



## Near Unit Roots, Cointegration, and the Term Structure of Interest Rates

Markku Lanne

*Journal of Applied Econometrics*, Vol. 15, No. 5. (Sep. - Oct., 2000), pp. 513-529.

Stable URL:

<http://links.jstor.org/sici?sici=0883-7252%28200009%2F10%2915%3A5%3C513%3ANURCAT%3E2.0.CO%3B2-U>

*Journal of Applied Econometrics* is currently published by John Wiley & Sons.

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/jwiley.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

---

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

# NEAR UNIT ROOTS, COINTEGRATION, AND THE TERM STRUCTURE OF INTEREST RATES

MARKKU LANNE\*

*Research Unit on Economic Structures and Growth, Department of Economics, PO Box 54 (Unioninkatu 37), FIN-00014, University of Helsinki, Finland*

## SUMMARY

The term structure of interest rates is often modelled as a cointegrated system with the yield spreads forming the cointegrating vectors. Testing whether the yield spreads span the cointegration space is problematic because conventional tests on the cointegration vectors tend to overreject when the largest autoregressive roots deviate from unity, as is likely to be the case with interest rates. A new test that is robust w.r.t. deviations from the exact unit root assumption is developed and applied to monthly US interest rate data from 1952:1–1991:2. Taking into account the regime shift in 1979, the hypothesis of the yield spreads being the cointegrating vectors cannot be rejected using the robust test. Copyright © 2000 John Wiley & Sons, Ltd.

## 1. INTRODUCTION

Modelling the yield curve was one of the first applications of cointegration methods, already considered by Engle and Granger (1987) in their seminal paper on cointegration. The approach relies on the assumption that interest rates can be described as  $I(1)$  processes, and typically it is tested whether yields on bonds of different maturities move together in such a way that their spreads are stationary. This latter feature has sometimes been introduced as an implication of the expectations hypothesis of the term structure of interest rates. Early papers concentrating mostly on bivariate models include Campbell and Shiller (1987), Kugler (1990), MacDonald and Speight (1991) and Taylor (1992). Hall *et al.* (1992) and Shea (1992) extended the bivariate cointegration approach to the multivariate case to model the entire yield curve simultaneously. The empirical evidence is mixed. While the cointegration restrictions usually cannot be rejected in bivariate systems, tests in multivariate systems often indicate the presence of several common stochastic trends underlying the yield curve, at least in some time periods (e.g. Shea, 1992; Zhang, 1993; Hall *et al.*, 1992).

All the above results rely on the assumption that interest rates indeed are  $I(1)$  processes. Theoretically this property cannot strictly be justified since nominal interest rates are bounded below by zero whereas  $I(1)$  processes are unbounded. Thus this assumption is only based on the idea that as persistent series interest rates are better approximated by  $I(1)$  than stationary processes, and therefore cointegration methods are applicable. Furthermore, the null hypothesis of a unit root typically cannot be rejected for interest rate series. However, even if the properties of interest rates were indistinguishable from those of exact unit root processes for descriptive purposes, this need

---

\* Correspondence to: Markku Lanne, Department of Economics, PO Box 54 (Unioninkatu 37), FIN-00014, University of Helsinki, Finland; e-mail: markku.lanne@helsinki.fi

Contract/grant sponsor: Yrjö Jahnsson Foundation.

Contract/grant sponsor: Säästöpankkien tutkimussäätiö.

not be the case in testing in models consisting of a number of interest rates. Recent results (Elliott, 1995, 1998) suggest that with strong simultaneity in the system even minor deviations from the exact unit root can mean vast size distortions in tests concerning the parameters of the cointegrating space. In other words, the hypothesis of the spreads spanning the cointegration space can be spuriously rejected because the interest rates in fact are not exact  $I(1)$  processes.

In this paper we develop a new test of restrictions on the parameters of the cointegrating vectors that is robust with respect to deviations from the exact unit root assumption. The test is closely related to the approach of Cavanagh *et al.* (1995) who considered inference in a simple regression model where the order of integration of the regressor is unknown. There are by now a few other contributions in the literature that try to solve the problem of asymptotically valid inference in such cointegrated models (see the survey by Stock, 1997). The relative merits of the different methods are, however, largely unknown, and very few empirical applications exist so far. Moreover, typically only the case of a single cointegrating vector has been considered, so that those methods are not necessarily directly applicable in our model.

The data set consists of the US zero coupon yield curve constructed by McCulloch and Kwon (1993). The emphasis is on comparing the results from cointegration analysis based on the idea of exactly  $I(1)$  interest rates and the new test that allows deviations from this assumption. The empirical results indicate that in many cases the results obtained using these two methods differ, i.e. the assumption of exactly  $I(1)$  interest rates matters.

The plan of the paper is as follows. Section 2 presents a commonly applied basic model of the term structure of interest rates. In Section 3 the econometric methods are discussed, and the new robust test is derived. In Section 4 the results of the empirical analysis as well as the small sample simulation experiments are presented. Section 5 presents conclusions.

## 2. THEORETICAL FRAMEWORK

As a starting point for the analysis of the yield curve, we consider the following model (see e.g. Campbell and Shiller, 1991) that embodies the assumptions of rational expectations and no arbitrage opportunities:

$$r_t^n = \frac{1}{n} \sum_{i=0}^{n-1} E_t r_{t+i}^1 + L \quad (1)$$

where  $r_t^1$  and  $r_t^n$  are the one-period and  $n$ -period interest rates, respectively,  $L$  reflects the term premium, and  $E_t$  denotes mathematical expectation conditional on public information at time  $t$ . Thus according to equation (1) the longer-term interest rate is an average of the present and expected future one-period interest rates over the life of the longer-term bond plus a premium term  $L$ . Under the expectations hypothesis of the term structure of interest rates  $L$  is constant. By rearranging equation (1) we obtain

$$r_t^n - r_t^1 = \frac{1}{n} \sum_{i=1}^{n-1} \sum_{j=1}^i E_t \Delta r_{t+j}^1 + L \quad (2)$$

where  $\Delta r_{t+j}^1 = r_{t+j}^1 - r_{t+j-1}^1$ . Assuming that interest rates are  $I(1)$  processes, their first differences are stationary  $I(0)$  processes and because the premium term  $L$  is constant, it follows that as a sum

of stationary components the right-hand side of equation (2) must be stationary. For the term on the left-hand side to be stationary it is necessary that  $r_t^n$  be cointegrated with  $r_t^1$  such that their difference is stationary. In other words, it is required that the yield spread between the one-period yield and any  $n$ -period yield is  $I(0)$ . With  $p$  interest rates this means that the cointegration space is spanned by the  $p-1$  yield spreads.

The fact that the term structure can be expressed as the above-mentioned cointegrated system can be given at least two intuitive interpretations based on different representations of the system. First, the error-correction representation relates the changes in each yield series to past equilibrium errors and past changes in all yields. In this case the yield spreads are the equilibrium errors that adjust the yields on bonds of different maturity when they diverge, so that in the long run different yields move together. As Campbell and Shiller (1988) have shown, the existence of an error-correction model in this context does not necessarily reflect partial adjustment of one variable to another, but the cointegration relationship can also arise because agents are forecasting and have rational expectations. The yield spreads can be used to make more accurate forecasts of future short rates than would be possible using only the past observations of the short rate series, and the error-correction model results from agents' forward-looking behaviour. Second, the cointegrated system can also be interpreted in terms of common trends. More specifically, the existence of  $p-1$  cointegrating relationships is equivalent to there being one  $I(1)$  common trend underlying the interest rate series.

It is clear that while stationary yield spreads are necessary for the expectations hypothesis to hold, this requirement is by no means sufficient. Indeed, Miron (1991) has pointed out that the cointegration implications mentioned above continue to hold also under some alternative theories. For example, a model where the term premium  $L$  in equation (1) is time varying and stationary has the exactly same cointegration implications. Similarly, any model like equation (1) where the weights deviate from  $1/n$  still implies stationary yield spreads. Therefore, it seems that tests on the cointegration rank and the parameters of the cointegration space, that depend only on the time series properties of interest rates, cannot discriminate between alternative theories of the term structure of interest rates. However, although these tests probably do not have power against some plausible alternatives they can be useful in narrowing down the class of theories that fit the data. For instance, the market segmentation hypothesis does not imply stationary yield spreads. There is also a related literature on continuous-time general equilibrium models of the term structure with different numbers of factors underlying the yield curve. Common stochastic trends have also been interpreted as these factors, and cointegration analysis has been used to test for the number of factors (see e.g. Hiraki *et al.*, 1996). Although seemingly analogous to the expectations hypothesis, one-factor models do not necessarily imply stationary spreads either, as was recently shown by Pagan *et al.* (1996) for the Cox–Ingersoll–Ross (1985) square root model.

### 3. ECONOMETRIC METHODOLOGY

In this paper the term structure is modelled as the following cointegrated system in triangular form:

$$\begin{aligned} r_t^1 &= \rho r_{t-1}^1 + v_{1t} \\ r_t^* &= \mu + \Gamma r_t^1 + v_{2t} \quad t = 1, \dots, T \end{aligned} \quad (3)$$

where  $r_t^*$  is a vector consisting of the  $p-1$  longer-maturity interest rates,  $\mu$  is a constant term,  $\Gamma$  is a  $(p-1) \times 1$  coefficient vector and  $\Phi(L)(v_{1t}, v_{2t})' = \Phi(L)v_t = \varepsilon_t$  with  $\varepsilon_t$  a  $p$ -dimensional vector martingale difference sequence having four finite moments and a constant conditional covariance matrix  $E(\varepsilon_t \varepsilon_t' | \varepsilon_1, \dots, \varepsilon_{t-1}) = \Sigma$ , and  $\Phi(L)$  is a  $k$ th-order lag polynomial with all roots outside the unit circle and  $\Phi_0$  is the identity matrix  $I_p$ , so that the long-run covariance matrix of  $v_t$  is  $\Omega = \Phi(1)^{-1} \Sigma \Phi(1)^{-1'}$ .<sup>1</sup> Model (3) embodies the implication of model (1) that the cointegration rank is  $p-1$ , whereas the particular normalization of the system is arbitrary, and is not implied by model (1). For the following derivations it is useful to rewrite the model as

$$\Delta r_t = m + M r_{t-1} + P v_t \quad (4)$$

where

$$r_t = \begin{pmatrix} r_t^1 \\ r_t^* \end{pmatrix}, \quad m = \begin{pmatrix} 0 \\ \mu \end{pmatrix}, \quad M = \begin{pmatrix} \rho - 1 & 0 \\ \rho \Gamma & -I_{p-1} \end{pmatrix} \text{ and } P = \begin{pmatrix} 1 & 0 \\ \Gamma & I_{p-1} \end{pmatrix}$$

The hypothesis of the cointegration rank being  $p-1$  can be tested using any of the standard cointegration tests suggested in the literature (in Section 4 we employ the tests of Saikkonen and Luukkonen, 1997, and Johansen, 1991). Under the assumption of  $p-1$  cointegrating vectors, the restriction of the spreads spanning the cointegrating space can be tested as usual. The Wald statistic for the hypothesis that  $\Gamma$  is a vector of ones,  $\Gamma_0$ , takes the following form:

$$W_0 = T(\hat{\Gamma} - \Gamma_0)' \hat{V}^{-1} T(\hat{\Gamma} - \Gamma_0) \quad (5)$$

where  $\hat{\Gamma}$  and  $\hat{V}$  are consistent estimators of  $\Gamma$  and its covariance matrix  $V$ , respectively. In Section 4 we shall employ the Saikkonen (1992) estimator of  $\Gamma$ .

When the interest rates cannot be assumed to be exactly  $I(1)$  processes, it is advisable to simultaneously test the hypotheses that there are  $p-1$  cointegrating vectors and that these are the spreads. The simulation results of Elliott (1995) show that the joint tests derived by Horvath and Watson (1995) and Elliott (1995) have superior power properties compared to the separate tests in this case. We shall consider their Wald test based on the following regression model;

$$\Delta r_t = m^* + \Psi_1 r_{t-1}^1 + \Psi_2 u_{t-1} + \Pi(L) \Delta r_{t-1} + \varepsilon_t^* \quad (6)$$

where  $m^* = P\Phi(1)P^{-1}m$ ,  $\varepsilon_t^* = P\Phi(L)v_t = P\varepsilon_t$ , and  $u_t = r_t^* - \Gamma_0 r_t^1$  is a vector of stationary known error correction terms under the null. Implicit in model (6) are the restrictions on the constant term, which (following Horvath and Watson, 1995) are ignored to facilitate estimation by OLS. It is worth noting that allowing for unrestricted intercepts implies linear trends in the interest rates under the null hypothesis (see Pesaran *et al.*, 1998). Although highly implausible, this assumption was also adopted by Engsted and Tanggaard (1994). The reason we ignore these restrictions is mainly computational simplicity; in particular, the robust test to be introduced below would become cumbersome if the restrictions were imposed. Furthermore, the recent results of Horvath and Watson (1995) and Saikkonen and Lütkepohl (1999) also suggest that in vector error-correction models the cointegration rank test based on the unconstrained estimator has somewhat better local power than the test based on the constrained estimator. As far as

<sup>1</sup> The model is a special case of that considered by Elliott (1998).

testing the cointegrating rank is concerned, the results for Johansen's (1991) test imposing these restrictions are also reported in Section 4.

Denoting the vector of regressors by  $x_{t-1} = (r_{t-1}^1, 1, u'_{t-1}, \Delta r'_{t-1}, \dots, \Delta r'_{t-k+1})'$  and the corresponding coefficient matrix by  $\Theta = (\Psi_1 : m^* : \Psi_2 : \Pi)$ , the OLS estimator can be written as  $\hat{\Theta}' = (\sum_{t=k+1}^T x_{t-1} x'_{t-1})^{-1} \sum_{t=k+1}^T x_{t-1} \Delta r'_t$ . Defining a selection matrix  $H$  that picks up the vector  $\Psi_1$  from  $\Theta$ , as  $H = (I_p \otimes H_a)$ ,  $H_a = (1, 0'_{pk})$ , the joint Wald test statistic becomes

$$W_J = [H \text{vec}(\hat{\Theta}')] \left\{ H \left[ \hat{\Sigma}^* \otimes \left( \sum_{t=k+1}^T x_{t-1} x'_{t-1} \right)^{-1} \right] H' \right\}^{-1} [H \text{vec}(\hat{\Theta}')] \quad (7)$$

where

$$\hat{\Sigma}^* = \frac{1}{T-k} \sum_{t=k+1}^T \hat{\varepsilon}_t^* \hat{\varepsilon}_t^{*'}.$$

is a consistent estimator of  $\Sigma^*$ . The joint Wald test has a non-standard asymptotic distribution under the null, and some critical values are tabulated by Horvath and Watson (1995, Table I) and Elliott (1995, Table II).

Specifically, in the case of the term structure we will consider a test for the joint null of there being  $p-1$  cointegrating vectors between the  $p$  interest rates and the yield spreads spanning the cointegration space against the alternative that there are  $p$  cointegrating vectors, i.e. that the interest rate series are stationary. Since all the cointegrating vectors are known under the null, the testable restrictions can be expressed in terms of model (3), and they are  $\rho=1$  and  $\Gamma=\Gamma_0$ , a  $(p-1)$ -vector of ones.

The problem with both the Wald test (5) for the spread restrictions and the joint test (7) is how to interpret the rejection. Since both of these tests have power against deviations from the exact unit root, they may reject even if the spread restrictions are true, i.e. the spreads are stationary. The fact that the Wald test (5) has power against deviations from the exact unit root was recently pointed out by Elliott (1998), who showed that under such deviations, the Saikkonen (1992) estimator (and hence the asymptotically equivalent maximum likelihood estimator of Johansen, 1991) may have a bias that disappears at rate  $T$ . He employed the local-to-unity asymptotic theory, i.e. parameterized the largest autoregressive root,  $\rho$ , as  $\rho=1+c/T$ , where  $c$  is fixed. The bias depends both on the local-to-unity parameter  $c$  and the simultaneity in the system as captured by the long-run covariance matrix between the error terms of the long-term and one-period interest rates,  $\Omega_{21}$ . If either of these equals zero, the bias disappears. It follows from this that the Wald statistic (5) has the standard asymptotic  $\chi^2_{p-1}$  distribution under the null only when either  $c$  or  $\Omega_{21}$  equals zero.<sup>2</sup> Otherwise, the use of standard critical values leads to overrejection.

To be able to control size uniformly over the local-to-unity parameter  $c$ , we suggest robustifying the joint test (7). The general idea is to replace the exact unit root hypothesis,  $\rho=1$ , in the joint test (7) by the local-to-unity parameterization  $\rho=1+c/T$ . Under this parameterization the null distribution of the test statistic becomes dependent on the local-to-

<sup>2</sup> Elliot (1998) derived the non-standard null distribution of this statistic that could, of course, be employed, if the nuisance parameters were consistently estimable. While  $\Omega$  indeed can be consistently estimated, the local-to-unity parameter cannot (see, e.g. Phillips, 1987, Theorem 1).

Table I. 95% confidence intervals for the largest autoregressive root of some interest rates

Interest rate	$k$	ADF $\tau^\mu$	95% interval
1952:1–1991:2			
$r^1$	8	−1.985	0.969, 1.007
$r^2$	12	−1.912	0.971, 1.007
$r^3$	12	−1.897	0.971, 1.008
$r^6$	11	−2.078	0.967, 1.007
$r^{12}$	11	−1.694	0.976, 1.008
$r^{24}$	11	−1.849	0.972, 1.008
$r^{36}$	11	−1.755	0.975, 1.008
$r^{60}$	11	−1.729	0.975, 1.008
$r^{120}$	11	−1.658	0.977, 1.008
1952:1–1979:9			
$r^1$	8	−0.649	0.990, 1.014
$r^2$	12	−1.144	0.981, 1.013
$r^3$	6	−0.686	0.989, 1.014
$r^6$	7	−0.338	0.995, 1.015
$r^{12}$	11	−0.783	0.987, 1.014
$r^{24}$	7	−0.376	0.994, 1.014
$r^{36}$	7	−0.371	0.994, 1.014
$r^{60}$	7	−0.304	0.995, 1.015
$r^{120}$	5	−0.075	0.998, 1.015
1979:10–1991:2			
$r^1$	0	−2.100	0.885, 1.023
$r^2$	12	−1.200	0.950, 1.032
$r^3$	6	−1.172	0.951, 1.032
$r^6$	1	−2.177	0.878, 1.022
$r^{12}$	1	−1.707	0.917, 1.028
$r^{24}$	2	−1.556	0.927, 1.029
$r^{36}$	1	−1.822	0.908, 1.027
$r^{60}$	0	−1.366	0.940, 1.031
$r^{120}$	11	−1.940	0.898, 1.025

Note: The lag length  $k$  in the ADF tests was determined by step-down testing (Ng and Perron, 1995), using a 5% level for each lag length. An intercept is included in the ADF regression. The confidence intervals were obtained by linear interpolation based on the figures in Stock (1991, Table A.1).

unity parameter  $c$  which is not consistently estimable. The effect of this nuisance parameter is eliminated by evaluating the test statistic for a range of values of  $c$  to find all the values of  $\Gamma_0$  for which the null hypothesis cannot be rejected for at least one value of  $c$ , say, at the  $100\eta\%$  significance level. These values of  $\Gamma_0$  then belong to the  $100(1-\eta)\%$  confidence region for the vector  $\Gamma$ . In our model the joint null hypothesis to consider is  $\Gamma=\Gamma_0$  and  $c=c_0$ . Noting that in model (6)

$$\begin{aligned}\Psi_1 &= \Psi\left(\frac{1}{\Gamma_0}\right) = P\Phi(1)P^{-1}M\left(\frac{1}{\Gamma_0}\right) = P\Phi(1)P^{-1}\left(\frac{\rho-1}{\rho\Gamma-\Gamma_0}\right) \\ &= \left(\frac{(\rho-1)\Phi_{11}(1) + \Phi_{12}(1)(\Gamma-\Gamma_0)}{(\rho-1)[\Gamma\Phi_{11}(1) + \Phi_{21}(1)] + (\Gamma\Phi_{12}(1) + \Phi_{22}(1))(\Gamma-\Gamma_0)}\right)\end{aligned}\quad (8)$$

we get under this hypothesis, replacing  $\rho$  by  $1+c_0/T$ ,

Table II. Results of the cointegration tests of the term structure for the period 1952:1–1991:2

Interest rates in the system	Rank test <sup>a</sup>		Spread test <sup>b</sup>	Joint test <sup>c</sup>	Robust test <sup>d</sup>
	$W_v$	$\lambda_{trace}$			
$r^1, r^2$	5.66	4.68	32.10	26.23	<0.01
$r^1, r^3$	5.56	4.58	28.92	23.35	<0.01
$r^1, r^6$	5.58	5.25	15.29	15.29	<0.01
$r^1, r^{12}$	5.67	4.15	5.30	9.53	0.07
$r^1, r^{36}$	3.72	3.81	0.91	6.19	0.20
$r^1, r^{60}$	3.16	4.19	0.70	6.42	0.25
$r^1, r^{120}$	2.52	3.15	0.61	4.38	0.35
$r^1, r^2, r^3$	4.39	4.42	32.14	26.89	<0.01
$r^1, r^2, r^6$	5.53	4.59	31.23	26.38	<0.01
$r^1, r^6, r^{12}$	5.03	4.35	30.26	23.49	<0.01
$r^1, r^{12}, r^{24}$	4.80	3.88	17.59	15.00	<0.01
$r^1, r^{12}, r^{60}$	3.47	3.41	19.83	13.83	0.01
$r^1, r^{24}, r^{120}$	2.43	2.90	17.21	10.02	0.10
$r^1, r^2, r^3, r^6$	4.40	4.45	31.50	27.77	0.01
$r^1, r^3, r^6, r^{12}$	3.81	4.25	30.87	25.98	<0.01
$r^1, r^6, r^{12}, r^{24}$	4.87	3.89	27.03	20.68	<0.01
$r^1, r^{12}, r^{60}, r^{120}$	2.72	2.71	19.00	11.81	0.14
$r^1, r^2, r^3, r^6, r^{12}$	3.70	4.18	32.57	26.16	0.02
$r^1, r^3, r^6, r^{12}, r^{24}$	3.36	3.87	32.09	24.80	0.01
$r^1, r^6, r^{12}, r^{24}, r^{60}$	2.68	3.40	31.92	17.50	0.04
$r^1, r^{12}, r^{24}, r^{60}, r^{120}$	2.70	2.80	19.15	11.31	0.19

Note: The lag length was selected using the Bayes information criterion (Schwarz, 1978). In the rank test the null hypothesis is that the cointegrating rank is  $p-1$  with  $p$  interest rates. The rank test  $W_v$  is due to Saikkonen and Luukkonen (1997), and the  $\lambda_{trace}$  test is the trace test of Johansen (1991). The spread restriction test is the Wald test (5) of the hypothesis that the yield spreads are cointegrating vectors. The joint test is the simultaneous Wald test (7) of both these hypotheses. The robust test is based on statistic (9).

<sup>a</sup>The 5% and 10% critical values are 9.13 and 7.50, respectively.

<sup>b</sup>The 5% and 10% critical values are 3.84 and 2.71 (two-variable system), 5.99 and 4.61 (three-variable system), 7.81 and 6.25 (four-variable system), and 9.49 and 7.78 (five-variable system), respectively.

<sup>c</sup>The 5% and 10% critical values are 9.66 and 8.00 (two-variable system), 11.22 and 9.37 (three-variable system), 12.79 and 10.82 (four-variable system), and 14.23 and 12.17 (five-variable system), respectively.

<sup>d</sup>Maximal observed significance levels.

$$T\Psi_1 = \begin{pmatrix} c_0\Phi_{11}(1) \\ c_0\Gamma_0\Phi_{11}(1) + c_0\Phi_{21}(1) \end{pmatrix}$$

and the test can be based on comparing this to the estimate of  $T\Psi_1$ . Correspondingly the Wald statistic (7) is modified as

$$W = [TH\text{vec}(\hat{\Theta}') - h]' \left\{ T^2 H \left[ \hat{\Sigma}^* \otimes \left( \sum_{t=k+1}^T x_{t-1} x_{t-1}' \right)^{-1} \right] H' \right\}^{-1} [TH\text{vec}(\hat{\Theta}') - h] \quad (9)$$

where  $h = (c_0\hat{\Phi}_{11}(1), (c_0\Gamma_0\hat{\Phi}_{11}(1) + c_0\hat{\Phi}_{21}(1)))'$ , and  $\hat{\Phi}_{11}(1)$  and  $\hat{\Phi}_{21}(1)$  are the consistent estimators of  $\Phi_{11}(1)$  and  $\Phi_{21}(1)$  under the null, respectively. It is shown in the Appendix that



under the null hypothesis

$$W \Rightarrow \left( \int_0^1 J_{c_0}^{\mu}(s) dB_1(s) \right)^2 \left( \int_0^1 J_{c_0}^{\mu^2}(s) \right)^{-1} + \chi_{p-1}^2 \quad (10)$$

where  $\Rightarrow$  denotes weak convergence on  $D[0,1]$ ,  $J_c(s)$  is an Ornstein–Uhlenbeck diffusion process defined by  $dJ_c(s) = cJ_c(s)ds + dB_1(s)$ ,  $J_c(0) = 0$ , and  $J_c^{\mu}(s) = J_c(s) - \int_0^1 J_c(r)dr$  with  $B_1(s)$  a univariate standard Brownian motion, and the two random variables are independently distributed. The asymptotic null distribution of  $W$  can be obtained by simulation. For the case  $p=2$  some critical values are tabulated in Cavanagh *et al.* (1995, Table II, panel A). By inverting the test, the at least  $100(1-\eta)\%$  confidence region for the vector  $\Gamma$  can be computed. In larger than two-dimensional systems it is, however, more convenient and computationally less demanding to find the maximum observed level of significance for the test instead of the explicit confidence region. In other words,  $\Gamma$  is fixed at  $\Gamma_0$  and for each fixed value of  $c_0$ , the value of the test statistic (9) and the corresponding observed level of significance are computed. Finally, the maximum (over  $c_0$ ) of these observed significance levels is reported as the observed level of significance of the test.<sup>3</sup>

For  $p=2$ ,  $W$  has the same asymptotic null distribution as the related statistic of Cavanagh *et al.* (1995) so that their results concerning asymptotic size and power are readily available. According to their computations the procedure is rather conservative but has fair power against local alternatives. Asymptotic size and power depend on  $\Omega$  so that reporting them beyond the bivariate case would require extensive tabulations. However, some exemplary calculations show similar results also for  $p > 2$ . Finite sample simulation results with models estimated from actual interest rate data are presented later in Table V.

#### 4. EMPIRICAL RESULTS

In this section the cointegrating implications of model (1) will be tested using the typical two-step procedure, the joint test and the new robust test. The results shed light on whether the conventional cointegration methods are robust with respect to deviations from the exact unit root assumption, whose violations are to be expected on theoretical grounds as discussed in the Introduction. The data consist of the monthly US zero-coupon yield curves constructed by McCulloch (1990) and updated by McCulloch and Kwon (1993), covering the period 1952:1–1991:2. Unfortunately, such comprehensive yield curve data are not available for more recent years. For the most part this is the same data set that was also examined by Shea (1992), Engsted and Tanggaard (1994), and Johnson (1994). Shea's analysis covers several subsets of the yield curve and various time periods, and the results are more or less inconclusive. The approach of Engsted and Tanggaard (1994), on the other hand, is more systematic in that they consider only their entire sample period (1952:1–1987:2) and the period 1952:1–1979:9 preceding the 'new operating procedures' of the Federal Reserve System. With the exception of the shortest-term yields (at most a year) the cointegration implications of model (1) cannot be rejected in either case. The model of Engsted and Tanggaard (1994) differs from the ones typically used in this literature in that they allowed for deterministic linear trends. Johnson (1994) rejected the restriction that the spreads are cointegrating but argued that the economic significance of the

<sup>3</sup> The empirical results in the next section are based on 160 different values of  $c_0$  ranging from  $-40.0$  to  $9.5$ .

rejection is small, because the  $R^2$  values of the equations of the restricted error-correction model were almost as high as those of the corresponding unrestricted model. Here we follow the approach of Engsted and Tanggaard, and report the results separately for the period 1952:1–1979:9, when the Federal Reserve System had an explicit interest rate target, and the rest of the sample, 1979:10–1991:2, consisting of the period of the new operating procedures and the gradual abandonment of these procedures (see e.g. Meulendyke, 1989). This division is common in empirical studies in monetary economics concerning the US. Because of lack of data on longer maturities for the entire period only at most 10-year interest rates are considered.<sup>4</sup>

In order to give an idea of how well the interest rates are approximated by an  $I(1)$  process the 95% confidence intervals for the largest autoregressive root for some maturities and the two samples are reported in Table I. The null hypothesis of a unit root cannot be rejected in any case at the 5% level. In general, the lower limits of the confidence intervals are rather close to one, and they seem to be the closer the longer the maturity. The difference between the samples is also clear. In the interest rate targeting period the intervals are narrower with lower limits corresponding to values of the local-to-unity parameter  $c$  ranging approximately from  $-7$  to  $-1$  while the corresponding figures for the latter subsample are  $-17$  and  $-7$ , and for the entire sample  $-15$  and  $-11$ , respectively. Thus there really seems to be a structural break in the time series of interest rates. The lower limits closer to unity in the targeting period are well in accordance with the prediction in the literature that interest rates are random walks due to central bank smoothing behaviour (e.g. Mankiw and Miron, 1986; McCallum, 1994). The wider intervals in the latter subsample period are probably to some extent due to the small number of observations. The value of  $c$  corresponding to the lower limit of the confidence interval for the one-month rate is about  $-14.5$  in the entire sample and  $-3.5$  in the former and  $-16$  in the latter subsample which are all prone to cause size distortions in the tests on the cointegrating vectors in system (3).

Summary measures for simultaneity in the systems cannot conveniently be reported beyond the case of two interest rates, as it would require the tabulation of the estimates of the large  $\Omega$  matrices. In bivariate systems the estimated long-run correlation coefficients vary between 0.969736 (in the latter subsample) and 0.999998 (in the targeting period). Although simultaneity seems to be high in all samples, the figures are systematically somewhat lower for the latter period. Hence, the nonrobustness of cointegration methods is indeed likely to be a problem in this data set.

The one-month rate will be used as the base rate in what follows so that the system is actually normalized as model (3), because the one-period rate seems to be the natural rate to include in each of the systems to be studied. Other normalizations were also considered, but the normalization has in general little effect on the results. The normalization in (3) was also tested using the  $\tilde{\lambda}_{\text{tr}}$  statistic due to Luukkonen *et al.* (1999) and it could not be rejected in any of the cases at reasonable significance levels. Moreover, the tests for the cointegrating rank due to Saikkonen and Luukkonen (1997) and Johansen (1991) are invariant to normalization.

The results of testing the cointegration implications of model (1) are provided in Tables II through IV for selected systems of interest rates. The null hypothesis of there being  $p-1$  cointegrating vectors among the  $p$  interest rates cannot be rejected for any of the systems in any sample period at the 5% significance level. In the entire sample (Table II) the cointegration

<sup>4</sup> The period of the abandonment of the new operating procedures was also analyzed separately, but the results were similar to those for the entire 1979:10–1991:2 period, and they are not reported.

implications are in almost all cases rejected at the 5% level both using the two-step procedure and the joint test. In view of the potential deviations from the exact unit root in this sample period it seems reasonable to interpret the rejections in favour of the alternative hypothesis of the joint test, i.e. stationarity of the interest rates rather than non-stationarity of the yield spreads. Still, the robust test also rejects in most cases (as a matter of fact, in exactly the same cases as the joint test does), which indicates that the primary reason for the rejections might not be the deviation from the exact unit root but probably the regime shift that has not been taken into account.

The results for the period of interest rate targeting (Table III) differ from those for the entire sample. With the exception of two cases ( $r^1$ ,  $r^{12}$ ,  $r^{60}$ , and  $r^1$ ,  $r^6$ ,  $r^{12}$ ,  $r^{24}$ ,  $r^{60}$ , where the spread restrictions alone are rejected at the 5% level, while the joint null hypothesis cannot be rejected) the two-step procedure and the joint test lead to the same conclusion. The rejections in the spread test and joint test occur in higher-dimensional systems involving long-term interest rates. The robust test rejects in only three cases at the 5% level. All in all, the general conclusion is that the cointegration implications cannot be rejected in the interest rate targeting period whereas they are rejected in the entire sample.

Also in the second subsample (Table IV) both the spread test and joint test reject in higher-dimensional systems at the long end of the maturity spectrum. At the 5% level the robust test rejects only in two cases. Still, it must be borne in mind that the second subsample is rather short

Table III. Results of the cointegration tests of the term structure for the period 1952:1–1979:9

Interest rates in the system	Rank test		Spread test	Joint test	Robust test
	$W_v$	$\lambda_{trace}$			
$r^1, r^2$	0.17	1.99	2.44	1.97	0.72
$r^1, r^3$	0.20	1.99	2.76	2.12	0.69
$r^1, r^6$	0.48	2.01	3.01	2.44	0.65
$r^1, r^{12}$	0.56	2.07	0.09	0.57	0.99
$r^1, r^{36}$	0.33	2.33	1.75	1.25	0.84
$r^1, r^{60}$	0.15	2.61	1.50	0.83	0.92
$r^1, r^{120}$	0.001	3.24	0.56	0.18	0.98
$r^1, r^2, r^3$	1.42	2.59	1.98	1.77	0.90
$r^1, r^2, r^6$	0.62	2.13	1.99	1.84	0.88
$r^1, r^6, r^{12}$	0.20	2.09	14.68	11.66	0.18
$r^1, r^{12}, r^{24}$	0.18	2.41	19.43	11.44	0.21
$r^1, r^{12}, r^{60}$	0.11	2.42	14.35	8.52	0.32
$r^1, r^{24}, r^{120}$	0.01	3.44	2.20	0.73	0.98
$r^1, r^2, r^3, r^6$	1.29	2.43	3.63	3.13	0.35
$r^1, r^3, r^6, r^{12}$	0.61	1.96	25.13	25.84	0.02
$r^1, r^6, r^{12}, r^{24}$	0.20	2.41	20.90	13.77	0.28
$r^1, r^{12}, r^{60}, r^{120}$	0.01	4.65	26.19	16.42	0.06
$r^1, r^2, r^3, r^6, r^{12}$	1.14	2.39	27.30	27.13	0.03
$r^1, r^3, r^6, r^{12}, r^{24}$	0.47	2.53	30.65	23.82	0.04
$r^1, r^6, r^{12}, r^{24}, r^{60}$	0.03	2.62	18.48	10.82	0.76
$r^1, r^{12}, r^{24}, r^{60}, r^{120}$	0.04	4.45	28.34	16.01	0.40

See notes to Table II.

so that the procedure may lack power. In general, the results seem quite favourable to model (1) once the regime shift and subsequent changes in the time series properties of interest rates are taken into account.

In particular, the results of the non-robust test resemble those of Shea (1992) in that the implications are in general rejected for systems of four or five interest rates, whereas in lower-dimensional systems they cannot be rejected. One explanation for this finding might be that the actual size and power of the tests depend on the dimension of the system in finite samples. To examine this possibility some simulation experiments were run with two-, three-, four-, and five-dimensional systems estimated from the data for period 1952:1–1979:9, with the largest autoregressive root of the one month rate restricted to be  $1+c/333$  (for  $c=0, -3.5$  and  $-5$ , with the case  $c=-3.5$  corresponding to the point estimate of the largest autoregressive root of  $r_t^1$  from the data). Thus the simulated data loosely match the properties of the actual data from the interest rate targeting period. The simulation results in Table V show that the power of the joint test increases with the deviation from the exact unit root of the one-month rate, as would be expected. Interestingly, the actual size indeed seems to increase with the dimension of the system. Thus the common finding in both this and earlier studies could result from this kind of varying finite-sample size of the test. These findings cast doubts on the reliability of the conventional cointegration methods relying on the exact unit root assumption in models for interest rates. The new robust test, on the other hand, seems to control size very well. In no case is the nominal size of 5% exceeded. However, also here the rejection rate somewhat increases with both the

Table IV. Results of the cointegration tests of the term structure for the period 1979:10–1991:2

Interest rates in the system	Rank test		Spread test	Joint test	Robust test
	$W_v$	$\lambda_{trace}$			
$r^1, r^2$	4.71	5.60	1.58	6.96	0.72
$r^1, r^3$	4.38	5.26	1.61	6.05	0.40
$r^1, r^6$	4.61	5.44	1.16	5.33	0.24
$r^1, r^{12}$	4.59	5.70	0.01	3.96	0.61
$r^1, r^{36}$	3.09	4.69	0.02	2.68	0.69
$r^1, r^{60}$	2.22	3.82	0.001	2.22	0.74
$r^1, r^{120}$	1.75	3.02	0.58	2.93	0.63
$r^1, r^2, r^3$	4.51	4.35	5.75	10.96	0.16
$r^1, r^2, r^6$	5.14	6.14	4.86	9.52	0.20
$r^1, r^6, r^{12}$	2.43	4.39	9.73	7.63	0.09
$r^1, r^{12}, r^{24}$	2.29	2.98	3.74	4.29	0.37
$r^1, r^{12}, r^{60}$	1.62	1.96	2.76	3.57	0.39
$r^1, r^{24}, r^{120}$	1.66	1.87	24.91	8.70	0.18
$r^1, r^2, r^3, r^6$	5.07	7.11	2.99	8.99	0.27
$r^1, r^3, r^6, r^{12}$	2.20	2.91	33.15	12.88	0.12
$r^1, r^6, r^{12}, r^{24}$	2.18	2.37	8.72	6.32	0.17
$r^1, r^{12}, r^{60}, r^{120}$	1.82	1.84	156.28	21.95	<0.01
$r^1, r^2, r^3, r^6, r^{12}$	4.84	4.85	56.85	15.82	0.40
$r^1, r^3, r^6, r^{12}, r^{24}$	2.37	2.35	34.06	13.42	0.08
$r^1, r^6, r^{12}, r^{24}, r^{60}$	1.77	2.77	33.11	14.61	0.10
$r^1, r^{12}, r^{24}, r^{60}, r^{120}$	2.03	3.36	175.38	27.76	<0.01

See notes to Table II.

Table V. Rejection rates for the joint and robust tests with nominal size  $\leq 5\%$ 

Dimension of the system	$c$		
	0	-3.5	-5
		Joint test	
2	0.057	0.096	0.132
3	0.081	0.124	0.155
4	0.091	0.143	0.173
5	0.112	0.172	0.198
		Robust test	
2	0.013	0.019	0.022
3	0.021	0.024	0.027
4	0.032	0.029	0.032
5	0.036	0.037	0.043

*Notes:* The entries are based on 5000 Monte Carlo replications of model (3) with parameter values loosely matched to the actual interest rate data from the period 1952:1–1979:9. Lag length was selected by BIC with a maximum of six lags. The interest rates used to estimate the systems are  $r^1$  and  $r^{24}$ ;  $r^1$ ,  $r^{24}$  and  $r^{36}$ ;  $r^1$ ,  $r^{24}$ ,  $r^{36}$  and  $r^{60}$ ; and  $r^1$ ,  $r^{24}$ ,  $r^{36}$ ,  $r^{60}$  and  $r^{120}$ , respectively. The simulated model is a VAR for the quasi difference of the short term rate,  $r_t^1 - (1 + c/333)r_{t-1}^1$ , and the spreads between the one-month rate and the other rates. The error terms are generated from a multivariate normal distribution with the estimated covariance matrix.

dimension of the system and the deviation from the exact unit root. In interpreting Table IV one must keep in mind, though, that each row corresponds to a model estimated from a different set of interest rates, and thus the dimension is not necessarily the only reason for the differences between the rejection rates.<sup>5</sup>

## 5. CONCLUSIONS

In this paper we have evaluated the usefulness of conventional cointegration methods in modelling the term structure of interest rates. This approach is based on the assumption that interest rates are well described as unit root processes although this assumption is not theoretically justified. Using monthly US term structure data from the period 1952:1–1991:2 it is demonstrated that conventional tests on the cointegrating vectors are not, in general, robust w.r.t. deviations from the exact unit root, and size distortions prevail. A new testing procedure is developed that is robust in this respect.

In the entire sample period the cointegration implications of model (1) are rejected using any of the tests considered. However, there seems to be a structural break in the time series of the interest rates such that in the targeting period they are more persistent, which may explain the rejections. Taking this break into account by dividing the period into subsamples, the implications cannot be rejected using the robust test. Simulation experiments also indicate that the joint test tends to overreject in larger systems even when there is an exact unit root. This may to some extent explain the common finding in previous empirical studies that the cointegration implications are rejected especially in higher-dimensional systems.

<sup>5</sup> The power of the joint test was also examined against several alternatives using the same kind of set-up as in Table V, and it was found to be excellent. Because of the multitude of possible alternatives, the results are not reported. They are available upon request.

Although the non-rejection of the cointegration implications of model (1) is often interpreted in favour of the expectations hypothesis, this approach is not a proper way to test that theory, because these implications also follow from several plausible alternative theories. However, there are many theories (e.g. the market segmentation hypothesis or continuous-time one-factor models) that do not share the implication of stationary yield spreads, and so the tests are useful in narrowing down the class of theories that fit the data. Furthermore, improved forecasts can be obtained upon pretesting and building a model incorporating the subsequent restrictions, as the theoretical and simulation results in Stock (1996) suggest; for empirical comparisons between forecasts see Hall, Anderson, and Granger (1992) and Bradley and Lumpkin (1992).

#### APPENDIX: DERIVATION OF RESULT (10)

We shall first show the consistency of the OLS estimator of the parameters of model (6) using the approach of Sims, Stock and Watson (1990) applied to the local-to-unity case and then derive the asymptotic null distribution of test static (9). To simplify notation, write the error-correction model (6) in its stacked single equation form:

$$\text{vec}(\Delta r) = (I_p \otimes x)\theta + (P \otimes I_{T-1})\text{vec}(\varepsilon) \quad (\text{A1})$$

where  $\Delta r = (\Delta r_2 \ \Delta r_3 \ \dots \ \Delta r_T)'$ ,  $x = (x_1 \ x_2 \ \dots \ x_{T-1})'$  with  $x_{t-1} = (r_{t-1}^1, 1, u'_{t-1}, \Delta r'_{t-1}, \dots, \Delta r'_{t-k+1})'$ ,  $\theta = \text{vec}(\Theta')$ , and  $\varepsilon = (\varepsilon_2 \ \varepsilon_3 \ \dots \ \varepsilon_T)'$ . Next, write model (A1) in the 'canonical form' of Sims *et al.* (1990). Without loss of generality set  $\mu = 0$  in model (3) (a constant is included in the error-correction model to be estimated). Since from model (3)  $v_t = \Phi(L)^{-1}\varepsilon_t$ ,  $\tilde{\Delta}r_t^1 = r_t^1 - \rho r_{t-1}^1$  and  $u_t = r_t^* - \Gamma r_t^1$  are stationary and have mean zero. The canonical form is thus

$$\text{vec}(\Delta r) = (I_p \otimes Z)\delta + (P \otimes I_{T-1})\text{vec}(\varepsilon) \quad (\text{A2})$$

where  $Z = (Z_1 \ Z_2 \ \dots \ Z_{T-1})'$  and the vector of canonical regressors is  $Z_{t-1} = (Z_{t-1}^1, Z_{t-1}^2, Z_{t-1}^3)'$  with  $Z_{t-1}^1 = (\Delta r_{t-1}^1, u'_{t-1}, \Delta r_{t-2}^1, u'_{t-2}, \dots, \tilde{\Delta}r_{t-k+1}^1, u'_{t-k+1})'$ ,  $Z_{t-1}^2 = 1$  and  $Z_{t-1}^3 = r_{t-1}^1$ . The transformation from the original regressors  $x_{t-1}$  to the canonical regressors is given by

$$Z_{t-1} = Dx_{t-1} = \begin{pmatrix} 1-\rho & 0 & 0 & \rho & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & I_{p-1} & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 1-\rho & 0 & 0 & \rho-1 & 0 & \rho & 0 & \dots & 0 & 0 \\ 0 & 0 & I_{p-1} & \Gamma & -I_{p-1} & 0 & 0 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1-\rho & 0 & 0 & \rho-1 & 0 & \rho-1 & 0 & \dots & \rho & 0 \\ 0 & 0 & I_{p-1} & \Gamma & -I_{p-1} & \Gamma & -I_{p-1} & \dots & \Gamma & -I_{p-1} \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix} \begin{pmatrix} r_{t-1}^1 \\ 1 \\ u_{t-1} \\ \Delta r_{t-1}^1 \\ \Delta r_{t-1}^* \\ \Delta r_{t-2}^1 \\ \Delta r_{t-2}^* \\ \cdot \\ \cdot \\ \cdot \\ \Delta r_{t-k+1}^1 \\ \Delta r_{t-k+1}^* \end{pmatrix}$$

Our assumptions concerning  $\Phi(L)$  and  $\varepsilon$  guarantee that Condition 1 in Sims *et al.* (1990) is satisfied, so that their Theorem 1 combined with Lemma 1 in Phillips (1987) yields the following

convergence result for the OLS estimator of  $\delta$ :  $(I_p \otimes \Upsilon_T)(\hat{\delta} - \delta) \Rightarrow \delta^*$ , where  $\Upsilon_T = \text{diag}(T^{1/2}I_{p(k+1)} T)$  is a scaling matrix and

$$\delta^* = \begin{pmatrix} I_p \otimes EZ_t'Z_t' & 0 & 0 \\ 0 & I_p & I_p \otimes \Omega_{11}^{1/2} \int_0^1 J_c(s)ds \\ 0 & I_p \otimes \Omega_{11}^{1/2} \int_0^1 J_c(s)ds & I_p \otimes \Omega_{11} \int_0^1 J_c^2(s)ds \end{pmatrix}^{-1} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{pmatrix}$$

with

$$\begin{aligned} \phi_1 &\sim N(0, P\Sigma P' \otimes EZ_t'Z_t') \\ \phi_2 &= \text{vec}(B(1)' \Sigma^{1/2'} P') \\ \phi_3 &= \text{vec}(\Omega_{11}^{1/2} \int_0^1 J_c(s)dB(s)' \Sigma^{1/2'} P') \end{aligned}$$

and  $B(s)$  is a  $p$ -dimensional standard Brownian motion, and  $\Omega_{11}$  is the long-run variance of  $v_{1t}$  in model (3). Since the OLS estimator  $\hat{\theta} = (I_p \otimes D')\hat{\delta}$ , it follows that  $\hat{\theta} - \theta \xrightarrow{P} 0$ .

For the coefficients of  $r_{t-1}^1$  in model (A1) the above result implies that under the null hypothesis  $\Gamma = \Gamma_0$  and  $c = c_0$ :

$$\begin{aligned} T\hat{\Psi}_1 - \left( c_0\Gamma_0\Phi_{11}(1) + c_0\Phi_{21}(1) \right) &= \frac{1}{T} \sum_{t=k+1}^T \varepsilon_t^* r_{t-1}^{1\mu} \left( \frac{1}{T^2} \sum_{t=k+1}^T r_{t-1}^{1\mu 2} \right)^{-1} + o_p(1) \\ &\Rightarrow \Omega_{11}^{-1/2} P_0 \Sigma^{1/2} \int_0^1 dB(s) J_{c_0}^\mu(s) \left( \int_0^1 J_{c_0}^{\mu 2}(s) ds \right)^{-1} \end{aligned} \quad (\text{A3})$$

where  $\mu$  refers to OLS demeaning.

Noting that the test static (9) can be written as

$$\begin{aligned} W &= \text{vec}(T\hat{\Psi}_1 - h)' \left( \hat{\Sigma}^* \otimes H_a \left( T^{-2} \sum_{t=k+1}^T x_{t-1} x_{t-1}' \right)^{-1} H_a' \right)^{-1} \text{vec}(T\hat{\Psi}_1 - h) \\ &= \text{tr} \left[ \left( H_a \left( T^{-2} \sum_{t=k+1}^T x_{t-1} x_{t-1}' \right)^{-1} H_a' \right)^{-1} (T\hat{\Psi}_1 - h)' \hat{\Sigma}^{*-1} (T\hat{\Psi}_1 - h) \right] \end{aligned}$$

and that result (A3) continues to hold, when  $\Phi_{11}(1)$  and  $\Phi_{21}(1)$  are replaced by their consistent estimators, applying again Lemma 1 in Phillips (1987), consistency of  $\hat{\Sigma}^*$  and the continuous mapping theorem, we get the asymptotic null distribution of  $W$  as follows (the arguments of functionals and integration limits are omitted for brevity):

$$\begin{aligned}
W &\Rightarrow \text{tr} \left[ \Omega_{11} \int J_{c_0}^{\mu 2} \Omega_{11}^{-1/2} \left( \int J_{c_0}^{\mu 2} \right)^{-1} \int J_{c_0}^{\mu} dB' \Sigma^{1/2} P_0' P_0'^{-1} \Sigma^{-1} P_0^{-1} \Omega_{11}^{-1/2} P_0 \Sigma^{1/2} \int dB J_{c_0}^{\mu} \left( \int J_{c_0}^{\mu 2} \right)^{-1} \right] \\
&= \text{tr} \left[ \int J_{c_0}^{\mu} dB' \int dB J_{c_0}^{\mu} \left( \int J_{c_0}^{\mu 2} \right)^{-1} \right] \\
&= \text{tr} \left[ \int J_{c_0}^{\mu} (dB_1 \dots dB_p) \begin{pmatrix} \int dB_1 \\ \vdots \\ \int dB_p \end{pmatrix} J_{c_0}^{\mu} \left( \int J_{c_0}^{\mu 2} \right)^{-1} \right] \\
&= \text{tr} \left[ \left( \int J_{c_0}^{\mu} dB_1 \right)^2 \left( \int J_{c_0}^{\mu 2} \right)^{-1} + \sum_{j=2}^p \left( \int J_{c_0}^{\mu} dB_j \right)^2 \left( \int J_{c_0}^{\mu 2} \right)^{-1} \right] \\
&= \left( \int J_{c_0}^{\mu} dB_1 \right)^2 \left( \int J_{c_0}^{\mu 2} \right)^{-1} + \xi
\end{aligned} \tag{A4}$$

where  $\xi$  is a random variable distributed as  $\chi_{p-1}^2$ , and independent of  $B_1$  and  $J_{c_0}$ . This can be seen by noting that the conditional distribution of  $\int J_{c_0}^{\mu} dB_j$  ( $j = 2, \dots, p$ ) conditional on a realization of  $J_{c_0}^{\mu}$  is  $N(0, \int J_{c_0}^{\mu 2})$  so that the conditional distribution of  $\int J_{c_0}^{\mu} dB_j (\int J_{c_0}^{\mu 2})^{-1/2}$  conditional on a realization of  $J_{c_0}^{\mu}$  is standard normal, and thus being independent of the realizations of  $J_{c_0}^{\mu}$ , is also the unconditional distribution (cf. Park and Phillips, 1988, Lemma 5.1). The consistency of  $\hat{\Sigma}^*$  follows by the law of large numbers. The consistency of the OLS estimator  $\hat{\Theta}$  implies that  $\Phi_{11}(1)$  and  $\Phi_{21}(1)$  are consistently estimated by the respective blocks of  $I_p - P_0^{-1} \hat{\Pi}(1) P_0$  under the null hypothesis.

#### ACKNOWLEDGEMENTS

I would like to thank Graham Elliott, Heikki Kauppi, Helmut Lütkepohl, Anne Mikkola, Pentti Saikkonen, Timo Teräsvirta, the participants of the 1998 European Winter Meeting of the Econometric Society, and especially the three anonymous referees and M. Hashem Pesaran (the editor) for useful comments. The usual disclaimer applies. This paper is a part of the research program of the Research Unit on Economic Structures and Growth (RUESG) in the Department of Economics of the University of Helsinki. Financial support from the Yrjö Jahnsson Foundation and Säästöpankkien tutkimussäätiö is gratefully acknowledged.

#### REFERENCES

Bradley MG, Lumpkin SA. 1992. The Treasury yield curve as a cointegrated system. *Journal of Financial and Quantitative Analysis* **27**: 449–463.



- Campbell JY, Shiller RJ. 1987. Cointegration and tests of present value models. *Journal of Political Economy* **95**: 1062–1088.
- Campbell JY, Shiller RJ. 1988. Interpreting cointegrated models. *Journal of Economic Dynamics and Control* **12**: 505–522.
- Campbell JY, Shiller RJ. 1991. Yield spreads and interest rate movements: A bird's eye view. *Review of Economic Studies* **58**: 495–514.
- Cavanagh CL, Elliott G, Stock JH. 1995. Inference models with nearly integrated regressors. *Econometric Theory* **11**: 1131–1147.
- Cox JC, Ingersoll JE, Ross SA. 1985. A theory of the term structure of interest rates. *Econometrica* **53**: 385–408.
- Elliott G. 1995. Tests for the correct specification of cointegrating vectors and the error correction model. Discussion paper 95-42, University of California, San Diego.
- Elliott G. 1998. On the robustness of cointegration methods when regressors almost have unit roots. *Econometrica* **66**: 149–158.
- Engle RF, Granger CWJ. 1987. Co-integration, and error correction: representation estimation and testing. *Econometrica* **55**: 251–276.
- Engsted T, Tanggaard C. 1994. Cointegration and the US term structure. *Journal of Banking and Finance* **18**: 167–181.
- Hall AD, Anderson HM, Granger CWJ. 1992. A cointegration analysis of Treasury bill yields. *Review of Economics and Statistics* **74**: 116–126.
- Hiraki T, Shiraishi N, Takezawa N. 1996. Cointegration, common factors, and the term structure of Yen offshore interest rates. *Journal of Fixed Income* **6**: 69–75.
- Horvath MTK, Watson MW. 1995. Testing for cointegration when some of the cointegrating vectors are prespecified. *Econometric Theory* **11**: 984–1014.
- Johansen S. 1991. Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica* **59**: 1551–1580.
- Johnson PA. 1994. On the number of common unit roots in the term structure of interest rates. *Applied Economics* **26**: 815–820.
- Kugler P. 1990. The term structure of Euro interest rates and rational expectations. *Journal of International Money and Finance* **9**: 234–244.
- Luukkonen R, Ripatti A, Saikkonen P. 1999. Testing for a valid normalization of cointegrating vectors in vector autoregressive processes. *Journal of Business & Economic Statistics* **17**: 195–204.
- McCallum BT. 1994. Monetary policy and the term structure of interest rates. NBER Working Paper no. 4938.
- MacDonald R, Speight AEH. 1991. The term structure of interest rates under rational expectations: Some international evidence. *Applied Financial Economics* **1**: 211–221.
- McCulloch JH. 1990. US government term structure data. In *The Handbook of Monetary Economics*, Friedman B, Hahn F (eds). North-Holland: Amsterdam.
- McCulloch JH, Kwon H-C. 1993. US term structure data, 1947–1991, Working paper 93-6, Ohio State University.
- Mankiw NG, Miron JA. 1986. The changing behavior of the term structure of interest rates. *Quarterly Journal of Economics* **101**: 211–228.
- Miron JA. 1991. Comment. In *NBER Macroeconomics Annual 1991*, Blanchard O, Fischer S (eds). MIT Press: Cambridge, MA.
- Meulendyke A-M. 1989. *US Monetary Policy and Financial Markets*. Federal Reserve Bank of New York: New York.
- Ng S, 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.
- Pagan AR, Hall AD, Martin V. 1996. Modeling the term structure. In *Handbook of Statistics 14: Statistical Methods in Finance*, Maddala GS, Rao CR (eds). Elsevier: Amsterdam.
- Park JY, Phillips PCB. 1988. Statistical inference in regressions with integrated processes: Part 1. *Econometric Theory* **4**: 468–497.
- Pesaran MH, Shin Y, Smith RJ. 1998. Structural analysis of vector error correction models with exogenous I(1) variables. *Journal of Econometrics* (forthcoming)
- Phillips PCB. 1987. Towards a unified asymptotic theory for autoregression. *Biometrika* **74**: 535–547.

- Saikkonen P. 1992. Estimation and testing of cointegrated systems by an autoregressive approximation. *Econometric Theory* **8**: 1–27.
- Saikkonen P, Luukkonen R. 1997. Testing cointegration in infinite order vector autoregressive processes. *Journal of Econometrics* **81**: 93–126.
- Saikkonen P, Lütkepohl H. 1999. Local power of likelihood ratio tests for the cointegrating rank of a VAR process. *Econometric Theory* **15**: 50–78.
- Schwarz G. 1978. Estimating the dimension of a model. *Annals of Statistics* **4**: 461–464.
- Shea GS. 1992. Benchmarking the expectations hypothesis of the interest-rate term structure: An analysis of cointegration vectors. *Journal of Business & Economic Statistics* **10**: 347–366.
- Sims CA, Stock JH, Watson MW. 1990. Inference in linear time series models with some unit roots. *Econometrica* **58**: 113–144.
- Stock JH. 1991. Confidence intervals for the largest autoregressive root in US macroeconomic time series. *Journal of Monetary Economics* **28**: 435–459.
- Stock JH. 1996. VAR, error correction and pretest forecasts at long horizons. *Oxford Bulletin of Economics and Statistics* **58**: 685–701.
- Stock JH. 1997. Cointegration, long-run comovements, and long-horizon forecasting. In *Advances in Economics and Econometrics: Theory and Applications, Seventh World Congress, Volume III*, Kreps DM, Wallis KF (eds). Cambridge University Press: Cambridge.
- Taylor MP. 1992. Modelling the yield curve. *Economic Journal* **102**: 524–537.
- Zhang H. 1993. Treasury yield curves and cointegration. *Applied Economics* **25**: 361–367.

## LINKED CITATIONS

- Page 1 of 4 -



You have printed the following article:

### **Near Unit Roots, Cointegration, and the Term Structure of Interest Rates**

Markku Lanne

*Journal of Applied Econometrics*, Vol. 15, No. 5. (Sep. - Oct., 2000), pp. 513-529.

Stable URL:

<http://links.jstor.org/sici?sici=0883-7252%28200009%2F10%2915%3A5%3C513%3ANURCAT%3E2.0.CO%3B2-U>

---

*This article references the following linked citations. If you are trying to access articles from an off-campus location, you may be required to first logon via your library web site to access JSTOR. Please visit your library's website or contact a librarian to learn about options for remote access to JSTOR.*

### **[Footnotes]**

#### <sup>2</sup> **Towards a Unified Asymptotic Theory for Autoregression**

P. C. B. Phillips

*Biometrika*, Vol. 74, No. 3. (Sep., 1987), pp. 535-547.

Stable URL:

<http://links.jstor.org/sici?sici=0006-3444%28198709%2974%3A3%3C535%3ATAUATF%3E2.0.CO%3B2-S>

### **References**

#### **The Treasury Yield Curve as a Cointegrated System**

Michael G. Bradley; Stephen A. Lumpkin

*The Journal of Financial and Quantitative Analysis*, Vol. 27, No. 3. (Sep., 1992), pp. 449-463.

Stable URL:

<http://links.jstor.org/sici?sici=0022-1090%28199209%2927%3A3%3C449%3ATTYCAA%3E2.0.CO%3B2-O>

#### **Cointegration and Tests of Present Value Models**

John Y. Campbell; Robert J. Shiller

*The Journal of Political Economy*, Vol. 95, No. 5. (Oct., 1987), pp. 1062-1088.

Stable URL:

<http://links.jstor.org/sici?sici=0022-3808%28198710%2995%3A5%3C1062%3ACATOPV%3E2.0.CO%3B2-6>

**NOTE:** *The reference numbering from the original has been maintained in this citation list.*

## LINKED CITATIONS

- Page 2 of 4 -



### **Yield Spreads and Interest Rate Movements: A Bird's Eye View**

John Y. Campbell; Robert J. Shiller

*The Review of Economic Studies*, Vol. 58, No. 3, Special Issue: The Econometrics of Financial Markets. (May, 1991), pp. 495-514.

Stable URL:

<http://links.jstor.org/sici?sici=0034-6527%28199105%2958%3A3%3C495%3AYSARM%3E2.0.CO%3B2-U>

### **A Theory of the Term Structure of Interest Rates**

John C. Cox; Jonathan E. Ingersoll, Jr.; Stephen A. Ross

*Econometrica*, Vol. 53, No. 2. (Mar., 1985), pp. 385-407.

Stable URL:

<http://links.jstor.org/sici?sici=0012-9682%28198503%2953%3A2%3C385%3AATOTTS%3E2.0.CO%3B2-B>

### **Co-Integration and Error Correction: Representation, Estimation, and Testing**

Robert F. Engle; C. W. J. Granger

*Econometrica*, Vol. 55, No. 2. (Mar., 1987), pp. 251-276.

Stable URL:

<http://links.jstor.org/sici?sici=0012-9682%28198703%2955%3A2%3C251%3ACAECRE%3E2.0.CO%3B2-T>

### **A Cointegration Analysis of Treasury Bill Yields**

Anthony D. Hall; Heather M. Anderson; Clive W. J. Granger

*The Review of Economics and Statistics*, Vol. 74, No. 1. (Feb., 1992), pp. 116-126.

Stable URL:

<http://links.jstor.org/sici?sici=0034-6535%28199202%2974%3A1%3C116%3AACAOTB%3E2.0.CO%3B2-Z>

### **Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models**

Søren Johansen

*Econometrica*, Vol. 59, No. 6. (Nov., 1991), pp. 1551-1580.

Stable URL:

<http://links.jstor.org/sici?sici=0012-9682%28199111%2959%3A6%3C1551%3AEAHTOC%3E2.0.CO%3B2-P>

### **Testing for a Valid Normalization of Cointegrating Vectors in Vector Autoregressive Processes**

Ritva Luukkonen; Antti Ripatti; Pentti Saikkonen

*Journal of Business & Economic Statistics*, Vol. 17, No. 2. (Apr., 1999), pp. 195-204.

Stable URL:

<http://links.jstor.org/sici?sici=0735-0015%28199904%2917%3A2%3C195%3ATFAVNO%3E2.0.CO%3B2-X>

**NOTE:** The reference numbering from the original has been maintained in this citation list.

## LINKED CITATIONS

- Page 3 of 4 -



### **The Changing Behavior of the Term Structure of Interest Rates**

N. Gregory Mankiw; Jeffrey A. Miron

*The Quarterly Journal of Economics*, Vol. 101, No. 2. (May, 1986), pp. 211-228.

Stable URL:

<http://links.jstor.org/sici?sici=0033-5533%28198605%29101%3A2%3C211%3ATCBOTT%3E2.0.CO%3B2-%23>

### **Unit Root Tests in ARMA Models with Data-Dependent Methods for the Selection of the Truncation Lag**

Serena Ng; Pierre Perron

*Journal of the American Statistical Association*, Vol. 90, No. 429. (Mar., 1995), pp. 268-281.

Stable URL:

<http://links.jstor.org/sici?sici=0162-1459%28199503%2990%3A429%3C268%3AURTAM%3E2.0.CO%3B2-M>

### **Towards a Unified Asymptotic Theory for Autoregression**

P. C. B. Phillips

*Biometrika*, Vol. 74, No. 3. (Sep., 1987), pp. 535-547.

Stable URL:

<http://links.jstor.org/sici?sici=0006-3444%28198709%2974%3A3%3C535%3ATAUATF%3E2.0.CO%3B2-S>

### **Estimating the Dimension of a Model**

Gideon Schwarz

*The Annals of Statistics*, Vol. 6, No. 2. (Mar., 1978), pp. 461-464.

Stable URL:

<http://links.jstor.org/sici?sici=0090-5364%28197803%296%3A2%3C461%3AETDOAM%3E2.0.CO%3B2-5>

### **Benchmarking the Expectations Hypothesis of the Interest-Rate Term Structure: An Analysis of Cointegration Vectors**

Gary S. Shea

*Journal of Business & Economic Statistics*, Vol. 10, No. 3. (Jul., 1992), pp. 347-366.

Stable URL:

<http://links.jstor.org/sici?sici=0735-0015%28199207%2910%3A3%3C347%3ABTEHOT%3E2.0.CO%3B2-N>

### **Inference in Linear Time Series Models with some Unit Roots**

Christopher A. Sims; James H. Stock; Mark W. Watson

*Econometrica*, Vol. 58, No. 1. (Jan., 1990), pp. 113-144.

Stable URL:

<http://links.jstor.org/sici?sici=0012-9682%28199001%2958%3A1%3C113%3AIIILTSM%3E2.0.CO%3B2-P>

**NOTE:** The reference numbering from the original has been maintained in this citation list.

## LINKED CITATIONS

- Page 4 of 4 -



### **Modelling the Yield Curve**

Mark P. Taylor

*The Economic Journal*, Vol. 102, No. 412. (May, 1992), pp. 524-537.

Stable URL:

<http://links.jstor.org/sici?sici=0013-0133%28199205%29102%3A412%3C524%3AMTYC%3E2.0.CO%3B2-H>