January 31, 1996

Rank Tests for Unit Roots*

Jörg Breitung

Humboldt University, Berlin
Institut für Statistik und Ökonometrie,
Humboldt Universität zu Berlin,
Spandauer Strasse 1, D-10178 Berlin

Christian Gouriéroux

CREST–CEPREMAP,
Bâtiment Malakoff 2
15, Boulevard Gabriel Péri
92245 Malakoff Cedex, France

ABSTRACT:

In order to obtain exact distributional results without imposing restrictive parametric assumptions, several rank counterparts of the Dickey-Fuller statistic are considered. In particular, a rank counterpart of the score statistic is suggested which appears to have attractive theoretical properties. Assuming i.i.d. errors, an exact test is obtained for a random walk model with drift and under assumptions similar to Phillips & Perron (1988) the test is asymptotically valid. In a Monte Carlo study the rank tests are compared with their parametric counterparts.

* The research for this paper was carried out while the second author was a guest of the "Sonderforschungsbereich 373", Humboldt University Berlin. We wish to thank the "Deutsche Forschungsgemeinschaft" (DFG) for financial support. Furthermore, the comments and suggestions of the Editor and an anonymous referee are gratefully acknowledged.

1 Introduction

Unit root tests play an important role in the analysis of economic time series. The performance of such tests depends on a number of assumptions, which are often questionable in empirical applications. As argued by Granger & Hallman (1991), the assumption that the generating process is linear seems too restrictive in many circumstances. In fact, the time series to be tested are often transformed to logarithms before the unit root test is applied. As another example, Franses and McAleer (1995) consider the Box–Cox class of transformations for unit root testing. Nevertheless, such transforms are often chosen for convenience and alternative transforms may be considered plausible as well. Therefore, a test which is unaffected by the choice of the initial transformation is highly desirable.

Most unit root tests are based on the assumption of normally distributed errors. Although the asymptotic theory applies also for a wide class of alternate assumptions provided that the second moment exists, the critical values for small samples are computed using normally distributed data. However, it is well known that in particular the distributions of financial data exhibit much fatter tails than is expected by assuming normality and in some cases not even the second moments seem to exist.

Another reason to question the use of parametric unit root tests is that in many cases the economic development reveals evidence for a structural change. In a random walk model the shocks have a persistent effect so that structural breaks can be modeled allowing for outlying observations (Perron 1989). Since rank tests reduce the influence of outlying observations, they are expected to perform better than parametric tests in such situations.

To overcome these difficulties it is interesting to consider robust versions of standard unit root tests, such as tests based on the ranks of the observa-

tions. Using ranks instead of the original observations has two major advantages over parametric procedures. First, ranks are maximal invariants under monotonic transformations of the data and, thus, their distribution does not change if a monotonic transformation is applied to the original data. Second, ranks are invariant with respect to the distribution if they are applied to a sequence of exchangeable observations. As a consequence, rank tests are robust against a wide class of "outliers". In this paper both advantages are exploited by two different rank procedures.

The first test is the natural rank counterpart of the conventional Dickey-Fuller test (Dickey and Fuller 1979). The observations of the series are replaced by their ranks and the first differences of the ranks are regressed on the sequence of ranks and a constant. This test was suggested by Granger and Hallman (1991) and its asymptotic properties are considered in Section 2. Clearly, the test is invariant with respect to a monotonic transformation of the data. However, since under the null the observations are not exchangeable, the null distribution depends on the distribution of the errors. Moreover, it is not clear how to generalise such a test to a random walk model with drift and correlated errors.

Campbell and Dufour (1995) suggest a different approach. Let $\{y_t\}_{t\in\{1,\ldots,T\}}$ denote the observed time series of sample size T. Instead of the original series the products $\tau_t = (y_t - y_{t-1})y_{t-1}$ are ranked and it is shown that for series generated by the simple random walk model $y_t = y_{t-1} + \epsilon_t$, where ϵ_t is a strong white noise series symmetrically distributed around zero, the Wilcoxon test for a zero median can be applied to τ_t. In contrast to the ranked Dickey-Fuller test this test is invariant with respect to the error distribution whereas a monotonic transformation of the time series will affect the null distribution, in general. Again, such a test has no straightforward

generalisation to random walk models with drift and correlated errors.

In section 3 a version of a rank test for a unit root is suggested which can also be applied for a random walk with drift and correlated errors. In the former case the exact finite sample distribution can be computed while in the latter case an asymptotic correction similar to the one suggested by Phillips and Perron (1988) can be used. This test is invariant with respect to the error distribution but is affected by a monotonic transformation of the data, in general. Moreover, the test can be seen to be robust against outliers and a structural break in the intercept of the trend function.

To assess the small sample properties of the test, several Monte Carlo experiments are performed. The results, which are presented in section 4, suggest that the ranked score test performs well in some situations but lacks power in others. Section 5 concludes.

2 The ranked Dickey-Fuller test

First we consider the ranked Dickey-Fuller (rDF) test as suggested by Granger and Hallman (1991). We assume the existence of a monotonic function, such that $z_t = h(y_t)$ is generated by a first order autoregressive model

$$z_t = \alpha z_{t-1} + \epsilon_t, \qquad t = 1, 2 \ldots, T,$$

where ϵ_t is an i.i.d. white noise sequence (i.e. "strong white noise") with p.d.f. f. It is important to consider the identification of the two functional parameters (h, f) and the scalar parameter α. Assume that there exist two representations $h(y_t) = \alpha h(y_{t-1}) + \epsilon_t$ and $h^*(y_t) = \alpha^* h^*(y_{t-1}) + \epsilon_t^*$, where h, h^* are two monotonic functions and ϵ_t, ϵ_t^* are strong white noise sequences. If h and h^* are diffentiable we deduce that

$$f[h(y_t) - \alpha h(y_{t-1})] \cdot \left| \frac{\partial h(y_t)}{\partial y} \right|$$

$$= \quad f^*[h^*(y_t) - \alpha^* h^*(y_{t-1})] \cdot \left| \frac{\partial h^*(y_t)}{\partial y} \right| \quad , \quad \forall y_t, y_{t-1}$$

by considering the form of the conditional p.d.f. of y_t given its past. It follows that h and h^* are in a one to one affine relationship and $\alpha^* = \alpha$.[1]

Under the null hypothesis it is assumed that the transformed time series z_t is generated by a random walk process with $\alpha = 1$, i.e.

$$H_0 = \{ \ \exists h \text{ monotonic}, \epsilon_t \text{ i.i.d.} : \ h(y_t) = h(y_{t-1}) + \epsilon_t \ \}$$

while the alternative is

$$H_1 = \{ \ \exists h \text{ monotonic}, |\alpha| < 1, \epsilon_t \text{ i.i.d.} : \ h(y_t) = \alpha h(y_{t-1}) + \epsilon_t \ \}$$

From the previous considerations it is seen that α is an identified parameter and, consequently, H_0 is an identifiable hypothesis on α.

A testing procedure invariant with respect to monotonic transformations is necessarily based on the ranks of $h(y_t)$ or, equivalently, on the ranks of y_t:

$$r_{t,T} = \text{Rank}[\text{ of } y_t \text{ among } y_0, \ldots, y_T \ ] - \frac{T+2}{2},$$

An ordinary Dickey-Fuller t-test can be applied to the sequence of ranks with

$$t_\alpha = \hat{\sigma}_*^{-1}(c_1 - c_0)/\sqrt{c_0},$$

where $c_j = \sum_{t=0}^{T-1} r_{t+j,T} r_{t,T}$ and $\hat{\sigma}_*^2 = (T-1)^{-1} \sum_{t=1}^{T} [r_{t,T} - (c_1/c_0) r_{t-1,T}]^2$. This version is slightly different from the test used by Granger and Hallman (1991), who include a constant in the regression in order to correct for the mean of the ranks. The difference is, however, negligible even in fairly small samples.

Since the ranks of (y_t) are the same as the ranks of (z_t) the distributional properties of the test statistic may be deduced from the ones of the strong

[1] A possible way to show this result is to make use of the fact that for a function $f(y - \alpha x)$ the partial derivatives with respect to x and y are proportional. To retain proportional derivatives of an alternative functional representation $g[h(y) - \alpha^* h(x)]$, it follows that the derivative of h must be constant and $\alpha = \alpha^*$.

random walk. The asymptotic theory for the distribution of the *parametric* Dickey-Fuller (DF) test is based on the well known fact that

$$T^{-1/2} z_{[aT]} \;\Rightarrow\; \sigma W(a) \tag{1}$$

where $\sigma^2 = \lim T^{-1} \sum E(\epsilon_t^2)$, $0 \leq a \leq 1$, $[\cdot]$ represents the integer part, "$\Rightarrow$" means weak convergence of the associated probability measure and $W(a)$ stands for the Brownian motion defined on $[0,1]$.

Unfortunately, this asymptotic result does not carry over to the sequence of ranks. From (1) we deduce that

$$\begin{aligned}
T^{-1} r_{[aT],T} &= T^{-1} \sum_t \mathbb{1}(z_t < z_{[aT]}) \\
&= T^{-1} \sum_t \mathbb{1}\left(\frac{1}{\sqrt{T}} z_t < \frac{1}{\sqrt{T}} z_{[aT]} \right) \\
&= \sum_t \mathbb{1}\left(\frac{1}{\sqrt{T}} z_{[\frac{t}{T}T]} < \frac{1}{\sqrt{T}} z_{[aT]} \right) \left[\frac{t}{T} - \frac{t-1}{T} \right] \\
&\Rightarrow R(a) = \int_0^1 \mathbb{1}[\sigma W(u) < \sigma W(a)] \mathrm{d}u = \int_0^1 \mathbb{1}[W(u) < W(a)] \mathrm{d}u.
\end{aligned}$$

Therefore the limit ranks define a stochastic process indexed by $a \in [0,1]$, such that $R(a)$ is the occupation time of the set $]-\infty, W(a)]$ by the Brownian motion.

We may now easily deduce some distributional properties of this process. The first property is well known and may be found for instance in Revuz and Yor (1991).

PROPERTY 1: *The distribution of* $R(0) = \int_0^1 \mathbb{1}[W(u) < 0] \mathrm{d}u$ *is the arcsine distribution with c.d.f.* $\frac{2}{\pi} arcsin[\sqrt{R(0)}]$, *and p.d.f.* $\frac{1}{\pi} \frac{1}{\sqrt{R(0)}\sqrt{1-R(0)}}$.

This distribution equals the beta distribution with parameters $(0.5, 0.5)$ and its p.d.f. is depicted in figure 1. Note that as $R(0)$ tends to zero or one the density goes to infinity.

insert Fig. 1 about here

With this result we can deduce the marginal distribution of $R(a)$, which is stated in the following property.

PROPERTY 2: *The distribution of $R(a)$ is the same as the distribution of $aR_1(0) + (1-a)R_2(0)$, where $R_1(0)$ and $R_2(0)$ are two independent random variables with an arcsine distribution.*

PROOF: We have:

$$\begin{aligned} R(a) &= \int_0^1 \mathbb{1}[W(u) < W(a)]\mathrm{d}u \\ &= \int_0^a \mathbb{1}[W(u) < W(a)]\mathrm{d}u + \int_a^1 \mathbb{1}[W(u) < W(a)]\mathrm{d}u. \end{aligned}$$

$\{W(u) < W(a)\ ,\ u < a\}$ and $\{W(u) < W(a)\ ,\ u > a\}$ are independent events since the Brownian motion is with independent increments. Then we have with obvious notation $R(a) = R_1(a) + R_2(a)$, where $R_1(a)$ and $R_2(a)$ are independent random variables. Now

$$\begin{aligned} R_1(a) &= \int_0^a \mathbb{1}[W(u) < W(a)]\mathrm{d}u \\ &= \int_0^a \mathbb{1}[W(a) - W(u) > 0]\mathrm{d}u \\ &\stackrel{d}{=} \int_0^a \mathbb{1}[W(a-u) > 0]\mathrm{d}u \\ &\stackrel{d}{=} \int_0^a \mathbb{1}[W(u) > 0]\mathrm{d}u \\ &\stackrel{d}{=} a\int_0^1 \mathbb{1}[W(u) > 0]\mathrm{d}u \\ &\stackrel{d}{=} aR(0), \end{aligned}$$

where $\stackrel{d}{=}$ stands for the equality in distribution. □

insert Fig. 2 about here

TABLE 1: Empirical sizes for the DF and the rDF test

	$T = 50$		$T = 100$		$T = 500$	
f	DF	rDF	DF	rDF	DF	rDF
$\mathcal{N}(0,1)$	0.055	0.054	0.053	0.059	0.048	0.075
$\chi^2(1)$	0.042	0.033	0.045	0.044	0.046	0.065
$t(2)$	0.061	0.036	0.055	0.042	0.052	0.052

Note: Rejection frequencies computed from 10,000 replications of the random walk model without drift. DF indicates the parametric Dickey-Fuller t_μ test including a constant term and rDF signifies the ranked Dickey-Fuller test for the mean adjusted ranks. $\chi^2(1)$ and $t(2)$ denote χ^2 and t distributed errors with one and two degrees of freedom, respectively.

In Appendix 1 some properties of the distribution of $R(a)$ are considered and Fig. 2 presents the p.d.f. for selected values of a. In order to assess the impact of these properties on the null distribution of the ranked Dickey-Fuller test, a Monte Carlo simulation was conducted. For the sample sizes $T = 50$, $T = 100$ and $T = 500$ we generate 10,000 artificial random walk paths with different error distributions and $z_t = y_t$. To represent skewed and heavy-tailed distributions we generate errors having a (centered) χ^2 distribution with one degree of freedom and a t distribution with two degrees of freedom. The critical values of a Dickey-Fuller test with a constant term and a significance level of 0.05 as tabulated in Fuller (1976) were used for both the parametric and the ranked test.

From the results presented in Table 1 it emerges that the actual size of the ranked Dickey-Fuller test depends on the error distribution in finite samples, while the asymptotic distribution of the ranks is independent of the distribution of the errors. It is seen that the large sample distribution appears to be (slightly) different from the one derived by Dickey and Fuller (1979). Even for $T = 500$, the application of Fuller's (1976) critical values yields a significant over-rejection of the null hypothesis, when the errors are normally

distributed. This was also observed by Granger and Hallman (1991).

It is important to note that in our simulations we have assumed h to be known. As was shown by Granger and Hallman (1991), the parametric Dickey-Fuller test applied to the raw series may perform poorly if the series is generated by a nonlinear unit root process. In particular, a misspecification of h affects the rate of divergence of the parametric Dickey-Fuller test statistic, so that the test may fail to be consistent.

To conclude, the analysis shows that the sequence of ranks on a random walk behave asymptotically as a weighted average of two independent random variables with arcsine distribution. Since this distribution does not depend on nuisance parameters, an asymptotic test can be constructed using the rank version of the Dickey–Fuller test. However, such a test possesses some undesirable properties. First, the results of the Monte Carlo simulations suggest that the limiting behaviour of the test may provide a poor approximation to the small sample properties. Second, there is no simple way to extend the test procedure to random walk models with drift and correlated errors. In the next section we therefore consider an alternative test procedure.

3 The ranked score test

In this section we consider a rank counterpart of the score statistic suggested by Schmidt and Phillips (1992). This test can be used to test the "difference stationary model" given by

$$y_t = b + \alpha y_{t-1} + \epsilon_t, \quad \text{with } \alpha = 1, \tag{2}$$

against the "trend stationary model"

$$y_t = c + bt + \alpha y_{t-1} + \epsilon_t, \quad \text{with } |\alpha| < 1. \tag{3}$$

In what follows we assume that the errors are independent and identically distributed with $E(\epsilon_t) = 0$ and c.d.f. F. As will be discussed below, it is possible to relax this assumption in order to allow for heteroskedastic or serially correlated errors. In these cases, however, it is no longer possible to obtain the exact null distribution.

As shown by Schmidt and Phillips (1992) the score (or Lagrange Multiplier) principle gives rise to the following statistic[2]

$$\tilde{\phi}_T = \frac{\sum_{t=2}^{T} x_t S_{t-1}}{\sum_{t=2}^{T} S_{t-1}^2}, \tag{4}$$

where

$$\begin{aligned} x_t &= \Delta y_t - \hat{b} \\ \hat{b} &= T^{-1} \sum_{t=1}^{T} \Delta y_t = (y_T - y_0)/T \\ S_t &= \sum_{i=1}^{t} x_i. \end{aligned}$$

Under the null hypothesis of a random walk with drift, $-(2T\tilde{\phi}_T)^{-1}$ is asymptotically distributed as $\int_0^1 \overline{W}(a)^2 \mathrm{d}a$, where $\overline{W}(a) = W(a) - aW(1)$ represents the standard Brownian bridge (cf Schmidt and Lee 1991).

Letting

$$\begin{aligned} \tilde{r}_{t,T} &= \text{Rank[of } \Delta y_t \text{ among } \Delta y_1, \ldots, \Delta y_T \text{]} - \frac{T+1}{2} \\ S_{t,T}^R &= \sum_{s=1}^{t} \tilde{r}_{s,T}, \end{aligned}$$

[2]There are two possible forms of the score statistic. Here we use the form (15B) of Schmidt and Phillips (1992). The form (15A) which they prefer is not appropriate to construct a rank statistic. The properties of the test according to (4) is also considered in Schmidt and Lee (1991).

a rank counterpart of the score statistic is

$$\tilde{\phi}_T^R = \frac{\sum_{t=2}^{T} \tilde{r}_{t,T} S_{t-1,T}^R}{\sum_{t=2}^{T} (S_{t-1,T}^R)^2}. \tag{5}$$

Note that the ranks of the observations are not affected by subtracting the mean of the series so that the mean of the differences $\hat{b}$ can be neglected.

The numerator of (5) does not depend on the data but is merely a function of T. Indeed

$$\sum_{t=2}^{T} \tilde{r}_{t,T} S_{t-1,T}^R = \sum_{i=2}^{T} \sum_{j=1}^{i-1} \tilde{r}_{i,T} \tilde{r}_{j,T}$$

so that the expression is the sum of the products given by all non-redundant combinations of the ranks. This quantity does not change by permutation of the ranks $\tilde{r}_{1,T}, \ldots, \tilde{r}_{T,T}$ and, thus,

$$\sum_{t=2}^{T} \tilde{r}_{t,T} S_{t-1,T}^R = \sum_{i=2}^{T} \sum_{j=1}^{i-1} ij,$$

which is a function of T only. As a consequence, it is sufficient to consider the statistic

$$\tilde{\lambda}_T = \sum_{t=1}^{T} (S_{t,T}^R)^2,$$

where we add $(S_{T,T}^R)^2$ in order to symmetrise the expression. As in the previous subsection, the ranked score statistic is defined by analogy with some standard statistic based on the y_t's themselves. The test statistic only depends on the ranks of the observed differences and, thus, it is invariant with respect to monotonic transformations of Δy_t. Similarly, the null hypothesis can be shown to have such an invariance property. Let us consider the strong random walk hypothesis corresponding to (2):

$$H_0^* \;=\; \{\ \exists \text{ strong white noise } \epsilon \text{ such that: } \Delta y_t = \epsilon_t\ \},$$

where ϵ_t is not necessarily centered. It is immediately seen that H_0^* satisfies the invariance property and, as a consequence, the ranked score statistic has a null distribution independent of F, even in finite samples.

It should be noted that the hypothesis H_0^* is strictly included in the hypothesis H_0 considered in the previous section. Therefore the two rank tests are associated with different implicit null hypotheses. In the following theorem, the limiting distribution of the test statistic is presented and several remarks are made which may be useful for practical applications.

THEOREM: *Let $\{\epsilon_1, \ldots, \epsilon_T\}$ be a sequence of independently and identically distributed random variables. Then, under H_0^* and $T \to \infty$:*

$$\frac{1}{T^4}\tilde{\lambda}_T \;\Rightarrow\; \frac{1}{12}\int_0^1 \overline{W}(a)^2 \mathrm{d}a$$

where $\overline{W}(a)$ denotes the standard Brownian bridge on $[0,1]$.

PROOF: Let the normalized ranks be defined as $\varrho_{t,T} = T^{-1}\tilde{r}_{t,T}$. We have $\tilde{\lambda}_T = \sum_{t=1}^{T}(S_{t,T}^R)^2 = \sum_{t=1}^{T}(\sum_{s=1}^{T}\tilde{r}_{s,T})^2 = T^2\sum_{t=1}^{T}(\sum_{s=1}^{T}\varrho_{s,T})^2$. Furthermore, let $\xi_{t,T}$ be a sequence of independent random variables with

$$P\Big\{\xi_{t,T} = T^{-1}[j-(T+1)/2]\Big\} = T^{-1} \qquad \text{for } j = 1,\ldots,T \text{ and } t = 1,\ldots,T.$$

Then, the distribution of the normalized ranks $(\varrho_{1,T},\ldots,\varrho_{T,T})$ is asymptotically equivalent to the distribution of $(\xi_{1,T},\ldots,\xi_{T,T})$ conditional on $\sum_{t=1}^{T}\xi_{t,T} = 0$, since the probability of equality of two variables $\xi_{t,T}$ and $\xi_{s,T}$ $(t \neq s)$ becomes asymptotically negligible. Asymptotically, $\xi_{t,T},\ldots,\xi_{T,T}$ tend to independently uniformly $U[-0.5, 0.5]$ distributed random variables variables and, thus, $T^{-1/2}\sum_{t=1}^{[aT]}\xi_{t,T} \Rightarrow (12)^{-1/2}W(a)$. Since the distribution of $T^{-1/2}\sum_{t=1}^{[aT]}\varrho_{t,T}$ is asymptotically the same as the distribution of $T^{-1/2}\sum_{t=1}^{[aT]}\xi_{t,T}$ conditioned by $T^{-1/2}\sum_{t=1}^{T}\xi_{t,T} = 0$, we have $T^{-1/2}\sum_{t=1}^{[aT]}\varrho_{t,T} \Rightarrow (12)^{-1/2}[W(a) - aW(1)] =$

$(12)^{-1/2}\overline{W}(a)$. Finally, it follows that $T^{-2}\sum_{t=1}^{T}(\sum_{s=1}^{T}\varrho_{s,T})^2 \Rightarrow (12)^{-1}\int_0^1 \overline{W}(a)^2 \mathrm{d}a$.
□

REMARK A: This result suggests to use the test statistic

$$\lambda_T(\text{uni}) = T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{t}\left(\sqrt{12}\varrho_{s,T}\right)^2, \tag{6}$$

where $\varrho_{t,T} = T^{-1}\tilde{r}_{t,T}$ is the normalized rank. Under the null hypothesis λ_T converges to $\int_0^1 \overline{W}(a)^2 \mathrm{d}a$. In Appendix 2 critical values for λ_T are given.

REMARK B: From the literature on nonparametric tests it is known that the test might be improved by using nonlinear transformations of the ranks such as the "inverse normal scores" (INS) transformation:

$$\tilde{r}_t^* = \Phi^{-1}\left(\varrho_{t,T} + 0.5\right), \tag{7}$$

where $\Phi(\cdot)$ is the c.d.f. of the standard normal distribution (cf van der Waerden 1952). Let us assume that we know a priori that (2) is satisfied with a Gaussian white noise, i.e. that we know that $F(z)$ is the standard normal c.d.f.. Then, the idea behind the INS transformation is that

$$\varrho_{t,T} = \widehat{F}_T(\Delta y_t) - \frac{T+1}{2T} \simeq \widehat{F}_T(\Delta y_t) - 0.5,$$

where $\widehat{F}_T(z)$ denotes the empirical distribution function of Δy_t given by $\widehat{F}_T(\delta) = T^{-1}\sum_{t=1}^{T}\mathbb{1}(\Delta y_t \leq \delta)$. Since $\widehat{F}_T$ converges pointwise to the distribution of the errors, $F(z)$, the transformation $F^{-1}[\varrho_{t,T} + 0.5]$ renders a series with similar distributional properties as Δy_t. Therefore, for normally distributed errors the INS transformation renders a test which behaves similar to the parametric test. Critical values for the modified statistic can be found in Appendix 2. Of course, this procedure may be extended by introducing any other transformation in order to approximate the inverse of the error

distribution as closely as possible. However, since in most applications the error distribution is unknown such a transformation is difficult to choose.

REMARK C: As usual in the literature on rank tests, the i.i.d. assumption on the errors may be relaxed by allowing for "exchangeable" errors, i.e., $F(\epsilon_1, \ldots, \epsilon_T) = F(\epsilon_{d_1}, \ldots, \epsilon_{d_T})$ for all permutations $d_1, \ldots, d_T$, where $F(\cdot)$ is the joint distribution function of $\epsilon_1, \ldots, \epsilon_T$ (e.g. McCabe 1989). This assumption would allow for the inclusion of a stochastic drift term. However, since the distribution is not affected by the way the drift term is specified, such a generalisation is of no practical use here.

REMARK D: It is easy to see that the test is robust against a structural break in the intercept of the trend function. For the parametric version of the test this property was noted by Amsler and Lee (1995). Assume that the value of the parameter c in model (3) changes to $c + \delta$ at time $1 < T_0 < T$. Then, under the null hypothesis we have $\Delta y_t = b + \epsilon_t$ for $t \neq T_0$ and $\Delta y_{T_0} = b + \delta + \epsilon_{T_0}$. If $\delta < \infty$ it is not difficult to verify that the limiting distribution given in the theorem is the same for all $-\infty < \delta < +\infty$. Hence the test may allow for a structural break in c.

Similarly it can be seen that the limiting distribution of the test does not change in the presence of the "additive" and the "innovative" outliers (see, e.g., Tsay (1988) for a discussion of this outlier classification).

REMARK E: It is possible to relax the assumption of i.i.d. errors in order to allow for a wide class of heteroskedastic and serially correlated errors. However, in such cases only asymptotic results are available. Since the sequence $\rho_{t,T} + 0.5$ converges to $F(\Delta y_t)$ as $T \to \infty$, the normalized ranks can be seen as a monotonic mapping of the differences of the original series onto $[0, 1]$. Under the assumptions given in Phillips (1987) it can be shown that

$\lambda_T \Rightarrow \sigma_\varrho^2 \int_0^1 \overline{W}(a)^2 \mathrm{d}a$, where $\sigma_\varrho^2 = \lim_{T\to\infty} E[T^{-1}(S_{T,T}^R)^2]$. A valid estimate for σ_ϱ^2 can be obtained by using sample autocovariances and applying a Newey-West weighting scheme, for example.

4 Small sample properties

In this section the results of several Monte Carlo simulations are presented in order to compare the small sample properties of the rank statistics with their parametric counterparts. The first test of the comparison is the (parametric) Dickey-Fuller t-statistic (DF) computed from a first order autoregression with a linear time trend as given in (3). Second, the score type statistic suggested by Schmidt and Phillips (SP) is computed as given in (4). For the DF statistic, the critical values of Fuller (1976) are used and for the SP test, we apply the critical values for the INS transformation given in table A of the appendix to $-(2T\tilde{\phi}_T)^{-1}$.

Two versions of the score type rank statistic are considered. First, we compute $\lambda_T(\text{uni})$ as in (6). Second, the ranks are transformed using the INS transformation and are used instead of the normalized ranks. The resulting test statistic is denoted by $\lambda_T(\text{INS})$. For these statistics , the critical values presented in Appendix 2 are used.

For all experiments the rejection frequencies are computed from 5000 Monte Carlo replications and the nominal significance level for the tests is 0.05.

4.1 Small sample properties for different error distributions

In the first Monte Carlo experiment we compute the rejection frequencies for the case that $z_t = y_t$ is a random walk with drift letting $\alpha = 1$ (the

TABLE 2: Rejection frequencies (T=50)

α	DF	SP	λ_T(uni)	λ_T(INS)
		$N(0,1)$		
1.00	0.056	0.050	0.048	0.049
0.90	0.081	0.097	0.097	0.092
0.80	0.190	0.230	0.200	0.220
		$\chi^2(2)$		
1.00	0.053	0.042	0.054	0.053
0.90	0.077	0.090	0.072	0.074
0.90	0.189	0.252	0.120	0.127
		$t(2)$		
1.00	0.053	0.045	0.049	0.047
0.90	0.079	0.088	0.062	0.074
0.80	0.174	0.250	0.101	0.137
		ADD$(T/2, 5\sigma)$		
1.00	0.315	0.227	0.065	0.085
0.90	0.445	0.360	0.122	0.151
0.80	0.684	0.603	0.247	0.310
		INNO$(T/2, 5\sigma)$		
1.00	0.049	0.051	0.051	0.045
0.90	0.082	0.105	0.084	0.087
0.80	0.214	0.257	0.158	0.181

Note: Rejection frequencies for the random walk with drift under different error distributions. $\chi^2(2)$ denotes the central χ^2 distribution with two degrees of freedom and $t(2)$ indicates the t distribution with two degrees of freedom. "DF" denotes the Dickey-Fuller t-statistic and "SP" is the Schmidt-Phillips statistic given in (4).

null hypothesis), as well as for the trend stationary model with $\alpha = 0.9$ and $\alpha = 0.8$ (the alternative). Three different error distributions are considered. First the normal distribution for which the parametric tests are asymptotically optimal. As an example for a skewed distribution we compute errors generated by a (centered) χ^2 distribution with two degrees of freedom, and t distributed errors with two degrees of freedom are used to represent a heavy tailed distribution.

To investigate the relative performance in the presence of outliers we

TABLE 3: Rejection frequencies (T=100)

α	DF	SP	λ_T(uni)	λ_T(INS)
		$N(0,1)$		
1.00	0.047	0.044	0.048	0.048
0.90	0.186	0.227	0.198	0.216
0.80	0.641	0.623	0.510	0.596
		$\chi^2(2)$		
1.00	0.050	0.044	0.048	0.050
0.90	0.192	0.236	0.099	0.092
0.80	0.660	0.663	0.208	0.228
		$t(2)$		
1.00	0.055	0.041	0.052	0.053
0.90	0.177	0.256	0.072	0.110
0.80	0.665	0.725	0.129	0.241
		ADD$(T/2, 5\sigma)$		
1.00	0.174	0.145	0.057	0.066
0.90	0.509	0.543	0.263	0.338
0.80	0.927	0.890	0.595	0.725
		INNO$(T/2, 5\sigma)$		
1.00	0.049	0.047	0.047	0.041
0.90	0.190	0.236	0.168	0.194
0.80	0.681	0.678	0.413	0.521

Note: See Table 2.

contaminate the data with normally distributed errors by an additive and an innovative outlier at $T_0 = T/2$ of size 5σ. The experiment with an additive outlier is indicated by ADD$(T/2, 5\sigma)$ and an innovative outlier in the data is indicated by INNO$(T/2, 5\sigma)$.

Tables 2 and 3 present the results for $T = 50$ and $T = 100$, respectively. It turns out that the score type test of Schmidt and Phillips (1992) is more powerful than the Dickey-Fuller test in small samples. Under normality, the power of the rank tests is similar to their parametric counterpart. For alternate distributions a different picture emerges. While the Dickey-Fuller test turns out to be quite robust against violations of the normality assumption,

TABLE 4: Empirical sizes for some nonlinear transformations

$h(y_t)$	DF	SP	λ_T(uni)	λ_T(INS)
$z_t = (y_t)^3$	0.402	0.211	0.049	0.062
$z_t = (y_t)^{1/3}$	0.159	0.196	0.058	0.081
$z_t = \ln(y_t)$	0.848	0.754	0.051	0.082
$z_t = \tan(y_t)$	0.421	0.172	0.047	0.067

Note: The series are generated by $z_t = h^{-1}(y_t)$ and $y_t = y_{t-1} + \epsilon_t$, where $\epsilon_t \sim N(0,1)$. The nominal size is 0.05 and the sample size is $T = 100$. The rejection frequencies are based on 5000 replications each. The test statistics are labeled as in Table 1.

the parametric Schmidt-Phillips statistic tends to be conservative.

In case of additive outliers the parametric tests possess a substantial size bias while the rank tests turn out to be much more robust. By contrast, in case of innovative outliers all tests have empirical sizes close to the nominal ones. The reason for this finding is that a random walk with an additive outlier can be seen as a random walk with two consecutive innovative outliers of the same size but different sign. Accordingly, the errors at these time periods are negatively correlated and imply a negative bias in the estimation of the autoregressive coefficient.

With respect to the performance under the alternative, the rank statistic reveals a substantial loss of power, in particular under heavy tailed distributions. The λ_T(INS) statistic performs better than λ_T(uni) but a considerable loss of power remains.

4.2 Small sample properties for some nonlinear transformations

The data are generated such that $z_t = h(y_t)$ is a random walk without a drift. We apply the (monotonic) transformations $z_t = (y_t)^3$, $z_t = (y_t)^{1/3}$, $z_t = \ln(y_t)$, and $z_t = \tan(y_t)$. The increments of z_t are normally distributed.

From the results in Table 4 it appears that the parametric tests suffer substantially from the misspecification of the process as a linear random walk, while the rank tests perform much better. As an extreme case, the parametric tests reject the unit root hypothesis for the logarithmic transformation in more than 70% of the cases while the actual sizes of the rank tests are close to the nominal ones.

On a less rigorous level the performance of the tests can be understood using a linear approximation for $y_t = h^{-1}(z_t)$ around z_{t-1}:

$$y_t \simeq y_{t-1} + [\mathrm{d}h(z_{t-1})/\mathrm{d}z_{t-1}]^{-1}\epsilon_t$$

and, thus, the approximation suggests that the first differences of the observed series behave approximately like conditionally heteroskedastic white noise. For the rank tests to be valid it is required that the differences behave like a strong white noise process. Although, in general this assumption is not satisfied here, it appears that for a wide range of possible transformations the sizes of the rank tests are affected only marginally.

4.3 The case of correlated errors

It is well known that parametric unit root tests run into serious problems, if the errors are generated by a MA process with a root close to one (Schwert 1989, Agiakloglou and Newbold 1992). In particular the Phillips-Perron test has been shown to suffer from a tremendous size bias in this case. Thus, it is interesting to consider the performance of the rank test in a model with MA(1) errors.

In order to correct for nuisance parameters both the parametric and the rank test employ the Newey-West weighting scheme to the covariances as suggested by Phillips (1987) and Phillips and Perron (1988). In analogy to the parametric tests the "long run variance" σ_ϱ^2 is estimated using the

TABLE 5: Rejection frequencies for MA(1) errors

β	DF-PP	SP	λ_T(uni)	λ_T(INS)
		$\alpha = 1.0$ (size)		
-0.50	0.013	0.007	0.005	0.004
0.00	0.063	0.037	0.035	0.036
0.20	0.218	0.104	0.107	0.111
0.40	0.594	0.289	0.288	0.314
0.60	0.951	0.644	0.595	0.663
0.80	1.000	0.943	0.846	0.925
		$\alpha = 0.8$ (power)		
-0.50	0.241	0.235	0.179	0.219
0.00	0.691	0.575	0.461	0.564
0.20	0.975	0.779	0.650	0.769
0.40	1.000	0.915	0.789	0.899
0.60	1.000	0.970	0.866	0.949
0.80	1.000	0.984	0.900	0.968

Note: Rejection frequencies from 5000 replication of model (1) with MA(1) errors generated by $\epsilon_t = u_t - \beta u_{t-1}$, where u_t is generated by a Gaussian white noice process. The sample size is T=100 and the nominal significance level is 0.05.

residuals from a regression of the ranks $\tilde{r}_{t,T}$ on $S^R_{t-1,T}$. Table 5 presents the rejection frequencies of the ranked score statistic for $\alpha = 1$ (i.e. the actual size) and $\alpha = 0.8$ (i.e. the empirical power) and using normally distributed errors. The truncation lag for the Phillips-Perron type of correction is eight. It turns out that the Phillips-Perron test (DF-PP) is seriously biased even for moderate values of the moving average parameter β. The Schmidt-Phillips test performs slightly better although it still has a severe size distortion for $\beta \geq 0.4$. With respect to the size, the rank versions of the Schmidt-Phillips test perform similar to the parametric counterpart.

For trend stationary alternatives with $\alpha = 0.8$, the power of the rank statistic λ_T(INS) is somewhat smaller than the parametric counterpart, but the differences are not very large for moderate values of β. In summary, even

in the case of a known transform h, the rank tests appear to perform roughly similar to the parametric tests so that theses tests should be used with care whenever the errors have a substantial negative correlation.

5 Concluding remarks

In this paper several aspects of rank tests are addressed. First, we consider a rank version of the Dickey-Fuller test suggested by Granger and Hallman (1991) to test the hypothesis that the series is a monotonic transformation of a random walk. It is shown, that the sequence of ranks built from the levels of the time series does not converge to a Brownian motion. Consequently, the asymptotic properties of the ranked Dickey-Fuller test are different from its parametric counterpart.

Second, for testing the null hypothesis of a strong random walk with drift a new rank test is suggested which can be seen as a ranked counterpart of the score test suggested by Schmidt and Phillips (1992). An asymptotic correction is suggested in order to account for serially correlated errors. We also consider a modification of the rank statistic by transforming the ranks using the inverse of the normal c.d.f.. This approach, known as "inverse normal scores" transformation, seems to improve the power of the test in many cases.

Concluding the results of the Monte Carlo experiments it turns out that if the transform h is known and the errors are i.i.d. there may be a substantial loss in power when using ranks instead of the original observations. While the size of the rank test is robust against departures from the normality assumption, the power of the test may deteriorate considerably when the error distribution is skewed or heavy tailed. In order to improve the robustness with respect to the power, alternative transformations of the ranks may be

useful. For example, applying the inverse of a heavy tailed distribution, such as the t-distribution with a small number of degrees of freedom, may help to improve the power of the rank test.

On the other hand, when the monotonic transform h is unknown or the errors are correlated, the parametric tests may perform poorly and the ranked score tests are attractive competitors to standard testing procedures. Accordingly, rank tests may serve as a useful tool for specifying nonlinear time series processes. Once the unit root hypothesis is accepted for the transformed process $z_t = h(y_t)$, it is interesting to estimate the function h from the data. For a Markov process of order one we have $E[h(y_t) - h(y_{t-1})|y_{t-1}, y_{t-2}, \ldots] = 0$ and, thus, the function h is an eigenfunction of the transition operator $h(\cdot) \longrightarrow E[h(y_t)|y_{t-1}, y_{t-2}, \ldots]$ associated with a unit eigenvalue (see Conley, Hansen, Luttmer, Scheinkman (1995) for a related problem arising in the estimation of stochastic diffential equations). Such problems, however, are beyond the scope of this paper and are left for research.

Appendix 1: The distribution of $R(a)$

(i) Symmetry with respect to 1/2

By the symmetry of the arcsine distribution we have:

$$\begin{aligned} R(a) &\stackrel{d}{=} aR_1(0) + (1-a)R_2(0) \\ &\stackrel{d}{=} a[1-R_1(0)] + (1-a)[1-R_2(0)] \\ &\stackrel{d}{=} 1 - [aR_1(0) + (1-a)R_2(0)] \\ &\stackrel{d}{=} 1 - R_1(a), \end{aligned}$$

which is the required result.

(ii) Form of the p.d.f. of R(a)

We get:

$$\begin{aligned} f(y) &\simeq \frac{1}{\mathrm{d}y} P\{R(a) \in [y, y+\mathrm{d}y]\} \\ &= \frac{1}{\mathrm{d}y} P\{aR_1(0) + (1-a)R_2(0) \in [y, y+\mathrm{d}y]\} \\ &= \frac{1}{\mathrm{d}y} P\left\{R_1(0) \in \left[\frac{y-(1-a)R_2(0)}{a}, \frac{y-(1-a)R_2(0)}{a} + \frac{\mathrm{d}y}{a}\right]\right\} \\ &= \int_{max[0,(y-a)/(1-a)]}^{min[1,y/(1-a)]} \frac{1}{\pi^2} \frac{1}{\sqrt{z}\sqrt{1-z}} \frac{1}{\sqrt{\frac{y-(1-a)z}{a}}} \frac{1}{\sqrt{1-\frac{y-(1-a)z}{a}}} \frac{\mathrm{d}z}{a} \ . \end{aligned}$$

The plots of $f(y)$ presented in Fig. 2 are computed by numerical integration.

Let us assume $a < 1/2$. The p.d.f. takes different forms depending on the position of the argument y. Three cases have to be considered: $y < a$, $a < y < 1-a$, and $y > 1-a$, where the third one is the symmetric counterpart of the first one from (i).

Case: $y < a$

We get:

$$f(y) = \int_0^{y/(1-a)} \frac{1}{a\pi^2} \frac{1}{\sqrt{z}\sqrt{1-z}} \frac{1}{\sqrt{\frac{y-(1-a)z}{a}}} \frac{1}{\sqrt{1-\frac{y-(1-a)z}{a}}}\, \mathrm{d}z.$$

Let us introduce the change of variable: $u = (1-a)z/y$. We obtain:

$$f(y) = \frac{1}{\pi^2 y} \int_0^1 \frac{1}{\sqrt{u}\sqrt{1-u}} \frac{1}{\sqrt{\frac{1-a}{y} - u}} \frac{1}{\sqrt{\frac{a-y}{y} + u}}\, \mathrm{d}u$$

Case: $a < y < 1-a$

We get:

$$f(y) = \int_{(y-a)/(1-a)}^{y/(1-a)} \frac{1}{a\pi^2} \frac{1}{\sqrt{z}\sqrt{1-z}} \frac{1}{\sqrt{\frac{y-(1-a)z}{a}}} \frac{1}{\sqrt{1-\frac{y-(1-a)z}{a}}}\, \mathrm{d}z$$

Let us introduce the change of variable: $z = \frac{y-a}{1-a} + u\frac{a}{1-a}$. Then,

$$f(y) = \frac{1}{a\pi^2} \int_0^1 \frac{1}{\sqrt{u}\sqrt{1-u}} \frac{1}{\sqrt{u + (y-a)/a}} \frac{1}{\sqrt{(1-y)/a - u}}\, \mathrm{d}u$$

(iii) Behaviour of the p.d.f. in a neighbourhood of $y = a$

It is directly seen that if y is close to a we get:

$$f(y) \simeq \frac{1}{a\pi^2} \int_0^1 \frac{1}{u} \frac{1}{\sqrt{1-u}} \frac{1}{\sqrt{(1-a)/a - u}}\, \mathrm{d}u,$$

which is equal to $+\infty$.

(iv) Behaviour of the p.d.f. for $y = 0$

For this case we get:

$$\begin{aligned} f(y) &= \frac{1}{\pi^2}\int_0^1 \frac{1}{\sqrt{u}\sqrt{1-u}}\frac{1}{\sqrt{a}\sqrt{1-a}}\mathrm{d}u \\ &= \frac{1}{\pi}\frac{1}{\sqrt{a}\sqrt{1-a}}. \end{aligned}$$

Appendix 2: Computation of critical values

The critical values for the statistic using uniform ranks, $\tilde{\lambda}_T(\text{uni})$ and the "inverse normal scores transformation" $\tilde{\lambda}_T(\text{INS})$ are computed from 10,000 Monte Carlo replications of the random walk model (2) with $\epsilon_t \sim$IN(0,1) for a grid of 116 sample sizes letting $T = 15, 16, 17, ..., 100, 105, 110, ..., 250$.

For the estimated critical values the following regression model was fitted by OLS:

$$c_T^\alpha = b_0 + b_1 \frac{1}{\sqrt{T}} + b_2 \frac{1}{T} + e_T,$$

where c_T^α denotes the critical value for the significance values α =0.01, 0.05, 0.10 observed at sample size T. Table A presents the fitted values $\hat{c}_T^\alpha = \hat{b}_0 + \hat{b}_1 \frac{1}{\sqrt{T}} + \hat{b}_2 \frac{1}{T}$ for selected sample sizes.

TABLE A: Critical values

			λ(uni)				
T	0.01	0.05	0.10	T	0.01	0.05	0.10
20	0.0296	0.0404	00491	70	0.0261	0.0376	0.0469
25	0.0286	0.0397	0.0486	80	0.0260	0.0375	0.0468
30	0.0279	0.0392	0.0482	90	0.0258	0.0374	0.0467
35	0.0274	0.0388	0.0479	100	0.0257	0.0373	0.0466
40	0.0271	0.0385	0.0477	150	0.0254	0.0370	0.0463
45	0.0268	0.0383	0.0475	200	0.0253	0.0368	0.0462
50	0.0266	0.0381	0.0473	250	0.0252	0.0367	0.0461
60	0.0263	0.0378	0.0471	∞	0.0250	0.0362	0.0453

			λ(INS)				
T	0.01	0.05	0.10	T	0.01	0.05	0.10
20	0.0311	0.0424	00516	70	0.0264	0.0381	0.0475
25	0.0298	0.0412	0.0504	80	0.0262	0.0379	0.0473
30	0.0289	0.0404	0.0496	90	0.0260	0.0377	0.0471
35	0.0283	0.0400	0.0490	100	0.0259	0.0376	0.0470
40	0.0278	0.0393	0.0486	150	0.0255	0.0373	0.0467
45	0.0274	0.0390	0.0483	200	0.0253	0.0371	0.0466
50	0.0271	0.0388	0.0481	250	0.0252	0.0370	0.0465
60	0.0267	0.0384	0.0477	∞	0.0250	0.0368	0.0464

Note: Entries report estimated critical values for the significance levels 0.01, 0.05, 0.10 resulting from the response surface analysis. λ(uni) and λ(INS) denote the rank statistic based on the original ranks and the "inverse normal scores" transformation given in (7). The estimated standard deviations are roughly 0.0005.

References

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

Amsler, C. and Lee, J., 1995, An LM test for a unit root in the presence of a structural break, Econometric Theory 11, 359–368.

Campbell, B. and Dufour, J.-M., 1995, Exact nonparametric orthogonality and random walk tests, The Review of Economics and Statistics, 77, 1–16.

Conley, T.G., Hansen, L.P., Luttmer, E.G.J. and Scheinkman, J.A., 1995, Short-term interest rates as subordinated diffusions, manuscript.

Davidson, R. and MacKinnon, J.G., 1993, Estimation and inference in econometrics, New York: Oxford University Press.

Franses, P.H. and McAleer, M., 1995, Testing for unit roots and non-linear transformations, Paper presented at the 7th World Congress of the Econometric Society, Tokyo.

Granger C.W.J. and J. Hallman, 1991, Nonlinear transformations of integrated time series, Journal of Time Series Analysis 12, 207–224.

McCabe, B.P.M., 1989, Misspecification tests in econometrics based on ranks, Journal of Econometrics 40, 261–278.

Perron, P., 1989, The great crash, the oil price shock, and the unit root hypothesis, Econometrica 57, 1361–1401.

Phillips, P.C.B., 1987, Time series regressions with a unit root, Econometrica 55, 277–301.

Phillips, P.C.B. and Perron, P., 1988, Testing for a unit root in time series regressions, Biometrika 75, 335–346.

Revuz, D. and Yor, M., 1991, Continuous martingales and Brownian motion (Springer, New York).

Schmidt, P. and J. Lee, 1991, A modification of the Schmidt-Phillips unit root test, Economics Letters 36, 285–289.

Schmidt, P. and P.C.B. Phillips, 1992, LM test for a unit root in the presence of deterministic trends, Oxford Bulletin of Economics and Statistics 54, 257–287.

Schwert, G.W., 1989, Tests for unit roots: a Monte Carlo investigation, Journal of Business and Economic Statistics 7, 147–159.

Tsay, R.S., 1988, Outliers, level shifts, and variance changes in time series, Journal of Forecasting, 7, 1–20.

van der Waerden, B.L., 1952, Order tests for the two-sample problem and their power, Proceedings Koninklijke Nederlandse Akademie van Wetenschappen (A) 55, 453–458.