Working Paper 95-39
Statistics and Econometrics Series 11
September 1995

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (341) 624-9849

NO LACK OF RELATIVE POWER OF THE DICKEY-FULLER TESTS FOR UNIT ROOTS

Jesús Gonzalo and Tae-Hwy Lee*

Abstract

This paper shows numerically that the lack of power and size distortions of the Dickey-Fuller type test for unit roots (very well documented in the unit root literature) are similar to and in many situations even smaller than the lack of power and size distortions of the standard Student-t tests for stationary roots of an AR model.

Key Words
DF Tests; Student-t Tests; Unit Roots.

*Gonzalo, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid; Lee, Department of Economics, University of California, Riverside. We thank In Choi, Miguel Herce, Pierre Perron, Peter Phillips, James Stock, Mark Watson, Farshid Vahid, participants at the 1995 Winter Econometric Society Meetings, and an anonymous referee for helpful comments and suggestions. Any errors are ours.

1 Introduction

If we open a contest to select the most mentioned sentence in time series econometrics or even applied macroeconomics, during the last 10 years, almost certainly the winner would be something like "the lack of power of the unit root test". Many papers have shown numerically this lack of power and also the size distortions of the unit root tests. A partial list includes: Agiakloglou and Newbold (1992), Bierens (1993), DeJong, Nankervis, Savin and Whiteman (1992a, 1992b), Dickey and Fuller (1979, 1981), Elliot, Rothemberg and Stock (1992), Hall (1992), Ng and Perron (1993, 1995), Perron (1989), Said and Dickey (1985), Schmidt and Phillips (1992), Schwert (1989), and a survey by Stock (1995). To the best of our knowledge, no paper has considered whether this lack of power is typical only in the unit root tests or it can also be found in any standard test for stationary roots (for instance, in tests for zero first order correlation). This article claims that the Student-t tests for stationary roots of an autoregressive (AR) model have as bad performance as and sometimes even worse performance than the Dickey-Fuller (DF) t-type tests for unit roots.

This paper does not claim that testing for a unit root is the same as testing for a stationary root. There are two main differences. First, the limit distribution in the unit root case is non-standard and it depends on the specification of the deterministic component of the analyzed variable. Second, in economics, to be able to tell between 1.0 and 0.9 is more important than to be able to distinguish between 0.0 and non-zero or between 0.5 and 0.4. Nonrejection of the hypothesis of a unit root implies the existence of permanent shocks and also the possibility for having spurious regressions. Therefore it is understandable to see all the effort and interest that has been dedicated to the unit root case, but these two differences are not enough to explain the huge amount of papers that have been trying to convince the profession that the DF test has a big lack of power and awful size distortions. This paper only tries to show that the sentence "lack of power of the unit root tests", although accurate in absolute terms, is totally inaccurate in relative terms. The power of a test should be judged in both dimensions.

In order to show our claim, we re-do the experiments produced in the unit root literature, but this time not only to test for a root equals one but also to test for a stationary root of an AR model. We compare the power and size distortions of the t-tests in both scenarios, I(1) and I(0). The analysis concentrates only on the t type tests because they are the most commonly used in practice, specially the DF test for unit roots.

This paper is neither a survey on unit roots nor proposes a new unit root test. The paper is organized as follows. Section 2 introduces the notation and the models used in the Monte Carlo experiment. Section 3 investigates the power of the t tests for the null hypotheses of I(1) and I(0), in a model with AR errors. Section 4 analyzes the size distortion of the t tests for the same null hypotheses as in section 3, first in a model with moving average (MA) errors, and second in a model with heteroskedastic (GARCH) errors. Section 5 concludes.

2 Notation and Models

Let the time series $\{y_t\}$ be the stochastic process generated by the linear model

$$y_t = d_t + x_t \tag{1}$$

$$d_t = \beta_0 + \beta_1 t \tag{2}$$

$$(1 - \alpha L)x_t = u_t \tag{3}$$

$$(1 - \rho L)u_t = (1 - \theta L)e_t \tag{4}$$

where L is the lag operator. We assume that $x_0 = 0$, e_t i.i.d. N(0, σ^2), $\rho \leq 1$ and $\theta < 1$.

Lemma 1 : *Let* $C(L) = \sum_{j=0}^{\infty} c_j L^j$. *Then:*

- *(a)*

$$C(L) = C(1) - (1 - L)\tilde{C}(L) \tag{5}$$

where $\tilde{C}(L) = \sum_{j=0}^{\infty} \tilde{c}_j L^j$, *with* $\tilde{c}_j = \sum_{k=j+1}^{\infty} c_k$.

- *(b) Provided* $\alpha_0 \neq 0$,

$$C(L) = C(1/\alpha_0) - (1 - \alpha_0 L)\tilde{C}^*(L) \tag{6}$$

where $C^*(L) = \sum_{j=0}^{\infty} c_j^* L^{*j}$, *with* $c_j^* = (1/\alpha_0)^j c_j$ *and* $L^* = \alpha_0 L$.

A proof of the algebraic decomposition (a) can be found in Gelfand (1989, p.160). This decomposition was used by Beveridge and Nelson (1981) to decompose economic variables into permanent and transitory components.

The proof of part (b) follows from applying (a) to the polynomial $C^*(L^*)$.

Phillips and Solo (1992) show that a sufficient condition for (5) to make sense ($\sum_{j=0}^{\infty} \tilde{c}_j^2 < \infty$) is that the polynomial $C(L)$ is 1/2-summable,

$$\sum_{j=1}^{\infty} j^{1/2}|c_j| < \infty. \tag{7}$$

It can be shown that condition (7) is sufficient for $\sum_{j=0}^{\infty}(\tilde{c}_j^*)^2 < \infty$. All the cases treated in this paper satisfy condition (7).

Substituting (4), (3), and (2) into (1), using Lemma 1, and rearranging gives the above data generating process (DGP) as a single equation model

$$(1-\alpha_0 L)y_t = \mu_0 + \mu_1 t + \lambda y_{t-1} + \sum_{i=1}^{\infty} \delta_i (1-\alpha_0 L) y_{t-i} + e_t, \tag{8}$$

where the parameter of interest is

$$\lambda = (1-\theta/\alpha_0)^{-1}(1-\rho/\alpha_0)(\alpha-\alpha_0). \tag{9}$$

The null hypothesis that $\{y_t\}$ has an autoregressive root equals α_0,

$$H_o: \ \alpha = \alpha_0$$

can be tested by the t-ratio

$$t_i = \hat{\lambda}/sd(\hat{\lambda}), \quad (i = 1,2,3), \tag{10}$$

where $\hat{\lambda}$ and $sd(\hat{\lambda})$, the standard error of $\hat{\lambda}$, are obtained by applying ordinary least squares (OLS) to the following regressions R1, R2 and R3, respectively,

$$\begin{aligned}
R1: &\quad Dy_t = \lambda y_{t-1} + \textstyle\sum_{i=1}^{p} \delta_i D y_{t-i} + e_t \\
R2: &\quad Dy_t = \mu_0 + \lambda y_{t-1} + \textstyle\sum_{i=1}^{p} \delta_i D y_{t-i} + e_t \\
R3: &\quad Dy_t = \mu_0 + \mu_1 t + \lambda y_{t-1} + \textstyle\sum_{i=1}^{p} \delta_i D y_{t-i} + e_t
\end{aligned}$$

where $D = (1 - \alpha_0 L)$.

The 5% critical values of the t_1, t_2, and t_3 tests, for T=100, are tabulated in Table 1 for different values of α_0. To generate these critical values it is assumed $d_t = 0$, $\rho = \theta = 0$ and p=0. Under the null hypothesis of $\alpha = 1.0$, the tests t_1, t_2 and t_3 have non-standard asymptotic distributions, see Hamilton (1994, p.486) for a complete summary. These distributions depend on the specification of the deterministic component included in the regressions. The distributions are skewed to the left and have too many negative values relative to the Student-t distribution. The asymptotic distribution of t_1, t_2 and t_3 is Normal (0, 1), for testing for a hypothesis of a stationary root ($|\alpha| < 1.0$).

Table 1 about here

The critical values of Table 1 will be used in the next two sections to evaluate the power and size of the t-tests.

3 Power Comparisons

In this section we compare numerically the power of the one sided t-tests (t_1, t_2 and t_3) for non-stationary (NS) and stationary (S) roots. This comparison has been done for different values of the null hypothesis (α_0=1.0, 0.9, ..., 0.0), and different alternatives, ($\alpha_0 - \alpha = 0.3, 0.2, 0.1, 0.05$ and 0.01),

$$\text{(NS) } H_o: \ \alpha_0 = 1 \ versus \ H_a: \ \alpha < 1$$
$$\text{(S) } H_o: \ \alpha = \alpha_0, \ \ \alpha_0 < 1, versus \ H_a: \ \alpha < \alpha_0.$$

In Table 2 the DGP is

$$(1 - \alpha L)y_t = e_t \tag{11}$$

and the regressions R1, R2 and R3 do not contain lags (p=0) of $(1 - \alpha_0 L)y_t$. We only report results for $\alpha_0 = 1.0, 0.9, ..., 0.0$ and $(\alpha_0 - \alpha) = 0.1$ and 0.05. Other results are available upon request.

Table 2 about here

The main feature of Table 2 is the drastic decrease in power of the DF test when the regression contains a trend and/or a constant term. This is a well known result (see Stock (1995)). The reason of this decrease in power is the random collinearity that there exists in the unit root case between a constant and y_{t-1} and between a deterministic trend and y_{t-1}. For the stationary roots the power is very uniform across regressions, and although there is also a random collinearity between a deterministic trend and y_{t-1}, this collinearity does not show up in the limit distribution.

Comparing both situations (NS and S) the only case where the DF test has clearly lower power than the t-test for a stationary root α is when there is a trend in the regression (t_3). Many will argue (see Campbell and Perron (1992)) that this is the relevant case in practice because it is never known whether y_t has a drift or not, and the t_3 test is invariant to that. Therefore in applied research we are forced to run the regression R3. In Table 3, we investigate if this finding of lower power of the DF-t_3 test is robust to other standard misspecifications that occur in practice, like misspecifications in the number of lags p. In the stationary case, the inclusion of irrelevant lags of $(1-\alpha_0 L)y_{t-1}$ introduces collinearity in the regression models. This collinearity is not random but causes similar power problems (under local alternatives) to the ones created by a deterministic trend in the DF-t_3 test.

To make the paper shorter and without loss of generality, we only report (through the rest of the paper) results for the following two cases

$$\text{(NS) } H_o:\ \alpha_0 = 1 \textit{ versus } H_a:\ \alpha = .9$$
$$\text{(S) } H_o:\ \alpha_0 = .5 \textit{ versus } H_a:\ \alpha = .4.$$

The selection of a metric, the distance between the null and the alternative hypotheses, to compare power in NS and S is not clear. There are several distance measures for stochastic processes in the literature (see Zinde-Walsh (1992) for a comparison of some of them). We have selected the one proposed by Piccolo (1990), that in our case is equivalent to the absolute distance ($|\alpha_0 - \alpha| = 0.1$). We could also have chosen as a metric, the standard deviation of the finite sample distribution of $\hat{\alpha}$ under the different null hypotheses. In this case the metric would have depended on T as well as on α_0. The main conclusions of this paper are invariant to the chosen metric.

In Table 3, the DGP is

$$(1-\alpha L)y_t = (1-\rho L)^{-1} e_t, \tag{12}$$

with $\rho = (0.0,\ 0.1,\ ...,\ 1.0)$. By Lemma 1, (12) can be re-written as

$$(1-\alpha_0 L)y_t = (1-\rho/\alpha_0)(\alpha-\alpha_0)y_{t-1} + (\rho\alpha/\alpha_0)(1-\alpha_0 L)y_{t-1} + e_t. \quad (13)$$

Table 3 considers situations where the number of lags p is correctly specified (p=1), underparametrized (p=0) and overparametrized (p=4). The case of p=4 is extremely important because as we mentioned before, the irrelevant variables $(1-\alpha_0 L)y_{t-i}$ (i=2, 3, 4) are correlated with y_{t-1}, if y_t is stationary. This case is in some sense similar to the DF-t_3 where the deterministic trend is correlated with y_{t-1}.

Table 3 about here

Several features of Table 3 stand out. For example, the power of the DF type tests is always higher than the power of the t-ratio for the null hipothesis of a 0.5 root, when there are some lags in the regression model (p=1 or p=4). This is true even for ρ=0. When there are no lags in the regressions, the only situation where the power of the DF is lower is for t_3 with $\rho = 0$ (the result previously mentioned in Table 2 and so much cited in the unit root literature). These results indicate the strong conclusion, that in general the DF-t type tests do not have less power than the t-ratio tests for a stationary root, even if we include a deterministic trend in the regression model.

4 Size Comparisons

The most influential Monte Carlo study in the unit root literature is Schwert's (1989), who found large size distortions in several unit root tests (DF is one of them) when the errors u_t have an MA component. In this section following Schwert's DGP we compare the size distortions of DF tests with the t-tests for the null hipothesis that $\alpha_0 = 0.5$.

The DGP in Table 4 is

$$(1-\alpha_0 L)y_t = (1-\theta L)e_t, \quad (14)$$

with θ = (0.1, 0.3, 0.5, 0.7 and 0.9). Only positive values of θ have been considered, not because we think they are the most relevant values in practice (maybe it is the opposite) but because these are the ones that produce higher size distortions in Schwert's (1989). For the same reason Agiakloglou and Newbold (1992) consider only positive values for θ.

Following Schwert (1989) and Agiakloglou and Newbold (1992) the tests are based on the OLS estimates of the approximating autoregressive regressions R1, R2 and R3. We also report the results of the tests based on the regression models selected by the AIC. The most striking feature of Table 4 is that the t-test for the stationary root ($\alpha_0 = 0.5$) has even larger size distortions than the DF test. This is even true for $\theta = 0.9$, if we choose the autoregressive model selected by the AIC criteria.

Table 4 about here

Recently there has been some concern about the size distortions when there is conditional heteroskedasticity in the errors. Kim and Schmidt (1993) show that the DF tests tend to overreject in the presence of GARCH errors. The DGP in Table 5 is

$$(1 - \alpha_0 L)y_t = e_t \tag{15}$$

$$e_t | I_{t-1} \ is \ N(0, h_t) \tag{16}$$

$$h_t = \phi_0 + \phi_1 e_{t-1}^2 + \phi_2 h_{t-1}, \tag{17}$$

where I_{t-1} is the information available at time t-1. Let $z_t \equiv e_t / h_t^{1/2}$ be i.i.d. N(0, 1) and h_0=1. Nelson (1990) shows that the h_t process has a strictly stationary and ergodic distribution if and only if $\phi_0 > 0$ and $E[ln(\phi_2 + \phi_1 z_t^2)] < 0$. By Jensen's inequality and the strict concavity of $ln(x)$, $E[ln(\phi_2 + \phi_1 z_t^2)] < ln(\phi_2 + \phi_1 E(z_t^2))$. If $E(z_t^2) = 1$, then $\phi_0 > 0$ and $\phi_1 + \phi_2 \leq 1$ are sufficient conditions for the process h_t to be strictly stationary and ergodic. When $\phi_0 > 0$ and $\phi_1 + \phi_2 \leq 1$, it is seen from Table 5 that the size distortions are slightly larger for the stationary root tests than for the unit root tests.

Table 5 about here

Summarizing, the results in this section show that in general the size distortions of the DF type tests are similar or even smaller than the size distortions of the standard t-ratio tests for a stationary root.

5 Conclusions

In this paper we compare numerically the lack of power and size distortions of the DF-t type tests, with the lack of power and size distortions of the standard t-ratio tests for stationary AR roots.

Two clear conclusions emerge from our analysis. First, the DF-t type tests do not have less power than the t-ratio tests for a stationary root, when the number of lags is unknown (in practice always). This is true even if we include a deterministic trend in the regression. In other words, the well known result of lack of power of the DF test when there is a deterministic trend in the regression model is not robust (in relative terms) to correlated errors. Second, the size distortions of the t-tests for stationary roots are as big as the ones of the DF tests for unit roots. This is the case with MA errors as well as GARCH errors.

References

Agiakloglou, C. and Newbold, P. (1992) Empirical evidence on Dickey-Fuller type tests. *Journal of Time Series Analysis* **13**, 471-483.

Beveridge, S and Nelson, C.R. (1981) A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle'. *Journal of Monetary Economics* **7**, 151-174.

Bierens, H.J. (1993) Higher-order sample autocorrelations and the unit root hypothesis. *Journal of Econometrics* **57**, 137-160.

Campbell, J.Y. and Perron, P. (1991) Pitfalls and opportunities: what macroeconomists should know about unit roots. *NBER Macroeconomics Annual*, 141-200.

DeJong, D.N., Nankervis, J.C., Savin, N.E., and Whiteman, C.H. (1992a) Integration versus trend-stationarity in macroeconomic time series. *Econometrica* **60**, 423-434.

DeJong, D.N., Nankervis, J.C., Savin, N.E., and Whiteman, C.H. (1992b) The power problems of unit roots tests for time series with autoregressive errors. *Journal of Econometrics* **53**, 323-343.

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

Dickey, D.A. and Fuller, W.A. (1981) Likelihood ratio tests for autoregressive time series with a unit root. *Econometrica* **49**, 1057-1072.

Elliot, G., Rothemberg, T.J., and Stock J.H. (1992) Efficient tests for an autoregressive unit root. NBER Technical Working Paper No. 130.

Gelfand, I.M. (1961) *Lectures on Linear Algebra.* New York: Dover Publications.

Hall, A. (1992) Testing for a unit root in time series with pretest data-based model selection. *Journal of Business & Economic Statistics* **12**, 461-470.

Hamilton, J. (1994), *Time Series Analysis.* Princeton University Press.

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

Nelson, D.B. (1990) Stationary and Persistence in the GARCH(1, 1) Model. *Econometric Theory* **6**, 318-334.

Ng, S. and Perron, P. (1993) Useful modifications to some unit root tests with dependent errors and their local asymptotic properties. Manuscript, C.R.D.E, University of Montreal, Quebec.

Ng, S. and Perron, P. (1995) Unit root tests in ARMA models with data-dependent methods for the selection of the truncation lag. *Journal of the American Statitical Association* **90**, 268-281.

Perron, P. (1989) Testing for a random walk: a simulation experiment of power when the sampling interval is varied. Ch. 3 in Raj, B. (ed.) *Advances in Econometrics and Modelling.* Boston: Kluwer Academic Publishers.

Phillips, P.C.B. and Solo, V. (1992) Asymptotics for linear processes. *The Annals of Statistics* **20**, 971-1001.

Piccolo, D. (1990) A distance measure for classifying ARIMA models. *Journal of Time Series Analysis* **11**, 153-164.

Said, E.S. and Dickey, D.A. (1984) Testing for unit roots in autoregressive moving average models of unknown order. *Biometrika* **71**, 599-607.

Schmidt, P. and Phillips, P.C.B. (1992) Testing for unit root in the presence of deterministic trends. *Oxford Bulletin of Economics and Statistics* **54**, 257-288.

Schwert, G.W. (1989) Tests for unit roots: a monte carlo investigation. *Journal of Business & Economic Statistics* **7**, 147-159.

Stock, J.H. (1995) Unit roots, structural breaks and trends. *Handbook of Econometrics*, Volume **4**, Engle R.F. and McFadden, D. (ed.). North Holland.

Zinde-Walsh, V. (1992) Relations between some distance measures for stochastic processes. McGill Univeristy Working Paper, 6-92.

Table 1. Critical Values (5% level, T = 100)

	t_1	t_2	t_3
α = 1.0	-1.933	-2.889	-3.451
α = 0.9	-1.795	-2.209	-2.609
α = 0.8	-1.764	-2.050	-2.339
α = 0.7	-1.739	-1.965	-2.197
α = 0.6	-1.724	-1.915	-2.113
α = 0.5	-1.716	-1.878	-2.048
α = 0.4	-1.704	-1.847	-1.993
α = 0.3	-1.689	-1.819	-1.948
α = 0.2	-1.674	-1.792	-1.912
α = 0.1	-1.664	-1.770	-1.880
α = 0.0	-1.657	-1.756	-1.846

Notes: $\{y_t\}$ is generated by $y_t = \alpha y_{t-1} + e_t$, e_t is i.i.d. N(0, 1), $t = 1, \ldots, T = 100$, and $y_1 = 0$. Three t-ratios for the OLS estimate of λ are computed: t_1 from the regression $(1 - \alpha L)y_t = \lambda y_{t-1} + e_t$, t_2 from the regression $(1 - \alpha L)y_t = \mu_0 + \lambda y_{t-1} + e_t$, and t_3 from $(1 - \alpha L)y_t = \mu_0 + \mu_1 t + \lambda y_{t-1} + e_t$. The 5% critical values are computed from 50000 replications.

Table 2. Power of testing various autoregressive roots

$\alpha - \alpha_0 = -0.1$	t_1	t_2	t_3	$\alpha - \alpha_0 = -0.05$	t_1	t_2	t_3
α_0=1.0, α=0.9	.7766	.3083	.1826	α_0=1.0, α=0.95	.3356	.1150	.0836
α_0=0.9, α=0.8	.4659	.3789	.3032	α_0=0.9, α=0.85	.1966	.1601	.1334
α_0=0.8, α=0.7	.3605	.3277	.2846	α_0=0.8, α=0.75	.1569	.1430	.1314
α_0=0.7, α=0.6	.3140	.2964	.2691	α_0=0.7, α=0.65	.1416	.1351	.1273
α_0=0.6, α=0.5	.2878	.2737	.2564	α_0=0.6, α=0.55	.1316	.1305	.1237
α_0=0.5, α=0.4	.2693	.2610	.2470	α_0=0.5, α=0.45	.1269	.1255	.1208
α_0=0.4, α=0.3	.2608	.2531	.2425	α_0=0.4, α=0.35	.1263	.1235	.1200
α_0=0.3, α=0.2	.2534	.2466	.2435	α_0=0.3, α=0.25	.1248	.1226	.1217
α_0=0.2, α=0.1	.2513	.2490	.2445	α_0=0.2, α=0.15	.1241	.1236	.1205
α_0=0.1, α=0.0	.2534	.2508	.2470	α_0=0.1, α=0.05	.1262	.1231	.1197
α_0=0.0, α=-0.1	.2564	.2495	.2515	α_0=0.0, α=-0.05	.1254	.1231	.1224

Notes: 5% level. T = 100. 10000 replications. $d_t = 0$ ($\beta_0 = \beta_1 = 0$). $\rho = \theta = 0$. α is the true autoregressive root in the DGP (i.e., under the alternative hypothesis), and α_0 is the value of α claimed under the null hypothesis, H_0: $\alpha = \alpha_0$. The difference $(\alpha - \alpha_0)$ measures the departure from the null hypothesis. The critical values in Table 1 are used.

Table 3. Power with AR(1) errors

	$\alpha = 0.9$, $\alpha_0 = 1.0$			$\alpha = 0.4$, $\alpha_0 = 0.5$		
	t_1	t_2	t_3	t_1	t_2	t_3
Using $p = 0$						
ρ = 0.0	.7766	.3083	.1826	.2693	.2610	.2470
0.1	.6251	.1711	.0856	.0648	.0627	.0606
0.3	.2767	.0273	.0120	.0005	.0005	.0004
0.5	.0526	.0018	.0013	.0000	.0000	.0000
0.7	.0026	.0000	.0011	.0000	.0000	.0000
0.9	.0000	.0012	.0092	.0000	.0000	.0000
1.0	.0000	.1345	.0481	.0000	.0000	.0000
Using $p = 1$						
ρ = 0.0	.7409	.2997	.1804	.1347	.1333	.1297
0.1	.7320	.2961	.1778	.1219	.1199	.1155
0.3	.7071	.2824	.1716	.0862	.0861	.0851
0.5	.6666	.2589	.1624	.0513	.0549	.0540
0.7	.5855	.2230	.1364	.0263	.0290	.0302
0.9	.3428	.1308	.0961	.0088	.0098	.0112
1.0	.0551	.0673	.0614	.0033	.0043	.0070
Using $p = 4$						
ρ = 0.0	.6277	.2361	.1405	.0535	.0561	.0549
0.1	.6270	.2342	.1394	.0531	.0542	.0563
0.3	.6044	.2246	.1363	.0491	.0528	.0530
0.5	.5743	.2105	.1310	.0473	.0501	.0513
0.7	.5123	.1880	.1185	.0445	.0473	.0494
0.9	.3185	.1221	.0899	.0421	.0452	.0469
1.0	.0552	.0703	.0704	.0354	.0359	.0428

Notes: 5% level. $T = 100$. 10000 replications. $d_t = 0$ ($\beta_0 = \beta_1 = 0$). $\theta = 0$. α is the true autoregressive root in the DGP (i.e., under the alternative hypothesis), and α_0 is the value of α claimed under the null hypothesis, H_0: $\alpha = \alpha_0$. The difference $(\alpha - \alpha_0)$ measures the departure from the null hypothesis. Three t-ratios for λ are computed: t_1 from the regression $Dy_t = \lambda y_{t-1} + \Sigma_{i=1}^{p}\delta_i Dy_{t-i} + e_t$, t_2 from the regression $Dy_t = \mu_0 + \lambda y_{t-1} + \Sigma_{i=1}^{p}\delta_i Dy_{t-i} + e_t$, and t_3 from $Dy_t = \mu_0 + \mu_1 t + \lambda y_{t-1} + \Sigma_{i=1}^{p}\delta_i Dy_{t-i} + e_t$, where $D = 1 - \alpha_0 L$. p is the number of lagged Dy_{t-i} in the regressions. If $p = 0$ is used when $\rho > 0$, the number of lags is under-parameterized. If $p = 1$, the number of lags is correctly parameterized. If $p > 1$, it is over-parameterized. The critical values in Table 1 are used.

Table 4. Size with MA(1) errors

	α_0 = 1.0			α_0 = 0.5		
	t_1	t_2	t_3	t_1	t_2	t_3
Using $p = p_{aic}$						
θ = 0.1	.0723	.0775	.1007	.2498	.2538	.2730
	(0.98)	(1.00)	(1.03)	(0.77)	(0.82)	(0.90)
	[1.72]	[1.80]	[1.91]	[1.67]	[1.72]	[1.86]
0.3	.0810	.0963	.1486	.7105	.7062	.7264
	(1.81)	(1.78)	(1.66)	(0.80)	(0.84)	(0.89)
	[1.68]	[1.74]	[1.85]	[1.67]	[1.72]	[1.83]
0.5	.1060	.1301	.2077	.8840	.8816	.8964
	(2.57)	(2.48)	(2.16)	(0.77)	(0.81)	(0.92)
	[1.69]	[1.74]	[1.79]	[1.66]	[1.71]	[1.85]
0.7	.1735	.2519	.4343	.9532	.9523	.9531
	(3.62)	(3.22)	(2.37)	(1.62)	(1.66)	(1.89)
	[1.89]	[1.95]	[1.96]	[2.01]	[2.04]	[2.19]
0.9	.4364	.7596	.9136	.9766	.9729	.9673
	(3.79)	(2.00)	(1.02)	(4.13)	(4.14)	(4.29)
	[2.92]	[2.48]	[1.88]	[2.47]	[2.47]	[2.51]
Using p = 0						
θ = 0.1	.0813	.0832	.0954	.2248	.2239	.2177
0.3	.2137	.2376	.3551	.8863	.8794	.8722
0.5	.4465	.5847	.7987	.9996	.9996	.9994
0.7	.7916	.9371	.9952	1.0000	1.0000	1.0000
0.9	.9946	1.0000	1.0000	1.0000	1.0000	1.0000
Using p = 1						
θ = 0.1	.0531	.0506	.0578	.0658	.0650	.0636
0.3	.0829	.0788	.1039	.2092	.2046	.2007
0.5	.1889	.2100	.3082	.7226	.7080	.6868
0.7	.4686	.6203	.8170	.9902	.9872	.9833
0.9	.8771	.9986	1.0000	.9998	.9997	.9994
Using p = 4						
θ = 0.1	.0531	.0542	.0536	.0467	.0505	.0507
0.3	.0553	.0543	.0534	.0499	.0527	.0525
0.5	.0648	.0618	.0645	.0791	.0773	.0778
0.7	.1311	.1334	.1711	.2696	.2549	.2423
0.9	.4429	.7259	.7891	.7430	.6882	.6302
Using p = 9						
θ = 0.1	.0585	.0619	.0621	.0581	.0615	.0621
0.3	.0598	.0614	.0605	.0586	.0585	.0613
0.5	.0591	.0617	.0578	.0598	.0612	.0614
0.7	.0652	.0689	.0682	.0743	.0770	.0779
0.9	.1694	.2486	.2506	.2697	.2387	.2077

Notes (for Table 4): 5% level. T = 100. 10000 replications. $d_t = 0$ ($\beta_0 = \beta_1 = 0$). $\rho = 0$. $\alpha = \alpha_0$ to examine the size. The number of lags p_{aic} is chosen using the AIC among p = 0 to 9. When $p = p_{aic}$ is used, the mean of p_{aic} and the standard deviation of p_{aic} in 10000 replications are reported in ( ) and [ ], respectively. The 95% confidence interval of the empirical size is (0.0456, 0.0544), since if the true nominal size is s (s = 0.05), the observed size follows the asymptotic normal distribution with mean s and variance $s(1 - s)/10000$ for 10000 replications. The critical values in Table 1 are used.

Table 5. Size with GARCH(1, 1) errors

$(\phi_0 \quad \phi_1 \quad \phi_2)$	$\alpha_0 = 1.0$ t_1	t_2	t_3	$\alpha_0 = 0.5$ t_1	t_2	t_3
When $\phi_1 + \phi_2 < 1$ and $\phi_0 = 1 - \phi_1 - \phi_2$						
(1 0 0)	.0493	.0500	.0503	.0481	.0490	.0483
(0.1 0.3 0.6)	.0518	.0761	.0777	.1011	.1007	.0998
(0.05 0.3 0.65)	.0499	.0825	.0824	.1063	.1055	.1041
(0.05 0.1 0.85)	.0506	.0620	.0586	.0610	.0617	.0626
When $\phi_1 + \phi_2 = 1$ and $\phi_1 = 0.3$						
(0.0001 0.3 0.7)	.0412	.0956	.0939	.1131	.1112	.1074
(0.01 0.3 0.7)	.0505	.0956	.0939	.1131	.1112	.1074
(1 0.3 0.7)	.0535	.0956	.0939	.1131	.1112	.1074
(100 0.3 0.7)	.0535	.0956	.0939	.1131	.1112	.1074
When $\phi_1 + \phi_2 = 1$ and $\phi_1 = 0.1$						
(0.0001 0.1 0.9)	.0469	.0710	.0657	.0647	.0642	.0647
(0.01 0.1 0.9)	.0503	.0627	.0597	.0635	.0642	.0646
(1 0.1 0.9)	.0493	.0623	.0599	.0637	.0642	.0642
(100 0.1 0.9)	.0497	.0623	.0599	.0637	.0642	.0642

Notes: 5% level. $T = 100$. 10000 replications. $d_t = 0$ ($\beta_0 = \beta_1 = 0$). $\rho = \theta = 0$. $\alpha = \alpha_0$ to examine the size. $h_t = \phi_0 + \phi_1 e_{t-1}^2 + \phi_2 h_{t-1}$. $h_0 = 1$. The critical values in Table 1 are used. ϕ_0 is simply a scale parameter when $\phi_0 > 0$. This is true only if a fairly large number of initial observations are discarded. When ϕ_0 is very small, $h_0 = 1$ is initialized too far in the right tail of the stationary distribution of the process h_t, and so h_t tends to decline as t gets large. Thus, in order for the results not to depend on ϕ_0 for fixed h_0, the first 500 observations were discarded.