CIRJE-F- 1216

Johansen Test with Fourier-Type Smooth Nonlinear Trends in Cointegrating Relations

Takamitsu Kurita
Kyoto Sangyo University
Mototsugu Shintani
The University of Tokyo

June 2023

Johansen test with Fourier-type smooth nonlinear trends in cointegrating relations

Takamitsu Kurita*
Faculty of Economics, Kyoto Sangyo University

Mototsugu Shintani†
Faculty of Economics, The University of Tokyo

June 2023

Abstract

We develop methodology for testing cointegrating rank in vector autoregressive (VAR) models in the presence of Fourier-type smooth nonlinear deterministic trends in cointegrating relations. The limiting distribution of log-likelihood ratio test statistics is derived and approximated limit quantiles are tabulated. A sequential procedure to select cointegrating rank is evaluated by Monte Carlo simulations. Our empirical application to economic data also demonstrates the usefulness of the proposed methodology in a practical context.

Keywords: Vector autoregressive systems; Smooth nonlinear deterministic trends; Trigonometric functions; Cointegrating rank; Log-likelihood ratio test statistics; Monte Carlo experiments.
JEL classifications: C12; C32; C50.

*Faculty of Economics, Kyoto Sangyo University, Motoyama, Kamigamo, Kita-ku, Kyoto 603-8555 E-mail: tkurita@cc.kyoto-su.ac.jp

†Faculty of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. Email: shintani@e.u-tokyo.ac.jp

1 Introduction

A likelihood ratio testing approach proposed by Johansen (1988, 1996) has been widely used in selecting the cointegrating rank in vector autoregressive (VAR) systems. In the recent literature of a unit root test in univariate time series, unknown smooth nonlinear deterministic trends are often approximated with a trigonometric function used as the base of a Fourier expansion. Examples include Pascalau (2010), Harvey, Leybourne and Xiao (2010), Enders and Holt (2012), Enders and Lee (2012a,b), Rodrigues and Taylor (2012), Astill, Harvey, Leybourne and Taylor (2015) and Perron, Shintani and Yabu (2017, 2021), among others. While a unit root testing with such a nonlinear trend modeling strategy has become very popular in practice, to the best of our knowledge, there are no systematic analyses of incorporating trigonometric trends into cointegrating relations in the VAR framework.[1] We aim to fill the gap between the two strands of literature.

The objective of this paper is to develop a methodology for testing the cointegrating rank in VAR models allowing for trigonometric trends included in cointegrating relations. We derive the limiting distributions of the proposed test statistics and provide statistical tables of their approximate limit quantiles, which have not been available. The tables include empirically relevant cases with various choices of frequency of the trigonometric trend function. The usefulness of our method has been confirmed through a simulation experiment as well as an empirical application to economic data.

To illustrate our modeling approach, let us consider the simple example of a term structure of interest rates. Suppose that both a long-term interest rate, x_{1t}, and a short-term interest rate, x_{2t}, are nonstationary. The theory of term structure implies that their linear combination $x_{1t} - \gamma - \beta x_{2t}$ becomes stationary so that two series are cointegrated. Such a theoretical restriction can be described by a bivariate error

[1] A recent paper by Pascalau, Lee, Nazlioglu, and Lu (2022) also consider the presence of trigonometric trends in a cointegrated VAR system. However, unlike our analysis, they do not allow trigonometric trends in cointegrating relations.

correction system with its cointegrating rank set at one, given by

$$\begin{aligned}\Delta x_{1t} &= \alpha_1 EC_{t-1} + \sum_{i=1}^{k-1} \gamma_{11i} \Delta x_{1t-i} + \sum_{i=1}^{k-1} \gamma_{12i} \Delta x_{2t-i} + \varepsilon_{1t}, \\ \Delta x_{2t} &= \alpha_2 EC_{t-1} + \sum_{i=1}^{k-1} \gamma_{21i} \Delta x_{1t-i} + \sum_{i=1}^{k-1} \gamma_{22i} \Delta x_{2t-i} + \varepsilon_{2t}\end{aligned}$$

where

$$EC_{t-1} = x_{1t-1} - \gamma - \beta x_{2t-1}$$

is an error correction term; k is the lag length of the VAR model in term of level variables; and ε_{1t} and ε_{2t} are zero-mean disturbance terms. By generalizing the model to include eleven yields on a US Treasury bill with different maturities, Hall, Anderson, and Granger (1992, Table 1) employed a Johansen's cointegrating rank test and found ten cointegrating relations consistent with the theory. Their analysis of term structure relies on the assumption that the equilibrium term premium, γ, is constant so that term spreads are stationary. However, if the equilibrium term premium is time-varying and nonstationary, a long-run relationship may not be appropriately detected, even if the term structure of interest rate holds. For example, if there is an abrupt shift in the equilibrium term premium, γ, the cointegrating rank test may be modified to allow for a break in the level of cointegrating relations along the line of Johansen, Mosconi and Nielsen (2000) and others.[2] In this paper, we take a more flexible approach and utilize trigonometric trends to describe time-varying equilibrium term premium with a new error correction term given by

$$EC_{t-1} = x_{1t-1} - \gamma - \beta x_{2t-1} - \sum_{j=1}^{n} \left\{ \delta_{1j} \sin\left(\frac{2\pi jt}{T}\right) + \delta_{2j} \cos\left(\frac{2\pi jt}{T}\right) \right\},$$

where n denotes the total number of frequencies and $\{\delta_{1j}, \delta_{2j}\}_{j=1}^{n}$ is a set of trend coefficients. To approximate the underlying nonlinear trends in cointegrating relations, we employ a Fourier expansion with an arbitrary number of frequencies, as in Gallant

[2] Other studies that allow for discrete changes and shifts in a linear trend and constant deterministic non-linearity in cointegrated VAR systems include Inoue (1999), Hendry (2000, 2006), Lütkepohl, Saikkonen and Trenkler (2004), Hungnes (2010), Kurita, Nielsen and Rahbek (2011), Guo and Shintani (2013) and Kurita and Nielsen (2019).

(1981) and Gallant and Souza (1991), among others. This approach allows for a very general class of nonlinear functions, including a case of the smooth transition of the regimes sometimes referred to as smooth breaks rather than abrupt breaks. We use, as an empirical example of our procedure, monthly time series data for Japanese government bond (JGB) yields of five maturities, along with the overnight call rate. Under the assumption of the constant equilibrium term premium, the original test of Johansen suggests that cointegrating rank is one. In contrast, our procedure, which allows for a flexible time-varying equilibrium term premium, suggests that the cointegrating rank is five, which is consistent with the theory of term structure.

It should be noted that previous works by Saikkonen (2001a, b) and Rippati and Saikkonen (2001), who incorporate smooth nonlinear deterministic trends in cointegrated VAR systems, are very close to ours in terms of motivation. In particular, Saikkonen (2001a) considers a general smooth nonlinear trend function in cointegrating relations and derives a version of Granger's representation theorem. While Saikkonen's (2001a) framework is similar to ours, his results cannot be directly applied to our model. For this reason, we develop another version of Granger's representation theorem suitably modified to our model.

This study is composed of seven sections. In Section 2, a class of cointegrated VAR models, with trigonometric deterministic trends included in cointegrating relations, are introduced. In Section 3, the moving-average representations of the models are presented. In Section 4, the limit distributions of log-likelihood ratio test statistics for the cointegrating rank in models with trigonometric trends are derived, and approximate limit quantiles of the test statistics are tabulated. The procedures to test the null hypothesis of no cointegration in the presence of nonlinear trends are discussed in Section 5. A Monte Carlo simulation experiment to evaluate the sequential procedure to select the cointegrating rank in finite sample is conducted in Section 6. An empirical application to the term structure of interest rates is provided in Section 7. Lastly,

concluding remarks are made in Section 8.[3]

Before proceeding to the next section, we present a set of notational conventions used throughout the paper. Suppose that ψ is a certain matrix with full column rank so that $(\psi'\psi)^{-1}$ exists and $\bar{\psi} = \psi(\psi'\psi)^{-1}$. Let $\psi_\perp$ be the orthogonal complement of ψ, such that the equality $\psi'_\perp\psi = 0$ holds with the matrix $(\psi, \psi_\perp)$ being of full rank. The outer product of two identical vectors is represented by $\otimes 2$ in superscript when treating large-dimensional vectors. Let $I(d)$ be a stochastic process integrated of order d. The symbols $[b]$ and $\Rightarrow$ represent the integer part of a real number b and weak convergence, respectively. For any two vector time series H_t and J_t for $t = 1, ..., T$, a residual series from regression of H_t on J_t is denoted by

$$(H_t|J_t) = H_t - \sum_{i=1}^{T} H_i J_i' \left(\sum_{i=1}^{T} J_i J_i' \right)^{-1} J_t.$$

Similarly, for any two vector processes H_u and J_u on $u \in [0, 1]$, a process derived from regression of H_u on J_u is expressed as

$$(H_u|J_u) = H_u - \int_0^1 H_v J_v' dv \left(\int_0^1 J_v J_v' dv \right)^{-1} J_u.$$

2 Cointegrated VAR systems with smooth nonlinear trends

The cointegrated VAR analysis was pioneered by Johansen (1988, 1996) and has since been one of the most important econometric methods for the study of nonstationary time series data.[4] Let us first present the unrestricted VAR(k) model for a p-dimensional $I(1)$ time series X_t given by

$$\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \Phi D_t + \varepsilon_t \quad \text{for} \quad t = 1, ..., T, \tag{1}$$

where ε_t is a $p \times 1$ vector of mean-zero i.i.d. Gaussian sequence with a $p \times p$ positive-definite variance matrix Ω; D_t is a $s \times 1$ vector of deterministic variables; Π, Γ_i for

[3] All computational analyses were carried out using *Ox* (Doornik, 2013) and *PcGive* (Doornik and Hendry, 2013).

[4] See Hunter, Burke and Canepa (2017), for example, for details of recent developments in the cointegrated VAR methodology.

$i = 1, \ldots, k-1$ and Φ are $p \times p$, $p \times p$, and $p \times s$ coefficient matrices, respectively; and $\{X_{-k+1}, ..., X_0\}$ is a set of starting values. The cointegrating rank of the system (1) is $r(< p)$ with a reduced rank condition $\Pi = \alpha\beta'$ where α and β are $p \times r$ full column rank matrices.

In practice, the cointegrated VAR(k) model is often reformulated by incorporating either a restricted constant or a restricted linear trend (see Johansen, 1994). When a constant term is placed inside the cointegrating relations, we have a *constant restricted model* defined as

$$\Delta X_t = \alpha\left(\beta' X_{t-1} + \gamma'\right) + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \varepsilon_t. \tag{2}$$

The model (2) imposes restrictions $\Phi D_t = \alpha\gamma'$ with $\Phi = \alpha\gamma'$ and $D_t = 1$ on (1) where γ is an $1 \times r$ vector. Alternatively, when a linear trend term is placed inside the cointegrating relations (but a constant term is placed outside them), we have a *linear trend restricted model* defined as

$$\Delta X_t = \alpha\left(\beta' X_{t-1} + \gamma' t\right) + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \mu + \varepsilon_t. \tag{3}$$

The model (3) imposes restrictions $\Phi D_t = \alpha\gamma' t + \mu$ with $\Phi = (\alpha\gamma', \mu)$ and $D_t = (t, 1)'$ on (1) where γ and μ are an $1 \times r$ vector and a $p \times 1$ vector μ, respectively.

In an important study, Johansen, Mosconi and Nielsen (2000) extended the models (2) and (3) to allow for shifts in the level and in the slope of the linear trend. For example, a model with a break in the restricted constant at $t = T_1(< T)$ can be defined as

$$\Delta X_t = \alpha\left(\beta' X_{t-1} + \gamma_1' 1_{(0<t\leq T_1)} + \gamma_2' 1_{(T_1<t\leq T)}\right) + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \sum_{j=1}^{k} \Psi_j 1_{(t=T_1+j)} + \varepsilon_t, \tag{4}$$

where $1_{(\cdot)}$ denotes an indicator function that is assigned 1 if a condition in the parentheses is met and 0, otherwise, and $\Psi_i \in \mathbf{R}^p$ for $j = 1, \ldots, k$. Here, the model (2) has been reformulated into (4), such that (i) the restricted constant is now subject to a discrete break dividing the total sample period into two sub-samples, and (ii) the period over $X_{T_1+1}, \ldots, X_{T_1+k}$ is treated as a transition from the first sub-sample and the

second sub-sample. This type of model is certainly useful in modeling and analyzing time series data influenced by an abrupt regime change. In contrast, we will extend the models (2) and (3) in a different direction from Johansen, Mosconi and Nielsen (2000), such that the smooth nonlinear trending features can be incorporated. In particular, we propose a class of two cointegrated VAR models containing trigonometric functions as additional deterministic components.

Let n denote the total number of frequencies used in the trigonometric function. We introduce a $2n \times 1$ vector of nonlinear deterministic terms, denoted by $\widetilde{D}_t$, along with its parameter $\Upsilon \in \mathbf{R}^{p\times 2n}$. Adding $\Upsilon\widetilde{D}_t$ to (1) under $\Pi = \alpha\beta'$ yields

$$\Delta X_t = \alpha\beta' X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \Phi D_t + \Upsilon\widetilde{D}_t + \varepsilon_t. \tag{5}$$

The first model extends the model (2) and is defined by

$$\Delta X_t = \alpha\left(\beta' X_{t-1} + \delta' F_{t,T} + \gamma'\right) + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \varepsilon_t, \tag{6}$$

where

$$\delta' F_{t,T} = \sum_{j=1}^{n} \left[\delta_{s,j} \sin\left(\frac{2\pi jt}{T}\right) + \delta_{c,j} \cos\left(\frac{2\pi jt}{T}\right)\right]$$

with $\delta' = (\delta_{s,1}, \delta_{c,1}, \cdots, \delta_{s,n}, \delta_{c,n}) \in \mathbf{R}^{r\times 2n}$ and

$$F_{t,T} = \left[\sin\left(\frac{2\pi t}{T}\right), \cos\left(\frac{2\pi t}{T}\right), \cdots, \sin\left(\frac{2\pi nt}{T}\right), \cos\left(\frac{2\pi nt}{T}\right)\right]'.$$

The model (6), which we refer to a *constant and nonlinear trend restricted (CNR) model*, imposes restrictions on (5), such that ΦD_t is the same as that in (2) and $\Upsilon\widetilde{D}_t = \alpha\delta' F_{t,T}$. The CNR model excludes a linear trend but includes the smooth nonlinear components and the intercept inside the cointegrating space. It should be noted that this type of nonlinear trend function has long been considered in time series literature (see e.g., Anderson, 1971).

The second model extends the model (3) and is defined by

$$\Delta X_t = \alpha\left(\beta' X_{t-1} + \delta' F_{t,T} + \gamma' t\right) + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \mu + \varepsilon_t \tag{7}$$

The model (7), which we refer to as a *linear and nonlinear trend restricted (LNR) model*, imposes restrictions on (5), such that ΦD_t is the same as that in model (3) and $\Upsilon \widetilde{D}_t$ is the same as that in model (6). The LNR model has all the linear and nonlinear deterministic trends restricted within the cointegrating space; only the constant is placed in the model in an unrestricted manner.

3 Granger-Johansen representations

In order to examine the asymptotic properties of X_t, we derive a class of moving-average representations, often called Granger-Johansen representations, according to the specifications of deterministic terms in the CNR and LNR models.

Assumption 3.1 *Define a characteristic equation based on (1) as*

$$\det\left\{(1-z)\, I_p - \Pi z - \sum_{i=1}^{k-1} \Gamma_j\,(1-z)\, z^i\right\} = 0,$$

and the roots of this equation are either $z = 1$ or $|z| > 1$ in the complex plane. Define the decomposition $\Pi = \alpha\beta'$ for $\alpha, \beta \in \mathbf{R}^{p\times r}$, and both α and β are of full column rank $r < p$. Define $\Gamma = I_p - \sum_{i=1}^{k-1} \Gamma_i$, and the square matrix $\alpha'_\perp \Gamma \beta_\perp$ is of full rank $p - r$.

Under Assumption 3.1, we follow Hansen (2005) and introduce a set of new composite parameters

$$C = \beta_\perp(\alpha'_\perp\Gamma\beta_\perp)^{-1}\alpha'_\perp, \quad \xi = (I_p - C\Gamma)\,\bar{\beta}, \quad \eta' = \bar{\alpha}'\,(I_p - \Gamma C)\,,$$
$$\Psi = \textstyle\sum_{i=1}^{k-1} i\Gamma_i, \quad \tau_c = -\xi\gamma' \quad \text{and} \quad \tau_l = C\mu - \xi\gamma'.$$

As shown in Lemma 4.1 in Johansen (1996), the expansion of the polynomial $C(z) = \sum_{i=0}^{\infty} C_i z^i$ around $z = 1$ leads to

$$C(z) = C(1) + (1-z)\, C^*(z),$$

where $C^*(z)$ is also a polynomial, such that $C^*(1) = -\frac{dC(z)}{dz}\,|_{z=1}$ (see also Phillips and Solo, 1992). In this section, X_t is denoted by $X_{t,T}$ to stress the fact that the process

contains $F_{t,T}$ as its constituent, which is a function of t as well as the sample size T. The next proposition provides the representation of the CNR model, which is seen as a reformulation of the representation given in Theorem 2.1 in Saikkonen (2001a).

Proposition 3.2 *Suppose Assumption 3.1 is satisfied in the CNR model (6). The Granger-Johansen representation of the CNR model for* $t = 1, ..., T$ *is*

$$X_{t,T} = C \sum_{i=1}^{t} \varepsilon_i + C(L)\varepsilon_t - \xi\delta' F_{t,T} + (1-L)\, C^*(L)\alpha\delta' F_{t,T} + \tau_c + A, \tag{8}$$

where $A = C\left(X_0 - \Gamma_1 X_{-1} - \cdots - \Gamma_{k-1} X_{-k+1}\right)$.

The term consisting of the initial values was given in Saikkonen (2001a, Theorem 2.1) as $A_0 = \beta_\perp(\beta_\perp'\beta_\perp)^{-1}\beta_\perp' X_0$, which has been replaced with A as shown above and in Hansen (2005, Theorem 1), although the difference can be asymptotically marginal due to $\beta' A_0 = \beta' A = 0$. In other words, A is expressed as $A = A_0 + \beta_\perp Z_0$, in which Z_0 is composed of the initial values of the underlying stationary components; see the proof of Lemma 2 in Hansen (2005) for further details.

Proof. See Appendix A. ■

Next, we provide the representation of the LNR model, which is proven in a manner similar to Proposition 3.2, although we have to treat the linear trend with care. It should be noted that Saikkonen (2001a, Theorem 2.1) did not address the representation including the linear trend.

Proposition 3.3 *Suppose Assumption 3.1 is satisfied in the LNR model (6). The Granger-Johansen representation of the LNR model for* $t = 1, ..., T$ *is*

$$\begin{aligned} X_{t,T} = C \sum_{i=1}^{t} \varepsilon_i + C(L)\varepsilon_t - \xi\delta' F_{t,T} + (1-L)\, C^*(L)\alpha\delta' F_{t,T} + \tau_l t \\ - \left(\xi\eta' + C\Psi C\right)\mu - \left(\xi + \xi\eta'\Gamma\xi - C\Psi\xi\right)\gamma' + A, \end{aligned} \tag{9}$$

where $A = C\left(X_0 - \Gamma_1 X_{-1} - \cdots - \Gamma_{k-1} X_{-k+1}\right)$.

Proof. See Appendix A. ■

Note that $F_{t,T}$ does not accumulate in both representations, in contrast to ε_t, as a result of the restriction that $F_{t,T}$ is placed within the cointegrating space. It should also be emphasized that the representation of the CNR model is free of the linear trend, which is one of the critical differences from that of the LNR model.

Both of the representations contain $(1-L)\,C^*(L)\alpha\delta' F_{t,T}$, which is $\mathrm{O}(T^{-1})$ because of a set of trigonometric identities given by

$$\begin{aligned}
\Delta \sin\left(\frac{2\pi jt}{T}\right) &= 2\sin\left(\frac{\pi j}{T}\right)\cos\left(\frac{2\pi jt+\pi j}{T}\right) = \mathrm{O}(T^{-1}),\\
\Delta \cos\left(\frac{2\pi jt}{T}\right) &= -2\sin\left(\frac{\pi j}{T}\right)\sin\left(\frac{2\pi jt+\pi j}{T}\right) = \mathrm{O}(T^{-1}),
\end{aligned}$$

for $j = 1, \ldots, n$. Thus, we can argue that $(1-L)\,C^*(L)\alpha\delta' F_{t,T}$ plays no role in the asymptotic theory developed in the next section.

4 Asymptotic distributions of test statistics

A log-likelihood ratio test statistic for the choice of cointegrating rank r against p, denoted by $-2\log LR\,(r|p;n)$, can be obtained from the reduced rank regression of two models for a given number of frequencies n. See Johansen (1996, Chs. 6, 10 and 11) for further details about reduced rank regression and related limit theorems. Reduced rank regression leads to a class of squared sample canonical correlations, $1 \geq \hat{\lambda}_1 \geq \cdots \geq \hat{\lambda}_p \geq 0$, which are the solutions to the generalized eigenvalue problem given by

$$\det(\lambda S_{11} - S_{10}S_{00}^{-1}S_{01}) = 0,$$

where S_{01}, S_{10} and S_{11} represent the sample product moment matrices defined below. The test statistic can then be written as

$$\begin{aligned}
-2\log LR\,(r|p;n) &= -T\sum_{i=r+1}^{p}\log(1-\hat{\lambda}_i)\\
&= T\,\mathrm{tr}\left\{P_{\alpha_\perp}S_{01}N_T\left(N_T'S_{11}N_T\right)^{-1}N_T'S_{10}\right\} + \mathrm{o_P}(1), \qquad (10)
\end{aligned}$$

where $P_{\alpha_\perp} = \alpha_\perp \left(\alpha_\perp' \Omega \alpha_\perp\right)^{-1} \alpha_\perp'$ and N_T denotes a transformation matrix defined below.

Since the CNR model is nested in the LNR model, we first examine the LNR model. For the LNR model, after running auxiliary regressions, the sample product moment matrices are given by

$$S_{01} = \frac{1}{T}\sum_{t-1}^{T} \left(\Delta X_t \left| \begin{array}{c} 1 \\ \Delta \mathbb{X}_{t-1} \end{array} \right. \right) \left(\begin{array}{c|c} X_{t-1} & 1 \\ F_{t,T} & \\ t & \Delta \mathbb{X}_{t-1} \end{array} \right)',$$

$$S_{00} = \frac{1}{T}\sum_{t-1}^{T} \left(\Delta X_t \left| \begin{array}{c} 1 \\ \Delta \mathbb{X}_{t-1} \end{array} \right. \right)^{\otimes 2}, \text{ and } \quad S_{11} = \frac{1}{T}\sum_{t=1}^{T} \left(\begin{array}{c|c} X_{t-1} & 1 \\ F_{t,T} & \\ t & \Delta \mathbb{X}_{t-1} \end{array} \right)^{\otimes 2},$$

where $\Delta \mathbb{X}_{t-1} = \left(\Delta X_{t-1}', \cdots, \Delta X_{t-k+1}'\right)'$. By using the transformation

$$N_T = \begin{pmatrix} \Gamma' \alpha_\perp & 0 & 0 \\ \delta \xi' \Gamma' \alpha_\perp & T^{1/2} I_{2n} & 0 \\ -\tau_l' \Gamma' \alpha_\perp & 0 & T^{-1/2} \end{pmatrix},$$

which is based on the Granger-Johansen representation (9) and $\alpha_\perp' \Gamma C = \alpha_\perp'$, both $F_{t,T}$ and t can be removed from X_{t-1} as

$$\begin{aligned} N_T' \left(\begin{array}{c|c} X_{t-1} & 1 \\ F_{t,T} & \\ t & \Delta \mathbb{X}_{t-1} \end{array} \right) &= \begin{pmatrix} \alpha_\perp' \Gamma & \alpha_\perp' \Gamma \xi \delta' & -\alpha_\perp' \Gamma \tau_l \\ 0 & T^{1/2} I_{2n} & 0 \\ 0 & 0 & T^{-1/2} \end{pmatrix} \left(\begin{array}{c|c} X_{t-1} & 1 \\ F_{t,T} & \\ t & \Delta \mathbb{X}_{t-1} \end{array} \right) \\ &= \left(\begin{array}{c|c} \alpha_\perp' \sum_{i=1}^{t} \varepsilon_i & 1 \\ T^{1/2} F_{t,T} & \\ T^{-1/2} t & \Delta \mathbb{X}_{t-1} \end{array} \right) + \mathrm{O_P}(1). \end{aligned} \tag{11}$$

As a consequence of this transformation, the random-walk component remains in X_{t-1} as the dominant term. The statistic $-2 \log LR\left(r|p;n\right)$ can then be seen as asymptotically similar with respect to the parameters of the deterministic components; see Cox and Hinkley (1974, Ch.5) as well as Nielsen and Rahbek (2000) for further details of similarity. Normalizing the first $p-r$ coordinates of (11) then leads to the following weak convergence:

$$T^{-1/2} \left(\alpha_\perp' \sum_{i=1}^{t} \varepsilon_i \left| \begin{array}{c} 1 \\ \Delta \mathbb{X}_{t-1} \end{array} \right. \right) \Rightarrow \left(\alpha_\perp' W_u | \, 1\right) \text{ on } u \in [0,1], \tag{12}$$

where $W_u = (W_{1,u}, \ldots, W_{p-r,u})'$ is a $(p-r) \times 1$ vector Brownian motion with variance Ω.

Let us get back to (10) and note the following identity:

$$\begin{aligned} & T \operatorname{tr}\left\{P_{\alpha_\perp} S_{01} N_T \left(N_T' S_{11} N_T\right)^{-1} N_T' S_{10}\right\} \\ = \; & \operatorname{tr}\left\{P_{\alpha_\perp}^* S_{01} N_T \Lambda \left(T^{-1} \Lambda' N_T' S_{11} N_T \Lambda\right)^{-1} \Lambda' N_T' S_{10} P_{\alpha_\perp}^{*\prime}\right\}, \end{aligned} \tag{13}$$

where $P_{\alpha_\perp}^* = \left(\alpha_\perp' \Omega \alpha_\perp\right)^{-1/2} \alpha_\perp'$ and Λ is an invertible matrix defined as

$$\Lambda = \begin{pmatrix} \left(\alpha_\perp' \Omega \alpha_\perp\right)^{-1/2} & 0 \\ 0 & I_{2n+1} \end{pmatrix}.$$

Utilizing a class of limit theorems in Johansen (1996, Chs. 10 and 11), we can derive the limit distribution of each product moment in (13):

$$P_{\alpha_\perp}^* S_{01} N_T \Lambda \Rightarrow \int_0^1 dB_u G_u' \quad \text{and} \quad T^{-1} \Lambda' N_T' S_{11} N_T \Lambda \Rightarrow \int_0^1 G_u G_u' du, \tag{14}$$

where $B_u = \left(B_{1,u}, \ldots, B_{p-r,u}\right)'$ on $u \in [0,1]$ is a $(p-r) \times 1$ standard Brownian motion, and G_u on $u \in [0,1]$ is given as

$$G_u = \left(B_u', \mathcal{F}_u', u \middle| 1\right)', \tag{15}$$

where $\mathcal{F}_u$ is $2n \times 1$ vector function defined by

$$\mathcal{F}_u = \left[\sin\left(2\pi u\right), \cos\left(2\pi u\right), \ldots, \sin\left(2\pi n u\right), \cos\left(2\pi n u\right)\right]'.$$

For the CNR model, after running auxiliary regressions, the sample product moment matrices are given by

$$S_{01} = \frac{1}{T} \sum_{t-1}^{T} \left(\Delta X_t \middle| \Delta \mathbb{X}_{t-1}\right) \left(\begin{array}{c|c} X_{t-1} & \\ F_{t,T} & \Delta \mathbb{X}_{t-1} \\ 1 & \end{array} \right)',$$

$$S_{00} = \frac{1}{T} \sum_{t-1}^{T} \left(\Delta X_t \middle| \Delta \mathbb{X}_{t-1}\right)^{\otimes 2}, \text{ and } S_{11} = \frac{1}{T} \sum_{t=1}^{T} \left(\begin{array}{c|c} X_{t-1} & \\ F_{t,T} & \Delta \mathbb{X}_{t-1} \\ 1 & \end{array} \right)^{\otimes 2}.$$

By using the transformation

$$N_T = \begin{pmatrix} \Gamma' \alpha_\perp & 0 & 0 \\ \delta \xi' \Gamma' \alpha_\perp & T^{1/2} I_{2n} & 0 \\ -\tau_c' \Gamma' \alpha_\perp & 0 & T^{1/2} \end{pmatrix},$$

which is based on (8), the vector $F_{t,T}$ is removed as

$$
\begin{aligned}
N_T' \left(\begin{array}{c} X_{t-1} \\ F_{t,T} \\ 1 \end{array} \middle| \Delta \mathbb{X}_{t-1} \right) &= \begin{pmatrix} \alpha_\perp' \Gamma & \alpha_\perp' \Gamma \xi \delta' & -\alpha_\perp' \Gamma \tau_c \\ 0 & T^{1/2} I_{2n} & 0 \\ 0 & 0 & T^{1/2} \end{pmatrix} \left(\begin{array}{c} X_{t-1} \\ F_{t,T} \\ 1 \end{array} \middle| \Delta \mathbb{X}_{t-1} \right) \\
&= \left(\begin{array}{c} \alpha_\perp' \sum_{i=1}^{t} \varepsilon_i \\ T^{1/2} F_{t,T} \\ T^{1/2} \end{array} \middle| \Delta \mathbb{X}_{t-1} \right) + \mathrm{O_P}(1).
\end{aligned}
$$

Again, $-2 \log LR\,(r|p;n)$ can be viewed as asymptotically similar with respect to the parameters of the deterministic components. With the use of Λ, we arrive at a set of limit expressions of product moments in (14), both of which are now composed of

$$
G_u = (B_u', \mathcal{F}_u', 1)',
$$

in place of (15).

Proposition 4.1 *Suppose Assumption 3.1 is satisfied in the CNR model (6) and the LNR model (7). The asymptotic distribution of the log-likelihood ratio test statistic for the null of cointegrating rank r is*

$$
-2 \log LR\,(r|p;n) \Rightarrow \mathrm{tr} \left\{ \int_0^1 dB_u G_u' \left(\int_0^1 G_u G_u' du \right)^{-1} \int_0^1 G_u dB_u' \right\},
$$

where G_u on $u \in [0,1]$ is defined as

$$
G_u = (B_u', \mathcal{F}_u', 1)'
$$

for the CNR model, and

$$
G_u = (\, B_u', \mathcal{F}_u', u | \, 1)'
$$

for the LNR model.

Since both the limit distributions in Proposition 4.1 are free from nuisance parameters, the distributions can be approximated by simulation, given a set of values for $p-r$ and n. We follow Nielsen (1997) and Doornik (1998) in approximating the limit distributions. Let us introduce a $(p-r) \times 1$ pseudo normalized standard Gaussian process ϵ_t and its partial sum process $Z_{t-1} = \sum_{i=1}^{t-1} \epsilon_i$ with $Z_0 = 0$. We construct

$$
\mathsf{Q} = -T \sum_{i=1}^{p-r} \log(1 - \tilde{\lambda}_i),
$$

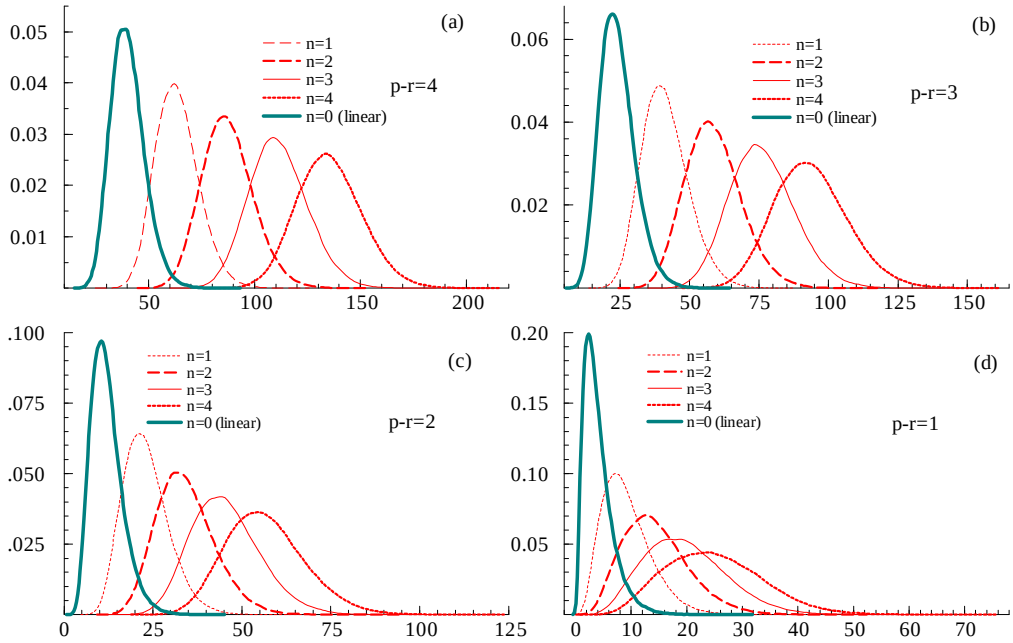

Figure 1: Simulated limit distributions: CNR and standard (constant restricted) models

where $\tilde{\lambda}_1, ..., \tilde{\lambda}_{p-r}$ represent the eigenvalues of $T^{-1}\mathsf{G}_{\epsilon 1}\left(\mathsf{G}_{11}\right)^{-1}\mathsf{G}_{1\epsilon}$, which consists of

$$\mathsf{G}_{\epsilon 1} = \sum_{t=1}^{T} \epsilon_t R_{1t}', \quad \mathsf{G}_{1\epsilon} = \mathsf{G}_{\epsilon 1}' \quad \text{and} \quad \mathsf{G}_{11} = \sum_{t=1}^{T} R_{1t} R_{1t}',$$

with $R_{1t} = (Z_{t-1}', F_{t,T}', 1)'$ for the CNR model and with $R_{1t} = \left(Z_{t-1}', F_{t,T}', t \middle| 1\right)'$ for the LNR model. See Appendix B for a set of tables that present the approximate limit quantiles, the first and second moments of the rank test statistics for the CNR and LNR models, for various combinations of $p-r$ and n.

Figure 1 compares simulated (approximate) limit distributions for CNR model (6) and standard constant restricted model (2), given $p-r \ = 4, 3, 2, 1$ and $n = 4, 3, 2, 1$. Similarly, Figure 2 makes a comparison of those distributions for LNR model (7) and standard linear trend restricted model (3), given the same settings for $p-r$ and n as in Figure 1. The numbers of observations and Monte Carlo replications are the same as those adopted in the simulation for Appendix B. As n increases, the distributions shift to the right and their shapes become less skewed to the right. These comparisons allow

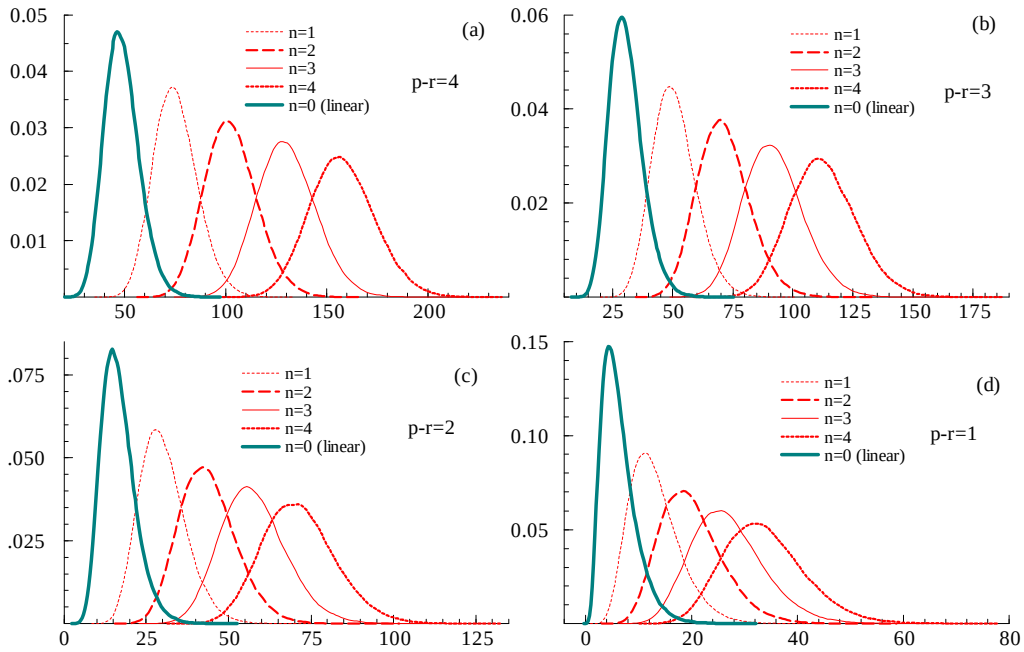

Figure 2: Simulated limit distributions: LNR and standard (linear trend restricted) models

us to predict that the use of standard models, when studying those data which are subject to Fourier-type trends, will lead to a misleading inference about the underlying cointegrating rank. The next section demonstrates this issue, along with other aspects, through Monte Carlo simulations.

It is known that a class of gamma distributions can be used to approximate limit distributions of log-likelihood ratio tests for the cointegrating rank under various circumstances; see Nielsen (1997), Doornik (1998), Johansen, Mosconi and Nielsen (2000) and Kurita and Nielsen (2019) for further details. We apply this method to the CNR and LNR models, resulting in success in approximating the underlying limit distributions. It is therefore feasible, by running a class of response surface regressions, to calculate p-values of observed test statistics according to the corresponding gamma distributions. An *Ox* code for the purpose of calculating p-values is available in the supplementary material for this paper. The empirical illustration given in Section 7 below utilizes this code in the determination of the cointegrating rank.

5 Nonlinear trend in the absence of cointegration

We have thus far considered a class of models in which Fourier-type nonlinear deterministic trends are contained in a cointegrating relation. These models are advantageous in that we can explicitly allow for various smooth nonlinearities, approximated by a set of smooth trigonometric functions, embedded in the underlying long-run economic relationships. In our specification, however, nonlinear trends disappear from the model under the null hypothesis of no cointegration. In the literature of the unit root test with structural breaks, some studies including Kim and Perron (2009) and Carrion-i-Silvestre, Kim and Perron (2009) allow a break under both the null and alternative hypotheses. Along this line, Enders and Lee (2012a) and Enders and Lee (2012b) respectively propose a Lagrange Multiplier (LM) type and Dickey-Fuller type unit root tests allowing for a flexible nonlinear trend using a Fourier approximation under both the null and alternative hypotheses. Rodrigues and Taylor (2012) also consider the same nonlinear trend in their local GLS detrended test for a unit root. In particular, for a scalar time series y_t, the null hypothesis of $\rho = 0$ is tested in the model given by

$$\Delta y_t = \rho y_{t-1} + \sum_{j=1}^{n} \left\{ \delta_{1j} \sin\left(\frac{2\pi jt}{T}\right) + \delta_{2j} \cos\left(\frac{2\pi jt}{T}\right) \right\} + \varepsilon_t$$

where ε_t is a mean-zero i.i.d. random sequence with a finite variance. Therefore, nonlinear trends are present in the model, both under the null of $\rho = 0$ and the alternative of $\rho < 0$. We can also generalize this specification to the cointegrated VAR system so that the coefficients on the nonlinear trend terms are identified under the null hypothesis of no cointegration. In this section, we briefly address testing procedures under such specifications that complement the main models considered in the previous sections.[5]

In order to permit the presence of Fourier functions in a model with no cointegration,

[5] A similar framework is also considered in a recent work of Pascalau, Lee, Nazlioglu, and Lu (2022).

the model (5) in Section 2 is now replaced by

$$\Delta X_t = \alpha\beta' X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \Phi D_t + \Upsilon F_{t,T} + \varepsilon_t. \tag{16}$$

In this specification, Υ can be estimated under the null hypothesis of no cointegration because the Fourier function does not disappear, even if $r = 0$; that is,

$$\Delta X_t = \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \Phi D_t + \Upsilon F_{t,T} + \varepsilon_t.$$

At the same time, however, the model (16) rules out the possibility of long-run economic relationships described by smooth nonlinear trends. In the reduced rank regression framework, the effects of nonlinear terms are considered only through auxiliary regressions. In contrast, in the CNR model (6) and LNR model (7), the method of reduced rank regression explicitly yields the maximum likelihood estimator of δ, the parameters for the Fourier-type nonlinear terms embedded in the cointegrating relations. Therefore, the CNR and LNR models are more suitable for the analysis of long-run economic relationships subject to smooth nonlinear trends, which is infeasible when employing the specification (16).

With regard to an intercept and linear trend term in (16), it is convenient to follow the standard specifications as we did for the CNR and LNR models in Section 2. We thus introduce the following class of models

$$\begin{aligned}
\Delta X_t &= \alpha\left(\beta' X_{t-1} + \gamma'\right) + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \Upsilon F_{t,T} + \varepsilon_t, \\
\Delta X_t &= \alpha\left(\beta' X_{t-1} + \gamma' t\right) + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \Upsilon F_{t,T} + \mu + \varepsilon_t.
\end{aligned}$$

The former model is referred to as a *constant restricted and nonlinear trend unrestricted (CNU) model* with $\Phi = \alpha\gamma'$ and $D_t = 1$, while the latter is called a *linear trend restricted and nonlinear trend unrestricted (LNU) model* with $\Phi = (\alpha\gamma', \mu)$ and $D_t = (t, 1)'$. Let us refer back to the last section for a set of of notational conventions; by running auxiliary regressions, the sample product moment matrices for the CNU model

are then specified as follows:

$$S_{01} = \frac{1}{T}\sum_{t-1}^{T}\left(\Delta X_t \,\middle|\, \begin{matrix} F_{t,T} \\ \Delta\mathbb{X}_{t-1} \end{matrix}\right)\left(\begin{matrix} X_{t-1} \\ 1 \end{matrix} \,\middle|\, \begin{matrix} F_{t,T} \\ \Delta\mathbb{X}_{t-1} \end{matrix}\right)', \quad S_{00} = \frac{1}{T}\sum_{t-1}^{T}\left(\Delta X_t \,\middle|\, \begin{matrix} F_{t,T} \\ \Delta\mathbb{X}_{t-1} \end{matrix}\right)^{\otimes 2},$$

$$S_{11} = \frac{1}{T}\sum_{t=1}^{T}\left(\begin{matrix} X_{t-1} \\ 1 \end{matrix} \,\middle|\, \begin{matrix} F_{t,T} \\ \Delta\mathbb{X}_{t-1} \end{matrix}\right)^{\otimes 2} \quad \text{and} \quad N_T = \begin{pmatrix} \Gamma'\alpha_\perp & 0 \\ 0 & T^{1/2} \end{pmatrix}.$$

Similarly, for the LNU model, those matrices are

$$S_{01} = \frac{1}{T}\sum_{t-1}^{T}\left(\Delta X_t \,\middle|\, \begin{matrix} 1 \\ F_{t,T} \\ \Delta\mathbb{X}_{t-1} \end{matrix}\right)\left(\begin{matrix} X_{t-1} \\ t \end{matrix} \,\middle|\, \begin{matrix} 1 \\ F_{t,T} \\ \Delta\mathbb{X}_{t-1} \end{matrix}\right)', \quad S_{00} = \frac{1}{T}\sum_{t-1}^{T}\left(\Delta X_t \,\middle|\, \begin{matrix} 1 \\ F_{t,T} \\ \Delta\mathbb{X}_{t-1} \end{matrix}\right)^{\otimes 2},$$

$$S_{11} = \frac{1}{T}\sum_{t=1}^{T}\left(\begin{matrix} X_{t-1} \\ t \end{matrix} \,\middle|\, \begin{matrix} 1 \\ F_{t,T} \\ \Delta\mathbb{X}_{t-1} \end{matrix}\right)^{\otimes 2} \quad \text{and} \quad N_T = \begin{pmatrix} \Gamma'\alpha_\perp & 0 \\ 0 & T^{-1} \end{pmatrix}.$$

Suppose Assumption 3.1 is satisfied with respect to both models. The limit distribution of the log-likelihood ratio test statistic for the null of cointegrating rank r is expressed as

$$-2\log LR\left(r|p;n\right) \Rightarrow \mathrm{tr}\left\{\int_0^1 dB_u G_u'\left(\int_0^1 G_u G_u' du\right)^{-1}\int_0^1 G_u dB_u'\right\},$$

where G_u on $u \in [0,1]$ is defined as

$$G_u = \left(B_u' \middle| \mathcal{F}_u'\right)'$$

for the CNU model, and

$$G_u = \left(B_u', u \middle| 1, \mathcal{F}_u'\right)'$$

for the LNU model.

The *Ox* code available in the supplementary material simulates a class of approximates to the above limit distributions, along with those based on the proposed models that were studied in the previous section. A series of response surface regressions incorporated in the code also allows us to calculate the p-values of observed test statistics, according to the corresponding gamma distributions.

6 Monte Carlo simulation

6.1 Non-Fourier DGPs and CNR model

Let us introduce a nonlinear bivariate DGP for $X_t = (X_{1,t}, X_{2,t})'$ with $r = 1$ and $k = 1$ by

$$\Delta X_t = \alpha \left(\beta' X_{t-1} + H_{t,T} + \gamma \right) + \varepsilon_t, \tag{17}$$

where

$$\alpha = \begin{pmatrix} -0.2 \\ 0.1 \end{pmatrix}, \ \beta = \begin{pmatrix} 1.0 \\ -0.5 \end{pmatrix}, \ \gamma = -2.3,$$

and $H_{t,T}$ represents a deterministic nonlinear term that does not fall in a class of Fourier-type smooth nonlinear trends. The starting values are set at log(100) and the innovation sequence ε_t is a pseudo mean-zero $N(0, s\Omega)$ sequence with each diagonal element of Ω being unity and each off-diagonal element of Ω being 0.25, along with a scaling factor $s = 0.01$.

Following Doornik, Hendry and Nielsen (1998), we use a class of approximate 95% quantiles tabulated in the previous section and calculate a set of recursive rejection probabilities of the test statistic $-2 \log LR(r|p; n)$. The numbers of observations (T) range from 50 to 400, increasing by 25 steps each, with the first 20 observations discarded to allow for the influences of the initial values. The number of Monte Carlo replications is 10,000. The standard (constant restricted) model (2) and the CNR model (6) are estimated. A set of 5% critical values (95% asymptotic quantiles) for the former model is taken from Doornik (2003), while 5% critical values for the latter are those tabulated in the previous section. The thick lines in the figures presented below correspond to recursive rejection frequencies when using the proposed models, whereas dotted lines denote rejection rates with the use of the standard (constant restricted) linear model. These rejection rates are sorted according to the null hypotheses of $r = 0$ and $r \leq 1$ and displayed in the upper two panels of each figure, along with a set of sample paths of X_t and $\beta' X_t$ for $T = 400$ in the lower panels.

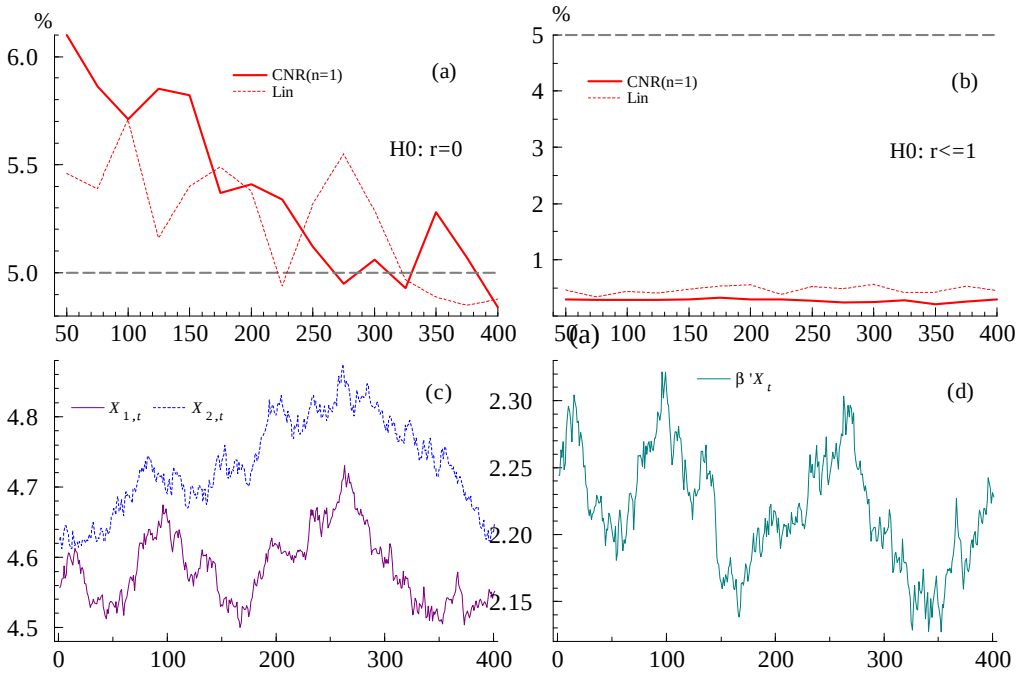

Figure 3: Recursive rejection frequencies and sample paths: NF-DGP-1

We first check the performance of the procedure using a simple DGP with no cointegration.

1. NF-DGP-1: DGP (17), in which α is replaced with $(0,0)'$, that is,

$$\Delta X_t = \varepsilon_t.$$

Models: CNR model ($n = 1$) and standard (constant restricted) model

The results shown in Figure 3 allow us to argue that the finite sample performance of the CNR model is not inferior to that of the standard model, in the analysis of the data from which cointegration and nonlinear components are absent. Next, we move on to a comparative analysis of CNR and standard constant restricted models in the presence of non Fourier-type components in the DGPs.

2. NF-DGP-2: DGP (17), where

$$H_{t,T} = -0.05E\left(0.4T \leq t\right),$$

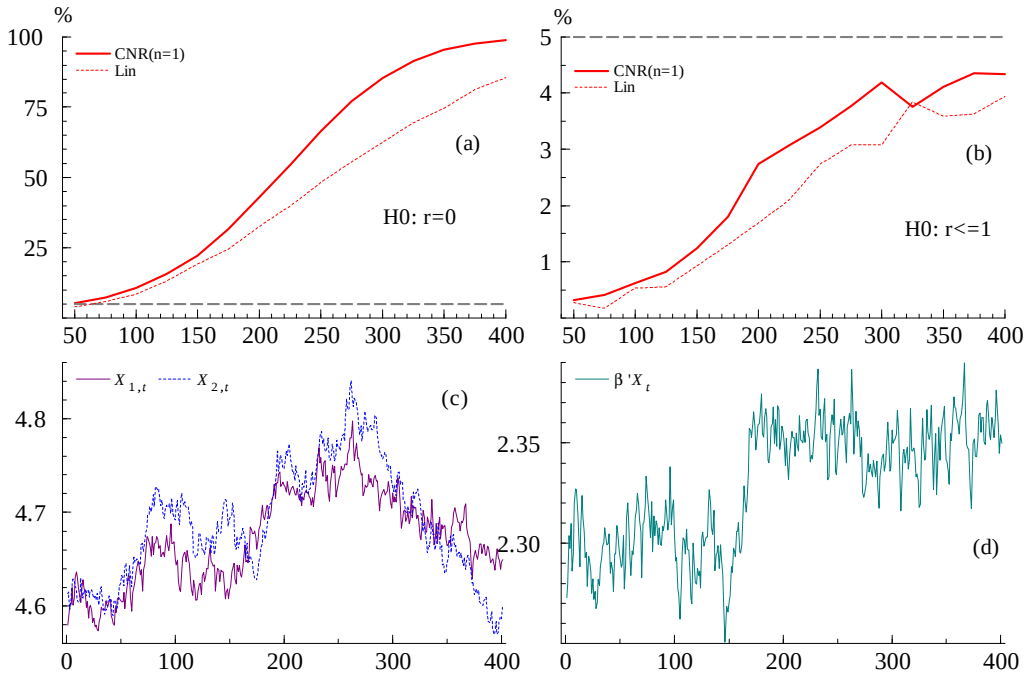

Figure 4: Recursive rejection frequencies and sample paths: NF-DGP-2

and $E(\cdot)$ denotes an indicator function.

Models: CNR model ($n = 1$) and standard (constant restricted) model

See Figure 4 for the simulation results from the use of NF-DGP-2. Although the CNR model does not contain a level-shift term that is included in the underlying DGP, its performance is judged to be better than that of the standard model in terms of the rejection of the false hypothesis $r = 0$.

3. NF-DGP-3: DGP (17), where

$$H_{t,T} = -0.1 \left\{ 1 - \exp \left[-10 \left(\frac{t}{T} - 0.4 \right)^2 \right] \right\}.$$

Models: CNR model ($n = 2$) and standard (constant restricted) model

4. NF-DGP-4: DGP (17), where

$$H_{t,T} = -0.2 \left\{ 1 - \exp \left[-30 \left(\frac{t}{T} - 0.8 \right)^2 \right] \right\}.$$

Models: CNR model ($n = 2$) and standard (constant restricted) model

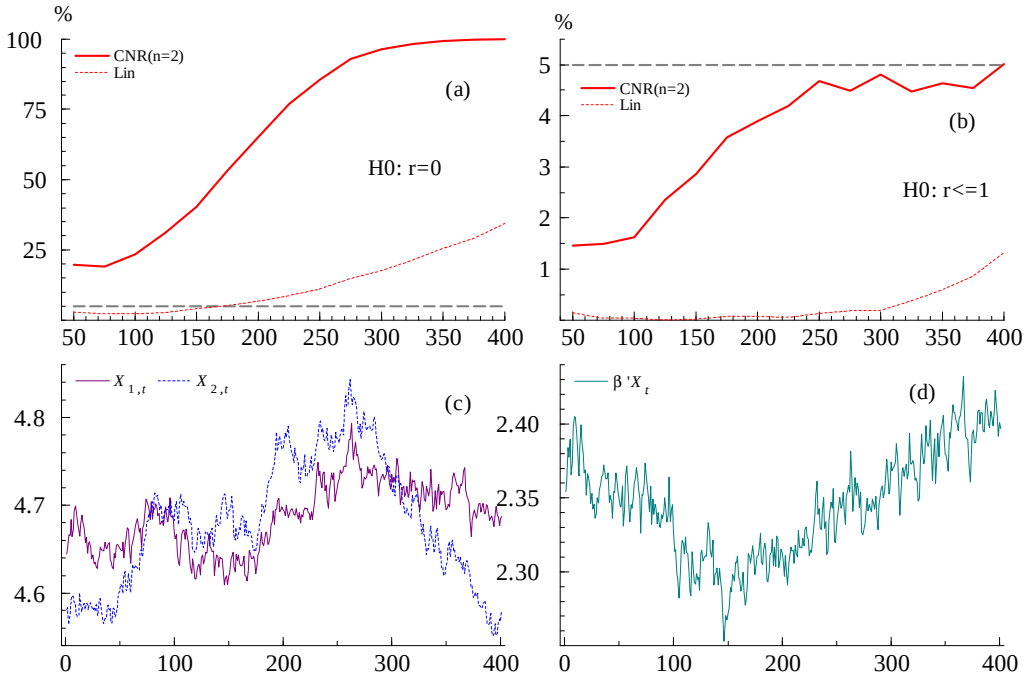

Figure 5: Recursive rejection frequencies and sample paths: NF-DGP-3

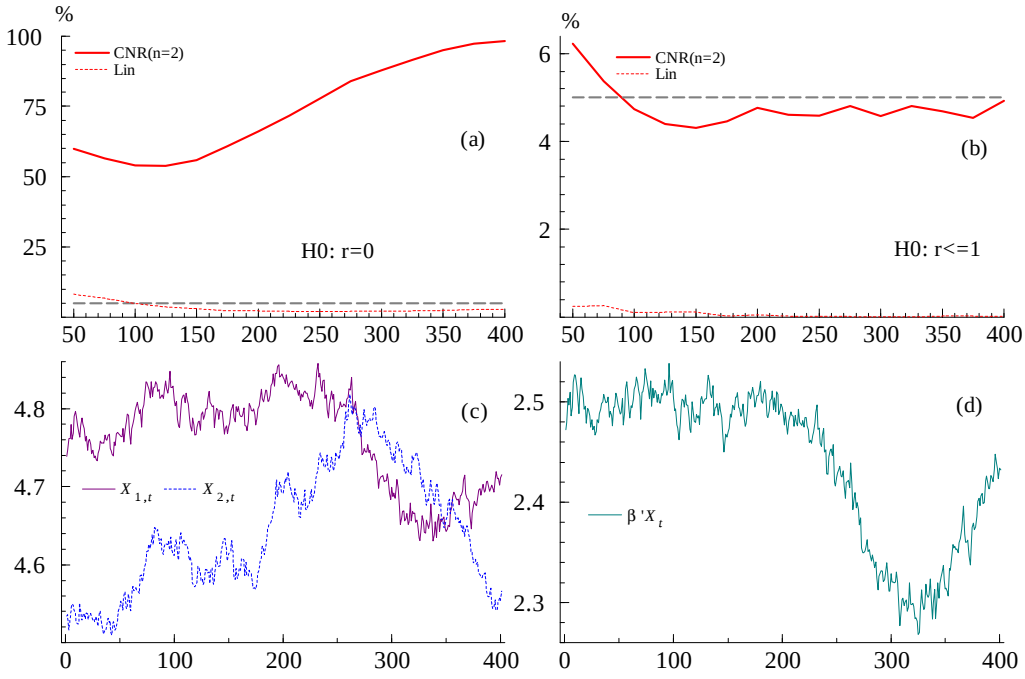

Figure 6: Recursive rejection frequencies and sample paths: NF-DGP-4

See Figures 5 and 6, for which the dynamics of the underlying DGPs (NF-DGP-3 and NF-DGP-4) are driven by exponential smooth transition functions. According to these figures, the CNR models have performed much better than the standard model from the standpoints of the selection of the correct rank, despite the fact that both DGPs are distinctively different from Fourier-type processes.

Simulation results from the use of various other Non-Fourier DGPs are reported in the supplementary material of this paper. Overall, we are justified in concluding that, in comparison with the standard model, the proposed models can be useful in addressing various other nonlinear trends than the Fourier-type trends.

6.2 Fourier DGPs and the CNR model

We proceed to the study of a class of DGPs that explicitly include Fourier-type nonlinear components. Let us introduce a 4-dimensional vector of time series $X_t = (X_{1,t}, X_{2,t}, X_{3,t}, X_{4,t})'$, which is driven by a system in the form

$$\Delta X_t = \alpha \left(\beta' X_{t-1} + \delta' F_{t,T} + \gamma'\right) + \varepsilon_t,$$

where the starting values are set at $\log(100)$ and ε_t is a pseudo mean-zero $N(0, s\Omega)$ sequence with each diagonal element of Ω being unity and each off-diagonal element of Ω being 0.25, along with a scaling factor $s = 0.01$. We first study a class of two DGPs according to the specification of the parameters and periodic frequencies in the system.

The first Fourier-type DGP (denoted as F-DGP-1) based on this system is specified as

$$F_{t,T} = \left[\sin\left(\frac{2\pi t}{T}\right), \cos\left(\frac{2\pi t}{T}\right)\right]',$$

$$\begin{aligned} \alpha &= \begin{pmatrix} -0.2 \\ 0.1 \\ 0.0 \\ 0.0 \end{pmatrix}, \ \beta = \begin{pmatrix} 1.0 \\ -1.0 \\ -1.0 \\ 0.5 \end{pmatrix}, \\ \delta' &= \begin{pmatrix} 0.1 & -0.1 \end{pmatrix}, \ \gamma' = 2.3. \end{aligned}$$

while the second DGP, called F-DGP-2, is given as

$$F_{t,T} = \left[\sin\left(\frac{2\pi t}{T}\right), \cos\left(\frac{2\pi t}{T}\right), \sin\left(\frac{4\pi t}{T}\right), \cos\left(\frac{4\pi t}{T}\right)\right]',$$

$$\begin{aligned} \alpha &= \begin{pmatrix} -0.4 & 0.0 \\ 0.0 & -0.3 \\ -0.2 & -0.1 \\ 0.0 & 0.0 \end{pmatrix}, \ \beta = \begin{pmatrix} 1.0 & 0.0 \\ 0.0 & 1.0 \\ 0.5 & 0.5 \\ -0.5 & 0.5 \end{pmatrix}, \\ \delta' &= \begin{pmatrix} 0.08 & 0.06 & -0.03 & -0.06 \\ 0.00 & -0.08 & 0.04 & 0.04 \end{pmatrix}, \ \gamma' = \begin{pmatrix} -4.6 \\ -9.2 \end{pmatrix}. \end{aligned}$$

We will explore below various characteristics of cointegrating rank tests that are applied to nonstationary time series data generated from these DGPs.

We adopt the same Monte Carlo setting to conduct a comparative analysis as that in Section 6.1, except for one critical difference: the CNR model employed here nests the DGP used in each experiment in such a way that the model is correctly specified in terms of the structure of lag length and trigonometric frequency. Before discussing the results, we present a set of sample paths of artificial data ($T = 400$), generated from F-DGP-1 and F-DGP-2, in Figure 7 (a) and (b), respectively. These figures show that the generated data resemble economic time series data in practice.

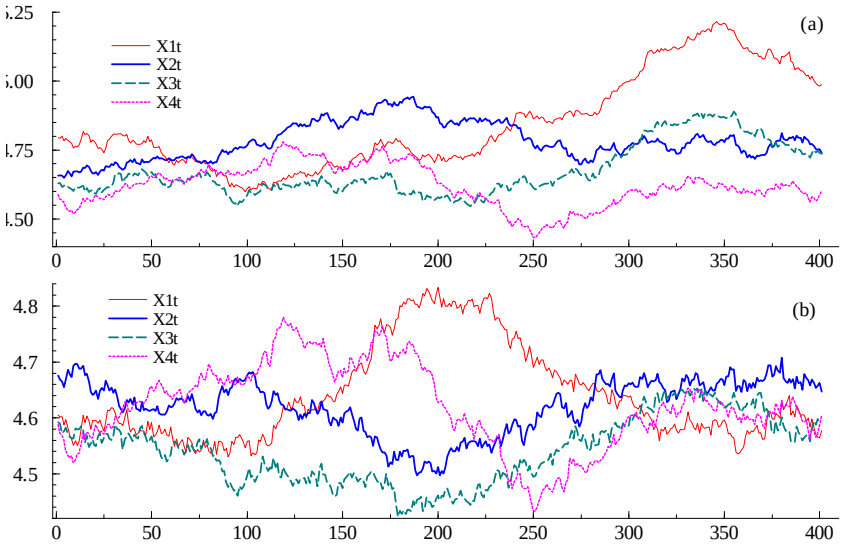

Figure 7: Sample paths of artificial data generated from F-DGP-1 and F-DGP-2

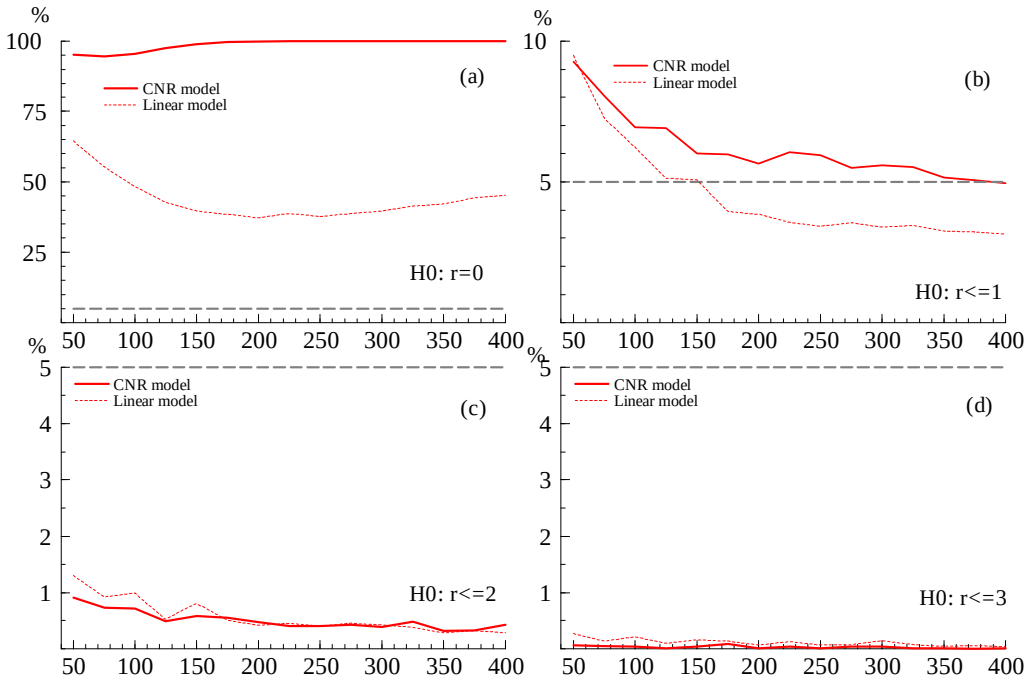

Figure 8: Recursive rejection frequencies: F-DGP-1

We are in a position to study the results of the comparative simulation analysis. First, F-DGP-1 is used as a device to generate a series of artificial data. Recursive rejection rates of the rank test statistics are sorted according to the null hypotheses of $r \leq 0,\ 1,\ 2$ and 3; they are displayed in Panels (a), (b), (c) and (d) in Figure 8. Panel (a) shows the advantage of the proposed nonlinear model over the standard model in terms of selecting the correct cointegrating rank ($r = 1$). The log-likelihood ratio test based on the CNR model rejects almost completely the null hypothesis of $r = 0$ in Panel (a), and the test tends to approach the nominal 5% level as T increases in Panel (b). In contrast, the test fails in many times to detect a cointegrating relation in the standard (constant restricted) model, and the rates of rejecting $r = 0$ in Panel (a) remain around 40 to 50%. The rejection rates for the hypothesis $r \leq 2$, with the use of the standard (constant restricted) model, have gone below the nominal 5% level with an increase in the sample size; see Panel (b).

Next, we proceed to the analysis of F-DGP-2. In line with Figure 8 for F-DGP-1, Panels (a), (b) and (c) in Figure 9 are in favor of the CNR model over the standard

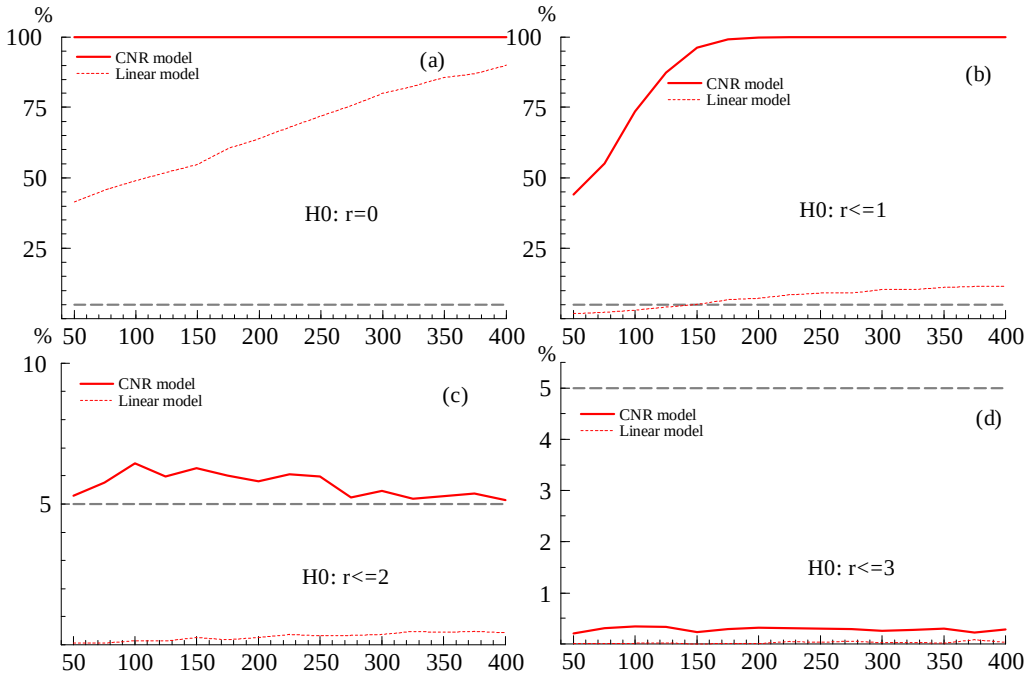

Figure 9: Recursive rejection frequencies: F-DGP-2

(constant restricted) model in the selection of correct cointegrating rank ($r = 2$). The overall performance of the standard (constant restricted) model is much worse than that of the CNR model, reflecting the fact that the DGP is not nested in the former model.

By utilizing the results of this experiment, Table 1 summarizes how frequent we arrive at each of all five conceivable choices $r = 0, 1, 2, 3$ and 4 at each of the following five observation points: $T = 50, 100, 200, 300$ and 400 in the analysis of data generated from F-DGP-1. The use of the CNR model indicates that the selection rates for $r = 1$ approach the level of 95% with an increase in the sample size (see Panel (1)), whereas the use of the standard (constant restricted) model leads to a misleading conclusion, even when the sample size is 400 (see Panel (2)).

Similarly, the analysis of data generated from F-DGP-2 has led to a set of results in support of the CNR model; see Table 2. An approximate 95% selection rate has been attained when $T \geq 200$, while the standard (constant restricted) linear model has failed to reach the same level, even at $T = 400$.

Panel (1) CNR model

	$r=0$	$r=1$	$r=2$	$r=3$	$r=4$	(%)
$T=50$	4.83	85.90	8.36	0.85	0.06	
$T=100$	4.56	88.51	6.22	0.67	0.04	
$T=200$	0.11	94.23	5.19	0.46	0.01	
$T=300$	0	94.41	5.20	0.35	0.04	
$T=400$	0	95.05	4.52	0.42	0.01	

Panel (2) Linear model

	$r=0$	$r=1$	$r=2$	$r=3$	$r=4$	(%)
$T=50$	35.35	55.15	8.20	1.06	0.24	
$T=100$	51.54	42.22	5.25	0.78	0.21	
$T=200$	62.78	33.38	3.46	0.33	0.05	
$T=300$	60.25	36.36	2.99	0.27	0.13	
$T=400$	54.54	42.31	2.88	0.23	0.04	

Table 1: Cointegrating rank selection frequencies: F-DGP-1

The overall phenomena recorded in the above figures and tables are attributable to the underlying non-negligible model misspecifications. The findings also indicate the problem of *spurious stochastic trends*, which should be contrasted with *spurious cointegration*, as discussed by Leybourne and Newbold (2003).

Lastly, Figure 10 displays a class of rejection frequencies of the rank tests based on over-specified, correctly-specifed and under-specified models in terms of the frequency n. Recall $n = 1$ for F-DGP-1($r = 1$), so that the models for Panels (a) and (c) are either over-specified ($n = 5, 4, 3$ and 2) or correctly-specifed ($n = 1$). It should be noted that the power properties of the tests, as shown in Panel (a), are improved as the model's frequency approaches to the true frequency $n = 1$. Recalling $n = 2$ for F-DGP-2($r = 2$), we find that the models for Panels (b) and (d) are classified as over-specified ($n = 5, 4$ and 3), correctly-specified ($n = 2$) and under-specified ($n = 1$). What is distinguishing in Panel (b) is the inferiority of the model with $n = 1$ to all the other models, a result clearly attributed to the under-specification of the frequency. It is thus safer to proceed from a VAR model with a relatively large n when conducting a series of tests for the selection of the cointegrating rank. This procedure is also justifiable in terms of the principle of general-to-specific econometric modeling (see

Panel (1) CNR model

	$r=0$	$r=1$	$r=2$	$r=3$	$r=4$	(%)
$T=50$	0	55.85	38.85	5.09	0.21	
$T=100$	0	26.53	67.01	6.12	0.34	
$T=200$	0	0.11	94.08	5.49	0.32	
$T=300$	0	0	94.54	5.20	0.26	
$T=400$	0	0	94.86	4.86	0.28	

Panel (2) Linear model

	$r=0$	$r=1$	$r=2$	$r=3$	$r=4$	(%)
$T=50$	58.48	39.62	1.83	0.06	0.01	
$T=100$	50.99	45.83	3.04	0.12	0.02	
$T=200$	36.10	56.63	7.02	0.24	0.01	
$T=300$	19.95	69.66	10.04	0.33	0.02	
$T=400$	9.83	78.65	11.09	0.39	0.04	

Table 2: Cointegrating rank selection frequencies: F-DGP-2

Hendry and Doornik, 2014, *inter alia*).

7 Empirical illustration

In this section, we present the results of an empirical application of our proposed procedure to interest rate data. We consider a system of 6 variables given by $X_t = (i_{20yr,t}, i_{10yr,t}, i_{5yr,t}, i_{3yr,t}, i_{1yr,t}, i_{call,t})'$, in which $i_{myr,t}$ for $m = 20, 10, 5, 3$ and 1 represents JGB yield to maturity of m years, while $i_{call,t}$ is Japan's overnight call rate. All data were downloaded from the website of Japan's Ministry of Finance (https://www.mof.go.jp). Because the data are available on a daily basis, they were converted into respective monthly averages. The term structure of interest rates implies that the yield to the shortest maturity in X_t should drive all the other yields to longer maturities, so that the number of cointegrating relations in the VAR system for X_t is expected to be five ($r = 5$), or a single common stochastic trend ($p - r = 1$).

An overview of the data spanning the period of 1986.12 - 1995.11 is provided in Figure 11. The terminal point, 1995.11, was selected on the basis of the observation that the call rate had nearly reached the effective lower bound of 0.5% around the end of 1995. All the interest series in Figure 11 appear to behave in a fairly similar

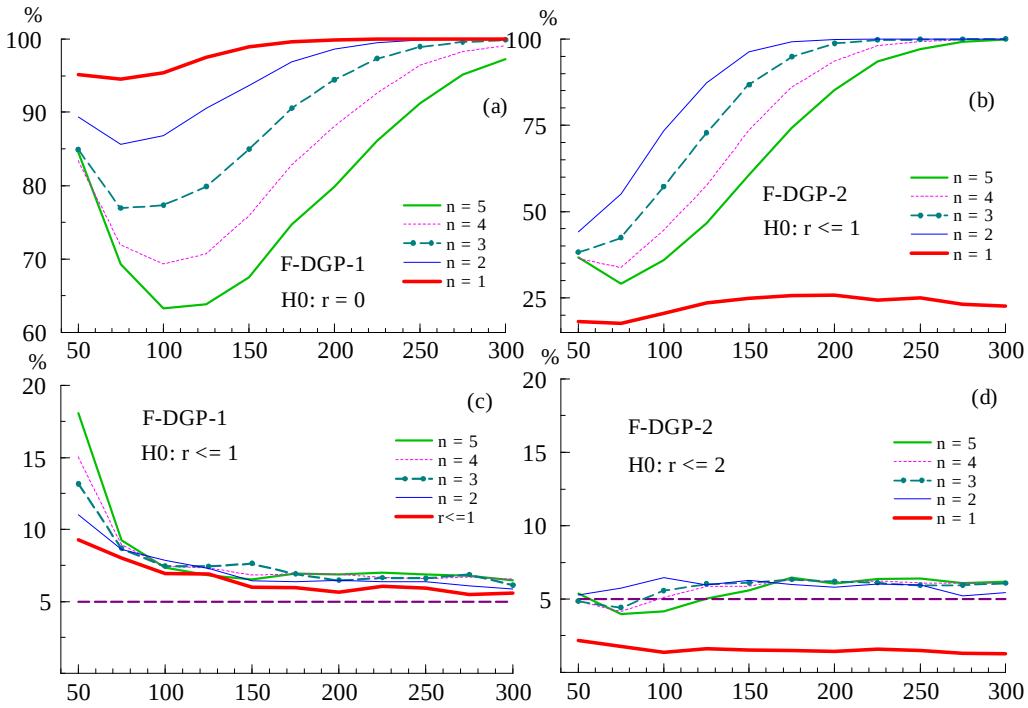

Figure 10: Recursive rejection rates based on a range of models

manner, indicating the possibility that they share a single stochastic trend over the sample period.

In order to test the hypothesis of $r = 5$, we first need to determine which model we should choose among a class of various models conceivable for the analysis of the interest rate data. We opt for the CNR model (6), instead of the LNR model and standard models, on the grounds that interest rate series are in general considered to be free of a linear trend but can be influenced by domestic and international business cycles.[6] Next, given the choice of a CNR model, we need to determine the frequencies of trigonometric function, since the theory implies the critical values of the test depend on the frequencies. To this end, we adapt a sequential procedure to select the frequencies of trigonometric function proposed by Perron, Shintani and Yabu (2021). Since their method is valid for both $I(0)$ and $I(1)$ cases, it is appropriate not only as the preliminary

[6] Johansen (Ch.12, 1996) applied an idea of Pantula (1989) to a set of procedures for selection between unrestricted and restricted deterministic terms in standard CVAR models. Since both CNR and LNR models have Fourier-type trends restricted in the cointegrating space, we assign importance here to economic grounds and choose a CNR model over a LNR model; see Doornik, Hendry and Nielsen (1998) for a detailed discussion on this type of issue.

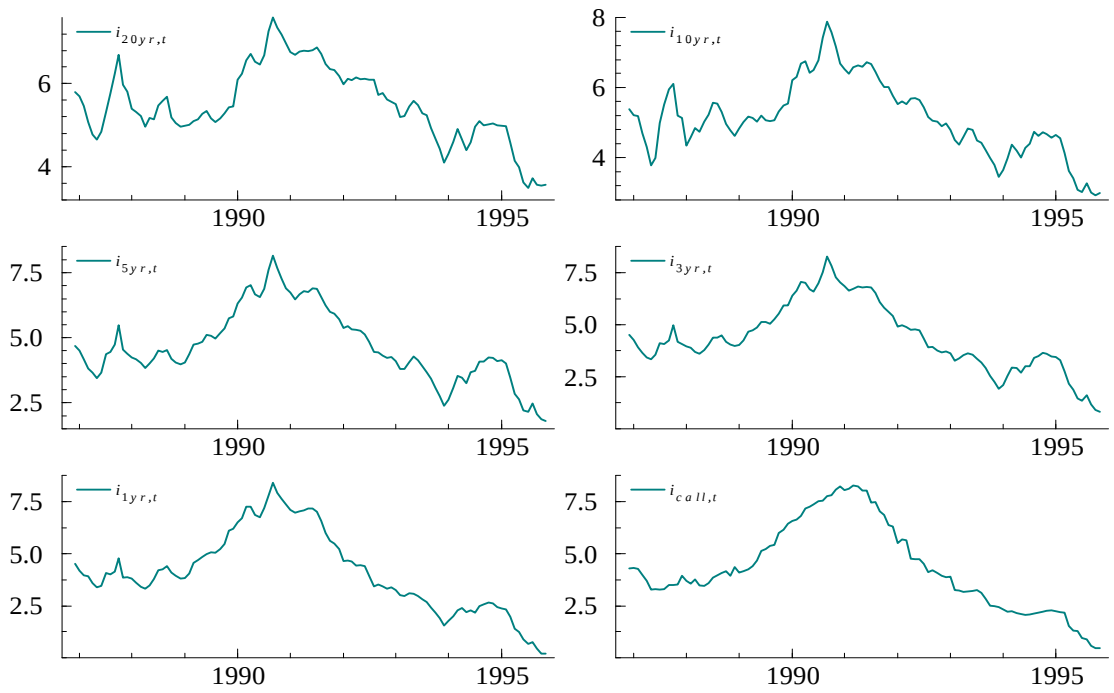

Figure 11: An overview of monthly data for Japan's bond yields

analysis for the unit root tests with Fourier-type nonlinear deterministic trends but also for our cointegrating rank test.

Their procedure is based on what they call sup-$W(\ell+1)$ test and $Mean$-$W(\ell+1)$ test for the coefficients on trigonometric functions. Let $n_{\max}$ be the maximum frequency allowed. Their sequential method of estimating the number of the frequencies proceeds as follows. First, starting with $\ell = 0$, use the sup-$W(\ell+1)$ or $Mean$-$W(\ell+1)$ test. If the null hypothesis of zero coefficient is not rejected, conclude for a model with ℓ frequencies estimated as $(\widehat{m}_1, \ldots, \widehat{m}_\ell) = \arg\min SSR$ where SSR is sum of the squared residuals and $\widehat{m}_j$'s are estimated frequencies with total of ℓ frequencies. Second, if the null is rejected, update ℓ to $\ell+1$ and repeat until a non-rejection occurs or the maximal allowed value $\ell = n_{\max} - 1$ is attained. This specific-to-general procedure will result in a consistent model selection if the size of the test converges to 0 slowly enough for the tests to be consistent. We apply this proposed procedure to each of JGB yield series and call rate. The significant level is set at 10 percent, $n_{\max}$ is set at 5, and the lag

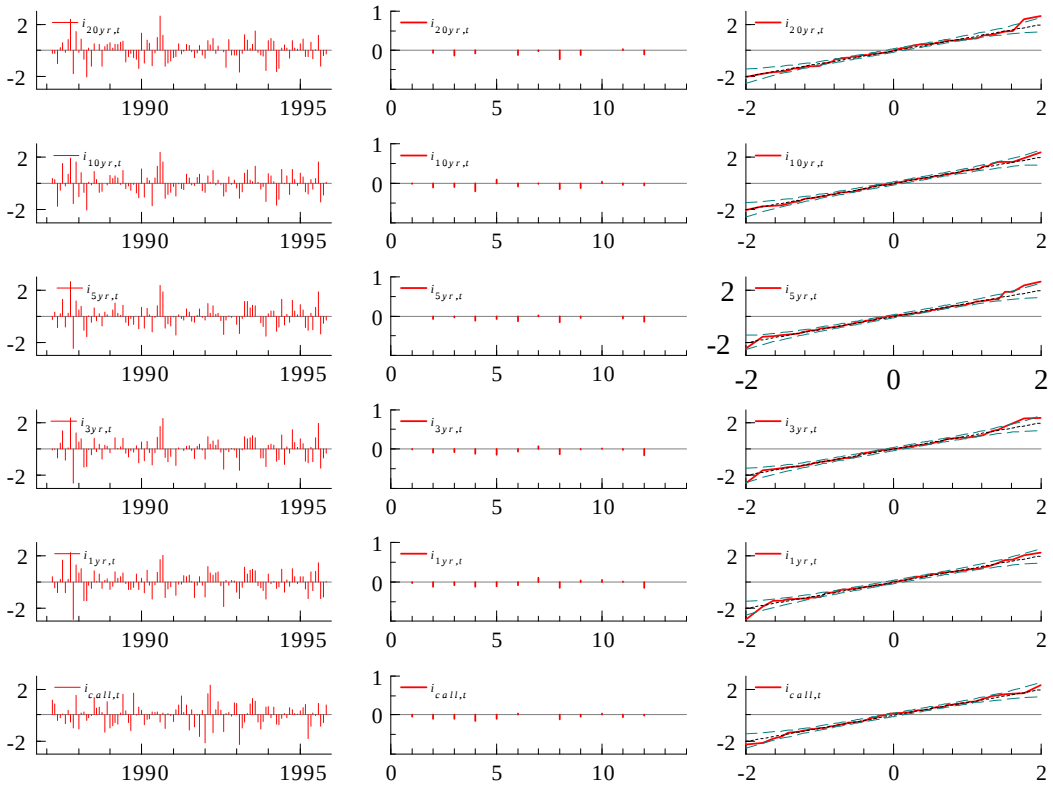

Figure 12: Diagnostic tests of the residuals from the non-linear model
First column: Scaled residuals for $i_{20yr,t}, ..., i_{call,t}$; Second column: Residual autocorrelation functions for $i_{20yr,t}, ..., i_{call,t}$; Third column: Residual quantile-quantile plots for $i_{20yr,t}, ..., i_{call,t}$.

length is selected by AIC. As a result, the largest estimated frequency was 5 with the call rate, and thus we employ $n = 5$ in subsequent analysis.

Given the selection of $n = 5$, we are in a position to study the details of an empirical CNR model. An initial unrestricted VAR(k) model, which provides a basis for the CNR model, should be formulated as a general model with sufficiently large orders for lag length k, so that the model can be subject to subsequent model reduction. The starting model should also be able to represent the data sufficiently well to justify subsequent likelihood-based inference for the cointegrating rank. See Johansen (1996, Ch.2) for further details of several checkpoints for the initial model in cointegrated VAR analysis. We selected $k = 3$ for the starting model in our study. The F test statistic for the hypothesis of model reduction from $k = 3$ to $k = 2$ is 2.48, which

rejects the hypothesis at the 5% level, based on standard inference. Furthermore, a diagnostic analysis of residuals from the VAR model was performed and the results are displayed in Figure 12. The first column in the figure records scalded residuals, while the second column and the third column provide residual autocorrelation plots and residual quantile-quantile plots (against normal distributions), respectively. We are able to argue that there is no strong evidence indicating outliers, serial correlation and non-normality in the residual series. This graphic diagnostic analysis indicates that the initial VAR(3) model has succeeded in representing the data over the effective sample period 1987.3 - 1995.11. The validity of this model allows us to proceed to the likelihood-based inference for the underlying cointegrating rank.

Panel (1) CNR model

	$r=0$	$r\leq 1$	$r\leq 2$	$r\leq 3$	$r\leq 4$	$r\leq 5$
$-2\log LR\,(r\|p=6;n=5)$	431.10	318.74	232.10	149.81	91.17	36.91
	$[0.00]^{**}$	$[0.00]^{**}$	$[0.00]^{**}$	$[0.01]^{**}$	$[0.04]^{*}$	[0.23]

Panel (2) Linear model

	$r=0$	$r\leq 1$	$r\leq 2$	$r\leq 3$	$r\leq 4$	$r\leq 5$
$-2\log LR\,(r)$	131.80	75.58	45.46	23.70	9.41	2.07
	$[0.00]^{**}$	[0.06]	[0.24]	[0.49]	[0.70]	[0.76]

Note. Figures in square brackets in Panel (1) are p-values calculated from response surface regressions using the gamma-approximation method, while those in Panel (2) are p-values taken from *PcGive* outputs (see Doornik, 1998).

Table 3: Log likelihood ratio tests for the cointegrating rank

Panel (1) in Table 3 presents a sequence of log-likelihood ratio test statistics for the choice of cointegrating rank in the CNR model. Judging from p-values derived from the gamma-approximation method (with 10,000 Monte Carlo replications) by using the *Ox* code in the supplementary material, we argue that the null hypotheses of $r=0,\ r\leq 1,\ ...,\ r\leq 4$ are rejected at the 5% level while the hypothesis of $r\leq 5$ is not rejected, so that $r=5$ or $p-r=1$ is supported by the Japanese data. For comarison, we calculated test statistics from a standard (constant restricted) model (2) for $k=3$, the parameters of which were estimated from the same data by using

reduced rank regression. The results are reported in Panel (2) in Table 3 and in favor of the selection of $r = 1$, that is, $p - r = 5$, which is not compatible with the term structure of interest rates. This result also appears to be puzzling in terms of the observation that all the interest rate series exhibit fairly similar behavior in Figure 11, presumably driven by the underlying single common stochastic trend. Furthermore, the difference in the test results in Table 3 is viewed as evidence consistent with the conceivable problem of finding spurious stochastic trends, as discussed at the end of Section 6, when applying the standard models to nonstationary data that are subject to smooth nonlinear deterministic trends.

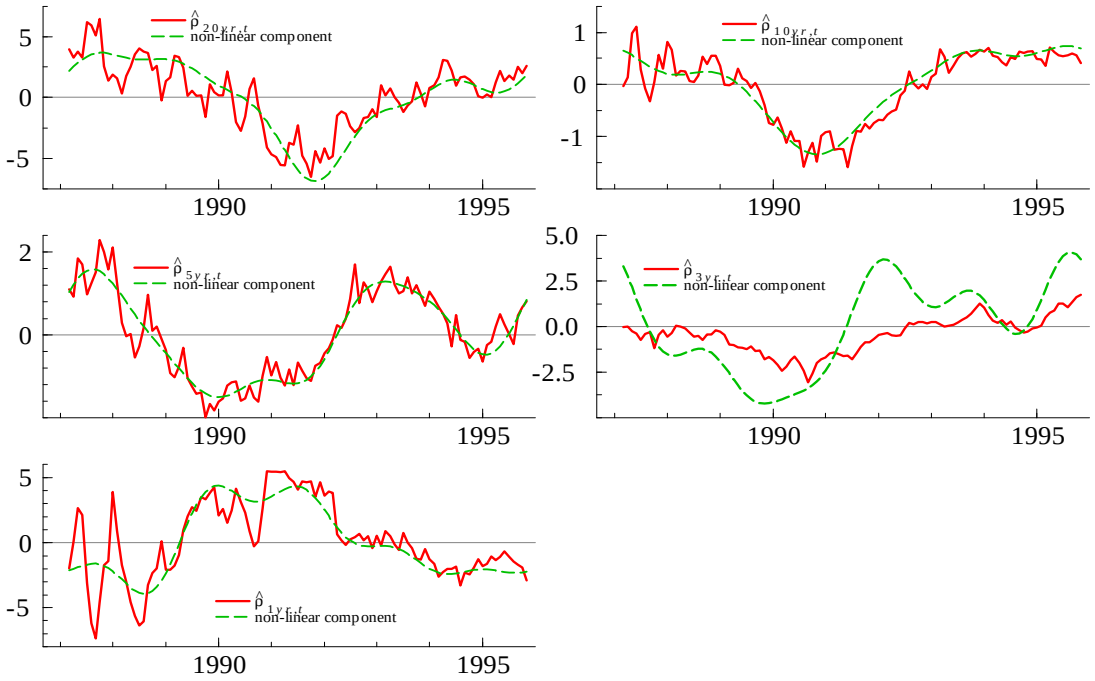

Figure 13: Implied risk premium terms

In addition, if we consider that the combination of nonlinear deterministic trends and stationary components in the cointegrating relations represents the underlying risk premium, it is possible to calculate a vector sequence of risk premiums $\hat{\rho}_t$ implied in

the bond yields as

$$\begin{aligned}\hat{\rho}_t &= \hat{\delta}' F_{t,T} + \hat{\upsilon}_t \\ &= -\hat{\beta}' X_{t-1} - \hat{\gamma}',\end{aligned}$$

for $\hat{\upsilon}_t = -\hat{\beta}' X_{t-1} - \hat{\delta}' F_{t,T} - \hat{\gamma}'$, in which $\hat{\beta}$, $\hat{\delta}$ and $\hat{\gamma}$ denote the maximum likelihood estimators of β, δ and γ, respectively. Figure 13 shows a set of the implied risk premium terms (thick lines)

$$\hat{\rho}_t = \left(\hat{\rho}_{20yr,t}, \hat{\rho}_{10yr,t}, \hat{\rho}_{5yr,t}, \hat{\rho}_{3yr,t}, \hat{\rho}_{1yr,t}\right)', \tag{18}$$

along with each of their nonlinear components (dotted lines), that is, each of $\hat{\delta}' F_{t,T}$ corresponding to a constituent of $\hat{\rho}_t$ given in (18). Note that these values were calculated from the CNR model estimated under the restriction of $r = 5$, with $\hat{\beta}$ normalized with respect to each of the five treasury bill yields. According to Figure 13, most of the implied premium terms (thick lines) appear to exhibit cyclical movements, reflecting the underlying deterministic trends (dotted lines) attributable to domestic and international business cycles. Overall, the results obtained in this section could justify us in arguing the advantage of the proposed nonlinear model over the standard model, in that the nonlinear model has selected multiple cointegrating relationships rather than a single cointegrating combination, a finding in line with the theory of the term structure of interest rates.

8 Summary and conclusion

Both nonlinear and nonstationary trending features observed in real-life time series data need to be reflected in a detailed econometric study providing a basis for economic policy analysis and forecasting. In view of this objective, this study has developed a methodology for testing the cointegrating rank in the VAR systems in the presence of smooth nonlinear deterministic trends, which are represented by a set of trigonometric functions. The asymptotic distributions of the proposed rank test statistics have been approximated by using simulation, and the statistics' limit quantiles have

been tabulated as a set of statistical tables for inference purposes. The *Ox* for the purpose of calculating approximate p-values of observed rank test statistics is available for use in empirical applications. Lastly, both the Monte Carlo and empirical studies have demonstrated the usefulness of the suggested methodology in a practical context. This paper has paved the way for various empirical studies aimed at simultaneously exploring the nonlinear and nonstationary characteristics of multivariate time series data.

Acknowledgements:

We thank Masato Kagihara, Kozo Ueda, Tomoyoshi Yabu, participants at the RCAST Macroeconomic Analysis Workshop (University of Tokyo), the 27th Kansai Econometrics Workshop (Hitotsubashi University) and the 5th International Conference on Econometrics and Statistics (Ryukoku University) for helpful comments and suggestions. We gratefully acknowledge financial support from JSPS KAKENHI 17H0251020, 18K01600 and 20H01482.

References

[1] Anderson, T. W. (1971). *The Statistical Analysis of Time Series*. Wiley, New York, New York.

[2] Astill, S., Harvey, D. I., Leybourne, S. J. and Taylor, A. M. R. (2015). Robust and powerful tests for nonlinear deterministic components, *Oxford Bulletin of Economics and Statistics*, 77, 780–799.

[3] Becker, R., Enders, W. and Lee, J. (2006). A stationarity test in the presence of an unknown number of smooth breaks. *Journal of Time Series Analysis*, 27, 381–409.

[4] Carrion-i-Silvestre, J. L., Kim, D., and Perron, P. (2009). GLS-based unit root tests with multiple structural breaks under both under the null and the alternative hypotheses, *Econometric Theory*, 25(6) 1754-1792.

[5] Cox, D. R. and Hinkley, D. V. (1979). *Theoretical Statistics*. CRC Press.

[6] Doornik, J. A. (1998). Approximations to the asymptotic distributions of cointegration tests. *Journal of Economic Surveys*, 12, 573–593.

[7] Doornik, J.A. (2003) Asymptotic tables for cointegration tests based on the gamma distribution approximation. https://www.doornik.com/research/coigamma_tables.pdf

[8] Doornik, J.A. (2013). *An Object-Oriented Matrix Programming Language - Ox7*. Timberlake Consultants Ltd.

[9] Doornik, J. A. and Hendry, D. F. (2013). *Modelling Dynamic - PcGive 14*: Volume 2. Timberlake Consultants Ltd.

[10] Doornik, J. A., Hendry, D. F. and Nielsen, B. (1998). Inference in cointegrating models: UK M1 revisited. *Journal of Economic Surveys*, 12, 533–572.

[11] Kim D., and Perron, P. (2009). Unit root tests allowing for a break in the trend function under both the null and alternative hypotheses, *Journal of Econometrics*, 148(1) 1-13.

[12] Enders, W. and Holt, M. T. (2012). Sharp breaks or smooth shifts? An investigation of the evolution of primary commodity prices. *American Journal of Agricultural Economics*, 94, 659–673.

[13] Enders, W. and Lee, J. (2012a). A unit root test using a Fourier series to approximate smooth breaks. *Oxford Bulletin of Economics and Statistics*, 74, 574–599.

[14] Enders, W. and Lee, J. (2012b). The flexible Fourier form and Dickey–Fuller type unit root tests. *Economics Letters*, 117, 196–199.

[15] Gallant, A. R. (1981). On the bias in flexible functional forms and an essentially unbiased form: the Fourier flexible form, *Journal of Econometrics*, 15, 211–245.

[16] Gallant, A. R. and Souza, G. (1991). On the asymptotic normality of Fourier flexible form estimates, *Journal of Econometrics*, 50, 329–353.

[17] Guo, Z.-F., and Shintani, M. (2013) Consistent co-trending rank selection when both stochastic and nonlinear deterministic trends are present. *Econometrics Journal*, 16(3), 473–484.

[18] Hall, A. D., Anderson, H. M. and Granger, C. W. J. (1992). A cointegration analysis of treasury bill yields. *Review of Economics and Statistics*, 74, 116–126.

[19] Hansen, P. R. (2005). Granger's representation theorem: A closed-form expression for I(1) processes. *Econometrics Journal*, 8, 23–38.

[20] Harvey, D. I., Leybourne, S. J. and Xiao, L. (2010). Testing for nonlinear deterministic components when the order of integration is unknown, *Journal of Time Series Analysis*, 31, 379–391.

[21] Hendry, D. F. (2000). On detectable and non-detectable structural change. *Structural Change and Economic Dynamics*, 11, 45–65.

[22] Hendry, D. F. (2006). Robustifying forecasts from equilibrium-correction systems. *Journal of Econometrics*, 135, 399–426.

[23] Hendry, D. F. and Doornik, J.A. (2014). *Empirical Model Discovery and Theory Evaluation: Automatic Selection Methods in Econometrics*. MIT Press.

[24] Hungnes, H. (2010). Identifying structural breaks in cointegrated vector autoregressive models. *Oxford Bulletin of Economics and Statistics*, 72, 551–565.

[25] Hunter, J., Burke, S.P. and Canepa, A. (2017). *Multivariate Modelling of nonstationary Economic Time Series*. Palgrave Macmillan.

[26] Inoue, A. (1999). Tests of cointegrating rank with a trend-break. *Journal of Econometrics*, 90, 215–237.

[27] Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12, 231–254.

[28] Johansen, S. R. (1994). The role of the constant and linear terms in cointegration analysis of nonstationary variables. *Econometric Reviews*, 13, 205–229.

[29] Johansen, S. (1996). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, 2nd Printing. Oxford University Press.

[30] Johansen, S., Mosconi, R. and Nielsen, B. (2000). Cointegration analysis in the presence of structural breaks in the deterministic trend. *Econometrics Journal*, 3, 216–249.

[31] Kurita, T. and Nielsen, B. (2019). Partial cointegrated vector autoregressive models with structural breaks in deterministic terms. *Econometrics*, 7, 42, in special issue *Celebrated Econometricians: Katarina Juselius and Søren Johansen*.

[32] Kurita, T., Nielsen, H. B., and Rahbek, A. (2011). An I(2) cointegration model with piecewise linear trends. *Econometrics Journal*, 14, 131–155.

[33] Leybourne, S. J. and Newbold, P. (2003). Spurious rejections by cointegration tests induced by structural breaks. *Applied Economics*, 35, 1117–1121.

[34] Lütkepohl, H., Saikkonen, P. and Trenkler, C. (2004). Testing for the cointegrating rank of a VAR process with level shift at unknown time. *Econometrica*, 72, 647–662.

[35] Nielsen, B. (1997). Bartlett correction of the unit root test in autoregressive models. *Biometrika*, 84, 500–504.

[36] Nielsen, B. and Rahbek, A. (2000). Similarity issues in cointegration analysis. *Oxford Bulletin of Economics and Statistics*, 62, 5–22.

[37] Pantula, S. G. (1989). Testing for unit roots in time series data. *Econometric Theory*, 5, 256-271.

[38] Pascalau, R. (2010). Unit root tests with smooth breaks: an application to the Nelson–Plosser data set. *Applied Economics Letters*, 17, 565–570.

[39] Pascalau, R., Lee, J., Nazlioglu, S. and Lu Y. (2022). Johansen-type cointegration tests with a Fourier function. *Journal of Time Series Analysis*, 43, 828-852.

[40] Perron, P., Shintani, M. and Yabu, T. (2017). Testing for flexible nonlinear trends with an integrated or stationary noise component. *Oxford Bulletin of Economics and Statistics*, 79, 822–850.

[41] Perron, P., Shintani, M. and Yabu, T. (2021). Trigonometric trend regressions of unknown frequencies with stationary or integrated noise. Working Paper, University of Tokyo.

[42] Phillips, P. C. and Solo, V. (1992). Asymptotics for linear processes. *Annals of Statistics*, 971–1001.

[43] Rippati, A. and Saikkonen, P. (2001) Vector autoregressive processes with nonlinear time trends in cointegrating relations. *Macroeconomic Dynamics*, 5, 577–597.

[44] Rodrigues, P. M. and Robert Taylor, A. M. (2012). The Flexible Fourier Form and Local Generalised Least Squares De-trended Unit Root Tests. *Oxford Bulletin of Economics and Statistics*, 74, 736–759.

[45] Saikkonen, P. (2001a). Consistent estimation in cointegrated vector autoregressive models with nonlinear time trends in cointegrating relations. *Econometric Theory*, 17, 296–326.

[46] Saikkonen, P. (2001b). Statistical inference in cointegrated vector autoregressive models with nonlinear time trends in cointegrating relations. *Econometric Theory*, 17, 327–356.

A Proofs

A.1 Proof of Proposition 3.2

We follow the proof of Theorem 1 in Hansen (2005) in order to prove Proposition 3.2. First, define $X_t^* = \left(X_t', X_{t-1}', \ldots, X_{t-k+1}'\right)'$ so that we can transform (6) into

$$\Delta X_t^* = \alpha^* \beta^{*\prime} X_{t-1}^* + \Phi_t^* + \varepsilon_t^* \quad \text{for} \quad t = 1, ..., T, \tag{19}$$

where

$$\alpha^* = \begin{pmatrix} \alpha & \Gamma_1 & \cdots & \Gamma_{k-1} \\ 0 & I_p & & 0 \\ \vdots & & \ddots & \\ 0 & & & I_p \end{pmatrix}, \; \beta^* = \begin{pmatrix} \beta & I_p & 0 & \cdots & 0 \\ 0 & -I_p & I_p & & \\ \vdots & & & \ddots & \\ & & & & I_p \\ 0 & \cdots & 0 & & -I_p \end{pmatrix},$$

$$\varepsilon_t^* = (\varepsilon_t', 0, \ldots, 0)' \quad \text{and} \quad \Phi_t^* = ((\alpha\delta' F_{t,T} + \alpha\gamma')', 0, \ldots, 0)'.$$

Given the parameters α^* and β^*, it is also found that

$$\alpha_\perp^* = (\alpha_\perp', -\alpha_\perp'\Gamma_1, \cdots, -\alpha_\perp'\Gamma_{k-1})' \quad \text{and} \quad \beta_\perp^* = (\beta_\perp', \ldots, \beta_\perp')',$$

so that the inverse of the full rank matrix $(\beta^*, \alpha_\perp^*)'$ is, with the notational convention $\Gamma_i^s = \Gamma_i + \cdots + \Gamma_{k-1}$,

$$\begin{pmatrix} \beta^{*\prime} \\ \alpha_\perp^{*\prime} \end{pmatrix}^{-1} = \begin{pmatrix} \xi & -C\Gamma_1^s & & \cdots & & -C\Gamma_{k-1}^s & C\bar{\alpha}_\perp \\ \xi & -C\Gamma_1^s - I_p & \ddots & & & -C\Gamma_{k-1}^s & C\bar{\alpha}_\perp \\ \xi & -C\Gamma_1^s - I_p & \ddots & \ddots & & -C\Gamma_{k-1}^s & C\bar{\alpha}_\perp \\ \vdots & \vdots & & \ddots & \ddots & \vdots & \vdots \\ \xi & -C\Gamma_1^s - I_p & & & \ddots & -C\Gamma_{k-1}^s & C\bar{\alpha}_\perp \\ \xi & -C\Gamma_1^s - I_p & & \cdots & & -C\Gamma_{k-1}^s - I & C\bar{\alpha}_\perp \end{pmatrix}.$$

Next, premultiplying (19) with $\beta^{*\prime}$ and $\alpha_\perp^{*\prime}$ leads to

$$\beta^{*\prime} X_t^* = \sum\nolimits_{i=0}^{\infty} (I + \beta^{*\prime}\alpha^*)\, \beta^{*\prime} \left(\varepsilon_{t-i}^* + \Phi_{t-i}^*\right) \text{ and } \alpha_\perp^{*\prime} X_t^* = \alpha_\perp^{*\prime} X_0^* + \sum\nolimits_{i=1}^{t} \alpha_\perp^{*\prime} (\varepsilon_i^* + \Phi_i^*),$$

and it is thus possible to express X_t as

$$\begin{aligned} X_t &= (I, 0, \cdots, 0) \begin{pmatrix} \beta^{*\prime} \\ \alpha_\perp^{*\prime} \end{pmatrix}^{-1} \begin{pmatrix} \sum_{i=0}^{\infty} (I + \beta^{*\prime}\alpha^*)\, \beta^{*\prime} (\varepsilon_t^* + \Phi_t^*) \\ \alpha_\perp^{*\prime} X_0^* + \sum_{i=1}^{t} \alpha_\perp^{*\prime} (\varepsilon_i^* + \Phi_i^*) \end{pmatrix} \\ &= \left(\xi, -C\Gamma_1^s, \cdots, -C\Gamma_{k-1}^s, C\bar{\alpha}_\perp\right) \begin{pmatrix} \sum_{i=0}^{\infty} (I + \beta^{*\prime}\alpha^*)\, \beta^{*\prime} \left(\varepsilon_{t-i}^* + \Phi_{t-i}^*\right) \\ \alpha_\perp^{*\prime} X_0^* + \sum_{i=1}^{t} \alpha_\perp^{*\prime} (\varepsilon_i^* + \Phi_i^*) \end{pmatrix}. \end{aligned}$$

It then follows that

$$X_t = C\sum_{i=1}^{t}\varepsilon_i + C(L)\varepsilon_t + C(L)\alpha\delta' F_{t,T} + C(L)\alpha\gamma' + A, \tag{20}$$

where the coefficients of $C(z) = \sum_{i=0}^{\infty} C_i z^i$ are

$$C_i = \left(\xi, -C\Gamma_1^s, \cdots, -C\Gamma_{k-1}^s\right)\left(I + \beta^{*\prime}\alpha^*\right)^i \beta^{*\prime}\left(I, 0, \cdots, 0\right)', \tag{21}$$

and A consists of the initial values such that

$$A = C\bar{\alpha}_\perp \alpha_\perp^{*\prime} X_0^* = C\left(X_0 - \Gamma_1 X_{-1} - \cdots - \Gamma_{k-1} X_{-k+1}\right). \tag{22}$$

Lastly, recall Lemma 4.1 in Johansen (1996), which says the expansion of the polynomial $C(z)$ around $z = 1$ leads to

$$C(z) = C(1) + (1-z)\, C^*(z), \tag{23}$$

where $C^*(z)$ is also a polynomial, such that $C^*(1) = -\frac{dC(z)}{dz}\,|_{z=1}$. In addition, Lemma A.6 in Hansen (2005) shows $C(1) = -\xi\eta' - C\Psi C$, which is substituted into (23) to find

$$C(z) = -\xi\eta' - C\Psi C + (1-z)\, C^*(z). \tag{24}$$

This expression is then applied to $C(L)\alpha\delta' F_{t,T}$ and $C(L)\alpha\gamma'$ in (20), and noting

$$C\alpha = \beta_\perp(\alpha_\perp'\Gamma\beta_\perp)^{-1}\alpha_\perp'\alpha = 0 \quad \text{and} \quad \eta'\alpha = (\alpha'\alpha)^{-1}\alpha'\left(I_p - \Gamma C\right)\alpha = I_r, \tag{25}$$

and recalling that X_t is denoted by $X_{t,T}$ in Section 3, we have obtained the desired representation.

A.2 Proof of Proposition 3.3

Follow the first and second steps of the proof of Proposition 3.2 above by noting that Φ_t^* is redefined here as

$$\Phi_t^* = \left((\alpha\delta' F_{t,T} + \alpha\gamma' t + \mu)', 0, \ldots, 0\right)',$$

so that we arrive at

$$X_t = C\sum_{i=1}^{t}(\varepsilon_i + \mu) + C(L)\alpha\delta' F_{t,T} + C(L)\varepsilon_i + C(L)\mu + C(L)\alpha\gamma' t + A, \tag{26}$$

in which $C(L)$ and A are specified by (21) and (22), respectively. Apply the expansion (24) to $C(L)\alpha\delta' F_{t,T}$ and $C(L)\mu$ in (26) to find

$$\begin{aligned} X_t \;=\; & C\sum_{i=1}^{t}\varepsilon_i + C(L)\varepsilon_t - \xi\delta' F_{t,T} + (1-L)\, C^*(L)\alpha\delta' F_{t,T} \\ & - \left(\xi\eta' + C\Psi C\right)\mu + C\mu t + C(L)\alpha\gamma' t + A. \end{aligned} \tag{27}$$

Use Lemma 4.1 in Johansen (1996) twice to derive the expansion of the polynomial $C(z)$

$$C(z) = C(1) + (1-z)\, C^*(1) + (1-z)^2\, C^{**}(z), \tag{28}$$

where $C^{**}(z)$ is also a polynomial, such that $C^{**}(1) = \frac{1}{2}\frac{d^2C(z)}{dz^2}\,|_{z=1}$. According to Lemma A.6 in Hansen (2005), we have $C(1) = -\xi\eta' - C\Psi C$ and

$$C^*(1) = -\xi\eta' - C\Psi C - \xi\eta'\Gamma\xi\eta' + \xi\eta'\Psi C + C\Psi\xi\eta' + C\Psi C\Psi C + C\left(\sum_{i=1}^{k-1}\frac{(i+1)i}{2}\Gamma_i\right)C,$$

both of which are plugged in (28). The resultant expansion is then applied to $C(L)\alpha\gamma' t$ in (27) to find the stated representation as a consequence of (25).

B Tables for approximate limit quantiles

p-r	n	90%	95%	97.5%	99%	Mean	Var
8	1	215.32	222.28	228.39	235.77	192.90	295.22
8	2	268.05	275.96	282.82	290.72	242.91	375.90
8	3	321.21	329.61	337.26	345.96	293.33	460.70
8	4	374.39	383.76	391.68	401.24	344.06	546.96
8	5	427.82	437.64	446.32	456.55	395.07	637.83
7	1	174.61	180.95	186.53	193.44	154.54	239.04
7	2	220.83	227.88	234.03	241.92	197.95	310.73
7	3	267.49	275.45	282.33	290.44	241.83	387.45
7	4	313.83	322.43	329.86	338.80	285.85	463.97
7	5	360.16	369.35	377.52	387.06	330.09	540.59
6	1	137.94	143.58	148.89	154.96	120.12	187.99
6	2	177.63	183.99	189.77	196.59	157.05	249.99
6	3	217.37	224.55	230.76	238.17	194.29	314.19
6	4	257.08	264.82	271.64	279.67	231.62	380.54
6	5	296.54	304.89	312.21	320.98	269.15	445.16
5	1	105.52	110.64	115.08	120.46	89.84	143.00
5	2	138.56	144.37	149.51	155.31	120.20	196.05
5	3	171.38	177.87	183.62	190.44	150.75	250.61
5	4	204.31	211.44	217.89	225.56	181.45	308.67
5	5	237.12	244.96	251.88	259.95	212.28	364.95
4	1	76.78	81.19	85.07	89.90	63.50	102.27
4	2	103.09	108.22	112.83	118.40	87.33	145.31
4	3	129.38	135.03	140.10	145.75	111.28	189.19
4	4	155.48	161.64	167.20	174.07	135.44	234.74
4	5	181.56	188.25	194.36	201.30	159.55	282.14
3	1	52.28	56.06	59.38	63.39	41.25	69.30
3	2	71.88	76.30	80.22	84.92	58.64	101.30
3	3	91.49	96.46	100.83	106.24	76.15	136.47
3	4	111.01	116.55	121.57	127.57	93.72	173.72
3	5	130.39	136.70	141.97	148.50	111.27	211.64
2	1	31.63	34.52	37.34	40.84	23.10	40.81
2	2	44.99	48.68	52.02	55.89	34.23	65.36
2	3	58.11	62.44	66.39	71.07	45.40	92.53
2	4	71.25	76.14	80.42	85.54	56.61	122.21
2	5	84.36	89.75	94.61	100.53	67.88	155.42
1	1	15.12	17.37	19.52	22.19	9.25	19.00
1	2	22.31	25.17	27.79	30.85	14.48	34.20
1	3	29.49	32.71	35.69	39.16	19.71	53.05
1	4	36.42	40.03	43.40	47.48	24.94	74.89
1	5	43.52	47.51	51.20	55.57	30.19	100.66

Table B1: Approximate limit quantiles for the CNR model

Notes: (1) The number of observations is 2,000 and the number of Monte Carlo replications is 100,000.
(2) An Ox code calculating p-values by using response surface regressions is available in the supplementary material.

p-r	n	90%	95%	97.5%	99%	Mean	Var
8	1	236.63	243.78	249.96	257.15	213.36	317.30
8	2	293.43	301.34	308.09	316.69	267.26	403.42
8	3	350.54	359.03	366.68	375.66	321.64	493.66
8	4	407.84	417.25	425.50	434.99	376.45	585.37
8	5	465.40	475.57	484.10	495.01	431.56	682.38
7	1	194.03	200.65	206.47	213.04	172.88	260.91
7	2	244.00	251.37	257.85	265.43	220.13	336.12
7	3	294.55	302.73	310.09	318.35	267.96	417.41
7	4	345.14	353.89	361.65	371.02	316.00	500.60
7	5	395.73	405.14	413.42	423.13	364.28	583.02
6	1	155.08	161.03	166.49	172.88	136.42	207.36
6	2	198.77	205.37	211.30	218.63	177.05	275.02
6	3	242.40	249.80	256.23	263.71	218.15	344.27
6	4	285.96	294.03	301.15	309.54	259.40	415.36
6	5	329.35	338.30	345.93	355.12	300.86	484.17
5	1	120.45	125.85	130.73	136.41	104.03	159.87
5	2	157.22	163.35	168.69	175.15	138.00	216.99
5	3	193.76	200.58	206.68	213.84	172.25	276.28
5	4	230.51	237.83	244.62	252.55	206.69	337.42
5	5	267.14	275.19	282.38	290.42	241.30	396.44
4	1	89.62	94.25	98.41	103.50	75.45	117.00
4	2	119.42	124.78	129.61	135.18	102.68	163.87
4	3	149.09	154.96	160.14	166.05	130.14	209.93
4	4	178.83	185.38	191.10	197.79	157.84	258.31
4	5	208.31	215.19	221.47	228.63	185.50	304.67
3	1	62.66	66.66	70.26	74.67	50.78	81.48
3	2	85.45	89.99	94.06	99.09	71.28	115.65
3	3	108.06	113.23	117.83	123.29	91.94	151.22
3	4	130.56	136.19	141.39	147.19	112.67	187.58
3	5	152.97	159.01	164.42	170.70	133.41	222.20
2	1	39.34	42.63	45.47	49.22	30.06	49.23
2	2	55.00	58.75	62.10	66.17	43.73	72.93
2	3	70.54	74.75	78.81	83.47	57.53	97.44
2	4	85.88	90.57	94.84	100.01	71.36	121.64
2	5	101.20	106.35	110.87	116.09	85.28	145.76
1	1	19.46	21.92	24.13	26.99	13.11	22.59
1	2	27.85	30.62	33.24	36.45	19.99	34.46
1	3	35.93	39.09	41.91	45.51	26.90	46.35
1	4	43.91	47.39	50.60	54.59	33.79	58.71
1	5	51.81	55.59	59.11	63.26	40.75	70.67

Table B2: Approximate limit quantiles for the LNR model

Note: (1) The number of observations is 2,000 and the number of Monte Carlo replications is 100,000.
(2) An Ox code calculating p-values by using response surface regressions is available in the supplementary material.