A Monte Carlo Analysis of Unit Root Testing with GARCH(1,1) Errors

Shiqing Ling[1], W.K. Li[2], and Michael McAleer[1]
[1] Department of Economics, University of Western Australia
[2] Department of Statistics and Actuarial Science, University of Hong Kong

Abstract: Least squares (LS) estimation and maximum likelihood (ML) estimation are considered for unit root processes with GARCH(1,1) errors. The asymptotic distributions of LS and ML estimators are given under the second moment condition. The former has the usual unit root distribution and the latter is a functional of a bivariate Brownian motion, as in Ling and Li [1998]. Several unit root tests based on the LS estimator and on mixing LS and ML estimators are constructed. Simulation results show that tests based on mixing LS and ML estimators perform better than Dickey-Fuller tests based on LS estimators.

1. INTRODUCTION

Consider two unit root processes,

$$y_t = \phi y_{t-1} + \varepsilon_t \qquad (1.1)$$

$$y_t = \mu + \phi y_{t-1} + \varepsilon_t \qquad (1.2)$$

where $\phi = 1$, $\mu = 0$, ε_t follows the first-order generalized autoregressive conditional heteroscedasticity GARCH(1,1) model given by

$$\varepsilon_t = \eta_t \sqrt{h_t}, \qquad h_t = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1} \qquad (1.3)$$

where $\omega > 0$, $\alpha \geq 0$, $\beta \geq 0$, and the η_t are a sequence of independently and identically distributed (i.i.d.) random variables with zero mean and unit variance. For model (1.3), we make the following assumptions:

Assumption 1. $\alpha + \beta = 1$.
Assumption 2. The parameter vector $\delta = (\omega, \alpha, \beta)' \in \Theta$, where $\Theta = \{(\omega, \alpha, \beta) | \omega > 0, \underline{\alpha} \leq \alpha \leq 1 - \underline{\alpha}, \ \underline{\beta} \leq \beta \leq 1 - \underline{\beta}, \text{ for some } \underline{\alpha}, \underline{\beta} > 0\}$.
Assumption 3. η_t has a symmetric distribution and $E\eta_t^4 < \infty$.

The GARCH model was proposed by Bollerslev [1986] and has had many important applications in financial and econometric time series. Some recent reviews can be found in Bollerslev et al. [1992], Bollerslev et al. [1994] and Li et al. [1999]. When $\alpha = \beta = 0$, the ε_t defined by model (1.3) reduce to i.i.d. white noise and, for this case, the unit root process has been investigated extensively. Motivated by practical applications, in recent decades many statisticians and econometricians have considered various unit root processes with non-i.i.d. errors. Some related results on estimating and testing unit roots can be found in Phillips and Durlauf [1986], Phillips [1987], Chan and Wei [1988], Lucas [1995], and Herce [1996], and the references cited therein. When the error term follows a GARCH process, estimation and testing for a unit root involves intrinsic problems, an issue that was first raised by Pantula [1989]. He derived the asymptotic distribution of least squares (LS) estimators for a unit root process with a first-order ARCH error (i.e. model (1.3) with $\beta = 0$), and showed that Dickey-Fuller tests could still be employed in this case. Pantula [1986, p. 73] also stated without proof that Dickey-Fuller tests could be used for unit root processes with GARCH errors.

Peters and Veloce [1988] and Kim and Schmidt [1993] provided simulation results to show that Dickey-Fuller tests based on LS estimators are often sensitive and, when $\alpha + \beta$ is close to 1, the problem can be very serious. It seems that this phenomenon can be explained partly by the loss of efficiency of the LS estimator. Ling and Li [1998] derived the limiting distribution of the maximum likelihood (ML) estimator for a general nonstationary autoregressive moving average time series process with general-order GARCH errors, and demonstrated that it is more efficient than the LS estimator. As for stationary time series with GARCH errors [see Weiss, 1986, and Ling and Li, 1997a], Ling and Li's [1998] results are obtained under the assumption that the fourth moment is finite. However, for the GARCH(1,1) process, the condition for strict stationarity is $E(\ln(\alpha\eta_t^2 + \beta)) < 0$ [see Nelson, 1990], the condition for a finite variance is $\alpha+\beta<1$, and the condition for a finite fourth moment is $3\alpha^2+2\alpha\beta+\beta^2<1$. The fourth moment condition is clearly the most stringent.

For the pure GARCH(1,1) model, Lee and Hansen [1994] and Lumsdaine [1996] proved that ML estimators are consistent and asymptotically normal under the condition that $E\left(\ln(\alpha\eta_t^2+\beta)\right)<0$. A challenging problem is whether the limiting distribution of ML estimators can be derived under weaker conditions for the unit root process with GARCH errors.

Ling and Li [1997b] obtained the asymptotic distribution of the ML estimators for models (1.1) and (1.2) under Assumptions 1-3. The limiting distribution of the estimated unit root is a functional of a bivariate Brownian motion and is the same as that obtained in Ling and Li [1998]. Based on these asymptotic results, we can construct several new unit root tests. Simulation results reported in the paper show that tests based on mixing LS and ML estimators perform better than those based on LS estimators alone.

This paper proceeds as follows. Section 2 presents LS estimation and its asymptotic properties. Section 3 considers ML estimation and its asymptotic properties. Section 4 reports some unit root tests. Section 5 presents some simulation results. The proofs of all theorems can be found in Ling et al. [1999].

Throughout the paper, we use the following notation. U' denotes the transpose of the vector U; $o(1)$ ($o_p(1)$) denotes a series of numbers (random numbers) converging to zero (in probability); $O(1)$ ($O_p(1)$) denotes a series of numbers (random numbers) that are bounded (in probability); $\xrightarrow{p}$ and $\xrightarrow{L}$ denote convergence in probability and in distribution, respectively; $D=D[0,1]$ denotes the space of functions $f(s)$ on $[0,1]$, which is defined and equipped with the Skorokhod topology [Billingsley, 1968]; $\|\cdot\|$ denotes the Euclidean norm.

2. PRELIMINARY ESTIMATION

Consider the observations $y_t, \ldots, y_n$ with initial value $y_0=0$, generated by model (1.1) or (1.2). Denote $\hat{\phi}_{LS}$ as the LS estimator of ϕ in model (1.1) and $(\tilde{\mu}_{LS}, \tilde{\phi}_{LS})$ as the LS estimator of (μ,ϕ) in model (1.2). Then

$$\hat{\phi}_{LS}=\left(\sum y_{t-1}^2\right)^{-1}\left(\sum y_t y_{t-1}\right) \tag{2.1}$$

$$\begin{pmatrix}\tilde{\mu}_{LS}\\ \tilde{\phi}_{LS}\end{pmatrix}=\begin{pmatrix} n & \sum y_{t-1}\\ \sum y_{t-1} & \sum y_{t-1}^2\end{pmatrix}^{-1}\begin{pmatrix}\sum y_t\\ \sum y_t y_{t-1}\end{pmatrix} \tag{2.2}$$

where $\sum$ denotes $\sum_{t=1}^n$.

Theorem 2.1. Suppose that Assumptions 1-3 hold. Then

(a) $n\left(\hat{\phi}_{LS}-1\right)\xrightarrow{L}\dfrac{\int B_1(t)dB_1(t)}{\int B_1^2(t)dt}$;

(b) $N_n\begin{pmatrix}\tilde{\mu}_{LS}\\ \tilde{\phi}_{LS}-1\end{pmatrix}\xrightarrow{L}$

$$\begin{pmatrix}1 & \int B_1(t)dt\\ \int B_1(t)dt & \int B_1^2(t)dt\end{pmatrix}^{-1}\begin{pmatrix}B_1(1)\\ \int B_1(t)dB_1(t)\end{pmatrix}$$

where $N_n=diag\left(\sqrt{n},n\right)$, $B_1(t)$ is a standard Brownian motion, and $\int$ denotes $\int_0^1$.

Let $\hat{\varepsilon}_t=y_t-\hat{\phi}_{LS}y_{t-1}$ or $y_t-\tilde{\mu}_{LS}-\tilde{\phi}_{LS}y_{t-1}$. Then $\{\hat{\varepsilon}_1,\hat{\varepsilon}_2,\ldots,\hat{\varepsilon}_n\}$ is a sequence of artificial observations from model (1.3). Denote

$$l_t\left(\varepsilon_t|\delta\right)=-\frac{1}{2}\ln h_t-\frac{\varepsilon_t^2}{2h_t} \tag{2.3}$$

$$l_t\left(\hat{\varepsilon}_t|\delta\right)=-\frac{1}{2}\ln \hat{h}_t-\frac{\hat{\varepsilon}_t^2}{2\hat{h}_t} \tag{2.4}$$

where $\hat{h}_t=\omega+\alpha\hat{\varepsilon}_t^2+\beta\hat{h}_{t-1}$, with $\hat{h}_0=1$ and $\delta=(\omega,\alpha,\beta)'$.

The first-order conditions seek the supremum of $\delta\in\Theta$. Suppose that δ_0 is the true value of $\delta\in\Theta$ and let

$$\hat{\delta}_n=\arg\max_{\delta\in\Theta}\left[\frac{1}{n}\sum l_t\left(\hat{\varepsilon}_t|\delta\right)\right].$$

The following shows the asymptotic properties of $\hat{\delta}_n$.

Theorem 2.2. Under Assumptions 1-3:

(a) $\hat{\delta}_n-\delta_0\xrightarrow{p}0$

(b) $\sqrt{n}(\hat{\delta}_n-\delta_0)\xrightarrow{L}N(0,V_0)$

where V_0 is given in Ling et al. [1999].

Remark. From Theorems 2.1 and 2.2, ϕ or (μ,ϕ) and δ can be estimated separately, with $\hat{\phi}_{LS} - \phi = O(n^{-1})$, $(\tilde{\mu}_{LS}, \tilde{\phi}_{LS}) - (\mu, \phi) = O(N_n^{-1})$, and $\hat{\delta} - \delta_0 = O(n^{-1/2})$. However, under GARCH errors, the LS estimator of ϕ or (μ,ϕ) loses some efficiency (see the next section) and hence, for finite samples, this will result in a loss of efficiency for the ML estimator of δ. A more efficient estimation procedure will be shown in the next section, so that all the estimators in this section can be viewed as preliminary estimators.

3. ML ESTIMATION

To simplify notation, in this section the true parameter δ_0 is denoted as δ. Let $\lambda = (\phi,\delta')'$, $\lambda_\mu = (\phi_\mu',\delta')'$ and $\phi_\mu = (\mu,\phi)'$. The ML estimators of λ and λ_μ are denoted by $\hat{\lambda} = (\hat{\phi}_{ML}, \hat{\delta}'_{ML})'$ and $\tilde{\lambda}_\mu = (\tilde{\phi}'_{\mu.ML}, \tilde{\delta}'_{\mu.ML})$, respectively, with $\tilde{\phi}_{\mu,ML} = (\tilde{\mu}_{ML}, \tilde{\phi}_{ML})'$, that maximise the log-likelihood

$$L = \frac{1}{n}\sum l_t \qquad (3.1)$$

where l_t is defined as in (2.3). The following theorem gives the asymptotic properties of the information matrix.

Theorem 3.1. Under Assumptions 1-3:

(a) $\frac{1}{n^2}\sum \frac{\partial^2 l_t}{\partial \phi^2} \xrightarrow{L} F\int w_1^2(t)dt$;

(b) $N_n^{-1}\sum \frac{\partial^2 l_t}{\partial \phi_\mu \partial \phi_\mu'} N_n^{-1} \xrightarrow{L}$
$F\begin{pmatrix} 1 & \int w_1(t)dt \\ \int w_1(t)dt & \int w_1^2(t)dt \end{pmatrix}$;

(c) $\frac{1}{n^{3/2}}\sum \frac{\partial^2 l_t}{\partial \phi \partial \delta} \xrightarrow{p} 0$;

(d) $(n^{1/2}N_n)^{-1}\sum \frac{\partial^2 l_t}{\partial \phi_\mu \partial \delta'} \xrightarrow{p} 0$;

(e) $\frac{1}{n}\sum \frac{\partial^2 l_t}{\partial \delta \partial \delta'} \xrightarrow{p} -E\left(\frac{\partial^2 l_t}{\partial \delta \partial \delta'}\right)$,

where $w_1(t)$ is a Brownian motion with covariance $t\sigma^2$, N_n is defined in Theorem 2.1, and $F = E(1/h_t) + 2\alpha^2 \sum_{k=1}^{\infty} \beta^{2(k-1)} E(\varepsilon_{t-k}^2 / h_t^2)$.

Since the likelihood equations $\delta l_t/\delta\lambda = 0$ and $\delta l_t/\delta\lambda_\mu = 0$ are nonlinear in λ and λ_μ, respectively, an iterative numerical procedure is required to obtain the solutions to these equations. By Theorem 3.1, ϕ (or ϕ_μ) and δ can be estimated separately without loss of efficiency. Thus, we can define an algorithm by the iterative approximate Newton-Raphson relation for $\hat{\phi}_{ML}$ and $\hat{\delta}_{ML}$.

Similarly, define a scheme for $\tilde{\phi}_{\mu.ML}$ and $\tilde{\delta}_{\mu.ML}$. As in the argument of Yap and Reinsel [1995], the estimators of λ and λ_μ obtained by the Newton-Raphson relation will be consistent if the initial estimators are consistent. By Theorem 2.1 and 2.2 in Section 2, we can obtain the initial estimator of λ (or λ_μ), such that $\hat{\phi} - \phi = O_p(n^{-1})$ (or $\hat{\phi}_\mu - \phi_\mu = O_p(N_n^{-1})$) and $\tilde{\delta} - \delta = O_p(n^{-1/2})$. With these consistent initial estimators, we can obtain the asymptotic representations of $\hat{\phi}_{ML}$ and $\hat{\delta}_{ML}$ by standard arguments. The asymptotic representations of $n(\tilde{\phi}_{\mu.ML} - \phi)$ and $\sqrt{n}(\tilde{\delta}_{\mu.ML} - \delta)$ can be obtained analogously. The following theorem gives the asymptotic distributions of $(\tilde{\phi}_{ML}, \hat{\delta}')'$ and $(\tilde{\phi}'_{\mu.ML}, \tilde{\delta}'_{ML})'$.

Theorem 3.2. Let $(\hat{\phi}_{ML}, \hat{\delta}')'$ and $(\tilde{\phi}'_{\mu,ML}, \tilde{\delta}'_{\mu.ML})'$ be the estimators of $(\phi,\delta')'$ and $(\phi_\mu',\delta')'$ obtained from the Newton-Raphson relation by using initial estimators $\hat{\phi}$ or $\hat{\phi}_\mu$ and $\hat{\delta}$, with $\hat{\phi} - \phi = O_p(n^{-1})$ or $\hat{\phi}_\mu - \phi_\mu = O_p(N_n^{-1})$ and $\tilde{\delta} - \delta = O_p(n^{-1/2})$, respectively. Then, under Assumptions 1-3:

(a) $n(\hat{\phi}_{ML} - 1) \xrightarrow{L} \frac{\int w_1(t)dw_2(t)}{F\int w_1^2(t)dt}$;

(b) $N_n\begin{pmatrix} \tilde{\mu}_{ML} \\ \tilde{\phi}_{\mu,ML} - 1 \end{pmatrix} \xrightarrow{L}$
$F^{-1}\begin{pmatrix} 1 & \int w_1(t)dt \\ \int w_1(t)dt & \int w_1^2(t)dt \end{pmatrix}^{-1}\begin{pmatrix} w_2(1) \\ \int w_1(t)dw_2(t) \end{pmatrix}$

where N_n is defined in Theorem 2.1, and $(w_1(t),w_2(t))$ is a bivariate Brownian motion with covariance $t\Omega$, $\kappa = E\eta_t^4 - 1$ and, when η_t is normal, $\kappa = 2$. For $\sqrt{n}(\hat{\delta}_{ML} - \delta)$ and $\sqrt{n}(\tilde{\delta}_{\mu,ML} - \delta)$, their asymptotic distributions are the same as those given in Theorem 2.2(b).

Remark. The asymptotic distributions of $\hat{\phi}_{ML}$ and $\tilde{\phi}_{\mu,ML}$ can be represented, respectively, as combinations of those of $\hat{\phi}_{LS}$ and $\tilde{\phi}_{\mu,LS}$ and a scale mixture of normals. These properties are similar to hose of the least absolute deviation estimators of unit roots given in Herce [1996]. The ML estimator of ϕ or ϕ_μ is more efficient than the LS estimator given in the previous section [see Ling and Li, 1998].

4. UNIT ROOT TESTS

Based on the asymptotic results in Sections 2 and 3, we can report some new unit root tests for the nonstationary models (1.1) and (1.2) with GARCH error (1.3). First, we define the test statistics, L_ϕ, L_t, $L_{\mu,\phi}$ and $L_{\mu,t}$, based on LS estimators, as given in Ling et al. [1999]. The limiting distributions of the LS-based tests are the same as those given in Dickey and Fuller [1979]. The critical values of these distributions can be found in Tables 8.5.2 and 8.5.3 of Fuller [1976].

In order to apply ML estimators for unit root tests, we need to modify $n(\hat{\phi}_{ML}-1)$ and $n(\tilde{\phi}_{\mu,ML}-1)$ in Theorem 3.2 because these limiting distributions depend on nuisance parameters. These nuisance parameters should be replaced by consistent estimators, an approach that was first used by Phillips [1987]. Recently, a similar approach was employed by Lucas [1995] and Herce [1996]. We can define new test statistics by mixing LS and ML estimators to yield M_ϕ, M_t, $M_{\mu,\phi}$ and $M_{\mu,t}$, as given in Ling et al. [1999].

We call these new test statistics ML-based tests. The limiting distributions of the ML-based tests are the same as those based on the least absolute deviations estimators of Herce [1996]. However, the test statistics themselves are quite different. The empirical critical values of these distributions will be given in the next section. When $\alpha = 0$, h_t is a constant, and hence $\kappa\sigma^2 = 1$. In this case, Assumption 2 is violated and the above tests cannot be used. Therefore, it is necessary to check whether or not the coefficient α is equal to zero before deciding to employ the ML-based test statistics. This can be done easily by applying the diagnostic checking method in Li and Mark [1994] for the pure GARCH model (1.3), with the artificial observations, $\hat{\varepsilon}_t$, in Section 2.

5. SIMULATION STUDY

In this section, we obtain some critical values of the ML-based tests, and report the empirical sizes and powers of the new test statistics in Section 4.

The finite sample distributions of the ML-based tests are obtained by simulation. First, we generate a series of GARCH(1,1) processes $\{\varepsilon_t\}$, where $\varepsilon_t = \eta_t\sqrt{h_t}$, $h_t = 0.1 + 0.3\varepsilon_t^2 + 0.6h_{t-1}$ and $\eta_t \sim$ i.i.d.$N(0,1)$, and use the series $\{\varepsilon_t\}$, to generate the unit root process:

$$y_t = y_{t-1} + \varepsilon_t, \qquad t = 1, \ldots, n,$$

with $y_0 = 0$. Then the ML estimators of $(\phi,\omega,\alpha,\beta)$ and $(\mu,\phi,\omega,\alpha,\beta)$ can be obtained by the estimation procedures in Section 3. Finally, we compute the values of the ML-based test statistics. Such a procedure is repeated for 20,000 independent replications. Using the 20,000 simulated values, the empirical quantiles of the distributions of the ML-based tests are estimated. For $n = 200$, 300 and 5000, some of the empirical quantiles are summarised in Table 1. When $n = 200$ and 300, the critical values are generally smaller than those given in Herce [1996]. The differences are understandable since the finite sample distributions depend on the test statistics themselves, the distribution of errors and the estimation methods, all of which are different from those in Herce [1996]. When $n = 5000$, the critical values of M_ϕ and $M_{\mu,\phi}$ are almost the same as those in Herce [1996], and for M_t and $M_{\mu,t}$, their critical values are very close to those of the standard normal distribution.

In order to investigate the empirical sizes and powers of the test statistics in Section 4, we generate data sets from the following model:

$$y_t = \phi y_{t-1} + \varepsilon_t, \qquad \varepsilon_t = \eta_t\sqrt{h_t},$$
$$h_t = \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1}, \qquad \eta_t \sim \text{i.i.d.N}(0,1),$$

with $\phi = 0.9, 0.95, 0.99, 1.0$, $\omega = 1 - \alpha - \beta$, and $(\alpha,\beta) = (0.2,0.7)$, $(0.3,0.6)$ and $(0.4,0.5)$. Each data set is estimated by model (1.1) with GARCH error (1.3) and model (1.2) with GARCH error (1.3). For model (1.1) with GARCH error (1.3), we first estimate ϕ by LS and then obtain a series of artificial observations of ε_t which are used to estimate (ω,α,β) by the IMSL subroutine DBCOAH. Using these estimators as the initial values, we obtain the ML estimator of $(\phi,\omega,\alpha,\beta)$ by the estimation procedure given in Section 3. A similar estimation procedure is employed for model (1.2) with GARCH error (1.3). For each parameter vector (ϕ,α,β) and (μ,ϕ,α,β), we use 1000 independent replications. The empirical sizes and powers of the eight test statistics, L_ϕ, L_t, $L_{\mu,\phi}$, $L_{\mu,t}$, M_ϕ, M_t, $M_{\mu,\phi}$ and $M_{\mu,t}$, are summarised in Table 2 for $n = 200$ for the 5% significance level. Results are also available for $n = 300$, but are not reported in tabular form.

When $n = 200$, the empirical sizes of the LS-based tests, especially L_t and $L_{\mu,t}$, tend to overreject a true null hypothesis. For the ML-based tests, the sizes are closer to the nominal 5% level, and powers are also acceptable as compared with those reported in other studies under i.i.d. errors (see, for example, Dickey and Fuller [1979]). When $n = 300$ (not reported here), all the test statistics for the fitted model (1.1) with GARCH error (1.3) have similar sizes and powers. However, for the fitted model (1.2) with GARCH error (1.3), the LS-based tests still tend to overreject, which is consistent with the findings in Kim and Schmidt [1993]. In this case, the ML-based tests basically solve the overrejection problem. From Table 2 and the results for $n = 300$ (not reported here), we see that when α increases, the LS-based tests became more sensitive, with increasing sizes and decreasing powers. This phenomenon can be explained by the fact that, when α increases, the values of y_t from the unit root process have increasingly heavy-tailed innovations. Meanwhile, when α increases, the power of the ML-based tests improves because, in this case, $\hat{\sigma}^2\hat{\kappa} - 1$ (or $\tilde{\sigma}^2\tilde{\kappa} - 1$) in the ML-based tests can be evaluated more accurately. All of these results suggest clearly that the ML-based tests are more robust and perform better than their LS-based counterparts.

ACKNOWLEDGEMENTS

The first and third authors wish to acknowledge the financial support of the Australian Research Council.

REFERENCES

Billingsley, P., *Convergence of Probability Measures*, Wiley, New York, 1968.

Bollerslev, T., Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, 31, 307-327, 1986.

Bollerslev, T., R.Y. Chou and K.F. Kroner, ARCH modelling in finance: a review of the theory and empirical evidence, *Journal of Econometrics*, 52, 5-59, 1992.

Bollerslev, T., R.F. Engle and D.B. Nelson, ARCH models, in *Handbook of Econometrics IV*, R.F. Engle and D.L. McFadden (eds.), Elsevier Science, Amsterdam, pp. 2961-3038, 1994.

Chan, N.H., and C.Z. Wei, Limiting distributions of least squares estimates of unstable autoregressive processes, *Annals of Statistics*, 16, 267-401, 1988.

Dickey, D.A., and W.A. Fuller, Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association*, 74, 427-431, 1979.

Fuller, W.A., *Introduction to Statistical Time Series*, Wiley, New York, 1976.

Herce, M.A., Asymptotic theory of LAD estimation in a unit root process with finite variance errors, *Econometric Theory*, 12, 129-153, 1996.

Kim, K., and P. Schmidt, Unit root tests with conditional heteroskedasticity, *Journal of Econometrics*, 59, 287-300, 1993.

Lee, S.-W., and B.E. Hansen, Asymptotic theory for GARCH(1,1) quasi-maximum likelihood estimator, *Econometric Theory*, 10, 29-52, 1994.

Li, W.K., S. Ling and M. McAleer, A survey of recent theoretical results for time series models with GARCH errors, 1999 (submitted).

Li, W.K., and T.K. Mak, On the squared residual autocorrelations in nonlinear time series modelling, *Journal of Time Series Analysis*, 15, 627-636, 1994.

Ling, S., and W.K. Li, Fractional ARIMA-GARCH time series models, *Journal of the American Statistical Association*, 92, 1184-1194, 1997a.

Ling, S., and W.K. Li, Estimation and testing for unit root processes with GARCH(1,1) errors, Technical Report, Department of Statistics, University of Hong Kong, 1997b.

Ling, S., and W.K. Li, The limiting distributions of maximum likelihood estimators for unstable autoregressive moving-average time series with GARCH errors, *Annals of Statistics*, 26, 84-125, 1998.

Ling, S., W.K. Li and M. McAleer, Efficient estimation and testing for unit root processes with GARCH(1,1) errors: a Monte Carlo analysis, 1999 (submitted).

Lucas, A., Unit root tests based on M estimators, *Econometric Theory*, 11, 331-346, 1995.

Lumsdaine, R., Consistency and asymptotic normality of the quasi-maximum likelihood estimator in IGARCH(1,1) and covariance stationary GARCH(1,1) models, *Econometrica*, 64, 575-596, 1996.

Nelson, D.B., Stationarity and persistence in the GARCH(1,1) model, *Econometric Theory*, 6, 318-334, 1990.

Pantula, S.G., Comment on modelling the persistence of conditional variance, *Econometric Reviews*, 5, 71-74, 1986.

Pantula, S.G., Estimation of autoregressive models with ARCH errors, *Sankhya B*, 50, 119-138, 1989.

Peters, T.A., and W. Veloce, Robustness of unit root tests in ARMA models with generalized ARCH errors, Unpublished manuscript, Brock University, Canada, 1988.

Phillips, P.C.B., Time series regression with a unit root, *Econometrica*, 55, 277-301, 1987.

Phillips, P.C.B., and S.N. Durlauf, Multiple time series regression with integrated processes, *Review of Economic Studies*, LIII, 473-495, 1986.

Weiss, A.A., Asymptotic theory for ARCH models: estimation and testing, *Econometric Theory*, 2, 107-131.

Yap, S.F., and G.C. Reinsel, Estimation and testing for unit roots in a partially nonstationary vector autoregressive moving average model, *Journal of the American Statistical Association*, 90, 253-267, 1995.

TABLE 1
Empirical Critical Values for the M_ϕ, M_t, $M_{\mu,\phi}$ and $M_{\mu,t}$ Tests

		Empirical Quantiles							
Statistic	n	.010	.025	.050	.100	.900	.950	.975	.990
M_ϕ	200	-9.78	-6.71	-4.77	-3.09	2.42	3.43	4.40	5.59
	300	-8.96	-6.17	-4.36	-2.92	2.47	3.49	4.52	5.87
	5000	-6.53	-4.98	-3.68	-2.60	2.56	3.73	5.01	6.70
M_t	200	-2.37	-2.00	-1.66	-1.28	1.20	1.56	1.90	2.27
	300	-2.39	-1.98	-1.64	-1.28	1.24	1.60	1.91	2.28
	5000	-2.31	-1.96	-1.64	-1.28	1.24	1.60	1.92	2.31
$M_{\mu,\phi}$	200	-16.41	-11.93	-9.03	-6.25	2.96	4.20	5.30	6.83
	300	-15.04	-11.08	-8.43	-5.73	3.12	4.36	5.56	7.09
	5000	-8.79	-6.90	-5.44	-3.95	3.86	5.31	6.69	8.74
$M_{\mu,t}$	200	-3.54	-2.88	-2.33	-1.81	1.05	1.42	1.75	2.18
	300	-3.38	-2.69	-2.22	-1.70	1.10	1.46	1.80	2.24
	5000	-2.31	-1.95	-1.67	-1.29	1.28	1.64	1.94	2.31

TABLE 2
Sizes and Powers for Unit Root Processes with GARCH(1,1) Errors
n = 200, 1000 Replications, $\omega = 1 - \alpha - \beta$

			ϕ					ϕ			
α	β	Test	0.90	0.95	0.99	1.0	Test	0.90	0.95	0.99	1.0
0.2	0.7	L_ϕ	.996	.848	.132	.064	$L_{\mu,\phi}$	.901	.466	.104	.064
		L_t	.994	.747	.129	.065	$L_{\mu,t}$	.830	.325	.073	.066
		M_ϕ	.799	.556	.133	.061	$M_{\mu,\phi}$	.648	.335	.064	.037
		M_t	.460	.283	.132	.057	$M_{\mu,t}$	.356	.215	.070	.041
0.3	0.6	L_ϕ	.992	.739	.134	.074	$L_{\mu,\phi}$	.892	.470	.113	.077
		L_t	.987	.748	.135	.069	$L_{\mu,t}$	.819	.328	.083	.079
		M_ϕ	.890	.700	.168	.055	$M_{\mu,\phi}$	.786	.783	.089	.058
		M_t	.626	.470	.170	.057	$M_{\mu,t}$	.543	.330	.097	.049
0.4	0.5	L_ϕ	.986	.737	.138	.075	$L_{\mu,\phi}$	.891	.487	.122	.083
		L_t	.981	.743	.144	.070	$L_{\mu,t}$	.810	.355	.093	.084
		M_ϕ	.934	.777	.211	.048	$M_{\mu,\phi}$	.863	.578	.111	.063
		M_t	.741	.528	.212	.049	$M_{\mu,t}$	.653	.440	.119	.059