International Scholarly Research Network
ISRN Probability and Statistics
Volume 2012, Article ID 649134, 12 pages
doi:10.5402/2012/649134

Research Article

Simulation Analysis of Threshold Autoregressive Unit Root Tests

Steve Cook

Department of Economics, Swansea University, Singleton Park, Swansea SA2 8PP, UK

Correspondence should be addressed to Steve Cook, s.cook@swan.ac.uk

Received 9 March 2012; Accepted 24 April 2012

Academic Editors: D. Fiems, J. Perelló, and M. Scotto

Using numerical simulation, the finite-sample properties of threshold autoregressive (TAR) and momentum-threshold (MTAR) autoregressive-based unit root tests under both deterministic and consistent methods of threshold estimation are examined in the presence of generalised autoregressive conditional heteroskedasticity (GARCH). Previous research is extended by considering both the impact of alternative robust methods of covariance matrix estimation and the behaviour of the secondary tests of asymmetry associated with the TAR and MTAR models. The results obtained reveal many interesting features, in particular the distortionary effects of consistent-threshold estimation. In summary, the findings presented indicate that caution should be exercised when interpreting the results of these frequently employed threshold-based testing methods.

1. Introduction

Following the work of Engle [1], Bollerslev [2], and Taylor [3], the analysis of volatility has become a central feature of both empirical and theoretical finance. This in turn has led to the emergence of a large literature examining the properties and behaviour of unit root tests in the presence of volatility in its commonly considered form of generalised autoregressive conditional heteroskedasticity (GARCH). (The GARCH model has become a cornerstone of empirical research in economics and finance, receiving widespread application and proving to provide accurate forecasts (see [4]).) While some authors have considered joint estimation of unit root testing equations and GARCH (1,1) processes (see, *inter alia*, [5–7] and references contained therein), others have examined the size properties of unit root tests in the presence of neglected GARCH behaviour (see [8–10]). (Following the emergence of a number of studies considering the relevance and implications of heavy-tailed distributions in economic and financial time series data (see, *inter alia*, [11–14]), a further related literature has emerged examining the properties of unit root tests when applied to series with heavy-tailed disturbances (see [15, 16])). In this paper, this latter literature is extended by considering the properties of the threshold autoregressive (TAR) and momentum-threshold autoregressive (MTAR) unit root tests of Enders and Granger [17] when applied to time series with GARCH (1,1) disturbances. This issue clearly merits attention given the noted prevalence of GARCH behaviour in economic and financial data. While Cook [9] provides an initial examination of the properties of the TAR and MTAR unit root tests in the presence of GARCH, the current study develops this work in two crucial ways. First, the impact of alternative methods of variance-covariance estimation is considered. An analysis of this is of clear importance as while the covariance matrix estimators of White [18] and Newey and West [19] are routinely adopted by econometricians and time series analysts as a general correction for problems encountered with heteroskedasticity, their impact in the presently considered circumstances is unknown. Second, the subsequent tests of the hypothesis of symmetry associated with the TAR and MTAR tests are examined. Again, this is an important issue as the results of these subsequent tests of symmetry are often reported by investigators using Fisher's F-distribution to reinforce their empirical analyses, despite the individual properties of the tests in these circumstances, and the applicability of the F-distribution, being unknown. (The use of the F-distribution to provide critical values for the symmetry tests is apparent in the empirical analysis conducted in the seminal work of Enders and Granger [17]. More recent studies employing the F-distribution include, *inter alia*, Payne and Mohammadi [20] and Thompson [21]). Indeed, it will be seen that the use of the F-distribution can generate misleading results for the symmetry tests, irrespective of whether GARCH is present. The present work addresses these issues using Monte Carlo simulation, considering the TAR and MTAR tests under both deterministic and consistent methods of threshold estimation.

This paper proceeds as follows. In the following section the TAR and MTAR tests are outlined. Using Monte Carlo methods, Section 3 considers the finite-sample distributions of the TAR and MTAR unit root tests under alternative methods of covariance matrix estimation and the empirical sizes of the symmetry tests under use of the F-distribution. The simulation analysis is extended in Section 4 to consider the behaviour of the TAR and MTAR unit root tests and the subsequent tests of symmetry, in the presence of GARCH. Section 5 concludes.

2. TAR and MTAR Unit Root Tests

Consider a time series process $\{y_t\}_{t=0}^{T}$. Given this series, the familiar Dickey-Fuller [22] (DF) test in its simplest form can be expressed as the t-ratio of $\widehat{\phi}$ in the following regression:

$$\Delta y_t = \phi y_{t-1} + \varepsilon_t. \tag{2.1}$$

Comparison of the calculated test statistic with specifically derived nonstandard critical values allows examination of the unit root hypothesis $H_0 : \phi = 0$ against an alternative of asymptotic stationarity ($H_0 : \phi < 0$). However it is apparent that (2.1) is an implicitly symmetric specification. To allow for the possibility of stationary asymmetric adjustment about an underlying attractor, Enders and Granger [17] draw upon the threshold autoregressive methods of Tong [23, 24]. Adopting this approach, (2.1) is generalised via the introduction of the Heaviside indicator function to partition the lagged level term y_{t-1} with a resulting testing equation given as

$$\Delta y_t = I_t \rho_1 y_{t-1} + (1 - I_t)\rho_2 y_{t-1} + \xi_t, \tag{2.2}$$

where the zero-one Heaviside indicator function (I_t) can be specified as either

$$I_t = \begin{cases} 1 & \text{if } y_{t-1} \geq 0, \\ 0 & \text{if } y_{t-1} < 0 \end{cases} \tag{2.3}$$

or:

$$I_t = \begin{cases} 1 & \text{if } \Delta y_{t-1} \geq 0, \\ 0 & \text{if } \Delta y_{t-1} < 0. \end{cases} \tag{2.4}$$

The combination of (2.2) and (2.3) results in a TAR model, while combination of (2.2) and (2.4) leads to an MTAR model. The testing equation (2.2) therefore provides the basis of the asymmetric unit root tests of Enders and Granger [17], with the speed of adjustment about the stationary attractor given by ρ_1 when above equilibrium and ρ_2 when below. The unit root hypothesis is then tested via the null $\mathrm{H}_0 : \rho_1 = \rho_2 = 0$, with the resulting test statistic denoted as Φ when using (2.3) and Φ^* when using (2.4). Due to the nonstandard distributions of the Φ and Φ^* statistics, their use requires the derivation of appropriate critical values via Monte Carlo experimentation. To further examine potential asymmetric stationarity, investigators often conduct a further test of the symmetry hypothesis via the null $\rho_1 = \rho_2$ using a conventional F-statistic. The appropriateness of the F-distribution when conducting this subsequent test of symmetry will be questioned later in this paper.

The above analysis adopts an implicit assumption of zero attractor ($y_t = 0$). However, in practice it is more realistic to consider a nonzero attractor. In these circumstances, it is possible to derive the required threshold either deterministically or via consistent-threshold estimation. Under the first option, a deterministic threshold is imposed via regressing the original series $\{y_t\}$ upon a constant to derive a new series $\{\tilde{y}_t\}$. (In this analysis, the intercept-only model is considered for the TAR and MTAR tests as it is this form alone which is available for the tests under consistent-threshold estimation.) The derived series $\{\tilde{y}_t\}$ is then employed in a modified version of (2.2) as given below:

$$\Delta \tilde{y}_t = I_t \rho_1 \tilde{y}_{t-1} + (1 - I_t)\rho_2 \tilde{y}_{t-1} + \zeta_t. \tag{2.5}$$

In a similar manner, the specification of the Heaviside indicator function in (2.3) or (2.4) is modified by using the appropriate revised series $\{\tilde{y}_t\}$ rather than $\{y_t\}$. The resulting TAR and MTAR tests are denoted as Φ_μ and Φ_μ^*, respectively. Under consistent-threshold estimation, the above two-step procedure is replaced by an optimisation procedure to select the threshold. To illustrate this, consider the Φ_μ and Φ_μ^* tests above. To construct these tests statistics, prior to regression upon a constant term effectively, this demeans the series of interest, implying that the version of (2.2) actually estimated is given as

$$\Delta y_t = I_t \rho_1 \left(y_{t-1} - \overline{y}\right) + (1 - I_t)\rho_2 \left(y_{t-1} - \overline{y}\right) + \xi_t. \tag{2.6}$$

However, as noted by Enders [25], in the presence of asymmetric adjustment ρ_1 and ρ_2 differ, and consequently the mean is a biased estimator of the threshold. To overcome this and obtain a consistent value of the threshold, the approach of Chan [26] is followed with a grid search procedure employed to select the threshold. The previously defined Heaviside indicator functions of (2.3) and (2.4) are then revised as follows:

$$I_t = \begin{cases} 1 & \text{if } y_{t-1} \geq \tau, \\ 0 & \text{if } y_{t-1} < \tau, \end{cases} \tag{2.7}$$

$$I_t = \begin{cases} 1 & \text{if } \Delta y_{t-1} \geq \tau\Delta, \\ 0 & \text{if } \Delta y_{t-1} < \tau\Delta, \end{cases} \tag{2.8}$$

with the thresholds τ and τ_Δ selected by grid search over a range of values of $\{y_t\}$ and $\{\Delta y_t\}$ respectively. The appropriate asymmetric unit root testing equations are therefore given as

$$\Delta y_t = I_t \rho_1 (y_{t-1} - \psi) + (1 - I_t)\rho_2 (y_{t-1} - \psi) + \eta_t, \quad \psi = \tau, \tau_\Delta. \tag{2.9}$$

Application of the consistent-threshold versions of the TAR and MTAR tests therefore requires selection of the appropriate threshold value (ψ). Considering TAR adjustment, the series of interest $\{y_t\}$ is reordered as $y_1^o < y_2^o < \cdots < y_T^o$, with the central 70% of observations ($y_i^o, i = 0.15T, \ldots, 0.85T$) within this range of values considered as potential thresholds. Each of these values ($y_i^o = \tau$) is employed in turn in the indicator function of (2.7) with (2.9) then estimated. The value y_i^o delivering the minimum residual sum of squares ($\sum \eta_t^2$) for (2.9) is then deemed to be the consistent estimate of the threshold (τ) with the resulting TAR asymmetric unit root test labelled Φ_μ^c. Under MTAR adjustment, a similar approach is followed with the central 70% of observations from the reordered sequence $\{\Delta y_1^o, \Delta y_2^o, \ldots, \Delta y_T^o\}$ now considered as potential thresholds. Again the selected threshold τ_Δ is that value minimising the residual sum of squares of (2.9) with (2.8) now employed as the appropriate indicator function. The resultant MTAR test under consistent-threshold estimation is labelled as $\Phi_\mu^{*,c}$. In the subsequent research of Cook and Manning [27], the use of consistent-threshold estimation was shown to substantially increase the power of both TAR and MTAR asymmetric unit root tests in the presence of asymmetrically stationary processes. In the subsequent sections of this paper, the behaviour of the above tests will be examined in the presence of GARCH, with particular attention paid to the role of alternative variance-covariance matrix estimators and the secondary test of symmetry.

3. Monte Carlo Experimentation I

3.1. Derivation of Critical Values for the Asymmetric Unit Root Tests

Before considering the properties of the $\Phi_\mu, \Phi_\mu^*, \Phi_\mu^c$, and $\Phi_\mu^{*,c}$ tests and their associated subsequent tests of symmetry in the presence of GARCH, critical values for these tests are derived under the use of alternative variance-covariance estimators. In addition to considering the standard OLS covariance estimator, the covariance matrix estimators of White [18] and Newey-West [19] are also employed for the four tests. The use of White and Newey-West covariance matrix estimators is denoted by the addition of the subscripts w and

Table 1: Critical values for asymmetric unit root tests.

$\Phi_{\mu,w}^{*,c}$	4.87	5.98	8.61	4.60	5.60	7.78	4.49	5.38	7.48
$\Phi_{\mu,n}^{*,c}$	5.63	7.02	10.47	5.12	6.33	9.06	4.79	5.85	8.20

The reported results represent critical values for the TAR and MTAR tests under deterministic and consistent threshold estimation using alternative covariance matrix estimators. The results were derived over 50,000 simulations.

n, respectively, to the test statistic, while no additional subscript is employed when using the standard OLS covariance estimator. To obtain the required critical values, the following data generation process (DGP) is employed:

$$y_t = y_{t-1} + \varepsilon_t \quad t = 1, \ldots, T, \tag{3.1}$$

where the innovation series $\{\varepsilon_t\}$ is generated using pseudo *i.i.d.* $\mathsf{N}(0,1)$ random numbers from the NRND procedure in the EViews 6. The experiments are performed over 50,000 simulations for three sample sizes: $T = \{200, 400, 800\}$. Critical values are reported in Table 1 for three levels of significance ($\alpha = 0.10, 0.05, 0.01$). The consideration of four tests, three sample sizes, three covariance matrix estimators, and three levels of significance results in the calculation of ($4 \times 3 \times 3 \times 3 \ =$) 108 critical values. For all of the four tests examined, the results presented in Table 1 show that finite-sample critical values are clearly dependent upon the covariance matrix estimator employed. Although the impact of alternative covariance estimators is negligible asymptotically, its finite-sample impact is apparent and is noticeable for even the relatively large sample of 800 observations considered in the present experiments. In short, it can be seen that movement from the OLS covariance matrix estimator to the White estimator and then Newey-West estimator leads to inflation of the critical values of all tests.

3.2. *Empirical Sizes of the Symmetry Tests*

To consider the finite-sample sizes of the subsequent associated tests of symmetry ($\mathsf{H}_0 : \rho_1 = \rho_2$) associated with the $\Phi_\mu, \Phi_\mu^*, \Phi_\mu^c$ and $\Phi_\mu^{*,c}$ tests under alternative methods of covariance matrix estimation, the DGP of the previous subsection is employed but with (3.1) respecified as follows:

$$y_t = \rho y_{t-1} + \varepsilon_t \quad t = 1, \ldots, T. \tag{3.2}$$

This modification is introduced to allow consideration of stationary processes via $|\rho| < 1$, as the symmetry test is valid for stationary processes only. To examine the finite-sample empirical sizes of the symmetry tests associated with the $\Phi_\mu, \Phi_\mu^*, \Phi_\mu^c$ and $\Phi_\mu^{*,c}$ statistics, $\rho = 0.85$ is employed in the following Monte Carlo analysis. The empirical rejection frequencies for the symmetry tests are derived at the 5% level of significance ($\alpha = 0.05$) over 50,000 replications with the innovation series $\{\varepsilon_t\}$ again generated as a standard normal variable using the NRND procedure.

The empirical rejection frequencies for the symmetry tests are reported in Table 2. From inspection of this table, it is clear that there are a number of interesting findings to consider. First, under deterministic estimation of the threshold, the symmetry test for the TAR model is substantially undersized irrespective of the covariance matrix estimator employed.

Table 2: Empirical sizes of symmetry tests.

	$T = 200$	$T = 400$	$T = 800$
Φ_{μ}	0.08	0.09	0.12
$\Phi_{\mu,w}$	0.12	0.10	0.13
$\Phi_{\mu,n}$	0.22	0.19	0.16
Φ_{μ}^{c}	33.15	29.42	26.70
$\Phi_{\mu,w}^{c}$	34.78	30.07	26.99
$\Phi_{\mu,n}^{c}$	36.02	31.33	27.73
Φ_{μ}^{*}	4.35	4.45	4.39
$\Phi_{\mu,w}^{*}$	5.00	4.70	4.60
$\Phi_{\mu,n}^{*}$	6.16	5.46	5.07
$\Phi_{\mu}^{*,c}$	26.69	24.37	20.69
$\Phi_{\mu,w}^{*,c}$	29.94	25.68	21.21
$\Phi_{\mu,n}^{*,c}$	33.33	27.77	22.43

The reported results represent empirical rejection frequencies at the 5% level of significance of symmetry tests for the TAR and MTAR models under deterministic and consistent threshold estimation using alternative covariance matrix estimators. The results were derived over 50,000 simulations.

In contrast to this, the symmetry test for the MTAR model with a deterministic threshold has near nominal size. Second, under consistent-threshold estimation, the symmetry tests for both the TAR and MTAR models are dramatically oversized, with the former exhibiting greater size distortion than the latter. Third, movement from the standard OLS covariance matrix estimator to the White and the Newey-West estimator results in an increase in empirical size irrespective of the test and method of threshold estimation employed.

4. Monte Carlo Experimentation II

4.1. Empirical Sizes of Asymmetric Unit Root Tests in the Presence of GARCH

To examine the properties of the above-mentioned asymmetric unit root tests in the presence of GARCH (1,1) errors, the following DGP is employed:

$$\begin{gathered} y_t = y_{t-1} + w_t \quad t = 1, \dots, T, \\ h_t^2 = \phi_0 + \phi_1 w_{t-1}^2 + \phi_2 h_{t-1}^2, \\ w_t = h_t v_t, \\ v_t \sim N(0,1). \end{gathered} \tag{4.1}$$

This DGP closely follows that of Kim and Schmidt [8] and considers GARCH processes generated for a range of realistic values of $\{\phi_1, \phi_2\}$ corresponding to near integration, with $\phi_0 = 1 - \phi_1 - \phi_2$ in all cases. (In this paper the less empirically realistic, or relevant, cases of degenerate GARCH ($\phi_0 = 0$) and integrated GARCH ($\phi_1 + \phi_2 = 1$) are not considered.) The precise values of the volatility parameter (ϕ_1) and the degree of persistence ($\phi_1 + \phi_2$) employed are based upon estimated values observed in empirical research, with

Table 3: TAR unit root testing in the presence of GARCH.

		(ϕ_1, ϕ_2)				
		(0.01, 0.98)	(0.05, 0.94)	(0.10, 0.89)	(0.15, 0.84)	(0.25, 0.70)
	Φ_μ	5.24	6.27	7.79	8.94	8.51
	$\Phi_{\mu,w}$	5.07	5.22	5.23	5.15	4.95
$T = 200$	$\Phi_{\mu,n}$	5.26	5.23	5.18	4.91	4.38
	Φ_μ^c	5.32	6.47	8.76	10.83	12.25
	$\Phi_{\mu,w}^c$	4.97	5.16	5.20	5.23	5.39
	$\Phi_{\mu,n}^c$	5.02	5.10	5.24	5.24	5.11
	Φ_μ	5.28	6.31	8.00	9.25	8.29
	$\Phi_{\mu,w}$	5.15	5.18	5.22	5.00	4.84
$T = 400$	$\Phi_{\mu,n}$	5.24	5.25	5.31	5.04	4.38
	Φ_μ^c	5.42	7.09	10.38	13.15	13.49
	$\Phi_{\mu,w}^c$	5.28	5.27	4.95	4.65	4.60
	$\Phi_{\mu,n}^c$	5.37	5.30	5.12	4.64	4.44
	Φ_μ	5.02	5.99	7.42	8.94	7.62
	$\Phi_{\mu,w}$	4.85	4.80	4.69	4.67	4.90
$T = 800$	$\Phi_{\mu,n}$	4.92	4.86	4.91	4.77	4.63
	Φ_μ^c	5.16	7.13	10.53	14.11	12.89
	$\Phi_{\mu,w}^c$	4.93	4.86	4.34	4.32	4.41
	$\Phi_{\mu,n}^c$	4.94	4.82	4.59	4.28	3.75

The reported results represent empirical sizes of the test of the unit root hypothesis at the 5% level of significance for the TAR model under deterministic and consistent threshold estimation using alternative covariance matrix estimators. The results were derived over 20,000 simulations.

the distinction between typical values of $\{\phi_1, \phi_2\}$ observed at differing data frequencies noted (see [28, 29]). Throughout, the initial value of the conditional variance is set equal to one ($h_0 = 1$), with an additional, initial 400 observations of the generation of the GARCH process discarded prior to generation of the series of interest y_t to remove the impact of the initial condition. The initial value of y_t is set to zero ($y_0 = 0$) without loss of generality. The innovation series $\{v_t\}$ is generated using pseudo *i.i.d.* $\mathsf{N}(0,1)$ random numbers from the NRND procedure in EViews 6 with all experiments performed over 20,000 simulations for the previously considered sample sizes: $T = \{200, 400, 800\}$. The empirical sizes of the alternative tests are derived as the rejection frequencies for the alternative tests at the 5% level of significance under the use of the OLS, White and Newey-West covariance estimators using the critical values derived in the previous section.

Considering the results presented in Table 3 for the TAR unit root test for the standard OLS covariance estimator, it can be seen that the presence of GARCH does result in size inflation, with this becoming more noticeable for larger values of the volatility parameter. From inspection of the results for $\{\phi_1, \phi_2\} = \{0.15, 0.84\}$ and $\{\phi_1, \phi_2\} = \{0.25, 0.70\}$ it is also apparent that the degree of near integration as well as the size of the volatility parameter influences size distortion, as empirical size is lower for the second pairing of parameters with a lower sum for the GARCH parameters, but a higher degree of volatility. While these results hold under both deterministic and consistent threshold estimation, it can be seen that greater size distortion results from the use of consistent-threshold estimation. Conversely, it is apparent that the use of either the White or Newey-West covariance covariance matrix estimators dramatically reduces oversizing for the tests to a near negligible level.

Table 4: MTAR unit root testing in the presence of GARCH.

		(ϕ_1, ϕ_2)				
		(0.01, 0.98)	(0.05, 0.94)	(0.10, 0.89)	(0.15, 0.84)	(0.25, 0.70)
	Φ^*_μ	4.99	5.99	7.61	9.09	8.78
	$\Phi^*_{\mu,w}$	4.98	5.07	4.92	4.90	4.74
$T = 200$	$\Phi^*_{\mu,n}$	5.06	5.23	5.22	5.01	4.50
	$\Phi^{*,c}_\mu$	5.30	6.33	7.83	9.13	9.66
	$\Phi^{*,c}_{\mu,w}$	4.94	4.97	4.59	4.26	3.95
	$\Phi^{*,c}_{\mu,n}$	4.94	5.07	4.81	4.41	3.83
	Φ^*_μ	5.21	6.21	8.31	9.84	8.59
	$\Phi^*_{\mu,w}$	5.06	4.92	4.78	4.63	4.28
$T = 400$	$\Phi^*_{\mu,n}$	4.96	4.92	4.82	4.68	4.05
	$\Phi^{*,c}_\mu$	4.96	6.54	8.42	10.12	9.74
	$\Phi^{*,c}_{\mu,w}$	5.00	4.63	4.21	3.80	3.43
	$\Phi^{*,c}_{\mu,n}$	4.95	4.52	4.09	3.68	3.14
	Φ^*_μ	5.09	6.29	8.15	10.04	8.04
	$\Phi^*_{\mu,w}$	5.03	4.76	4.61	4.41	4.48
$T = 800$	$\Phi^*_{\mu,n}$	5.25	5.02	4.82	4.59	4.43
	$\Phi^{*,c}_\mu$	5.21	6.63	8.94	10.56	10.05
	$\Phi^{*,c}_{\mu,w}$	5.21	4.88	4.27	3.60	3.26
	$\Phi^{*,c}_{\mu,n}$	5.04	4.72	4.09	3.59	3.24

The reported results represent empirical sizes of the test of the unit root hypothesis at the 5% level of significance for the MTAR model under deterministic and consistent threshold estimation using alternative covariance matrix estimators. The results were derived over 20,000 simulations.

From inspection of the results for the MTAR unit root test in Table 4, it can be seen that the behaviour of the MTAR test closely mimics that of the TAR test. However, it is also apparent that under consistent-threshold estimation and use of the standard OLS covariance matrix estimator, the oversizing of the MTAR test is less than that of TAR test for the identical values of the GARCH parameters. However, despite this, empirical sizes more than twice the nominal size are obtained.

4.2. Empirical Sizes of the Symmetry Tests in the Presence of GARCH

In Tables 5 and 6, the empirical sizes of the symmetry tests associated with the TAR and MTAR unit root tests are reported. These results are obtained using the following DGP:

$$\begin{gathered} y_t = \rho y_{t-1} + w_t \quad t = 1, \dots, T, \\ h_t^2 = \phi_0 + \phi_1 w_{t-1}^2 + \phi_2 h_{t-1}^2, \\ w_t = h_t v_t, \\ v_t \sim N(0,1), \end{gathered} \tag{4.2}$$

where $\rho = 0.85$ in (4.2). Aside from the consideration of stationary time series via the use of the AR(1) specification $y_t = 0.85 y_{t-1} + w_t$, the Monte Carlo analysis conducted in this

Table 5: TAR symmetry testing in the presence of GARCH.

		(ϕ_1, ϕ_2)				
		(0.01, 0.98)	(0.05, 0.94)	(0.10, 0.89)	(0.15, 0.84)	(0.25, 0.70)
$T = 200$	Φ_μ	0.09	0.17	0.36	0.63	0.99
	$\Phi_{\mu,w}$	0.13	0.16	0.18	0.20	0.20
	$\Phi_{\mu,n}$	0.28	0.26	0.37	0.44	0.42
	Φ_μ^c	34.01	35.87	39.03	41.37	44.73
	$\Phi_{\mu,w}^c$	35.52	35.39	35.05	34.37	34.90
	$\Phi_{\mu,n}^c$	36.59	37.04	37.84	38.27	39.25
$T = 400$	Φ_μ	0.14	0.25	0.90	1.86	2.61
	$\Phi_{\mu,w}$	0.15	0.15	0.21	0.24	0.34
	$\Phi_{\mu,n}$	0.22	0.26	0.34	0.45	0.57
	Φ_μ^c	29.37	31.84	35.54	38.72	41.60
	$\Phi_{\mu,w}^c$	29.82	29.41	28.71	28.05	28.82
	$\Phi_{\mu,n}^c$	30.85	31.14	31.46	31.90	33.18
$T = 800$	Φ_μ	0.11	0.39	2.09	4.76	5.38
	$\Phi_{\mu,w}$	0.09	0.13	0.22	0.39	0.44
	$\Phi_{\mu,n}$	0.11	0.15	0.37	0.63	0.67
	Φ_μ^c	26.83	29.26	33.95	38.42	41.23
	$\Phi_{\mu,w}^c$	26.52	25.20	23.73	23.01	24.20
	$\Phi_{\mu,n}^c$	27.51	26.93	26.56	27.21	28.85

The reported results represent empirical sizes of the test of the symmetry hypothesis at the 5% level of significance for the TAR model under deterministic and consistent threshold estimation using alternative covariance matrix estimators. The results were derived over 20,000 simulations.

subsection follows that of the previous subsection in terms of generation and replication. While a range of results is reported in Tables 5 and 6, the findings can be summarised straightforwardly.

Considering the TAR model with a deterministically imposed threshold value, substantial undersizing is apparent. Conversely, when consistent-threshold estimation is employed, the resulting symmetry test is massively oversized, with the use of OLS covariance matrix inducing greater size inflation than the corrected covariance matrix estimators of White and Newey-West. Despite undersizing (oversizing) of the test with a deterministic (consistent) threshold being noted previously in the absence of GARCH, it is apparent that neglected GARCH behaviour does increase size distortion with size inflation being positively related to the degree of volatility as given by ϕ_1. Turning to the results for the MTAR test, under the use of a deterministic threshold and corrected covariance matrix estimator, the test has relatively good size, particularly when the White covariance matrix is used. However, when the standard OLS covariance is employed, substantial size distortion is exhibited for volatile GARCH processes. Considering the symmetry test for the MTAR model under consistent threshold estimation, size distortion is apparent irrespective of the covariance matrix estimator employed, although the corrected covariance estimators, and the White covariance matrix estimator in particular, do reduce this. The results obtained therefore indicate that use of the preferred method of consistent-threshold estimation results in substantial oversizing of the symmetry test for both the TAR and MTAR models with all covariance matrix estimators, leading to a spurious detection of asymmetric behaviour.

Table 6: MTAR symmetry testing in the presence of GARCH.

		(ϕ_1, ϕ_2)				
		(0.01, 0.98)	(0.05, 0.94)	(0.10, 0.89)	(0.15, 0.84)	(0.25, 0.70)
$T = 200$	Φ_μ^*	4.35	5.41	7.43	9.80	11.62
	$\Phi_{\mu,w}^*$	4.98	4.94	5.29	5.42	5.46
	$\Phi_{\mu,n}^*$	6.11	6.22	6.71	7.23	7.43
	$\Phi_\mu^{*,c}$	26.71	30.73	35.95	40.28	43.13
	$\Phi_{\mu,w}^{*,c}$	29.73	30.09	29.37	28.93	27.49
	$\Phi_{\mu,n}^{*,c}$	33.15	33.93	33.98	34.06	33.42
$T = 400$	Φ_μ^*	4.50	6.12	9.35	12.64	14.41
	$\Phi_{\mu,w}^*$	4.68	4.82	4.90	5.09	5.10
	$\Phi_{\mu,n}^*$	5.68	5.87	6.18	6.55	6.56
	$\Phi_\mu^{*,c}$	25.11	30.24	37.19	43.18	44.91
	$\Phi_{\mu,w}^{*,c}$	25.72	26.05	25.66	24.81	23.72
	$\Phi_{\mu,n}^{*,c}$	28.06	28.62	29.15	29.14	28.95
$T = 800$	Φ_μ^*	4.73	7.14	12.12	17.05	18.63
	$\Phi_{\mu,w}^*$	4.68	4.94	5.04	5.16	4.94
	$\Phi_{\mu,n}^*$	5.20	5.40	5.84	6.30	6.29
	$\Phi_\mu^{*,c}$	21.68	27.70	37.28	44.45	44.76
	$\Phi_{\mu,w}^{*,c}$	21.49	21.27	21.02	20.70	20.04
	$\Phi_{\mu,n}^{*,c}$	22.60	23.19	24.03	24.32	24.09

The reported results represent empirical sizes of the test of the symmetry hypothesis at the 5% level of significance for the MTAR model under deterministic and consistent threshold estimation using alternative covariance matrix estimators. The results were derived over 20,000 simulations.

5. Concluding Remarks

The popularity of threshold-based unit root tests as a means of detecting asymmetric stationarity in time series processes has increased in recent years, with their analysis and application now widespread. This paper has examined the practically important issue of the finite-sample size properties of these TAR and MTAR tests in the presence of GARCH behaviour, with particular attention paid to the use of alternative covariance matrix estimators and the subsequent test of the symmetry hypothesis. The results of the Monte Carlo analysis undertaken show the TAR and MTAR tests to be oversized when examining the unit root hypothesis in the presence of GARCH, unless a corrected covariance matrix estimator (White or Newey-West) is employed. This result is particularly apparent when considering more highly volatile GARCH processes. Interestingly, it was seen also that the use of consistent-threshold estimation increases the degree of size distortion exhibited. However, it was the examination of the symmetry hypothesis which provided the more dramatic and important results. The most crucial result concerned the substantial oversizing of the symmetry test for the TAR and MTAR models irrespective of both the covariance matrix estimator employed and the presence of GARCH. This finding indicated that a spurious inference of asymmetric adjustment could be noted with a high degree of probability when applying these tests and strongly rejected the use of the F-distribution for this test as routinely employed by practitioners (see [17, 20, 21]). Under the use of a deterministic threshold it was found that the symmetry test under the TAR model was undersized in the absence of GARCH, while the test under the MTAR model had approximately correct size.

In the presence of GARCH behaviour, the most noticeable finding concerned the subsequent oversizing of the symmetry test for the MTAR model, although this was noted for the model using the standard OLS, rather than corrected, covariance matrix estimator.

In summary, the results of the present paper indicate that the routine application of TAR and MTAR unit root tests to economic and financial time series can result in incorrect or spurious inferences. While the probability of a spurious inference concerning the unit root hypothesis is most likely to occur when a corrected covariance matrix estimator is employed in the examination of time series processes exhibiting GARCH behaviour, the symmetry test has been shown to exhibit spurious rejection under consistent-threshold estimation even in the absence of GARCH. The results obtained indicate that practitioners should exercise care when applying TAR and MTAR unit root tests to economic and financial time series, particularly when considering the secondary test of symmetry where use of the F-distribution is clearly inappropriate despite its routine use in empirical analyses. As result of the above findings, an obvious future avenue of research concerns the development of asymmetric unit root tests which capture and draw upon any GARCH behaviour exhibited by data. This work, which can be viewed as a development of the studies of Seo [5] and Cook [6], is the subject of ongoing research.

References

[1] R. F. Engle, "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation," *Econometrica*, vol. 50, no. 4, pp. 987–1007, 1982.

[2] T. Bollerslev, "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, vol. 31, no. 3, pp. 307–327, 1986.

[3] S. Taylor, *Modelling Financial Time Series*, John Wiley & Sons, New York, NY, USA, 1986.

[4] T. G. Andersen and T. Bollerslev, "Answering the skeptics: yes, standard volatility models do provide accurate forecasts," *International Economic Review*, vol. 39, no. 4, pp. 885–905, 1998.

[5] B. Seo, "Distribution theory for unit root tests with conditional heteroskedasticity," *Journal of Econometrics*, vol. 91, no. 1, pp. 113–144, 1999.

[6] S. Cook, "Joint maximum likelihood estimation of unit root testing equations and GARCH processes: some finite-sample issues," *Mathematics and Computers in Simulation*, vol. 77, no. 1, pp. 109–116, 2008.

[7] S. Cook, "A re-examination of the stationarity of inflation," *Journal of Applied Econometrics*, vol. 24, no. 6, pp. 1047–1053, 2009.

[8] K. Kim and P. Schmidt, "Unit root tests with conditional heteroskedasticity," *Journal of Econometrics*, vol. 59, no. 3, pp. 287–300, 1993.

[9] S. Cook, "The impact of GARCH on asymmetric unit root tests," *Physica A*, vol. 369, no. 2, pp. 745–752, 2006.

[10] S. Cook, "The robustness of modified unit root tests in the presence of GARCH," *Quantitative Finance*, vol. 6, no. 4, pp. 359–363, 2006.

[11] C. Granger and D. Orr, "Infinite variance and research strategy in time series analysis?" *Journal of the American Statistical Association*, vol. 67, pp. 275–285, 1972.

[12] M. Loretan and P. C. B. Phillips, "Testing the covariance stationarity of heavy-tailed time series: an overview of the theory with applications to several financial datasets," *Journal of Empirical Finance*, vol. 1, no. 2, pp. 211–248, 1994.

[13] R. S. Deo, "On estimation and testing goodness of fit for m-dependent stable sequences," *Journal of Econometrics*, vol. 99, no. 2, pp. 349–372, 2000.

[14] S. I. Resnick, *Heavy-Tail Phenomena*, Springer Series in Operations Research and Financial Engineering, Springer, New York, NY, USA, 2007.

[15] S. T. Rachev, S. Mittnik, and J.-R. Kim, "Time series with unit roots and infinite-variance disturbances," *Applied Mathematics Letters*, vol. 11, no. 5, pp. 69–74, 1998.

[16] K. D. Patterson and S. M. Heravi, "The impact of fat-tailed distributions on some leading unit roots tests," *Journal of Applied Statistics*, vol. 30, no. 6, pp. 635–667, 2003.

[17] W. Enders and C. W. J. Granger, "Unit-root tests and asymmetric adjustment with an example using the term structure of interest rates," *Journal of Business and Economic Statistics*, vol. 16, no. 3, pp. 304–311, 1998.

[18] H. White, "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity," *Econometrica*, vol. 48, no. 4, pp. 817–838, 1980.

[19] W. K. Newey and K. D. West, "Automatic lag selection in covariance matrix estimation," *Review of Economic Studies*, vol. 61, no. 4, pp. 631–653, 1994.

[20] J. E. Payne and H. Mohammadi, "Are adjustments in the U.S. budget deficit asymmetric? Another look at sustainability," *Atlantic Economic Journal*, vol. 34, no. 1, pp. 15–22, 2006.

[21] M. A. Thompson, "Asymmetric adjustment in the prime lending-deposit rate spread," *Review of Financial Economics*, vol. 15, no. 4, pp. 323–329, 2006.

[22] D. A. Dickey and W. A. Fuller, "Distribution of the estimators for autoregressive time series with a unit root," *Journal of the American Statistical Association*, vol. 74, no. 366, part 1, pp. 427–431, 1979.

[23] H. Tong, *Threshold Models in Nonlinear Time Series Analysis*, vol. 21 of *Lecture Notes in Statistics*, Springer, New York, NY, USA, 1983.

[24] H. Tong, *Non-Linear Time-Series: A Dynamical Approach*, Oxford University Press, Oxford, UK, 1990.

[25] W. Enders, "Improved critical values for the Enders-Granger unit-root test," *Applied Economics Letters*, vol. 8, no. 4, pp. 257–261, 2001.

[26] K. S. Chan, "Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model," *The Annals of Statistics*, vol. 21, no. 1, pp. 520–533, 1993.

[27] S. Cook and N. Manning, "The power of asymmetric unit root tests under threshold and consistent-threshold estimation," *Applied Economics*, vol. 35, no. 14, pp. 1543–1550, 2003.

[28] F. C. Drost and T. E. Nijman, "Temporal aggregation of GARCH processes," *Econometrica*, vol. 61, no. 4, pp. 909–927, 1993.

[29] R. Engle and A. Patton, "What makes a good volatility model?" *Quantitative Finance*, vol. 1, pp. 237–245, 2001.