed to the null value of $\hat{d}=1$ used in the LM test. LV conjectured that using $\hat{d}$ rather than 1 may provide superior finite-sample power. All results assume that ε_t is i.i.d., and hence heteroskedasticity has not been considered.

DGM suggested the fractional Dickey–Fuller test, based on an ad hoc choice of $\hat{\beta}(L;d_1)=\Delta^{d_1-1}L$ for some fixed d_1. They suggested a particular trimmed estimator for d_1 that we denote as $\hat{d}_1$. Lobato and Velasco (2006) further considered the choice of d_1. For $p=0$ they showed that the optimal choice of d_1 when testing against local alternatives is $d_1=0.69145$. For the case of a fixed alternative they suggested $d_1=-0.031+0.719\hat{d}$, where $\hat{d}$ is a consistent estimator of d_0. Again all results relating to the fractional Dickey–Fuller test assume that ε_t is i.i.d.

For convenience, Table 1 summarizes the differing choices of $\hat{\beta}(L;\hat{d})$ for the various existing regression-based LM and Wald tests described earlier.

Various authors have derived asymptotic local power results for these tests. To briefly summarize, asymptotic efficiency under normality has been proved for the

https://doi.org/10.1017/S0266466609990314 Published online by Cambridge University Press

TABLE 1. Comparison of different choices of $\hat{\beta}(L;\hat{d})$ in the regression-based tests

	$\hat{\beta}(L;\hat{d})$	
	$p=0$	$p\geq 1$
Regression-based LM tests		
Agiakloglou and Newbold (1994) and Breitung and Hassler (2002)	$\lambda(L;1)$	$\hat{\alpha}(L;1)\lambda(L;1)$
Demetrescu, Kuzin, and Hassler (2008)	$\lambda(L;1)$	$\lambda(L;1)$
Regression-based Wald tests		
Dolado, Gonzalo, and Mayoral (2002)	$\Delta^{\hat{d}_1-1}L$	$\Delta^{\hat{d}_1-1}L$
Lobato and Velasco (2007)	$\lambda(L;\hat{d})$	$\hat{\alpha}(L;\hat{d})\lambda(L;\hat{d})$

LM test by Robinson (1994) and Tanaka (1999) and for the Wald test by LV. The fractional Dickey–Fuller tests (for any choice of d_1) and the DKH test are asymptotically inefficient. Unless there is compelling simulation evidence, we would therefore prefer the LM or Wald approach in practice. We provide asymptotic local power results subsequently that hold for all tests based on (4) under both conditional and unconditional heteroskedasticity.

3. ASYMPTOTIC DISTRIBUTION THEORY

In this section we consider using the test equation (4) to test (3) in the model (1) where now ε_t satisfies the following assumption.

Assumption E. Define $\mathcal{F}_t=\sigma(\varepsilon_1,\ldots,\varepsilon_t)$ and let ε_t satisfy

(a) $\mathrm{E}(\varepsilon_t|\mathcal{F}_{t-1})=0$;
(b) $\mathrm{E}\left(\varepsilon_t^2\right)=\sigma_t^2$ where $0<\underline{\sigma}\leq\sigma_t\leq\bar{\sigma}<\infty$;
(c) $\|\mathrm{E}\left(\varepsilon_t^2-\sigma_t^2|\mathcal{F}_{t-s}\right)\|_4\leq c_t\psi_s$ where $\psi_s\to 0$ and $c_t\leq\bar{c}<\infty$;
(d) $0<\mathrm{E}\left(\varepsilon_t^2\varepsilon_{t+j}^2\right)$ uniformly in t and j;
(e) for any integer q, $2\leq q\leq 8$, and for q nonnegative integers s_i, $\mathrm{E}\left(\prod_{i=1}^q\varepsilon_{t_i}^{s_i}\right)=0$ when at least one s_i is exactly one and $\sum_{i=1}^q s_i\leq 8$;
(f) the fourth cumulant $\kappa(t,t+i,t+j,t+k)$ of $\left(\varepsilon_t^2-\sigma_t^2,\varepsilon_{t+i}^2-\sigma_{t+i}^2,\varepsilon_{t+j}^2-\sigma_{t+j}^2,\varepsilon_{t+k}^2-\sigma_{t+k}^2\right)$ satisfies $\sup_t\sum_{i,j,k=-\infty}^{\infty}\kappa(t,t+i,t+j,t+k)<\infty$.

Assumption E relaxes the i.i.d. condition by using Assumption E(a) to replace independence and allowing unconditional heteroskedasticity in Assumption E(b). The unconditional variance process σ_t^2 allowed in (b) is essentially unrestricted, except for ruling out explosive (by $\bar{\sigma}<\infty$) and degenerate (by $\underline{\sigma}>0$) cases. As in Cavaliere (2004) and Xu and Phillips (2008), for example, we can allow

https://doi.org/10.1017/S0266466609990314 Published online by Cambridge University Press

a trend in the variance by allowing the variance to have an array structure of the form $\sigma_t^2 = \omega_1^2 + (\omega_2^2 - \omega_1^2)(t/T)^r$. However the cost of the relaxation of the i.i.d. assumption is the need to impose additional higher moment conditions E(c), (d), (e) that are not required by, for example, DGM, LV, or Robinson (1991). Assumption E(c) requires that ε_t^2 is an L_4 mixingale, which limits the memory allowed in the conditional variance, although we impose no condition on the rate at which ψ_s must disappear. Note that $\bar{c} < \infty$ implies that we are requiring that ε_t have uniformly bounded eighth moment. Assumption E(d) is a nondegeneracy requirement on $\varepsilon_t^2 \varepsilon_{t+j}^2$. Assumption E(e) is from Condition A(vii) of Deo (2000, p. 293) and extends up to eight moments the condition (11) of Robinson (1991), which applies only to fourth moments. Assumption E(f) controls the extent of the nonnormality in the joint distribution of $\varepsilon_t^2 - \sigma_t^2$ and its lags. Most likely these latter two conditions could be replaced with a mixing and moment condition, a mixingale-type assumption like E(c) for higher moments, or a summability condition on the eighth cumulants of ε_t (such as a version of Brillinger, 1981, Assump. 2.6.1), all at the cost of greater complication in the proofs. We use them as simplifying assumptions to allow us to focus on the task of allowing for nonconstant second moments.

To state the asymptotic properties of t_ϕ we require some structure to be placed on the choice of $\hat{\beta}(L; \hat{d})$ in (6). We impose the following assumption.

Assumption B. The choices of $\hat{d}$, d_1, and $\hat{\beta}(L; d)$ are such that

$$\sum_{j=1}^{T} (\hat{\beta}_j(\hat{d}) - \beta_j(d_1))^2 \xrightarrow{p} 0, \tag{8}$$

where

$$\hat{d} \xrightarrow{p} d_1, \qquad 0 < \sum_{j=1}^{\infty} \beta_j(d_1)^2 < \infty. \tag{9}$$

This is a high-level condition that can be shown to hold for the various practical implementations of the test that have been considered in the literature; see, for example, Lemma 1 in the Appendix of Kew and Harris (2008). Note that we use d_1 in (8) instead of the true value d_0 so that we can allow for the use of inconsistent choices of $\hat{d}$ in the fractional Dickey–Fuller test.

Theorem 1, which follows, establishes the asymptotic properties of t_ϕ for the data generating process (1) under the sequence of local alternatives $H_{1,T} : d_0 = 1 - \delta T^{-1/2}$ for some $\delta > 0$. To express the noncentrality parameter for t_ϕ we define the notation

$$u_t = \alpha(L)^{-1} \varepsilon_t, \tag{10}$$

where the stability of $\alpha(L)$ guarantees that $\alpha(L)^{-1}$ is well defined, and

$$z_{t-1}(d_1) = \beta(L; d_1) u_t, \tag{11}$$

https://doi.org/10.1017/S0266466609990314 Published online by Cambridge University Press

where $\beta(L;d_1) = \sum_{j=1}^{t-1} \beta_j(d_1)L^j$ and $\beta_j(d_1)$ is defined in (8). We then define $U_{p,t-1} = (u_{t-1}, \ldots, u_{t-p})'$, $\Gamma_T = T^{-1}\sum_{t=1}^{T} \mathrm{E}(U_{p,t-1}U'_{p,t-1})$, and $\gamma_T = T^{-1}\sum_{t=1}^{T} \mathrm{E}(U_{p,t-1}z_{t-1}(d_1))$ and

$$z^*_{t-1}(d_1) = z_{t-1}(d_1) - U'_{p,t-1}\Gamma_T^{-1}\gamma_T. \tag{12}$$

THEOREM 1. *Let the data generating process be given by (1) under* $H_{1,T}$: $d_0 = 1 - \delta T^{-1/2}$ *where* ε_t *satisfies Assumption E and* $\hat{\beta}(L;\hat{d})$ *satisfies Assumption B. Then the OLS t-statistic satisfies*

$$v_T^{-1} t_\phi^{OLS} + \delta\mu_T \xrightarrow{d} N(0,1),$$

and the White (1980) t-statistic satisfies

$$t_\phi + \delta\mu_T \xrightarrow{d} N(0,1),$$

where

$$\mu_T = \frac{T^{-1}\sum_{t=1}^{T} \mathrm{E}(\lambda(L;1)\varepsilon_t \cdot z^*_{t-1}(d_1))}{\sqrt{T^{-1}\sum_{t=1}^{T} \mathrm{E}(\varepsilon_t^2 \cdot z^*_{t-1}(d_1)^2)}}$$

and

$$v_T^2 = \frac{T^{-1}\sum_{t=1}^{T} \mathrm{E}\left(\varepsilon_t^2 \cdot z^*_{t-1}(d_1)^2\right)}{T^{-1}\sum_{t=1}^{T} \mathrm{E}\left(\varepsilon_t^2\right) \cdot T^{-1}\sum_{t=1}^{T} \mathrm{E}(z^*_{t-1}(d_1)^2)}$$

satisfy $0 < \mu_T < \infty$ *and* $0 < v_T^2 < \infty$ *uniformly in* T.

In Theorem 1 we show that, under the null hypothesis (i.e., when $\delta = 0$), the White (1980) t-statistic $t_\phi \xrightarrow{d} N(0,1)$. Therefore, as expected, the use of the White's standard errors in t_ϕ produces a test that is asymptotically robust to heteroskedasticity. Under local alternatives, μ_T is neither degenerate nor divergent because Assumption E bounds the variances and fourth moments of ε_t above and below, and the coefficients of $\beta(L;d_1)$ are assumed nondegenerate and square summable in (9).

The results of Theorem 1 can be used to show that the t_ϕ test with $p = 0$ and $\hat{\beta}_j(\hat{d}) = j^{-1}$ is asymptotically equivalent to the robustified LM statistic for strong serial correlation given in Robinson (1991, p. 74).

3.1. Asymptotic Size Distortion of the OLS Test

In Theorem 1 we give the asymptotic distribution of the OLS t-statistic. The additional term v_T can be seen to arise from the relationship $t_\phi = \left(\hat{\omega}_{\hat{\phi}}^{OLS}/\hat{\omega}_{\hat{\phi}}\right)t_\phi^{OLS}$; that is, we can show $\hat{\omega}_{\hat{\phi}}^{OLS}/\hat{\omega}_{\hat{\phi}} - v_T^{-1} \xrightarrow{p} 0$. In general v_T does not converge to 1

https://doi.org/10.1017/S0266466609990314 Published online by Cambridge University Press

under heteroskedasticity so that the OLS t-statistic, when $\delta = 0$, does not have an asymptotic null $N(0,1)$ distribution and hence will suffer from asymptotic size distortions. To illustrate this we provide the following specific result under the empirically relevant assumption that the variance of ε_t contains a single variance shift.

COROLLARY 1. *Under the conditions of Theorem 1, suppose that*

(i) ε_t *is independently distributed across all* t,
(ii) $p = 0$ *with* $\beta(L;d) = \lambda(L;1)$,
(iii) $\sigma_t^2 = \omega_1^2$ *for* $t \leq \lfloor \tau T \rfloor$ *and* $\sigma_t^2 = \omega_2^2$ *for* $t > \lfloor \tau T \rfloor$ *for some* $\tau \in (0,1)$.

Then

$$\nu_T^2 \to \nu(\tau,\eta)^2 = \frac{\tau + \eta^2(1-\tau)}{(\tau + \eta(1-\tau))^2},$$

where $\eta = \omega_2^2/\omega_1^2$ *and hence under* H_0 *(i.e.,* $\delta = 0$*)*

$$t_\phi^{OLS} \xrightarrow{d} N(0, \nu(\tau,\eta)^2).$$

The parameter η measures the magnitude or size of the variance shift. Note that when $\eta = 1$, the error ε_t is homoskedastic and so the asymptotic variance $\nu(\tau,\eta)^2 = 1$. This means that under H_0, $t_\phi^{OLS} \xrightarrow{d} N(0,1)$, which is a well-established result in the fractional unit root literature. Conversely, when $\eta \neq 1$ and for some $\tau \in (0,1)$, the error ε_t is heteroskedastic, and so $\nu(\tau,\eta)^2 \neq 1$. Consequently, t_ϕ^{OLS} does not have an asymptotic null $N(0,1)$ distribution, and thus critical values obtained from the $N(0,1)$ distribution will result in misleading statistical inferences. Note that $\nu(\tau,\eta)^2 > 1$ for $\eta \neq 1$ and $\tau \in (0,1)$ and thus the t_ϕ^{OLS} test will tend to overreject the unit root null hypothesis, even asymptotically. For a fixed value of τ, the asymptotic variance $\nu(\tau,\eta)^2$ is strictly increasing in η when $\eta > 1$ (an upward variance shift) and is strictly decreasing in η when $\eta < 1$ (a downward variance shift), which implies that as the magnitude of the variance shift increases, the variance $\nu(\tau,\eta)^2$ will move further away from 1, inducing greater size distortions. Also note that there is the symmetry property $\nu(\tau,\eta)^2 = \nu((1-\tau),1/\eta)^2$, which implies that (for $\tau < 0.5$, say) an early upward variance shift of magnitude η gives the same asymptotic size distortion as a late downward shift of magnitude $1/\eta$. To provide some indication of the possible severity of the asymptotic size distortions of the t_ϕ^{OLS} test, Figure 1 plots the asymptotic size at the 5% significance level against $\tau \in (0,1)$ for different values of η. Without loss of generality we set $\omega_1^2 = 1$ and choose $\eta = 5, 9$. These choices of η are standard in the literature; see, for example, Cavaliere and Taylor (2007). When $\eta = 9$ and as τ moves from 0 to 1, the asymptotic size of the t_ϕ^{OLS} test moves further away from 5% and achieves a maximum[6] at $\tau = 9/(9+1) = 0.9$.

https://doi.org/10.1017/S0266466609990314 Published online by Cambridge University Press

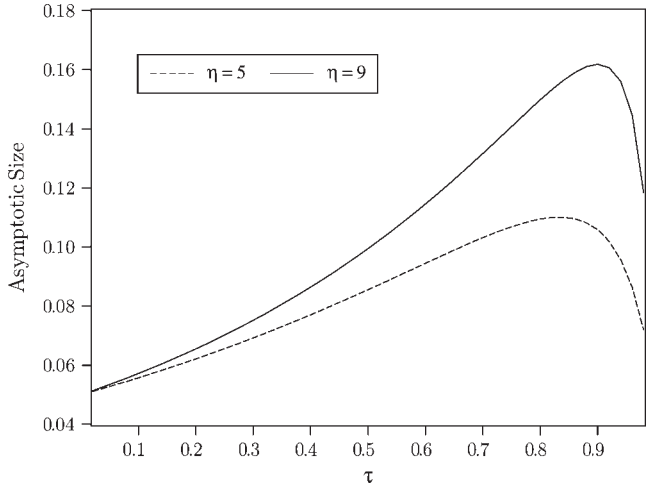


FIGURE 1. Asymptotic size of the OLS t-statistic under a single variance shift model.

Therefore substantial size distortions are observed in Figure 1 when there is a late upward shift in the variance. This phenomenon is also reflected in the Monte Carlo simulation results reported in Table 2 of Section 4.

3.2. Asymptotic Inefficiency of the White Test

We can use the results of Theorem 1 to conclude that the t_ϕ tests, although asymptotically robust to heteroskedasticity, are not asymptotically efficient. Under heteroskedasticity, the asymptotically efficient (albeit infeasible) test is obtained by computing t_ϕ (or t_ϕ^{OLS}) using the time series y_t *weighted* by its conditional variance, and using $\beta(L;d) = \alpha(L;1)\lambda(L;1)$. This follows immediately from the optimality results of Robinson (1994) and LV. It also follows from their results that the asymptotic distribution of this infeasible test statistic, denoted t_ϕ^*, under local alternatives would be

$$t_\phi^* + \delta\sqrt{\frac{\pi^2}{6}} \xrightarrow{d} N(0,1). \tag{13}$$

That is, the noncentrality parameter of the efficient test is $\sqrt{\pi^2/6}$, and any test with a different noncentrality parameter is not efficient. Straightforward but tedious derivations show that μ_T in Theorem 1 cannot equal $\sqrt{\pi^2/6}$ under heteroskedasticity, implying that a choice of $\beta(L;d)$ cannot be made to achieve the same effect as appropriately weighting the observations. For example, if ε_t is independent but heteroskedastic and we use the LM test with $\beta(L;d) = \lambda(L;1)$ and $p = 0$, then $z_{t-1}^*(d) = \lambda(L;1)\varepsilon_t$, giving $\mathrm{E}(\lambda(L;1)\varepsilon_t \cdot z_{t-1}^*(d)) = \sum_{j=1}^{t-1} j^{-2}\sigma_{t-j}^2$

https://doi.org/10.1017/S0266466609990314 Published online by Cambridge University Press

and $\mathrm{E}(\varepsilon_t^2 \cdot z_{t-1}^*(d)^2) = \sum_{j=1}^{t-1} j^{-2}\sigma_t^2\sigma_{t-j}^2$. Thus

$$\mu_T = \frac{\sum_{j=1}^{T-1} j^{-2}\left(T^{-1}\sum_{t=1}^{T-j}\sigma_t^2\right)}{\sqrt{\sum_{j=1}^{T-1} j^{-2}\left(T^{-1}\sum_{t=1}^{T-j}\sigma_t^2\sigma_{t+j}^2\right)}}.$$

This does not reduce to $\sqrt{\sum_{j=1}^{T-1} j^{-2}}\left(\text{which converges to}\sqrt{\pi^2/6}\right)$ unless σ_t^2 is constant for all t. The loss of asymptotic local power of t_ϕ relative to t_ϕ^* will vary depending on the pattern of heteroskedasticity. We next explore this for a specific example.

In the heteroskedastic case, Theorem 1 shows that the asymptotic local power of the t_ϕ test is dependent not only on δ and the choice of $\beta(L;d)$ (as in the i.i.d. case) but is also dependent on the variances and the fourth moments of ε_t, which makes interpretation difficult. To give some illustration of the possible effect of unconditional heteroskedasticity on the asymptotic local power of the test, we give the following result based on the same single variance shift model as in Corollary 1.

COROLLARY 2. *Under the conditions of Corollary 1,*

$$t_\phi \xrightarrow{d} N\left(-\delta\sqrt{\frac{\pi^2}{6}}\nu(\tau,\eta)^{-1}, 1\right). \tag{14}$$

The effect of a variance shift on the asymptotic local power of the t_ϕ test is captured by $\nu(\tau,\eta)^{-1}$ in the mean of the limiting normal distribution. This $\nu(\tau,\eta)$ term is the same term that appeared in Corollary 1 in the asymptotic variance of the t_ϕ^{OLS} statistic. We can conclude from this that a variance shift that causes large size distortions for the t_ϕ^{OLS} test will result in large power losses for the t_ϕ test as reflected in the Monte Carlo results reported in Table 2 of Section 4. In particular we can see that asymptotic local power is nonnegligible (because $\nu(\tau,\eta)^{-1} > 0$), that heteroskedasticity causes power losses (because $\nu(\tau,\eta)^{-1} < 1$), and that the power losses become larger for larger variance shifts (because $\nu(\tau,\eta)^{-1}$ is decreasing in $|\eta - 1|$).

To quantify the possible power loss due to heteroskedasticity in a single practical situation, we note that the growth rate of U.S. real quarterly GDP has standard deviation 1.179 over the period March 1947 to December 1984 and standard deviation 0.489 for March 1985 to March 2007.[7] This implies $\eta = 0.172$, $\tau = 0.631$, and hence $\nu(\tau,\eta)^{-1} = 0.867$. In Figure 2, the asymptotic local power curve of the t_ϕ test implied by these values of η and τ is contrasted with the asymptotic local power curve in the homoskedastic case $\eta = 1$. The power loss caused by such a variance shift is not dramatic but can be as much as 10% in this case. Also Figure 2 shows the asymptotic local power curve for $\tau = 0.05$ and $\eta = 0.05$, to demonstrate how strongly a variance break can affect power. In general the worst

https://doi.org/10.1017/S0266466609990314 Published online by Cambridge University Press

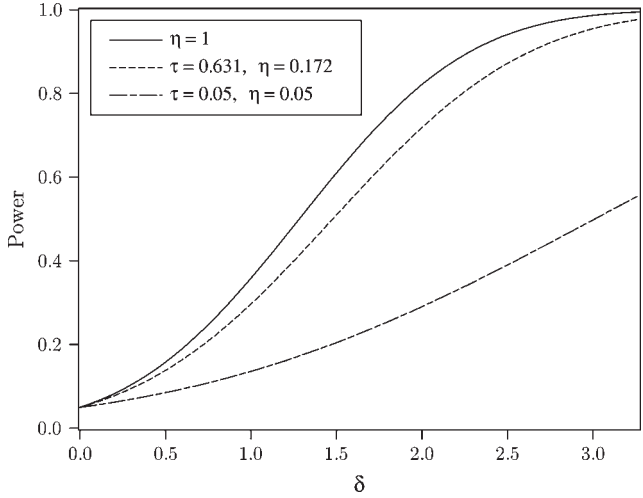


FIGURE 2. Asymptotic local power curves in the presence of a single variance shift model.

power results from a pattern of variance that has a small fraction of the sample exhibiting a very large variance relative to the remainder of the sample. The values $\tau = 0.05$ and $\eta = 0.05$ correspond to $\omega_1^2 = 20\omega_2^2$ with ω_1^2 being sustained for only 1/20th of the sample.

We can also interpret the power curves in Figure 2 as providing evidence on the power loss of the asymptotically inefficient t_ϕ test relative to the asymptotically efficient but infeasible weighted t_ϕ^* test (see eqn. (13)). The asymptotic local power curve for the t_ϕ^* test is the same as that shown in Figure 2 under homoskedasticity ($\eta = 1$). This curve therefore defines the asymptotic local power envelope, and hence the other two power curves (under heteroskedasticity) illustrate the power loss relative to this envelope.

We do not explore in this paper the possibility of estimating the variance function for the purpose of making the asymptotically efficient t_ϕ^* test feasible, though this is an obvious topic for further research that is currently under consideration. As kindly pointed out by a referee, the consideration of estimating the variance function raises a number of interesting questions, such as the relative performance of parametric versus nonparametric variance estimation and the effect of such estimation on the properties (asymptotic and finite samples) of the resulting feasible tests. For example, as proposed in general terms by Wooldridge (1991), one could specify a parametric model for the variance (e.g., variance shift at unknown time and/or a generalized autoregressive conditionally heteroskedastic model) and estimate the parameters of this model jointly with the testing equation in (4) to develop a feasible weighted efficient test. Alternatively the variance process could be estimated nonparametrically by smoothing methods as Hansen (1995) and Xu and Phillips (2008) suggested for regression models. All of these possibilities are substantial topics that lie beyond the scope of the current paper.

https://doi.org/10.1017/S0266466609990314 Published online by Cambridge University Press

3.3. Local Power Properties of the DKH Test

If $p = 1$ then we find a difference in asymptotic local power between the two-step LM test, as initially implemented by Agiakloglou and Newbold (1994) and Breitung and Hassler (2002), and the one-step variant recently proposed by DKH. This latter one-step test is not the full LM test, and, as shown in the next corollary, it is not asymptotically efficient, even under homoskedasticity.

COROLLARY 3. *Under the conditions of Theorem 1, suppose* $\varepsilon_t \sim$ i.i.d. $(0, \sigma^2)$ *and* $p = 1$ *with* $\alpha(L) = 1 - \alpha L$, $\alpha \neq 0$, $|\alpha| < 1$. *Then*

(i) for the two-step LM test with $\beta(L; d) = \hat{\alpha}(L; 1)\lambda(L; 1)$

$$\mu_T \to \mu^{LM} = \sqrt{\frac{\pi^2}{6} - \frac{1-\alpha^2}{\alpha^2}(\ln(1-\alpha))^2};$$

(ii) for the one-step DKH test with $\beta(L; d) = \lambda(L; 1)$

$$\mu_T \to \mu^{DKH} = \frac{\frac{\pi^2}{6} + \mathrm{Li}_2(\alpha) + (\ln(1-\alpha))^2\left(\frac{1}{2} - \frac{1}{\alpha^2}\right)}{\sqrt{(1-\alpha^2)^{-1}\left(\frac{\pi^2}{6} + 2\mathrm{Li}_2(\alpha) + (\ln(1-\alpha))^2(1-\alpha^{-2})\right)}},$$

where $\mathrm{Li}_2(\alpha) = \sum_{j=1}^{\infty} j^{-2}\alpha^j$ *is the dilogarithm function.*

Because the two-step LM test is asymptotically efficient, it follows that $\mu^{DKH} \leq \mu^{LM}$ for all α, and the extent of the resulting power loss is illustrated in Figure 3.

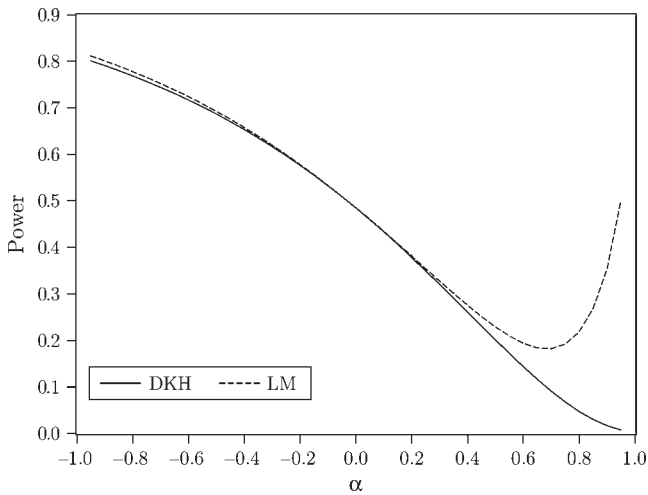


FIGURE 3. Asymptotic local power for DKH and LM tests with AR(1) dynamics and $\delta = 2$.

https://doi.org/10.1017/S0266466609990314 Published online by Cambridge University Press

The power loss is greatest for large positive values of α. This provides a strong motivation to prefer the two-step LM test over its more convenient but suboptimal one-step variant.

It should be noted that the results of Theorem 1 apply to the DKH test by setting $\beta(L;d) = \lambda(L;1)$, so that the DKH test with White's standard errors (White, 1980) is robust to unconditional heteroskedasticity. Moreover the expression given in that theorem for μ_T could be used to derive the local power for the test in the presence of any form of heteroskedasticity desired. The finite-sample properties of the DKH test in the presence of a variance shift can be found in Table 4 later in the paper. As predicted by the asymptotic theory, the DKH test is effective in removing the observed size distortions, but it displays poor finite-sample power properties compared to the two-step LM test when α is large ($\alpha = 0.8$).

4. SIMULATION EVIDENCE

This section examines, via Monte Carlo simulations, the effect of heteroskedasticity on the finite-sample properties of the regression-based tests outlined in Section 2. When an estimated value of d_0 is required, we follow DGM's suggestion and use the parametric estimator of Mayoral (2007).[8] Kew (2007) demonstrated that Mayoral's estimator remains $\sqrt{T}$-consistent even when the errors are heteroskedastic. Our simulation experiments thus reflect the feasible implementations of these tests in practice. The simulated data for y_t are generated according to (1), and we first consider $\alpha(L) = 1$ (no short-run dynamics) and then let $\alpha(L) = 1 - \alpha L$ to allow for the presence of first-order serial dependence. In each specification of $\alpha(L)$, d_0 is chosen from the interval $[0.6, 1]$ in steps of 0.05, $T = 100, 250$, and 500, and the number of replications is 10,000. We report the percentage of rejections at the 5% critical value of -1.645.

The heteroskedasticity-robust tests defined in Section 2 do not require a priori specification of a parametric functional form for σ_t^2. Thus, in our simulations, we consider two models for σ_t^2 :

1. Single variance shift model: $\sigma_t^2 = \omega_1^2 \mathbf{1}(t \leq \lfloor \tau T \rfloor) + \omega_2^2 \mathbf{1}(t > \lfloor \tau T \rfloor)$ with $\tau \in (0, 1)$.
2. Autoregressive conditionally heteroskedastic (ARCH) model: $\varepsilon_t = g_t v_t$, with $g_t^2 = \rho_0 + \rho_1 \varepsilon_{t-1}^2$ and where v_t is i.i.d. $N(0, 1)$.

Model 1 is taken from Kim et al. (2002) and Cavaliere and Taylor (2007). As in Corollary 1, we let $\eta = \omega_2^2/\omega_1^2$ measure the size of the variance shift. It follows that $\eta > 1$ implies an upward variance shift, and conversely. When $\eta = 1$, y_t is homoskedastic. Without loss of generality we set $\omega_1^2 = 1$. We only report simulation results for the case of an upward shift by setting $\eta = 9$, and this shift occurs either early in the sample with $\tau = 0.2$ or late in the sample with $\tau = 0.8$. The unconditional variance in Model 1 is not constant over time, but this is not true for

https://doi.org/10.1017/S0266466609990314 Published online by Cambridge University Press

Model 2. The latter model is taken from Engle (1982) and is only conditionally heteroskedastic. We consider $\rho_1 = 0.25, 0.5, 0.75$ and $\rho_0 = 1$. The choices of $\rho_1 = 0.5, 0.75$ do not satisfy our moment existence conditions in Assumption E, but these choices provide an interesting additional robustness check for the tests with White's standard error (White, 1980).

In Tables 2 and 3, we set $p = 0$ in (1) and (4). We computed five test statistics with White's standard errors (denoted t_ϕ) and OLS standard errors (denoted t_ϕ^{OLS}). They were as follows: (a) DGM: $\hat{\beta}(L;\hat{d}) = \Delta^{\hat{d}-1}L$;[9] (b) LV: $\hat{\beta}(L;\hat{d}) = \lambda(L;\hat{d})$; (c) LM: $\hat{\beta}(L;\hat{d}) = \lambda(L;1)$; (d) LV-a: $\hat{\beta}(L;\hat{d}) = \Delta^{d_1-1}L$ with $d_1 = 0.69145$; and (e) LV-b: $\hat{\beta}(L;\hat{d}) = \Delta^{d_1-1}L$ with $d_1 = -0.031 + 0.719\hat{d}$. Tables 2 and 3 display the empirical sizes (when $d_0 = 1$) of the t_ϕ^{OLS} and t_ϕ tests and also the empirical (raw) power of the t_ϕ tests.

We first discuss the finite-sample properties of the t_ϕ^{OLS} tests. Whereas the first panel of Table 2 assumes homoskedasticity ($\eta = 1$), the second and third panels assume the presence of an upward shift in the variance process ($\eta = 9$). Under homoskedasticity, the t_ϕ^{OLS} tests have slightly better size properties than t_ϕ in all cases. In unreported simulations, it was found that the power loss resulting from the use of White's correction when it is not required is negligible even when $T = 100$. Although the t_ϕ^{OLS} tests have approximately correct size under homoskedasticity, they tend to overreject the unit root null hypothesis under heteroskedasticity, and substantial size distortions are observed when there is a late upward shift ($\tau = 0.8$) rather than an early upward shift ($\tau = 0.2$). These finite-sample results follow from the asymptotic property of the t_ϕ^{OLS} tests established in Corollary 1. In unreported simulations, we also considered the case of a downward shift with $\eta = 1/9$, and, as expected from Corollary 1, we found substantial size distortions when the shift happens early rather than late in the sample. Thus, we can conclude that the presence of a small burst of a very large variance relative to the rest of the sample can cause severe size distortions. In the ARCH model, Table 3 shows that the degree of overrejection increases as ρ_1 approaches 1 and hence suggests that the t_ϕ^{OLS} tests are not robust to the presence of conditional heteroskedasticity.

Looking at the t_ϕ tests, Tables 2 and 3 show that, in all cases, the t_ϕ tests can effectively reduce the observed size distortions displayed by the t_ϕ^{OLS} tests. For example, Table 2 shows that when $\eta = 9$ and $\tau = 0.8$, the White's correction reduces the empirical sizes of all the tests from about 15% to 6% for a 5% nominal size when $T = 500$. Turning to the raw power, Tables 2 and 3 show that when the t_ϕ^{OLS} tests suffer from substantial size distortions, the t_ϕ tests suffer from large power losses. This is expected from the asymptotic results given in Corollaries 1 and 2.

In Table 4, we allowed for the presence of first-order serial correlation by setting $\alpha(L) = 1 - \alpha L$ in (1) with $\alpha = \{0, 0.3, 0.8\}$ and $p = 1$ in (4). We computed the four test statistics listed in Table 1: (a) ADGM: $\hat{\beta}(L;\hat{d}) = \Delta^{\hat{d}-1}L$; (b) 2-LV: $\hat{\beta}(L;\hat{d}) = \hat{\alpha}(L;\hat{d})\lambda(L;\hat{d})$; (c) LM-AN&BH: $\hat{\beta}(L;\hat{d}) = \hat{\alpha}(L;1)\lambda(L;1)$;

https://doi.org/10.1017/S0266466609990314 Published online by Cambridge University Press

TABLE 2. Single variance shift model: Empirical size of t_ϕ^{OLS} and t_ϕ tests and empirical power of t_ϕ tests

				d_0									
				t_ϕ^{OLS}	t_ϕ								
τ	η	T		1	1	0.95	0.90	0.85	0.80	0.75	0.70	0.65	0.60
	1	100	DGM	5.81	6.37	15.32	29.35	47.69	65.56	80.48	90.80	96.21	98.40
			LV	6.09	6.78	17.00	34.43	58.12	78.25	92.09	97.80	99.49	99.88
			LM	5.99	6.64	16.84	34.08	56.98	76.26	90.59	97.06	99.27	99.77
			LV-a	5.62	6.07	15.91	32.10	52.17	71.32	85.62	93.76	97.91	99.16
			LV-b	6.05	6.56	16.88	34.00	56.55	75.91	90.47	96.78	99.20	99.67
		250	DGM	5.19	5.48	21.72	48.87	77.42	93.12	98.63	99.78	99.98	100.00
			LV	5.23	5.39	26.27	62.25	90.73	99.12	99.97	100.00	100.00	100.00
			LM	5.22	5.55	26.30	61.91	90.18	98.87	99.97	100.00	100.00	100.00
			LV-a	5.13	5.29	24.94	58.49	87.01	97.61	99.86	99.99	100.00	100.00
			LV-b	5.01	5.38	25.40	60.69	89.17	98.56	99.96	99.99	100.00	100.00
		500	DGM	5.29	5.40	31.80	72.80	95.06	99.65	99.99	100.00	100.00	100.00
			LV	5.24	5.37	40.84	88.82	99.52	100.00	100.00	100.00	100.00	100.00
			LM	5.43	5.47	40.77	88.52	99.47	100.00	100.00	100.00	100.00	100.00
			LV-a	5.16	5.28	39.00	85.67	99.16	99.98	100.00	100.00	100.00	100.00
			LV-b	5.27	5.33	39.80	86.99	99.32	100.00	100.00	100.00	100.00	100.00
0.2	9	100	DGM	7.29	6.49	14.83	27.28	43.88	59.56	75.26	87.04	93.58	97.12
			LV	7.89	7.03	15.92	31.32	51.94	71.47	87.46	95.53	98.78	99.55
			LM	7.56	6.99	15.97	30.91	51.00	69.58	85.37	94.24	98.12	99.34
			LV-a	7.14	6.30	15.08	28.59	46.92	64.27	79.61	90.02	95.34	97.98
			LV-b	7.67	6.97	15.87	30.98	50.98	69.99	85.39	94.39	98.10	99.18
		250	DGM	6.84	5.69	20.15	44.46	71.65	88.69	97.46	99.49	99.93	99.99
			LV	6.97	5.75	23.54	55.90	85.99	97.61	99.83	99.99	100.00	100.00
			LM	7.11	5.74	23.65	55.51	85.26	97.03	99.77	99.97	100.00	100.00
			LV-a	6.80	5.60	22.47	52.39	81.45	94.73	99.38	99.85	100.00	100.00
			LV-b	7.02	5.69	23.20	54.75	84.15	96.56	99.70	99.97	100.00	100.00
		500	DGM	6.86	5.53	28.80	67.32	92.05	99.19	99.92	100.00	100.00	100.00
			LV	7.13	5.64	36.60	83.05	98.72	99.99	100.00	100.00	100.00	100.00
			LM	7.12	5.50	36.60	82.51	98.53	99.99	100.00	100.00	100.00	100.00
			LV-a	6.99	5.33	35.34	79.45	97.67	99.93	100.00	100.00	100.00	100.00
			LV-b	6.96	5.55	36.17	81.10	98.24	99.98	100.00	100.00	100.00	100.00
0.8	9	100	DGM	16.40	8.68	14.68	22.83	33.17	44.05	55.41	67.48	76.45	83.87
			LV	15.61	8.58	14.96	24.64	37.98	49.43	63.24	75.33	84.73	90.93
			LM	15.76	8.46	14.91	23.80	36.28	47.39	60.30	72.40	81.63	88.26
			LV-a	15.28	8.11	13.92	22.19	32.86	43.56	54.40	66.44	75.18	82.84
			LV-b	15.66	8.67	14.87	24.48	37.16	48.63	61.96	74.06	82.99	89.73
		250	DGM	15.80	7.15	16.36	30.67	49.43	67.18	82.81	91.94	96.48	98.78
			LV	15.64	6.93	19.11	36.80	59.15	79.46	91.89	97.63	99.24	99.83
			LM	16.02	6.95	18.67	36.25	58.23	78.22	90.85	96.87	99.02	99.82
			LV-a	15.34	6.81	17.28	33.40	54.11	72.86	86.97	94.41	97.87	99.33
			LV-b	15.52	6.97	18.50	35.62	57.86	77.87	90.81	96.88	98.90	99.76
		500	DGM	15.59	6.50	20.00	44.57	70.60	88.16	96.67	99.13	99.83	100.00
			LV	15.71	6.50	24.87	55.53	84.02	96.65	99.59	99.98	99.99	100.00
			LM	15.94	6.54	24.66	54.99	83.38	96.16	99.52	99.95	99.99	100.00
			LV-a	15.49	6.40	23.00	51.31	79.58	94.10	99.03	99.83	99.98	100.00
			LV-b	15.61	6.64	24.48	54.11	82.43	95.82	99.53	99.92	99.99	100.00

Note: The t_ϕ^{OLS} and t_ϕ columns refer to the tests with OLS and White standard errors, respectively, $\tau \in (0, 1)$ determines the location of the shift, and η measures the size of the shift.

https://doi.org/10.1017/S0266466609990314 Published online by Cambridge University Press

TABLE 3. ARCH model: Empirical size and power of t_ϕ^{OLS} and t_ϕ tests and empirical power of t_ϕ tests

			d_0									
			t_ϕ^{OLS}	t_ϕ								
ρ_1	T		1	1	0.95	0.90	0.85	0.80	0.75	0.70	0.65	0.60
0.25	100	DGM	9.99	6.92	13.30	24.45	38.85	53.93	68.87	79.63	88.91	94.12
		LV	8.40	6.69	15.92	29.75	47.65	67.98	82.32	90.72	96.13	98.44
		LM	8.66	6.61	15.04	28.61	45.81	64.53	78.82	88.03	94.71	97.48
		LV-a	9.00	6.56	13.71	25.77	41.43	57.79	71.96	82.05	90.37	94.84
		LV-b	8.85	6.82	15.01	28.43	45.90	65.03	79.50	88.62	95.11	97.64
	250	DGM	9.31	5.66	18.13	38.68	62.53	81.55	92.96	98.09	99.39	99.83
		LV	7.98	5.88	22.54	52.48	79.60	94.02	98.84	99.79	99.87	99.99
		LM	8.46	5.89	22.43	50.82	77.74	92.83	98.32	99.65	99.83	99.99
		LV-a	8.78	5.47	20.45	46.19	71.69	88.86	96.74	99.16	99.70	99.91
		LV-b	8.47	5.74	21.46	49.55	76.33	92.14	98.17	99.59	99.84	99.99
	500	DGM	9.87	5.55	24.34	57.30	85.14	96.85	99.51	99.93	99.98	100.00
		LV	8.25	5.64	32.98	76.46	96.34	99.64	99.97	99.98	100.00	100.00
		LM	8.59	5.67	32.61	75.41	95.86	99.54	99.95	99.98	100.00	100.00
		LV-a	8.86	5.58	30.16	70.13	93.29	99.01	99.91	99.98	99.99	100.00
		LV-b	8.65	5.53	31.45	72.98	95.02	99.39	99.96	99.98	100.00	100.00
0.5	100	DGM	13.79	7.45	12.50	21.29	33.58	45.55	59.07	68.99	79.20	86.70
		LV	12.22	7.41	15.24	26.35	42.13	57.80	72.34	81.52	88.78	93.31
		LM	12.71	7.20	14.45	24.71	39.50	53.65	67.69	77.09	85.23	90.71
		LV-a	12.68	6.98	12.94	22.25	35.07	47.32	60.52	69.78	79.27	86.07
		LV-b	12.55	7.34	14.37	24.96	39.86	54.85	69.43	79.10	86.78	92.05
	250	DGM	15.42	6.28	15.57	30.02	49.33	67.10	81.60	90.14	94.80	97.54
		LV	12.80	6.74	19.20	41.88	66.73	83.33	92.12	96.49	98.06	99.00
		LM	13.39	6.56	18.79	39.98	63.70	80.39	90.27	95.18	97.34	98.60
		LV-a	13.99	6.36	17.37	35.12	56.94	73.81	85.99	92.28	95.71	97.78
		LV-b	13.56	6.62	18.38	39.20	62.74	79.83	90.52	95.53	97.51	98.75
	500	DGM	15.88	5.76	19.37	42.52	67.79	85.08	93.52	97.67	98.98	99.45
		LV	13.31	5.78	26.72	60.89	85.19	94.54	97.84	99.11	99.61	99.69
		LM	13.78	5.75	26.12	58.62	83.09	93.54	97.27	98.89	99.46	99.60
		LV-a	14.50	5.68	23.19	52.72	77.72	90.78	95.81	98.25	99.22	99.42
		LV-b	14.02	5.62	24.79	56.79	81.98	93.28	97.29	98.99	99.54	99.68
0.75	100	DGM	18.47	8.08	13.47	20.82	29.66	39.67	50.74	61.15	68.89	77.02
		LV	16.50	8.09	15.26	24.88	36.91	50.51	62.42	72.58	79.67	85.26
		LM	17.16	7.93	14.38	23.22	33.99	45.92	56.84	66.96	74.08	80.37
		LV-a	17.06	7.49	13.21	20.90	30.04	40.06	50.59	59.60	66.61	73.70
		LV-b	17.16	8.02	14.73	23.80	35.10	47.89	59.35	70.10	77.44	83.28
	250	DGM	22.30	7.56	15.01	25.52	39.33	53.21	66.46	76.81	84.31	89.75
		LV	19.67	7.31	18.08	33.67	53.49	68.50	79.59	87.20	91.13	94.08
		LM	20.55	7.38	17.30	31.80	50.03	64.30	75.36	83.41	88.31	91.86
		LV-a	20.98	7.16	15.97	28.19	43.68	57.13	69.00	77.81	84.28	89.02
		LV-b	20.44	7.28	17.18	31.29	49.94	65.24	76.75	84.99	90.04	93.39
	500	DGM	25.31	6.98	16.55	31.63	49.15	66.68	79.06	86.24	91.72	94.66
		LV	22.70	7.10	21.53	44.74	65.95	80.53	88.59	92.65	95.22	96.78
		LM	23.22	7.11	20.62	42.47	62.44	77.36	86.11	90.61	93.76	95.69
		LV-a	23.63	6.76	18.63	37.21	55.94	72.14	81.93	87.48	91.63	94.04
		LV-b	23.46	6.79	20.17	41.70	62.09	77.73	86.99	91.63	94.53	96.39

Note: The t_ϕ^{OLS} and t_ϕ columns refer to the tests with OLS and White standard errors, respectively, and ρ_1 is the coefficient of an ARCH(1) model.

https://doi.org/10.1017/S0266466609990314 Published online by Cambridge University Press

TABLE 4. Single variance shift ($\eta = 9, \tau = 0.8$) model with AR(1) errors: Empirical size and power of t_ϕ^{OLS} and t_ϕ tests and empirical power of t_ϕ tests

			d_0									
			t_ϕ^{OLS}	t_ϕ								
α	T		1	1	0.95	0.90	0.85	0.80	0.75	0.70	0.65	0.60
0	100	ADGM	16.39	9.18	13.33	20.08	26.20	33.83	42.57	51.35	60.25	67.34
		2-LV	13.46	7.80	11.63	17.96	24.33	32.21	41.32	50.49	60.28	68.10
		LM-AN&BH	17.68	9.79	13.54	19.59	26.52	32.44	41.12	48.50	57.40	64.56
		LM-DKH	17.61	9.52	13.62	19.65	26.75	32.93	42.51	49.88	59.18	66.98
	250	ADGM	15.87	7.62	14.45	23.21	36.19	50.65	64.37	76.95	86.31	92.60
		2-LV	13.97	6.58	13.16	22.27	35.66	50.52	65.19	78.67	88.24	94.28
		LM-AN&BH	16.56	7.55	14.50	24.32	35.88	50.40	64.48	75.92	85.72	91.77
		LM-DKH	16.48	7.40	14.40	24.56	36.40	51.33	65.98	77.65	87.32	93.27
	500	ADGM	15.75	6.35	16.66	31.89	53.14	72.43	87.31	94.99	98.22	99.63
		2-LV	14.98	6.07	16.26	32.25	54.21	74.34	89.36	96.36	99.00	99.75
		LM-AN&BH	16.67	7.15	16.69	32.66	53.72	74.03	88.47	95.63	98.71	99.75
		LM-DKH	16.71	7.11	16.61	33.00	54.22	74.94	89.67	96.39	99.06	99.79
0.3	100	ADGM	16.50	9.11	11.92	15.33	19.68	25.54	31.51	37.87	45.10	53.38
		2-LV	12.09	7.87	10.66	14.31	19.22	24.44	30.64	37.71	45.00	53.74
		LM-AN&BH	16.56	9.35	12.59	15.16	19.86	25.17	30.14	36.69	42.10	49.72
		LM-DKH	17.16	9.46	12.24	15.25	19.64	25.05	29.95	36.61	42.56	50.57
	250	ADGM	16.57	7.82	11.26	17.72	25.39	35.04	47.03	57.26	68.20	77.46
		2-LV	12.34	6.53	9.71	16.17	23.72	33.17	46.20	58.06	69.83	79.86
		LM-AN&BH	16.89	8.00	12.24	18.44	27.14	36.63	45.92	56.68	68.37	76.90
		LM-DKH	17.01	7.85	12.03	17.73	26.46	36.13	45.41	56.85	69.08	77.67
	500	ADGM	17.10	7.07	13.48	23.64	37.92	52.08	66.65	80.83	89.42	94.71
		2-LV	12.65	5.68	11.62	21.75	35.78	52.27	68.90	84.37	92.55	96.83
		LM-AN&BH	17.06	7.09	13.80	24.35	38.49	53.54	69.09	81.69	90.60	95.93
		LM-DKH	16.91	6.93	13.46	23.55	37.57	52.78	68.57	81.58	90.79	96.27
0.8	100	ADGM	15.99	8.92	7.85	7.37	6.51	6.38	6.66	6.72	7.64	8.96
		2-LV	8.33	4.17	5.89	7.88	10.33	13.33	16.72	20.33	24.33	28.06
		LM-AN&BH	9.78	5.71	6.70	7.26	8.11	8.72	9.52	10.30	11.11	12.32
		LM-DKH	15.58	8.86	7.96	7.65	6.93	6.81	7.16	7.17	8.10	9.24
	250	ADGM	16.49	7.40	6.11	5.28	4.92	4.55	4.62	5.31	6.30	8.18
		2-LV	11.79	4.72	7.24	10.34	14.49	20.23	26.25	31.98	37.45	42.44
		LM-AN&BH	11.71	5.14	7.04	8.95	10.82	12.59	13.74	14.46	15.33	15.94
		LM-DKH	16.06	7.51	6.43	5.79	5.65	5.33	5.50	6.11	7.05	8.37
	500	ADGM	16.70	6.74	5.26	4.40	3.91	3.80	3.87	4.72	6.60	10.15
		2-LV	14.04	5.38	8.89	14.30	21.09	30.16	38.48	46.12	53.10	59.95
		LM-AN&BH	13.03	5.06	8.06	12.13	15.35	18.56	19.87	20.35	21.03	22.18
		LM-DKH	16.37	6.75	5.71	5.01	4.63	4.74	4.95	5.68	7.00	8.87

Note: The t_ϕ^{OLS} and t_ϕ columns refer to the tests with OLS and White standard errors, respectively, and α is the coefficient of a first-order short-run dependence.

https://doi.org/10.1017/S0266466609990314 Published online by Cambridge University Press

and (d) LM-DKH: $\hat{\beta}(L;\hat{d}) = \lambda(L;1)$. The variance of ε_t is assumed to contain a late upward shift ($\eta = 9$ and $\tau = 0.9$). Table 4 shows again that the t_ϕ^{OLS} tests are unreliable but the t_ϕ tests have desirable size and power properties in finite samples. When $T = 100$, the t_ϕ tests are slightly oversized, but this problem rapidly disappears as T increases. For each sample size, the empirical power of the t_ϕ tests increases as d_0 moves away from 1 except when the errors display strong serial dependence. For example, when $\alpha = 0.8$ the power becomes nonmonotonic, and this is true even under homoskedasticity as reported in LV (Table 3, p. 587).

We have also carried out Monte Carlo simulations for the following heteroskedastic models:[10]

1. Two variance symmetric shifts model: $\sigma_t^2 = \omega_1^2 \mathbf{1}(t \le \lfloor \tau T \rfloor) + \omega_2^2 \mathbf{1}(\lfloor \tau T \rfloor < t \le 1 - \lfloor \tau T \rfloor) + \omega_1^2 \mathbf{1}(1 - \lfloor \tau T \rfloor < t \le T)$ with $\tau \in (0,1)$.
2. Linear trending variance model: $\sigma_t^2 = \omega_1^2 + (\omega_2^2 - \omega_1^2)(t/T)$.
3. Stochastic volatility model: $\varepsilon_t = \exp(0.5 g_t)\epsilon_t$, where $g_t = \vartheta g_{t-1} + \sigma_v \nu_t$ and ϵ_t and ν_t are independent sequences of i.i.d. $N(0,1)$.

We briefly discuss here the finite-sample performance of the t_ϕ^{OLS} tests. Just like the preceding single variance shift model we find that, depending on the size and the timing of the variance shifts, the two variance symmetric shifts model can cause severe size distortions. When σ_t^2 follows a linear trend, we find that the size distortions are less dramatic than those observed in Table 2 with empirical sizes of about 7%. For the stochastic volatility model, the degree of overrejection is quite small when $\vartheta < 1$ but can be quite substantial when $\vartheta = 1$ whereas σ_v is held fixed.[11] The t_ϕ tests, by contrast, are found to deliver good size properties in all of the heteroskedastic models considered.

5. CONCLUSION

In an LM-type framework, Robinson (1991) and DKH suggested test statistics that are robust to the presence of conditional heteroskedasticity, while assuming unconditional homoskedasticity. We extend their results by relaxing the latter assumption. In addition we show that their proposed method can be used to provide a robustified version of the Wald-type tests under both conditional and unconditional heteroskedasticity. Specifically, under heteroskedasticity, we demonstrate that the limiting null distributions of all the existing regression-based LM and Wald-type tests that use OLS standard errors suffer from asymptotic size distortions, but it is not the case for tests that use the White's standard errors. We also analyze in detail the asymptotic local power of these tests under a specific form of heteroskedasticity. Monte Carlo results show that White's correction is very effective in removing the observed size distortions and delivers tests with desirable size and power properties in finite samples.

https://doi.org/10.1017/S0266466609990314 Published online by Cambridge University Press

NOTES

1. Depending on the choice of $\hat{z}_{t-1}(\hat{d})$, φ_j and e_t in this testing equation may match α_j and ε_t, respectively.
2. The lag polynomial $\hat{\beta}(L;d)$ depends on t, and so more complete notation would be $\hat{\beta}_t(L;\hat{d})$. We drop the subscript t so as not to overburden the notation.
3. Agiakloglou and Newbold (1994) truncate the sum in $\hat{\beta}(L;\hat{d})$ at some fixed value m instead of $t-1$.
4. Tanaka (1999) also discussed the LM test for parametric autoregressive moving average short-run dynamics but not explicitly in the regression framework.
5. The notation reflects the choice of $\hat{d}=1$, consistent with the LM principle.
6. For a fixed value of η (other than 1), the asymptotic variance of the t_{ϕ}^{OLS} test, $\nu(\tau,\eta)$, achieves its maximum when $\tau=\eta/(\eta+1)$, and thus at this value of τ, the asymptotic size distortion is at its peak.
7. This is the type of variance shift found by McConnell and Perez-Quiros (2000).
8. Mayoral suggested that the number of autocorrelations, k, used in the estimation framework is chosen such that $k=T^{1/4}$.
9. The parameter c in the trimming rule was set to 0.02; see page 1977 of DGM for details.
10. To save space, simulation results for these models are not reported but are available on request.
11. The choice of $\vartheta=1$ violates the moment existence condition imposed by Assumption E.

REFERENCES

Agiakloglou, C. & P. Newbold (1994) Lagrange multiplier tests for fractional difference. *Journal of Time Series Analysis* 15, 253–262.

Beare, B. (2008) Unit Root Testing with Unstable Volatility. Yale University, mimeo.

Breitung, J. & U. Hassler (2002) Inference on the cointegration rank in fractionally integrated processes. *Journal of Econometrics* 110, 167–185.

Brillinger, D.R. (1981) *Time Series: Data Analysis and Theory*. McGraw-Hill.

Cavaliere, G. (2004) Unit root tests under time-varying variances. *Econometric Reviews* 23, 259–292.

Cavaliere, G. & A.M.R. Taylor (2007) Testing for unit roots in time series models with non-stationary volatility. *Journal of Econometrics* 140, 919–947.

Cavaliere, G. & A.M.R. Taylor (2008a) Bootstrap unit root tests for time series with nonstationary volatility. *Econometric Theory* 24, 43–71.

Cavaliere, G. & A.M.R. Taylor (2008b) Time-transformed unit root tests for models with non-stationary volatility. *Journal of Time Series Analysis* 29, 300–330.

Demetrescu, M., V. Kuzin, & U. Hassler (2008) Long memory testing in the time domain. *Econometric Theory* 24, 176–215.

Deo, R. (2000) Spectral tests of the martingale hypothesis under conditional heteroscedasticity. *Journal of Econometrics* 99, 291–315.

Dolado, J.J., J. Gonzalo, & L. Mayoral (2002) A fractional Dickey–Fuller test for unit roots. *Econometrica* 70, 1963–2006.

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

Hamori, S. & A. Tokihisa (1997) Testing for a unit root in the presence of a variance shift. *Economics Letters* 57, 245–253.

Hansen, B.E. (1995) Regression with nonstationary volatility. *Econometrica* 63, 1113–1132.

Kew, H. (2007) Estimation and testing of a fractionally integrated model under heteroskedasticity. Ph.D. Dissertation, University of Melbourne.

Kew, H. & D. Harris (2008) Heteroskedasticity Robust Testing for a Fractional Unit Root. Working paper, University of Melbourne. Available at SSRN: http://ssrn.com/abstract=1184333.

https://doi.org/10.1017/S0266466609990314 Published online by Cambridge University Press

Kim, T.H., S. Leybourne, & P. Newbold (2002) Unit root tests with a break in innovation variance. *Journal of Econometrics* 109, 365–387.

Lobato, I.N. & C. Velasco (2006) Optimal fractional Dickey–Fuller tests. *Econometrics Journal* 9, 492–510.

Lobato, I.N. & C. Velasco (2007) Efficient Wald tests for fractional unit roots. *Econometrica* 75, 575–589.

Marinucci, D. & P.M. Robinson (1999) Alternative forms of fractional Brownian motion. *Journal of Statistical Planning and Inference* 80, 111–122.

Mayoral, L. (2007) Minimum distance estimation of stationary and non-stationary ARFIMA processes. *Econometrics Journal* 10, 124–148.

McConnell, M.M. & G. Perez-Quiros (2000) Output fluctuations in the United States: What has changed since the early 1980s? *American Economic Review* 90, 1464–1476.

Nicholls, D.F. & A.R. Pagan (1983) Heteroscedasticity in models with lagged dependent variables. *Econometrica* 51, 1233–1242.

Phillips, P.C.B. & K.-L. Xu (2006) Inference in autoregression under heteroskedasticity. *Journal of Time Series Analysis* 27, 289–308.

Robinson, P.M. (1991) Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. *Journal of Econometrics* 47, 67–84.

Robinson, P.M. (1994) Efficient tests of nonstationary hypotheses. *Journal of the American Statistical Association* 89, 1420–1437.

Tanaka, K. (1999) The nonstationary fractional unit root. *Econometric Theory* 15, 549–582.

van Dijk, D., D.R. Osborn, & M. Sensier (2002) Changes in Variability of the Business Cycles in the G7 Countries. Econometric Institute Report EI 2002–28.

Watson, M.W. (1999) Explaining the increased variability in long-term interest rates. *Federal Reserve Bank of Richmond Economic Quarterly* 85, 71–96.

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

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

Xu, K.-L. & P.C.B. Phillips (2008) Adaptive estimation of autoregressive models with time-varying variances. *Journal of Econometrics* 142, 265–280.

https://doi.org/10.1017/S0266466609990314 Published online by Cambridge University Press