Rank Tests for Nonlinear Cointegration

Jörg Breitung
Humboldt-Universität zu Berlin, Institut für Statistik und Ökonometrie, D-10178 Berlin, Germany
(*breitung@wiwi.hu-berlin.de*)

A test procedure based on ranks is suggested to test for nonlinear cointegration. For two (or more) time series it is assumed that monotonic transformations exist such that the normalized series can asymptotically be represented as Wiener processes. Rank-test procedures based on the difference between the sequences of ranks are suggested. If there is no cointegration between the time series, the sequences of ranks tend to diverge, whereas under cointegration the sequences of ranks evolve similarly. Monte Carlo simulations suggest that for a wide range of nonlinear models the rank tests perform better than their parametric competitors. To test for nonlinear cointegration, a variable addition test based on ranks is suggested. In an empirical illustration, the rank statistics are applied to test the relationship between bond yields with different times to maturity.

KEY WORDS: Arcsine distribution; Stochastic trends; Unit roots.

Since the introduction of the concept of cointegration by Granger (1981), the analysis of cointegrated models has been intensively studied in a linear context, whereas the work on the extension to nonlinear cointegration is still comparatively limited. Useful reviews of recent developments in the analysis of nonlinear cointegration were provided by Granger and Teräsvirta (1993), Granger (1995), and Granger, Inoue, and Morin (1997).

In many cases, economic theory suggests a nonlinear relationship, as for the production function or the Phillips curve, for example. However, theory does not always provide a precise specification of the functional form so that it is desirable to have nonparametric tools for estimation and inference. In this article, rank-test procedures are considered to test whether a (possible nonlinear) cointegration relationship among the variables exists and whether this relationship is in fact nonlinear.

To illustrate the problems of nonlinear cointegration, it is helpful to consider a simple example. Let x_t be a nonlinear random walk given by $x_t^3 = \sum_{j=1}^{t} v_j$, where $\{v_t\}_1^T$ is a white-noise sequence with $v_t \sim \mathrm{N}(0, \sigma_v^2)$. Furthermore, y_t is given by

$$y_t = x_t^3 + u_t \tag{1}$$

where u_t is white noise with $u_t \sim \mathrm{N}(0, \sigma_u^2)$. Figure 1 presents a realization of the sequences x_t and y_t, where $\sigma_v^2 = \sigma_u^2 = .01$. The sample size is $T = 200$. Apparently, there is a fairly strong comovement between both series suggesting a stable, long-run relationship. However, applying an augmented Dickey–Fuller test (with four lagged differences and a constant) to the residuals of the linear cointegration regression yields a t statistic of -2.77, which is insignificant with respect to the .05 significance level. In contrast, applying the rank tests suggested in Section 2 yields significant test statistics with respect to all reasonable significance levels. This example illustrates that ignoring the nonlinear nature of the cointegration relationship may lead to the misleading conclusion that no long-run relationship exists between the series.

On the other hand, one may argue that there is no problem with a test that fails to reject in the presence of a nonlinear alternative because I am interested in detecting a *linear* cointegrating relationship. In many applications, however, it is not clear whether the variables must be transformed (e.g., to logarithms) to achieve a linear cointegrating relationship (cf. Franses and McAleer 1998), and thus the robustness of the test against such monotonic transformation is a desirable property of a cointegration test.

The rest of the article is organized as follows. In Section 1, I demonstrate that the parametric cointegration test based on the residuals of a linear cointegration regression is inconsistent against a given class of nonlinear alternatives. Two test statistics based on ranks are suggested in Section 2, and in Section 3 the power of the tests is considered. Generalizations of the test procedures are proposed in Section 4, and a rank test for neglected nonlinearities in the cointegration relationship is suggested in Section 5. Section 6 presents some results on the small-sample properties of the test procedures, and Section 7 provides an application of the rank tests to the term structure of interest yields. Finally, Section 8 concludes.

1. THE PROPERTIES OF LINEAR COINTEGRATION TESTS

For a theoretical analysis of a nonlinear cointegration relationship, different concepts are used. Granger and Hallman (1991a) and Granger (1995) considered time series that are long memory in mean but have a nonlinear relationship that is short memory in mean. Corradi (1995) employed the concept of nonstrong mixing processes (processes with a long memory) and strong mixing processes (short memory). In this article, I adopt the definition of an integrated process due to Davidson (1998). The degree of integration is defined as follows:

Definition 1. (1) A time series z_t is I(0) if, as $T \to \infty$,

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[rT]} z_t \Rightarrow \bar{\sigma}_z W(r),$$

where $\bar{\sigma}_z^2 = \lim_{T\to\infty} E(T\bar{z}^2)$, $\bar{z} = T^{-1}\sum_{t=1}^{T} z_t$, and $W(r)$ represents a standard Brownian motion. (2) If $(1-B)^d z_t \sim \mathrm{I}(0)$, then z_t is integrated of order d, denoted as $z_t \sim \mathrm{I}(d)$.

Journal of Business & Economic Statistics
July 2001, Vol. 19, No. 3

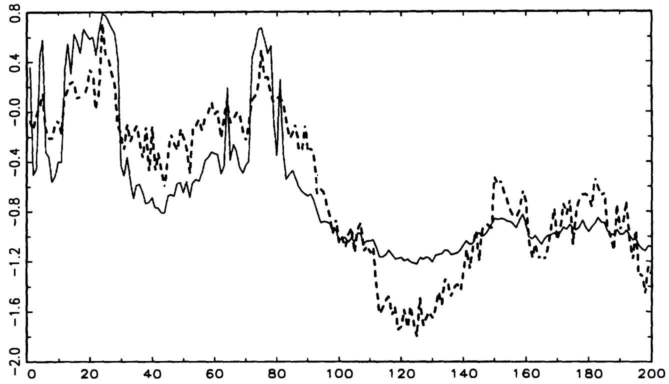

Figure 1. Realization of y_t and x_t Generated as in (1): ——— , x(t); . . . ,y(t).

Different sets of sufficient conditions for $z_t \sim \mathrm{I}(0)$ were given by Phillips (1987), Phillips and Solo (1992), and Gallant and White (1988).

In this section, I shall consider the nonlinear relationship between two real-valued time series $\{x_t\}_1^T$ and $\{y_t\}_1^T$,

$$y_t = f(x_t) + u_t, \tag{2}$$

where $y_t \sim \mathrm{I}(1)$ and $z_t = f(x_t) \sim \mathrm{I}(1)$. Under the null hypothesis, u_t is assumed to be I(1), whereas under the alternative of a cointegration relationship, $u_t \sim \mathrm{I}(0)$. A similar framework was used by Abadir and Taylor (1999).

As demonstrated by Granger and Hallman (1991b), the Dickey–Fuller test may perform poorly when it is applied to a nonlinear transformation of a random walk. To investigate the effects of a nonlinear cointegration relationship on the power of a residual-based cointegration test, it is convenient to consider a variant of the Dickey–Fuller type of test due to Sargan and Bhargava (1983) and Phillips and Ouliaris (1990). The statistic is given by

$$S_T^2 = \frac{1}{\omega_{11\cdot 2} T^2} \sum_{t=1}^{T} (y_t - \hat{\beta} x_t)^2, \tag{3}$$

where the $\hat{\beta}$ is the least squares estimator from a regression of y_t on x_t. The parameter $\omega_{11\cdot 2}$ was defined by Phillips and Ouliaris (1990, below eq. 12).

In the following theorem, it is stated that a test against linear cointegration may be inconsistent for some class of nonlinear functions. Further results can be obtained by using the framework of Park and Phillips (1999). However, since the latter approach requires some special concepts that are beyond the scope of this article, I shall confine myself to a simple class of nonlinear functions, which includes the function $y_t = x_t^a$ as a special case.

Theorem 1. Assume that the function $f(x)$ is monotonically increasing and that a function $h(a)$ exists such that $f^{-1}(az) = h(a) f^{-1}(z)$, where $f^{-1}(z)$ denotes the inverse function. Furthermore, $z_t = f(x_t) = \sum_{i=1}^{t} v_i$ is integrated of order 1 and y_t is generated as in (2), where $u_t \sim \mathrm{I}(0)$. Then, a test based on the statistic S_T^2 given in (3) is consistent iff $f(x)$ is a linear function.

Proof. Using

$$\frac{1}{T^2}\sum_{t=1}^{T} \hat{u}_t^2 = \frac{1}{T^2}\sum_{t=1}^{T} y_t^2 - \frac{\left(T^{-2}\sum_{t=1}^{T} y_t x_t\right)^2}{T^{-2}\sum_{t=1}^{T} x_t^2},$$

the test is seen to be consistent if the difference on the right side converges to 0 as $T \to \infty$. From the continuous mapping theorem, it follows that

$$T^{-2}\sum_{t=1}^{T} y_t^2 = T^{-2}\sum_{t=1}^{T} (z_t + u_t)^2 \Rightarrow \bar{\sigma}_z^2 \int_0^1 W(r)^2 dr. \tag{4}$$

Furthermore,

$$\begin{aligned} T^{-2}\sum_{t=1}^{T} x_t^2 &= T^{-2}\sum_{t=1}^{T} f^{-1}(z_t)^2 \\ &= h(\bar{\sigma}_z^{-1} T^{-1/2})^{-2} T^{-2} \sum_{t=1}^{T} f^{-1}(\bar{\sigma}_z^{-1} T^{-1/2} z_t)^2 \\ &\Rightarrow h(\bar{\sigma}_z^{-1} T^{-1/2})^{-2} \int_0^1 f^{-1}[W(r)]^2 dr \end{aligned}$$

and

$$\begin{aligned} T^{-2}\sum_{t=1}^{T} y_t x_t &= T^{-2}\sum_{t=1}^{T} z_t f^{-1}(z_t) + o_p(1) \\ &= h(\bar{\sigma}_z^{-1} T^{-1/2})^{-1} T^{-2} \sum_{t=1}^{T} z_t f^{-1}(\bar{\sigma}_z^{-1} T^{-1/2} z_t) \\ &\Rightarrow \bar{\sigma}_z h(\bar{\sigma}_z^{-1} T^{-1/2})^{-1} \int_0^1 W(r) f^{-1}[W(r)] dr. \end{aligned}$$

It follows that

$$\begin{aligned} \frac{1}{T^2}\sum_{t=1}^{T} \hat{u}_t^2 \Rightarrow \bar{\sigma}_z^2 \int_0^1 W(r)^2 dr - \bar{\sigma}_z^2 \\ \times \frac{\left\{\int_0^1 W(r) f^{-1}[W(r)] dr\right\}^2}{\int_0^1 f^{-1}[W(r)]^2 dr}. \end{aligned} \tag{5}$$

Since $x = f^{-1}(z)$ is an affine mapping, it is seen that the right side of (5) is 0 iff $f(x) = bx$ with some constant b.

This theorem shows that residual-based cointegration tests are inconsistent at least for the class of functions considered here.

An example may help to illustrate the result. Let y_t be generated as in (1). Using (5) and $f^{-1}(z) = x^{1/3}$, I have, for $\beta \neq 0$,

$$S_T^2 \Rightarrow \int_0^1 W(r)^2 dr - \frac{[\int_0^1 W(r)^{4/3} dr]^2}{\int_0^1 W(r)^{2/3} dr}.$$

Thus, under the alternative of nonlinear cointegration, the test statistic is $O_p(1)$. Accordingly, a test based on S_T is inconsistent against nonlinear alternatives as given in (1).

2. RANK TESTS FOR COINTEGRATION

To overcome the difficulties of standard unit-root tests in detecting nonlinear cointegration, I apply a rank transformation to the time series. The resulting test is valid for more general situations as considered in the previous section. Specifically, I assume under the alternative that a nonlinear cointegration relationship exists, given by

$$u_t = g(y_t) - f(x_t), \tag{6}$$

where $f(x_t) \sim \mathrm{I}(1)$, $g(y_t) \sim \mathrm{I}(1)$, and $u_t \sim \mathrm{I}(0)$. The functions $g(y)$ and $f(x)$ are monotonically increasing. If it is not known whether these functions are monotonically increasing or decreasing, a two-sided test is available. A similar framework was considered by Granger and Hallman (1991a).

I define the ranked series as $R_T(x_t) = \text{Rank}$ [of x_t among $x_1, \dots, x_T$] and construct $R_T(y_t)$ accordingly. Breitung and Gouriéroux (1997) developed an asymptotic theory for a ranked random walk. Here I give a slightly more general version of their main result, which allows for serial correlation of the increments v_t.

Theorem 2. Let $x_t = \mu + \sum_{i=1}^{T} v_i$ be I(1) as defined in Definition 1. Then, as $T \to \infty$, the limiting distribution of the sequence of ranks can be represented as $T^{-1}R_T(x_{[aT]}) \Rightarrow a\mathcal{A}_1 + (1-a)\mathcal{A}_2$, where $\mathcal{A}_1$ and $\mathcal{A}_2$ are two independent random variables with an arcsine distribution.

Proof. Consider

$$\begin{aligned} T^{-1}R_T(x_{[aT]}) &= T^{-1}\sum_{t=1}^{T} \mathbb{1}(x_t < x_{[aT]}) \\ &= \sum_t \mathbb{1}\left(\frac{1}{\sqrt{T}}z_{[\frac{t}{T}T]} < \frac{1}{\sqrt{T}}z_{[aT]}\right)\left[\frac{t}{T} - \frac{t-1}{T}\right] \\ &\Rightarrow \int_0^1 \mathbb{1}[W(u) < W(a)]du \\ &= \int_0^a \mathbb{1}[W(u) < W(a)]du \\ &\quad + \int_a^1 \mathbb{1}[W(u) < W(a)]du. \end{aligned}$$

Since the increments of the Brownian motion are independent, the two parts of the integral are independent as well.

Using $\stackrel{d}{=}$ to indicate equality in distribution, I have

$$\begin{aligned} \int_0^a \mathbb{1}[W(u) < W(a)]du &= \int_0^a \mathbb{1}[W(a) - W(u) > 0]du \\ &\stackrel{d}{=} \int_0^a \mathbb{1}[W(a-u) > 0]du \\ &\stackrel{d}{=} a\int_0^1 \mathbb{1}[W(u) > 0]du \\ &\stackrel{d}{=} a\mathcal{A}_1, \end{aligned}$$

where $\mathcal{A}_1 = \int_0^1 \mathbb{1}[W(u) > 0]du$ is a random variable with an arcsine distribution (cf. Breitung and Gouriéroux 1997). Similarly, I find

$$\int_a^1 \mathbb{1}[W(u) < W(a)]du = (1-a)\mathcal{A}_2,$$

where $\mathcal{A}_2$ is another random variable with an arcsine distribution independent of $\mathcal{A}_1$.

The rank statistic is constructed by replacing $f(x_t)$ and $g(y_t)$ by the ranked series

$$R_T[f(x_t)] = R_T(x_t) \tag{7}$$

and

$$R_T[g(y_t)] = R_T(y_t). \tag{8}$$

The attractive property of the rank transformation is that the sequence of ranks is invariant to a monotonic transformation of the data. Therefore, if x_t (or y_t) is a random walk, then $R_T[f(x_t)]$ behaves like the ranked random walk $R_T(x_t)$.

I consider two "distance measures" between the sequences $R_T(x_t)$ and $R_T(y_t)$ that are similar to the ones used by Acosta and Granger (1995):

$$\kappa_T = T^{-1}\sup_t |d_t| \tag{9}$$

and

$$\xi_T = T^{-3}\sum_{t=1}^{T} d_t^2, \tag{10}$$

where $d_t = R_T(y_t) - R_T(x_t)$. Note that d_t is $O_p(T)$ and, thus, the normalization factors are different from other applications of these measures. The statistic κ_T is a Kolmogorov–Smirnov type of statistic considered by Lo (1991) and ξ_T is a Cramer–von Mises type of statistic used by Sargan and Bhargava (1983). The null hypothesis of no (nonlinear) cointegration between x_t and y_t is rejected if the test statistics are too small.

The statistic ξ_T is related to the rank correlation coefficient, which is known as "Spearman's rho." This statistic is defined as

$$r_s = 1 - \frac{6}{T^3 - T}\sum_{t=1}^{T} d_t^2 \tag{11}$$

(e.g., Kendall and Gibbons 1990, p. 8). The statistic r_s can therefore be seen as a mapping of ξ_T into the interval $[-1, 1]$. If x_t and y_t are cointegrated, Spearman's rho converges in probability to 1 as $T \to \infty$.

It is also interesting to relate the statistic ξ_T to the least squares estimate $\tilde{b}_T$ from a regression of $R_T(y_t)$ on $R_T(x_t)$. Using $\sum R_T(x_t)^2 = \sum R_T(y_t)^2 = T^3/3 + O(T^2)$, I have

$$\begin{aligned} \xi_T &= \frac{1}{T^3}\sum_{t=1}^{T}[R_T(y_t)^2 - 2R_T(y_t)R_T(x_t) + R_T(x_t)^2] \\ &= \frac{2 - 2\tilde{b}_T}{T^3}\sum_{t=1}^{T} R_T(x_t)^2 \\ &= \frac{2}{3}(1 - \tilde{b}_T) + o_p(1). \end{aligned}$$

If y_t and x_t are not cointegrated, then $\tilde{b}_T$ has a nondegenerate limiting distribution [see Phillips (1986) for the linear case]. On the other hand, if y_t and x_t are cointegrated, then $\tilde{b}_T$ converges in probability to 1 and therefore ξ_T converges to 0.

Table 1. Critical Values

T	.10	.05	.01
κ	.6442	.5524	.4220
ξ	.0573	.0423	.0238
κ^*	.3941	.3635	.3165
ξ^*	.0232	.0188	.0130
$\Xi^*[1]$	.0248	.0197	.0136
$\Xi^*[2]$	.0197	.0165	.0119
$\Xi^*[3]$	.0160	.0137	.0100
$\Xi^*[4]$	.0136	.0117	.0092
$\Xi^*[5]$	.0118	.0104	.0083
$\Xi^*[6]$	.0104	.0093	.0077

NOTE: Critical values are computed from 10,000 realizations of independent random-walk sequences with $T = 500$.

A two-sided version of the test statistic can be constructed using the residuals of a cointegration regression on the ranks:

$$\begin{aligned}\Xi_T &= \frac{1}{T^3}\sum_{t=1}^{T}[R_T(y_t) - \tilde{b}_T R_T(x_t)]^2 \\ &= \frac{1}{T^3}\sum_{t=1}^{T}[R_T(y_t)^2 - 2\tilde{b}_T R_T(y_t)R_T(x_t) + \tilde{b}_T^2 R_T(x_t)^2] \\ &= \frac{1-\tilde{b}_T^2}{T^3}\sum_{t=1}^{T} R_T(x_t)^2 \\ &= \frac{1}{3}(1-\tilde{b}_T^2) + o_p(1).\end{aligned}$$

Since this test statistic is a function of the squared regression coefficient, the test rejects the null hypothesis if $\tilde{b}_T$ tends to plus or minus infinity.

The critical values for the statistics ξ_T and κ_T are computed by using 10,000 Monte Carlo replications of the test statistics with $T = 500$. From the results presented in Table 1, it turns out that the critical values corresponding to different significance levels are fairly close. This suggests that a slight change of the test statistic (for example, due to an outlier in the data) may have an important effect on the test decision. On the other hand, test statistics computed from the ranks of the observations are known to be more robust against outliers (e.g., Breitung and Gouriéroux 1997). In other words, the rank transformation limits the influence of extreme values so that the distribution of a test statistic based on ranks tends to be more concentrated around its mean.

To investigate the robustness of the rank statistic, a simulation experiment is performed. Let x_t and y_t be two mutually and serially uncorrelated random walks with $T = 200$. At time period $T^* = T/2$, the series y_t is affected by an additive outlier of size $\lambda\sigma$, where σ denotes the standard deviation of Δy_t. Table 2 reports the empirical rejection frequencies of the rank statistics compared to the residual Dickey–Fuller test for cointegration (CDF). The results suggest that the rank statistics are robust against additive outliers, whereas the CDF statistic is quite sensitive against outlying observations.

3. POWER

Theorem 2 implies that, if $f(x_t)$ and $g(y_t)$ are independent I(1) series, it follows that $T^{-1}d_{[aT]} \Rightarrow a(\mathcal{A}_1 - \mathcal{A}_3) + (1-a) \times (\mathcal{A}_2 - \mathcal{A}_4)$, where $\mathcal{A}_1, \ldots, \mathcal{A}_4$ are independent arcsine-distributed random variables. Notice that the increments of random-walk sequences x_t and y_t are allowed to be serially correlated.

Table 2. Robustness Against Additive Outliers

λ	κ_T	ξ_T	CDF
5	.051	.049	.090
10	.045	.048	.240
15	.041	.048	.465
20	.038	.048	.672

NOTE: Rejection frequencies are computed from 10,000 replications of two uncorrelated random-walk sequences with unit variances. One of the series is affected by an additive outlier of size λ at time period $T^* = T/2$. The nominal size is .05, and the sample size is $T = 200$.

Under the alternative of a cointegration relationship as given in (6), it is seen that

$$\begin{aligned}T^{-1}d_{[aT]} &= T^{-1}\{R_T[T^{-1/2}g(y_t)] - R_T[T^{-1/2}f(x_t)]\} \\ &= T^{-1}\{R_T[T^{-1/2}f(x_t) + o_p(1)] - R_T[T^{-1/2}f(x_t)]\} \\ &\Rightarrow 0.\end{aligned}$$

Hence, κ_T and ξ_T converge to 0 as $T \to \infty$; that is, both rank tests are consistent.

Apart from this general statement, it is quite difficult to obtain analytical results for the (local) power of the test. Nevertheless, some interesting properties of the rank test can be derived when the parametric analog of the ranked differences is considered. Let the normalized difference of the series be defined as

$$\delta_t = \frac{y_t}{\bar{\sigma}_y} - \frac{x_t}{\bar{\sigma}_x},$$

where y_t and x_t are I(1), $E(y_t) = E(x_t) = 0$ for all t, and $\bar{\sigma}_x^2$, $\bar{\sigma}_y^2$ denote the respective long-run variances as defined in Definition 1. Accordingly, a parametric analog of the statistic ξ_T is constructed as

$$D_T = \frac{1}{T^2}\sum_{t=1}^{T}\delta_t^2, \qquad (12)$$

and under the null hypothesis of two uncorrelated random-walk sequences, the statistic is asymptotically distributed as $\int_0^1 [W_1(r) - W_2(r)]^2\,dr$, where $W_1(r)$ and $W_2(r)$ are uncorrelated standard Brownian motions. Under the alternative hypothesis, assume that $u_t = y_t - \beta x_t$ is stationary and $\beta > 0$. In this case, $\bar{\sigma}_y = \beta\bar{\sigma}_x$ so that

$$\begin{aligned}D_T &= \frac{1}{T^2}\sum_{t=1}^{T}\left(\frac{\beta x_t + u_t}{\beta\bar{\sigma}_x} - \frac{x_t}{\bar{\sigma}_x}\right)^2 \\ &= \frac{1}{\beta^2\bar{\sigma}_x^2 T^2}\sum_{t=1}^{T} u_t^2 \\ &= \frac{\sigma_u^2}{\beta^2\bar{\sigma}_x^2 T} + o_p(T^{-1}).\end{aligned}$$

It is seen that for large T the power of the test depends on the "signal-to-noise ratio" $\beta^2\bar{\sigma}_x^2/\sigma_u^2$, where σ_u^2 is the variance of u_t. As a consequence, the power of the test is a monotonically increasing function of the parameter β. In contrast, β does not affect the power of the residual Dickey–Fuller test. As a result, one expects that a test based on D_T or its ranked counterpart ξ_T will have more (less) power than the residual Dickey–Fuller test if β is large (small).

4. EXTENSIONS

So far I have assumed that the $f(x_t)$ and $g(y_t)$ are independent I(1) series. Of course, this assumption is quite restrictive, and in many applications it is reasonable to assume that the series are correlated. I therefore relax this assumption and assume instead that $f(x_t)$ and $g(y_t)$ converge to two correlated Brownian motions $W_1(r)$ and $W_2(r)$ with correlation coefficient $\rho = E[W_1(1)W_2(1)]$. Since $f(x_t)$ and $g(y_t)$ are not observed, it is not possible to estimate ρ directly. Rather I shall investigate the relationship between ρ and the expected correlation coefficient of the rank differences

$$\rho_T^R = \frac{\sum_{t=2}^T \Delta R_T(x_t)\Delta R_T(y_t)}{\sqrt{\left(\sum_{t=2}^T \Delta R_T(x_t)^2\right)\left(\sum_{t=2}^T \Delta R_T(y_t)^2\right)}}. \quad (13)$$

If there exists an (asymptotic) one-to-one relationship between ρ and $E(\rho_T^R)$, then it is possible to derive the limiting distributions of the test statistics. Unfortunately, the relationship between ρ and $E(\rho_T^R)$ is very complicated, and an analytical evaluation appears intractable. Therefore, Monte Carlo, simulations are employed to approximate the functional relationship between the two parameters.

Figure 2 presents the estimated function between ρ and $E(\rho_T^R)$ using 5,000 Monte Carlo replications with $T = 100$. It is seen that ρ_T^R tends to underestimate ρ in absolute value. However, the differences are small for moderate values of ρ. Therefore, ρ_T^R can be used as a first guess of ρ. This suggests that for small values of ρ, the test statistic can be corrected in a similar manner as in the linear case

$$\xi_T^* = \xi_T/\hat{\sigma}_{\Delta d}^2 \quad \text{and} \quad \kappa_T^* = \kappa_T/\hat{\sigma}_{\Delta d}, \quad (14)$$

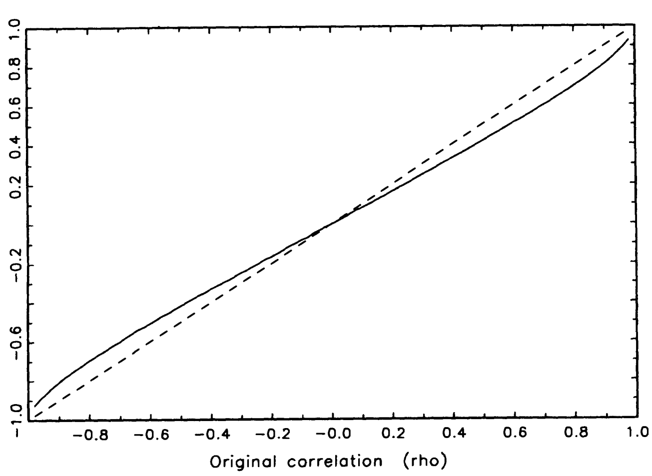

Figure 2. The Relationship Between ρ and $E(\rho_T^*)$.

where

$$\hat{\sigma}_{\Delta d}^2 = T^{-2}\sum_{t=2}^T (d_t - d_{t-1})^2.$$

Critical values for the corrected statistics ξ_T^* and κ_T^* can be found in Table 1.

If the absolute value of ρ is close to 1, this correction may perform poorly and a more accurate correction is required. In general, the critical values of ξ_T^* and κ_T^* depend on the correlation coefficient ρ or—by using the relationship between ρ and $E(\rho_T^R)$—on $E(\rho_T^R)$. Let $c_\xi^\alpha(\rho)$ denote the critical value of ξ_T^* with respect to the significance level of α. Using the relationship between ρ and $E(\rho_T^R)$, the critical value may alternatively be expressed as $c_\xi^\alpha(\rho) = c_\xi^\alpha(0)\lambda_\xi^\alpha(E\rho_T^R)$, where $\lambda_\xi^\alpha(\cdot)$ is an (unknown) function and $c_\xi^\alpha(0)$ is the critical value of ξ_T^* as presented in Table 1. Accordingly, a test with a correct size in the case of a substantial value of the correlation coefficient is obtained as

$$\xi_T^{**} = \xi_T^*/\lambda_\xi^\alpha(E\rho_T^R) \quad \text{and} \quad \kappa_T^{**} = \kappa_T^*/\lambda_\kappa^\alpha(E\rho_T^R). \quad (15)$$

Unfortunately, the determination of the function $\lambda_\xi^\alpha(E\rho_T^R)$ seems intractable so that Monte Carlo simulations were used to approximate the function. I generated 5,000 random draws with $T = 100$ and $\alpha = .05$ in the range $\rho = [-.98, -.96, \ldots, .96, .98]$. A fourth-order polynomial was estimated, and the most important regressors were used to obtain the following approximations:

$$\lambda_\kappa^{.05} \simeq 1 - .174(\rho_T^R)^2, \qquad R^2 = .985$$

and

$$\lambda_\xi^{.05} \simeq 1 - .462\rho_T^R, \qquad R^2 = .929.$$

The statistics ξ_T^{**} and κ_T^{**} have the same limiting distributions as ξ_T^* and κ_T^* (see Table 1 for the critical values).

Furthermore, it is possible to generalize the test to cointegration among $k+1$ variables $y_t, x_{1t}, \ldots, x_{kt}$, where it is assumed that $g(y_t)$ and $f_j(x_{jt})$ $(j = 1, \ldots, k)$ are monotonic functions. Let $R_T(\mathbf{x}_t) = [R_T(x_{1t}), \ldots, R_T(x_{kt})]'$ be a $k \times 1$ vector and $\tilde{\mathbf{b}}_T$ be the least squares estimate from a regression of $R_T(y_t)$ on $R_T(\mathbf{x}_t)$. Using the residuals $\tilde{u}_t^R = R_T(y_t) - \tilde{\mathbf{b}}_T' R_T(\mathbf{x}_t)$, a multivariate rank statistic is obtained from the normalized sum of squares:

$$\Xi_T[k] = T^{-3}\sum_{t=1}^T (\tilde{u}_t^R)^2.$$

Using

$$T^{-3}\sum_{t=1}^T R_T(y_t)R_T(x_{jt}) = \frac{1}{2}\left[\frac{2}{3} - \xi_{jt}\right] + o_p(1),$$

where ξ_{jt} is the bivariate rank statistic for y_t and x_{jt} defined as $\xi_{jt} = T^{-3}\sum_{t=1}^T [R_T(y_t) - R_T(x_{jt})]^2$, it is not difficult to show that the multivariate test statistic can be represented as

$$\Xi_T[k] = \frac{1}{3} - \frac{1}{4}\delta_T'\Psi_T\delta_T,$$

where

$$\delta_T = \begin{bmatrix} \xi_{1T} - 2/3 \\ \xi_{2T} - 2/3 \\ \vdots \\ \xi_{kT} - 2/3 \end{bmatrix}$$

and

$$\Psi_T = T^{-3} \sum_{t=1}^{T} R_T(\mathbf{x}_t) R_T(\mathbf{x}_t)'.$$

To account for a possible correlation between the series, a modified statistic can be constructed:

$$\Xi_T^*[k] = \Xi_T[k]/\hat{\sigma}_{\Delta u}^2, \tag{16}$$

where

$$\hat{\sigma}_{\Delta u}^2 = T^{-2} \sum_{t=2}^{T} (\tilde{u}_t^R - \tilde{u}_{t-1}^R)^2.$$

Critical values for the test statistic $\Xi_T^*[k]$ are presented in Table 1.

5. A RANK TEST FOR NEGLECTED NONLINEARITY

Whenever the rank test for cointegration indicates a stable long-run relationship, it is interesting to know whether the cointegration relationship is linear or nonlinear. For a convenient representation of such null and alternative hypotheses I follow Granger (1995) and write the nonlinear relationship as

$$y_t = \gamma_0 + \gamma_1 x_t + f^*(x_t) + u_t, \tag{17}$$

where $\gamma_0 + \gamma_1 x_t$ is the linear part of the relationship. Under the null hypothesis, it is assumed that $f^*(x_t) = 0$ for all t. If $f^*(x_t)$ is unknown, it may be approximated by Fourier series (Gallant 1981) or a neural network (Lee, White, and Granger 1993). Here the multiple of the rank transformation $\theta R_T(x_t)$ is used instead of $f^*(x_t)$.

It is interesting to note that the rank transformation is to some extent related to the neural-network approach suggested by Lee et al. (1993). If $\mathbf{x_t}$ is a $k \times 1$ vector of "input variables" and α is a corresponding vector of coefficients, the neural-network approach approximates $f^*(\mathbf{x_t})$ by $\sum_{j=1}^{q} \beta_j \psi(\mathbf{x}_t' \alpha_j)$, where $\psi(\cdot)$ has the properties of a cumulated distribution function. A function often used in practice is the logistic $\psi(x) = x/(1-x)$. In this context, x_t is a scalar variable so that the neural-network term simplifies to $\beta\psi(\alpha x_t)$. Using $T^{-1} R_T(x_t) = \widehat{F}_T(x_t)$, where $\widehat{F}_T(x_t)$ is the empirical distribution function, the rank transformation can be motivated as letting $\psi(\alpha x_t)$ be the empirical distribution function with the attractive property that the parameter α can be dropped due to the invariance of the rank transformation.

If it is assumed that x_t is exogenous and u_t is white noise with $u_t \sim \mathrm{N}(0, \sigma^2)$, a score test statistic is obtained as the $T \cdot R^2$ statistic of the least squares regression

$$\tilde{u}_t = c_0 + c_1 x_t + c_2 R_T(x_t) + e_t, \tag{18}$$

where $\tilde{u}_t = y_t - \tilde{\gamma}_0 - \tilde{\gamma}_1 x_t$ and $\tilde{\gamma}_0$ and $\tilde{\gamma}_1$ are the least squares estimates from a regression of y_t on x_t and a constant.

A problem with applying the usual asymptotic theory to derive the limiting null distribution of the test statistic is that the regression (18) involves the nonstationary variables x_t and $R_T(x_t)$. However, under some (fairly restrictive) assumptions, Theorem 3 shows that under the null hypothesis $c_2 = 0$ the score statistic is asymptotically χ^2 distributed.

Theorem 3. Let $x_t = \sum_{j=1}^{t} v_j$ and $y_t = \gamma_0 + \gamma_1 x_t + u_t$, where it is assumed that v_t is I(0) according to Definition 1 and u_t is white noise with $E(u_t) = 0$, $E(u_t^2) = \sigma_u^2$, and $E(u_t v_s) = 0$ for all t and s. As $T \to \infty$, the score statistic for $H_0: c_2 = 0$ in Regression (18) has an asymptotic χ^2 distribution with 1 df.

Proof. It is convenient to introduce the matrix notation

$$\mathbf{X}_1 = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_T \end{bmatrix} \quad \text{and} \quad \mathbf{X}_2 = \begin{bmatrix} R_T(x_1) \\ \vdots \\ R_T(x_T) \end{bmatrix},$$

$\mathbf{y} = [y_1, \ldots, y_T]'$ and $\tilde{\mathbf{u}} = [\tilde{u}_1, \ldots, \tilde{u}_T]'$. With this notation, the score statistic can be written as

$$T \cdot R^2 = \frac{1}{\tilde{\sigma}^2} (\hat{\beta}_2)^2 [\mathbf{X}_2'\mathbf{X}_2 - \mathbf{X}_2'\mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{X}_2],$$

where $\hat{\beta}_2$ is the least squares estimator of β_2 in the regression $\mathbf{y} = \mathbf{X}_1\beta_1 + \mathbf{X}_2\beta_2 + \mathbf{u}$ and $\tilde{\sigma}^2 = \tilde{\mathbf{u}}'\tilde{\mathbf{u}}/T$. As shown by Park and Phillips (1988), the least squares estimator in a regression with strictly exogenous I(1) regressors is conditionally normally distributed so that conditional on $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2]$, $\hat{\beta}_2$ is asymptotically distributed as $\mathrm{N}(\mathbf{0}, \mathbf{V}_2)$, where

$$\mathbf{V}_2 = \sigma_u^2 [\mathbf{X}_2'\mathbf{X}_2 - \mathbf{X}_2'\mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{X}_2]^{-1}.$$

From $\tilde{\sigma}_u^2 \to^p \sigma_u^2$ it follows that $T \cdot R^2$ has an asymptotic χ^2 distribution with 1 df.

Unfortunately, the assumptions for Theorem 3 are too restrictive to provide a useful result for practical situations. In many situations, the errors u_t are found to be serially correlated and the regressor x_t may be endogenous. However, using standard techniques for cointegration regressions (Saikkonen 1991; Stock and Watson 1993), the test can be modified to accommodate serially correlated errors and endogenous regressors. For this purpose, assume that

$$\begin{aligned} u_t &= E(u_t | \Delta x_t, \Delta x_{t\pm 1}, \Delta x_{t\pm 2}, \ldots) + \eta_t \\ &= \sum_{j=-\infty}^{\infty} \pi_j \Delta x_{t-j} + \eta_t \end{aligned}$$

and η_t admits the autoregressive representation

$$\eta_t = \sum_{j=1}^{\infty} \alpha_j \eta_{t-j} + \varepsilon_t,$$

where the lag polynomial $\alpha(B) = 1 - \alpha_1 B - \alpha_2 B^2 - \cdots$ has all roots outside the complex unit circle.

Under the null hypothesis of linear cointegration, the following representation results:

$$y_t = \gamma_0^* + \sum_{j=1}^{\infty} \alpha_j y_{t-j} + \gamma_1^* x_t + \sum_{j=-\infty}^{\infty} \pi_j^* \Delta x_{t-j} + \varepsilon_t \qquad (19)$$

(see Stock and Watson 1993; Inder 1995).

A test for nonlinear cointegration can be obtained by truncating the infinite sums appropriately and forming $T \cdot R^2$ for the regression of the residuals $\tilde{\varepsilon}_t$ on the regressors of (19) and $R_T(x_t)$. Along the lines of Theorem 3, it can be shown that the resulting score statistic is asymptotically χ^2 distributed under the null hypothesis of a linear cointegration relationship.

6. SMALL-SAMPLE PROPERTIES

To investigate the small-sample properties of the rank tests, I generate two time series according to the model equations

$$y_t = \beta f(x_t) + u_t, \qquad u_t = \alpha u_{t-1} + \varepsilon_t$$
$$f(x_t) = f(x_{t-1}) + v_t, \qquad (20)$$

where

$$\begin{bmatrix} \varepsilon_t \\ v_t \end{bmatrix} \sim \text{iid N}\left(\mathbf{0}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right).$$

The variable x_t is obtained from the random walk z_t by using the inverse function $x_t = f^{-1}(z_t)$.

Under the null hypothesis $H_0 : \alpha = 1$, there is no cointegration relationship between the series. If, in addition, $\beta = 0$ and $\rho = 0$, then x_t and y_t are two independent random walks. For this specification, Table 3 reports the rejection frequencies of different cointegration tests. The rank tests κ_T and ξ_T are computed as in (9) and (10), and "CDF" indicates the Dickey–Fuller t test applied to the residuals of a linear regression of y_t on x_t and a constant. The results for the linear process are given in the left half of Table 3, indicated by $f(x) = x$. It might be surprising that for $\beta = 1$ the rank test is much more powerful than the CDF test if α is close to 1. As argued in Section 3, this is because the rank statistics depend on the parameter β, whereas the Dickey–Fuller type of test for cointegration does not depend on β. In fact, the simulation results indicate that the power of the rank tests is that they are quite sensitive to the value of β. For $\beta = .5$, the Dickey–Fuller test performs better, whereas for $\beta = 1$ the rank tests clearly outperform the Dickey–Fuller type of tests.

It should also be noted that the rank tests do not require estimating the cointegration parameter β. Accordingly, this test has the same power as for the case of a known cointegration relationship. Furthermore, the rank-test procedures impose the *one-sided* hypothesis that $f(x_t)$ is an *increasing* function.

Since the rank tests are invariant to a monotonic transformation of the variables, the power function is the same as for the linear case. Comparing the power of the CDF test with the rank counterparts, it turns out that the power of the CDF test may drop dramatically for nonlinear alternatives (see also Granger and Hallman 1991b), while the rank test performs as well as in the linear case. In particular, for the case $f(x) = \log(x)$ the parametric CDF test performs quite poorly.

To investigate the power of the rank test against nonlinear alternatives, which cannot be written as in (6), I generate data using the threshold cointegrated process given by

$$u_t = \begin{cases} u_{t-1} + \varepsilon_t & \text{if } |u_{t-1}| \le \lambda\sigma_u \\ a u_{t-1} + \varepsilon_t & \text{if } |u_{t-1}| > \lambda\sigma_u, \end{cases} \qquad (21)$$

where σ_u is the standard deviation of u_t and λ is a positive constant. Furthermore, I generate z_t as a random walk independent of u_t and let $\beta = 1$. Such nonlinear processes were investigated by Balke and Fomby (1997) and Rozada and Gonzalo (1998). I consider two different threshold values given by $\lambda = .5$ and $\lambda = 1.5$. The results for various values of a are presented in Table 4. It turns out that the rank tests possess much more power against that kind of nonlinear alternatives than the CDF test.

Table 3. Size and Power

$f(x) =$	x			x^3	$\log(x)$	$\tan(x_t)$
α	κ_T	ξ_T	CDF	CDF	CDF	CDF
			Size ($\beta = 0$)			
1.00	.049	.049	.050	.098	.051	.077
			Power ($\beta = .5$)			
.98	.232	.248	.080	.310	.156	.186
.95	.288	.332	.225	.491	.229	.295
.90	.387	.484	.698	.730	.323	.474
.80	.551	.708	.999	.866	.410	.616
			Power ($\beta = 1$)			
.98	.594	.616	.080	.310	.156	.186
.95	.733	.792	.225	.491	.229	.295
.90	.861	.930	.698	.730	.323	.474
.80	.953	.993	.999	.866	.410	.616

NOTE: Rejection frequencies resulting from 10,000 replications of the process given in (20). The sample size is $T = 200$. Under the null hypothesis ($\alpha = 1$) I let $\beta = 0$, and under the alternative ($|\alpha| < 1$), $\beta = 1$. The nominal size is .05. The test statistic κ_T and ξ_T are defined in (9) and (10). CDF indicates a Dickey–Fuller t test on the residuals of a cointegrating regression including a constant term.

Table 4. Power Against Threshold Cointegration

	$\lambda = .5$			$\lambda = 1.5$		
a	κ_T	ξ_T	CDF	κ_T	ξ_T	CDF
.98	.600	.629	.081	.547	.571	.083
.95	.727	.797	.201	.642	.688	.141
.90	.854	.929	.655	.783	.840	.327
.80	.948	.992	.999	.913	.961	.960
.60	.989	1.000	1.000	.974	.998	1.000

NOTE: Empirical powers are computed from 10,000 replications of the threshold cointegrated process given in (20) and (21), where $T = 200$. The parameter values are $\beta = 1$ and $\sigma_u^2 = 1/(1-a^2)$. The significance level is .05. The test statistic κ_T and ξ_T are defined in (9) and (10). CDF indicates a Dickey–Fuller t test on the residuals of a cointegrating regression including a constant term.

To study the ability of the modified statistics suggested in Section 4 to account for correlated random walks, I simulate correlated data by varying the correlation coefficient in the range $\rho = -.9, \ldots, .9$. The statistics κ_T^* and ξ_T^* use a correction that is similar to the correction in the linear case. As argued in Section 4, this test statistic should perform well if the correlation is moderate. For more substantial correlation coefficients, the improved statistics κ_T^{**} and ξ_T^{**} defined in (15) should be used.

Table 5 presents the empirical sizes for testing the null hypothesis of no cointegration with $\beta = 0$ and $\alpha = 1$. It turns out that the statistic κ_T^* performs well in the range $\rho \in [-.4, .4]$, whereas the statistics ξ_T^* and $\Xi_T^*[1]$ should only be used for a small correlation in the range $\rho \in [-.2, .2]$. In contrast, the statistics κ_T^{**} and ξ_T^{**} turn out to be very robust against a correlation between x_t and y_t.

Next I consider the small-sample properties of the rank test for nonlinear cointegration suggested in Section 5. It is assumed that y_t and x_t are cointegrated so that $y_t - \beta f(x_t)$ is stationary. By setting $\alpha = .5$, I generate serially correlated errors and, letting $\rho = .5$, the variable x_t is correlated with the errors u_t; that is, x_t is endogenous. The rank test for nonlinear cointegration is obtained by regressing y_t on $x_t, y_{t-1}, \Delta x_{t+1}, \Delta x_t, \Delta x_{t-1}$, and a constant. The score statistic is computed as $T \cdot R^2$ from a regression of the residuals on the same set of regressors and the ranks $R_T(x_t)$.

To study the power of the test, I consider three different nonlinear functions. As a benchmark, I perform the tests using $f(x_t)$ instead of the ranks $R_T(x_t)$. Of course, using the true functional form, which is usually unknown in practice, I expect the test to have better power than the test based on the ranks. Surprisingly, the results of the Monte Carlo simulations (see Table 6) suggest that the rank test may even be (slightly) more powerful than the parametric test, whenever the nonlinear term enters the equation with a small weight ($\beta = .01$). However, the gain in power is quite small and falls in the range of the simulation error. In any case, the rank test performs very well and seems to imply no important loss of power in comparison to the parametric version of the test.

Table 5. Testing Correlated Random Walks

ρ	κ_T^*	ξ_T^*	κ_T^{**}	ξ_T^{**}	$\Xi_T^*[1]$
−.900	.130	.007	.054	.033	.003
−.600	.073	.016	.048	.039	.012
−.400	.058	.023	.048	.042	.020
−.200	.052	.030	.049	.042	.029
.000	.048	.041	.047	.040	.042
.200	.050	.053	.047	.040	.056
.400	.052	.070	.043	.037	.074
.600	.062	.096	.039	.036	.107
.900	.105	.234	.038	.053	.255

NOTE: Rejection frequencies resulting from 10,000 realizations of two random walks with $\text{corr}(\Delta x_t, \Delta y_t) = \rho$. The sample size is $T = 200$. The statistics κ_T^* and ξ_T^* are defined in (14) and κ_T^{**} and ξ_T^{**} are given in (15). The statistic $\Xi_T^*[1]$ is the two-sided test statistic given in (16).

7. AN EMPIRICAL APPLICATION

The rank tests are applied to test for a possible nonlinear cointegration between interest rates with different times to maturity. Recent empirical work suggests that interest rates with different times to maturity are nonlinearly related (e.g., see Campbell and Galbraith 1993; Pfann, Schotman, and Tschering 1996, and the references therein). The dataset consists of monthly yields of government bonds with different times to maturity as published by the German *Bundesbank*. The sample runs from 1967(1) through 1995(12) yielding 348 monthly observations for each variable.

The nonlinear relationship between yields for different times to maturity can be motivated as follows. Let r_t denote the yield of a one-period bond and R_t represent the yield of a two-period bond at time t. The expectation theory of the term structure implies that

$$R_t = \phi_t + .5r_t + .5E_t(r_{t+1}), \tag{22}$$

where E_t denotes the conditional expectation with respect to the relevant information set available at period t and ϕ_t represents the risk premium. Letting $r_{t+1} = E_t(r_{t+1}) + 2\nu_t$ and subtracting r_t from both sides of (22) gives $R_t - r_t = .5(r_{t+1} - r_t) + \phi_t + \nu_t$. Assuming that r_t is I(1) and $\phi_t + \nu_t$ is stationary implies that R_t and r_t are (linearly) cointegrated (e.g., Wolters 1995). However, if the risk premium depends on r_t such that $\phi_t = f^*(r_t) + \eta_t$ with η_t stationary, then the yields are nonlinearly cointegrated: $R_t - f(r_t) = u_t \sim \text{I}(0)$, where $f(r_t) = r_t + f^*(r_t)$ and $u_t = .5(r_{t+1} - r_t) + \eta_t + \nu_t$. Note that u_t is correlated with r_t and, therefore, r_t is endogenous. Furthermore, if the sampling interval is shorter than the time to maturity, then the errors are serially correlated even if ν_t and η_t are white noise.

Table 6. Power Against a Nonlinear Cointegration Relationship

	$f(x) =$					
	x^3		$log(x)$		$tan(x)$	
Regressor	$R_T(x_t)$	$f(x_t)$	$R_T(x_t)$	$f(x_t)$	$R_T(x_t)$	$f(x_t)$
$\beta = .01$	.267	.252	.246	.216	.237	.226
$\beta = .05$	.473	.485	.701	.676	.549	.548
$\beta = .1$	.714	.746	.957	.955	.834	.855
$\beta = .5$	.974	.988	1.000	1.000	.999	1.000

NOTE: Simulated power from a score test using $R_T(x_t)$ and $f(x_t)$ as additional regressors. The sample size is $T = 200$.

Table 7. The Cointegration Relationship With R1

Var.	CDF	$\xi_T^*[1]$	κ_T^{**}	ξ_T^{**}	ρ_T^R	Nonlin
R2	-4.138^+	$.010^+$	.407	$.017^+$	.893	.003
R3	-3.591^+	$.014^+$	.405	.022	.808	.012
R4	-3.373^+	$.016^+$	.404	.023	.730	.034
R5	-3.213	$.018^+$	.413	.025	.678	.107
R10	-2.702	.025	.498	.033	.582	.489

NOTE: "CDF" denotes the Dickey–Fuller t test applied to the residuals of the cointegration regression. $\xi_T^*[1]$ is the two-sided test statistic given in (16), and κ_T^{**} and ξ_T^{**} are defined in (15). "Nonlin" indicates the variable addition test for nonlinearity suggested in Section 5 with the ranks of x_t and y_{t-1}, Δx_{t-1}, Δx_t, Δx_{t+1} as regressors. ρ_T^R is the correlation coefficient between the differenced ranks. "+" indicates significance with respect to the significance level of .05.

To test whether interest rates possess a (nonlinear) cointegration relationship, I first apply various tests for unit roots to the series. Neither the conventional Dickey–Fuller t test nor the ranked counterpart suggested by Breitung and Gouriéroux (1997) reject the null hypothesis that the interest rates are I(1) (not presented).

The results of the different cointegration test statistics are presented in Table 7. The (parametric) Dickey–Fuller test applied to the residuals of a linear cointegrating regression (CDF) indicates a cointegration relationship for (R1,R2), (R1,R3), and (R1,R4). A similar result is obtained by applying the (two-sided) test statistic $\Xi_T^*[1]$. However, this test is also able to (marginally) reject the null hypothesis for the case (R1,R5). With respect to the high correlation coefficients ρ_T^R, the application of this test is problematical. The test statistics κ_T^{**} and ξ_T^{**} are more appropriate in this case. With the exception of (R1,R2), where ξ_T^{**} rejects the null hypothesis of no cointegration, the rank statistics are not able to detect a cointegration relationship in a bivariate model. Furthermore, the insignificant test statistics suggested in Theorem 3 (see the column "Nonlin" in Table 7) suggest that, if there is a cointegration relationship, then this relationship is linear.

When the rank test is applied to the six-dimensional vector $y_t = [R1, \ldots, R5, R10]_t'$, the rank statistic results as $\Xi_T^*[5] = .00475$, which is clearly significant with respect to the critical values presented in Table 1. Thus, there is strong evidence that there is at least one cointegration relationship among the six interest yields.

8. CONCLUDING REMARKS

This article considers rank tests for a nonlinear cointegration relationship. Under the hypothesis that $T^{-1/2}g(y_{[aT]})$ and $T^{-1/2}f(x_{[aT]})$ converge weakly to two independent Brownian motions $W_1(a)$ and $W_2(a)$, the limiting distribution of the rank differences $d_{[aT]} = R_T(y_{[aT]}) - R_T(x_{[aT]})$ can be derived. Tests based on the sequence d_t appear to have good power properties against cointegration relationships of the form $g(y_t) - f(x_t)$. However, if the cointegration relationship has a more general form—for example, $h(y_t, x_t) = x_t/y_t \sim I(0)$—the rank test may lack power. Therefore, it seems desirable to construct nonparametric tests against a more general form of nonlinear cointegration.

To allow for correlated time series, a correction is suggested that can be computed by using the correlation coefficient of the differenced ranks. The results of my simulation experiments suggest that the correction performs well. Furthermore, the rank tests may clearly outperform their parametric counterparts if the cointegration relationship is indeed nonlinear.

The rank-test procedures were applied to test for nonlinear cointegration between interest yields with different times to maturity. In all, my results do not reveal important nonlinearities in the long-run relationship.

ACKNOWLEDGMENTS

I am indebted to Filipe Aparicio Acosta, Rolf Tschernig, Halbert White, two referees, an associate editor, and the editor for helpful comments and suggestions. Of course, I am responsible for all remaining errors and deficiencies. Furthermore, I gratefully acknowledge financial support from the Sonderforschungsbereich 373 at the Humboldt University Berlin.

[Received September 1998. Revised December 2000.]

REFERENCES

Abadir, K. M., and Taylor, A. M. R. (1999), "On the Definitions of Cointegration," *Journal of Time Series Analysis*, 20, 129–137.

Acosta, F. M. A., and Granger, C. W. J. (1995), "Nonlinear Cointegration and Some New Tests for Comovements," Discussion Paper 95-15, University of California, San Diego, Dept. of Economics.

Balke, N. S., and Fomby, T. B. (1997), "Threshold Cointegration," *International Economic Review*, 38, 627–645.

Breitung, J., and Gouriéroux, C. (1997), "Rank Tests for Unit Roots," *Journal of Econometrics*, 81, 7–28.

Campbell, B., and Galbraith, J. W. (1993), "Inference in Expectations Models of the Term Structure: A Non-parametric Approach," *Empirical Economics*, 18, 623–638.

Corradi, V. (1995), "Nonlinear Transformations of Integrated Time Series: A Reconsideration," *Journal of Time Series Analysis*, 16, 537–549.

Davidson, J. (1998), "When Is a Time Series I(0)? Evaluating the Memory Properties of Nonlinear Dynamic Models," working paper, University of Cardiff, Cardiff Business School, Economics Section.

Franses, P. H., and McAleer, M. (1998), "Testing for Unit Roots and Nonlinear Transformations," *Journal of Time Series Analysis*, 19, 147–164.

Gallant, A. R. (1981), "Unbiased Determination of Production Technologies," *Journal of Econometrics*, 30, 149–169.

Gallant, A. R., and White, H. (1988), "A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models," Oxford, U.K. Basil Blackwell.

Granger, C. W. J. (1981), "Some Properties of Time Series Data and Their Use in Econometric Model Specification," *Journal of Econometrics*, 16, 121–130.

——— (1995), "Modelling Nonlinear Relationships Between Extended-memory Variables," *Econometrica*, 63, 265–279.

Granger, C. W. J., and Hallman, J. (1991a), "Long Memory Processes With Attractors," *Oxford Bulletin of Economics and Statistics*, 53, 11–26.

——— (1991b), "Nonlinear Transformations of Integrated Time Series," *Journal of Time Series Analysis*, 12, 207–224.

Granger, C. W. J., Inoue, T., and Morin, N. (1997), "Nonlinear Stochastic Trends," *Journal of Econometrics*, 81, 65–92.

Granger, C. W. J., and Teräsvirta, T. (1993), *Modelling Nonlinear Economic Relationships*, Oxford, U.K.: Oxford University Press.

Inder, B. (1995), "Finite Sample Arguments for Appropriate Estimation of Cointegrating Vectors," discussion paper, Monash University, Australia, Dept. of Econometrics.

Kendall, M., and Gibbons, J. D. (1990), *Rank Correlation Methods*, London: Edward Arnold.

Lee T.-H., White, H., and Granger, C. W. J. (1993), "Testing for Neglected Nonlinearity in Time Series Models: A Comparison of Neural Network Methods and Alternative Tests," *Journal of Econometrics*, 56, 269–290.

Lo, A. W. (1991), "Long-term Memory in Stock Market Prices," *Econometrica*, 59, 1279–1314.

Park, J. Y., and Phillips, P. C. B. (1988), "Statistical Inference in Regressions With Integrated Processes: Part 1," *Econometric Theory*, 4, 468–497.

——— (1999), "Asymptotics for Nonlinear Transformations of Integrated Time Series," *Econometric Theory*, 15, 269–298.

Pfann, G., Schotman, P., and Tschernig, R. (1996), "Nonlinear Interest Rate Dynamics and Implications for the Term Structure," *Journal of Econometrics*, 74, 149–176.

Phillips, P. C. B (1986), "Understanding Spurious Regressions in Econometrics," *Journal of Econometrics*, 33, 311–340.

——— (1987), "Time Series Regression With a Unit Root," *Econometrica*, 55, 277–301.

Phillips, P. C. B., and Oularis, S. (1990), "Asymptotic Properties of Residual Based Tests for Cointegration," *Econometrica*, 58, 165–193.

Phillips, P. C. B., and Solo, V. (1992), "Asymptotics for Linear Processes," *The Annals of Statistics*, 20, 971–1001.

Rozada, M. G., and Gonzalo, J. (1998), "Threshold Unit Root Models," paper presented at the European Meeting of the Econometric Society, Berlin, August 29–September 2.

Saikkonen, P. (1991), "Asymptotically Efficient Estimation of Cointegration Regressions," *Econometric Theory*, 7, 1–21.

Sargan, J. D., and Bhargava, A. (1983), "Testing Residuals From Least Squares Regression for Being Generated by the Gaussian Random Walk," *Econometrica*, 51, 153–174.

Stock, J. H., and Watson, M. W. (1993), "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems," *Econometrica*, 61, 783–820.

Wolters, J. (1995), "On the Term Structure of Interest Rates—Empirical Results for Germany," *Statistical Papers*, 36, 193–214.