

# Can Choice of Reference Density Improve Power of M-Estimation Based Unit Root Tests?

# Tapan Kar<sup>1</sup>, Malay Bhattacharyya<sup>2</sup> 💿

<sup>1</sup>Indian Institute of Management Nagpur, Maharashtra, India <sup>2</sup>Indian Institute of Management Kashipur, Uttarakhand, India Email: tapan@iimnagpur.ac.in, malay.bhattacharyya@iimkashipur.ac.in

How to cite this paper: Kar, T. and Bhattacharyya, M. (2022) Can Choice of Reference Density Improve Power of M-Estimation Based Unit Root Tests? *Journal of Mathematical Finance*, **12**, 340-355. https://doi.org/10.4236/jmf.2022.122019

**Received:** February 23, 2022 **Accepted:** May 9, 2022 **Published:** May 12, 2022

Copyright © 2022 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

http://creativecommons.org/licenses/by/4.0/

CC O Open Access

Abstract

In this paper, we investigate if the choice of the reference density could improve the power of M-estimation-based unit root tests. For this investigation, we consider models where the AR-coefficient is very close to one (local-to-unity) in the true data generating process. Motivated by the stylized facts that empirical return distributions have large skewness and high leptokurtosis, we explore if Johnson SU and Pearson Type IV distributions can be used as the reference densities to improve the power of the M-estimation based unit root tests. Through extensive simulations, we find that the proposed procedure, in finite samples, is as powerful as the Dickey-Fuller test for normal errors and is significantly more powerful than several existing tests for non-normal errors. We apply the proposed test to the Nelson and Plosser data set and to the nominal monthly interest rate of India.

# **Keywords**

Unit Root, M-Estimation, Local-to-Unity, Johnson SU Distribution, Reference Density

# **1. Introduction**

Following the seminal work of Dickey and Fuller [1], there has been immense interest in the statistical test for the unit root hypothesis in time series data. Stock [2], Phillips, and Xiao [3] have surveyed many such tests. Elliot, Rothenberg, and Stock [4] proposed a family of tests and also a modified Dickey-Fuller test for Gaussian time series with an unknown mean or a trend. Most of these tests are based on the OLS method assuming Gaussian errors. Monte Carlo procedure indicates that these tests do not have high statistical power when the errors show a heavy-tailed behaviour. Many scholars explored tests with higher power for non-Gaussian errors based on M-estimation. Cox and Llatas [5], Lucas [6], Rothenberg and Stock [7], Knight [8], Herce [9], Xiao [10], Thompson [11] and [12] and others have made significant contributions in the M-estimation domain.

They all assumed a known density function on the error process for the model to estimate the model parameters using the M-estimation procedure. Henceforth we shall call this assumed density function on the error process the "reference density".

For example, Lucas [6] investigated student t distribution. Herce [9] used the absolute value function (double exponential density). Hansen [13] proposed a unit root test using a covariate approach. However, identifying the covariates may be difficult in practice. Hasan and Koenker [14] proposed rank-based tests. Thompson [11] and [12] showed that for each rank-based test, there exists an M-estimation test with the same asymptotic power. Shin and So [15] used a nonparametric method to estimate the unknown error process. The method developed by them is difficult to implement. Moreover, for the Cauchy errors, their test has very poor power [11]). Koenker and Xiao [16] apply quantile regression to analyse the unit root process. In his monograph, Choi [17] has explained (page 97) that the LAD estimation-based test of Herce [9] has lower power than that of the quantile regression test. Using the approach proposed by Potscher and Prucha [18], Lima and Xiao [19] used a partially adaptive estimation method (PADF) to estimate the unknown error density. Hallin, Van Den Akker, and Werker [20] proposed a class of tests using the ranks of the samples. The rank-based test, however, requires an independently and identically distributed (IID) error process which may not be feasible in many practical applications.

Johnson SU (JHSU) and Pearson Type IV (PIV) distributions have been used to model financial time series data and in the risk management literature. Nagahara [21] used PIV density to model the stock return distribution. Bhattacharyya and Madhav [22] used the Johnson SU distribution and other methods to estimate the VaR for leptokurtic equity index returns. Bhattacharyya, Chaudhary, and Yadav [23] used PIV distribution to obtain the conditional VaR. Bhattacharyya, Misra, and Kodase [24] used PIV distribution to obtain the conditional MaxVaR.

Johnson SU distribution covers a wide range of shapes depending on its parameter values. It may also be a good approximation of Pearson type IV distribution [25]. The main advantage of this distribution is its ability to capture high kurtosis and skewness which is commonly observed in financial and economic time series data.

The main objective of this paper is to explore if the use of Johnson SU distribution (see Section 2) as a reference density (hereafter we call this test the JHSU test) could improve power. This is, perhaps, the first use of Johnson SU distribution in M-estimation literature. Using some results in Lucas [6], Xaio [10], and Thompson [11] [12], we obtain our test statistic and its asymptotic properties.

In our simulation studies, we generate data assuming various time series models such as AR (1), MA (1), etc. with error processes assumed as standard normal, t and chi-square distributions. Please note that, in practice, the data generating process, and therefore, the error process will be unknown. We use the M-estimation method to estimate the parameters of a selected model from the data thus generated. In the M-estimation method, we need to assume a probability density for the error process, unlike OLS estimation. This assumed density is called the "reference density". We explore Johnson SU and Pearson Type IV as the reference densities. These choices are especially explored to investigate if there is any improvement in the local power. We compare our choices with two other well-known tests such as ADF, PADF and test with reference density as t distribution with 3 degrees of freedom.

From the Monte Carlo simulations, we observe that the JHSU test, in finite samples, is as efficient as the augmented Dickey-Fuller (ADF) test for normal errors, and more powerful than many existing traditional tests for non-normal errors. The JHSU test is surprisingly easy to implement.

In Section 2, we sketch a brief outline of the Johnson SU distribution. Section 3 presents the time series model that we study in this paper and obtain the test statistic and its asymptotic distribution. The method of estimation of parameters, and the calculation of the test statistic, and the critical value have been explained in Section 4. Section 5 presents Monte Carlo simulation results. Section 6 describes the empirical studies, and we conclude in Section 7.

#### 2. Johnson SU Distribution

Johnson [25] proposed three transformations, *f*, of the following form:

$$Z = \gamma + \delta f\left(\frac{X-\xi}{\lambda}\right),$$

where Z is a standard normal variable and X is a continuous random variable whose distribution is unknown with shape parameters  $\gamma$  and  $\delta$ , scale parameter  $\lambda$  and location parameter  $\xi$ .  $\lambda > 0$  and  $\delta > 0$ .

$$f(y)$$
 may be chosen as  $\log(y)$ ,  $\sinh^{-1}(y)$  or  $\log\left(\frac{y}{1-y}\right)$ .

For the Johnson SU distribution, f is chosen as  $\sinh^{-1}$  so that

$$Z = \gamma + \delta \sinh^{-1} \left( \frac{X - \xi}{\lambda} \right), \quad -\infty < X < \infty$$
$$X = \xi + \lambda f^{-1} \left( \frac{Z - \gamma}{\delta} \right), \text{ where } f^{-1} \left( x \right) = \left( \frac{e^x - e^{-x}}{2} \right)$$

The parameters  $\gamma$  and  $\delta$  control skewness and kurtosis. The distribution is positively (negatively) skewed if  $\gamma$  is negative (positive). Increasing  $\delta$ , holding  $\gamma$  constant, reduces the kurtosis. Johnson SU distribution can capture a wide range of shapes depending on its parameter values.

The probability density function of Johnson SU distribution is given by

$$g(x) = \frac{\delta}{\lambda \sqrt{2\pi}} R\left(\frac{x-\zeta}{\lambda}\right) \exp\left\{-\frac{1}{2}\left[\gamma + \delta V\left(\frac{x-\zeta}{\lambda}\right)\right]^2\right\}, -\infty < x < \infty$$
(1)

The mean  $\mu = \xi - \lambda \theta^{1/2} \sinh(\Phi)$ , where  $\theta = e^{\delta^{-2}}$  and  $\Phi = \frac{\gamma}{\delta}$ . When mean is zero  $\xi = \lambda \theta^{1/2} \sinh(\Phi)$ .

### 3. The Model and Asymptotic Distribution of the Test Statistic

We assume the following data generating process (DGP) for our analysis.

$$y_{t} = a_{0} + a_{1}t + u_{t},$$
$$u_{t} - u_{t-1} = \pi u_{t-1} + v_{t}, \ F(L)v_{t} = \varepsilon_{t}$$

The errors  $\varepsilon_t$  are independently and identically distributed with expectation zero and finite variance  $\sigma_{\varepsilon}^2$  with  $u_0 = 0$ .

The term F(L) is the lag polynomial  $\left[1-F_1-F_2L^2-\dots-F_pL^p\right]$ . All the roots of F(L)=0 lie outside the unit circle. First, we re-write this model in the augmented Dicky and Fuller format, which is defined below

$$\Delta y_t = y_{t-1}\beta + \mu_0 + \mu_1 t + \sum_{i=1}^p \mu_{i+1} \Delta y_{t-i} + \varepsilon_t$$

where  $\Delta y_t = y_t - y_{t-1}$ ,  $\beta = \pi F(1)$ . For the DGP, the null and alternative hypotheses are

$$H_0: \pi = 0$$
 (*i.e.*,  $\beta = 0$ )  $H_1: \pi < 0$ 

Let  $\hat{\beta}$  be the M-estimator of  $\beta$  using Johnson SU as a reference density. Then  $\hat{\beta}$  will minimize the following objective function

$$\sum_{i=p+2}^{N} h \Big( \Delta y_{t} - y_{t-1} \beta - \mu_{0} - \mu_{1} t - \sum_{i=1}^{p} \mu_{i+1} \Delta y_{t-i} \Big),$$

where  $h = -\log(g)$  (g = Johnson SU density function).

For Johnson SU density, h(x) is given below

$$h(x) = \left(\frac{1}{2}\right)\log(G(x)) + \left(\frac{1}{2}\right)\left[\gamma + \delta\log\left[\frac{K(x)}{\lambda}\right]\right]^2 - \log(\delta) + \text{constant}$$
(2)

where K(x) and G(x) are given by

$$K(x) = \left[H(x) + \sqrt{G(x)}\right]$$
(3)

$$G(x) = \lambda^{2} + (x - \xi)^{2}$$
(4)

and

$$H(x) = (x - \xi) \tag{5}$$

Assumption 1. The function h(x) is continuously differentiable and its second and higher order derivatives are bounded.

We denote the first and the second derivatives of h(x) by  $\psi(x)$  and  $\varphi(x)$  respectively. Further, we define the following

$$\omega = E\left[\varphi(\varepsilon_t)\right], \ \rho = \operatorname{Corr}\left(\varepsilon_t, \psi(\varepsilon_t)\right), \ \sigma_{\psi}^2 = \operatorname{Var}\left[\psi(\varepsilon_t)\right] \text{ and } \ \vartheta = \frac{\sigma_{\varepsilon}\omega}{\sigma_{\psi}}$$

Following Thompson (2004) an approximate estimate of  $\hat{\beta}$  is

$$\hat{\beta} \approx \frac{\left[X^{t}P\right]\psi\left(\Delta Y - Z_{1}\tilde{M}\right)}{\omega\hat{\sigma}_{\psi}\left[X^{t}PX\right]^{1/2}}$$
(6)

where  $X = (y_{p+1}, y_{p+2}, \dots, y_{N-1})^t$ .

 $Z_1 \text{ is the matrix with row equal to } (1, t, \Delta y_{t-1}, \Delta y_{t-2}, \dots, \Delta y_{t-p}).$  $t = p + 2, p + 3, \dots, N.$ 

When time trend is not present (*i.e.*,  $a_1 = 0$ )  $Z_1$  is defined by  $(1, \Delta y_{t-1}, \Delta y_{t-2}, \dots, \Delta y_{t-p})$ .

 $\tilde{M} = (\tilde{\mu}_0, \tilde{\mu}_1, \tilde{\mu}_2, \cdots, \tilde{\mu}_{p+1})^t \text{ is a } (p+2) \text{ dimensional parameter vector that}$ minimizes the objective function  $\sum_{i=p+2}^N h \Big( \Delta y_i - \mu_0 - \mu_1 t - \sum_{i=1}^p \mu_{i+1} \Delta y_{t-i} \Big). \quad \hat{\sigma}_{\psi}$ is an estimate of  $\sigma_{\psi}$ .

 $\psi(x)$  is (N-p-1) dimensional column vector with components  $\psi(x_i)$ . For the case of Johnson SU distribution  $\psi(x_i)$  is defined in equation (2.3.15). *P* is the projection matrix defined by  $\left[I - Z_1((Z_1^t Z_1))^{-1} Z_1^t\right]$  where *I* is the Identity matrix. For Johnson SU as a reference density  $\psi(x)$  is defined below

$$\psi(x) = \frac{H(x)}{G(x)} + \left[\gamma + \delta \log\left(\frac{K(x)}{\lambda}\right)\right] \left(\frac{\delta}{\sqrt{G(x)}}\right)$$
(7)

K(x), G(x) and H(x) are given in Equations (3)-(5), respectively.

Following Thompson [26], we can remove  $\omega$  from the above Equation (6) since it does not affect the asymptotic power but, in small samples, can affect the size of the test because of poor estimation of  $\omega$ . After removing  $\omega$  test statistic in the *t*-ratio format is  $\hat{\beta} = \frac{\left[X^{t}P\right]\psi\left(\Delta Y - Z_{1}\tilde{M}\right)}{\hat{\sigma}_{\psi}\left[X^{t}PX\right]^{1/2}}$ . We reject null for small val-

ues of  $\hat{\beta}$ . Our next task is to find the asymptotic distribution of  $\hat{\beta}$ .

In this paper, we restrict our alternative hypothesis to the AR coefficient  $\pi$  being very close to 1. We set  $\pi = C/N$ , where *C* a constant, when making a limiting argument. *N* is the sample size. So, the parameter space is a shrinking neighbourhood of zero (see Chan and Wei [27] and Phillips [28]). In the presence of unit root, *C* is equal to zero, obviously.

Assumption 2  $E\left[\psi(\varepsilon_t)\right] = 0$ .

#### Asymptotic of the Test Statistic

The asymptotic distribution function of the test statistic is represented in terms of the function of Brownian motion. Let  $W_0$  be a standard Brownian motion defined on [0,1] and  $W_c(.)$  be a related diffusion process

$$W_{c}(t) = \int_{0}^{t} \exp(c(t-s)) dW_{0}(s)$$

that satisfies the stochastic differential equation

$$\mathrm{d}W_{c}\left(t\right)=cW_{c}\left(t\right)+\mathrm{d}W_{0}\left(t\right).$$

Let  $D_c$  be another process defined, for the intercept only model, by

$$D_{C}(r) = W_{C}(r) - \int_{0}^{1} W_{c}(s) ds$$

and for the model with time trend by

$$D_{C}(r) = W_{C}(r) - 2\int_{0}^{1} (2 - 3s - r(3 - 6s))W_{C}(s) ds$$

Thompson [12], under the assumptions (1) and (2), proved that the limiting distribution of  $\hat{\beta}$  converges weakly to the random variable  $\mathcal{F}_c$ , which is defined below

$$\mathcal{F}_{C} \equiv \rho \frac{T_{C}}{\sqrt{R_{C}}} + \sqrt{1 - \rho^{2}} \left( \frac{\int_{0}^{1} D_{C}(r) dW_{1}(r)}{\sqrt{R_{C}}} \right) + \mathcal{O}C\sqrt{R_{C}}$$

From the above, it is clear that, while  $\rho$  controls the null distribution (under null, C = 0), the power is determined by both  $\rho$  and  $\vartheta$ , because power of the test is  $\operatorname{Prob}\left[\mathcal{F}_{c} \leq q_{t}(\rho)\right]$ ,  $q_{t}(\rho)$  is obtained from  $\operatorname{Pr}\left[\mathcal{F}_{0} \leq q_{t}(\rho)\right] = \alpha$ , and  $\alpha$  is the size of the test. Thompson [12] argues that  $\vartheta$  dominates the power function relative to  $\rho$ .

As the value of  $\mathcal{G}$  increases, the asymptotic distribution shifts to the left, because *C* is negative when the alternative hypothesis is true. Since the alternate hypothesis is one-sided (left), the rejection zone is on the left tail of the distribution of the test statistic. Therefore, a left shift of the asymptotic distribution gives the source of power improvement.

**Table 1** gives the lists of the functional form of h(x) and  $\psi(x)$  for different reference densities.

### 4. Calculation of the Test Statistic and Critical Value

To compute the test statistic, we adopt the following steps.

1) Select lag length  $= k_{\text{maic}}$ , using MAIC criterion developed by Ng and Per-

ron [29] setting the maximum lag at  $\left[12\left(\frac{N}{100}\right)^{1/4}\right]$ .

<b>Table 1.</b> Expressions of $h(x)$ and $\psi(x)$	for different reference density functions
---	---

Reference density	h(x)	$\psi(x)$
Student <i>t</i> <sub>3</sub>	$\log(1+x^2/3)$	$(x/3)/(1+x^2/3)$
Pearson IV	$m\log\left\{1+\left[\left((x-\lambda)/a\right)^2\right]\right\}+\tan^{-1}\left[(x-\lambda)/a\right]$	$2m\left[\frac{x-\lambda}{a^2+x^2}\right] + \left[\frac{av}{a^2+x^2}\right]$
Normal	$x^{2}/2$	X

2) Run the following regression:

$$\Delta y_t^{\text{det}} = \pi^1 y_{t-1}^{\text{det}} + \sum_{p=1}^{k_{\text{maic}}} \Delta y_{t-p}^{\text{det}} + \varepsilon_t$$

 $y_t^{det}$  is the de-trended series according to Elliot, Rothenberg and Stock [4].

3) Estimate parameters of Johnson SU density from the estimated residuals of the above regression equation by maximum likelihood method with the condition  $\xi = \lambda \theta^{\overline{2}} \sinh(\Phi)$ . (Since mean of the true error process is zero) where  $\Phi$  and  $\theta$  are as in Section 2.

4) Find  $\hat{M}$  by minimizing  $h(\Delta Y - Z_1 M)$  and define  $\hat{\varepsilon} = \Delta Y - Z\hat{M}$ , where  $\hat{\varepsilon}$  is a  $(N - k_{\text{maic}} - 1)$  dimensional row vector, because the errors are estimated under the null hypothesis (see Hasan and Koenker [14]).

5) The estimates of the nuisance parameters  $\rho$  and  $\sigma_{\mu}$  are calculated as:

$$\hat{\sigma}_{\varepsilon}^{2} = \frac{\left(\hat{\varepsilon} - \overline{\hat{\varepsilon}}\right)^{t} \left(\hat{\varepsilon} - \overline{\hat{\varepsilon}}\right)}{N}, \quad \hat{\sigma}_{\psi}^{2} = \frac{\left(\psi\left(\hat{\varepsilon}\right) - \overline{\psi}\right)^{t} \left(\psi\left(\hat{\varepsilon}\right) - \overline{\psi}\right)}{N}, \\ \hat{\rho} = \left(\hat{\varepsilon} - \overline{\hat{\varepsilon}}\right)^{t} \left(\psi\left(\hat{\varepsilon}\right) - \overline{\psi}\right) / \left(N\hat{\sigma}_{\varepsilon}\hat{\sigma}_{\psi}\right)$$

where  $\overline{\hat{\varepsilon}}$  and  $\overline{\psi}$  is the sample means of  $\hat{\varepsilon}$  and  $\psi(\hat{\varepsilon})$  respectively.

6) Calculate test statistic  $\hat{\beta} = \frac{\left[X'P\right]\psi\left(\Delta Y - Z\hat{M}\right)}{\left[X'PX\right]^{1/2}}\hat{\sigma}_{\psi}^{-1}.$ 

7) Compute approximate  $\alpha$ % critical value for model *M* by the polynomial in  $(1-\rho)$  given below.

$$Q_{\alpha,M}(\rho) = A_{0,\alpha,M} + A_{1,\alpha,M}(1-\rho) + A_{2,\alpha,M}(1-\rho)^{2} + A_{3,\alpha,M}(1-\rho)^{3}$$

The coefficients of the polynomial  $Q_{\alpha,M}(\rho)$  are adapted from Thompson [12] (p. 368) and are reported in Table 2 for two different models, for ready reference.

8) Reject null hypothesis  $H_0: \pi = 0$  for model *M*, if  $\hat{\beta} < Q_{\alpha,M}(\hat{\rho})$  at  $\alpha$ % level of significance.

## 5. Monte Carlo Evidence

In this section, using Monte Carlo Simulation, we evaluate the small sample performance of the tests for sample size 100. For our simulations, we have considered

Table 2. Estimated coefficients\*.

α	$A_{_{0,lpha}}$	$A_{\mathrm{l},lpha}$	$A_{2,lpha}$	$A_{\scriptscriptstyle 3,lpha}$
5	-2.853	0.780	0.548	-0.118
5	-3.389	1.213	0.658	-0.126
1	-3.418	0.556	0.636	-0.098
1	-3.934	0.998	0.656	-0.044
	a 5 5 1 1	$ \begin{array}{c cccc} a & A_{0,\alpha} \\ \hline 5 & -2.853 \\ 5 & -3.389 \\ 1 & -3.418 \\ 1 & -3.934 \\ \end{array} $	$\alpha$ $A_{0,\alpha}$ $A_{1,\alpha}$ 5         -2.853         0.780           5         -3.389         1.213           1         -3.418         0.556           1         -3.934         0.998	$\alpha$ $A_{0,\alpha}$ $A_{1,\alpha}$ $A_{2,\alpha}$ 5         -2.853         0.780         0.548           5         -3.389         1.213         0.658           1         -3.418         0.556         0.636           1         -3.934         0.998         0.656

\*The suffix *M* is removed for notational simplicity. Source: Thompson [12] (p. 368).

two sets of values for  $(a_0, a_1)$ . For each such set, we have assumed three different error processes for  $v_t$ . Finally, we have assumed four different distributions for the errors  $\epsilon_t$ . Thus, a total of 24 different models have been used in the simulations. On all these 24 models, five different tests, based on different reference densities, have been investigated. All the calculations have been performed in R-studio.

Data have been generated according to the model defined below.

$$y_t = a_0 + a_1 t + u_t, \ u_t - u_{t-1} = \pi u_{t-1} + v_t$$

Two sets of values considered for  $(a_0, a_1)$  are (1,0) and (1,1). Three selected error processes for  $v_t$  are:

1) IID  $v_t = \epsilon_t$ 

2) AR (1): 
$$v_t = 0.5v_{t-1} + \varepsilon_t$$
, and

3) MA (1):  $v_t = \varepsilon_t + 0.5\varepsilon_{t-1}$ .

We set the initial condition  $\varepsilon_1 = 0$ .

The error process  $\varepsilon_t$  has been generated from the following four distributions.

1) Standard normal distribution.

2) Student *t* distribution with df = 3.

3) Lognormal distribution with mean centred at zero.

4) Chi-square distribution with df = 1 with mean centred at zero.

Five tests, with the following notations, have been used for all the 24 models described above.

1) ADF—Augmented Dicky-Fuller test.

2) T3—Student *t* distribution with df = 3 as the reference density.

3) PADF—Partially adaptive estimation method proposed by Lima and Xaio (2010).

4) PIV—Pearson Type IV distribution as reference density.

5) JHSU—Johnson SU distribution as reference density.

We have used the package *Urca* and the function *ur.df has* been used to perform the test setting the lag length at  $k_{\text{maic}}$ . JHSU test has been performed according to the steps from 1 - 4 described in Section 4. To estimate the parameters of the Johnson SU density by the maximum likelihood method, we use the *constroptim* function.

We have performed 1000 replications of each test for a sample size of 100. All the results have been reported at a 5% significance level. The numbers in the tables below represent the rejection ratios of the null hypothesis by different tests among 1000 replications. We have also investigated the ERS (Elliot *et al.* [4]) test and compared it with ours. We have found that the power of the JHSU test is better than that of the ERS test for the asymmetric error process. Hence, we have not reported the results of the ERS test.

**Tables 3-5** report the results for *only intercept* cases with error process  $v_t$  as IID, AR (1), and MA (1) respectively. **Tables 6-8** report the results when the *time trend* is included in the model and the error process  $v_t$  as IID, AR (1), and MA (1) respectively. The first column in each table represents the assumed distribution

	С	N	ADF	Т3	PADF	PIV	JHSU
	0	100	0.0490	0.0380	0.0390	0.0445	0.0445
N	$^{-1}$	100	0.0460	0.0630	0.0480	0.0510	0.0420
Normai	-5	100	0.0790	0.0940	0.0650	0.0680	0.0920
	-10	100	0.2050	0.1830	0.1660	0.2145	0.2220
	0	100	0.0450	0.0540	0.0450	0.0510	0.0485
,	-1	100	0.0380	0.0720	0.0720	0.0825	0.0685
$t_3$	-5	100	0.0730	0.2670	0.2760	0.2915	0.2935
	-10	100	0.2150	0.6400	0.6280	0.6480	0.6105
	0	100	0.0500	0.0590	0.0690	0.0680	0.0605
T NT 1	-1	100	0.0440	0.1240	0.1120	0.3160	0.4200
Log Normal	-5	100	0.0860	0.6670	0.6300	0.9805	0.9900
	-10	100	0.2370	0.9510	0.9250	1.0000	0.9965
	0	100	0.0640	0.0630	0.0650	0.0580	0.0615
01.	-1	100	0.0460	0.0850	0.0800	0.3815	0.7545
Cn1-square	-5	100	0.0780	0.5010	0.4860	0.9945	0.9875
	-10	100	0.2340	0.8550	0.8650	0.9995	0.9975

**Table 3.** Rejection ratios of the null hypothesis of different tests among 1000 replications with drift only model *i.e.*  $(a_0, a_1) = (1, 0)$  taking i.i.d error process, 5% significance level.

**Table 4.** Rejection ratios of the null hypothesis of different tests among 1000 replications with drift only model *i.e.*  $(a_0, a_1) = (1, 0)$  taking AR (1) error process, 5% significance level.

	С	N	ADF	Т3	PADF	PIV	JHSU
	0	100	0.0340	0.0370	0.0410	0.0400	0.0490
Normal	-1	100	0.0380	0.0340	0.0270	0.0460	0.0520
Normai	-5	100	0.0850	0.0810	0.0690	0.0930	0.0830
	-10	100	0.1950	0.1670	0.1610	0.2000	0.1920
	0	100	0.0480	0.0420	0.0380	0.0520	0.0410
4	-1	100	0.0470	0.0760	0.0770	0.0780	0.0810
l <sub>3</sub>	-5	100	0.0590	0.2540	0.2770	0.2930	0.2710
	-10	100	0.1930	0.5740	0.5640	0.5840	0.6000
	0	100	0.0540	0.0740	0.0560	0.0690	0.0530
Log Normal	-1	100	0.0430	0.1310	0.1090	0.3110	0.4070
Log Normai	-5	100	0.0710	0.6430	0.6030	0.9770	0.9750
	-10	100	0.2050	0.9200	0.9040	0.9990	0.9990
	0	100	0.0440	0.0770	0.0680	0.0560	0.0580
Chi-square	-1	100	0.0360	0.0880	0.0830	0.3800	0.6880
	-5	100	0.0580	0.4690	0.4520	0.9870	0.9980
	-10	100	0.1830	0.8320	0.8190	0.9980	1.0000

	С	N	ADF	Т3	PADF	PIV	JHSU
	0	100	0.0430	0.0350	0.0570	0.0430	0.0590
Normal	-1	100	0.0390	0.0430	0.0420	0.0370	0.0430
Normai	-5	100	0.0560	0.0570	0.0680	0.0770	0.0870
	-10	100	0.1380	0.1350	0.1320	0.1650	0.1700
	0	100	0.0470	0.0480	0.0410	0.0430	0.0610
4	-1	100	0.0470	0.0610	0.0680	0.0840	0.0660
$I_3$	-5	100	0.0530	0.2420	0.2650	0.2590	0.2590
	-10	100	0.1380	0.5500	0.5250	0.5270	0.5290
	0	100	0.0370	0.0480	0.0520	0.0560	0.0470
T NT	-1	100	0.0360	0.0980	0.0970	0.2500	0.3020
Log Normai	-5	100	0.0500	0.5500	0.5250	0.9550	0.9660
	-10	100	0.1570	0.5200	0.5250	0.9970	0.9990
	0	100	0.0520	0.0560	0.0520	0.0570	0.0570
Chiaman	-1	100	0.0320	0.0610	0.0630	0.2660	0.4220
Uni-square	-5	100	0.0500	0.4080	0.3620	0.9740	0.9860
	-10	100	0.1390	0.7530	0.7520	0.9980	1.0000

**Table 5.** Rejection ratios of the null hypothesis of different tests among 1000 replications with drift only model *i.e.*  $(a_0, a_1) = (1, 0)$  taking MA (1) error process, 5% significance level.

**Table 6.** Rejection ratios of the null hypothesis of different tests among 1000 replications with time trend model *i.e.*  $(a_0, a_1) = (1, 1)$  taking i.i.d error process significance level at 5%.

	С	N	ADF	T3	PADF	PIV	JHSU
	0	100	0.032	0.038	0.025	0.031	0.0395
NT	-1	100	0.036	0.04	0.03	0.05	0.035
Normai	-5	100	0.055	0.046	0.055	0.079	0.064
	-10	100	0.103	0.091	0.08	0.136	0.1105
	0	100	0.045	0.039	0.047	0.056	0.042
4	-1	100	0.045	0.046	0.058	0.06	0.059
l <sub>3</sub>	-5	100	0.061	0.205	0.186	0.221	0.218
	-10	100	0.098	0.468	0.449	0.495	0.483
	0	100	0.046	0.054	0.026	0.048	0.045
Log Normal	-1	100	0.045	0.106	0.107	0.226	0.3235
Log Normai	-5	100	0.063	0.556	0.512	0.894	0.9475
	-10	100	0.108	0.882	0.877	0.921	0.973
	0	100	0.043	0.039	0.043	0.057	0.0555
Chi aquara	-1	100	0.05	0.07	0.071	0.243	0.7295
Chi-square	-5	100	0.063	0.349	0.35	0.865	0.9185
	-10	100	0.114	0.767	0.738	0.852	0.9265

	С	N	ADF	Т3	PADF	PIV	JHSU
	0	100	0.0370	0.0320	0.0210	0.0390	0.0370
N	-1	100	0.0370	0.0360	0.0330	0.0440	0.0420
Normai	-5	100	0.0570	0.0510	0.0440	0.0550	0.0440
	-10	100	0.0880	0.1010	0.0600	0.1340	0.0990
	0	100	0.0230	0.0450	0.0450	0.0470	0.0460
4	-1	100	0.0400	0.0520	0.0500	0.0640	0.0560
$I_3$	-5	100	0.0540	0.1990	0.1930	0.1960	0.1980
	-10	100	0.1030	0.4480	0.4280	0.4200	0.4170
	0	100	0.0450	0.0390	0.0460	0.0950	0.0530
Log Normal	-1	100	0.0410	0.1010	0.0960	0.1610	0.2500
Log Normai	-5	100	0.0540	0.5320	0.4690	0.6300	0.8480
	-10	100	0.1110	0.8500	0.8220	0.6770	0.8790
	0	100	0.0400	0.0500	0.0400	0.1050	0.0510
Chi aguana	-1	100	0.0270	0.0680	0.0640	0.1980	0.2220
Cm-square	-5	100	0.0490	0.3350	0.3380	0.5470	0.7940
	-10	100	0.1200	0.7040	0.6860	0.5600	0.8580

**Table 7.** Rejection ratios of the null hypothesis of different tests among 1000 replications with time trend model *i.e.*  $(a_0, a_1) = (1,1)$  taking AR (1) error process significance level at 5%.

**Table 8.** Rejection ratios of the null hypothesis of different tests among 1000 replications with time trend model *i.e.*  $(a_0, a_1) = (1, 1)$  taking MA (1) error process significance level at 5%.

	С	N	ADF	Т3	PADF	PIV	JHSU
	0	100	0.0320	0.0410	0.0210	0.0300	0.0370
Normal	-1	100	0.0250	0.0410	0.0330	0.0360	0.0420
Normai	-5	100	0.0350	0.0470	0.0440	0.0470	0.0630
	-10	100	0.0680	0.0860	0.0600	0.0890	0.0940
	0	100	0.0310	0.0370	0.0330	0.0310	0.0470
4	-1	100	0.0330	0.0560	0.0480	0.0600	0.0420
$I_3$	-5	100	0.0510	0.1870	0.1630	0.1560	0.1620
	-10	100	0.0760	0.3860	0.3670	0.3340	0.3590
	0	100	0.0380	0.0440	0.0210	0.0870	0.0430
I NI 1	-1	100	0.0410	0.0780	0.0620	0.1610	0.1890
Log Normai	-5	100	0.0540	0.3860	0.4090	0.6030	0.8190
	-10	100	0.1110	0.7760	0.7500	0.6740	0.9190
	0	100	0.0290	0.0260	0.0350	0.0880	0.0510
Chiaman	-1	100	0.0350	0.0510	0.0520	0.1590	0.2220
Uni-square	-5	100	0.0360	0.2740	0.2390	0.4900	0.7940
	-10	100	0.0750	0.6010	0.5920	0.5630	0.8580

Journal of Mathematical Finance

of  $\varepsilon_i$ , the second gives the values of *C*, the third column shows the sample size, and the fourth to eighth columns report the rejection ratios when the tests 1 - 5 (described above) respectively applied to the DGP. The boldface numbers in the table indicate the highest rejection ratios. The R-code of the simulation studies can be available upon request.

The above results clearly suggest that the JHSU test, in respect of power, has a very good small sample performance. JHSU test is as powerful as the ADF test when the innovation process is Gaussian and has substantial gain in power when the errors are non-Gaussian. In terms of power JHSU test also has higher power than that of PIV, T3, and PADF tests for Lognormal and chi-square error distribution processes.

### 6. Empirical Evidence

We have considered two data sets for application. The first is the extended Nelson and Plosser (1982) data set and the second is the nominal monthly interest rate of India from January 2005 to March 2017. Nelson and Plosser extended data set are openly available at NPEXT: Nelson and Plosser extended data in urca: Unit Root and Cointegration Tests for Time Series Data (rdrr.io). The second data are collected from International Monetary Fund International Financial Statistics (IMF-IFS) database

(https://data.imf.org/?sk=4c514d48-b6ba-49ed-8ab9-52b0c1a0179b&sId=140915 1240976).

First, we consider the case of the Nelson and Plosser data set. Many researchers have used the data of Nelson and Plosser [30] to investigate whether macroeconomic time series are random walks or stationary processes around a level or a trend. As their data set is considered a testing ground for new procedures, we also implement our proposed method on Nelson and Plosser data. Lag length is

obtained by MAIC criterion setting maximum lag at  $\left[12\left(\frac{N}{100}\right)^{1/4}\right]$ , where N is

#### the sample size.

Table 9 reports the Jarque-Bera test statistic for all the series of Nelson and Plosser data sets. Table 10 contains the unit root analysis. We considered three tests, viz., ADF, ERS, and JHSU. We have taken the time trend model for our analysis.

We observe from **Table 10**, that the JHSU test rejects the hypothesis of unit root for "GNP per Capita" at 1% and for "Unemployment" at 5% level. For "Real GNP and "Unemployment", ERS test rejects only at 5% level but for no series at 1% level. For "GNP per Capita" ADF test rejects at 5% level.

Normality assumption is rejected at 5% significance level for the "GNP per capita series". JHSU test rejects the unit root hypothesis at 1% level and ADF test rejects it at 5% significance level. ERS test is not able to reject the null hypothesis. This clearly shows the power improvement in the power of the JHSU test.

From **Table 10**, we note that for the "Real GNP" series, ERS rejects the null but ADF cannot reject the null at 5% significance level. Also, for the "Real GNP"

series, the Normality assumption is rejected (See **Table 9**) at 5% significance level. JHSU test cannot reject the null for the "Real GNP" series. This supports the finding of the JHSU test for the Real GNP series.

Macroeconomic Variables	Jarque-Bera Statistic	<i>p</i> -value
Real GNP	6.6436	0.0360
Nominal GNP	5.4668	0.0650
GNP deflator	7.3095	0.0258
GNP per capita	6.5721	0.0374
Unemployment Rate	0.3426	0.8426
CPI	24.9690	0.0000
Nominal interest rate	60.5820	0.0000
Velocity	19.4730	0.0000
Employment	3.4307	0.1799
Industrial Production	6.7842	0.0336
Real Wages	9.0637	0.0107
Nominal Wages	5.5135	0.0635
Money Stock	4.2857	0.1173
Stock Prices	13.1330	0.0014

Table 9. Jarque-Bera statistic of Nelson Plosser data set.

#### Table 10. Unit root analysis of Nelson Plosser data set.

	N	Lags	JHSU	ADF	ERS
Real GNP	80	1	0.2099	-3.4545	-3.0400*
Nominal GNP	80	1	-2.0597	-2.0204	-1.9205
GNP deflator	100	1	-0.4035	-1.5902	-1.1553
GNP per capita	80	1	-3.969**	-3.5225*	-3.0149
Unemployment Rate	99	2	-3.2545*	-3.4031	-3.2594*
CPI	129	3	3.6686	-1.1985	-1.1247
Nominal interest rate	89	5	-0.7742	-1.1068	-1.2283
Velocity	120	1	7.0521	-1.5944	-0.9363
Employment	99	2	0.5982	-2.8744	-2.7711
Industrial Production	129	9	-2.6359	-2.3291	-1.7402
Real Wages	89	1	-1.69	-1.6841	-1.8075
Nominal Wages	89	2	-2.5657	0.5716	-1.8808
Money Stock	100	2	-0.6342	-2.507	-2.5815
Stock Prices	118	5	-1.5731	-1.1877	-0.8581

Note: \* 5%; \*\* 1%.

Series		N K	urtosis	Jarque	-Bera sta	tistic
Nominal interest	rate	147	4.3642		0.0000	
<b>Fable 12.</b> Unit root and	ılysis o	f interest rate o	f India.			
Series	Lags	Deterministic	component	t ADF	ERS	JHSU

Table 11. Descriptive statistics of interest rate of India.

JHSU and ERS tests give the same result (reject the null at 5% significance level) when the normality assumption has not been rejected, e.g., for the "Unemployment" series.

In our second study, we consider the case of the nominal monthly interest rate of India from January 2005 to March 2017. **Table 11** gives the descriptive statistics and **Table 12** reports the unit root analysis.

Here, we use the drift-only model. In this study, the series has excess kurtosis (**Table 11**) and also rejects the normality assumption. JHSU test rejects the unit root hypothesis at 5% significance level (**Table 12**) but other tests do not. Again, this shows that the JHSU test has higher power than that of others.

#### 7. Conclusions

In this paper, we have explored the unit root test based on Johnson SU distribution as a reference density. A step-by-step method for computing the test statistic is detailed. Monte Carlo evidence shows a significant power improvement over the ADF test when the innovations are non-Gaussian. The choice of JHSU is much better than other reference densities that are used in literature. JHSU test is very powerful for asymmetric error processes. JHSU test dominates the partially adaptive estimation method proposed by Lima and Xiao (2010). It also dominates Pearson Type IV and student  $t_3$  density-based tests when the error follows asymmetric distributions. For symmetric errors, the JHSU test performs 89as well as most other traditional procedures.

We have also obtained very satisfactory results when the proposed test procedure has been applied to real data sets. Therefore, the JHSU test can be a viable and much better option for the practitioners and researchers.

Apart from the application of this test procedure, future studies can be carried out on whether one should use a deterministic time trend model or drift while testing stationarity using the JHSU test.

### Acknowledgements

The authors wish to thank the editor and the referee for their valuable comments that helped improve the presentation of the paper.

# **Conflicts of Interest**

The authors declare no conflicts of interest regarding the publication of this paper.

#### References

- Dickey, D.A. and Fuller, W.A. (1979) Distribution of the Estimators for 5 Autoregressive Time Series with a Unit Root. *Journal of American Statistical Association*, 74, 427-431. <u>https://doi.org/10.1080/01621459.1979.10482531</u>
- [2] Stock, J.H. (1994) Unit Roots, Structural Breaks and Trends. Handbook of Econometrics, 4, 2739-2841. <u>https://doi.org/10.1016/S1573-4412(05)80015-7</u>
- [3] Phillips, P.C.B. and Xiao, Z. (1998) A Primer on Unit Root Testing. *Journal of Economic Surveys*, 12, 423-467. <u>https://doi.org/10.1111/1467-6419.00064</u>
- [4] Elliot, G., Rothenberg, T.J. and Stock, J.H. (1996) Efficient Tests for an Autoregressive Unit Root. *Econometrica*, 64, 813-836. <u>https://doi.org/10.2307/2171846</u>
- [5] Cox, D. and Llatas, I. (1991) Maximum Likelihood Type Estimation for Nearly Nonstationary Autoregressive Time Series. *Annals of Statistics*, 19, 1109-1128. https://doi.org/10.1214/aos/1176348240
- [6] Lucas, A. (1995) Unit Root Test Based on M Estimators. *Econometric Theory*, 11, 331-346. <u>https://doi.org/10.1017/S0266466600009191</u>
- [7] Rothenberg, T.J. and Stock, J.H. (1997) Inference in a Nearly Integrated Autoregressive Model with Nonnormal Innovations. *Journal of Econometrics*, 80, 269-286. https://doi.org/10.1016/S0304-4076(97)00040-7
- [8] Knight, K. (1991) Limit Theory for M Estimators in an Integrated Infinite Variance Process. *Econometric Theory*, 7, 200-212. https://doi.org/10.1017/S0266466600004400
- Herce, M.A. (1996) Asymptotic Theory of LAD Estimation in a Unit Root Process with Finite Variance Errors. *Econometric Theory*, 12, 129-153. https://doi.org/10.1017/S0266466600006472
- Xiao, Z. (2001) Likelihood Based Inference in Trending Time Series with a Root near Unity. *Econometric Theory*, 17, 1082-1112. https://doi.org/10.1017/S0266466601176036
- [11] Thompson, S.B. (2004) Optimal versus Robust Inference in Nearly Integrated Non-Gaussian Models. *Econometric Theory*, 20, 23-55. https://doi.org/10.1017/S0266466604201025
- Thompson, S.B. (2004) Robust Tests of the Unit Root Hypothesis Should Not Be "Modified". *Econometric Theory*, 20, 360-381. https://doi.org/10.1017/S0266466604202055
- [13] Hansen, B.E. (1995) Rethinking the Univariate Approach to Unit Root Testing, Using Covariates to Increase Power. *Econometric Theory*, 11, 1148-1171. https://doi.org/10.1017/S0266466600009993
- [14] Hasan, M.N. and Koenker, R.W. (1997) Robust Rank Test of Unit Root Hypothesis. *Econometrica*, 65, 133-161. <u>https://doi.org/10.2307/2171816</u>
- Shin, D.W. and So, B.S. (1999) Unit Root Test Based on Adaptive Maximum Likelihood Estimation. *Econometric Theory*, **15**, 1-23. https://doi.org/10.1017/S0266466699151016
- [16] Koenker, R. and Xiao, Z. (2004) Unit Root Quantile Autoregression Inference. *Journal of the American Statistical Association*, 99, 775-787.

https://doi.org/10.1198/016214504000001114

- [17] Choi, I. (2015) Almost All about Unit Roots. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9781316157824
- [18] Potscher, B.I.P. (1986) A Class of Partially Adaptive One-Step M Estimators for the Nonlinear Regression Model with Dependent Observations. *Journal of Econometrics*, **32**, 219-251. <u>https://doi.org/10.1016/0304-4076(86)90039-4</u>
- [19] Lima, L.R. and Xiao, Z. (2010) Testing Unit Root Based on Partially Adaptive Estimation. *Journal of Time Series Econometrics*, 2, 1-32. <u>https://doi.org/10.2202/1941-1928.1038</u>
- [20] Hallin, M., Van Den Akker, R. and Werker, B.J. (2011) A Class of Simple Distribution Free Rank-Based Unit Root Tests. *Journal of Econometrics*, 163, 200-214. https://doi.org/10.1016/j.jeconom.2011.03.007
- [21] Nagahara, Y. (1999) The PDF and CF of Pearson Type IV Distributions and the ML Estimation of the Parameters. *Statistics and Probability Letters*, **43**, 251-264. <u>https://doi.org/10.1016/S0167-7152(98)00265-X</u>
- [22] Bhattacharyya, M. and Siddarth, M. (2012) A Comparison of VaR Estimation Procedures for Leptokurtic Equity Index Returns. *Journal of Mathematical Finance*, 2, 13-30. https://doi.org/10.4236/jmf.2012.21002
- [23] Bhattacharyya, M., Chaudhary, A. and Yadav, G. (2008) Conditional VaR Estimation Using Pearson Type IV Distribution. *European Journal of Operational Research*, **191**, 386-397. <u>https://doi.org/10.1016/j.ejor.2007.07.021</u>
- [24] Bhattacharyya, M., Nityanand, M. and Bharat, K. (2009) MaxVaR for Non-Normal Heteroskedastic Returns. *Quantitative Finance*, 9, 925-935. https://doi.org/10.1080/14697680802595684
- [25] Johnson, N.L. (1949) Systems of Frequency Curves Generated by Method of Translation. *Biometrica*, 36, 149-176. <u>https://doi.org/10.1093/biomet/36.1-2.149</u>
- [26] Thompson, S.B. (2004) Robust Confidence Interval for Autoregressive Coefficients near One. In: Andrews, D. and Stock, J., Eds., *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, Cambridge University Press, Cambridge, 375-402. <u>https://doi.org/10.1017/CBO9780511614491.017</u>
- [27] Chan, N.Z. and Wei, C.Z. (1987) Asymptotic Inference for Nearly Nonstationary AR (1) Process. Annals of Statistics, 15, 1050-1063. https://doi.org/10.1214/aos/1176350492
- [28] Phillips, P.C.B. (1987) Towards a Unified Asymptotic Theory for Autoregression. *Biometrika*, 74, 535-547. https://doi.org/10.1093/biomet/74.3.535
- [29] Ng, S. and Perron, P. (2001) Lag Length Selection and Construction of Unit Root Tests with Good Size and Power. *Econometrica*, 69, 1519-1554. https://doi.org/10.1111/1468-0262.00256
- [30] Nelson, C.R. and Plosser, C.R. (1982) Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implication. *Journal of Monetary Economics*, 10, 139-162. <u>https://doi.org/10.1016/0304-3932(82)90012-5</u>