Hitotsubashi Journal of Economics 34 (1993) 147-164. © The Hitotsubashi Academy

NORMAL TESTS FOR A UNIT ROOT IN THE AUTOREGRESSIVE TIME SERIES MODEL*

TAKU YAMAMOTO

Abstract

Test for a unit root has been widely used to investigate the dynamic nature of economic time series. However, it is known to be rather cumbersome, since the asymptotic distributions of the conventional test statistics are non-normal and we have to resort to the special tables for critical values. Recently, Choi derived the test statistic whose asymptotic distribution is normal. This paper proposes similar test statistics and examine their empirical size and power through the Monte Carlo experiment, and compare them with those of the conventional Dickey-Fuller test statistics.

I. *Introduction*

Let $\{x_t\}$ be generated from

$$x_t=\gamma_0+\beta_1x_{t-1}+\beta_2x_{t-2}+\ \dots\ +\beta_px_{t-p}+\eta_t\ (t=1,2,\ \dots\ ,T), \tag{1.1}$$

where $\gamma_0\neq 0$, and η_t $(t=1,2,\ \dots\ ,T)$ are independently identically distributed as $N(0,\sigma^2)$. We are interested in testing for a unit root hypothesis H_0: $\sum_{i=1}^{p}\beta_i=1$ against the alternative hypothesis H_1: $\sum_{i=1}^{p}\beta_i<1$ or H_2: $\sum_{i=1}^{p}\beta_i\neq 1$. Under H_0, the process $\{x_t\}$ has a unit root as well as drift. We may note that the existence of the constant term $\gamma_0\neq 0$ is not essential in the discussion of the paper.

The well-known Dickey-Fuller tests for a unit root developed by Fuller (1976) or Dickey and Fuller (1979) are based upon the following transformed model

$$x_t=\gamma_0+\rho x_{t-1}+\beta_1^*\Delta x_{t-1}+\ \dots\ +\beta_{p-1}^*\Delta x_{t-p+1}+\gamma_1t+\eta_t, \tag{1.2}$$

where $\rho=\sum_{i=1}^{p}\beta_i$, $\beta_j^*=-\sum_{i=j+1}^{p}\beta_i$ $(j=1,2,\dots,p-1)$, and Δ is the difference operator. The time trend t is introduced since the process $\{x_t\}$ is a trend plus stationary process under the alternative hypothesis H_1. Let $\hat{\rho}_\tau$ be the least squares estimator of ρ. The Dickey-Fuller test statistics, $T(\hat{\rho}_\tau-1)$ and the usual t-statistic for testing the hypothesis $\rho=1$, are known to be non-normal and necessary critical values are tabulated in Fuller (1976, Ch.8).

* I am grateful to Ms. Motoko Yuasa for an able research assistance for the computations reported in the paper. The research was in part supported by Grant-in-Aid 04630013 of the Ministry of Education, science and Culture of Japan.

Recently, Choi (1991) has proposed a new test statistic which is asymptotically normal. Yamamoto (1992a) has also proposed a simple and convenient method for circumventing the difficulties encountered in the statistical inference in a vector autoregressive model which may have unit roots. It was shown that the usual Wald test statistic of constraints is asymptotically chi-square. See also Toda and Yamamoto (1993).

In this paper, as a special case of Yamamoto (1992a), we propose test statistics for a unit root whose asymptotic distributions are normal. We also propose a simple way to enhance the power of the tests. We examine empirical sizes and powers of those tests through the Monte Carlo experiment. Further, we compare their powers with the Dickey-Fuller tests.

II. *Test for a Unit Root Based Upon the Asymptotic Normality*

In this section, we briefly explain the test procedure. In order to test for a unit root based upon asymptotic normality, we consider the following regression model:

$$x_t=\gamma_0+\beta_1 x_{t-1}+\beta_2 x_{t-2}+\ \dots\ +\beta_p x_{t-p}+\beta_{p+1}x_{t-p-1}+\gamma_1 t+\eta_t. \tag{2.1}$$

The key elements of the above regression model are that it is expressed in levels not in transformed form and it includes an extra lagged variable x_{t-p-1}. Since x_{t-p-1} and t are redundant variables, we have $\beta_{p+1}=\gamma_1=0$.

Following Dickey and Fuller (1979) or Sims, Stock and Watson (hereafter SSW, 1990) the regression model (2.1) can be expressed as

$$x_t=\delta_{1,1}z^1{}_{t-1}+\delta_{1,2}z^1{}_{t-2}+\ \dots\ +\delta_{1,p}z^1{}_{t-p}+\delta_2+\delta_3 z_t{}^3+\delta_4(t-1)+\eta_t, \tag{2.2}$$

where $z^1{}_{t-j}=\Delta x_{t-j}-\mu\ (j=1,2,\ \dots\ ,p)$, $z_t{}^3=x_{t-1}-\mu(t-1)$,

$$\delta_{1,j}=-\textstyle\sum_{i=j+1}^{p+1}\beta_i\ \ (j=1,2,\ \dots\ ,p),\ \delta_2=\mu+\gamma_1,$$

$$\delta_3=\textstyle\sum_{i=1}^{p+1}\beta_i,\ \delta_4=\delta_3\mu+\gamma_1,\ \text{and}\ \mu=\gamma_0/(1-\sum_{j=1}^{p}\delta_{1,j}).$$

Under H_0, that is, $\sum_{i=1}^{p}\beta_i=1$, and the assumptions $\beta_{p+1}=\gamma_1=0$, we have $\delta_{1,p}=0$, $\delta_2=\mu$, $\delta_3=1$, and $\delta_4=\mu$. Using the above representation, we have the following result:

Theorem: Let $\hat\beta_j$ *the least squares estimator of* β_j $(j=1,\ \dots\ ,p+1)$ *in the regression model* (2.1). *We have*

(a)
$$\hat\tau_{\tau N}\xrightarrow{D}N(0,1), \tag{2.3}$$

and

(b)
$$\sqrt{T}(\hat\rho_{\tau N}-1)\xrightarrow{D}N(0,1), \tag{2.4}$$

where $\hat\rho_{\tau N}=\sum_{i=1}^{p}\hat\beta_i$, $\hat\tau_{\tau N}$ *is the conventional t statistic for testing* $\sum_{i=1}^{p}\beta_i=1$, *that is,* $\hat\tau_{\tau N}=(\hat\rho_{\tau N}-1)/SE(\hat\rho_{\tau N})$, $SE(\hat\rho_{\tau N})$ *is the relevant estimate of asymptotic standard error of* $\hat\rho_{\tau N}$, *and*

$\xrightarrow{D}$ *implies the convergence in the distribution.*

The proof is essentially a special case of Yamamoto (1992a), and is given in Appendix A in detail. We may note that the proof of (b) needs an extra explanation.

Remark 1: Choi (1991) derived a similar result from the following regression model, by adding x_{t-p-1} variable to (1.2) not to (1.1):

$$x_t = \gamma_0 + \rho_1 x_{t-1} + \beta_1^* \Delta x_{t-1} + \dots + \beta_{p-1}^* \Delta x_{t-p+1} + \rho_{p+1} x_{t-p-1} + \gamma_1 t + \eta_t. \tag{2.5}$$

He showed that the conventiooal t statistic for testing $\rho_1 = 1$ based upon the OLS estimate of ρ_1 is asymptotically normal. It corsesponds to (a) of the above theorem, but they are not algebraically the same in general.

Remark 2: It is important to note that the proposed test statistics $\hat{\rho}_{\tau N}$ and $\hat{\tau}_{\tau N}$ include only the first p coefficient estimates and intentionally drop $\hat{\beta}_{p+1}$. If we included $\hat{\beta}_{p+1}$, we have $\sum_{i=1}^{p+1} \hat{\beta}_i = \hat{\delta}_3$. Then $\sqrt{T}(\sum_{i=1}^{p+1} \hat{\beta}_i - 1)$ becomes degenerate, and $T(\sum_{i=1}^{p+1} \hat{\beta}_i - 1)$ has a non-normal limiting distribution as in the Dickey-Fuller statistics. This is easily seen in the argument in Appendix A.

Remark 3: The proposed method can be extended to more general models. Consider now that the model (1.1) has double unit roots. For simplicity, we here assume that we know a priori that $\gamma_0 = 0$. In this case, we consider the following regression model:

$$x_t = \beta_1 x_{t-1} + \beta_2 x_{t-2} + \dots + \beta_p x_{t-p} + \beta_{p+1} x_{t-p-1} + \beta_{p+2} x_{t-p-2} + \eta_t. \tag{2.6}$$

We now have two additional lagged variables x_{t-p-1} and x_{t-p-2}. The true values of β_{p+1} and β_{p+2} are zero. The transformed representation is given as follows:

$$x_t = \delta_{1,1} z^1{}_{t-1} + \delta_{1,2} z^1{}_{t-2} + \dots + \delta_{1,p} z^1{}_{t-p} + \delta_2 z_t{}^2 + \delta_3 z_t{}^3 + \eta_t, \tag{2.7}$$

where

$$z^1{}_{t-j} = \Delta^2 x_{t-j} \ (j=1, 2, \dots, p),\ z_t{}^2 = \Delta x_{t-1},\ z_t{}^3 = x_{t-1},$$

$$\delta_{1,j} = \sum_{i=j+2}^{p+2} (i-2)\beta_i \quad (j=1, 2, \dots, p);$$

$$\delta_2 = -\sum_{i=1}^{p+2} (i-1)\beta_i, \quad \text{and} \quad \delta_3 = \sum_{i=1}^{p+2} \beta_i.$$

Following Hasza and Fuller (1979), the null hypothesis for double unit roots is given by $\delta_2 = \delta_3 = 1$. We can formulate the hypothesis in terms of the original parameter β_i's as follows:

$$H_0': R\beta = r, \tag{2.8}$$
$$H_1': R\beta \neq r,$$

where $\beta = [\beta_1, \beta_2, \dots, \beta_{p+2}]'$, $r = [1, 1]'$, and R is the $2 \times (p+2)$ matrix such that

$$R = \begin{bmatrix} 1 & 1 & 1 & 1 & \dots & 1 & 0 & 0 \\ 0 & -1 & -2 & -3 & \dots & -(p-1) & 0 & 0 \end{bmatrix}.$$

We can easily construct the Wald statistic for the hypothesis. Yamamoto (1992a) has shown that it is asymptotically ch-isquare. It corresponds to the test statistic $\Phi_1(2)$ in Hasza and Fuller (1979) which is asymptotically non-normal. The method can be further generalized to regression models which have the constant term and/or trend terms. However, if we are interested in the Wald tests which also involve coefficients of the constant term, the present method will no longer produces a chi-square distribution. See Yamamoto (1992b) for treatment of the case.

III. *Empirical Size of the Normal Tests*

3.1 Estimators to be Examined

While the proposed tests are convenient, it is important to examine their properties in small and moderate samples before we use them in practice. In this experiment, we consider two types of very simple data generating processes:

$$x_t = x_{t-1} + \eta_t, \tag{3.1}$$

and

$$x_t = 1 + x_{t-1} + \eta_t, \tag{3.2}$$

where η_t is independently identically distributed as $N(0, 1)$.

We now consider test statistics which are asymptotically normal. For data from the first process (3.1), two relevant regression models are given by

$$x_t = \beta_1 x_{t-1} + \beta_2 x_{t-2} + \eta_t, \tag{3.3}$$

and

$$x_t = \gamma_0 + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \eta_t. \tag{3.4}$$

The least squares estimates of β_1 are denoted as $\hat{\rho}_N$ and $\hat{\rho}_{\mu N}$, respectively. The conventional t statistics are denoted as $\hat{\tau}_N$ and $\hat{\tau}_{\mu N}$, respectively. For the data from the second process (3.2), the regression model is given by

$$x_t = \gamma_0 + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \gamma_1 t + \eta_t. \tag{3.5}$$

The least squares estimate of β_1 is denoted as $\hat{\rho}_{\tau N}$ and the concentional t statistic is denoted as $\hat{\tau}_{\tau N}$. Here, the subscript N indicates that these test statistics are asymptotically standard normal. Their asymptotic distributions are summarized as follows:

$$\sqrt{T}(\hat{\rho}_i - 1) \xrightarrow{D} N(0, 1) \qquad (i = N, \mu N, \tau N), \tag{3.6}$$

and

$$\hat{\tau}_i \xrightarrow{D} N(0, 1) \qquad (i = N, \mu N, \tau N). \tag{3.7}$$

We call these test statistics as the *unmodified test statistics* in order to distinguish from the ones introduced below. Since the data generating processes (3.1) and (3.2) do not include higher order stationary parts, that is, $p=1$, the test statistics $\hat{\tau}_i$ ($i=N$, μN, τN) exactly coincide with the ones proposed by Choi (1991). Thus, these experiments also serve as ones for his method.

Next, we introduce a simple modification to the above test statistics, by adding a variable x_{t-10} instead of x_{t-2} in the regression models. The modification is expected to give more power to the tests. For data from the first process (3.1), the regression models are, instead of (3.3) and (3.4), given respectively as

$$x_t=\beta_1 x_{t-1}+\beta_2 x_{t-10}+\eta_t, \tag{3.8}$$

and

$$x_t=\gamma_0+\beta_1 x_{t-1}+\beta_2 x_{t-10}+\eta_t. \tag{3.9}$$

The least squares estimates of β_1 are denoted as $\hat{\rho}_{Nm}$ and $\hat{\rho}_{\mu Nm}$, respectively. The conventional t statistics are denoted as $\hat{\tau}_{Nm}$ and $\hat{\tau}_{\mu Nm}$, respectively. For data from the second process (3.2), the regression model is, instead of (3.5), given by

$$x_t=\gamma_0+\beta_1 x_{t-1}+\beta_2 x_{t-10}+\gamma_1 t+\eta_t. \tag{3.10}$$

The least squares estimate of β_1 is denoted as $\hat{\rho}_{\tau Nm}$ and corresponding t statistic is denoted as $\hat{\tau}_{\tau Nm}$. Here, the added subscript m indicates that these test statistics based upon the modified regression models. We call them the *modified test statistics*. The asymptotic distributions of these statistics are summarized as follows:

$$3\sqrt{T}(\hat{\rho}_i-1)\xrightarrow{D}N(0,1) \qquad (i=Nm,\ \mu Nm,\ \tau Nm), \tag{3.11}$$

and

$$\hat{\tau}_i\xrightarrow{D}N(0,1) \qquad (i=Nm,\ \mu Nm,\ \tau Nm). \tag{3.12}$$

The explanation for the normalizing factor $3\sqrt{T}$ is given in Appendix B.

3.2 Design of Experiment

Sample sizes are $T=25$, 50, 100, 250, 500, and 1000. For each process, 40,000 series are generated. Normal deviates are obtained from RANN2 of Facom Library Function SSLII, which is based upon the Box-Muller procedure.

3.3 Small Sample Characteristics of the Tests

For illustration, Figs. 1–4 show the histograms of density functions of $\sqrt{T}(\hat{\rho}_{\mu N}-1)$, $\hat{\tau}_{\mu N}$, $3\sqrt{T}(\hat{\rho}_{\mu Nm}-1)$, and $\hat{\tau}_{\mu Nm}$, respectively. The range of each cell of the histogram is 0.08, and each curve is drawn by averaging three neighbouring cells for smoothing out an otherwise erratic curve. These figures generally show the pattern of convergence to $N(0, 1)$ as T increases. Figures of other statsitics show similar patterns and are omitted for the sake of space.

FIGURE 1. $\sqrt{T}(\hat{\rho}_{\mu N}-1)$ IN (3.4)

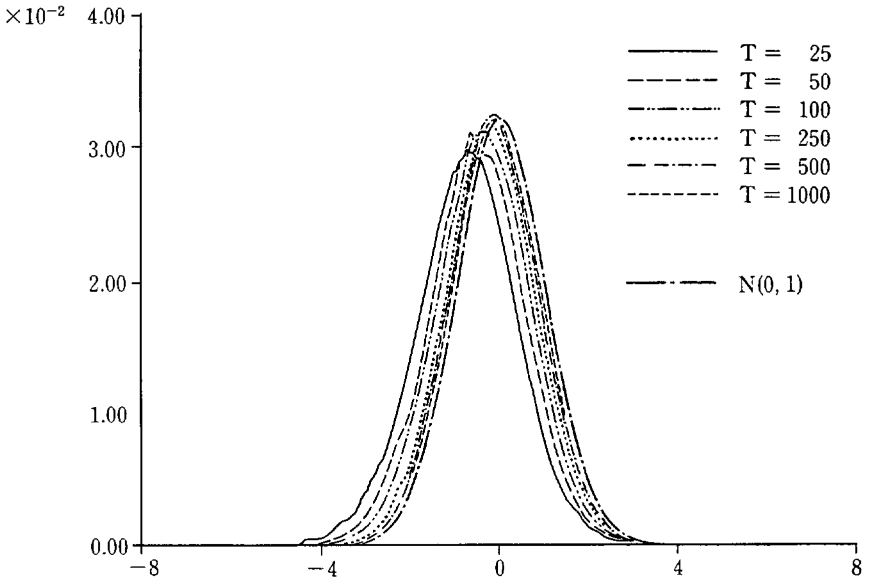


FIGURE 2. $\hat{\tau}_{\mu N}$ IN (3.4)

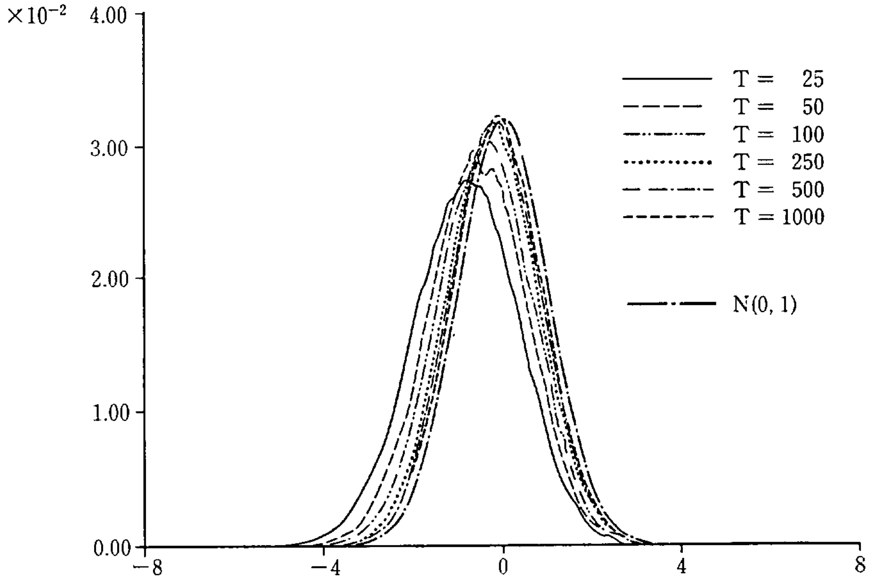


FIGURE 3. $3\sqrt{T}(\hat{\rho}_{\mu Nm}-1)$ IN (3.9)

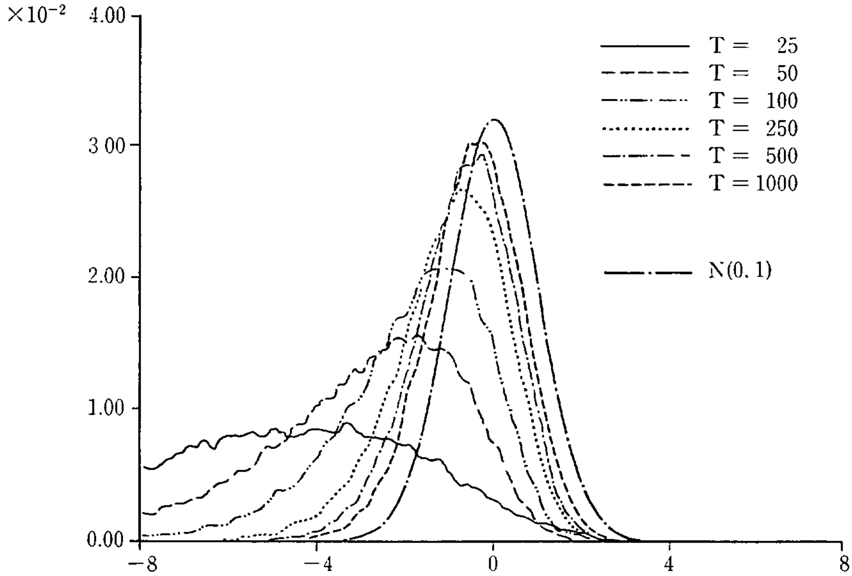
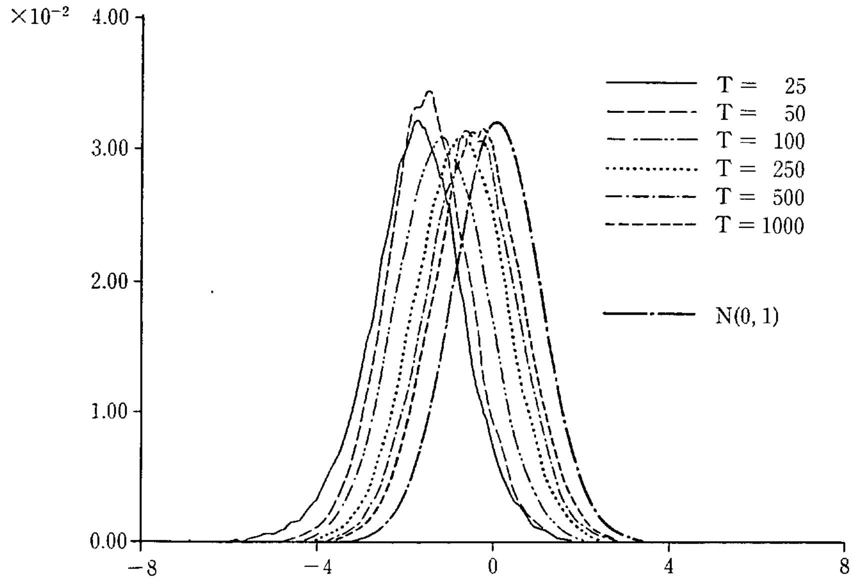


FIGURE 4. $\hat{\tau}_{\mu Nm}$ IN (3.9)



Basic characteristics are tabulated in Table 1. As explained earlier, all test statistics which appropriately normalized converge to the standard normal variate as T goes to infinity. Thus, we expect that, in the limit, bias and skewness are zero, variance is unity, and kurtosis is three.

Major findings are summarized as follows:

(1) Generally, biases are negative, variance are greater than unity, skewnesses are negative, and kurtosises are greater than three. Negative baises are more profound than negative skewnesses. As T becomes large, they indicate the convergence to $N(1, 0)$.

TABLE 1. SMALL SAMPLE PROPERTIES OF $\sqrt{T}(\hat{\rho}_i-1)$ AND $\hat{\tau}_i$*

T	Mean	Variance	Skewness	Kurtosis	Mean	Variance	Skewness	Kurtosis
	$\hat{\rho}_N$				$\hat{\tau}_N$			
25	−0.237	1.20	−0.109	3.06	−0.218	1.07	−0.073	3.40
50	−0.157	1.07	−0.058	2.96	−0.151	1.04	−0.051	3.15
100	−0.117	1.06	−0.037	2.97	−0.114	1.04	−0.036	3.08
250	−0.069	1.02	−0.001	3.01	−0.068	1.02	−0.002	3.06
500	−0.048	1.01	−0.009	2.98	−0.048	1.00	−0.010	3.00
1000	−0.036	1.01	−0.003	3.05	−0.036	1.01	−0.003	3.06
	$\hat{\rho}_{\mu N}$				$\hat{\tau}_{\mu N}$			
25	−0.852	1.36	−0.137	3.08	−0.795	1.27	−0.130	3.39
50	−0.581	1.19	−0.089	3.02	−0.561	1.15	−0.088	3.18
100	−0.413	1.11	−0.042	2.99	−0.406	1.09	−0.043	2.09
250	−0.255	1.04	−0.000	3.01	−0.253	1.04	−0.000	3.05
500	−0.179	1.02	−0.008	2.98	−0.178	1.01	−0.007	3.01
1000	−0.128	1.02	−0.001	3.05	−0.128	1.02	−0.001	3.06
	$\hat{\rho}_{\tau N}$				$\hat{\tau}_{\tau N}$			
25	−1.524	1.64	−0.160	3.07	−1.430	1.50	−0.167	3.38
50	−1.063	1.36	−0.118	3.02	−1.026	1.30	−0.107	3.13
100	−0.747	1.17	−0.066	3.00	−0.733	1.15	−0.066	3.08
250	−0.470	1.08	−0.005	2.96	−0.466	1.07	−0.004	3.00
500	−0.329	1.03	−0.007	2.95	−0.328	1.02	−0.007	2.97
1000	−0.234	1.02	−0.018	2.97	−0.234	1.02	−0.018	2.99
	$\hat{\rho}_{Nm}$				$\hat{\tau}_{Nm}$			
25	−1.684	6.81	−1.489	6.77	−0.641	1.18	−0.069	3.69
50	−1.028	3.86	−1.345	6.14	−0.549	1.02	0.022	3.17
100	−0.673	2.20	−0.996	5.00	−0.430	1.04	0.002	3.05
250	−0.388	1.26	−0.580	3.69	−0.277	1.03	−0.018	3.07
500	−0.280	1.11	−0.387	3.25	−0.211	1.01	−0.016	3.01
1000	−0.195	1.06	−0.249	3.08	−0.149	1.02	−0.006	2.99
	$\hat{\rho}_{\mu Nm}$				$\hat{\tau}_{\mu Nm}$			
25	−5.695	15.68	−0.620	3.47	−1.808	1.21	−0.223	3.93
50	−3.089	5.93	−1.076	4.94	−1.566	0.95	−0.018	3.34
100	−1.830	2.93	−0.950	4.59	−1.221	1.06	0.057	3.00
250	−1.009	1.62	−0.617	3.72	−0.808	1.08	−0.012	3.04
500	−0.693	1.28	−0.404	3.29	−0.592	1.04	−0.013	3.01
1000	−0.480	1.15	−0.262	3.12	−0.422	1.04	0.002	3.01
	$\hat{\rho}_{\tau Nm}$				$\hat{\tau}_{\tau Nm}$			
25	−8.889	19.03	−0.311	3.08	−2.437	1.28	−0.453	4.22
50	−5.502	8.33	−0.771	3.72	−2.388	0.77	−0.115	3.40
100	−3.235	4.44	−0.841	4.03	−1.980	0.98	0.119	3.06
250	−1.746	2.16	−0.653	3.82	−1.375	1.15	0.020	3.02
500	−1.164	1.50	−0.427	3.41	−1.004	1.10	−0.001	3.00
1000	−0.813	1.24	−0.274	3.13	−0.735	1.07	−0.001	3.01

* The normalizing factor is $3\sqrt{T}$ for the modified ρ-type statistics.

(2) Generally, τ-type statistics are closer to $N(0, 1)$ than the corresponding ρ-type statistics.
(3) The modification makes the statistics deviate from $N(0, 1)$. Especially, the deviation is larger for the ρ-type statistics. It is reflected in large negative biases and large variances, particularly when the sample size is small.

3.4 Empirical Sizes of the Tests

Tables 2a and 2b show empirical sizes for both two sided and lower one sided tests for

TABLE 2a. EMPIRICAL SIZES OF ρ-TYPE TESTS

	T	Size of 1% Lower Tail	Upper Tail	Sum	Size of 5% Lower Tail	Upper Tail	Sum	Lower 1% Tail	Lower 5% Tail
$\hat{\rho}_N$									
	25	0.015	0.003	0.018	0.055	0.017	0.072	0.027	0.095
	50	0.010	0.003	0.013	0.043	0.018	0.061	0.019	0.077
	100	0.009	0.004	0.013	0.037	0.020	0.057	0.016	0.069
	250	0.006	0.004	0.011	0.030	0.023	0.053	0.013	0.059
	500	0.006	0.004	0.010	0.027	0.023	0.050	0.012	0.056
	1000	0.006	0.005	0.011	0.028	0.024	0.052	0.012	0.055
$\hat{\rho}_{Nm}$									
	25	0.278	0.008	0.286	0.360	0.019	0.379	0.310	0.408
	50	0.151	0.000	0.151	0.230	0.002	0.232	0.179	0.283
	100	0.084	0.000	0.084	0.150	0.002	0.152	0.107	0.199
	250	0.03t	0.000	0.038	0.088	0.005	0.093	0.054	0.131
	500	0.023	0.000	0.023	0.065	0.008	0.073	0.036	0.104
	1000	0.016	0.001	0.017	0.050	0.012	0.062	0.026	0.085
$\hat{\rho}_{\mu N}$									
	25	0.071	0.001	0.072	0.170	0.006	0.176	0.103	0.245
	50	0.037	0.001	0.038	0.103	0.008	0.111	0.059	0.165
	100	0.021	0.002	0.023	0.072	0.011	0.083	0.036	0.122
	250	0.012	0.003	0.015	0.047	0.015	0.062	0.022	0.085
	500	0.009	0.003	0.012	0.038	0.017	0.055	0.016	0.074
	1000	0.008	0.004	0.011	0.035	0.020	0.055	0.015	0.067
$\hat{\rho}_{\mu Nm}$									
	25	0.773	0.004	0.777	0.831	0.007	0.838	0.797	0.857
	50	0.516	0.000	0.516	0.631	0.000	0.631	0.561	0.692
	100	0.285	0.000	0.285	0.409	0.000	0.409	0.332	0.484
	250	0.112	0.000	0.112	0.209	0.001	0.210	0.147	0.279
	500	0.058	0.000	0.058	0.132	0.003	0.135	0.082	0.192
	1000	0.032	0.001	0.033	0.088	0.006	0.094	0.049	0.139
$\hat{\rho}_{\tau N}$									
	25	0.201	0.000	0.201	0.358	0.002	0.360	0.260	0.453
	50	0.098	0.000	0.098	0.219	0.003	0.222	0.139	0.304
	100	0.048	0.001	0.049	0.231	0.005	0.137	0.073	0.204
	250	0.021	0.001	0.022	0.077	0.009	0.086	0.037	0.130
	500	0.013	0.002	0.015	0.054	0.012	0.066	0.024	0.099
	1000	0.010	0.003	0.013	0.044	0.015	0.059	0.019	0.082
$\hat{\rho}_{\tau Nm}$									
	25	0.939	0.002	0.941	0.958	0.003	0.961	0.947	0.965
	50	0.856	0.000	0.856	0.918	0.000	0.918	0.884	0.942
	100	0.570	0.000	0.570	0.701	0.000	0.701	0.624	0.764
	250	0.260	0.000	0.260	0.399	0.000	0.399	0.312	0.487
	500	0.126	0.000	0.126	0.245	0.001	0.246	0.167	0.327
	1000	0.064	0.000	0.064	0.150	0.003	0.153	0.092	0.220

TABLE 2b. EMPIRICAL SIZES OF τ-TYPE TESTS

	T	Size of 1% Lower Tail	Upper Tail	Sum	Size of 5% Lower Tail	Upper Tail	Sum	Lower 1% Tail	Lower 5% Tail
$\hat{\tau}_N$									
	25	0.014	0.004	0.018	0.047	0.017	0.064	0.024	0.083
	50	0.011	0.004	0.015	0.040	0.018	0.058	0.019	0.072
	100	0.009	0.004	0.013	0.036	0.020	0.056	0.017	0.068
	250	0.007	0.004	0.011	0.030	0.023	0.053	0.013	0.058
	500	0.006	0.004	0.010	0.027	0.023	0.050	0.012	0.056
	1000	0.006	0.005	0.011	0.028	0.024	0.052	0.012	0.055
$\hat{\tau}_{Nm}$									
	25	0.038	0.003	0.041	0.102	0.010	0.112	0.057	0.165
	50	0.022	0.002	0.024	0.079	0.008	0.087	0.038	0.136
	100	0.017	0.002	0.019	0.066	0.010	0.076	0.031	0.116
	250	0.013	0.003	0.016	0.048	0.014	0.062	0.022	0.089
	500	0.006	0.005	0.011	0.026	0.024	0.050	0.010	0.051
	1000	0.008	0.003	0.011	0.036	0.018	0.055	0.016	0.069
$\hat{\tau}_{\mu N}$									
	25	0.060	0.002	0.062	0.145	0.007	0.152	0.086	0.216
	50	0.033	0.002	0.035	0.096	0.009	0.105	0.053	0.153
	100	0.021	0.002	0.023	0.069	0.011	0.080	0.035	0.117
	250	0.012	0.003	0.015	0.047	0.015	0.062	0.022	0.085
	500	0.009	0.003	0.012	0.038	0.017	0.055	0.016	0.073
	1000	0.008	0.004	0.012	0.036	0.020	0.056	0.015	0.067
$\hat{\tau}_{\mu Nm}$									
	25	0.222	0.000	0.222	0.431	0.001	0.432	0.297	0.556
	50	0.142	0.000	0.142	0.342	0.000	0.342	0.209	0.472
	100	0.091	0.000	0.091	0.239	0.001	0.240	0.142	0.345
	250	0.046	0.001	0.047	0.133	0.004	0.137	0.073	0.209
	500	0.026	0.001	0.027	0.091	0.007	0.098	0.045	0.152
	1000	0.018	0.001	0.019	0.065	0.010	0.075	0.031	0.116
$\hat{\tau}_{\tau N}$									
	25	0.168	0.001	0.169	0.320	0.002	0.322	0.222	0.418
	50	0.086	0.001	0.087	0.204	0.004	0.208	0.125	0.286
	100	0.045	0.001	0.046	0.125	0.006	0.131	0.070	0.197
	250	0.021	0.002	0.023	0.076	0.010	0.086	0.037	0.128
	500	0.013	0.002	0.015	0.053	0.012	0.065	0.024	0.098
	1000	0.010	0.003	0.013	0.044	0.015	0.059	0.019	0.082
$\hat{\tau}_{\tau Nm}$									
	25	0.423	0.000	0.423	0.660	0.000	0.660	0.519	0.768
	50	0.401	0.000	0.401	0.694	0.000	0.694	0.522	0.811
	100	0.279	0.000	0.279	0.516	0.000	0.516	0.370	0.640
	250	0.131	0.000	0.131	0.244	0.001	0.245	0.188	0.399
	500	0.066	0.000	0.066	0.181	0.002	0.183	0.103	0.273
	1000	0.038	0.001	0.039	0.117	0.005	0.122	0.062	0.189

1% and 5% significance levels.

(1) Since the test statistics are negatively biased as explained earlier, lower tail empirical sizes are larger than the corresponding nominal sizes and upper tail empirical sizes are smaller, especially when the sample size is not large.

(2) The ρ-type tests are worse than the τ-type tests in the sense their sizes are much larger than the corresponding nominal sizes, as expected from the above results. For the unmodified test statistics, as the sample size increases, empirical sizes converge to the corresponding nominal sizes. The modified test statistics also exhibit the convergence, but their sizes remain much larger than the nominal sizes.

(3) In conclusion, in terms of size, the unmodified τ-type tests are the best among them, and they are close to $N(0, 1)$ when the sample size is 500 or larger.

IV. *Power of the Tests*

4.1 Test Statistics to be Compared

In addition to the test statistics proposed in this paper, we compare their power with that of the well-known test statistics by Dickey and Fuller (1979). For data from the first process (3.1), The Dickey-Fuller statistics are obtained by estimating following the models:

$$x_t = \beta_1 x_{t-1} + \eta_t, \tag{4.1}$$

and

$$x_t = \gamma_0 + \beta_1 x_{t-1} + \eta_t. \tag{4.2}$$

The least squares estimates of β_1 are denoted as $\hat{\rho}$ $\hat{\rho}_\mu$, respectively. The conventional t statistics are denoted as $\hat{\tau}$ and $\hat{\tau}_\mu$, respectively. For the data from the second process (3.2), the Dickey-Fuller statistic is obtained from

$$x_t = \gamma_0 + \beta_1 x_{t-1} + \gamma_1 t + \eta_t. \tag{4.3}$$

The least squares estimate of β_1 is denoted as $\hat{\rho}_\tau$, and the corresponding t statistic is denoted as $\hat{\tau}_\tau$. The notation for these statistics is exactly the same as used in Dickey and Fuller (1979), and necessary critical values are tabulated in Fuller (1976).

4.2 Design of Experiment

In this experiment, we consider two types of data generating processes for representing alternative hypotheses. Instead of (3.1) and (3.2), we have respectively

$$x_t = \rho x_{t-1} + \eta_t, \tag{4.4}$$

and

$$(x_t - t) = \rho(X_{t-1} - (t-1)) + \eta_t, \tag{4.5}$$

where η_t is independently identically distributed as $N(0, 1)$. Since we are generally interested in testing for a unit root hypothesis against a stationary alternative, we only consider the case where $\rho < 1$. In particular, we set $\rho = 0.95$, 0.90, and 0.80.

Since the proposed test statistics are generally negatively biased, we first obtain the size-corrected critical values from the experiments reported in section 3. These critical values are tabulated in Table 3.

Sample sizes are $T=25$, 50, 100, 250, 500, and 1,000. For each process, 40,000 series are generated. Normal deviates are obtained as in the previous experiment.

TABLE 3. CRITICAL VALUES OF 5% AND 1% LOWER TAIL TESTS

5%

T	$\hat{\rho}_N$	$\hat{\rho}_{Nm}$	$\hat{\rho}_{\mu N}$	$\hat{\rho}_{\mu Nm}$	$\hat{\rho}_{\tau N}$	$\hat{\rho}_{\tau Nm}$
25	−2.01	−6.73	−2.82	−12.94	−3.68	−16.56
50	−1.88	−4.19	−2.41	−7.67	−3.00	−10.89
100	−1.82	−3.08	−2.16	−4.98	−2.56	−7.16
250	−1.73	−2.38	−1.93	−3.29	−2.19	−4.39
500	−1.69	−2.13	−1.83	−2.67	−1.99	−3.28
1000	−1.69	−1.97	−1.79	−2.31	−1.90	−2.72

1%

T	$\hat{\rho}_N$	$\hat{\rho}_{Nm}$	$\hat{\rho}_{\mu N}$	$\hat{\rho}_{\mu Nm}$	$\hat{\rho}_{\tau N}$	$\hat{\rho}_{\tau Nm}$
25	−2.78	−10.53	−3.68	−16.63	−4.66	−19.90
50	−2.59	−6.60	−3.16	−10.51	−3.89	−13.80
100	−2.52	−4.73	−2.87	−7.01	−3.31	−9.39
250	−2.42	−3.49	−2.63	−4.57	−2.86	−5.88
500	−2.37	−3.06	−2.51	−3.65	−2.67	−4.39
1000	−2.41	−2.80	−2.49	−3.19	−2.56	−3.63

5%

T	$\hat{\tau}_N$	$\hat{\tau}_{Nm}$	$\hat{\tau}_{\mu N}$	$\hat{\tau}_{\mu Nm}$	$\hat{\tau}_{\tau N}$	$\hat{\tau}_{\tau Nm}$
25	−1.93	−2.41	−2.69	−3.65	−3.49	−4.39
50	−1.84	−2.19	−2.36	−3.15	−2.92	−3.86
100	−1.81	−2.09	−2.14	−2.89	−2.52	−3.57
250	−1.72	−1.94	−1.92	−2.53	−2.18	−3.13
500	−1.69	−1.87	−1.82	−2.28	−1.99	−2.71
1000	−1.68	−1.81	−1.79	−2.09	−1.90	−2.43

1%

T	$\hat{\tau}_N$	$\hat{\tau}_{Nm}$	$\hat{\tau}_{\mu N}$	$\hat{\tau}_{\mu Nm}$	$\hat{\tau}_{\tau N}$	$\hat{\tau}_{\tau Nm}$
25	−2.74	−3.34	−3.59	−4.66	−4.47	−5.54
50	−2.60	−2.91	−3.15	−3.91	−3.81	−4.55
100	−2.54	−2.82	−2.89	−3.56	−3.29	−4.18
250	−2.43	−2.69	−2.64	−3.25	−2.86	−3.84
500	−2.38	−2.56	−2.51	−2.97	−2.67	−3.44
1000	−2.41	−2.51	−2.48	−2.81	−2.56	−3.15

4.3 Power of the Tests

The experimental results of the powers of the tests for 5% and 1% significance levels are given respectively in Tables 4a and 4b. Major findings are summarized as follows:

(1) The modified test statistics are more powerful than the unmodified ones. It indicates that the modification is quite effective.

(2) The proposed test statistics, both the unmodified and the modified, are less powerful than the Dickey-Fuller tests.

V. *Conclusion*

In this paper, we proposed a few test statistics for a unit root which are asymptotically normal, and reported their experimental sizes and powers. We found that, when the sample

TABLE 4a. POWER OF 5% LOWER TEST: CORRECTED SIZE

ρ	T	$\hat{\rho}$	$\hat{\rho}_N$	$\hat{\rho}_{Nm}$	$\hat{\rho}_\mu$	$\hat{\rho}_{\mu N}$	$\hat{\rho}_{\mu Nm}$	$\hat{\rho}_\tau$	$\hat{\rho}_{\tau N}$	$\hat{\rho}_{\tau Nm}$
0.95										
	25	0.09	0.08	0.08	0.08	0.06	0.06	0.04	0.04	0.05
	50	0.14	0.09	0.13	0.10	0.07	0.07	0.02	0.04	0.05
	100	0.32	0.12	0.25	0.19	0.10	0.13	0.01	0.06	0.06
	250	0.90	0.19	0.58	0.62	0.16	0.41	0.12	0.12	0.24
	500	1.00	0.30	0.89	0.99	0.27	0.81	0.89	0.24	0.70
	1000	1.00	0.47	1.00	1.00	0.44	0.99	1.00	0.41	0.98
0.90										
	25	0.15	0.11	0.12	0.11	0.08	0.06	0.03	0.04	0.06
	50	0.32	0.16	0.26	0.19	0.11	0.12	0.03	0.06	0.07
	100	0.76	0.24	0.60	0.46	0.19	0.34	0.10	0.13	0.16
	250	1.00	0.46	0.98	0.99	0.41	0.93	0.90	0.35	0.80
	500	1.00	0.72	1.00	1.00	0.68	1.00	1.00	0.64	1.00
	1000	1.00	0.93	1.00	1.00	0.92	1.00	1.00	0.91	1.00
0.80										
	25	0.34	0.21	0.24	0.19	0.13	0.09	0.06	0.07	0.07
	50	0.78	0.36	0.55	0.32	0.25	0.21	0.16	0.17	0.16
	100	1.00	0.59	0.98	0.87	0.51	0.82	0.71	0.41	0.62
	250	1.00	0.93	1.00	1.00	0.91	1.00	1.00	0.87	1.00
	500	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	1000	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

ρ	T	$\hat{\tau}$	$\hat{\tau}_N$	$\hat{\tau}_{Nm}$	$\hat{\tau}_\mu$	$\hat{\tau}_{\mu N}$	$\hat{\tau}_{\mu Nm}$	$\hat{\tau}_\tau$	$\hat{\tau}_{\tau N}$	$\hat{\tau}_{\tau Nm}$
0.95										
	25	0.08	0.07	0.07	0.06	0.06	0.05	0.06	0.07	0.05
	50	0.14	0.09	0.12	0.07	0.07	0.06	0.06	0.04	0.05
	100	0.31	0.12	0.23	0.12	0.09	0.12	0.15	0.06	0.06
	250	0.90	0.19	0.58	0.44	0.16	0.40	0.70	0.12	0.23
	500	1.00	0.30	0.90	0.97	0.27	0.81	1.00	0.23	0.70
	1000	1.00	0.47	1.00	1.00	0.44	0.99	1.00	0.41	0.98
0.90										
	25	0.14	0.10	0.10	0.07	0.08	0.06	0.07	0.05	0.05
	50	0.31	0.15	0.22	0.11	0.11	0.10	0.12	0.07	0.07
	100	0.76	0.23	0.57	0.31	0.18	0.30	0.35	0.13	0.16
	250	1.00	0.46	0.98	0.97	0.41	0.92	0.96	0.34	0.79
	500	1.00	0.71	1.00	1.00	0.68	1.00	1.00	0.64	1.00
	1000	1.00	0.93	1.00	1.00	0.92	1.00	1.00	0.91	1.00
0.80										
	25	0.32	0.20	0.15	0.11	0.12	0.06	0.11	0.07	0.06
	50	0.77	0.36	0.55	0.32	0.25	0.21	0.29	0.17	0.12
	100	1.00	0.59	0.98	0.87	0.51	0.82	0.79	0.41	0.55
	250	1.00	0.93	1.00	1.00	0.91	1.00	1.00	0.87	1.00
	500	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	1000	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

size is large, say, $T=500$ or larger, empirical sizes of the unmodified τ-type tests are close to the corresponding nominal ones. On the other hand, the ρ-type tests are strongly negatively biased, and we have to use the size-corrected critical values for tesitng. The modification for improved power was quite effective. However, the modification has its cost, that is, it brings even larger negative biases and the size-corrected critical values are also needed for testing. Generally, powers of the proposed tests, even those of the modified ones, are weaker than the those of the Dickey-Fuller type tests.

TABLE 4b. POWER OF 1% LOWER TEST: CORRECTED SIZE

ρ	T	$\hat{\rho}$	$\hat{\rho}_N$	$\hat{\rho}_{Nm}$	$\hat{\rho}_\mu$	$\hat{\rho}_{\mu N}$	$\hat{\rho}_{\mu Nm}$	$\hat{\rho}_\tau$	$\hat{\rho}_{\tau N}$	$\hat{\rho}_{\tau Nm}$
0.95										
	25	0.02	0.02	0.02	0:02	0.01	0.01	0.01	0.01	0.01
	50	0.03	0.02	0.03	0.02	0.02	0.02	0.00	0.01	0.01
	100	0.08	0.03	0.06	0.05	0.03	0.03	0.00	0.01	0.01
	250	0.51	0.06	0.27	0.25	0.05	0.15	0.01	0.03	0.06
	500	0.99	0.11	0.66	0.87	0.10	0.53	0.46	0.08	0.37
	1000	1.00	0.21	0.97	1.00	0.20	0.94	1.00	0.19	0.90
0.90										
	25	0.03	0.03	0.03	0.02	0.02	0.01	0.01	0.01	0.01
	50	0.08	0.05	0.07	0.05	0.03	0.03	0.00	0.01	0.02
	100	0.31	0.08	0.24	0.15	0.06	0.10	0.02	0.03	0.04
	250	0.99	0.22	0.87	0.88	0.18	0.70	0.52	0.15	0.45
	500	1.00	0.46	1.00	1.00	0.43	1.00	1.00	0.38	0.99
	1000	1.00	0.78	1.00	1.00	0.76	1.00	1.00	0.75	1.00
0.80										
	25	0.09	0.07	0.06	0.05	0.04	0.02	0.01	0.01	0.02
	50	0.34	0.16	0.26	0.16	0.10	0.08	0.03	0.04	0.04
	100	0.93	0.33	0.84	0.68	0.26	0.51	0.31	0.18	0.27
	250	1.00	0.78	1.00	1.00	0.74	1.00	1.00	0.68	1.00
	500	1.00	0.98	1.00	1.00	0.98	1.00	1.00	0.97	1.00
	1000	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

ρ	T	$\hat{\tau}$	$\hat{\tau}_N$	$\hat{\tau}_{Nm}$	$\hat{\tau}_\mu$	$\hat{\tau}_{\mu N}$	$\hat{\tau}_{\mu Nm}$	$\hat{\tau}_\tau$	$\hat{\tau}_{\tau N}$	$\hat{\tau}_{\tau Nm}$
0.95										
	25	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	50	0.03	0.02	0.02	0.01	0.02	0101	0.01	0.01	0.01
	100	0.07	0.03	0106	0.03	0.02	0.03	0.04	0.01	0.02
	250	0.49	0.06	0.24	0.14	0.05	0.16	0.36	0.03	0.06
	500	0.99	0.11	0.66	0.73	0]10	0.52	0.92	0.08	0.38
	1000	1.00	0.21	0.96	1.00	0.20	0.94	1.00	0.19	0.89
0.90										
	25	0.03	0.03	0.02	0.01	0.02	0.01	0.02	0.01	0.01
	50	0.08	0.04	0.05	0.03	0.03	0.02	0.03	0.01	0.02
	100	0.31	0.08	0.20	0.08	0.06	0.08	0.12	0.03	0.04
	250	0.99	0.21	0.84	0.74	0.18	0.67	0.77	0.14	0.44
	500	1.00	0.46	1.00	1.00	0.42	1.00	1.00	0.38	0.98
	1000	1.00	0.78	1.00	1.00	0.77	1.00	1.00	0.75	1.00
0.80										
	25	0.08	0.06	0.03	0.02	0.03	0.01	0.03	0.02	0.01
	50	0.33	0.14	0.18	0.09	0.08	0.05	0.09	0.04	0.03
	100	0.92	0.31	0.77	0.49	0.24	0.43	0.44	0.17	0.23
	250	1.00	0.78	1.00	1.00	0.73	1.00	1.00	0.68	1.00
	500	1.00	0.98	1.00	1.00	0.98	1.00	1.00	0.97	1.00
	1000	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

In sum, the ρ-type tests and the modified ones are inferior to the Dickey-Fuller tests since they need the size-corrected critical values for testing as in the Dickey-Fuller tests and they are less powerful. The τ-type tests in the large sample may be useful in practice, since they do not need the size-corrected critical values, although they are also less powerful than the Dickey-Fuller tests.

HITOTSUBASHI UNIVERSITY

Appendix A: Proof of the Theorem

(a) The following proof is a special case of the one in Yamamoto (1992a) except the proof of (b) below. We express the model (2.1) in the matrix notation as follows:

$$x_t = \beta' Y_{t-1} + \eta_t, \tag{A.1}$$

where $Y_{t-1} = [x_{t-1}, x_{t-2}, \ldots, x_{t-p-1}, 1, t]'$, and $\beta' = [\beta_1, \beta_2, \ldots, \beta_{p+1}, \gamma_0, \gamma_1]$. We now introduce the following non-singular transformation matrix D:

$$Z_{t-1} = DY_{t-1}, \tag{A.2}$$

where $Z_{t-1} = [z^1{}_{t-1}, z^1{}_{t-2}, \ldots, z^1{}_{t-p}, 1, z_t{}^3, t]'$, and

$$D = \begin{pmatrix} & & & & -\mu & 0 \\ & & & & -\mu & 0 \\ & & D^+ & & \cdot & \cdot \\ & & & & \cdot & \cdot \\ & & & & -\mu & \cdot \\ 0 & & \cdots\cdots\cdots & 0 & 1 & 0 \\ 1 & 0 & \cdots\cdots\cdots & 0 & \mu & -\mu \\ 0 & & \cdots\cdots\cdots & 0 & -1 & 1 \end{pmatrix},$$

and D^+ is the $p \times (p+1)$ matrix such that

$$D^+ = \begin{pmatrix} 1 & -1 & 0 & \cdots\cdots & & 0 \\ 0 & 1 & -1 & \cdots\cdots & & 0 \\ \cdot & & \cdot & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot & \cdot \\ 0 & \cdots\cdots & & 0 & 1 & -1 \end{pmatrix}.$$

The model (2.2) is rewritten in the matrix form as

$$\begin{aligned} x_t &= \beta' D^{-1} D Y_{t-1} + \eta_t \\ &= \delta' Z_{t-1} + \eta_t, \end{aligned} \tag{A.3}$$

where $\delta' = [\delta_{1,1}, \delta_{1,2}, \ldots, \delta_{1,p}, \delta_2, \delta_3, \delta_4]' = \beta' D^{-1}$.

We may note that the variables $z^1{}_{t-k}$'s are stationary of order $0_p(1)$, the constant term is of $0_p(1)$, $z_t{}^3$ is stochastic of order $0_p(\sqrt{T})$, and the trend variable t is of $0_p(T)$. Moreover, for the limit of the moment matrix, SSW showed that

$$\gamma^{-1}(\textstyle\sum_{t=p+2}^{T} Z_{t-1} Z'_{t-1})\gamma^{-1} \xrightarrow{P} V = \begin{pmatrix} V_{11} & 0 & 0 & 0 \\ 0 & V_{22} & V_{23} & V_{24} \\ 0 & V_{32} & V_{33} & V_{34} \\ 0 & V_{42} & V_{43} & V_{44} \end{pmatrix}, \tag{A.4}$$

where γ is the scaling matrix such that

$$\gamma=\begin{pmatrix} \sqrt{T}I_p & 0 & 0 & 0 \\ 0 & \sqrt{T} & 0 & 0 \\ 0 & 0 & T & 0 \\ 0 & 0 & 0 & T^{3/2} \end{pmatrix}.$$

Further, they showed that while V_{11} is the fixed constant, and V_{ij} $(i, j\geqq 2)$ are stochastic and are appropriate integrals of Wiener processes.

Let $\hat{\delta}_{1,k}(k=1, 2, \ldots, p)$, $\hat{\delta}_2$, $\hat{\delta}_3$, $\hat{\delta}_4$ be the least squares estimates of $\delta_{1,p}(k=1, 2, \ldots, p)$, δ_2, δ_3, and δ_4, respectively. SSW and Park and Phillips (1989) showed that $\hat{\delta}_1-\delta_1=0_p(1/\sqrt{T})$ where $\delta_1=[\delta_{1,1}, \delta_{1,2}, \ldots, \delta_{1,p}]'$, $\hat{\delta}_2-\delta_2=0_p(1/\sqrt{T})$, $\hat{\delta}_3-\delta_3=0_p(1/T)$, and $\hat{\delta}_4-\delta_4=0_p(T^{-3/2})$. In particular,

$$\sqrt{T}(\hat{\delta}_1-\delta_1)\xrightarrow{D} N(0, \sigma^2 V_{11}^{-1}). \tag{A.5}$$

Further, it is important to recall that $\sqrt{T}\hat{\delta}_3\xrightarrow{P}\delta_3=1$.

The null hypothesis H_0 is expressed as $R\beta=1$ in the matrix form, where R is the $p+3$ element vector such that $R=[1, 1, \ldots, 1, 0, 0, 0]$. We now consider the conventional t test statistic as follows:

$$t=(R\hat{\beta}-1)/SE(R\hat{\beta})=\sqrt{T}(R\hat{\beta}-1)/\sqrt{T}SE(R\hat{\beta}) \tag{A.6}$$

where $SE(R\hat{\beta})=[R\hat{\sigma}^2(\sum_{t=p+2}^{T}Y_{t-1}Y'_{t-1})^{-1}R']^{1/2}$, and where $\sigma^2=\Sigma(x_t-\hat{\beta}'Y_{t-1})^2/(T-p-1)$.

Since we have $\beta=D'\delta$, it is easily seen that

$$R\beta=\sum_{i=1}^{p}\beta_i=\delta_{1,p}+\delta_3.$$

Since $\sqrt{T}(\hat{\delta}_3-1)\to 0$, we have

$$\begin{aligned}\sqrt{T}(R\hat{\beta}-1)&=\sqrt{T}(\hat{\delta}_{1,p}+\hat{\delta}_3-1)\\ &\xrightarrow{P}\sqrt{T}\hat{\delta}_{1,p}\\ &\xrightarrow{D}N(0, \sigma^2(V_{11}^{-1})_{pp}),\end{aligned} \tag{A.7}$$

where $(V_{11}^{-1})_{pp}$ is the (p, p)-element of V_{11}^{-1}. On the other hand, since $Z_t=DY_t$, we have

$$\begin{aligned}\sqrt{T}SE(R\hat{\beta})&=\sqrt{T}[\hat{\sigma}^2RD'\gamma^{-1}\gamma(\textstyle\sum_{t=p+2}^{T}Z_{t-1}Z'_{t-1})^{-1}\gamma\gamma^{-1}DR']^{1/2}\\ &=\sqrt{T}[\hat{\sigma}^2RD'\gamma^{-1}(\gamma^{-1}\textstyle\sum_{t=p+2}^{T}Z_{t-1}Z'_{t-1}\gamma^{-1})^{-1}\gamma^{-1}DR']^{1/2}.\end{aligned}$$

By the structures of R and D, we have

$$RD'=[0, 0, \ldots, 0, -1, 0, 1, 0].$$

Noting that V is block diagonal and $\hat{\sigma}^2\xrightarrow{P}\hat{\sigma}^2$, we have

$$\begin{aligned}\sqrt{T}SE(R\hat{\beta}) &\xrightarrow{P} \sqrt{T}[\sigma^2 RD'\gamma^{-1}V^{-1}\gamma^{-1}DR']^{1/2}\\ &= \sqrt{T}[T^{-1}\sigma^2/(V_{11}^{-1})_{pp}+0_p(T^{-2})]^{1/2}\\ &\xrightarrow{P} \sqrt{T}[T^{-1}\sigma^2/(V_{11}^{-1})_{pp}]^{1/2}\\ &=[\sigma^2/(V_{11}^{-1})_{pp}]^{1/2}\end{aligned} \tag{A.8}$$

From (A.6), (A.7) and (A.8), the conclusion follows.

(b) Here, by definition, $\delta_{1,p}=-\beta_{p+1}=0$ and it is regarded as the redundant parameter in the p-th order stationary autoregressive process $\{z_t^1\}$. Then, it is easily verified that $(V_{11}^{-1})_{pp}=1/\sigma^2$. (See, for example, Appendix 7.5 of Box and Jenkins (1976).) Thus, from (A.7), we have the desired result:

$$\sqrt{T}(R\hat{\beta}-1)\xrightarrow{D} N(0, 1). \tag{A.9}$$

Appendix B: Derivation of the Normalizing Factor $3\sqrt{T}$

For illustration, we derive the asymptotic distribution of $\hat{\rho}_{\tau Nm}$ in the regression model (3.10). We can develop the argument as a special case of the one in Appendix A, by setting $p=1$. For the transformed expression (A.3), we have $Y_{t-1}=[x_{t-1}, x_{t-10}, 1, t]'$, $Z_{t-1}=[x_{t-1}-x_{t-10}, 1, x_{t-1}-\mu(t-1), t-1]'$, and

$$D=\begin{pmatrix}1 & -1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 1 & 0 & \mu & -\mu\\ 0 & 0 & -1 & 1\end{pmatrix}. \tag{B.1}$$

For the limit of the moment matrix V, V_{11} becomes scalar, and in particular we have

$$V_{11}=9\sigma^2. \tag{B.2}$$

Thus, from (A.5) and $\delta_{1,1}=0$, we have

$$\sqrt{T}(\hat{\delta}_{1,1}-\delta_{1,1})=\sqrt{T}\hat{\delta}_{1,1}\xrightarrow{D} N(0, 1/9). \tag{B.3}$$

Following the similar argument in Appendix A, we have $\hat{\rho}_{\tau Nm}=\hat{\delta}_{1,1}+\hat{\delta}_3$ and $\sqrt{T}(\hat{\delta}_3-1)\xrightarrow{P}0$.

$$\sqrt{T}(\hat{\rho}_{\tau Nm}-1)\xrightarrow{P} \sqrt{T}\hat{\delta}_{1,1}. \tag{B.4}$$

Consequently, we have

$$3\sqrt{T}(\hat{\rho}_{\tau Nm}-1)\xrightarrow{D} N(0, 1). \tag{B.5}$$

In the above (B.2) is essential for the normalizing factor $3\sqrt{T}$.

REFERENCES

Box, G.E.P. and Jenkins, G. (1976), *Time Series Analysis: Forecasting and Control* (2nd ed.), San Francisco, Holden-Day.

Choi, I. (1991), "Asymptotic normality of the least squares estimates for autoregressive integrated process with some applications," forthcoming in *Econometric Theory.*

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, pp. 427–431.

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

Hasza, D.P. and Fuller, W.A. (1979), "Estimation for autoregressive processes with unit roots," *Annals of Statistics* 7, pp. 1106–1120.

Park, J.Y. and Phillips, P.C.B. (1989), "Statistical inference in regressions with integrated processes: Part 2," *Econometric Theory* 5, pp. 95–131.

Sims, C.A., Stock, J.H. and Watson, M.W. (1990), "Inference in linear time series models with some unit roots," *Econometrica* 58, pp. 113–144.

Toda, H.Y. and Yamamoto, T. (1993), "Statistical inference in vector antoregressions with possibly integrated processes," forthcoming in *Journal of Econometrics.*

Yamamoto, T. (1992a), "A simple test of Granger causality in VAR models which may have some unit roots," Mimeographed paper, Hitotsubashi University.

Yamamoto, T. (1992b), "A simple approach to the statistical inference in linear time series models which may have some unit roots," Mimeographed paper, Hitotsubashi University.