More Powerful Unit Root Tests with Non-normal Errors

Kyung So Im Junsoo Lee Margie A. Tieslau*

this version, February 12, 2011

Abstract

This paper proposes new unit root tests that can become more powerful when the error term follows a non-normal distribution. The improved power is gained by utilizing the additional moment conditions embodied in non-normal errors. We adopt a two-step procedure based on the "residual augmented least squares" (RALS) methodology which can make use of non-linear moment conditions. The testing procedure is simple and does not require information on functional forms or the distribution of the error term. The RALS-based unit root tests show significant power gains over traditional unit root tests which ignore valuable information contained in non-normal errors.

JEL Classification: C22, C12, C13.

Key Words: Unit root test, Generalized methods of moments (GMM), Residual augmented least squares (RALS), Non-normality.

*Kyung So Im, Researcher, Division of Insurance and Research, Federal Deposit Insurance Corporation (FDIC), 550 17th Street, NW, Washington, DC (e-mail: kim@fdic.gov); Junsoo Lee (Corresponding coauthor), Professor, Department of Economics, Finance and Legal Studies, University of Alabama, Box 870224, Tuscaloosa, AL 35487 (email: jlee@cba.ua.edu); and Margie Tieslau, Associate Professor, Department of Economics, University of North Texas, 1155 Union Circle #311457, Denton, TX 76203 (email: tieslau@unt.edu)

1 Introduction

In this paper, we suggest new unit root tests that utilize the information contained in non-normal errors. They are more powerful when the error term follows a non-normal distribution. The search for more powerful unit root tests is not a trivial concern since traditional unit root tests have relatively low power. Our approach deviates from the usual practice that ignores the information in non-normal errors. While it is true that the limiting distribution of the usual unit root tests is not affected by non-normality of the errors, this result does not necessarily imply that the information embodied in non-normal errors is useless. This paper shows that more powerful unit root tests can be obtained in the presence of non-normality by utilizing the information in the higher moments of the errors—moments that are not used in usual unit root tests.

To achieve the efficiency gains and improved power when constructing our tests, we extend the work of Im and Schmidt (2008) and adopt a simple two-step procedure based on the "residual augmented least squares" (RALS) methodology. Im and Schmidt (2008) considered a family of models for which the estimator is consistent at rate $\sqrt{T}$. However, we examine the issue of testing for a unit root and consider the estimator having a higher rate of convergence (T), which is associated with non-stationary data. We first consider generalized method of moments (GMM) estimators utilizing non-linear moment conditions to test for a unit root, and show that the linearized RALS estimators using the same moment conditions are asymptotically equivalent to the GMM estimators. This result is useful since it permits us to use a simplified testing procedure using least squares estimation. Then, we show that the new RALS-based unit root tests yield substantial power gains.

When dealing with real-world data, it is common to find non-normality. Indeed, since traditional unit root tests ignore the valuable information contained in non-normal errors, it would be prudent to search for ways to utilize the information and achieve improved power in this setting. Non-normal distributions can occur for a variety of reasons, and this phenomenon may not be easily distinguished from some forms of non-linearity. For example, many financial time series variables have fat-tailed or leptokurtic distributions, which often are modeled in a non-linear framework. In addition, some financial variables are characterized by skewed distributions, which can occur when an asymmetric relationship exists in the data. In such instances, the non-linear exponential smooth transition autoregressive (ESTAR) or logistic smooth transition autoregressive (LSTAR) models often are applied. Furthermore, some economic time series variables have a mixture of different distributions, which typically would be modeled in a non-linear framework such as regime switching models. Clearly, these examples illustrate that many cases of non-normality may be attributed to non-linearity. If a specific nonlinear form is known, it would be possible to utilize non-linear tests using the specific information. But, if such information is not available, one would be interested in looking for an alternative option. It is encouraging that our RALS-based tests are fairly robust to various forms of unknown non-linearity.

A handful of authors previously have investigated the possibility of utilizing the information contained in non-normal errors when testing for a unit root. For example, Cox and Llatas (1991) studied the asymptotic distribution of maximum likelihood estimators (MLE) in the DF regression assuming that the true error density is known. Lucas (1995) derived the asymptotic distribution of the unit root test statistic based on the M-estimator. Shin and So (1999) considered adaptive maximum likelihood estimators. The results in these papers are important. However, one potential limitation exists. These tests require specific information on a particular density function for the error term, the score function, or a specific non-linear functional form. This information is generally unknown *a priori*. Our suggested tests differ from these other approaches as we do not require such informations. In this regard, our approach offers a meaningful practical advantage. Another major advantage of our approach is the use of a linear framework that relies on least squares estimation. Our RALS-based unit root tests make use of non-linear moment conditions through a computationally simple procedure. Thus, we do not need a non-linear optimization procedure or convergence of iterations. Moreover, the RALS-based tests show comparable in most cases and better performance in some cases that were examined.

Our RALS-based tests are closely related to the pioneering work of Hansen (1995), who suggested augmenting the unit root testing equation with stationary covariates, if available, to gain increased power. In so doing, the error variance of the regression augmented with the stationary covariates will be smaller than that of the usual Dickey-Fuller regression. Wooldridge (1993) and Qian and Schmidt (1999) also noted that it is possible to increase efficiency of estimation by augmenting the testing equation with variables that are correlated with the error term. Hansen's methodology requires stationary covariates that are correlated with the error term, but uncorrelated with the regressors. However, one potential point of difficulty is that it often is not easy to find such variables. Our RALS-based tests are useful even in such cases since we can make use of information contained in the series itself. If outside stationary covariates satisfying the requirement are available, the RALS-based tests can additionally utilize them to further increase their power. Hansen (1995) shows that the asymptotic distribution of the unit root test using stationary covariates is a mixture of the Dickey-Fuller distribution and the standard normal distribution. The RALS-based test statistic has the same asymptotic distribution.

The rest of the paper is organized as follows. In Section 2, we derive the asymptotic distribution of the GMM-based unit root tests. In Section 3, we propose the RALS-based unit root tests and provide the asymptotic distribution when the errors are non-normal. In Section 4, we provide simulation results to examine the performance of the RALS-based unit root tests and compare them with other tests. Section 5 provides an empirical example and section 6 provides concluding remarks.

2 GMM Unit Root Test Statistics

We first consider GMM estimators utilizing some moment restrictions for unit root tests. We wish to examine the asymptotic distributions of the GMM estimators as well as their associated t-statistics. Consider a time series that follows

$$y_t = \phi y_{t-1} + \varepsilon_t,\ t = 1, 2, .., T, \tag{1}$$

where $\{\varepsilon_t\}_{t=1}^{\infty}$ is a sequence of innovations. For the unit root hypothesis, we are interested in testing $H_0 : \phi = 1$ against the alternative hypothesis $H_A : \phi < 1$. We assume:

Assumption 1. $\varepsilon_t = \sum_{j=1}^{p} a_j \varepsilon_{t-j} + e_t$, $t = 1, 2, .., T$, where $\{e_t\}_{t=1}^{\infty}$, is an *iid* sequence with zero mean and a finite second moment σ_e^2, and all roots of $a(z) = 1 - \sum_{j=1}^{p} a_j z^j$ lie outside of the unit circle.

When Assumption 1 is met, one may consider the Dickey-Fuller testing equation

$$\Delta y_t = \beta y_{t-1} + \sum_{j=1}^{p} \delta_j \Delta y_{t-j} + e_t, \; t = 1, 2, .., T, \tag{2}$$

where $\Delta y_t = y_t - y_{t-1}$. Let $\hat{\beta}_{LS}$ be the least squares estimator of β in regression (2). We denote t_{LS} as its t-statistic. Then it is well known, under the null hypothesis, that

$$T\hat{\beta}_{LS} \Rightarrow a(1) \left(\int_0^1 W(r)^2 dr \right)^{-1} \int_0^1 W(r) dW(r), \tag{3}$$

and

$$t_{LS} \Rightarrow \left(\int_0^1 W(r)^2 dr \right)^{-1/2} \int_0^1 W(r) dW(r) = DF, \tag{4}$$

where $a(1) = 1 - \sum_{j=1}^{p} a_j$, and $W(r)$ is the standard Brownian motion on $r \in [0, 1]$.

Let $\xi_t = (\Delta y_{t-1}, \Delta y_{t-2}, ..., \Delta y_{t-p})'$, and $z_t = (y_{t-1}, \xi_t')'$. Suppose we have $J \times (p+1)$ additional moment conditions

$$E\left[g(e_t) \otimes z_t\right] = 0, \; t = 1, 2, ..., \tag{5}$$

where $g(e_t)$ is a $J \times 1$ vector that satisfies the following assumption.

Assumption 2. $g(\cdot)$ is differentiable and satisfies the first-order Lipshitz condition

$\left|g_j'(x) - g_j'(y)\right| < M\left|x - y\right|$ for some constant M for all j, where $g_j(\cdot)$ is the j-th element of $g(\cdot)$. Also, $E\left[g(e_t)\right] = 0$, the second moment of $g(e_t)$ exists, and $E\left[g'(e_t)\right] < \infty$.

Define $C = E\left[g(e_t)g(e_t)'\right]$ and $D = E\left[g'(e_t)\right]$, and $\psi(e_t) = D'C^{-1}g(e_t)$, for $t = 1, 2, .., T$. Also define the correlation between e_t and $\psi(e_t)$ as

$$\rho = \frac{\sigma_{\psi e}}{\sigma_\psi \sigma_e} \tag{6}$$

where $\sigma_\psi^2 = Var\left[\psi(e_t)\right] = Var\left[D'C^{-1}g(e_t)\right] = D'C^{-1}D$, and $\sigma_{\psi e} = E\left[\psi(e_t)e_t\right] = DC^{-1}E\left[g(e_t)e_t\right]$.

We let $\tilde{\beta}_G$ denote the GMM estimator using the moments conditions (5) in the ADF regression (2). The asymptotic distributions of $\tilde{\beta}_G$ and its corresponding t-statistic are given as shown below.

Theorem 1. Suppose that a time series follows (1), and Assumptions 1 and 2 are satisfied. Under the null hypothesis,

$$T\tilde{\beta}_G \Rightarrow \frac{a(1)}{\sigma_e \sigma_\psi} \left(\int_0^1 W_1(r)^2 dr \right)^{-1} \int_0^1 W_1(r) dW_2, \tag{7}$$

where $[W_1(r), W_2(r)]'$ is a bivariate Brownian motion with correlation ρ. The corresponding t-statistic is given as $t_G = \tilde{\beta}_G / se(\tilde{\beta}_G)$, where

$$se(\tilde{\beta}_G) = \tilde{\sigma}_\psi^{-1} \sqrt{\left(\sum_{t=1}^T y_{t-1}^2 - \sum_{t=1}^T y_{t-1}\xi_t \left(\sum_{t=1}^T \xi_t \xi_t' \right)^{-1} \sum_{t=1}^T \xi_t' y_{t-1} \right)^{-1}},$$

with $\tilde{\sigma}_\psi^2 = \tilde{D}'\hat{C}^{-1}\tilde{D}$, $\tilde{D} = T^{-1}\sum_{t=1}^T g'(\tilde{e}_t)$, and $\tilde{C} = T^{-1}\sum_{t=1}^T g(\tilde{e}_t) g(\tilde{e}_t)'$; and where $\tilde{e}_t$ is the residual from GMM estimation of regression (2). Then, we have

$$t_G \Rightarrow \rho DF + \sqrt{1-\rho^2} Z, \tag{8}$$

where ρ is defined in (6), DF denotes the Dickey-Fuller distribution as defined in (4), and Z signifies the standard normal distribution.

proof. See the Appendix.

In the case where an intercept is allowed in the model, we use the regression

$$\Delta y_t = \alpha_1 + \beta y_{t-1} + \sum_{j=1}^p \delta_j \Delta y_{t-j} + e_t, \ t = 1, 2, .., T, \tag{9}$$

and we have the additional moment conditions $E\left[g(e_t) \otimes (1, z_t)'\right] = 0$. In view of the expression for the estimator of β in (A.9) of the Appendix, this produces the GMM estimator that is given by

$$T\tilde{\beta}_{G,\mu} = \left(\sigma_\psi^2 T^{-2} \sum_{t=1}^T \tilde{y}_{t-1}^2 \right)^{-1} T^{-1} \sum_{t=1}^T \tilde{y}_{t-1} \psi(e_t) + o_p(1),$$

where $\tilde{y}_{t-1} = y_{t-1} - T^{-1}\sum_{t=1}^{T} y_{t-1}$, $t = 1, 2, .., T$. Consequently, we have

$$T\tilde{\beta}_{G,\mu} \Rightarrow \frac{a(1)}{\sigma_\psi \sigma_e} \int_0^1 \tilde{W}_1(r)dW_2(r) / \int_0^1 \tilde{W}_1(r)^2 dr, \tag{10}$$

where $\tilde{W}_1(r)$ is the demeaned Brownian motion: $\tilde{W}_1(r) = W_1(r) - \int_0^1 W_1(r)dr$. Also, by construction, we have

$$t_{G,\mu} \Rightarrow \rho DF_\mu + \sqrt{1-\rho^2} Z, \tag{11}$$

where DF_μ denotes the limiting distribution of the t-statistic from least squares in regression (9).

Similarly, when the model includes a linear time trend and an intercept, we use the regression

$$\Delta y_t = \alpha_1 + \alpha_2 t + \beta y_{t-1} + \sum_{j=1}^{p} \delta_j \Delta y_{t-j} + e_t, \; t = 1, 2, .., T, \tag{12}$$

and this will result in the GMM estimator that follows

$$T\tilde{\beta}_{G,\tau} \Rightarrow \frac{a(1)}{\sigma_\psi \sigma_e} \left(\int_0^1 \breve{W}_1(r)^2 dr \right)^{-1} \int_0^1 \breve{W}_1(r) d\breve{W}_2(r), \tag{13}$$

where $\breve{W}(r)$ is the detrended Brownian motion. Again, we have

$$t_{G,\tau} \Rightarrow \rho DF_\tau + \sqrt{1-\rho^2} Z, \tag{14}$$

where DF_τ denotes the limiting distribution of the t-statistic for the OLS estimator of β in the regression (12).

Remark 1. Each of the asymptotic distributions of t_G, $t_{G,\mu}$, and $t_{G,\tau}$ depends on the nuisance parameter ρ. Hansen (1995) reports the critical values of the asymptotic distribution of these t-statistics for $\rho^2 = 0.1$ to 1.0, at increments of 0.1.

3 RALS Unit Root Tests

We next explain the RALS estimator. We first consider the model with an intercept as in (9), and use $x_t = (1, z_t')'$. We let $g(e_t) = \left(e_t, [h(e_t) - K]'\right)'$ and consider the moment condition $E\left[g\left(e_t\right)\otimes x_t\right]=0$. We can split this moment condition into two parts. The first part is the usual moment condition of least squares estimation

$$E\left(e_t\otimes x_t\right)=0. \tag{15}$$

The second part involves an additional $2(J-1)$ moment conditions given by

$$E\left[\left(h\left(e_t\right)-K\right)\otimes x_t\right]=0. \tag{16}$$

Therefore, we have:

$$C=\begin{bmatrix}\sigma_e^2 & C_{21}'\\ C_{21} & C_{22}\end{bmatrix},\text{ and } D=\begin{bmatrix}1\\ D_2\end{bmatrix}, \tag{17}$$

where $C_{21}=E\left[e_t h(e_t)\right]$, $C_{22}=E\left[h(e_t)h(e_t)'\right]$, and $D_2=E\left[h'\left(e_t\right)\right]$.Then, we define

$$\hat{w}_t=h(\hat{e}_t)-\hat{K}-\hat{e}_t\hat{D}_2,\ t=1,2,..,T, \tag{18}$$

where $\hat{e}_t$ is the OLS residual from regression (9), $\hat{K}=\frac{1}{T}\sum_{t=1}^{T}h(\hat{e}_t)$, and $\hat{D}_2=\frac{1}{T}\sum_{t=1}^{T}h'(\hat{e}_t)$. The RALS-based testing equation is given by

$$\Delta y_t=\alpha_1+\beta y_{t-1}+\sum_{j=1}^{p}\delta_j\Delta y_{t-j}+\hat{w}_t'\gamma+v_t,\ t=1,2,..,T. \tag{19}$$

The RALS estimator is obtained through least squares estimation applied to (19). We denote the estimator of β as $\tilde{\beta}_{R,\mu}$, and the corresponding t-statistic for $\beta=0$ is denoted as $t_{R,\mu}$. In the following, we show that the RALS estimator is asymptotically identical to the GMM estimator using moment conditions (15) and (16).

Theorem 2. Suppose that a time series follows (1) with $\phi=1$. Under Assumptions 1 and 2, the RALS estimator $\tilde{\beta}_{R,\mu}$ from (19) is asymptotically equivalent to the GMM estimator $\tilde{\beta}_{G,\mu}$ using moment conditions (15) and (16). In addition, the limiting distribution of the RALS-based t-statistic $t_{R,\mu}$ is the same as that of the corresponding GMM t-statistic $t_{G,\mu}$.

proof. See the Appendix.

When a linear time trend is included in the regression, we use

$$\Delta y_t = \alpha_1 + \alpha_2 t + \beta y_{t-1} + \sum_{j=1}^{p} \delta_j \Delta y_{t-j} + \hat{w}_t' \gamma + v_t, \ t = 1, 2, .., T. \tag{20}$$

By construction, the RALS estimator $\tilde{\beta}_{R,\tau}$ and t-statistic $t_{R,\tau}$ will have the same distributions as the corresponding GMM estimator $\tilde{\beta}_{G,\tau}$ and t-statistic $t_{G,\tau}$, which are given in (13) and (14), respectively. Note in passing that we can obtain similar results for the RALS estimator $\tilde{\beta}_R$ and t-statistic t_R, when a basic model without an intercept and a trend is employed.

Next, we provide some guidance on how to apply the RALS procedure in practice.

- ρ^2 is estimated by

  $$\hat{\rho}^2 = \hat{\sigma}_A^2 / \hat{\sigma}^2,$$

  where $\hat{\sigma}^2$ is the usual estimate of the error variance in the standard ADF regression, and $\hat{\sigma}_A^2$ is the estimate of the error variance in the RALS regression in (19) and (20). See the proof of Theorem 2 [equations (A.16) and (A.19)]. Using the estimated value $\hat{\rho}^2$, we can use the same critical values reported in Hansen (1995). Note that our method corresponds to the case where $\rho^2 = R^2$ in Hansen (1995, p. 1151).[1]

- When the sample size is small (e.g. $T \leq 50$), one may impose the restriction of $\beta = 0$ in the first step regression that yields the residuals for the augmented variables in $\hat{w}_t$. According to our simulations, this procedure improves the size property of the test

[1] The Gaussian power envelope for these unit root tests is discussed in Hansen (1995) and Elliott and Jansson's (2003). The limiting distribution of RALS tests using additional covariates is not changed under this situation. One may possibly consider an extended version of the RALS tests using an alternative detrending method, for example, such as a DF-GLS detrending method against a point alternative with $\phi = 1 + c_0/T$, just as in Elliott and Jansson's (2003) extension of Hansen's (1995) work. Additional power gains are expected for the RALS-based tests, although these power gains will depend on whether or not the initial value of a time series is large. In addition, there will be an issue of determining the value of c_0 depending on the estimate of $\hat{\rho}^2$ since these are two nuisance parameters upon which the test statistics will depend. Fixing the value of c_0 independently of the estimate of $\hat{\rho}^2$ is one option, as done in Elliott and Jansson. These issues are beyond the scope of this paper, and we leave them for future research.

with only minimal effects on power. When the sample is relatively big (e.g., $T = 100$), however, this effect, disappears quickly.

4 Simulation Results

In this section, we investigate small sample properties of the RALS unit root tests. For example, suppose we have a testing regression $\Delta y_t = \alpha_1 + \beta y_{t-1} + \sum_{j=1}^{p} \delta_j \Delta y_{t-j} + e_t$, $t = 1, 2, .., T$. In the first step, we take the residuals from this standard Dickey-Fuller regression and use them to construct $\hat{w}_t$ as a function of the residuals, $\hat{e}_t = \Delta y_t - \hat{\alpha}_1 - \tilde{\beta} y_{t-1} - \sum_{j=1}^{p} \hat{\delta}_j \Delta y_{t-j}$. When the sample size is small, we estimate α_1 and δ by imposing $\beta = 0$ and construct the augmented variable $\hat{w}_t$ from the residuals of the restricted regression. Then, in the second step, the t-statistic on $\beta = 0$ is computed using the augmented RALS regression: $\Delta y_t = \alpha_1 + \beta y_{t-1} + \sum_{j=1}^{p} \delta_j \Delta y_{t-j} + \gamma \hat{w}_t + v_t$. In our Monte Carlo study, we consider two RALS tests, RALS(2&3) and RALS(t5), for each of the models with or without time trend as in (19) or (20).

First, the "RALS(2&3)" test imposes the moment conditions that the second and third moments of the errors are not correlated with the lagged dependent variables. Therefore, we let $h(\hat{e}_t) = [\hat{e}_t^2, \hat{e}_t^3]'$. Letting $m_j = T^{-1} \sum_{t=1}^{T} \hat{e}_t^j$, for $j = 2, 3$, for RALS(2&3) we use

$$\hat{w}_t = \left[\hat{e}_t^2 - m_2,\ \hat{e}_t^3 - m_3 - 3m_2\hat{\varepsilon}_t\right]',\ t = 1, 2, .., T. \tag{21}$$

The first term in $\hat{w}_t$ is associated with the moment condition $E\left[\left(e_t^2 - \sigma_e^2\right) y_{t-1}\right] = 0$, which is the condition of no heteroskedasticity. This condition improves the efficiency of the estimator of β when the errors are not symmetric. The second term in $\hat{w}_t$ improves efficiency unless $\mu_4 = 3\sigma^4$, where $\mu_j = E(e_t^j)$. In general, knowledge of higher moments μ_{j+1} are uninformative if $\mu_{j+1} = j\sigma^2\mu_{j-1}$. This is the redundancy condition initially identified by MaCurdy (1982) and Breusch *et al.* (1999). The normal distribution is the only distribution that satisfies the redundancy condition. However, if the distribution of the error term is not normal, the condition is not satisfied. In such cases, one may increase efficiency by augmenting the testing regression with $\hat{w}_t$.

Second, the "RALS(t5)" test imposes the restrictions that arise from the score of the maximum likelihood procedure when the error density is assumed to be a t-distribution with 5 degrees of freedom. It may not be easy, perhaps, to justify using this particular density function for empirical applications. However, this density function is a popular choice for mimicking a fat-tailed distribution in the tests using the M-estimate for which a specific density function is assumed. Thus, RALS(t5) would achieve the efficiency gains when the distribution of the errors has fat-tails. In this case, we have $h(e_t) = (c+1)\, e_t/(c+e_t^2)$, and $D_2 = (c+1)\,(c-e_t^2)\,/(c+e_t^2)^2$ with $c = 5$. Therefore, in this scenario we have

$$\hat{w}_t = \frac{6\hat{e}_t}{5+\hat{e}_t^2} - \frac{1}{T}\sum_{t=1}^{T}\frac{6\hat{e}_t}{5+\hat{e}_t^2} - \hat{e}_t\frac{1}{T}\sum_{t=1}^{T}\frac{6\,(5-\hat{e}_t^2)}{(5+\hat{e}_t^2)^2} \tag{22}$$

There is no compelling reason behind choosing $c = 5$. However, it seems that the tests are quite robust to the selection of different values of c. For example, our simulations that use $c = 3$, which are not reported here to save space, indicate that the empirical size and power of the tests are almost identical to the case when $c = 5$. To examine the size property, we report the rejection ratio for $\alpha = 0.05$ when $\phi = 1$. To examine the power, we use $\phi = 0.9$. We simulated sample cases for $T = 50$ and 100. All results are based on $5,000$ replications.

We compare the RALS(2&3) and RALS(t5) tests with three other tests: (a) DF, the standard Dickey-Fuller test based on OLS; (b) AD, the test studied by Beelder (1996) and Shin and So (1999) based on adaptive estimation; and (c) M5, the test based on the M-estimate assuming that the true density is the student-t density with 5 degrees of freedom, as studied by Lucas (1995). To compare our RALS-based tests with other tests, we replicated four distributions examined by Shin and So (1999): (i) standard normal, (ii) t-distribution with df = 3, (iii) mixture normal: 0.5N(-3,1)+0.5N(3,1), and (iv) chi-square with df = 1.Then, we compare the perfomance of our RALS tests with the AD and M5 tests using the simulation results reported in Shin and So (1999) and Lucas (1995).

Table 1 reports the results for the basic case when the errors are serially uncorrelated. The number of ADF augmentation terms (p) is set to zero. As is seen throughout Table 1, the sizes of the tests based on RALS(2&3) and RALS(t5) are quite close to the nominal 5% size. When the error has a normal distribution, both RALS tests have correct sizes, and the power is close to that of the DF tests. RALS tests generally have a better size property compared with the AD or M5 tests. The power gain over the standard DF test is substantial when the errors are not normal. The overall power of the RALS(2&3) and RALS(t5) tests is fairly comparable to the power of the AD or M5 tests. This is an encouraging result. The performance of the RALS(t5) and M5 tests is similar when the true density is student-t with 3 degrees of freedom, which is a special case where higher moments do not exist but RALS can be still examined via simulations. RALS(t5) is viewed as more powerful when the density is mixture normal. When the true density is a chi-square distribution with one degree of freedom, RALS(2&3) dominates other tests in terms of power. RALS(2&3) explicitly uses the moment condition that is useful when the error is not symmetric. This result may look surprising given that the RALS moment conditions do not include the scores of the log-density, but the AD-based test does not seem to capture the possible efficiency gain from the non-symmetric feature of the error density. In general, RALS(t5) is marginally better than RALS(2&3) when the density is symmetric. However, as we can see for the case when the density is chi-square with one degree of freedom, RALS(2&3) is generally better than RALS (t5) when the error density is skewed. The difference in power is quite substantial in some cases.

In Tables 2 and 3, we compare the performance of the tests when the errors are serially correlated. In doing so, we compare only three tests, ADF, RALS(2&3) and RALS(t5), in two data generation processes

$$AR : \varepsilon_t = 0.5\varepsilon_{t-1} + e_t, \; t = 1, 2, ...,$$

and

$$MA : \varepsilon_t = e_t - 0.5e_{t-1}, \; t = 1, 2, ...$$

We report the size and power of the cases using a fixed ADF augmentation lag at $p = 2$ and $p = 4$ when $T = 50$, and $p = 3$ and $p = 6$ when $T = 100$. We also examine the cases when p is selected by various information criterion. We simulated the Akaike and Schwarz criteria, but report only the results from the Schwarz criterion since the results from the Akaike criterion were similar. We consider the case when the errors are generated from the standard normal, Cauchy, student-t distribution with 2 degrees of freedom, double exponential, chi-square distribution with 4 degrees of freedom, and beta(2,2) distribution. The Cauchy and the t-distribution with 2 degrees of freedom do not satisfy Assumptions 1 and 2. Thus, we do not know the asymptotic distributions of the statistics in this case. However, it will be interesting to see the performance of the tests in this situation. We report the results for the model with a constant term. To save space, we omit the results when a linear time trend is allowed, but the results are similar.

Table 2 presents the size and power properties of the ADF test, the RALS(2&3) test, and the RALS(t5) test when the errors follow an AR(1) model. The overall pattern of the results is similar. The sizes of all three tests are close to the 5% nominal size, even when the errors are generated from a Cauchy or t-distribution with two degrees of freedom. We note that the power difference between the two tests based on OLS and RALS is the greatest when the errors are generated from a Cauchy distribution. Also, as we observed in Table 1, RALS(t5) is more powerful than RALS(2&3) for all of the symmetric distributions. However, RALS(2&3) is, in general, more powerful when the errors are asymmetric. In particular, the power of the RALS(t5)-based test is lower than that of the OLS-based tests when the error is chi-square distributed with 4 degrees of freedom; the power of RALS(2&3) is 63% while the power of RALS(t5) is 22% in the model without time trend for $T = 100$ and $p = 3$.

Table 3 presents the size and power properties of the three tests when the errors follow an MA(1) model. When $T = 50$ and p is determined by the Schwarz criteria or when p is fixed at 2, all of the tests tend to over-reject the null hypothesis. However, when $p = 4$ (for $T = 50$), the size of all three tests is quite close to the 5% nominal size. By comparison, when $T = 100$ the size of all three tests is reasonably close to the 5% nominal size. Thus the overall size of both the RALS(2&3) and the RALS(t5) tests seem as robust as that of the traditional ADF test. With regard to the power of the tests, we observe a pattern similar to what was found in the case of AR(1) errors. Except for the case where the errors follow a normal distribution, the RALS-based tests are substantially more powerful than the OLS-based ADF tests, and the RALS(2&3) test compares favorably with the RALS(t5) test.

We next examine the degree to which our RALS-based unit root tests are robust to various forms of non-linearity. The RALS-based unit root tests exhibit efficiency gains with non-normal errors but these tests are not designed specifically to detect non-linearity. However, it is reasonable to suspect that some forms of non-linearity could be captured in non-normal errors. Thus, we are interested in examining the power of the RALS-based tests against various forms of non-linearity. For this purpose, we follow the work of Choi and Moh (2007) who considered sixteen non-linear model specifications and examined the power of four popular non-linear unit root tests: the KSS test of Kapetanios *et al.* (2003) using an ESTAR model; the sign test of So and Shin (2001); the M-TAR test of Enders and Granger (1998); and the Inf-t test of Park and Shintani (2005). To illustrate the data generating process of these non-linear models, Figure 1 presents the plot of the Gaussian kernel density function for each of these sixteen non-linear models. Using these various non-linear model specifications, we wish to examine the power of our RALS-based tests and compare them with four popular non-linear tests as well as the traditional DF test. In this case, we let $T = 50$ and 100, and we use the 10% significance level in order to compare the performance of RALS tests directly with the results for the KSS, Sign, M-TAR and Inf-t tests, which are reported in Table 2 of Choi and Moh (2008, p. 90, $\rho_1 = 0.9$).[2] Indeed, Table 4 provides encouraging results: our RALS-based tests are fairly comparable to or more powerful than the other nonlinear unit root tests in most (14 out of 16) cases. Only in two cases, the RALS-based tests are less powerful than at least one of these other tests (DGP 4 and 7), but the difference is negligible. Although these non-linear tests generally perform well against specific nonlinear forms which they are designed to detect, and they are less powerful against other forms of non-linearity, the RALS-based tests are fairly robust to various forms of non-linearity. There are only a few cases where the RALS tests are not powerful; these are the cases when structural changes occur in the data (DGP 14 and 15). However, all other tests also suffer from a loss of power in these cases. This result

[2] Our own simulation results for these tests are similar to their results except for the differences caused by negligible sampling errors and we cite their results in Table 4.

simply confirms the initial finding of Perron (1989) who noted that unit root tests will lose power if existing breaks are ignored. Thus, They are the cases where the models are clearly mis-specified, and all tests will be subject to size distortions and/or loss of power.

Overall, our simulation results show that the RALS-based unit root tests remain relatively powerful under various forms of non-normal errors and non-linear alternatives throughout.

5 An Application of the RALS Unit Root Test

We now present an empirical application of our new test by applying the RALS-based unit root test "RALS(2&3)" to the CPI inflation rate series of several member-countries of the Organization for Economic Co-operation and Development (OECD). Knowledge of the long-run properties of the inflation rate (or the aggregate price level) is a key component for policy makers, applied econometricians and financial analysts who seek to understand or affect the behavior of the macroeconomy. For example, forecasters who seek to project expected or future inflation rates must know whether or not inflation rates are stationary when building their models. Similarly, officials who seek to use monetary policy to affect the behavior of macroeconomic variables also must have knowledge of the long-run properties of inflation when constructing optimal commodity price rules or when engaging in inflation rate targeting. In addition, financial planners who, for example, rely on the capital asset pricing model also must understand the long-run behavior of inflation.

Yet the question of whether or not the inflation rate is stationary still is widely disputed in the literature. Numerous researchers, employing various methodologies applied to the inflation rates of several different countries, have found this series to be non-stationary (see, for example, Crowder and Hoffman (1996), Rapach and Weber (2004), Crowder and Phengpis (2007)). At the same time, several authors have concluded that inflation is stationary (see, for example, Baillie, Chung and Tieslau (1996), and Costantini and Lupi (2007)). This contradiction in the empirical results on the inflation rate might be due, in part, to the low power of traditional unit root tests. We wish to examine whether or not accounting for non-normality in the series will make a difference. Since our test will be more powerful in the face of departures from normality or apparent non-linearities, we seek to shed light on the issue of whether or not inflation is stationary through the application of our more powerful tests.

The series used in our analysis are the first-differences of the log of the monthly consumer price index series (all items) for 12 OECD countries.[3] The data were taken from the International Monetary Fund's "International Financial Statistics" CD rom (July 2009), and span the period from January of 1957 through April of 2009. We analyze these inflation rates applying the RALS(2&3) test to both (19) and (20). The first step of the procedure begins by conducting the traditional Dickey-Fuller unit root test while choosing the optimal number of augmentation terms to ensure non-correlated errors in the testing equation.[4] The OLS residuals from this equation are then retained for use in the second step. The second step involves estimation of the RALS unit root testing equation, which is an augmented version of the original Dickey-Fuller equation.

The results of the RALS unit root test are presented in Table 5. In the case where the testing equation includes only an intercept, the RALS unit root test rejects the null of a unit root in 8 out of 12 cases, while the Dickey-Fuller unit root test rejects the null in only 3 of 12 cases. Similarly, when allowing for both a constant and a trend in the testing equation, the RALS unit root test rejects the null in 7 of 12 cases, while the Dickey-Fuller test rejects the null in only 2 cases. The ability of the RALS unit root test to reject the null in significantly more cases may lend support to the notion that our test is better able to distinguish non-normality from non-stationarity.

[3] The countries are: Belgium, Canada, Finland, France, Italy, Japan, Luxembourg, the Netherlands, Norway, Spain, the UK and the USA.

[4] One may choose the optimal lag length following the usual practice. For example, one can determine the optimal number of augmentation terms using the sequential t-test, following Ng and Perron (1995), or through use of the traditional Akaike Information Criteria or Schwarz Criteria, or other similar methods. In our application, we followed the procedure of Ng and Perron (1995) with a maximum of 12 lags.

6 Concluding Remarks

This paper proposes new unit root tests that are more powerful when the error term follows a non-normal distribution. The improved power is gained by utilizing more moment conditions through a computationally simple procedure. Specifically, we extend the residual augmented least squares (RALS) estimator proposed by Im and Schmidt (2008) in order to use the information implied by non-normal errors when testing for a unit root. We show that the asymptotic distribution of our simple RALS-based estimator is the same as that of the GMM estimator. Our Monte Carlo simulation results show that the size of the RALS-based unit root tests is quite close to the asymptotic size, and the power is improved significantly over the usual Dickey-Fuller tests when the error is not normal. As such, our findings show significant efficiency gains, although this information is ignored in traditional unit root tests. In addition, it is encouraging to see that the RALS-based unit root tests remain powerful under most cases of non-linear alternatives.

A Appendix

Lemma A1. We let $z_t = (y_{t-1}, \xi_t')'$, as defined previously in equation (5). We define a $(p+1) \times (p+1)$ matrix, $\Upsilon_T = diag\left(T, \sqrt{T}, ..., \sqrt{T}\right)$. Assume that Assumptions 1 and 2 hold. Then, we have under the null hypothesis

$$\sum_{t=1}^{T} \left[g'(e_t) \otimes \Upsilon_T^{-1} z_t z_t' \Upsilon_T^{-1}\right] \Rightarrow D \otimes \int zz', \tag{A.1}$$

$$\sum_{t=1}^{T} g(e_t) g(e_t)' \otimes \Upsilon_T^{-1} z_t z_t' \Upsilon_T^{-1} \Rightarrow C \otimes \int zz', \tag{A.2}$$

where $\int zz' = diag\left(a(1)^{-2}\sigma_e^2 \int_0^1 W_1(r)^2 dr, E(\xi_t \xi_t')\right)$, and C and D are defined in (17). Also, we have

$$\sum_{t=1}^{T} \psi(e_t) \Upsilon_T^{-1} z_t = \begin{bmatrix} T^{-1} \sum_{t=1}^{T} \psi(\varepsilon_t) y_{t-1} \\ T^{-1/2} \sum_{t=1}^{T} \psi(\varepsilon_t) \xi_t \end{bmatrix} \Rightarrow \begin{bmatrix} \frac{\sigma_\psi \sigma_\varepsilon}{a(1)} \int_0^1 W_1(r) dW_2(r) \\ \Gamma \end{bmatrix}, \tag{A.3}$$

where Γ is a $p \times p$ multivariate normal variable with covariance matrix $\sigma_\psi^2 E(\xi_t \xi_t')$.

proof. Lucas (1995, Lemma 1 in Appendix). See also Hansen (1995, Lemma).

Lemma A2. ρ is defined as in equation (6). Then,

$$\rho = \frac{1}{\sigma_e \sigma_\psi}. \tag{A.4}$$

Also,

$$\frac{1}{\sigma_\psi^2} = \sigma_e^2 - \left(C_{21} - \sigma_e^2 D_2\right)' \left(C_{22} + \sigma_e^2 D_2 D_2' - C_{21} D_2' - D_2 C_{21}'\right)^{-1} \left(C_{21} - \sigma_e^2 D_2\right). \tag{A.5}$$

proof. The first result follows from routine matrix algebra using the partitioned inverse lemma. For the second result, straightforward algebra gives

$$\left(D'C^{-1}D\right)^{-1} = \sigma_e^2 \left(1 + \left(C_{21} - \sigma_e^2 D_2\right)' \left(\sigma_e^2 C_{22} - C_{21} C_{21}'\right)^{-1} \left(C_{21} - \sigma_e^2 D_2\right)\right)^{-1},$$

which is the same as $1/\sigma_\psi^2$; see Amemiya (1985, p. 461, Lemma 20).

PROOF OF THEOREM 1: We note that the entire proof follows immediately from Lucas (1995, Theorem 1) since the GMM estimator is obtained by solving the score $\sum_{t=1}^{T} \left[DC^{-1} g(e_t) z_t\right] = \sum_{t=1}^{T} \left[\psi(e_t) z_t\right] = 0$, and this score could be viewed as that of the M-estimate. Here, we provide more details. Let $\theta = (\beta, \delta_1, \delta_2, ..., \delta_p)'$. The GMM estimator is obtained by solving

$$\min_\theta \sum_{t=1}^{T} \left[g(e_t) \otimes z_t\right]' \hat{\Lambda}^{-1} \sum_{t=1}^{T} \left[g(e_t) \otimes z_t\right], \tag{A.6}$$

where $\hat{\Lambda} = \left(\sum_{t=1}^{T} g(\hat{e}_t) g(\hat{e}_t)' \otimes z_t z_t'\right)$, and $\hat{e}_t$ is the residual from an initial consistent estimator of θ. Taking the derivative with respect to θ, we obtain the score

$$\sum_{t=1}^{T} \left[g'(\tilde{e}_t) \otimes z_t z_t'\right]' \hat{\Lambda}^{-1} \sum_{t=1}^{T} \left[g(\tilde{e}_t) \otimes z_t\right] = 0, \tag{A.7}$$

where $\tilde{e}_t = \Delta y_t - z_t \tilde{\theta}$, and $\tilde{\theta}$ is the GMM estimator. The Taylor series expansion of the term $\sum_{t=1}^{T} \left[g(\tilde{e}_t) \otimes z_t\right]$ with respect to the true disturbance e_t and premultiplication of $I_J \otimes \Upsilon_T^{-1}$ yields

$$\begin{aligned} & \sum_{t=1}^{T} \left[g(\tilde{e}_t) \otimes \Upsilon_T^{-1} z_t\right] \qquad\qquad (A.8) \\ = \; & \sum_{t=1}^{T} \left[g(e_t) \otimes \Upsilon_T^{-1} z_t - g'(e_t) \otimes \Upsilon_T^{-1} z_t z_t' \Upsilon_T^{-1} \Upsilon_T \left(\tilde{\theta} - \theta\right)\right] + o_p(1). \end{aligned}$$

Solving (A.7) with respect to $\Upsilon_T\left(\tilde{\theta}-\theta\right)$, after substituting (A.8) into (A.7), we obtain

$$\Upsilon_T(\tilde{\theta}-\theta) = \tag{A.9}$$
$$\left\{\sum_{t=1}^{T}\left[g'(\tilde{e}_t)\otimes\Upsilon_T^{-1}z_tz_t'\Upsilon_T^{-1}\right]'\left[\sum_{t=1}^{T}g(\hat{e}_t)g(\hat{e}_t)'\otimes\Upsilon_T^{-1}z_tz_t'\Upsilon_T^{-1}\right]^{-1}\sum_{t=1}^{T}\left[g'(e_t)\otimes\Upsilon_T^{-1}z_tz_t'\Upsilon_T^{-1}\right]\right\}^{-1}$$
$$\times\left\{\sum_{t=1}^{T}\left[g'(\tilde{e}_t)\otimes\Upsilon_T^{-1}z_tz_t'\Upsilon_T^{-1}\right]'\left[\sum_{t=1}^{T}g(\hat{e}_t)g(\hat{e}_t)'\otimes\Upsilon_T^{-1}z_tz_t'\Upsilon_T^{-1}\right]^{-1}\sum_{t=1}^{T}\left[g(e_t)\otimes\Upsilon_T^{-1}z_t\right]\right\}+o_p(1).$$

Noting that

$$\sum_{t=1}^{T}\left\{\left[g'(\tilde{e}_t)-g'(e_t)\right]\otimes\Upsilon_T^{-1}z_tz_t'\Upsilon_T^{-1}\right\}=o_p\left(1\right),$$

and

$$\sum_{t=1}^{T}\left\{\left[g(\hat{e}_t)g(\hat{e}_t)'-g(e_t)g(e_t)'\right]\otimes\Upsilon_T^{-1}z_tz_t'\Upsilon_T^{-1}\right\}=o_p\left(1\right),$$

we have, from Lemma A1

$$T\tilde{\beta}_G\Rightarrow\frac{a(1)}{\sigma_\psi\sigma_e}\left(\int_0^1W_1(r)^2dr\right)^{-1}\int_0^1W_1(r)dW_2(r), \tag{A.10}$$

where $[W_1(r),W_2(r)]$ is a bivariate Brownian motion with correlation ρ. Then, we have for the t-statistic

$$t_G\Rightarrow\left(\int_0^1W_1(r)^2dr\right)^{-1/2}\int_0^1W_1(r)dW_2(r), \tag{A.11}$$

which is a mixture of the Dickey-Fuller and the standard normal distribution as described in (8). To see this, note

$$T^{-1/2}\sum_{t=1}^{[rT]}\begin{bmatrix}e_t\\ \psi(e_t)\end{bmatrix}\Rightarrow\begin{bmatrix}\sigma_eW_1\left(r\right)\\ \sigma_\psi W_2\left(r\right)\end{bmatrix}, \tag{A.12}$$

where $[rT]$ denotes the integer part of rT. Therefore,

$$W_2\left(r\right)=\rho W_1\left(r\right)+\sqrt{1-\rho^2}W_3\left(r\right), \tag{A.13}$$

where $W_3\left(r\right)$ is independent of $W_1(r)$. The result follows if we note that

$$\left(\int_0^1W_1(r)^2d(r)\right)^{-1/2}\int_0^1W_1(r)dW_3(r)$$

is standard normal.

PROOF OF THEOREM 2: Define a variable as a function of true disturbances

$$w_t = h(e_t) - K - e_t D_2,\ \ t = 1, 2, .., T.$$

The variables in w_t are not observable, but we momentarily assume that they are observed. Then we show that the augmentation of w_t or $\hat{w}_t$ asymptotically yields the same estimator of $T\beta$. Consider a regression

$$\Delta y_t = \alpha_1 + \beta y_{t-1} + \sum_{j=1}^{p} \delta_j \Delta y_{t-j} + w_t'\gamma + v_t,\ \ t = 1, 2, .., T. \tag{A.14}$$

Therefore,

$$e_t = w_t'\gamma + v_t,\ \ t = 1, 2, .., T. \tag{A.15}$$

Let $\hat{\beta}_A^*$ be the least squares estimator of β from regression (A.14), $\sigma_v^2 = Var(v_t)$, and

$$\lambda = \frac{\sigma_{ev}}{\sigma_e \sigma_v} = \frac{\sigma_v}{\sigma_e}, \tag{A.16}$$

where $\sigma_{ev} = E(\varepsilon_t v_t)$. The second equality of (A.16) follows since w_t and v_t are not correlated, so that $\sigma_{ev} = \sigma_v^2$. From Hansen (1995, Theorem 2 and 3), we have

$$T\hat{\beta}_A^* \Rightarrow \frac{\sigma_v}{\sigma_e} \left(\int_0^1 W_4(r)^2 \right)^{-1} \int_0^1 W_4(r) dW_5(r), \tag{A.17}$$

and for the t-statistic

$$t_A^* = \lambda DF_\mu + \sqrt{1 - \lambda^2} N(0, 1), \tag{A.18}$$

where $[W_4(r), W_5(r)]'$ is the bivariate Brownian motion with correlation λ. Next, we will show that

$$\rho = \lambda. \tag{A.19}$$

Note that $\gamma = E(w_t w_t')^{-1} E(w_t e_t)$, so we have

$$\sigma_v^2 = \sigma_e^2 - E(e_t w_t') E(w_t w_t')^{-1} E(w_t e_t). \tag{A.20}$$

Also, $E(w_t e_t) = C_{21} - \sigma_e^2 D_2$ and $E(w_t w_t') = C_{22} + \sigma_\varepsilon^2 D_2 D_2' - C_{21} D_2' - D_2 C_{21}'$. Therefore,

$$\sigma_v^2 = \sigma_e^2 - \left(C_{21} - \sigma_\varepsilon^2 D_2\right)' \left(C_{22} + \sigma_\varepsilon^2 D_2 D_2' - C_{21} D_2' - D_2 C_{21}'\right)^{-1} \left(C_{21} - \sigma_\varepsilon^2 D_2\right),$$

which becomes $1/\sigma_\psi^2$ from Lemma A1. Therefore, we have $\rho = \lambda$.

Now we let $\hat{\beta}_A$ be the OLS estimator of β in the regression (19). The proof is complete if we show that $T\hat{\beta}_A$ and $T\hat{\beta}_A^*$ are identical asymptotically. Let $\hat{\zeta}_t = \left(\tilde{\xi}_t', \hat{w}_t'\right)'$, where $\tilde{\xi}_t = \xi_t - T^{-1}\sum_{t=1}^T \xi_t$. Then we have

$$T\hat{\beta}_A = \frac{T^{-1}\left(\sum_{t=1}^T \tilde{y}_{t-1}e_t - \sum_{t=1}^T \tilde{y}_{t-1}\zeta_t'\left(\sum_{t=1}^T \tilde{\zeta}_t\tilde{\zeta}_t'\right)^{-1}\sum_{t=1}^T \tilde{\zeta}_t e_t\right)}{T^{-2}\left(\sum_{t=1}^T \tilde{y}_{t-1}^2 - \sum_{t=1}^T \tilde{y}_{t-1}\zeta_t'\left(\sum_{t=1}^T \tilde{\zeta}_t\tilde{\zeta}_t'\right)^{-1}\sum_{t=1}^T \zeta_t\tilde{y}_{t-1}\right)},$$

Since $T^{-1}\sum_{t=1}^T \hat{w}_t\xi_t' = o_p(1)$, and $T^{-1}\sum_{t=1}^T \tilde{\xi}_t e_t = o_p(1)$, we have

$$T\hat{\beta}_A = \frac{T^{-1}\left(\sum_{t=1}^T \tilde{y}_{t-1}e_t - \sum_{t=1}^T \tilde{y}_{t-1}\hat{w}_t'\left(\sum_{t=1}^T \hat{w}_t\hat{w}_t'\right)^{-1}\sum_{t=1}^T \hat{w}_t' e_t\right)}{T^{-2}\left(\sum_{t=1}^T \tilde{y}_{t-1}^2\right)} + o_p(1).$$

Similarly,

$$T\hat{\beta}_A^* = \frac{T^{-1}\left(\sum_{t=1}^T \tilde{y}_{t-1}e_t - \sum_{t=1}^T \tilde{y}_{t-1}w_t'\left(\sum_{t=1}^T \tilde{w}_t\tilde{w}_t'\right)^{-1}\sum_{t=1}^T \tilde{w}_t' e_t\right)}{T^{-2}\left(\sum_{t=1}^T \tilde{y}_{t-1}^2\right)} + o_p(1).$$

$T\hat{\beta}_A$ and $T\hat{\beta}_A^*$ are asymptotically identical if $T^{-1}\sum \tilde{y}_{t-1}\left(\hat{w}_t - w_t\right) = o_p\left(1\right)$. However,

$$T^{-1}\sum \tilde{y}_{t-1}\hat{w}_t = T^{-1}\sum \tilde{y}_{t-1}\left[h(\varepsilon_t) + (\hat{\varepsilon}_t - \varepsilon_t)h'\left(\varepsilon_t\right) - \hat{\varepsilon}_t\hat{D}_2\right] + o_p(1).$$

Therefore,

$$\begin{aligned} & T^{-1}\sum \tilde{y}_{t-1}\left(\hat{w}_t - w_t\right) \\ = \; & T^{-1}\sum \tilde{y}_{t-1}\left[(\hat{\varepsilon}_t - \varepsilon_t)h'\left(\varepsilon_t\right) - (\hat{\varepsilon}_t - \varepsilon_t)\hat{D}_2 - \varepsilon_t\left(\hat{D}_2 - D_2\right)\right] + o_p(1) \end{aligned} \tag{A.21}$$

but,

$$T^{-1}\sum \tilde{y}_{t-1}(\hat{\varepsilon}_t - \varepsilon_t)h'\left(\varepsilon_t\right) = T\left(\hat{\beta} - \beta\right)T^{-2}\sum \tilde{y}_{t-1}^2 h'(\varepsilon_t) + o_p(1), \tag{A.22}$$

$$T^{-1}\sum \tilde{y}_{t-1}(\hat{\varepsilon}_t - \varepsilon_t)\hat{D}_2 = \hat{D}_2 T\left(\hat{\beta} - \beta\right)T^{-2}\sum \tilde{y}_{t-1}^2 + o_p(1), \tag{A.23}$$

and

$$T^{-1}\sum \tilde{y}_{t-1}\varepsilon_t\left(\hat{D}_2 - D_2\right) = o_p(1). \tag{A.24}$$

The two terms (A.22) and (A.23) cancel each other in the limit in (A.21), so the proof is complete.

References

[1] Amemiya, T. (1985), Advanced Econometrics, Basil Blackwell Ltd.

[2] Baillie, R.T., C.F. Chung, and M.A. Tieslau (1996), "Analyzing inflation by the fractionally integrated ARFIMA-GARCH model," *Journal of Applied Econometrics*, 11, 23–40.

[3] Beelder, O. (1996), "Adaptive estimation and unit root tests," manuscript, Department of Economics, Rochester University.

[4] Breusch, T., H. Qian, P. Schmidt, and D.J. Wyhowski (1999), "Redundancy of moment conditions," *Journal of Econometrics,* 91, 89–112.

[5] Choi, C.Y. and Y. Moh (2007), "How useful are tests for unit-root in distinguishing unit-root processes from stationary but non-linear processes?" Econometrics Journal, 10, 82–112.

[6] Costantini, M. and C. Lupi (2007), "An analysis of inflation and interest rates. New panel unit root results in the presence of structural breaks," *Economics Letters,* 95, 408–414.

[7] Cox, D.D. and I. Llatas (1991), "Maximum likelihood type estimation for nearly non-stationary autoregressive time series," *Annals of Statistics*, 19, 1109–1128.

[8] Crowder, W.J. and D.L. Hoffman (1996), "The long-run relationship between nominal interest rates and inflation: the Fisher equation revisited," *Journal of Money, Credit and Banking*, 28, 102–118.

[9] Crowder, W.J. and C. Phengpis (2007), "A re-examination of international inflation convergence over the modern float," *Journal of International Financial Markets, Institutions and Money*, 17, 125–139.

[10] Elliott, G and M. Jansson (2003), "Testing for unit roots with stationary covariates," *Journal of Econometrics*, 115, 75-89.

[11] Enders, W. and C.W.J. Granger (1998), "Unit-root tests and asymmetric adjustment with an example using the term structure of interest rates," *Journal of Business and Economic Statistics,* 16, 304–311.

[12] Hansen, B.E. (1995), "Rethinking the univariate approach to the unit root testing: using covariates to increase the power," *Econometric Theory*, 11, 1148–1171.

[13] Im, K.S. and P. Schmidt (2008), "More efficient estimation under non-normality when higher moments do not depend on the regressors, using residual-augmented least squares," *Journal of Econometrics*, 144, 219–233.

[14] Kapetanios, G., Y. Shin and A. Snell (2003), "Testing for a unit root in the non-linear STAR framework," *Journal of Econometrics* 112, 359–79.

[15] Lucas, A. (1995), "Unit root tests based on M-estimators," *Econometric Theory*, 11, 331–346.

[16] MaCurdy, T.E., (1982), "Using information on the moments of disturbances to increase the efficiency of estimation," NBER Technical Paper 22, Cambridge, MA.

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

[18] Park, J. and M. Shintani (2005). "Testing for a unit root against transitional autoregressive models," mimeo, Vanderbilt University.

[19] Perron, P. (1989), "The great crash, the oil price shock and the unit root hypothesis," *Econometrica*, 57, 1361–1401.

[20] Qian, H. and P. Schmidt (1999), "Improved instrumental variables and generalized method of moments estimators," *Journal of Econometrics*, 91, 145–170.

[21] Rapach, D.E. and C.E. Weber (2004), "Are real interest rates really nonstationary? New evidence from tests with good size and power" *Journal of Macroeconomics*, 26 409–430.

[22] Shin, D.W. and B.S. So (1999), "Unit Root Tests Based on Adaptive Maximum Likelihood Estimation," *Econometric Theory*, 15, 1–23.

[23] So, B.S. and D.W. Shin (2001), "An invariant sign test for random walks based on recursive median adjustment," *Journal of Econometrics*, 102, 197–229.

[24] Wooldridge, J.M. (1993), "Efficient estimation with orthogonal regressors," *Econometric Theory*, 9, 687.

Table 1
Rejection Ratio of Various Tests
No Serial Correlations, 5% significance level

No Time Trend

Distributions		DF (T=50)	RALS (2&3) (T=50)	RALS (t5) (T=50)	AD (T=50)	M5 (T=50)	DF (T=100)	RALS (2&3) (T=100)	RALS (t5) (T=100)	AD (T=100)	M5 (T=100)
Normal	$\phi = 1$	0.060	0.054	0.059	0.043	0.094	0.051	0.051	0.051	0.049	0.069
	$\phi = 0.9$	0.146	0.132	0.136	0.091	0.198	0.352	0.311	0.330	0.263	0.346
Student t	$\phi = 1$	0.058	0.051	0.050	0.045	0.052	0.053	0.051	0.051	0.067	0.037
df=3	$\phi = 0.9$	0.139	0.270	0.296	0.197	0.291	0.358	0.615	0.676	0.535	0.649
Mixture	$\phi = 1$	0.055	0.044	0.045	0.040	0.178	0.058	0.045	0.044	0.049	0.130
Normal	$\phi = 0.9$	0.145	0.850	0.916	0.790	0.217	0.361	0.995	0.998	0.991	0.281
Chi-square	$\phi = 1$	0.046	0.043	0.045	0.048	0.058	0.052	0.041	0.045	0.047	0.036
df=1	$\phi = 0.9$	0.126	0.909	0.339	0.360	0.332	0.355	0.999	0.723	0.796	0.666

With Linear Trend

Distributions		DF (T=50)	RALS (2&3) (T=50)	RALS (t5) (T=50)	AD (T=50)	M5 (T=50)	DF (T=100)	RALS (2&3) (T=100)	RALS (t5) (T=100)	AD (T=100)	M5 (T=100)
Normal	$\phi = 1$	0.062	0.054	0.057	0.025	0.148	0.057	0.054	0.055	0.035	0.078
	$\phi = 0.9$	0.109	0.092	0.099	0.049	0.204	0.216	0.191	0.206	0.129	0.251
Student t	$\phi = 1$	0.064	0.060	0.052	0.026	0.062	0.054	0.049	0.049	0.039	0.036
df=3	$\phi = 0.9$	0.100	0.187	0.192	0.120	0.231	0.197	0.441	0.507	0.386	0.495
Mixture	$\phi = 1$	0.054	0.042	0.044	0.024	0.292	0.053	0.042	0.043	0.027	0.192
Normal	$\phi = 0.9$	0.097	0.669	0.784	0.628	0.258	0.219	0.981	0.996	0.981	0.255
Chi-square	$\phi = 1$	0.055	0.045	0.051	0.026	0.064	0.055	0.039	0.049	0.038	0.048
df=1	$\phi = 0.9$	0.081	0.797	0.224	0.251	0.277	0.202	0.991	0.529	0.647	0.506

The 5% significance level was used. AD denotes the test based on the adaptive MLE of Shin and So (1999) and M5 is the test of Lucas (1995) using the M-estimate assuming that the error density is the student-*t* with 5 degrees of freedom. Mixture normal is 0.5N(-3,1) + 0.5N(3,1).

Table 2
Rejection Ratio of Various Tests
AR(1) error with AR coefficient 0.5 (No Time Trend)

T = 50

Distributions		ADF			RALS(2&3)			RALS(t5)		
		p=2	p=4	SC	p=2	p=4	SC	p=2	p=4	SC
Normal	$\phi = 1$	0.056	0.055	0.071	0.053	0.054	0.061	0.054	0.053	0.067
	$\phi = 0.9$	0.100	0.080	0.116	0.087	0.074	0.097	0.093	0.073	0.110
Cauchy	$\phi = 1$	0.075	0.079	0.055	0.046	0.053	0.063	0.048	0.059	0.051
	$\phi = 0.9$	0.074	0.074	0.088	0.664	0.599	0.698	0.567	0.504	0.568
Student t	$\phi = 1$	0.063	0.060	0.056	0.050	0.054	0.064	0.049	0.056	0.055
df=2	$\phi = 0.9$	0.080	0.069	0.089	0.300	0.252	0.326	0.343	0.281	0.341
Double	$\phi = 1$	0.051	0.053	0.059	0.048	0.050	0.057	0.049	0.051	0.058
Exponential	$\phi = 0.9$	0.091	0.084	0.109	0.131	0.110	0.150	0.151	0.120	0.161
Chi-square	$\phi = 1$	0.051	0.058	0.061	0.050	0.045	0.032	0.052	0.051	0.061
4 df	$\phi = 0.9$	0.094	0.080	0.110	0.260	0.202	0.191	0.091	0.081	0.106
Beta(2,2)	$\phi = 1$	0.060	0.055	0.073	0.057	0.050	0.059	0.053	0.047	0.060
	$\phi = 0.9$	0.100	0.087	0.121	0.126	0.101	0.133	0.131	0.103	0.149

T = 100

Distributions		ADF			RALS(2&3)			RALS(t5)		
		p=3	p=6	SC	p=3	p=6	SC	p=3	p=6	SC
Normal	$\phi = 1$	0.055	0.053	0.061	0.056	0.048	0.052	0.055	0.051	0.056
	$\phi = 0.9$	0.217	0.163	0.243	0.196	0.142	0.217	0.207	0.145	0.230
Cauchy	$\phi = 1$	0.080	0.076	0.055	0.040	0.042	0.055	0.045	0.050	0.042
	$\phi = 0.9$	0.144	0.125	0.181	0.907	0.852	0.943	0.796	0.775	0.803
Student t	$\phi = 1$	0.053	0.053	0.050	0.049	0.052	0.067	0.050	0.047	0.049
df=2	$\phi = 0.9$	0.190	0.134	0.220	0.610	0.512	0.678	0.716	0.616	0.740
Double	$\phi = 1$	0.059	0.053	0.062	0.055	0.050	0.063	0.055	0.047	0.056
Exponential	$\phi = 0.9$	0.216	0.155	0.246	0.321	0.231	0.362	0.377	0.273	0.400
Chi-square	$\phi = 1$	0.052	0.048	0.052	0.046	0.048	0.025	0.050	0.045	0.047
df=4	$\phi = 0.9$	0.224	0.156	0.242	0.629	0.480	0.556	0.217	0.157	0.237
Beta(2,2)	$\phi = 1$	0.057	0.053	0.063	0.048	0.048	0.052	0.050	0.046	0.054
	$\phi = 0.9$	0.216	0.155	0.246	0.324	0.225	0.355	0.343	0.235	0.376

Table 3
Rejection Ratio of Various Tests
MA(1) error with MA coefficient -0.5 (No Time Trend)

T = 50

Distributions		ADF			RALS(2&3)			RALS(t5)		
		p=2	p=4	SC	p=2	p=4	SC	p=2	p=4	SC
Normal	$\phi = 1$	0.088	0.054	0.095	0.079	0.051	0.079	0.085	0.052	0.091
	$\phi = 0.9$	0.227	0.109	0.231	0.190	0.091	0.182	0.205	0.097	0.208
Cauchy	$\phi = 1$	0.102	0.077	0.079	0.180	0.092	0.194	0.109	0.068	0.108
	$\phi = 0.9$	0.163	0.087	0.177	0.853	0.712	0.867	0.625	0.557	0.622
Student t	$\phi = 1$	0.091	0.060	0.079	0.109	0.060	0.123	0.102	0.059	0.101
df=2	$\phi = 0.9$	0.198	0.089	0.199	0.538	0.347	0.547	0.553	0.379	0.539
Double	$\phi = 1$	0.083	0.053	0.085	0.088	0.050	0.094	0.087	0.054	0.090
Exponential	$\phi = 0.9$	0.219	0.104	0.216	0.285	0.149	0.289	0.320	0.168	0.313
Chi-square	$\phi = 1$	0.085	0.050	0.088	0.108	0.054	0.070	0.079	0.051	0.082
df=4	$\phi = 0.9$	0.223	0.102	0.227	0.522	0.282	0.425	0.209	0.098	0.212
Beta(2,2)	$\phi = 1$	0.095	0.054	0.101	0.086	0.051	0.084	0.087	0.050	0.090
	$\phi = 0.9$	0.237	0.112	0.241	0.259	0.135	0.250	0.276	0.140	0.277

T = 100

Distributions		ADF			RALS(2&3)			RALS(t5)		
		p=3	p=6	SC	p=3	p=6	SC	p=3	p=6	SC
Normal	$\phi = 1$	0.050	0.049	0.057	0.052	0.046	0.053	0.050	0.046	0.056
	$\phi = 0.9$	0.260	0.187	0.287	0.230	0.172	0.257	0.246	0.176	0.269
Cauchy	$\phi = 1$	0.078	0.073	0.049	0.041	0.044	0.057	0.040	0.042	0.036
	$\phi = 0.9$	0.171	0.134	0.207	0.938	0.889	0.962	0.785	0.776	0.796
Student t	$\phi = 1$	0.053	0.048	0.048	0.050	0.047	0.065	0.048	0.046	0.050
df=2	$\phi = 0.9$	0.235	0.159	0.266	0.682	0.570	0.745	0.771	0.666	0.798
Double	$\phi = 1$	0.058	0.051	0.059	0.052	0.050	0.058	0.053	0.048	0.054
Exponential	$\phi = 0.9$	0.263	0.187	0.303	0.379	0.280	0.444	0.440	0.321	0.486
Chi-square	$\phi = 1$	0.047	0.044	0.050	0.046	0.045	0.027	0.041	0.043	0.045
df=4	$\phi = 0.9$	0.254	0.183	0.293	0.705	0.543	0.650	0.248	0.174	0.286
Beta(2,2)	$\phi = 1$	0.051	0.047	0.059	0.047	0.046	0.053	0.049	0.045	0.052
	$\phi = 0.9$	0.253	0.186	0.296	0.372	0.265	0.407	0.391	0.280	0.438

Table 4
Rejection Ratio of Various Tests
Under Various Nonlinear Models (No Time Trend)

DGP	Model	RALS (2&3)	RALS (t5)	DF	*KSS*	*Sign*	*M-TAR*	*Inf-t*
					$T = 50$			
1	AR(1)	0.411	0.376	0.29	*0.19*	*0.19*	*0.08*	*0.09*
2	Generalized AR(1)	0.411	0.376	0.29	*0.19*	*0.19*	*0.08*	*0.19*
3	Bilinear	0.425	0.388	0.29	*0.22*	*0.17*	*0.10*	*0.25*
4	Nonlinear AR	0.897	0.903	0.95	*0.59*	*0.99*	*1.00*	*1.00*
5	Squared Relation	0.822	0.695	0.69	*0.44*	*0.59*	*0.50*	*0.66*
6	Exponential Relation	0.853	0.795	0.76	*0.76*	*0.76*	*0.67*	*0.67*
7	Bilinear	0.897	0.901	0.95	*0.59*	*0.99*	*1.00*	*1.00*
8	SETAR(1)	0.372	0.328	0.22	*0.15*	*0.16*	*0.06*	*0.14*
9	EQ-TAR	0.359	0.316	0.22	*0.16*	*0.17*	*0.06*	*0.16*
10	Band TAR	0.319	0.262	0.19	*0.12*	*0.12*	*0.04*	*0.13*
11	ESTAR	0.332	0.293	0.19	*0.13*	*0.14*	*0.05*	*0.13*
12	LSTAR	0.577	0.551	0.50	*0.26*	*0.35*	*0.25*	*0.37*
13	Markov-switching	0.505	0.510	0.46	*0.30*	*0.41*	*0.24*	*0.37*
14	Level Shift	0.144	0.114	0.08	*0.08*	*0.08*	*0.02*	*0.06*
15	Multiple Shift	0.008	0.004	0.00	*0.00*	*0.05*	*0.01*	*0.00*
16	Variance Shift	0.470	0.395	0.30	*0.28*	*0.23*	*0.09*	*0.30*
					$T = 100$			
1	AR(1)	0.590	0.569	0.52	*0.37*	*0.47*	*0.21*	*0.42*
2	Generalized AR(1)	0.590	0.569	0.52	*0.37*	*0.46*	*0.21*	*0.41*
3	Bilinear	0.577	0.547	0.56	*0.40*	*0.39*	*0.20*	*0.54*
4	Nonlinear AR	0.968	0.973	0.99	*0.77*	*1.00*	*1.00*	*1.00*
5	Squared Relation	0.949	0.863	0.86	*0.60*	*0.88*	*0.81*	*0.89*
6	Exponential Relation	0.970	0.920	0.76	*0.76*	*0.82*	*0.84*	*0.84*
7	Bilinear	0.968	0.973	0.99	*0.77*	*1.00*	*1.00*	*1.00*
8	SETAR(1)	0.442	0.423	0.36	*0.28*	*0.36*	*0.10*	*0.28*
9	EQ-TAR	0.409	0.398	0.33	*0.34*	*0.33*	*0.10*	*0.32*
10	Band TAR	0.307	0.284	0.21	*0.18*	*0.24*	*0.05*	*0.17*
11	ESTAR	0.343	0.313	0.25	*0.21*	*0.29*	*0.07*	*0.20*
12	LSTAR	0.803	0.791	0.84	*0.46*	*0.69*	*0.66*	*0.75*
13	Markov-switching	0.727	0.775	0.77	*0.49*	*0.75*	*0.57*	*0.70*
14	Level Shift	0.121	0.097	0.08	*0.12*	*0.19*	*0.03*	*0.09*
15	Multiple Shift	0.000	0.000	0.00	*0.00*	*0.24*	*0.00*	*0.00*
16	Variance Shift	0.587	0.536	0.47	*0.41*	*0.48*	*0.25*	*0.54*

The 10% significance level was used. DF denotes the standard Dickey-Fuller test. KSS denotes the test of Kapetanios *et al.* (2003) based on the ESTAR model. Sign denotes the sign test of So and Shin (2001). M-TAR denotes the test of Enders and Granger (1998) and Inf-t denotes the test of Park and Shintani (2005).

Table 5
Empirical Application of Inflation Rates

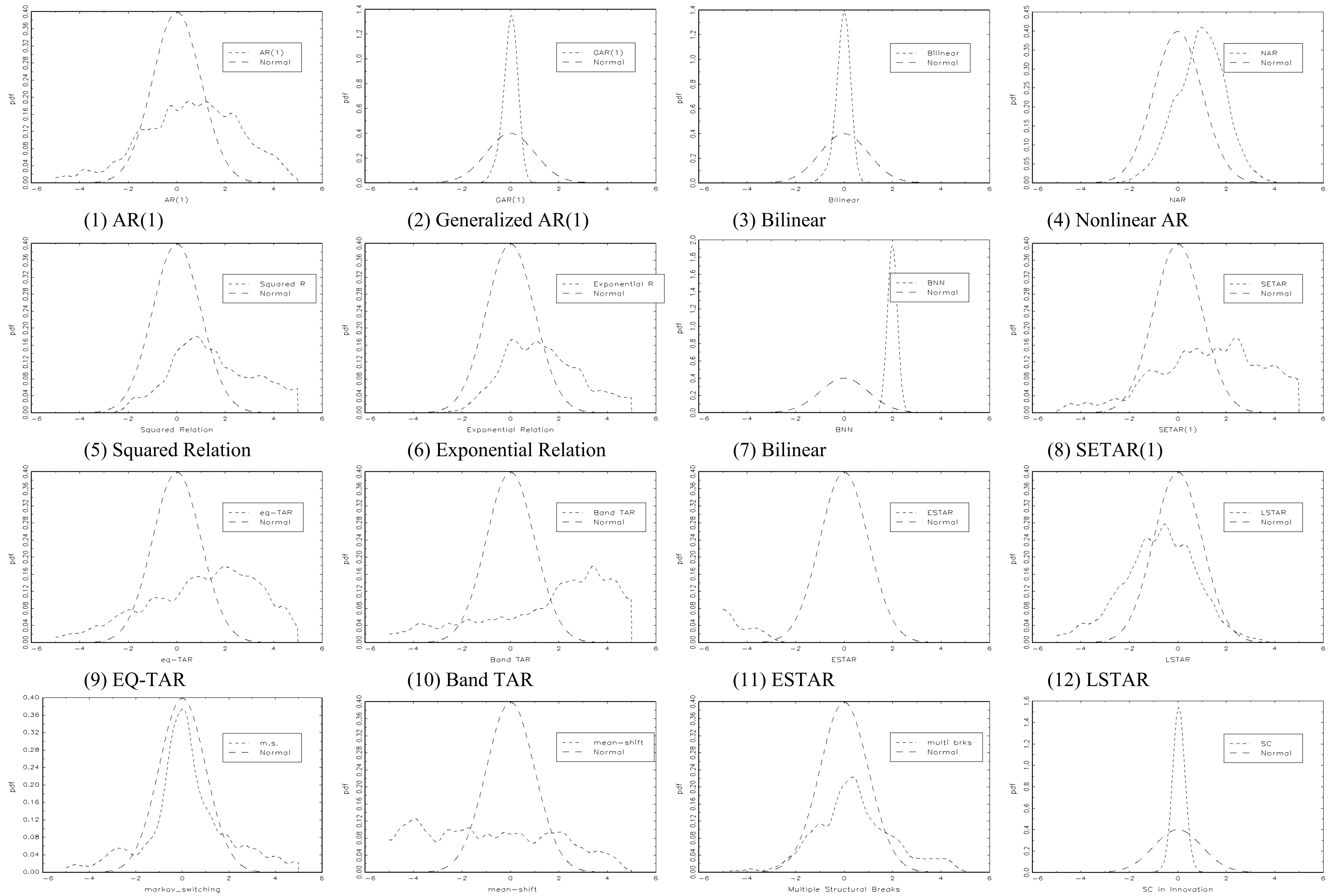
No Time Trend

Country	RALS(2&3)	$\hat{\rho}^2$	RALS 5% cv	ADF
Belgium	−2.967*	0.90	−2.810	−2.642
Canada	−2.167	0.92	−2.817	−2.013
Finland	−3.161*	0.78	−2.745	−2.240
France	−3.303*	0.76	−2.740	−3.296*
Italy	−4.544*	0.77	−2.741	−1.943
Japan	−6.346*	0.81	−2.758	−2.430
Luxembourg	−2.331	0.80	−2.752	−2.482
Netherlands	−3.140	0.59	−2.637	−3.201*
Norway	−3.359*	0.77	−2.745	−3.261*
Spain	−3.527*	0.83	−2.772	−2.311
UK	−3.746*	0.80	−2.754	−2.305
USA	−2.326	0.82	−2.764	−2.330

With Linear Trend

Country	RALS(2&3)	$\hat{\rho}^2$	RALS 5% cv	ADF
Belgium	−3.160	0.91	−3.338	−2.838
Canada	−2.324	0.92	−3.348	−2.189
Finland	−3.099	0.78	−3.246	−2.578
France	−3.328*	0.76	−3.235	−3.255
Italy	−4.424*	0.77	−3.244	−2.100
Japan	−6.329*	0.83	−3.287	−3.350
Luxembourg	−2.365	0.80	−3.268	−2.599
Netherlands	−3.867*	0.59	−3.079	−3.466*
Norway	−3.540*	0.77	−3.245	−3.594*
Spain	−3.774*	0.83	−3.290	−2.582
UK	−3.974*	0.80	−3.266	−2.528
USA	−2.302	0.82	−3.282	−2.401

*significant at 5%. The critical value for the ADF test is −2.87.



(1) AR(1) (2) Generalized AR(1) (3) Bilinear (4) Nonlinear AR

(5) Squared Relation (6) Exponential Relation (7) Bilinear (8) SETAR(1)

(9) EQ-TAR (10) Band TAR (11) ESTAR (12) LSTAR

(13) Markov-switching (14) Level Shift (15) Multiple Shift (16) Variance Shift

Figure 1. The Distribution of the Data Following Various Forms of Non-linear Models

** Additional Results for Tables 2 and 3 for the model with a linear trend (Not to be reported)

Table 2A
Rejection Ratio of Alternative Tests at $\alpha = 0.05$
AR(1) error with AR coefficient 0.5, with Linear Time Trend

T = 50

Distributions		ADF			RALS(2&3)			RALS(t5)		
		p=2	p=4	SC	p=2	p=4	SC	p=2	p=4	SC
Normal	$\phi = 1$	.053	.048	.074	.049	.044	.060	.050	.044	.070
	$\phi = 0.9$	.080	.069	.106	.069	.056	.081	.077	.059	.097
Cauchy	$\phi = 1$	.060	.058	.055	.064	.066	.078	.099	.103	.094
	$\phi = 0.9$	.073	.067	.067	.591	.516	.631	.370	.319	.383
Student t	$\phi = 1$	.057	.049	.060	.054	.060	.073	.058	.055	.063
df=2	$\phi = 0.9$	.078	.062	.081	.252	.201	.281	.248	.195	.255
Double	$\phi = 1$	.058	.053	.075	.055	.048	.074	.056	.049	.071
Exponential	$\phi = 0.9$	.083	.064	.098	.106	.085	.130	.115	.099	.135
Chi-square	$\phi = 1$	.058	.053	.073	.053	.049	.032	.053	.049	.065
df=4	$\phi = 0.9$	.074	.066	.094	.204	.133	.128	.074	.064	.091
Beta(2,2)	$\phi = 1$	.059	.050	.080	.050	.044	.058	.052	.044	.066
	$\phi = 0.9$	.082	.065	.111	.090	.072	.095	.092	.072	.112

T = 100

		p=3	p=6	SC	p=3	p=6	SC	p=3	p=6	SC
Normal	$\phi = 1$	.055	.053	.061	.053	.048	.057	.053	.053	.058
	$\phi = 0.9$	.155	.109	.175	.136	.094	.153	.139	.101	.167
Cauchy	$\phi = 1$	.061	.055	.031	.050	.057	.074	.120	.120	.108
	$\phi = 0.9$	.110	.100	.115	.867	.800	.925	.652	.627	.675
Student t	$\phi = 1$	.055	.045	.042	.049	.047	.071	.050	.050	.051
df=2	$\phi = 0.9$	.133	.090	.145	.525	.411	.608	.610	.495	.644
Double	$\phi = 1$	.059	.049	.061	.049	.045	.066	.051	.044	.059
Exponential	$\phi = 0.9$	.150	.105	.177	.223	.164	.276	.277	.196	.305
Chi-square	$\phi = 1$	.060	.049	.059	.050	.046	.021	.053	.044	.056
df=4	$\phi = 0.9$	.149	.102	.174	.524	.370	.414	.147	.100	.168
Beta(2,2)	$\phi = 1$	.052	.046	.064	.042	.036	.044	.043	.039	.047
	$\phi = 0.9$	.157	.106	.185	.235	.155	.259	.239	.159	.281

Table 3A
Rejection Ratio of Alternative Tests at $\alpha = 0.05$
MA(1) error with MA coefficient −0.5, with Linear Time Trend

<u>T = 50</u>

Distributions		ADF			RALS(2&3)			RALS(t5)		
		p=2	p=4	SC	p=2	p=4	SC	p=2	p=4	SC
Normal	$\phi = 1$	.111	.054	.127	.091	.046	.096	.096	.051	.108
	$\phi = 0.9$	.176	.076	.191	.145	.067	.148	.157	.072	.170
Cauchy	$\phi = 1$	.102	.061	.085	.262	.125	.270	.130	.094	.126
	$\phi = 0.9$	.138	.080	.126	.784	.608	.801	.423	.346	.424
Student t	$\phi = 1$	.101	.055	.099	.147	.076	.166	.115	.066	.120
df=2	$\phi = 0.9$	.158	.071	.157	.441	.263	.467	.401	.245	.397
Double	$\phi = 1$	.114	.059	.119	.116	.057	.124	.120	.061	.125
Exponential	$\phi = 0.9$	.177	.082	.184	.226	.109	.237	.250	.122	.248
Chi-square	$\phi = 1$	.111	.059	.118	.144	.059	.087	.102	.052	.108
df=4	$\phi = 0.9$	.168	.071	.171	.420	.183	.295	.153	.071	.154
Beta(2,2)	$\phi = 1$	.121	.055	.135	.101	.052	.099	.101	.051	.108
	$\phi = 0.9$	.181	.076	.199	.186	.086	.179	.201	.091	.211

<u>T = 100</u>

		p=3	p=6	SC	p=3	p=6	SC	p=3	p=6	SC
Normal	$\phi = 1$	.089	.052	.134	.080	.048	.122	.083	.051	.130
	$\phi = 0.9$	.260	.123	.396	.230	.112	.336	.243	.119	.359
Cauchy	$\phi = 1$	.081	.056	.067	.156	.075	.310	.134	.111	.144
	$\phi = 0.9$	.188	.111	.306	.934	.839	.972	.658	.620	.680
Student t	$\phi = 1$	.079	.044	.105	.098	.050	.195	.090	.052	.133
df=2	$\phi = 0.9$	.233	.106	.360	.684	.473	.823	.745	.556	.820
Double	$\phi = 1$	.093	.052	.135	.084	.048	.152	.086	.048	.136
Exponential	$\phi = 0.9$	.261	.125	.396	.371	.185	.533	.436	.228	.570
Chi-square	$\phi = 1$	.085	.049	.127	.095	.048	.091	.080	.049	.120
df=4	$\phi = 0.9$	.268	.128	.397	.708	.429	.725	.253	.121	.373
Beta(2,2)	$\phi = 1$	.084	.048	.136	.077	.040	.113	.078	.040	.125
	$\phi = 0.9$	.265	.135	.400	.362	.189	.483	.377	.192	.519