

UNIVERSITY OF NOTTINGHAM



Discussion Papers in Economics

Discussion Paper
No. 00/29

**A DIRECT TEST FOR COINTEGRATION BETWEEN A
PAIR OF TIME SERIES**

by S.J. Leybourne, P. Newbold, D. Vougas and T. Kim

December 2000

DP 00/29

ISSN 1360-2438

UNIVERSITY OF NOTTINGHAM



Discussion Papers in Economics

Discussion Paper
No. 00/29

**A DIRECT TEST FOR COINTEGRATION BETWEEN A
PAIR OF TIME SERIES**

by S.J. Leybourne, P. Newbold, D. Vougas and T. Kim

Steven Leybourne and Paul Newbold are Professors, Tae-Hwan Kim is
Lecturer, School of Economics, University of Nottingham, and Dimitrios
Vougas is Senior Lecturer, Department of Economics, University of
Portsmouth

December 2000

A DIRECT TEST FOR COINTEGRATION BETWEEN A PAIR OF TIME SERIES

S.J. Leybourne, P. Newbold, D. Vougas and T. Kim

School of Economics, University of Nottingham
University Park, Nottingham NG7 2RD
United Kingdom

September 1999

Abstract: In this paper we introduce a new test of the null hypothesis of no cointegration between a pair of time series. For a very simple generating model, our test compares favourably with the Engle-Granger/Dickey-Fuller test and the Johansen trace test. Indeed, shortcomings of the former motivated the development of our test. The applicability of our test is extended to series generated by low-order vector autoregressions. Again, we find evidence that this general version of our test is more powerful than the Johansen test. The paper concludes with an empirical example in which the new test finds strong evidence of cointegration, but the Johansen test does not.

1. Introduction

Following the pioneering work of Granger (1981) and Engle and Granger (1987), the topic of cointegration has been at the heart of time series econometrics. In addition to the substantial theoretical developments that have subsequently been reported, considering the possibility of cointegration, and the associated error-correction model representation, has become an integral part of the standard paradigm of applied time series econometrics.

In this paper we shall discuss a specific problem - testing the null hypothesis of no cointegration between a pair of time series. Let y_{1t} and y_{2t} both be generated by processes which are integrated of order one, $I(1)$; that is, the series are nonstationary, but their first differences are stationary, or $I(0)$. Then, the series are cointegrated if and only if there exists a scalar b , such that $(y_{1t} - by_{2t})$ is stationary. The general form of our test, applicable when the process generating the series can be well-approximated by a relatively low order vector autoregression, is introduced in Section 3. Simulation evidence reported there suggests that the new test has superior power compared with the most commonly applied extant test. Although generalisation of our test to the case of several series would be very difficult, the case of two series is of considerable interest in its own right: see, for example, Ng and Perron (1997).

Section 2 of the paper is, in effect, an extended introduction to, and motivation for, the general test procedure. In that section, we analyse a very simple generating model, for which our test has a straightforward and intuitively appealing structure. We argue that, for this simple model, the new test has desirable properties when compared with the two tests that have most often been applied in this case.

Finally, in Section 4, we briefly discuss an empirical example, where the new test finds strong evidence of cointegration when one of its competitors does not, in a situation where cointegration might reasonably be expected on *a priori* grounds .

2. The Basic Case

To motivate our test, we begin by reporting and discussing the results of a small simulation study. Pairs of time series $y_t = (y_{1t}, y_{2t})'$ were generated from the model

$$\begin{aligned}\Delta y_{1t} &= -\gamma(y_{1,t-1} - y_{2,t-1}) + v_{1t} \\ \Delta y_{2t} &= \gamma(y_{1,t-1} - y_{2,t-1}) + v_{2t}, \quad t = 1, \dots, T\end{aligned}\tag{1}$$

where $v_t = (v_{1t}, v_{2t})'$ is zero-mean vector white noise, taken to be Gaussian throughout our simulations, with covariance matrix

$$\Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

For $\gamma = 0$, the y_{it} are $I(1)$ pure random walks, possibly with contemporaneously correlated white noise innovations. For $0 < \gamma < 1$, the y_{it} are generated by ARIMA(1,1,1) processes of identical forms, and are cointegrated, with $(y_{1t} - y_{2t})$ following an $I(0)$ first order autoregression with parameter $\phi = 1 - 2\gamma$. Due to this symmetry, it is irrelevant which y_{it} is treated as the left-hand variable in Engle-Granger regressions, thus abstracting from an issue discussed by Ng and Perron (1997). The generating process (1) is a special case of processes used by Banerjee et al (1986), Blangiewicz and Charemza (1990), Hansen and Phillips (1990), Gonzalo (1994), and others.

The two most commonly applied tests of the null hypothesis of no cointegration are the Engle-Granger/Dickey-Fuller (EGDF) test (Engle and Granger 1987, Engle and Yoo 1987) and the Johansen trace (JT) test (Johansen 1988, 1991, Johansen and Juselius, 1990). We applied these tests to series of $T = 100, 200$ observations, generated from (1), allowing for non-zero intercept, but not trend, in the generating process. We also took the model order to be known, so that redundant autoregressive terms were not incorporated in the models to be estimated to generate the test statistics. Then, the EGDF test was based on the residuals from the regression of y_{1t} on y_{2t} , with intercept. The test statistic is the t -ratio from the Dickey-Fuller

regression on these residuals, with no intercept or lagged first differences. Critical values were obtained from the response surfaces of MacKinnon (1991). The JT test was based on the vector autoregression, with intercept, using critical values given by Osterwald-Lenum (1992). Here, and in the more general case of the next section, we tried also the maximal-eigenvalue variant of the Johansen test. However, the results were invariably very similar to those from the trace test, and so will not be reported.

The results of these experiments on the powers of the two tests are summarised in Table 1. The most striking feature of these results is that, while the power of the JT test is unaffected by the parameter ρ , the power of the EGDF test depends dramatically on that parameter. For example, for ρ close to 1, the EGDF test is much more powerful than the JT test, whereas for ρ close to -1 the performance of EGDF is relatively very poor. It is certainly interesting that the relative powers of the two tests depend so dramatically on a single parameter in such a simple model.

In fact, it is straightforward to show, *for the particular model considered here*, that the distribution of the JT test statistic is invariant to ρ . By contrast, the distribution of the EGDF test statistic is clearly influenced by the parameter ρ . This test statistic is based initially on least squares estimates from the regression model

$$y_{1t} = a + by_{2t} + u_t \tag{2}$$

Its distribution therefore depends on the sampling distribution of the least squares estimator \hat{b} . Table 2 reports simulation evidence on the mean and standard deviation of \hat{b} for series of 100 observations generated from (1). For large positive ρ this estimator has mean close to 1, with small standard deviation. However, as ρ decreases, the bias in \hat{b} rapidly increases, the extent of that bias being quite startling for large negative ρ . It is certainly not surprising that the EGDF test has low power in these circumstances. Gonzalo (1994) discusses in some detail bias in this and other estimators of the cointegrating parameter. Our interest lies in the impact of that bias on the EGDF test.

The results in Table 1 suggest that, for ρ close to -1 , the performance of the EGDF test is unacceptably poor. It is perhaps worth asking if such results would be

obtained in practice, for when b is poorly estimated the residuals \hat{u}_t from (2) may not appear to be first order autoregressive. Consequently, the user may apply the augmented version of the Dickey-Fuller test to these residuals. We simulated series of 100 observations from the model (1), applying now the augmented version of the EGDF test. We allowed a maximum of six lagged differences in the Dickey-Fuller regressions, selecting the number of lags actually employed each time through general-to-specific testing at the 0.05 level, in line with proposals of Ng and Perron (1995). In fact, the power of the EGDF test was not improved, and was slightly lower for ρ close to -1 . We conclude then that there exist circumstances - those in which the parameter of the cointegrating regression is poorly estimated - when EGDF has very poor performance. Unfortunately, in practical applications, which are likely to involve models that are different from and more elaborate than (1), it would be practically impossible to identify cases where such circumstances arose.

Although the performance of the EGDF test can be very poor, the results of Table 1 suggest also that there are cases where substantial gains in power can be achieved through using EGDF rather than the Johansen test. It is worth asking whether at least some of those gains can be captured through a modified test whose power does not fall dramatically where the power of EGDF does. We explore here a test motivated directly by the definition of cointegration, continuing to employ Dickey-Fuller-type regressions, but omitting the estimation of b through the regression (2), since this seems to be the source of the difficulties apparent in Table 1.

Let y_{1t} and y_{2t} be a pair of $I(1)$ time series. Then, these series are cointegrated if and only if there exists a scalar b such that $(y_{1t} - by_{2t})$ is stationary. We consider Dickey-Fuller regressions for *all possible* b . Specifically, we estimate regressions of the form

$$\Delta(y_{1t} - by_{2t}) = a + c(y_{1,t-1} - by_{2,t-1}) + \omega_t$$

Let τ_b be the t -ratio associated with the estimate of c . The lower is that t -ratio, the stronger is the evidence supporting cointegration for a given b . Our test statistic is the lowest value of τ_b for all possible b ; that is

$$\tau^* = \inf_{b \in \mathcal{S}} \tau_b$$

where \mathfrak{S} is a compact interval. Viewing τ_b as a function of the scalar b , τ^* can easily be found numerically.

Table 3 shows critical values of the τ^* statistic at the standard significance levels, obtained through simulations based on independent driftless random walks. We shall discuss the asymptotics in the next section, but note for now that the critical values appear to settle down quite quickly with increasing sample size. It is not necessary to verify the robustness of these critical values to correlation between the innovations generating the random walks. As we shall see, for the model (1), the finite sample distribution of τ^* is invariant to ρ for any γ .

We next show that, to assess the power of the τ^* test applied to the model (1), it is only necessary to consider the case of $\rho = 0$. We can write (1) as

$$\begin{aligned}(1 - \phi L)(y_{1t} - y_{2t}) &= \{2(1 - \rho)\}^{1/2} e_{1t} \\ (1 - L)(y_{1t} + y_{2t}) &= \{2(1 + \rho)\}^{1/2} e_{2t}\end{aligned}$$

where L is the lag operator and the white noise innovations $(e_{1t}, e_{2t})'$ have identity covariance matrix. Then, for any b , it follows that

$$\begin{aligned}(1 - \phi L)(1 - L)(y_{1t} - by_{2t}) &= \{(1 - \rho) / 2\}^{1/2} (1 + b)(1 - L)e_{1t} \\ &\quad + \{(1 + \rho) / 2\}^{1/2} (1 - b)(1 - \phi L)e_{2t}\end{aligned}\tag{3}$$

The required results follow directly from (3). Specifically we shall show that, in the case where $\rho = 0$, there exists b^* such that $(y_{1t} - b^* y_{2t})$ is precisely the same series as $(y_{1t} - by_{2t})$ in the $\rho \neq 0$ case. We need only consider $|b| \leq 1$, as for $|b| > 1$, we might just as well analyse $(y_{2t} - dy_{1t})$ for $|d| < 1$.

The $\rho = 0$ variant of (3) is given by

$$(1 - \phi L)(1 - L)(y_{1t} - b^* y_{2t}) = 2^{-1/2} \{(1 + b^*)(1 - L)e_{1t} + (1 - b^*)(1 - \phi L)e_{2t}\}\tag{4}$$

First, if $b = \pm 1$, the right-hand side of (3) is a constant multiple of the right-hand side of (4) with $b^* = \pm 1$. In the more general case, the right-hand side of (3) is a constant multiple of the right-hand side of (4) if

$$\frac{k(1+b)}{(1-b)} = \frac{(1+b^*)}{(1-b^*)}, \quad k = \left[\frac{1-\rho}{1+\rho} \right]^{1/2}$$

Then, since $k > 0$, it follows that for every b , with $|b| < 1$, there exists a corresponding b^* , with $|b^*| < 1$, and vice versa. This establishes that the set of τ_b Dickey-Fuller statistics is the same, whatever the value of ρ , for a specific realisation of the white noise generating series $(e_{1t}, e_{2t})'$. Therefore, both the size and power of the τ^* test are invariant to ρ for series generated from the model (1).

Table 4 shows powers of the τ^* test for model (1). Comparing these with the results of Table 1, we find that τ^* has appreciably higher power than the JT test. In fact, the new test has about the same power as the EGDF test when $\rho = 0$. For ρ close to 1, EGDF is noticeably more powerful than τ^* . This is to be expected, however, since in that case, as we see from Table 2, b is quite precisely estimated through the regression (2).

There is a sense in which the relatively strong performance of the EGDF test is fortuitous when ρ is not very close to one. For example, in the case $\rho = 0$, EGDF is noticeably more powerful than the Johansen test, and about as powerful as the τ^* test. Yet, in that case, as can be seen from Table 2, the least squares estimator of b in (2) is quite severely biased. To see why the EGDF test nevertheless has respectable power, let \hat{b} be any fixed number. Then precisely as in (3),

$$\begin{aligned} (1-\phi L)\Delta(y_{1t} - \hat{b}y_{2t}) \\ = 2^{-1/2} \{ (1-\rho)^{1/2} (1+\hat{b})(1-L)e_{1t} + (1+\rho)^{1/2} (1-\hat{b})(1-\phi L)e_{2t} \} \end{aligned}$$

It then follows that $(y_{1t} - \hat{b}y_{2t})$ is ARIMA (1,1,1), with a moving average parameter between ϕ and 1, when $|\hat{b}| < 1$. It is well known (see, for example, Schwert 1989 and Agiakloglou and Newbold 1992) that Dickey-Fuller regressions lead to frequent rejections of the unit root hypothesis for $I(1)$ generating models with large positive moving average parameters.

The τ^* statistic performs impressively for the model (1), which is of course a very special case of cointegration. In the next section we discuss an augmented

variant of the τ^* test, appropriate for general vector autoregressive generating models. However, even in the non-augmented case, the simulation results of Table 4 apply to a much broader set of models than simply (1). Starting with (1), consider the transformation

$$x_{1t} = cy_{1t} + dy_{2t}, \quad x_{2t} = (c-1)y_{1t} + (d+1)y_{2t}$$

Then $(x_{1t} - x_{2t}) = (y_{1t} - y_{2t})$. Also, the set of linear functions of (x_{1t}, x_{2t}) is the same as the set of linear functions of (y_{1t}, y_{2t}) , so the distribution of the τ^* test statistic is invariant to this transformation. From (1), we then have

$$\begin{aligned} \Delta x_{1t} &= -1/2(c-d)(1-\phi)(x_{1,t-1} - x_{2,t-1}) + u_{1t} \\ \Delta x_{2t} &= 1/2(d-c+2)(1-\phi)(x_{1,t-1} - x_{2,t-1}) + u_{2t} \end{aligned}$$

where the white noise innovations $(u_{1t}, u_{2t})'$ have covariance matrix

$$\Omega = M \Sigma M', \quad M = \begin{bmatrix} c & d \\ (c-1) & (d+1) \end{bmatrix}$$

Then, all values of (c, d, ρ) in this model yield τ^* statistics with identical sampling distributions. In particular, unidirectional casual models result from setting $c = d$. The usual error-correction interpretation in this model requires $0 \leq (c-d) < 2$.

3. The General Case

We begin by deriving a representation for the asymptotic distribution of τ^* in the case of no cointegration. The model we consider is more general than that in (1) as we allow v_t to follow a stationary vector process. This leads us to construct a test based on the infimum of a transformation of τ_b in which τ^* arises as a special case.

Let $y_t = (y_{1t}, y_{2t})'$ follow the $I(1)$ process given by

$$y_t = y_{t-1} + v_t, \quad t = 1, \dots, T,$$

where v_t is a stationary vector process with

$$G \equiv E(v_t v_t'), \quad \Omega \equiv \lim_{T \rightarrow \infty} T^{-1} E\left(\sum_{t=1}^T v_t \sum_{s=1}^T v_s'\right), \quad \Delta \equiv \lim_{T \rightarrow \infty} T^{-1} \sum_{t=2}^T \sum_{s=1}^{t-1} E(v_s v_t').$$

First, we define $x_t \equiv \beta' y_t$ where $\beta' = (1, -b)$. Hence, x_t is an $I(1)$ process for any b : $x_t = x_{t-1} + \xi_t$ where $\xi_t \equiv \beta' v_t$ and ξ_t is a serially correlated stationary process. Let τ_b be the t -statistic obtained from the regression of Δx_t on a constant and x_{t-1} . Then, the Phillips-Perron statistic denoted by $z(\tau_b)$ is given by

$$z(\tau_b) \equiv (\beta' G \beta)^{1/2} (\beta' \Omega \beta)^{-1/2} \tau_b - (\beta' \Omega \beta)^{-1/2} \left\{ T^{-2} \sum (x_{t-1} - \bar{x}_{-1})^2 \right\}^{-1/2} \beta' (\Omega - G) \beta / 2.$$

Theorem 1. Suppose that the sequence $\{v_t\}_{t=1}^{\infty}$ satisfies the conditions of Theorem 2.1 in Phillips and Durlauf (1986). Then, for any non-empty compact interval $\mathfrak{S} \equiv [b_1, b_2]$,

$$z^* \equiv \inf_{b \in \mathfrak{S}} z(\tau_b) \Rightarrow \inf_{b \in \mathfrak{S}} \ell_b$$

where

$$\ell_b \equiv \frac{(\beta' \Omega \beta)^{-1} \beta' K \left\{ \int_0^1 W(r) dW(r)' - \int_0^1 W(r) dr W(1)' \right\} K' \beta}{\left[(\beta' \Omega \beta)^{-1} \beta' K \left\{ \int_0^1 W(r) W(r)' dr - \int_0^1 W(r) dr \int_0^1 W(r)' dr \right\} K' \beta \right]^{1/2}}$$

and $W(r)$ is a bivariate vector Brownian motion defined as the limit of $K^{-1} (T^{-1/2} \sum_{t=1}^{\lfloor Tr \rfloor} v_t)$ with K defined as $KK' = \Omega$.

Proof: See Appendix.

The limiting random variable ℓ_b appears to contain nuisance parameters. However, note that it can be expressed more intuitively by defining

$$w_b(r) \equiv (\beta' \Omega \beta)^{-1/2} \beta' K W(r)$$

which is the limiting random variable of $(\beta'\Omega\beta)^{-1/2}T^{-1/2}\sum_{t=1}^{\lfloor Tr \rfloor}\xi_t$ and hence is a standard Brownian motion process. Using this notation, we have

$$\ell_b = \left[\int_0^1 w_b(r)^2 dr - \left\{ \int_0^1 w_b(r) dr \right\}^2 \right]^{-1/2} \left\{ \int_0^1 w_b(r) dw_b(r) - \int_0^1 w_b(r) dr w_b(1) \right\}$$

which is free of all nuisance parameters as in Phillips and Perron (1988). It should be noted that in the proof of Theorem 1 there is no restriction on the dimension of b . Hence, the extension of the theorem to a more general case where there are more than two cointegrated variables should be straightforward. In the case where $v_t \sim IID(0, \Sigma)$ then $G = \Omega = \Sigma$, $z(\tau_b) = \tau_b$ and so $z^* = \tau^*$, which is the simplified statistic considered in Section 2.

When v_t is a pure autoregressive process of finite order it is simple to construct an operational version of the test z^* . We suppose v_t follows the VAR(p) process

$$v_t = \sum_{i=1}^p A_i v_{t-i} + \eta_t, \quad \eta_t \sim IID(0, \Sigma)$$

where the roots of $|I - \sum_{i=1}^p A_i \lambda^i| = 0$ lie outside the unit circle and Σ is a positive definite matrix. Then G is implicitly defined by

$$G = \sum_{i=1}^p A_i G A_i' + \Sigma$$

which has the solution

$$vec(G) = \left\{ I - \sum_{i=1}^p (A_i \otimes A_i) \right\}^{-1} vec(\Sigma)$$

Further,

$$\Omega = \left(I - \sum_{i=1}^p A_i \right)^{-1} \Sigma \left(I - \sum_{i=1}^p A_i \right)'^{-1}$$

Now, since

$$y_t = y_{t-1} + \sum_{i=1}^p A_i \Delta y_{t-i} + \eta_t$$

then when the vector regression

$$\Delta y_t = \hat{\mu} + \sum_{i=1}^p \hat{A}_i \Delta y_{t-i} + \hat{\eta}_t$$

is estimated by least squares it follows from Hamilton (1994, Ch. 18) that $\hat{A}_i \rightarrow_p A_i$ and $\hat{\Sigma} = T^{-1} \sum_{t=1}^T \hat{\eta}_t \hat{\eta}_t' \rightarrow_p \Sigma$. Hence, G is consistently estimated by \hat{G} formed from

$$\text{vec}(\hat{G}) = \{I - \sum_{i=1}^p (\hat{A}_i \otimes \hat{A}_i)\}^{-1} \text{vec}(\hat{\Sigma})$$

and Ω is consistently estimated by

$$\hat{\Omega} = (I - \sum_{i=1}^p \hat{A}_i)^{-1} \hat{\Sigma} (I - \sum_{i=1}^p \hat{A}_i)'^{-1}$$

We now extend Theorem 1 to make the infimum statistic operational.

Theorem 2. Suppose that the sequence $\{v_t\}_{t=1}^{\infty}$ satisfies the conditions of Theorem 2.1 in Phillips and Durlauf (1986). Let \hat{G} and $\hat{\Omega}$ be any consistent estimators for G and Ω respectively, not depending on $b \in \mathfrak{S}$. Define $z_p(\tau_b)$ by

$$z_p(\tau_b) \equiv (\beta' \hat{G} \beta)^{1/2} (\beta' \hat{\Omega} \beta)^{-1/2} \tau_b - (\beta' \hat{\Omega} \beta)^{-1/2} \{T^{-2} \sum (x_{t-1} - \bar{x}_{-1})^2\}^{-1/2} \beta' (\hat{\Omega} - \hat{G}) \beta / 2.$$

Then, for any non-empty compact interval $\mathfrak{S} \equiv [b_1, b_2]$,

$$z_p^* \equiv \inf_{b \in \mathfrak{S}} z_p(\tau_b) \Rightarrow \inf_{b \in \mathfrak{S}} \ell_b$$

where ℓ_b is defined as in Theorem 1.

Proof: See Appendix.

In Theorem 2, we require \hat{G} and $\hat{\Omega}$ not to depend on b . This condition, however, is more than we actually need to prove the theorem. In fact, \hat{G} and $\hat{\Omega}$ could be allowed to depend on b as long as they can be expressed as continuous functions of b and

some sample moments which are $O_p(1)$. In this situation the argument in the proof still goes through without any major modification.

The main feature of the above approach is that it avoids weaknesses inherent with the variants of τ^* based on augmented Dickey-Fuller style regressions. As a simple example, suppose $p = 1$ so that we have the model

$$\Delta y_t = A_1 \Delta y_{t-1} + \eta_t$$

Then

$$\Delta x_t = \beta' A_1 \Delta y_{t-1} + \omega_t \neq a_1 \Delta x_{t-1} + \omega_t$$

except in the particular case when $A_1 = a_1 I$. So, in general, Δx_t will not follow an $AR(p)$ process and hence ADF-style corrections to unit root tests are, strictly speaking, not appropriate. Our suggested procedure, however, corrects for autocorrelation by fitting models explicitly to Δy_t rather than Δx_t and so avoids this problem. Moreover, since the actual process followed by Δx_t depends on b , so too will the degree of any autoregressive approximation used for ADF tests. Thus, we would need to allow the numbers of lagged terms in Δx_t included in the ADF regression to vary with b which would make finding the appropriate minimum ADF test over b computationally very expensive. On the other hand, in our approach, b does not influence the order of the model which needs to be estimated.

In this more general setup, we need to assess the effects of estimating the A_i and Σ on the finite sample size and power of the test z_p^* . To this end, we simulated the model (1) with $\gamma = 0$, assuming v_t follows

$$v_t = A_1 v_{t-1} + \eta_t$$

where $\eta_t = (\eta_{1t}, \eta_{2t})'$ is generated as Gaussian white noise with unit variances and covariance ρ . The size of z_1^* for different structures of A_1 and values of ρ is shown in Table 5; the critical values used for z_1^* being those given in Table 3. Also shown is the size of the Johansen trace statistic, assuming one lagged term in Δy_t is incorporated in the vector autoregression (denoted as JT_1). For $T = 100$, the sizes of

z_1^* are generally close to their nominal values, even with $A_1 = 0.8I$ (so that y_t begins to resemble an $I(2)$ process). It is only when each element of A_1 is 0.4 that z_1^* appears to be slightly over-sized. However, problems of over-size are rather more apparent for JT_1 , most noticeably when $A_1 = 0.8I$. With $T = 200$, the sizes of z_1^* are all broadly correct, whereas JT_1 remains over-sized in some cases. Finally, it is interesting to note that, for given A_1 , the size of z_1^* appears little affected by changing ρ . The results of the previous Section may well be relevant in explaining this phenomenon.

Table 6 examines the powers of z_1^* and JT_1 to detect cointegration for sample size $T = 100$. Since the size of the tests does not depend on ρ to any great extent, we set $\rho = 0$ throughout these simulations. In all the cases except when $A_1 = 0.8I$, z_1^* is generally seen to be more powerful than JT_1 , often considerably so, and with power gains of up to 22%. When $A_1 = 0.8I$ the powers appear roughly equal, but only because JT_1 is over-sized in this case. Table 7 reports results from a similar exercise using $T = 200$. Consistency of the test z_1^* is clearly evident, which is as we would expect given its origins in the Dickey-Fuller methodology. Once more, z_1^* is generally more powerful than JT_1 , with gains of around 10% in mid-range power.

4. An Empirical Example

As a simple example of our test procedure, we test for cointegration between short and long term U.K. interest rates. The short term rate (y_{1t}) is the 91 day U.K. Treasury Bill rate and the long term rate (y_{2t}) is the yield on 20 year U.K. gilts. The data are quarterly from 1952Q1-1988Q4 (148 observations), and were obtained from Mills (1993). Augmented Dickey-Fuller tests applied to each series suggest that both are $I(1)$ without drift, not rejecting the $I(1)$ null even at the 0.10 significance level (we omit the detailed results here). Table 8 gives the values of the cointegration test statistics z_p^* and JT_p for $p = 0, 1, \dots, 5$, where p is the order of the VAR model fitted

to Δy_t . Using standard t-ratios to assess significance of the estimated coefficients on lagged terms in Δy_t suggested $p = 3$ as the preferred order. Also shown, as b_{\min} in the table, is the value of b found to minimize $z_p(\tau_b)$, for each p . The Johansen test fails to reject at the 0.10-level the null hypothesis of no cointegration between the interest rates for any of the lag lengths considered here, whereas z_p^* rejects non-cointegration at the 0.05-level with $p = 1, 2, \dots, 4$, and at the 0.10-level with $p = 0, 5$. For this particular data set then, the z_p^* test suggests strong evidence of cointegration between short and long term interest rates, while the JT_p test does not.

5. Summary

In this paper we have introduced a new test of the null hypothesis of no cointegration between a pair of $I(1)$ time series. This test was motivated in Section 2 by observing the extremely variable power of the Engle-Granger/Dickey-Fuller (EGDF) test, relative to the Johansen trace (JT) test, over a wide range of values of a single parameter in a simple model. It was seen that the very poor performance of the EGDF test is associated with the severe bias in the least squares estimator of the parameter in the cointegrating regression. We therefore proposed circumventing the estimation of the cointegrating regression by, in effect, applying the Dickey-Fuller test to all linear combinations of the two time series. Our test statistic, in the simple case of Section 2, is then the Dickey-Fuller t-ratio that is least favourable to the null hypothesis. We saw, for the simple model, that neither the JT test nor the new test exhibit the volatility in power characteristic of the EGDF test. Moreover, simulation evidence indicated that the new test was a good deal more powerful than the JT test. A subsidiary benefit of our test, compared with EGDF, is that the test outcome does not depend on the arbitrary choice of one series as the dependent variable in the cointegrating regression.

Of course, for our test to have much practical value, it must be applicable in far more general situations than that of Section 2, where every linear combination of the series is a random walk under the null hypothesis, and a particular linear combination is a stationary first order autoregression under the alternative. An obvious possible

extension would be to base the new test on augmented Dickey-Fuller regression, by analogy with EGDF. However, we prefer an alternative generalisation, for two reasons. First, we wanted a direct comparison with the JT test for finite order vector autoregression. However, linear combinations of series generated by such models generally have infinite autoregressive order, requiring some truncation rule for fixing the number of lags in the Dickey-Fuller regressions. Second, since we are considering all possible linear combinations of the series, choosing the number of lags in this way is both problematic and computationally burdensome. Different linear combinations of the series will have different generating models. This suggests the necessity either of a different number of lags for every linear combination, which is practically impossible, or the fixing for every possibility of a very high number of lags, which would inevitably reduce the power of the test.

In Section 3 we proposed a general version of our test, which incorporates the test of Section 2 as a special case, based on the prior fitting of vector autoregressions to first differences of the time series. Simulation results confirmed that this general test has satisfactory size properties - more satisfactory in some cases than the JT test. Moreover, the relative superiority in power of the new test, observed in the simple case of Section 2, continues to hold for the vector autoregressions examined in a simulation experiment in Section 3. Finally, in Section 4 we applied the new test and the JT test to series of short-and long-term interest rates. While the former suggested strong evidence of cointegration, the latter did not, in a situation where several authors, including Engle and Granger (1987), have suggested that cointegration might reasonably be expected on *a priori* grounds.

Appendix

Proof of Theorem 1. The t -statistic τ_b is given by

$$\tau_b \equiv \frac{T(\hat{\rho} - 1)}{\hat{\sigma} \{T^{-2} \sum (x_{t-1} - \bar{x}_{-1})^2\}^{-1/2}}$$

where $\hat{\rho} \equiv \frac{\sum (x_{t-1} - \bar{x}_{-1})(x_t - \bar{x})}{\sum (x_{t-1} - \bar{x}_{-1})^2}$ and $\hat{\sigma}^2 \equiv T^{-1} \sum \{(x_t - \bar{x}) - \hat{\rho}(x_{t-1} - \bar{x}_{-1})\}^2$. Each

term in τ_b can be expressed as a continuous function of b and some sample moments as follows:

$$T(\hat{\rho} - 1) = \frac{\beta' A_T \beta}{\beta' B_T \beta}$$

$$\hat{\sigma}^2 = \beta'(T^{-1} \sum v_t v_t') \beta - \beta'(T^{-1} \sum v_t)(T^{-1} \sum v_t') \beta - T^{-1} \frac{(\beta' A_T \beta)^2}{\beta' B_T \beta}$$

$$T^{-2} \sum (x_{t-1} - \bar{x}_{-1})^2 = \beta' B_T \beta$$

where

$$A_T \equiv T^{-1} \sum y_{t-1} v_t' - T^{-3/2} \sum y_{t-1} T^{-1/2} \sum v_{t-1}'$$

$$B_T \equiv T^{-2} \sum y_{t-1} y_{t-1}' - T^{-3/2} \sum y_{t-1} T^{-3/2} \sum y_{t-1}'.$$

Therefore, we have

$$\begin{aligned} z(\tau_b) &= \frac{(\beta' G \beta)^{1/2}}{(\beta' \Omega \beta)^{1/2}} \left\{ \beta'(T^{-1} \sum v_t v_t') \beta - \beta'(T^{-1} \sum v_t)(T^{-1} \sum v_t') \beta - T^{-1} \frac{(\beta' A_T \beta)^2}{\beta' B_T \beta} \right\}^{-1/2} \frac{\beta' A_T \beta}{(\beta' B_T \beta)^{1/2}} \\ &\quad - (\beta' \Omega \beta)^{-1/2} (\beta' B_T \beta)^{-1/2} \beta' (\Omega - G) \beta / 2 \\ &\equiv u(Z_T, b) \end{aligned} \tag{5}$$

where Z_T is a $m \times 1$ random vector with $m = 19$ defined as:

$$\begin{aligned} Z_T' &\equiv [T^{-1} \text{vec}(\sum y_{t-1} v_t'), T^{-3/2} \sum y_{t-1}', T^{-1/2} \sum v_{t-1}', T^{-2} \text{vec}(\sum y_{t-1} y_{t-1}'), \\ &\quad T^{-1} \text{vec}(\sum v_t v_t'), T^{-1} \sum v_t', T^{-1}]. \end{aligned}$$

The identity in the second line of (5) defines the function $u(a,b)$ from $R^m \times R$ to R .

Next, we define a function $h(a)$ from R^m to R :

$$h(a) \equiv \inf_{b \in \phi(a)} u(a,b).$$

The function $\phi(a)$ is a set function from R^m to 2^R (power set of R) defined as:

$$\phi(a) \equiv [b_1, b_2]$$

for any $a \in R^m$ where b_1, b_2 are given in the theorem. First, note that $h(Z_T) = \inf_{b \in \mathcal{S}} z(\tau_b)$ whose limiting distribution we want to obtain. The next step is to show (i) the joint convergence of Z_T in distribution and (ii) the continuity of $h(a)$.

Since the last three elements in Z_T have degenerate limiting distributions and

$$[T^{-1} \text{vec}(\sum v_t v_t'), T^{-1} \sum v_t', T^{-1}] \rightarrow_p [\text{vec}(G)', 0_{1 \times 2}, 0] \equiv Z_2'$$

it is sufficient to show the joint convergence of the first four components. Using the functional central limit theorem and continuous mapping theorem, we have

$$[T^{-1} \text{vec}(\sum y_{t-1} v_t'), T^{-3/2} \sum y_{t-1}', T^{-1/2} \sum v_{t-1}', T^{-2} \text{vec}(\sum y_{t-1} y_{t-1}')] \Rightarrow \\ [\{\text{vec}(K \int_0^1 W(r) dW(r)' K' + \Delta)\}', \{K \int_0^1 W(r) dr\}', \{KW(1)\}', \{\text{vec}(K \int_0^1 W(r) W(r)' dr K')\}'] \equiv Z_1'.$$

See Sims, Stock and Watson (1990), Chan and Wei (1988) for more details. Hence, we have $Z_T \Rightarrow Z$ where $Z' \equiv (Z_1', Z_2')$.

Next, we need to invoke the maximum theorem in Berge (1963) to show that $h(a)$ is continuous. The four conditions we need to verify are:

- (i) The spaces R^m and R on which $u(a,b)$ is defined are Hausdorff.
- (ii) The function $u(a,b)$ is continuous in both arguments.
- (iii) For all $a \in R^m$, $\phi(a)$ is not empty.
- (iv) The set function $\phi(a)$ is continuous.

The first three conditions are trivially satisfied. The last condition is proved by showing that $\phi(a)$ is upper semicontinuous and lower semicontinuous (See Takayama (1985) for definitions). Let $a^0 \in R^m$ and V be an open set containing $\phi(a^0) = \mathfrak{S}$. Define $N(a^0) \equiv \{a \in R^m : \|a - a^0\| < 1\}$ which is not empty. Then, for any $a \in N(a^0)$, we have $\phi(a) \subset V$ which implies that $\phi(a)$ is upper semicontinuous at a^0 . Since a^0 is an arbitrary point in R^m , $\phi(a)$ is upper semicontinuous in R^m . Next, we show that $\phi(a)$ is lower semicontinuous. Let V and $N(a^0)$ be the same sets defined as above. Then, for any $a \in R^m$, $\phi(a) \cap V = \mathfrak{S} \neq \emptyset$ which implies that $\phi(a)$ is lower semicontinuous at a^0 . Therefore, $h(a)$ is continuous and we have

$$h(Z_T) \Rightarrow h(Z)$$

by the continuous mapping theorem. Note that $h(Z) = \inf_{b \in \mathfrak{S}} u(Z, b)$ where

$$u(Z, b) = (\beta' \Omega \beta)^{-1/2} (\beta' B \beta)^{-1/2} \beta' A \beta - (\beta' \Omega \beta)^{-1/2} (\beta' B \beta)^{-1/2} \beta' (\Omega - G) \beta / 2$$

$$A \equiv K \int_0^1 W(r) dW(r)' K' + \Delta - K \int_0^1 W(r) dr W(1)' K'$$

$$B \equiv K \int_0^1 W(r) W(r)' dr K' - K \int_0^1 W(r) dr \int_0^1 W(r)' dr K'.$$

Using $\beta' \Delta \beta = \beta' (\Delta + \Delta') \beta / 2 = \beta' (\Omega - G) \beta / 2$, $u(Z, b)$ can be simplified to

$$\begin{aligned} u(Z, b) &= \frac{(\beta' \Omega \beta)^{-1} \beta' K \{ \int_0^1 W(r) dW(r)' - \int_0^1 W(r) dr W(1)' \} K' \beta}{[(\beta' \Omega \beta)^{-1} \beta' K \{ \int_0^1 W(r) W(r)' dr - \int_0^1 W(r) dr \int_0^1 W(r)' dr \} K' \beta]^{1/2}} \\ &= \ell_b. \end{aligned}$$

Hence, $h(Z) = \inf_{b \in \mathfrak{S}} \ell_b$ which completes the proof. ■

Proof of Theorem 2. Basically the same arguments as used in Theorem 1 go through in the proof of Theorem 2. Note that

$$\begin{aligned}
z_p(\tau_b) &= \frac{(\beta' \hat{G} \beta)^{1/2}}{(\beta' \hat{\Omega} \beta)^{1/2}} \left\{ \beta' (T^{-1} \sum v_i v_i') \beta - \beta' (T^{-1} \sum v_i) (T^{-1} \sum v_i') \beta - T^{-1} \frac{(\beta' A_T \beta)^2}{\beta' B_T \beta} \right\}^{-1/2} \frac{\beta' A_T \beta}{(\beta' B_T \beta)^{1/2}} \\
&\quad - (\beta' \hat{\Omega} \beta)^{-1/2} \{ \beta' B_T \beta \}^{-1/2} \beta' (\hat{\Omega} - \hat{G}) \beta / 2 \\
&\equiv u_p(\hat{Z}_T, b)
\end{aligned}$$

where \hat{Z}_T is a $m \times 1$ random vector with $m = 23$ defined as:

$$\hat{Z}_T' \equiv [Z_T', \text{vec}(\hat{G})', \text{vec}(\hat{\Omega})'].$$

Define a function $h_p(a)$ from R^m to R : $h_p(a) \equiv \inf_{b \in \mathfrak{S}} u_p(a, b)$. It can be shown that

- (i) $\hat{Z}_T \Rightarrow Z_p$ where $Z_p' \equiv (Z', \text{vec}(G)', \text{vec}(\Omega)')$
- (ii) $h_p(a)$ is continuous.

Therefore, we have $h_p(\hat{Z}_T) \Rightarrow h_p(Z_p)$ by the continuous mapping theorem. Note that

- (i) $h_p(\hat{Z}_T) = \inf_{b \in \mathfrak{S}} z_p(\tau_b)$
- (ii) $h_p(Z_p) = \inf_{b \in \mathfrak{S}} \ell_b$

which completes the proof. ■

References

- AGIAKLOGLOU, C. and NEWBOLD, P. (1992), Empirical evidence on Dickey-Fuller-type tests, *Journal of Time Series Analysis*, 13, 471-483.
- BANERJEE, A., DOLADO, J.J., HENDRY, D.F. and SMITH, G.W. (1986), Exploring equilibrium relationships in econometrics through static models: Some Monte Carlo evidence, *Oxford Bulletin of Economics and Statistics*, 48, 253-277.
- BERGE, C. (1963), *Topological Spaces*, tr. By Patterson, New York, Macmillan, (French original, 1959).
- BLANGIEWICZ, M. and CHAREMZA, W. (1990), Cointegration in small samples: Empirical percentiles, drifting moments and customized testing, *Oxford Bulletin of Economics and Statistics*, 52, 303-315.
- CHAN, N.H. and WEI, C.Z. (1988), Limiting distributions of least squares estimates of unstable autoregressive processes, *Annals of Statistics*, 16, 367-401.
- ENGLE, R.F. and GRANGER, C.W.J. (1987), Cointegration and error correction: Representation, estimation and testing, *Econometrica*, 55, 251-276.
- ENGLE, R.F. and YOO, B.S. (1987), Forecasting and testing in cointegrated systems, *Journal of Econometrics*, 35, 143-159.
- GRANGER, C.W.J. (1981), Some properties of time series data and their use in econometric model specification, *Journal of Econometrics*, 16, 121-130.
- GONZALO, J. (1994), Five alternative methods of estimating long-run equilibrium relationships, *Journal of Econometrics*, 60, 203-233.
- HAMILTON, J.D. (1994), *Time Series Analysis*. Princeton, New Jersey: Princeton University Press.
- HANSEN, B.E and PHILLIPS, P.C.B. (1990), Estimation and inference in models of cointegration: A simulation study, *Advances in Econometrics*, 8, 225-248.
- JOHANSEN, S. (1988), Statistical analysis of cointegration vectors, *Journal of Economic Dynamics and Control*, 12, 231-254.
- JOHANSEN, S. (1991), Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models, *Econometrica*, 59, 1551-1580.
- JOHANSEN, S. and JUSELIUS, K. (1990), Maximum likelihood estimation and inference on cointegration - with applications to the demand for money, *Oxford Bulletin of Economics and Statistics*, 52, 169-210.
- MacKINNON, J.G. (1991), Critical values for cointegration tests, in Engle, R.F. and Granger, C.W.J. (eds), *Long-Run Economic Relationships* (Oxford: Oxford University Press).
- MILLS, T.C. (1993), *The Econometric Modelling of Financial Time Series*. Cambridge: Cambridge University Press.
- NG, S. and PERRON, P. (1995), Unit root tests in ARMA models with data-dependent methods for the selection of the truncation lag, *Journal of the American Statistical Association*, 90, 268-281.
- NG, S. and PERRON, P. (1997), Estimation and inference in nearly unbalanced nearly cointegrated systems, *Journal Econometrics*, 79, 53-81.
- OSTERWALD-LENUM, M. (1992), A note with fractiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics: Four cases, *Oxford Bulletin of Economics and Statistics*, 54, 461-472.

- PHILLIPS, P.C.B. and DURLAUF, S.N. (1986), Multiple time series regression with integrated processes, *Review of Economic Studies*, 53, 473-495.
- PHILLIPS, P.C.B and PERRON, P. (1988), Testing for a unit root in time series regression, *Biometrika*, 75, 335-346.
- SCHWERT, G.W. (1989), Tests for unit roots: A Monte Carlo investigation, *Journal of Business and Economic Statistics*, 7, 147-159.
- SIMS, C.A., STOCK, J.H. and WATSON, M.W. (1990), Inference in linear time series models with some unit roots, *Econometrica*, 58, 113-144.
- TAKAYAMA, A. (1985), *Mathematical Economics*, Cambridge, London: Cambridge University Press.

Table 1. Power of tests EGDF and JT for $T = 100$ and $T = 200$ at nominal 0.10 and 0.05 critical values (4000 replications).					
ρ	ϕ	$T = 100$		$T = 200$	
		EGDF	JT	EGDF	JT
-0.9	0.9	0.098	0.252	0.179	0.609
		0.051	0.148	0.096	0.420
-0.9	0.8	0.170	0.621	0.556	0.994
		0.095	0.433	0.408	0.970
-0.9	0.7	0.323	0.924	0.868	1.000
		0.196	0.809	0.767	1.000
-0.6	0.9	0.173	0.253	0.454	0.609
		0.086	0.140	0.279	0.421
-0.6	0.8	0.441	0.621	0.955	0.992
		0.292	0.434	0.893	0.967
-0.6	0.7	0.775	0.925	0.999	1.000
		0.663	0.821	0.995	1.000
-0.3	0.9	0.228	0.255	0.607	0.611
		0.120	0.139	0.417	0.421
-0.3	0.8	0.615	0.620	0.990	0.993
		0.420	0.432	0.973	0.967
-0.3	0.7	0.916	0.925	1.000	1.000
		0.821	0.828	0.999	1.000
0.0	0.9	0.266	0.252	0.703	0.604
		0.145	0.142	0.509	0.418
0.0	0.8	0.718	0.618	0.999	0.994
		0.525	0.428	0.990	0.968
0.0	0.7	0.962	0.922	1.000	1.000
		0.905	0.824	1.000	1.000
0.3	0.9	0.303	0.252	0.756	0.606
		0.167	0.142	0.579	0.414
0.3	0.8	0.781	0.620	0.999	0.993
		0.592	0.423	0.997	0.967
0.3	0.7	0.979	0.920	1.000	1.000
		0.935	0.821	1.000	1.000
0.6	0.9	0.329	0.249	0.798	0.601
		0.185	0.140	0.618	0.409
0.6	0.8	0.821	0.618	1.000	0.994
		0.646	0.434	0.999	0.969
0.6	0.7	0.987	0.922	1.000	1.000
		0.957	0.814	1.000	1.000
0.9	0.9	0.354	0.252	0.823	0.600
		0.203	0.141	0.651	0.418
0.9	0.8	0.842	0.613	1.000	0.994
		0.677	0.439	1.000	0.967
0.9	0.7	0.995	0.917	1.000	1.000
		0.965	0.814	1.000	1.000

Table 2. Means and standard deviations of \hat{b} from regression (2) for model (1) where $b = 1$, with $T = 100$ (2000 replications).

ρ	ϕ	Mean	Std. dev.
-0.9	0.9	-0.68	0.30
-0.9	0.8	-0.52	0.32
-0.9	0.7	-0.42	0.33
-0.6	0.9	-0.08	0.44
-0.6	0.8	0.13	0.39
-0.6	0.7	0.26	0.36
0.0	0.9	0.51	0.36
0.0	0.8	0.65	0.27
0.0	0.7	0.72	0.22
0.6	0.9	0.84	0.21
0.6	0.8	0.89	0.14
0.6	0.7	0.92	0.10
0.9	0.9	0.96	0.10
0.9	0.8	0.98	0.06
0.9	0.7	0.98	0.04

Table 3. Critical values of τ^* (20000 replications).

T	0.10	0.05	0.01
25	-3.61	-3.96	-4.75
50	-3.46	-3.77	-4.40
100	-3.41	-3.70	-4.24
200	-3.37	-3.65	-4.21
400	-3.36	-3.63	-4.16

Table 4. Power of test τ^* for $T = 100$ and $T = 200$ at nominal 0.10 and 0.05 critical values (4000 replications).

ϕ	$T = 100$	$T = 200$
0.9	0.266	0.695
	0.153	0.507
0.8	0.721	1.000
	0.516	0.997
0.7	0.975	1.000
	0.922	1.000

Table 5. Size of Tests z_1^* and JT_1 for $T = 100$ and $T = 200$ at nominal 0.10 and 0.05 critical values (4000 replications).						
A_1	ϕ	ρ	$T = 100$		$T = 200$	
			z_1^*	JT_1	z_1^*	JT_1
[0.0 0.0, 0.0 0.0]	0.0	0.0	0.118 0.059	0.112 0.056	0.111 0.056	0.110 0.053
[0.4 0.0, 0.0 0.4]	0.0	0.0	0.094 0.045	0.122 0.061	0.096 0.046	0.113 0.054
[0.4 0.2, 0.2 0.4]	0.0	0.0	0.101 0.049	0.126 0.061	0.096 0.048	0.116 0.056
[0.8 0.0, 0.0 0.8]	0.0	0.0	0.113 0.062	0.171 0.093	0.096 0.047	0.137 0.069
[0.4 0.4, 0.4 0.4]	0.0	0.0	0.125 0.069	0.144 0.077	0.107 0.056	0.125 0.064
[0.0 0.0, 0.0 0.0]	0.0	0.5	0.119 0.059	0.112 0.056	0.110 0.056	0.110 0.053
[0.4 0.0, 0.0 0.4]	0.0	0.5	0.094 0.045	0.122 0.061	0.095 0.045	0.113 0.054
[0.4 0.2, 0.2 0.4]	0.0	0.5	0.098 0.049	0.123 0.062	0.095 0.047	0.116 0.057
[0.8 0.0, 0.0 0.8]	0.0	0.5	0.108 0.058	0.171 0.093	0.097 0.045	0.137 0.069
[0.4 0.4, 0.4 0.4]	0.0	0.5	0.115 0.066	0.144 0.077	0.098 0.046	0.123 0.062
[0.0 0.0, 0.0 0.0]	0.0	-0.5	0.118 0.057	0.112 0.056	0.112 0.056	0.110 0.053
[0.4 0.0, 0.0 0.4]	0.0	-0.5	0.094 0.045	0.122 0.061	0.094 0.046	0.113 0.054
[0.4 0.2, 0.2 0.4]	0.0	-0.5	0.099 0.050	0.125 0.062	0.100 0.046	0.112 0.055
[0.8 0.0, 0.0 0.8]	0.0	-0.5	0.108 0.059	0.171 0.093	0.100 0.048	0.137 0.069
[0.4 0.4, 0.4 0.4]	0.0	-0.5	0.123 0.068	0.145 0.077	0.108 0.058	0.125 0.063

Table 6. Power of Tests z_1^* and JT_1 for $T = 100$, $\rho = 0$ at nominal 0.10 and 0.05 critical values (4000 replications).			
A_1	ϕ	z_1^*	JT_1
[0.0 0.0, 0.0 0.0]	0.90	0.311 0.192	0.261 0.140
[0.0 0.0, 0.0 0.0]	0.80	0.749 0.577	0.577 0.403
[0.0 0.0, 0.0 0.0]	0.70	0.970 0.922	0.843 0.696
[0.4 0.0, 0.0 0.4]	0.90	0.244 0.133	0.259 0.144
[0.4 0.0, 0.0 0.4]	0.80	0.550 0.372	0.490 0.341
[0.4 0.0, 0.0 0.4]	0.70	0.825 0.667	0.714 0.548
[0.4 0.2, 0.2 0.4]	0.90	0.274 0.166	0.276 0.153
[0.4 0.2, 0.2 0.4]	0.80	0.652 0.469	0.553 0.383
[0.4 0.2, 0.2 0.4]	0.70	0.919 0.815	0.793 0.643
[0.8 0.0, 0.0 0.8]	0.90	0.223 0.133	0.238 0.139
[0.8 0.0, 0.0 0.8]	0.80	0.340 0.218	0.330 0.196
[0.8 0.0, 0.0 0.8]	0.70	0.427 0.285	0.395 0.257
[0.4 0.4, 0.4 0.4]	0.90	0.345 0.222	0.313 0.180
[0.4 0.4, 0.4 0.4]	0.80	0.762 0.614	0.608 0.434
[0.4 0.4, 0.4 0.4]	0.70	0.975 0.923	0.858 0.727

Table 7. Power of Tests z_1^* and JT_1 for $T = 200$, $\rho = 0$ at nominal 0.10 and 0.05 critical values (4000 replications).			
A_1	ϕ	z_1^*	JT_1
[0.0 0.0, 0.0 0.0]	0.95	0.280 0.173	0.248 0.137
[0.0 0.0, 0.0 0.0]	0.90	0.707 0.533	0.595 0.404
[0.0 0.0, 0.0 0.0]	0.80	0.998 0.991	0.975 0.928
[0.4 0.0, 0.0 0.4]	0.95	0.239 0.134	0.250 0.138
[0.4 0.0, 0.0 0.4]	0.90	0.605 0.420	0.533 0.370
[0.4 0.0, 0.0 0.4]	0.80	0.981 0.936	0.941 0.848
[0.4 0.2, 0.2 0.4]	0.95	0.250 0.146	0.262 0.146
[0.4 0.2, 0.2 0.4]	0.90	0.655 0.473	0.576 0.396
[0.4 0.2, 0.2 0.4]	0.80	0.992 0.969	0.967 0.910
[0.8 0.0, 0.0 0.8]	0.95	0.220 0.123	0.244 0.130
[0.8 0.0, 0.0 0.8]	0.90	0.435 0.285	0.404 0.253
[0.8 0.0, 0.0 0.8]	0.80	0.752 0.584	0.657 0.488
[0.4 0.4, 0.4 0.4]	0.95	0.288 0.167	0.282 0.160
[0.4 0.4, 0.4 0.4]	0.90	0.708 0.544	0.608 0.437
[0.4 0.4, 0.4 0.4]	0.80	0.998 0.991	0.976 0.937

Table 8. Values of the tests z_p^* and JT_p applied to the Interest Rate series.			
p	z_p^*	JT_p	b_{\min}
0	-3.38*	13.64	0.831
1	-3.74**	14.32	0.828
2	-3.91**	14.33	0.833
3	-4.10**	14.60	0.846
4	-3.79**	13.43	0.865
5	-3.44*	11.55	0.834

* (**) denotes significant at the 0.10 (0.05) level.