Title
Local Power Functions of Tests for Double Unit Roots

Authors
Haldrup, Niels, Prof.
Lildholdt, Peter

Publication Date
2000-06-05

2000-12

<u>UNIVERSITY OF CALIFORNIA, SAN DIEGO</u>

DEPARTMENT OF ECONOMICS

LOCAL POWER FUNCTIONS OF TESTS FOR DOUBLE UNIT ROOTS

BY

NIELS HALDRUP

AND

PETER LILDHOLDT

DISCUSSION PAPER 2000-12
JUNE 2000

Local power functions of tests for double unit roots

Niels Haldrup & Peter Lildholdt
Department of Economics, University of Aarhus, and
Centre for Dynamic Modelling in Economics*

5. June, 2000

Abstract. The purpose of this paper is to characterize three commonly used double unit root tests in terms of their asymptotic local power. To this end, we study a class of nearly doubly integrated processes which in the limit will behave as a weighted integral of a double indexed Ornstein-Uhlenbeck process. Based on a numerical examination of the analytical distributions, a comparison of the tests is made via their asymptotic local power functions.

Keywords: Asymptotic local power function, Brownian motion, Ornstein-Uhlenbeck process.
JEL Classification: C12, C14, C22.

1. Introduction

Most economic time series have properties that mimic those characterizing unit root (integrated) processes. A characterization in terms of integration of order one, I(1), seems appropriate for the majority of the series, however, some variables like prices, wages, money balances[1], stock-variables *etc.*, appear to be smoother than normally observed for variables integrated of order one; such series are potentially integrated of order 2 whereby double differencing is needed to render the series stationary. We will refer to such series as having double unit roots. By now there is a growing literature focusing on the complications implied by double unit roots. This literature is not only concerned with univariate testing for I(2), (Hasza and Fuller (1979), Dickey and Pantula (1987), Sen and Dickey (1987), Shin and Kim (1999), and Haldrup (1994a)), but it also focuses on the rather complex dynamic interactions occurring in I(2) cointegrated models (compare Johansen (1995, 1997), Kitamura (1995), Choi,

*This paper was completed while the first author was visiting University of California, San Diego, in the spring of 2000. The Economics Department at UCSD is gratefully acknowledged for its hospitality. We also want to thank Søren Johansen for useful comments. Financial support from the Danish Social Sciences Research Council (SSF) and the Aarhus University Research Foundation is gratefully acknowledged. The usual disclaimer applies.

[1]For instance King *et. al.* (1991) find that both money and prices for the US can be described as I(2) processes.

Park, and Yu (1997), and Haldrup (1994b)). In Haldrup (1998) recent advances in the theoretical and empirical literature on I(2) are reviewed.

In the present paper, our attention is directed towards univariate testing for the order of integration, and the purpose is to compare the asymptotic local power functions of three commonly used unit root tests designed to test the null of I(2). The tests are: 1) the sequential Dickey-Pantula test, (which relies on prior differencing before testing for I(2), 2) the Hasza-Fuller test, (which is a joint test for the number of unit roots), and 3) the Sen-Dickey test, (which is a symmetric version of the Hasza-Fuller test). These tests have been compared in a number of settings via Monte Carlo simulation in terms of their power against a fixed alternative for a given sample size. However, as far as we know of, no systematical comparison has been conducted of the tests in terms of their asymptotic local power properties. In order to do so we define a class of nearly doubly integrated processes which extends similar analyses conducted for I(1) tests by Chan and Wei (1987) and Phillips (1987a) *inter alia.* Nearly doubly integrated processes have also been used in a different context by Nabeya and Perron (1994) and Perron and Ng (1996, 1998). It occurs that the limiting distributions can be described in terms of functionals of a twice indexed weighted integral of an Ornstein-Uhlenbeck process where the non-centrality parameters measure the (local) distance from the exact I(2) case. Consequently, the asymptotic distributions can be evaluated numerically by varying these parameters. This can serve different purposes; firstly, the simulated distributions may provide an approximation to the exact finite sample distributions when roots are close to but not exactly one, and secondly, the distributions can describe the asymptotic local power functions in terms of the two non-centrality parameters. The numerical results indicate that the superiority of the various tests to a large extent will depend upon the region where the non-centrality parameters are located. The symmetric test of Sen and Dickey is found to be superior against locally stationary alternatives, whereas the Dickey Pantula test has relative high powers against explosive alternatives. Also, the Dickey Pantula test performs rather well when one exact unit root is present. As a by product of the theoretical analysis we evaluate the adequacy of the analytical expressions to approximate the exact finite sample distributions of the various tests when a local discrepancy from the precise I(2) case applies.

The paper proceeds as follows. In section 2 the three double unit root tests scrutinized in section 4 are presented. Section 3 defines the class of nearly doubly integrated processes and the analytical distributions of the three test statistics are reported as a function of the non-centrality parameters. A numerical characterization of the local asymptotic power functions is subsequently provided in section 4 and in section 5 we compare exact finite sample distributions with some of the distributions that follow from the local to unity asymptotics. The final section concludes. All proofs are reported in a technical appendix.

2. Test statistics for double unit roots

In this section three commonly used procedures in testing for I(2) are briefly described. The main purpose is to present the statistics with the aim of comparing their asymptotic local power functions in section 3. All three tests can be represented in both parametric and semiparametric versions in order to remove the influence from possible nuisance parameters which otherwise would be harmful with respect to the limit theory. Since both the parametric and semiparametric tests have the same limiting distribution under the null of double unit roots, it appears that we can consider the simplifying assumption where the error term u_t is i.i.d in the data generating mechanism

$$(1-\alpha_1 L)(1-\alpha_2 L)y_t = u_t \qquad t = 1, 2, ..., T \tag{1}$$

This assumption will have no quantitative or qualitative implications with respect to the analytical findings of the paper but it will greatly simplify the analytical derivations. With no loss of generality, we also assume the initial conditions to be fixed; in particular, we let $y_0 = y_{-1} = 0$ to simplify the notation. To make the points clear, we will abstract from the presence of deterministic components in the regression models although we realize the importance of this aspect in unit root testing.

2.1. The Dickey-Pantula test. The first test is due to Dickey and Pantula (1987) who suggest a sequence of testing where initially I(2) is tested against I(1). This can be accomplished by prior differencing of the time series and then conducting a standard Dickey-Fuller test for an additional unit root. Hence the auxiliary regression reads (after imposing a single unit root, $\alpha_1 \equiv 1$) :

$$\Delta^2 y_t = (\widehat{a} - 1)\Delta y_{t-1} + \widehat{w}_t \tag{2}$$

where the t-ratio associated with the regressor Δy_{t-1} is used to test $\mathrm{H}_0 : a = 1$. This is nothing else than a standard I(1) problem for the first differenced data, and the test statistic follows the Dickey-Fuller distribution. Observe that under the null of double unit roots we have that $\alpha_1 = \alpha_2 = 1$ in accordance with (1) and $w_t = u_t$. When the null hypothesis is rejected, I(1) is subsequently tested against I(0) using standard I(1) tests. In the presence of non i.i.d. errors the augmented Dickey-Fuller, (Dickey and Fuller (1979)), or the Phillips (1987b) and Phillips-Perron (1988) tests can be used under rather general conditions[2].

2.2. The Hasza-Fuller test. An alternative way of writing the data generating mechanism (1) is $\Delta^2 y_t = (\alpha - 1)y_{t-1} + (\gamma - 1)\Delta y_{t-1} + u_t$ where $(\alpha - 1) = \alpha_1 + \alpha_2 -$

[2]The general conditions concerning u_t which are permitted in the semi-parametric class of tests can be found in e.g. Phillips (1987b), Assumption 2.1.

$\alpha_1\alpha_2 - 1$ and $(\gamma - 1) = \alpha_1\alpha_2 - 1$. Hence, a joint test of the hypothesis $\mathrm{H}_0 : \alpha_1 = \alpha_2 = 1$ (corresponding to $\alpha = \gamma = 1$) can be constructed from the auxiliary regression

$$\Delta^2 y_t = (\widehat{\alpha} - 1)y_{t-1} + (\widehat{\gamma} - 1)\Delta y_{t-1} + \widehat{u}_t. \tag{3}$$

Hasza and Fuller (1979) were the first to suggest an F-test for the double unit root hypothesis. Observe that since the F-test is two-sided, the alternative hypothesis is quite general as it includes situations where x_t is either explosive, I(0), or I(1). In the paper by Hasza and Fuller the non-standard limiting distribution of the F-test statistic is derived. By using a slightly different notation than in their paper, the distribution can be shown to read, see Haldrup (1994a):

$$F_{\alpha=\gamma=1} \Rightarrow \frac{1}{2}\int_0^1 G(r)'dW(\int_0^1 G(r)G(r)'dr)^{-1}\int_0^1 G(r)dW \tag{4}$$

where $G(r) = (W(r), \overline{W}(r))'$, and $W(r)$ is a standard Brownian motion on $C[0,1]$, *i.e.* the space of continuous functions on the unit interval, and $\overline{W}(r) = \int_0^r W(s)ds$. Empirical fractiles are reported in Hasza and Fuller's paper[3].

When u_t is not i.i.d. then under rather general conditions the Hasza-Fuller regression (3) can be augmented with lags of $\Delta^2 y_t$ to whiten the errors. As an alternative test, one of the authors, Haldrup (1994a) has constructed the semiparametric equivalent of this test.

2.3. The Sen and Dickey symmetric test. A different class of (non-sequential) tests is the so-called symmetric tests which, in an I(2) setting, were initially suggested by Sen (1986) and Sen and Dickey (1987). The test is a symmetric version of Hasza and Fuller's (1979) joint F-test. The motivation arises from the interesting property that if the stationary difference equation defining the time series is given by

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + u_t \tag{5}$$

where u_t is white noise with variance σ_u^2, then the series with the same difference equation equal to

$$y_t = \phi_1 y_{t+1} + \phi_2 y_{t+2} + ... + \phi_p y_{t+p} + v_t \tag{6}$$

will also have white noise errors v_t with the same error variance as for u_t, $\sigma_v^2 = \sigma_u^2$, see Fuller (1976). The basic idea is thus to jointly estimate (5) and (6), in a symmetric fashion, and use it to test for double unit roots.

[3]Note that the distribution result (4) also will apply for the asymptotically equivalent likelihood ratio test of the joint hypthesis. Hence we do not consider the likelihood ratio test in the present context.

Assuming y_t to be an AR(2) process, the symmetrized version of the Hasza-Fuller regression model (3) is given by the pair of regression equations

$$\Delta^2 y_t = (\hat{\alpha}-1)y_{t-1} + (\hat{\gamma}-1)\Delta y_{t-1} + \widehat{u}_t,\ t = 3, ..., T \quad (7)$$
$$\Delta^2 y_t = (\hat{\alpha}-1)y_{t-1} - (\hat{\gamma}-1)\Delta y_t + \widehat{v}_{t-2},\ t = 3, ..., T \quad (8)$$

where certain cross-equation restrictions on the parameters are seen to apply, and hence intuition suggests that under stationary alternatives this will provide more efficient estimates of the parameters of interest, α and γ.

From this 'extended' Hasza-Fuller regression the F-test of the hypothesis $\mathrm{H}_0 : \alpha = \gamma = 1$ can be constructed. Due to the symmetry of the regression model, the limiting distribution of the F-test statistic becomes somewhat simplified as cross terms appear to cancel. The distribution is still non-standard and reads

$$F^{Sym}_{\alpha=\gamma=1} \Rightarrow (\int_0^1 \overline{W}^2(r)\mathrm{d}r)^{-1}(\int_0^1 \overline{W}(r)\mathrm{d}W)^2 + \frac{1}{4}(\int_0^1 W^2(r)\mathrm{d}r)^{-1} \quad (9)$$

The empirical distributions reported in Sen and Dickey's article only consider the situation where an intercept and a possible time trend are included in the regression. Since we have deliberately decided to abstract from deterministic components in the present paper, the critical values for this case have been simulated, see Table 1.

Again the above test can be generalized to the case with non i.i.d. errors through augmentating lags of the second differenced series; which lags to include is non-trivial in this case, see Sen and Dickey's article for details. As an alternative test, Shin and Kim (1999) have suggested a semiparametric analogue of the symmetric test.

Insert Table 1 about here

3. Asymptotic Distributions of I(2) tests under local alternatives

In Bobkoski (1983), Phillips (1987a), Chan (1988), and Chan and Wei (1987), a class of autoregressive models with a root *local* to unity has been introduced. In particular, they consider the model

$$y_t = \exp(c/T)y_{t-1} + u_t$$

where u_t can satisfy rather general requirements. The model is quite general: The unit root model is encompassed by letting $c = 0$, mildly explosive processes occur when $c > 0$, and nearly integrated processes show up for $c < 0$. By letting $c \to \pm\infty$ (after letting $T \to \infty$) the stationary or explosive region is reached although only a heuristic description can be given in this case, see Phillips (1987a).

Here we study the double unit root tests presented in the previous section by assuming *nearly doubly integrated* processes. This class of models has previously been considered in a different context by Nabeya and Perron (1994), and Perron and Ng (1996, 1998). See also[4] Jeganathan (1991) and Swensen[5] (1993). The data generating mechanism is given by

$$y_t = \exp(c_1/T)y_{t-1} + v_t \tag{10}$$
$$v_t = \exp(c_2/T)v_{t-1} + u_t \tag{11}$$

where for simplicity u_t is assumed to be i.i.d. $(0,\sigma_u^2)$ and again we let $y_{-1} = y_0 = 0$. Alternatively, the process can be represented as

$$y_t = (\exp(c_1/T) + \exp(c_2/T))\, y_{t-1} - \exp((c_1 + c_2)/T)y_{t-2} + u_t \tag{12}$$

As $T \to \infty$, y_t is seen to have two unit roots hence suggesting the terminology of y_t being "nearly doubly integrated". By varying c_1 and c_2 a wide range of different processes can be examined including mildly explosive processes, single and double unit root processes, and nearly (doubly) integrated processes.

A few remarks on the limiting behavior of y_t will provide some insight. It can be shown, see Bobkoski (1983), Phillips (1987a), and Nabeya and Perron (1994), that as $T \to \infty$,

$$T^{-1/2}v_t \Rightarrow \sigma_u \int_0^r \exp((r-s)c_1)dW(s) \equiv \sigma_u J_{c_1}(r) \tag{13}$$
$$T^{-3/2}y_t \Rightarrow \sigma_u \int_0^r \exp((r-v)c_1)J_{c_2}(v)dv \equiv \sigma_u Q_{c_1}(J_{c_2}(r))$$

$J_c(r)$ is known as an Ornstein-Uhlenbeck process and corresponds to a Brownian motion for $c = 0$, i.e. $J_{c=0}(r) \equiv W(r)$. The expression $Q_{c_1}(J_{c_2}(r))$ is a double indexed weighted integral of an Ornstein-Uhlenbeck process, and hence, for $c_1 = c_2 = 0$ it is seen that $Q_{c_1=0}(J_{c_2=0}(r)) \equiv \int_0^r W(s)ds \equiv \overline{W}(r)$ as previously defined.

First we study the properties of the Dickey-Pantula test which relies on the Dickey-Fuller t-statistic from the regression (2). It occurs that the equation (12), can be rewritten in a different form by exploiting the series expansion

$$\exp\left(\frac{c}{T}\right) = 1 + \frac{c}{T} + R_T \tag{14}$$

The remainder R_T is of the order $O(T^{-2})$ which appears to vanish from the asymptotic expressions, and hence, in the sequel the approximation $\exp\left(\frac{c}{T}\right) \approx 1 + \frac{c}{T}$ will show

[4] Jeganathan (1991) presents a general approach to the asymptotic behaviour of least squares estimators in AR time series having roots close to the unit circle.

[5] Using contiguity arguments, Swensen (1993) shows how the local power function for the likelihood ratio test of the double unit root null can be expressed in terms of a Radon-Nikodym derivative of an Ito process with respect to a Brownian motion.

useful. It is shown in the appendix that the autoregressive model (12) is equivalent to writing

$$\Delta^2 y_t = \left(\frac{c_1+c_2}{T} + \frac{c_1 c_2}{T^2}\right)\Delta y_{t-1} - \frac{c_1 c_2}{T^2} y_{t-1} + u_t + O_p(T^{-3/2}) \qquad (15)$$

and hence, in terms of the notation (2)

$$a - 1 = \left(\frac{c_1+c_2}{T} + \frac{c_1 c_2}{T^2}\right)$$

and

$$\omega_t = -\frac{c_1 c_2}{T^2} y_{t-1} + u_t + O_p(T^{-3/2})$$

The following theorem holds:

Theorem 1. *Assume that y_t is nearly doubly integrated according to (12) and consider the Dickey-Pantula t-test of $H_0 : a = 1$ based on the regression (2). Then as $T \to \infty$:*

(a) $s_w^2 = \frac{1}{T}\sum_{t=2}^{T} \widehat{\omega}_t^2 \xrightarrow{p} \sigma_u^2$
(b) $T(\widehat{a} - 1) \Rightarrow c_1 + c_2 + \Psi_2^{-1}\left\{\Psi_1 - \frac{c_1 c_2}{2} Q_{c_1}(J_{c_2}(1))^2\right\}$
(c) $t_{a-1} \Rightarrow (c_1 + c_2)\,\Psi_2^{1/2} + \Psi_2^{-1/2}\left\{\Psi_1 - \frac{c_1 c_2}{2} Q_{c_1}(J_{c_2}(1))^2\right\}$

In the above expressions $\Psi_1 = \left\{c_1 \int_0^1 Q_{c_1}(J_{c_2}(r))dW(r) + \int_0^1 J_{c_2}(r)dW(r)\right\}$ *and* $\Psi_2 = \left\{c_1 Q_{c_1}(J_{c_2}(1))^2 - c_1^2 \int_0^1 Q_{c_1}(J_{c_2}(r))^2 dr + \int_0^1 J_{c_2}(r)^2 dr\right\}$.

Proof: See appendix.

Observe that the limiting results for t_{a-1} will mimic the usual Dickey-Fuller distribution when $c_1 = c_2 = 0$.

Next we shall focus our attention on the joint Hasza-Fuller F-test which is based on the regression (3). With the data generating mechanism (15) parameters are given as $\alpha - 1 = -c_1 c_2/T^2$ and $\gamma - 1 = (c_1 + c_2)/T + c_1 c_2/T^2$. If we write the regression equation more compactly as $\Delta^2 y_t = \mathbf{x}_{t-1}\widehat{\boldsymbol{\beta}} + \widehat{u}_t$ with $\widehat{\boldsymbol{\beta}}' = ((\hat{\alpha} - 1), (\hat{\gamma} - 1))$ and $\mathbf{x}_t = (y_t, \Delta y_t)'$, the $F-$statistic for the null hypothesis $H_0 : \alpha = \gamma = 1$ is defined as

$$F_{\alpha=\gamma=1} = \frac{\widehat{\boldsymbol{\beta}}'(\sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}_t')\widehat{\boldsymbol{\beta}}}{2 s_u^2}$$

The limiting distributions are given as follows:

Theorem 2. *Assume that y_t is nearly doubly integrated according to (12). Based upon the Hasza-Fuller regression (3) the following holds as $T \to \infty$:*

$$(a) \quad s_u^2 = \frac{1}{T}\sum_{t=2}^{T} \widehat{u}_t^2 \xrightarrow{p} \sigma_u^2$$

$$(b) \quad \begin{pmatrix} T^2(\widehat{\alpha}-1) \\ T(\widehat{\gamma}-1) \end{pmatrix} \Rightarrow k_{c_1c_2} + M^{-1}N$$

$$(c) \quad F_{\alpha=\gamma=1} \Rightarrow \frac{1}{2}\left\{k'_{c_1c_2}N + k'_{c_1c_2}Mk_{c_1c_2} + N'k_{c_1c_2} + N'M^{-1}N\right\}.$$

The following symbols have been used:

$$M = \begin{pmatrix} \int_0^1 Q_{c_1}(J_{c_2}(r))^2 dr & \frac{1}{2}Q_{c_1}(J_{c_2}(1))^2 \\ \frac{1}{2}Q_{c_1}(J_{c_2}(1))^2 & \Psi_2 \end{pmatrix}$$

$$N = \begin{pmatrix} \int_0^1 Q_{c_1}(J_{c_2}(r))dW(r) \\ \Psi_1 \end{pmatrix}, \text{ and } k_{c_1c_2} = \begin{pmatrix} -c_1c_2 \\ c_1+c_2 \end{pmatrix}.$$

Proof: See appendix.

Again the distributions simplify when $c_1 = c_2 = 0$; for instance, the Hasza and Fuller (1979) distribution reported in (4) follows as a special case of Theorem $2c$ as can be easily verified.

Finally, we focus our attention on the Sen-Dickey symmetric version of the Hasza-Fuller test. This is based on joint estimation of the equations (7) and (8) and construction of the $F-$test statistic for the null hypothesis $\mathrm{H}_0 : \alpha = \gamma = 1$. Likewise the Hasza-Fuller test, $\alpha - 1 = -c_1c_2/T^2$ and $\gamma - 1 = (c_1+c_2)/T + c_1c_2/T^2$. To construct the statistic we need to define the matrix of regressors $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)$ and the regressand vector $\mathbf{Y}$ where $\mathbf{X}_1 = (y_2, ..., y_{T-1}, y_{T-1}, ..., y_2)'$, $\mathbf{X}_2 = (\Delta y_{2,...,}\Delta y_{T-1}, -\Delta y_T, ..., -\Delta y_3)'$, and $\mathbf{Y} = (\Delta^2 y_3, ..., \Delta^2 y_T, \Delta^2 y_T, ..., \Delta^2 y_3)'$. Using this notation, $\mathbf{Y} = \mathbf{X}\widetilde{\boldsymbol{\beta}} + \widetilde{\mathbf{u}}$ with $\boldsymbol{\beta}' = (\alpha - 1, \gamma - 1)$ and where the errors are given by $\mathbf{u} = (u_3, ...u_T, v_{T-2}, ...v_1)'$. The $F-$statistic reads

$$F_{\alpha=\gamma=1}^{Sym} = \frac{\widetilde{\boldsymbol{\beta}}'(\mathbf{X}'\mathbf{X})\widetilde{\boldsymbol{\beta}}}{2\widetilde{\sigma}_u^2}$$

where $\widetilde{\sigma}_u^2 = (2T)^{-1}\widetilde{\mathbf{u}}'\widetilde{\mathbf{u}}$. The asymptotic results are given as follows:

Theorem 3. *Assume that y_t is nearly doubly integrated according to (12). Joint estimation of the symmetric regression equations (7) and (8) yield the following results as $T \to \infty$:*

(a) $\tilde{\sigma}_u^2 = \frac{1}{2T}\tilde{\mathbf{u}}'\tilde{\mathbf{u}} \xrightarrow{p} \sigma_u^2$

(b) $\begin{pmatrix} T^2(\widehat{\alpha}-1) \\ T(\widehat{\gamma}-1) \end{pmatrix} \Rightarrow k_{c_1c_2} + M_{sym}^{-1}N_{sym}$

(c) $F_{\alpha=\gamma=1}^{Sym} \Rightarrow \frac{1}{2}\left\{k'_{c_1c_2}N_{sym} + k'_{c_1c_2}M_{sym}k_{c_1c_2} + N'_{sym}k_{c_1c_2} + N'_{sym}M_{sym}^{-1}N_{sym}\right\}$

$k_{c_1c_2}$ has been defined in Theorem 2 and we use the following additional symbols:

$$M_{sym} = \begin{pmatrix} 2\int_0^1 Q_{c_1}(J_{c_2}(r))^2 dr & 0 \\ 0 & 2\Psi_2 \end{pmatrix}$$

$$N_{sym} = \begin{pmatrix} 2\int_0^1 Q_{c_1}(J_{c_2}(r))dW(r) + (c_1+c_2)Q_{c_1}(J_{c_2}(1))^2 \\ -1 - 2(c_1+c_2)\Psi_2 \end{pmatrix}$$

Proof: See appendix.

The semi-parametric analogies of the above tests, which were referred to in section 2, will have the same limiting distributions as those reported in Theorems 1-3.

4. A NUMERICAL COMPARISON OF POWER FUNCTIONS UNDER LOCAL ALTERNATIVES

It remains to compare and evaluate the quantitative implications of the derived analytical distributions under local alternatives for the three tests under scrutiny. In this section the tests are compared numerically with respect to their asymptotic local power functions.

The design of the numerical simulation was conducted as follows. The data generating process was given as

$$y_t = \left[\exp(c_1/T) + \exp(c_2/T)\right] y_{t-1} - \exp((c_1+c_2)/T)y_{t-2} + u_t.$$

with $u_t \sim$n.i.d.(0,1) and with the non-centrality parameters taking the values $\{c_1, c_2\} \in \{\pm 10, \pm 8, \pm 6, \pm 4, \pm 3, \pm 2, \pm 1.5, \pm 1, \pm .5, 0\}$. Note, that because the role of c_1 and c_2 enters symmetrically in the limiting distributions, the number of parameter combinations in the simulations can be remarkably reduced. For each of the test statistics in Theorems 1-3 the analytical distributions expressed as functionals of Ornstein-Uhlenbeck processes were approximated with discretized Riemann sums with $N = 1000$ steps. For instance, $\int_0^1 Q_{c_1}(J_{c_2}(r))dW(r)$ can be approximated by $N^{-2}\sum_1^N y_{t-1}u_t$. Hence, for the simulation purposes $T = N$. Chan (1988) found that this way of approximating stochastic integrals was superior, both in terms of better approximation compared to other methods but also with respect to speed of convergence. 5000 draws from each simulated distribution were registered and the rejection frequencies at a nominal 5% level were calculated. In Tables 2-4 the rejection frequencies corresponding to the asymptotic local powers are displayed. Table 2 concerns the case where two roots are locally stationary. Table 3 has two roots locally explosive, and Table 4 displays situations with one locally stationary root and one locally explosive root. Figures 1-3 display the power surfaces for each of the three tests.

Insert Table 2 about here

Insert Table 3 about here

Insert Table 4 about here

Insert Figure 1 about here

Insert Figure 2 about here

Insert Figure 3 about here

Generally, the comparative advantages of the various tests (in terms of asymptotic local power) depend upon the region being studied. For instance, when both roots are locally explosive the Dickey-Pantula test appears to have marginally better power than the two joint tests, (see Table 3); *e.g,* when $c_1 = c_2 = 1$ power is .50 as opposed to .43 and .37 for the other two tests[6]. When the process has one unit root and one explosive root, the Dickey Pantula test has superior power properties. This is not surprising because in this case the prior differencing of the data when conducting the Dickey Pantula test is a valid restriction. From Table 2 it can be seen that for processes with one unit root and one locally stationary root, the Sen-Dickey test performs quite well for alternatives close to the double unit root case, whereas for alternatives further away, *i.e.* for larger values of c_2, the Dickey-Pantula test has slightly higher power. On the other hand, diverging away from the null towards the locally stationary region for both roots, the Sen-Dickey symmetric test clearly outperforms the two other tests, *e.g.* compare $c_1 = c_2 = -6$ where the Sen-Dickey test has a power of .91 as opposed to .76 and .78. The superiority of the Sen-Dickey test in the (locally) stationary region is what we would expect a priori. In fact, in motivating their test, Sen and Dickey (1987) noted that the symmetric version of the model is permissable in the stationary region and hence it is in this region power gains can be expected. Our results show that this conjecture is indeed correct. In comparing the Dickey-Pantula test and the Hasza-Fuller test no clear ranking can be given unless one of the roots is very close to one (*e.g.* $c_1 \approx 0$) in which case the Dickey-Pantula test has better power.

[6]For the the number of replications conducted in the simulations a difference of powers exceeding .02 can be considered significant.

Finally, observe from Table 4 (*i.e.* the case with an explosive and a stationary root), that the power function is not everywhere monotonic in c_1 and c_2. For instance, when c_1 is positive whilst c_2 is negative, the influence from a locally explosive root and a locally stationary root may off-set the influence from each other and hence may lead to reduced power. No unambiguous ranking of the tests can be given in this case, however, the general impression is that when one root exceeds unity and the other is less than unity, then in most cases the Dickey-Pantula test performs the best.

5. APPROXIMATING THE EXACT DISTRIBUTIONS VIA LOCAL TO UNITY ASYMPTOTICS

In order to examine the adequacy of the local to unity asymptotic distributions to approximate distributions in finite samples a simulation experiment was conducted. Figures 4-7 display quantile-quantile (QQ) plots of exact finite sample distributions against the associated theoretical distributions obtained from simulating the various test distributions reported in Theorems 1-3 for particular values of c_1 and c_2. Obviously, points lying on a 45 degree line will indicate that the two distributions are quite similar. The local to unity distributions were simulated according to the procedure outlined in the previous section. For given values of c_1 and c_2, the corresponding finite sample distributions were found by noting that for a series of length T the associated values of α_1 and α_2 can be found as

$$\begin{aligned}\alpha_1 &= \exp\left(\frac{c_1}{T}\right)\\ \alpha_2 &= \exp\left(\frac{c_2}{T}\right)\end{aligned}$$

According to this scheme, AR(2) processess with $N(0,1)$ innovations and zero initial conditions were simulated for various combinations of c_1, c_2, and T. The three statistics were subsequently constructed for each data series and based upon 10000 replications an estimate of the exact distribution was found.

Three basic experimental designs implying autoregressive roots getting closer and closer to unity were considered ($\{c_1 = c_2 = -10\}, \{c_1 = -4, c_2 = -6\}, \{c_1 = -1,\ c_2 = -2\}$); one experiment was such that a stationary and an explosive root were allowed, ($\{c_1 = -4, c_2 = 1\}$). The sample sizes were $T = 25, 50, 100, 250, 500, 1000$ and the corresponding AR roots α_1 and α_2, can be seen from Table 5.

Insert Table 5 about here

Insert Table 6 about here

A goodness of fit measure in comparing the exact and asymptotic distributions can be made by use of the Smirnov test, see e.g. Conover (1980). Since both the exact and asymptotic distributions are simulated from two independent samples, (with a large number of replications n), the test can be constructed from the statistic

$$D_n = \sup_x |S_1(x) - S_2(x)|$$

where $S_1(x)$,and $S_2(x)$ are the empirical distribution functions of the exact and asymptotic distributions, respectively. The null hypothesis states equality of the two distributions (against a two-sided alternative) and critical values are given by $1.92n^{-1/2}$ at a 5% level and $2.30n^{-1/2}$ at a 1% level. $n = 10000$ replications is used in construction of the empirical distribution functions. The test values are reported in Table 6.

As can be seen from the Smirnov tests the limiting theory gives a far from complete characterization of the exact finite sample distributions, but this is what we would expect. Especially given the large number of Monte Carlo repetitions ($n = 10000$) upon which the comparison is made. In some situations, however, we cannot exclude that the the exact and limiting distributions are identical, even for as small sample sizes as 100-250; see in particular the Dickey-Pantula test for $\{c_1 = -1,\ c_2 = -2\}$ and $(\{c_1 = -4, c_2 = 1\}$. The QQ plots provide a visual impression in comparing the two distributions. Obviously, the limit theory approximations are fairly bad in most cases with a small sample size (smaller than 100), but also, the approximations tend to improve considerably as the number of observations increases. Generally, the exact distribution of the Dickey-Pantula test is described better than the Hasza-Fuller and Sen-Dickey tests which is also indicated in the Smirnov tests , i.e. the Dickey-Pantula Smirnov test statistics appear to be of a smaller magnitude than for the two other tests.

Insert Figure 4 about here

Insert Figure 5 about here

Insert Figure 6 about here

Insert Figure 7 about here

6. CONCLUSION

In this paper we have derived asymptotic local power functions for a number of commonly used tests for double unit roots. Asymptotic local powers for I(2) tests were compared using theory for nearly doubly integrated processes. This asymptotic theory is very useful in describing processes that have roots close to but not exactly on the unit circle and hence provides a convenient framework for bridging the apparent divergent theories describing stationary and non-stationary data. A numerical comparison of the asymptotic local power functions indicate that the symmetric test of Sen and Dickey is clearly advantageous compared to the Hasza-Fuller and the Dickey-Pantula test when the movement is towards (locally) stationary processes. The Dickey-Pantula test appears to be better when the explosive region is approached however.

We also used the limiting distributions of our asymptotic theory as a benchmark for comparison with exact finite sample distributions. The characterization was far from being complete as one would expect, but clearly the limiting theory was describing the distributions quite well when the number of observations exceeded 250. This points to the fact that finite sample distributions can be hard to describe properly by use of asymptotic theory, however, a local to unity approximation is clearly better than relying on asymptotic results for stationary processes, say, when the near unit root region of the parameter space is approached.

In the paper we have abstracted from the presence of deterministic components in the models. In practice, it is of immense importance to properly deal with the possible presence of drifts, trends, and quadratic trends, for instance. It remains for future work to extend the analysis of the present paper to these more realistic situations, but at least our analysis has provided an initial guidance toward the relative properties of the various tests.

7. Technical Appendix

Proof of results reported in section (3).

The following Lemma 4 contains results reported in Nabeya and Perron (1994) and Perron and Ng (1996,1998):

Lemma 4. *Let $\{y_t\}$ be generated by (12), or equally (15) where u_t satisfies the general conditions of Phillips (1987b), Assumption 2.1. Then for $T \to \infty$ the following limit results apply:*

(a) $T^{-4}\sum_1^T y_{t-1}^2 \Rightarrow \sigma_u^2 \int_0^1 Q_{c_1}(J_{c_2}(r))^2 dr$

(b) $T^{-2}\sum_1^T (\Delta y_{t-1})^2 \Rightarrow \sigma_u^2 \left\{ c_1 Q_{c_1}(J_{c_2}(1))^2 - c_1^2 \int_0^1 Q_{c_1}(J_{c_2}(r))^2 dr + \int_0^1 J_{c_2}(r)^2 dr \right\} \equiv \sigma_u^2 \Psi_2$

(c) $T^{-3}\sum_1^T y_{t-1}\Delta y_{t-1} \Rightarrow \frac{\sigma_u^2}{2} Q_{c_1}(J_{c_2}(1))^2$

(d) $T^{-2}\sum_1^T y_{t-1}u_t \Rightarrow \sigma_u^2 \int_0^1 Q_{c_1}(J_{c_2}(r))dW(r)$

(e) $T^{-1}\sum_1^T \Delta y_{t-1}u_t \Rightarrow \sigma_u^2 \left\{ c_1 \int_0^1 Q_{c_1}(J_{c_2}(r))dW(r) + \int_0^1 J_{c_2}(r)dW(r) \right\} \equiv \sigma_u^2 \Psi_1$

Proof of Theorem 1. First we prove (15). We use the series expansions

$$\exp\left(\frac{c_1}{T}\right) = 1+\frac{c_1}{T}+R_{1T} \qquad (16)$$

$$\exp\left(\frac{c_2}{T}\right) = 1+\frac{c_2}{T}+R_{2T} \qquad (17)$$

where R_{1t} and R_{2T} are both $O(T^{-2})$. It follows that (12) can be rewritten as

$$\begin{aligned}\Delta^2 y_t &= \left(\frac{c_1+c_2}{T}+\frac{c_1c_2}{T^2}+R_{1T}+R_{2T}+R_{1T}\frac{c_2}{T}+R_{2T}\frac{c_1}{T}+R_{1T}R_{2T}\right)\Delta y_{t-1}\\ &+\left(-\frac{c_1c_2}{T^2}-R_{1T}\frac{c_2}{T}-R_{2T}\frac{c_1}{T}-R_{1T}R_{2T}\right)y_{t-1}+u_t\end{aligned}$$

Since $R_{1T}\Delta y_{t-1}, R_{2T}\Delta y_{t-1}, R_{1T}\frac{c_2}{T}y_{t-1}$, and $R_{2T}\frac{c_1}{T}y_{t-1}$ are all of order $O_p(T^{-3/2})$ and the remaining terms involving R_{1t}, R_{2t} are of a lesser order, the result (15) follows.

Now, the regression model we consider is given by (2) with $a-1=\left[\frac{(c_1+c_2)}{T}+\frac{c_1c_2}{T^2}\right]$, and $\omega_t=-\frac{c_1c_2}{T^2}y_{t-1}+O_p(T^{-3/2})+u_t$.

First we prove (b): The least squares estimate of a (after appropriate scaling using Lemma (4)) is given by

$$\begin{aligned}T(\hat{a}-a) &= T^{-2}\left(\sum_1^T \triangle y_{t-1}^2\right)^{-1}\times\\ &\quad T^{-1}\left(\sum_1^T \triangle y_{t-1}\left(-\frac{c_1c_2}{T^2}y_{t-1}+O_p(T^{-3/2})+u_t\right)\right)\end{aligned}$$

From this, and by using Lemma 4, (c) and (e), the second term is seen to satisfy:

$$\begin{aligned}&T^{-1}\left(\sum_1^T \triangle y_{t-1}\left(-\frac{c_1c_2}{T^2}y_{t-1}+O_p(T^{-3/2})+u_t\right)\right)\\ &= -T^{-1}\sum_1^T \triangle y_{t-1}y_{t-1}\frac{c_1c_2}{T^2}+T^{-1}\sum_1^T\triangle y_{t-1}u_t+o_p(1)\\ &\Rightarrow -c_1c_2\frac{\sigma_u^2}{2}Q_{c_1}(J_{c_2}(1))^2+\sigma_u^2\Psi_1\end{aligned}$$

whilst the limit of the first term is given in Lemma $4(b)$.

Since $T(\widehat{a}-a)=T(\widehat{a}-1)-(c_1+c_2)+O(T^{-1})$, the result (b) follows.

Result (a) follows from the consistency of $\widehat{a}$ and the fact that

$$\frac{1}{T}\sum_2^T\omega_t^2=\frac{1}{T}\sum_2^T u_t^2+o_p(1)\xrightarrow{p}\sigma_u^2$$

The limiting result (c) can be deduced by using results (a), and (b) of Theorem 1, and Lemma 4(b) in evaluating $t_{a-1} = (\widehat{a}-1)/[s_\omega(\sum_{t=1}^T \triangle y_{t-1}^2)^{-1/2}]$.

Proof of Theorem 2. Result (a) holds trivially given the consistency of the parameter estimates reported in *(b)* below.

To prove (b), define $D_T = diag(T^{3/2}, T^{1/2})$ such that $D_T^{-1}(y_t, v_t)' \Rightarrow (\sigma_u Q_{c_1}(J_{c_2}(r)), \sigma_u J_{c_1}(r))'$. It follows from Lemma 4 that by letting $\mathbf{x}_t = (y_t, \Delta y_t)'$

$$\frac{1}{T}D_T^{-1}\sum_1^T \mathbf{x}_{t-1}\mathbf{x}_{t-1}' D_T^{-1} = \begin{pmatrix} T^{-4}\sum_1^T y_{t-1}^2 & T^{-3}\sum_1^T y_{t-1}\Delta y_{t-1} \\ T^{-3}\sum_1^T y_{t-1}\Delta y_{t-1} & T^{-2}\sum_1^T \Delta y_{t-1}^2 \end{pmatrix} \Rightarrow \sigma_u^2 M$$

with M defined in the Theorem. Similarly it holds that

$$\frac{1}{T^{1/2}}D_T^{-1}\sum_1^T \mathbf{x}_{t-1}\mathbf{u}_t = \begin{pmatrix} T^{-2}\sum_1^T y_{t-1}u_t \\ T^{-1}\sum_1^T \Delta y_{t-1}u_t \end{pmatrix} \Rightarrow \sigma_u^2 N$$

The (scaled) least squares estimator can now be written as:

$$T^{1/2}D_T(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}) = \left(\frac{1}{T}D_T^{-1}\sum_1^T \mathbf{x}_{t-1}\mathbf{x}_{t-1}'D_T^{-1}\right)^{-1}\frac{1}{T^{1/2}}D_T^{-1}\sum_1^T \mathbf{x}_{t-1}\mathbf{u}_t \Rightarrow M^{-1}N$$

Because

$$T^{1/2}D_T(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}) = \begin{pmatrix} T^2(\widehat{\alpha}-\alpha) \\ T(\widehat{\gamma}-\gamma) \end{pmatrix} = \begin{pmatrix} T^2(\widehat{\alpha}-1) \\ T(\widehat{\gamma}-1) \end{pmatrix} + \begin{pmatrix} c_1c_2 \\ -(c_1+c_2) \end{pmatrix} + O(T^{-1})$$

the desired result has been obtained with the definition $k_{c_1c_2} = (-c_1c_2, c_1+c_2)'$.

In deriving the distribution of the $F-$statistic (c), we exploit that

$$F_{\alpha=\gamma=1} = \frac{\widehat{\boldsymbol{\beta}}'\left(\sum_1^T \mathbf{x}_{t-1}\mathbf{x}_{t-1}'\right)\widehat{\boldsymbol{\beta}}}{2s_u^2} = \frac{T^{1/2}\widehat{\boldsymbol{\beta}}'D_T\left(T^{-1}D_T^{-1}\sum_1^T \mathbf{x}_{t-1}\mathbf{x}_{t-1}'D_T^{-1}\right)D_T T^{1/2}\widehat{\boldsymbol{\beta}}}{2s_u^2}.$$

The previously used procedure can be applied again, and hence we evaluate each component in this quantity, realizing that each term is non-degenerate. The result follows subsequently.

Proof of Theorem 3. We start by proving (b). Using $D_T = diag(T^{3/2}, T^{1/2})$ we have that

$$T^{1/2}D_T(\widetilde{\boldsymbol{\beta}}-\boldsymbol{\beta}) = \left(\frac{1}{T}D_T^{-1}\mathbf{X}'\mathbf{X}D_T^{-1}\right)^{-1}\left(\frac{1}{T^{1/2}}D_T^{-1}\mathbf{X}'\mathbf{u}\right) =$$

$$\left\{ \begin{array}{cc} 2T^{-4}\sum_3^T y_{t-1}^2 & T^{-3}\left\{\sum_3^{T-1}(\Delta y_t)^2 + y_2\Delta y_2 - y_{T-1}\Delta y_T\right\} \\ T^{-3}\left\{\sum_3^{T-1}(\Delta y_t)^2 + y_2\Delta y_2 - y_{T-1}\Delta y_T\right\} & T^{-2}\left\{2\sum_3^{T-1}(\Delta y_t)^2 + (\Delta y_2)^2 + (\Delta y_T)^2\right\} \end{array} \right\}^{-1} \times$$
$$\left\{ \begin{array}{c} T^{-2}\sum_3^T y_{t-1}u_t + T^{-2}\sum_3^T y_{t-1}v_{t-2} \\ T^{-1}\sum_3^T u_t\Delta y_{t-1} - T^{-1}\sum_3^T \Delta y_t v_{t-2} \end{array} \right\}$$

The first task is to compute the limiting distributions of the terms in $\left(\frac{1}{T}D_T^{-1}\mathbf{X}'\mathbf{X}D_T^{-1}\right)$. It can be seen from Lemma 4 that the off-diagonal terms in the matrix $\left(\frac{1}{T}D_T^{-1}\mathbf{X}'\mathbf{X}D_T^{-1}\right)$ will be $O_p(T^{-1})$ which is implied by the construction of the symmetric regression. Also,

$$T^{-2}\left\{2\sum_3^{T-1}(\Delta y_t)^2 + (\Delta y_2)^2 + (\Delta y_T)^2\right\} = 2T^{-2}\sum_3^{T-1}(\Delta y_t)^2 + o_p(1)$$

which together with Lemma $4(a)$ establishes that

$$\left(\frac{1}{T}D_T^{-1}\mathbf{X}'\mathbf{X}D_T^{-1}\right) \Rightarrow \sigma_u^2 M_{sym}$$

where M_{sym} is defined in the Theorem.

The second task is to compute the limiting distributions of the terms in

$$\left(\frac{1}{T^{1/2}}D_T^{-1}\mathbf{X}'\mathbf{u}\right) = \left\{ \begin{array}{c} T^{-2}\sum_3^T y_{t-1}u_t + T^{-2}\sum_3^T y_{t-1}v_{t-2} \\ T^{-1}\sum_3^T u_t\Delta y_{t-1} - T^{-1}\sum_3^T \Delta y_t v_{t-2} \end{array} \right\} \tag{18}$$

First, we will calculate the limit distribution of the first entry of this matrix which also reads

$$T^{-2}\sum_3^T y_{t-1}(u_t + v_{t-2}) \tag{19}$$

For that purpose we need to expand the coefficients in the "reverse regression" by the second order expansions (16) and (17) to obtain an expression for v_{t-2}. Using arguments similar to those in the proof of theorem 1, it can be easily shown that

$$v_{t-2} = \triangle^2 y_t + \left[\frac{c_1+c_2}{T} + \frac{c_1c_2}{T^2}\right]\triangle y_t + \frac{c_1c_2}{T^2}y_{t-1} + O_p(T^{-3/2}) \tag{20}$$

Using the expression for $\Delta^2 y_t$, (15), and inserting into (20), (19) can be rewritten

$$\begin{aligned} T^{-2}\sum_3^T y_{t-1}(u_t + v_{t-2}) &= 2T^{-2}\sum_3^T y_{t-1}u_t + T^{-2}\sum_3^T y_{t-1}\left[\frac{c_1+c_2}{T} + \frac{c_1c_2}{T^2}\right]\triangle y_{t-1} \\ &\quad + T^{-2}\sum_3^T y_{t-1}\left[\frac{c_1+c_2}{T} + \frac{c_1c_2}{T^2}\right]\triangle y_t + o_p(1) \end{aligned}$$

$$= \quad 2T^{-2}\sum_{3}^{T} y_{t-1}u_t + 2T^{-2}\sum_{3}^{T} y_{t-1}\frac{c_1+c_2}{T}\triangle y_{t-1} + o_p(1)$$
$$\Rightarrow \quad 2\sigma_u^2 \int Q_{c_1}(J_{c_2}(r))dW(r) + (c_1+c_2)\sigma_u^2 Q_{c_1}(J_{c_2}(1))^2$$

Now we will calculate the limit distribution of the second entry in (18), namely

$$T^{-1}\sum_{3}^{T} u_t\Delta y_{t-1} - T^{-1}\sum_{3}^{T}\Delta y_t v_{t-2} \tag{21}$$

Doing the same trick as above, that is, inserting (15) into (20) and again into

$$T^{-1}\sum_{3}^{T}\Delta y_t v_{t-2}$$

yields

$$T^{-1}\sum_{3}^{T}\Delta y_t v_{t-2} = T^{-1}\sum_{3}^{T}\triangle y_t\triangle y_{t-1}\frac{c_1+c_2}{T} + T^{-1}\sum_{3}^{T}\triangle y_t u_t + T^{-1}\sum_{3}^{T}(\triangle y_t)^2\frac{c_1+c_2}{T} + o_p(1)$$

Inserting the above expression into (21), and rearranging

$$\begin{aligned} T^{-1}\left(\sum_{3}^{T} u_t\Delta y_{t-1} - \sum_{3}^{T}\Delta y_t v_{t-2}\right) &= -T^{-1}\sum_{3}^{T} u_t\triangle^2 y_t - T^{-1}\frac{c_1+c_2}{T}\sum_{3}^{T}\triangle y_t\triangle y_{t-1} \\ &\quad -T^{-1}\frac{c_1+c_2}{T}\sum_{3}^{T}(\triangle y_t)^2 + o_p(1) \\ &= -T^{-1}\sum_{3}^{T} u_t\triangle^2 y_t - 2T^{-1}\frac{c_1+c_2}{T}\sum_{3}^{T}(\triangle y_t)^2 + o_p(1) \end{aligned}$$

where we have exploited that $\Delta y_t = \Delta^2 y_t + \Delta y_{t-1}$.

Now substitute $\triangle^2 y_t$ for (15) to get

$$\begin{aligned} T^{-1}\left(\sum_{3}^{T} u_t\triangle y_{t-1} - \sum_{3}^{T}\triangle y_t v_{t-2}\right) &= -T^{-1}\sum_{3}^{T} u_t^2 - 2T^{-1}\frac{c_1+c_2}{T}\sum_{3}^{T}(\triangle y_t)^2 + o_p(1) \\ &\Rightarrow -\sigma_u^2 - 2(c_1+c_2)\sigma_u^2\Psi_2 \end{aligned}$$

Hence we have established that

$$T^{1/2}D_T(\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta}) \Rightarrow M_{sym}^{-1}N_{sym}$$

where N_{sym} is defined in the Theorem. Because

$$T^{1/2}D_T(\widetilde{\boldsymbol{\beta}}-\boldsymbol{\beta}) = \begin{pmatrix} T^2(\hat{\alpha}-\alpha) \\ T(\hat{\gamma}-\gamma) \end{pmatrix} = \begin{pmatrix} T^2(\hat{\alpha}-1) \\ T(\hat{\gamma}-1) \end{pmatrix} + \begin{pmatrix} c_1c_2 \\ -(c_1+c_2) \end{pmatrix} + O(T^{-1})$$

the result (b) in the Theorem is now proven.

To derive (a), we need to establish that $\tilde{\sigma}_u^2 \to \sigma_u^2$:

$$\tilde{\sigma}_u^2 = \frac{1}{2T}\widetilde{\mathbf{u}}'\widetilde{\mathbf{u}} = \frac{1}{2}\frac{\sum_3^T \hat{u}_t^2}{T} + \frac{1}{2}\frac{\sum_3^T \hat{v}_{t-2}^2}{T}$$

As $\widetilde{\boldsymbol{\beta}}$ is consistent,

$$\frac{1}{2}\frac{\sum_3^T \hat{u}_t^2}{T} \to \frac{1}{2}\sigma_u^2$$

Inserting (15) into (20) and letting $T \to \infty$, it is clear that

$$v_{t-2} = u_t + O_p(T^{-1/2})$$

and

$$v_{t-2}^2 = u_t^2 + o_p(1)$$

Result (a) follows from the consistency of $\widetilde{\boldsymbol{\beta}}$, and the fact that

$$\frac{1}{2}\frac{\sum_3^T v_{t-2}^2}{T} = \frac{1}{2}\frac{\sum_3^T u_t^2}{T} + o_p(1) \to \frac{1}{2}\sigma_u^2$$

Result (c) can be deduced by using arguments analogous to those in the proof of Theorem 2. In particular, with the present notation, the symmetric $F-$statistic reads

$$F_{\alpha=\gamma=1}^{Sym} = \frac{T^{1/2}\widetilde{\boldsymbol{\beta}}'D_T\left(T^{-1}D_T^{-1}\mathbf{X}'\mathbf{X}D_T^{-1}\right)D_TT^{1/2}\widetilde{\boldsymbol{\beta}}}{2\widetilde{\sigma}_u^2}.$$

Taking appropriate limits the result (c) reported in Theorem 3 is seen to apply.

REFERENCES

[1] Bobkoski, M.J., (1983). Hypothesis testing in nonstationary time series. Unpublished PhD. thesis. (Department of Wisconsin, Madison, WI).

[2] Chan, N.H., (1988). The parameter inference for nearly nonstationary time series. *Journal of American Statistical Association* **83**, 857-862.

[3] Chan, N.H., and C.Z. Wei, (1987). Asymptotic Inference for nearly nonstationary AR(1) processes. *Annals of Statistics* **15**, 1050-1063.

[4] Choi, I. J. Y. Park and B. Yu, (1997). Canonical cointegration regression and testing for cointegration in the presence of I(1) and I(2) variables. *Econometric Theory* **13,** 850-876

[5] Conover, W. J., (1980). Practical Nonparametric Statistics, 2ed, John Wiley and Sons.

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

[7] Dickey, D. A. and S. G. Pantula, (1987). Determining the order of differencing in autoregressive processes. *Journal of Business and Economic Statistics* **5**, 455-461.

[8] Fuller, W. A, (1976). Introduction to statistical time series, New York, John Wiley & Sons.

[9] Haldrup, N., (1994a). Semiparametric tests for double unit roots. *Journal of Business and Economic Statistics* **12**, 109-122.

[10] Haldrup, N., (1994b). The asymptotics of single-equation cointegration regressions with I(1) and I(2) variables. *Journal of Econometrics,* **63**, 153-81.

[11] Haldrup, N. (1998). An econometric analysis of I(2) variables. *Journal of Economic Surveys* **12**, 595-650.

[12] Hasza, D. P. and W. A. Fuller, (1979). Estimation of autoregressive processes with unit roots. *The Annals of Statistics* **7**, 1106-1120.

[13] Jeganathan, P., (1991). On the asymptotic behavior of least-squares estimators in AR time series with roots near the unit circle. *Econometric Theory* **7**, 269-306.

[14] Johansen, S., (1995) A statistical analysis of cointegration for I(2) variables. *Econometric Theory* **11,** 25-29.

[15] Johansen, S., (1997) A likelihood analysis of the I(2) model. *Scandinavian Journal of Statistics,* **24,** 433-462.

[16] King, R. G., C. I. Plosser, J. H. Stock, and M. W. Watson, (1991). Stochastic trends and economic fluctuations. *American Economic Review* **81**, 819-840.

[17] Kitamura, Y., (1995) Estimation of cointegrated systems with I(2) processes. *Econometric Theory,* **11,** 1-24.

[18] Nabeya, S., and P. Perron, (1994). Local asymptotic distributions related to the AR(1) model with dependent errors. *Journal of Econometrics* **62**, 229-264.

[19] Perron, P., and S. Ng, (1996). Useful modifications to some unit root tests with dependent errors and their local asymptotic properties. *Review of Economic Studies* **63**, 435-463.

[20] Perron, P., and S. Ng, (1998). An autoregressive spectral density estimator at frequency zero for nonstationarity tests. *Econometric Theory* **14**, 560-603.

[21] Phillips, P. C. B., (1987a). Towards a unified asymptotic theory for autoregression. *Biometrika* **74,** 535-547.

[22] Phillips, P. C. B., (1987b). Time series regression with unit root. *Econometrica* **55**, 277-302.

[23] Phillips, P. C. B. and P. Perron, (1988). Testing for a unit root in time series regression. *Biometrika* **75**, 335-346.

[24] Sen, D. L., (1986). Robustness of single unit root test statistics in the presence of multiple unit roots. Unpublished Ph.D. thesis, North Carolina State University, Dept. of Statistics.

[25] Sen, D. L. and D. A. Dickey, (1987). Symmetric test for second differencing in univariate time series. *Journal of Business and Economic Statistics* **5**, 463-473.

[26] Shin, D. W., and H. J. Kim, (1999). Semiparametric tests for double unit roots based on symmetric estimators. *Journal of Business and Economic Statistics* **7**, 67-73.

[27] Swensen, A. R., (1993). A note on the asymptotic power calculations in nearly nonstationary time series. *Econometric Theory,* **9**, 659-667.

8. Tables

Table 1. Percentiles of the Sen-Dickey $F^{Sym}_{\alpha=\gamma=1}$ test without deterministic components. The fractiles are found from 250000 Monte Carlo replications.

	Probability of a smaller value								
T	.01	.025	.05	.10	.50	.90	.95	.975	.99
25	0.02	0.04	0.06	0.11	0.80	3.20	4.36	5.59	7.32
50	0.05	0.07	0.11	0.18	1.06	3.72	4.86	6.07	7.71
100	0.08	0.12	0.17	0.26	1.30	4.24	5.41	6.59	8.23
250	0.13	0.18	0.24	0.36	1.53	4.77	6.05	7.34	9.01
500	0.15	0.21	0.29	0.42	1.68	5.09	6.41	7.74	9.54
∞(1000)	0.17	0.23	0.31	0.45	1.78	5.28	6.68	8.03	9.82

Table 2. Asymptotic local powers. Two locally stationary roots. For each entry of c_1 and c_2 the powers are calculated respectively for the Dickey-Pantula, Hasza-Fuller, and Sen-Dickey tests.

c_2	c_1									
	-10	-8	-6	-4	-3	-2	-1.5	-1	-0.5	0
0	-	-	-	-	-	-	-	-	-	0.05 0.05 0.05
-0.5	-	-	-	-	-	-	-	-	0.04 0.05 0.07	0.04 0.05 0.05
-1	-	-	-	-	-	-	-	0.07 0.08 0.10	0.05 0.06 0.08	0.04 0.06 0.07
-1.5	-	-	-	-	-	-	0.11 0.10 0.14	0.08 0.10 0.13	0.06 0.08 0.10	0.05 0.06 0.08
-2	-	-	-	-	-	0.15 0.15 0.20	0.12 0.13 0.17	0.10 0.11 0.14	0.08 0.10 0.12	0.06 0.07 0.10
-3	-	-	-	-	0.28 0.26 0.36	0.21 0.20 0.27	0.18 0.16 0.23	0.15 0.14 0.20	0.12 0.12 0.15	0.09 0.10 0.12
-4	-	-	-	0.46 0.42 0.58	0.36 0.34 0.46	0.28 0.26 0.35	0.25 0.23 0.29	0.21 0.18 0.26	0.17 0.16 0.21	0.13 0.13 0.17
-6	-	-	0.76 0.78 0.91	0.62 0.61 0.76	0.55 0.51 0.65	0.46 0.39 0.52	0.41 0.34 0.46	0.36 0.30 0.37	0.31 0.27 0.32	0.25 0.22 0.28
-8	-	0.94 0.97 1.00	0.88 0.91 0.97	0.77 0.78 0.89	0.71 0.69 0.80	0.62 0.55 0.69	0.58 0.50 0.62	0.52 0.45 0.54	0.47 0.37 0.47	0.42 0.32 0.38
-10	0.99 1.00 1.00	0.98 0.99 1.00	0.94 0.97 0.99	0.88 0.88 0.96	0.83 0.81 0.91	0.76 0.70 0.82	0.73 0.65 0.76	0.70 0.58 0.70	0.65 0.51 0.62	0.57 0.46 0.53

Note. The asymptotic distributions of the test statistics as described in Theorems 1-3 have been simulated (see main text for details). The table reports the rejection frequencies after 5000 Monte Carlo draws. The rejection frequencies approximate the true powers.

If $\hat{p}_1$ and $\hat{p}_2$ denote the estimated powers of two different tests against a specific alternative, then, given the number of Monte Carlo replications, $\text{std}(\hat{p}_1 - \hat{p}_2) < .01$. A difference of .02 is thus to be considered significant.

Table 3. Asymptotic local powers. Two locally explosive roots. For each entry of c_1 and c_2 the powers are calculated respectively for the Dickey-Pantula, Hasza-Fuller, and Sen-Dickey tests.

c_2	c_1									
	0	0.5	1	1.5	2	3	4	6	8	10
10	-	-	-	-	-	-	-	-	-	1.00 1.00 1.00
8	-	-	-	-	-	-	-	-	1.00 1.00 1.00	1.00 1.00 1.00
6	-	-	-	-	-	-	-	1.00 1.00 1.00	1.00 1.00 1.00	1.00 1.00 1.00
4	-	-	-	-	-	-	0.99 0.99 0.98	1.00 1.00 1.00	1.00 1.00 1.00	1.00 1.00 1.00
3	-	-	-	-	-	0.96 0.96 0.94	0.98 0.98 0.98	1.00 1.00 1.00	1.00 1.00 1.00	1.00 1.00 1.00
2	-	-	-	-	0.85 0.84 0.83	0.93 0.92 0.90	0.97 0.97 0.96	0.99 1.00 0.99	1.00 1.00 1.00	1.00 1.00 1.00
1.5	-	-	-	0.72 0.69 0.65	0.81 0.79 0.75	0.91 0.90 0.88	0.96 0.96 0.95	1.00 0.99 0.99	1.00 1.00 1.00	1.00 1.00 1.00
1	-	-	0.50 0.43 0.37	0.65 0.57 0.54	0.75 0.71 0.68	0.90 0.88 0.86	0.96 0.95 0.94	0.99 0.99 0.99	1.00 1.00 1.00	1.00 1.00 1.00
0.5	-	0.22 0.14 0.11	0.38 0.28 0.24	0.54 0.45 0.42	0.68 0.63 0.60	0.86 0.85 0.83	0.95 0.94 0.94	0.99 0.99 0.99	1.00 1.00 1.00	1.00 1.00 1.00
0	0.05 0.05 0.05	0.11 0.07 0.06	0.24 0.15 0.13	0.44 0.33 0.31	0.62 0.54 0.50	0.83 0.83 0.81	0.94 0.94 0.93	0.99 0.99 0.99	1.00 1.00 1.00	1.00 1.00 1.00

See Note of Table 2.

Table 4. Asymptotic local powers. One root locally stationary and one root locally explosive. For each entry of c_1 and c_2 the powers are calculated respectively for the Dickey-Pantula, Hasza-Fuller, and Sen-Dickey tests.

c_2	c_1									
	0	0.5	1	1.5	2	3	4	6	8	10
	0.05	0.11	0.24	0.44	0.62	0.83	0.94	0.99	1.00	1.00
0	0.05	0.07	0.15	0.33	0.54	0.83	0.94	0.99	1.00	1.00
	0.05	0.06	0.13	0.31	0.50	0.81	0.93	1.00	1.00	1.00
	0.04	0.06	0.15	0.32	0.54	0.82	0.94	0.99	1.00	1.00
-0.5	0.05	0.06	0.09	0.23	0.46	0.80	0.93	0.99	1.00	1.00
	0.05	0.06	0.07	0.21	0.43	0.79	0.93	0.99	1.00	1.00
	0.04	0.05	0.09	0.24	0.47	0.79	0.93	0.99	1.00	1.00
-1	0.06	0.05	0.07	0.17	0.37	0.78	0.92	0.99	1.00	1.00
	0.07	0.06	0.06	0.13	0.36	0.76	0.91	0.99	1.00	1.00
	0.05	0.04	0.06	0.18	0.40	0.77	0.93	0.99	1.00	1.00
-1.5	0.06	0.05	0.05	0.10	0.31	0.75	0.92	0.99	1.00	1.00
	0.08	0.06	0.05	0.10	0.29	0.73	0.91	0.99	1.00	1.00
	0.06	0.05	0.05	0.12	0.35	0.76	0.91	0.99	1.00	1.00
-2	0.07	0.06	0.06	0.08	0.25	0.73	0.92	0.99	1.00	1.00
	0.10	0.08	0.05	0.07	0.23	0.70	0.90	0.99	1.00	1.00
	0.09	0.07	0.06	0.08	0.25	0.71	0.91	0.99	1.00	1.00
-3	0.10	0.08	0.07	0.07	0.19	0.68	0.90	0.98	1.00	1.00
	0.12	0.10	0.07	0.06	0.16	0.65	0.89	0.99	1.00	1.00
	0.13	0.11	0.09	0.08	0.19	0.64	0.89	0.99	1.00	1.00
-4	0.13	0.12	0.09	0.08	0.15	0.67	0.90	0.99	1.00	1.00
	0.17	0.13	0.11	0.09	0.11	0.62	0.88	0.99	1.00	1.00
	0.25	0.21	0.15	0.12	0.13	0.58	0.87	0.99	1.00	1.00
-6	0.22	0.18	0.15	0.13	0.16	0.62	0.89	0.99	1.00	1.00
	0.28	0.20	0.16	0.12	0.10	0.53	0.85	0.99	1.00	1.00
	0.42	0.34	0.27	0.19	0.16	0.53	0.85	0.99	1.00	1.00
-8	0.32	0.27	0.24	0.20	0.20	0.62	0.87	0.99	1.00	1.00
	0.38	0.30	0.25	0.17	0.14	0.46	0.84	0.98	1.00	1.00
	0.57	0.50	0.40	0.29	0.19	0.49	0.83	0.98	1.00	1.00
-10	0.46	0.41	0.35	0.30	0.28	0.62	0.88	0.98	1.00	1.00
	0.53	0.44	0.35	0.26	0.19	0.42	0.83	0.98	1.00	1.00

See Note of Table 2.

Table 5. Implied values of the AR parameters α_1 and α_2 in the AR(1) model (1-$\alpha_1 L)(1-\alpha_2 L)y_t = \varepsilon_t$ for various combinations of the sample size T, and non-centrality parameters c_1, and c_2.

	$(c_1, c_2) = (-10, -10)$	$(c_1, c_2) = (-4, -6)$	$(c_1, c_2) = (-1, -2)$	$(c_1, c_2) = (-4, 1)$
T	(α_1, α_2)	(α_1, α_2)	(α_1, α_2)	(α_1, α_2)
25	$(.670, 670)$	$(.852, .787)$	$(.961, .923)$	$(.852, 1.041)$
50	$.(819, .819)$	$(.923, .887)$	$(.980, .961)$	$(.923, 1.020)$
100	$(.905, 905)$	$(.961, .942)$	$(.990, .980)$	$(.961, 1.010)$
250	$(.961, 961)$	$(.984, .976)$	$(.996, .992)$	$(.984, 1.004)$
500	$(.980, 980)$	$(.992, .988)$	$(.998, .996)$	$(.992, 1.002)$
1000	$(.990, 990)$	$(.996, .994)$	$(.999, .998)$	$(.996, 1.001)$

Table 6. Smirnov goodness-of-fit test for equality of the exact distribution and the local to unity asymptotic distribution for various combinations of the sample size T, and the non-centrality parameters c_1, and c_2. The distributions were approximated via 10000 Monte Carlo repetitions. Equality of the distributions is rejected at 5% level (**) if the test value exceeds .019. At 1% level the critical value is .023 (***).

T	(c_1, c_2)			
	$(-10, -10)$	$(-4, -6)$	$(-1, -2)$	$(-4, 1)$
Dickey-Pantula				
25	.119***	.085***	.040***	.028***
50	.069***	.046***	.022**	.019
100	.028***	.022**	.012	.009
250	.033***	.017	.018	.009
500	.036***	.031***	.032***	.011
1000	.051***	.039***	.029***	.009
Hasza-Fuller				
25	.263***	.167***	.096***	.082***
50	.148***	.085***	.052***	.044***
100	.062***	.035***	.023**	.022**
250	.028***	.016	.012	.013
500	.047***	.033***	.024	.016
1000	.057***	.042***	.023**	.013
Sen-Dickey				
25	.376***	.341***	.242***	.212***
50	.258***	.214***	.137***	.111***
100	.170***	.126***	.073***	.070***
250	.086***	.047***	.022**	.030***
500	.043***	.023**	.016	.023**
1000	.047***	.024***	.016	.016

9. FIGURES

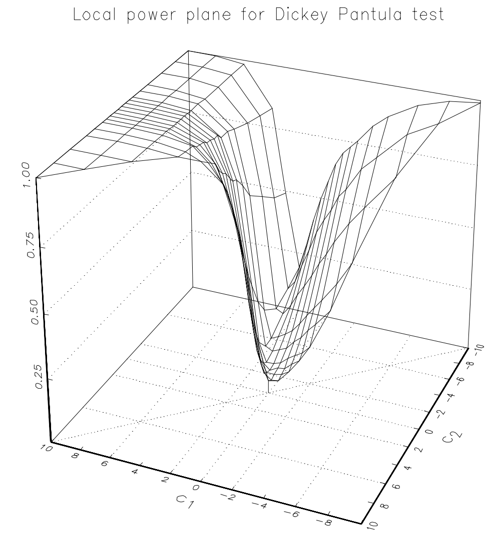


Figure 1. Asymptotic local power function for the Dickey-Pantula test.

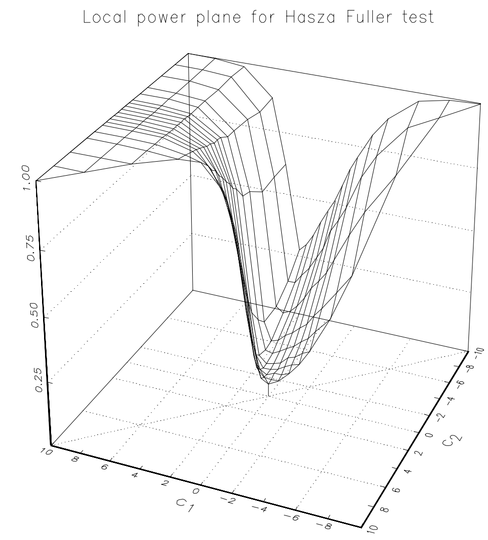


Figure 2. Asymptotic local power function for the Hasza-Fuller test.

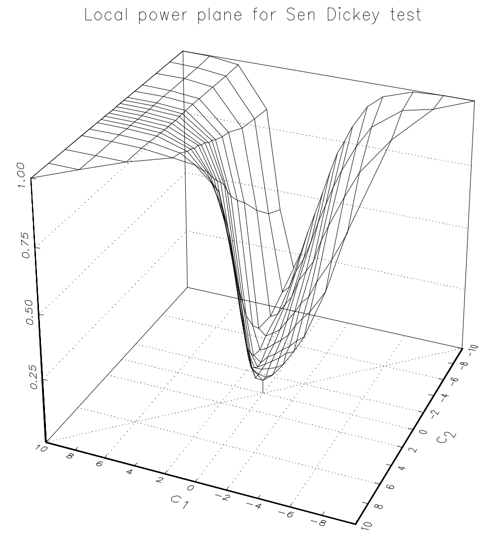
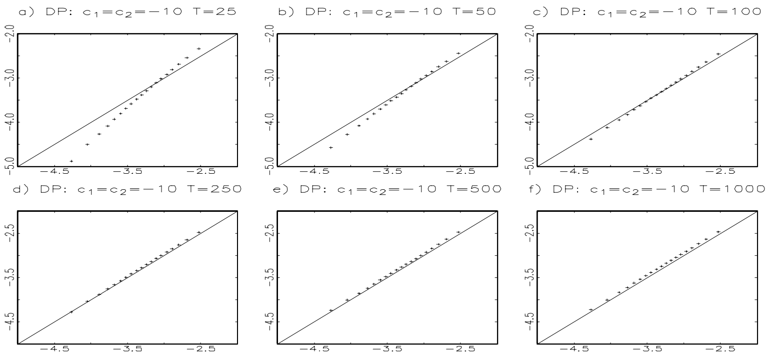
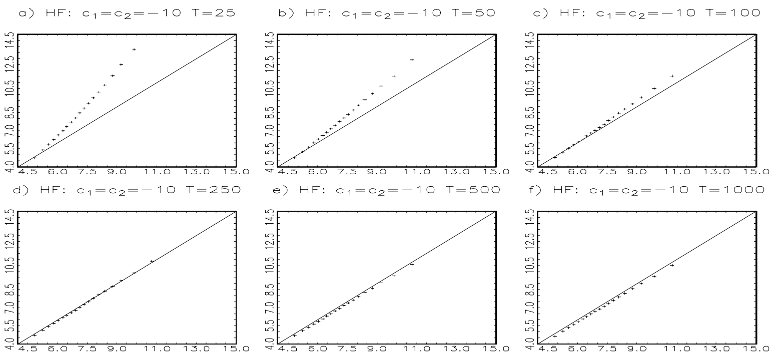


Figure 3. Asymptotic local power function for the Sen-Dickey test.

DICKEY-PANTULA



HASZA-FULLER



SEN-DICKEY

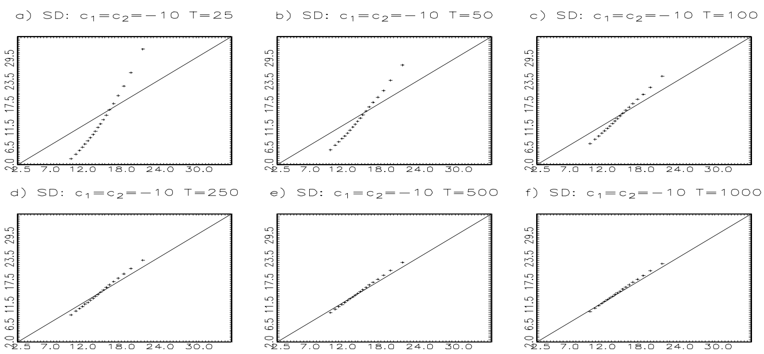
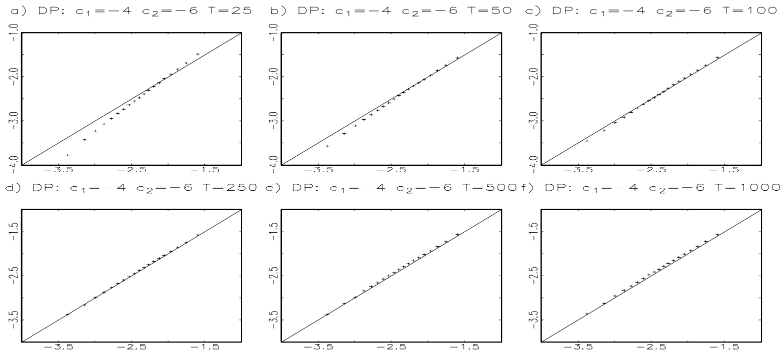
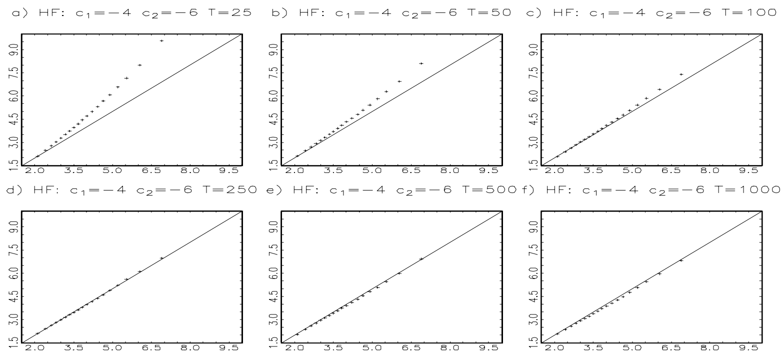


Figure 4. QQ plot: Exact finite sample distribution (vertical axis) versus asymptotic local to unity distribution (horizonzal axis). $c_1 = c_2 = -10$. $T = 25, 50, 100, 250, 500, 1000$.

DICKEY-PANTULA



HASZA-FULLER



SEN-DICKEY

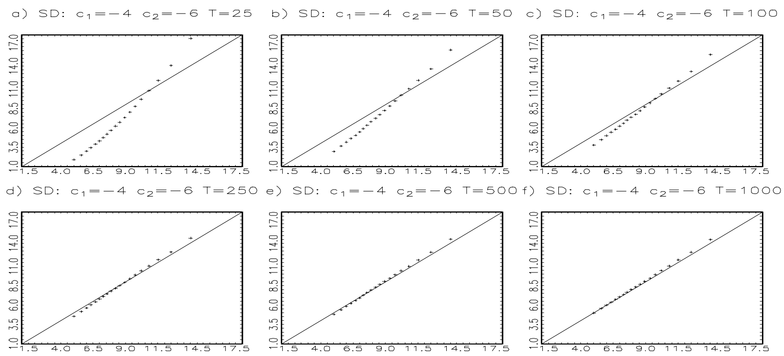
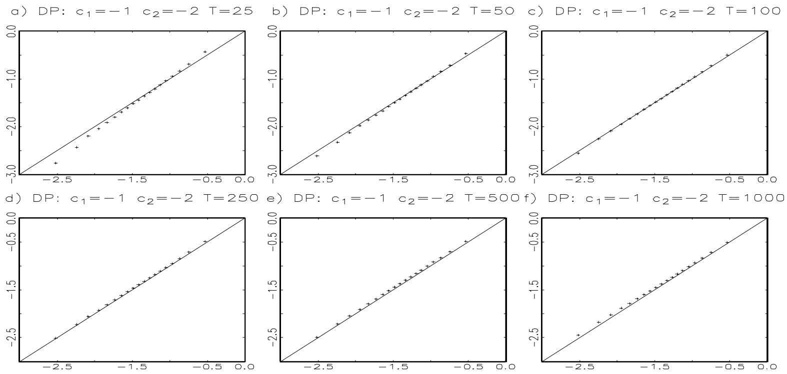
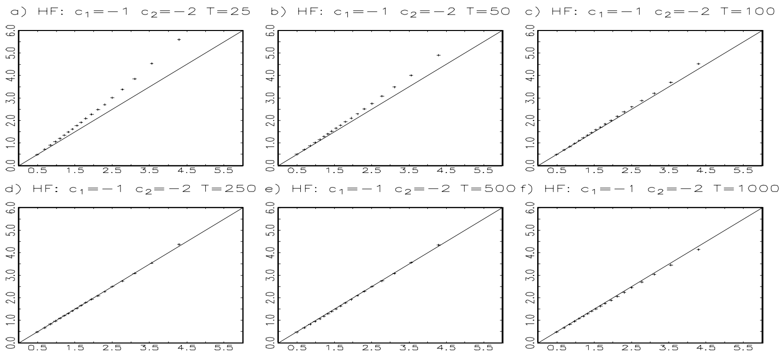


Figure 5. QQ plot: Exact finite sample distribution (vertical axis) versus local to unity distribution (horizontal axis). $c_1 = -4, c_2 = -6$. $T = 25, 50, 100, 250, 500, 1000$.

DICKEY-PANTULA



HASZA-FULLER



SEN-DICKEY

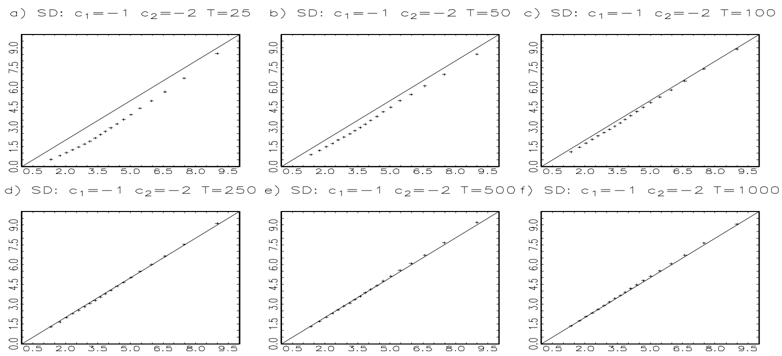
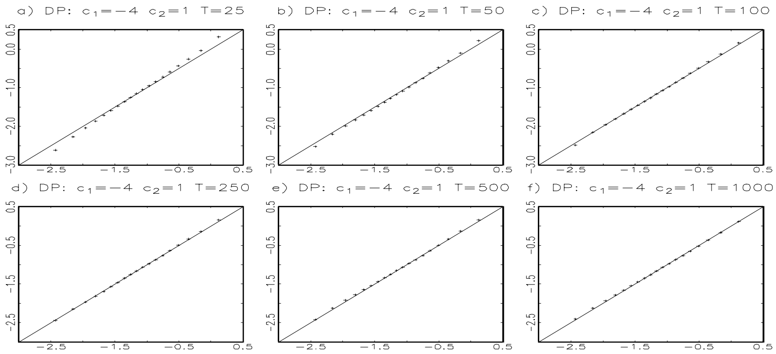
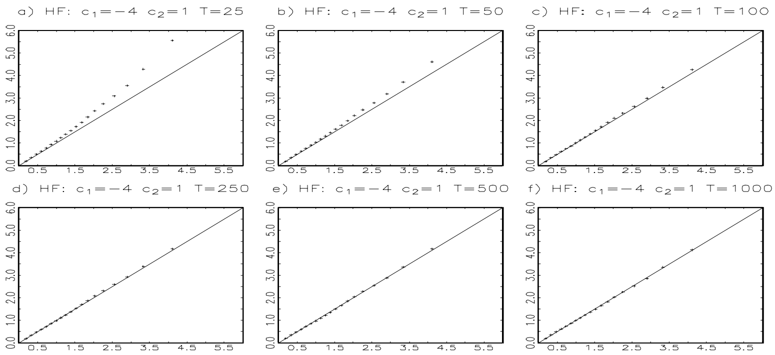


Figure 6. QQ plot: Exact finite sample distribution (vertical axis) versus local to unity distribution (horizontal axis). $c_1 = -1, c_2 = -2$. $T = 25, 50, 100, 250, 500, 1000$.

Dickey-Pantula



Hasza-Fuller



Sen-Dickey

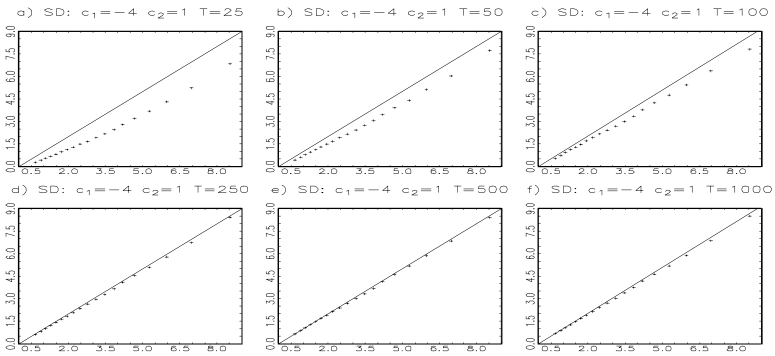


Figure 7. QQ plot: Exact finite sample distribution (vertical axis) versus local to unity distribution (horizontal axis). $c_1 = -4, c_2 = 1$. $T = 25, 50, 100, 250, 500, 1000$.