THE RANK OF A SUB-MATRIX OF COINTEGRATION

EIJI KUROZUMI *

Graduate School of Economics, Hitotsubashi University, Naka 2-1, Kunitachi, Tokyo, 186-8601, Japan

March, 2003

Abstract

This paper proposes a test of the rank of the sub-matrix of β, where β is a cointegrating matrix. In addition, the sub-matrix of $\beta_\perp$, an orthogonal complement to β, is investigated. We show that information on the rank of the sub-matrix of β and/or $\beta_\perp$ is useful in several situations. We construct the test statistic by using the eigenvalues of the quadratic form of the sub-matrix. We show that the test statistic has a limiting chi-squared distribution when the data is non-trending, and we propose a conservative test when the data is trending. Finite sample simulations show that, although the simulation settings are limited, the proposed test works well.

*Tel.: 81-42-580-8485; fax: 81-42-580-8195

E-mail address: kurozumi@stat.hit-u.ac.jp

I owe special thanks to Pierre Perron and Taku Yamamoto. All errors are my responsibility. This research was supported by the Ministry of Education, Culture, Sports, Science and Technology under Grants-in-Aid No. 13730023 and No. 14203003.

1. Introduction

A vector autoregressive (VAR) process has often been used to model a multivariate economic time series and, following the seminal work of Engle and Granger (1987), a cointegrating relation has been incorporated into the VAR model. A typical n-dimensional VAR model of order m is

$$x_t = d + A_1 x_{t-1} + \cdots + A_m x_{t-m} + \varepsilon_t, \tag{1}$$

for $t = 1, \cdots, T$, where $\{\varepsilon_t\}$ is independently and identically distributed (i.i.d.) with mean zero and a positive definite matrix Σ and $I_n - A_1 z - \cdots - A_m z^m = 0$ has roots outside the unit circle or equal to 1. The model (1) can be written in the error correction (EC) format,

$$\triangle x_t = d + \alpha\beta' x_{t-1} + \sum_{j=1}^{m-1} \Gamma_j \triangle x_{t-j} + \varepsilon_t, \tag{2}$$

where α and β are $n \times r$ matrices with rank r, $\triangle = 1 - L$, and L denotes the lag operator. We assume $0 < r < n$ and then there are r cointegrating relations. The exact condition of the existence of cointegration is given by Johansen (1992). We also assume that the cointegrating rank r is known or estimated by some testing procedure, such as the maximum likelihood (ML) test proposed by Johansen (1988, 1991) or the Lagrange multiplier (LM) test by Lütkepohl and Saikkonen (2000) and Saikkonen and Lütkepohl (2000). Other testing procedures of the cointegrating rank are reviewed by Hubrich, Lütkepohl, and Saikkonen (2001) and papers therein.

In this paper, we investigate the tests of the rank of β_1, the sub-matrix of β, and the rank of $\beta_{\perp,1}$, the sub-matrix of $\beta_\perp$, where $\beta = [\beta_1', \beta_2']'$ and $\beta_\perp = [\beta_{\perp,1}', \beta_{\perp,2}]'$, with $\beta_\perp$ being an orthogonal complement to β. In practical analysis, we sometimes encounter cases where we need to know the rank of β_1 and/or $\beta_{\perp,1}$. For example, the cointegrating matrix is sometimes normalized as $\beta^* = \beta(a'\beta)^{-1}$, as proposed by Johansen (1988, 1991) and Paruolo (1997), where a is an $n \times r$ matrix with full column rank, and the prototype normalization is represented by $a = [I_r, 0]'$. However, there is no guarantee that $a'\beta$ has full rank. In such a situation, we would like to know whether the first r rows of β have full rank. Another example is the Granger non-causality test. As shown in Toda and Phillips (1993), when there is a cointegrating relationship, in general the Wald statistic of the Granger non-causality test from the last n_3 variables of x_t to the first n_1 variables has a non-standard limiting distribution, depending on nuisance parameters. However, if the last n_3 rows of β have full row rank, the Wald statistic is asymptotically χ^2 distributed. In Section 2, we will illustrate these situations where information of the rank of β_1 and/or $\beta_{\perp,1}$ is useful in practical application.

Tests of the rank of a matrix have been investigated in the literature and recent econometric developments can be seen in works by Cragg and Donald (1996, 1997) and Robin and Smith (2000), among others. Although these papers proposed tests of the rank of a matrix, they assume that the estimator of the matrix is $T^{1/2}$ consistent and has a limiting normal distribution with a non-stochastic variance matrix. However, the estimator of the cointegrating matrix is T (or $T^{3/2}$) consistent and has an asymptotic non-standard distribution. As a result, we cannot apply existing testing procedures to the cointegrating matrix.

The paper is organized as follows. Section 2 illustrates situations in which we need to know the rank of β_1 and/or $\beta_{\perp,1}$. It is these situations that motivated us to investigate the test of the rank. In Section 3, we propose tests of the rank of β_1 and $\beta_{\perp,1}$ for non-trending data. We will show that the proposed test statistics have limiting χ^2 distributions. Section 4 considers the case of trending data. In this case, the test statistics do not necessarily converge to χ^2 distributions. To overcome this situation, we propose tests that are conservative. Section 5 investigates the finite sample properties of the tests. Section 6 concludes the paper.

In regard to notation, we use $vec(A)$ to stack the *rows* of a matrix A into a column vector, $[x]$ to denote the largest integer $\leq x$, $\bar{a} = a(a'a)^{-1}$ for a full column rank matrix a. $\xrightarrow{p}$, $\xrightarrow{d}$, and $\Rightarrow$ signify convergence in probability, convergence in distribution, and weak convergence of the associated probability measures. We denote the rank of A by $rk(A)$ and the column space of A by $sp(A)$. We write integrals like $\int_0^1 X(s)dY(s)'$ simply as $\int XdY'$ to achieve notational economy, and all integrals are from 0 to 1 except where otherwise noted.

2. Examples of Situations where Information on the Rank is Useful

2.1. Identifying normalizations

From the cointegrating matrix, β, we know directions that make the I(1) vector process stationary, but, in general, β cannot be identified because any pair of αc and $\beta c'^{-1}$ for a non-singular matrix c is equivalent to the pair of α and β in the model (2). Johansen (1988, 1991) and Paruolo (1997) proposed the identifying normalization such that $\beta^* = \beta(a'\beta)^{-1}$ and $\hat{\beta}^* = \hat{\beta}(a'\hat{\beta})^{-1}$ for the ML estimator $\hat{\beta}$, where a is an $n \times r$ matrix with full column rank. This normalization is useful in practical analysis because the limiting distribution of the normalized estimator has been derived by these authors and we can use it for statistical inference. However, there is no guarantee that $a'\beta$ has full rank. For example, the typical normalization is represented by $a = [I_r, 0]'$ and then β is normalized as $\beta(a'\beta)^{-1} = \beta\beta_1^{-1}$ where β_1 is the first r rows of β. Although β has full column rank r, there is no guarantee that β_1 has full rank, and then, as discussed in Paruolo (1997), we have to carefully chose the normalizing matrix a. In this case, the test of the rank of β_1 is useful to confirm that β_1 has full rank. If the rank of β_1 is decided to be r by the statistical test, we will use the normalizing matrix a. Otherwise, we have to choose another normalizing matrix. Since we may encounter the same identifying problem for $\beta_\perp$, we will also consider the test of the rank of the sub-matrix of $\beta_\perp$.

2.2. The Granger non-causality test

To test for Granger non-causality, we may use either the levels VAR model (1) or the EC format (2). Suppose that we are interested in whether the last n_3 variables in x_t are Granger-caused by the first n_1 variables. We write $x_t = [x'_{1t}, x'_{2t}, x'_{3t}]'$ and partition α, β and A_i for $i = 1, \cdots, m$ conformably with x_t. First, we consider the test with the levels VAR model (1). Then, the null hypothesis of non-causality is formulated as

$$A_1^{31} = \cdots = A_m^{31} = 0,$$

where A_i^{31} is the $n_3 \times n_1$ lower-left sub-matrix of A_i for $i = 1, \cdots, m$. If we estimate the model (1) by the least squares method and construct the Wald statistic in a usual form, the test statistic is shown by Toda and Phillips (1993) to have a non-standard limiting distribution and to depend on nuisance parameters in general. However, Toda and Phillips also showed that, if $rk(\beta_1) = n_1$ where β_1 is the first n_1 rows of β, the Wald statistic converges in distribution to $\chi^2_{n_1 n_3 m}$. Then, if we pretest the rank of β_1 and find that it has full row rank n_1, we can use the Wald statistic in a standard form.

On the other hand, if we estimate the model (2) by the ML method, the null hypothesis of non-causality is formulated as

$$\Gamma_1^{31} = \cdots = \Gamma_{m-1}^{31} = 0 \quad \text{and} \quad \bar{A}^{31} = 0,$$

where Γ_i^{31} and $\bar{A}^{31}$ are the $n_3 \times n_1$ lower-left matrices of Γ_i and $\bar{A}$ for $i = 1, \cdots, m-1$ with $\bar{A} = \alpha\beta'$. As in the case of the levels VAR, the Wald statistic for non-causality has a non-standard limiting distribution in general but Toda and Phillips (1993, Theorem 3) showed that if $rk(\beta_1) = n_1$ or $rk(\alpha_3) = n_3$, then the Wald statistic has a limiting $\chi^2_{n_1 n_3 m}$ distribution. Consequently, information on $rk(\beta_1)$ is useful in this case. Note that the existing testing procedure may be available for the test of $rk(\alpha_3)$ because the ML estimator of α has a limiting normal distribution with a variance matrix being a Kronecker product structure. See, for example, Corollary 3.1. of Robin and Smith (2000).

Although other testing procedures for Granger non-causality are proposed in the literature, such as the fully modified (FM) method by Phillips (1995) and Phillips and Hansen (1990), and the lag-augmented (LA) method by Dolado and Lütkepohl (1996), Toda and Yamamoto (1995) and Kurozumi and Yamamoto (2000), they have some deficiencies. For example, Yamada and Toda (1997, 1998) showed that the Granger non-causality test based on the FM method suffers from a large size distortion, while the LA method estimates the model with an artificially augmented lag that causes loss of power because of inefficiency. On the other hand, Yamamoto (2002) showed that the finite sample properties of the standard Wald statistic are fairly good when β_1 has full row rank. As a result, before testing for Granger non-causality, we recommend testing the rank of β_1 and, if it has full row rank, the standard Wald statistic should be used.

2.3. The test for long-run non-causality

A test for long-run non-causality was proposed in Bruneau and Jondeau (1999) and was developed into the test for block long-run non-causality by Yamamoto and Kurozumi (2001, 2002). The long-run causality is defined by considering the h-step ahead forecast with

$h \to \infty$. Let us consider the companion form of the model (1).

$$X_t = FX_{t-1} + E_t,$$

where $X_t = [x'_t, x'_{t-1}, \cdots, x'_{t-m+1}]'$, $E_t = [\varepsilon'_t, 0', \cdots, 0']'$ and F is defined consistently with the expression (1). The h-step ahead prediction of x_{t+h}, given X_t, is expressed as $x_{t+h|t} = MF^h X_t$, where $M = [I_n, 0, \cdots, 0]$, and the coefficient matrix of the long-run prediction is defined as

$$\bar{B} \equiv \lim_{h\to\infty} (MF^h).$$

Let us partition x_t in the same way as in the previous section. The hypothesis of the long-run non-causality of x_{3t} to x_{1t} is given by

$$R_L B R_R = 0,$$

where $R_L = [I_{n_1}, 0]$ and $R_R = I_m \otimes R_R^*$ with $R_R^* = [0, I_{n_3}]'$. Yamamoto and Kurozumi (2001, 2002) derived the limiting distribution of the ML estimator of $\bar{B}$, $\hat{B}$, which is asymptotically normal. The natural way to test the above hypothesis seems to be to construct the Wald statistic. However, the asymptotic variance matrix of $\hat{B}$ is singular and consequently the Wald statistic might also have a singular variance matrix. If a variance matrix has full rank, the usual Wald statistic is used to test the hypothesis, while we may construct the test statistic using a generalized inverse of the matrix if it is singular and we know the rank of the variance matrix. The important point is that singularity of the variance matrix depends only on $rk(\beta_3)$ and $rk(\beta_{\perp,1})$. In other words, we can identify the rank of the variance matrix if we know the rank of β_3 and $\beta_{\perp,1}$. Then, the tests of $rk(\beta_3)$ and $rk(\beta_{\perp,1})$ play an important role in the long-run non-causality test.

3. Test of the Rank of the Sub-Matrix for Non-Trending Data

3.1. The model with $d = 0$

In this section we consider a test of rank for non-trending data with $d = 0$. The model considered in this section is

$$\triangle x_t = \alpha\beta' x_{t-1} + \sum_{j=1}^{m-1} \Gamma_j \triangle x_{t-j} + \varepsilon_t. \tag{3}$$

We estimate the model (3) by the ML method assuming that $\{\varepsilon_t\}$ is Gaussian, although asymptotic properties are preserved under more general assumptions. We denote the ML estimator with "$\hat{}$". For example, the ML estimator of β is denoted by $\hat{\beta}$. Using the result that $T^{-1/2}\sum_{t=1}^{[Tr]}\varepsilon_t \Rightarrow W(r)$ for $0 \leq r \leq 1$ by the functional central limit theorem, where $W(\cdot)$ is an n-dimensional Brownian motion with a variance matrix Σ, Johansen (1988, 1996) showed that

$$\tilde{\beta} = \hat{\beta}(\bar{\beta}'\hat{\beta})^{-1} \xrightarrow{d} \left(\int G_0 G_0' ds\right)^{-1} \int G_0 dV', \tag{4}$$

where $G_0(\cdot) = \bar{\beta}_\perp' CW(\cdot)$ with $C = \beta_\perp(\alpha_\perp'\Gamma\beta_\perp)^{-1}\alpha_\perp'$, $\Gamma = I_n - \sum_{i=1}^{m-1}\Gamma_i$, $V(\cdot) = (\alpha'\Sigma^{-1}\alpha)^{-1}$ $\alpha'\Sigma^{-1}W(\cdot)$ and $G_0(\cdot)$ and $V(\cdot)$ are independent. He also showed that $\tilde{\alpha} = \hat{\alpha}\hat{\beta}'\bar{\beta}$, $\hat{\Sigma}$ and $\hat{\Gamma}_i$ $(i = 1, \cdots, m-1)$ are consistent estimators of α, Σ and Γ_i, respectively.

Let us partition β as $\beta' = [\beta_1', \beta_2']$ where β_1 and β_2 are $n_1 \times r$ and $(n - n_1) \times r$ matrices, respectively. Similarly, we partition $\beta_\perp' = [\beta_{\perp,1}', \beta_{\perp,2}']$ conformably. Our interest lies in finding the rank of β_1 and thus, we consider the following testing problem.

$$H_0 : rk(\beta_1) = f \quad v.s. \quad H_1 : rk(\beta_1) > f. \tag{5}$$

Note that the rank of β_1 is at most $p \equiv \min(n_1, r)$.

To test the rank of β_1, we follow the same strategy as Robin and Smith (2000), who test the rank of a matrix and investigate its quadratic form. In our situation, we construct a quadratic form of β_1. The advantage of considering a quadratic form is that the eigenvalues are non-negative real values, even if those of β_1 are complex values. Then, the null hypothesis H_0 becomes equivalent to the existence of f positive real and $n_1 - f$ zero eigenvalues.

Let Ψ and Φ be $r \times r$ and $n_1 \times n_1$ possibly stochastic matrices that are symmetric and positive definite almost surely (a.s.). Since they are full rank matrices (a.s.), the rank of β_1 is equal to the rank of $\Phi^{-1}\beta_1\Psi\beta_1'$ (a.s.). Therefore, the test of the rank of β_1 is equivalent to that of $\Phi^{-1}\beta_1\Psi\beta_1'$, and then we consider the rank of the latter matrix. Note that, although this strategy is basically the same as that followed by Robin and Smith (2000), we cannot directly use their result because they assume that the estimated matrix is asymptotically normally distributed with a convergence rate $T^{1/2}$, while $\hat{\beta}_1$ is shown to be T consistent and the limiting distribution is mixed Gaussian.

For the test of the rank of β_1, we define $\Psi = \alpha' \Sigma^{-1} \alpha$ and

$$\Phi = \left[\beta_1, \beta_{\perp,1} (\beta_\perp' \beta_\perp)^{-1} \right] \begin{bmatrix} (\beta' \beta)^{-1} & 0 \\ 0 & (\int G_0 G_0' ds)^{-1} \end{bmatrix} \begin{bmatrix} \beta_1' \\ (\beta_\perp' \beta_\perp)^{-1} \beta_{\perp,1}' \end{bmatrix}. \tag{6}$$

Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{n_1}$ be the ordered eigenvalues of $\Phi^{-1} \beta_1 \Psi \beta_1'$, which are the solution of the determinant equation

$$|\beta_1 \Psi \beta_1' - \lambda \Phi| = 0. \tag{7}$$

Then, under H_0, $\lambda_1 \geq \cdots \geq \lambda_f > 0$ and $\lambda_{f+1} = \cdots = \lambda_{n_1} = 0$ (a.s.).

We construct a sample analogue of (7) using the LM estimator and investigate the limiting distributions of the eigenvalues. The sample analogue of (7) is given by

$$|\hat{\beta}_1 \hat{\Psi} \hat{\beta}_1' - \hat{\lambda} \hat{\Phi}| = 0, \tag{8}$$

where $\hat{\beta}_1$ is the first n_1 rows of $\hat{\beta}$, $\hat{\Psi} = \hat{\alpha}' \hat{\Sigma}^{-1} \hat{\alpha}$ and

$$\begin{aligned} \hat{\Phi} &= \left[\hat{\beta}_1, \hat{\beta}_{\perp,1} (\hat{\beta}_\perp' \hat{\beta}_\perp)^{-1} \right] \begin{bmatrix} (\hat{\beta}' \hat{\beta})^{-1} & 0 \\ 0 & \left(\frac{1}{T} \bar{\hat{\beta}}_\perp' S_{11} \bar{\hat{\beta}}_\perp \right)^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta}_1' \\ (\hat{\beta}_\perp' \hat{\beta}_\perp)^{-1} \hat{\beta}_{\perp,1}' \end{bmatrix} && (9) \\ &= \hat{\beta}_1 (\hat{\beta}' \hat{\beta})^{-1} \hat{\beta}_1' + \hat{\beta}_{\perp,1} \left(\frac{1}{T} \hat{\beta}_\perp' S_{11} \hat{\beta}_\perp \right)^{-1} \hat{\beta}_{\perp,1}', && (10) \end{aligned}$$

where $S_{11} = T^{-1} \sum_{t=1}^T R_{1t} R_{1t}'$, with R_{1t} being the regression residual of X_{t-1} on $\triangle x_{t-1}, \cdots, \triangle x_{t-m+1}$, and we denote the ordered eigenvalues of (9) as $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \cdots \geq \hat{\lambda}_{n_1}$. Note that when $n_1 > r$, the smallest $n_1 - r$ eigenvalues are obviously equal to 0, that is, $\hat{\lambda}_{r+1} = \cdots = \hat{\lambda}_{n_1} = 0$. We can easily see from the expressions (6) and (9) that Φ and $\hat{\Phi}$ are positive definite (a.s.), while the expression (10) is simpler and may be used to construct $\hat{\Phi}$ in practice.

To test the rank of β_1, we consider the following test statistic.

$$\mathcal{L}_T = T^2 \sum_{i=f+1}^{p} \hat{\lambda}_i = T^2 \sum_{i=f+1}^{n_1} \hat{\lambda}_i,$$

which rejects the null hypothesis when $\mathcal{L}_T$ takes large values. The second equality is established because $p = \min(n_1, r)$ and $\hat{\lambda}_{p+1} = \cdots = \hat{\lambda}_{n_1} = 0$ when $n_1 > r$.

Theorem 1 *Let $\hat{\Psi} = \hat{\alpha}' \hat{\Sigma}^{-1} \hat{\alpha}$ and $\hat{\Phi}$ be given by (10). If $f < p$, under H_0, $\mathcal{L}_T \xrightarrow{d} \chi^2_{(n_1 - f)(r - f)}$.*

Remark 1: Since the determinant equation (8) converges to (7) in distribution, the estimated ordered eigenvalues of (8) also converge in distribution to those of (7). Then, under the alternative, $\hat{\lambda}_{f+1} \xrightarrow{d} \lambda_{f+1} > 0$ (a.s.), so that $T^2\hat{\lambda}_{f+1}$ goes to infinity. Therefore, the test statistic $\mathcal{L}_T$ is consistent.

Next, we consider a test of the rank of the sub-matrix of $\beta_\perp$. The testing problem is

$$H_{0\perp}:\ rk(\beta_{\perp,1}) = g \quad v.s. \quad H_{1\perp}:\ rk(\beta_{\perp,1}) > g.$$

For the same reason as in the test of β_1, we investigate the rank of $\ddot{\Phi}^{-1}\beta_{\perp,1}\ddot{\Psi}\beta'_{\perp,1}$, where $\ddot{\Phi}$ and $\ddot{\Psi}$ are $(n-r)\times(n-r)$ and $n_1 \times n_1$ full rank matrices (a.s.). Similar to (7), we consider the following determinant equation.

$$|\beta_{\perp,1}\ddot{\Psi}\beta'_{\perp,1} - \mu\ddot{\Phi}| = 0, \tag{11}$$

where $\ddot{\Psi} = \int G_0 G_0' ds$ and

$$\ddot{\Phi} = [\beta_{\perp,1}, \beta_1(\beta'\beta)^{-1}] \begin{bmatrix} (\beta'_\perp\beta_\perp)^{-1} & 0 \\ 0 & (\alpha'\Sigma^{-1}\alpha)^{-1} \end{bmatrix} \begin{bmatrix} \beta'_{\perp,1} \\ (\beta'\beta)^{-1}\beta'_1 \end{bmatrix},$$

and the sample analogue of (11) is given by

$$|\hat{\beta}_{\perp,1}\hat{\ddot{\Psi}}\hat{\beta}'_{\perp,1} - \hat{\mu}\hat{\ddot{\Phi}}| = 0, \tag{12}$$

where $\hat{\ddot{\Psi}} = T^{-1}\bar{\hat{\beta}}'_\perp S_{11}\bar{\hat{\beta}}_\perp$ and

$$\begin{aligned} \hat{\ddot{\Phi}} &= \left[\hat{\beta}_{\perp,1}, \hat{\beta}_1(\hat{\beta}'\hat{\beta})^{-1}\right] \begin{bmatrix} (\hat{\beta}'_\perp\hat{\beta}_\perp)^{-1} & 0 \\ 0 & (\hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha})^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta}'_{\perp,1} \\ (\hat{\beta}'\hat{\beta})^{-1}\hat{\beta}'_1 \end{bmatrix} \\ &= \hat{\beta}_{\perp,1}(\hat{\beta}'_\perp\hat{\beta}_\perp)^{-1}\hat{\beta}'_{\perp,1} + \hat{\beta}_1(\hat{\beta}'\hat{\beta})^{-1}(\hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha})^{-1}(\hat{\beta}'\hat{\beta})^{-1}\hat{\beta}'_1. \end{aligned} \tag{13}$$

Let $\mu_1 \geq \cdots \geq \mu_{n_1}$ and $\hat{\mu}_1 \geq \cdots \geq \hat{\mu}_{n_1}$ be ordered eigenvalues of (11) and (12), respectively. Under the null hypothesis, the smallest $(n_1 - g)$ eigenvalues, $\mu_{g+1}, \cdots, \mu_{n_1}$, are all zeros (a.s.) and we then construct the following test statistic.

$$\mathcal{L}_{\perp T} = T^2 \sum_{i=g+1}^{q} \hat{\mu}_i = T^2 \sum_{i=g+1}^{n_1} \hat{\mu}_i,$$

which rejects the null hypothesis when it takes large values, where $q = \min(n_1, n-r)$. The second equality is established because $\hat{\mu}_{n-r+1} = \cdots \hat{\mu}_{n_1} = 0$ are obvious solutions when $n_1 > (n-r)$. The following theorem gives the limiting distribution of $\mathcal{L}_{\perp T}$.

Theorem 2 *Let* $\hat{\bar{\Psi}} = T^{-1}\hat{\bar{\beta}}_{\perp}' S_{11}\hat{\bar{\beta}}_{\perp}$ *and* $\hat{\bar{\Phi}}$ *be given by (13). If* $g < q$*, under* $H_{0\perp}$*,* $\mathcal{L}_{\perp T} \xrightarrow{d} \chi^2_{(n_1-g)(n-r-g)}$.

Note that the consistency of the test is shown in a similar way as in Remark 1.

Given the above two theorems, we can test the rank of β_1 and $\beta_{\perp,1}$. In addition, we may consider the procedure to decide the rank of the sub-matrix, as the cointegrating rank is selected sequentially using the test of the cointegrating rank. For example, to decide the rank of β_1, we firstly test the null of $f = 0$. If the null hypothesis is accepted, the rank of β_1 is decided to be zero. Otherwise, we then test the hypothesis of $f = 1$. We sequentially continue to test the rank of β_1 until the null hypothesis is accepted. When the null of $f = p - 1$ is rejected, we consider that β_1 has full rank. Similarly, the rank of $\beta_{\perp,1}$ can be decided by the same procedure.

3.2. The model with $d \neq 0$

In the previous section, we considered the model with $d = 0$ for non-trending data. However, in practice, we sometimes consider the model (2) with $d \neq 0$ but the level of data has no linear trend. In this case, the constant term can be expressed as $d = \alpha\rho_0$ where ρ_0 is a $r \times 1$ coefficient vector, so that the model (2) becomes

$$\triangle x_t = \alpha\beta^{+\prime} x_{t-1}^{+} + \sum_{j=1}^{m-1} \Gamma_j \triangle x_{t-j} + \varepsilon_t, \tag{14}$$

where $\beta^{+} = [\beta', \rho_0]'$ and $x_{t-1}^{+} = [x_{t-1}', 1]'$. See Johansen (1991, 1996). The maximum likelihood estimator of β^{+} can be obtained by the reduced rank regression of $\triangle x_t$ on x_{t-1}^{+} corrected for $\triangle x_{t-1}, \cdots, \triangle x_{t-m+1}$, and the estimator of the cointegrating matrix is the first n rows of $\hat{\beta}^{+}$.

To test the rank of the sub-matrix of β for the model (14), we construct the test statistic $\mathcal{L}_T$ with $\hat{\Phi}$ defined by

$$\hat{\Phi} = \hat{\beta}_1(\hat{\beta}'\hat{\beta})^{-1}\hat{\beta}_1' + \hat{\beta}_{\perp,1}(\hat{\beta}_{\perp}'\hat{\beta}_{\perp})^{-1}L'(\Upsilon_T' S_{11}^{+}\Upsilon_T)^{-1}L(\hat{\beta}_{\perp}'\hat{\beta}_{\perp})^{-1}\hat{\beta}_{\perp,1}', \tag{15}$$

where L and Υ_T are $(n-r+1) \times (n-r)$ and $(n+1) \times (n-r+1)$ matrices defined by

$$L = \begin{bmatrix} I_{n-r} \\ 0 \end{bmatrix}, \quad \Upsilon_T = \begin{bmatrix} T^{-1/2}\hat{\bar{\beta}}_{\perp} & 0 \\ 0 & 1 \end{bmatrix},$$

and $S_{11}^{+} = T^{-1}\sum_{t=1}^{T} R_{1t}^{+}R_{1t}^{+\prime}$, with R_{1t}^{+} being the regression residual of x_{t-1}^{+} on $\triangle x_{t-1}, \cdots, \triangle x_{t-m+1}$.

Theorem 3 *Consider the model (14) and let $\hat{\Psi} = \hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha}$ and $\hat{\Phi}$ be given by (15). If $f < p$, under H_0, $\mathcal{L}_T \xrightarrow{d} \chi^2_{(n_1-f)(r-f)}$.*

For the test of the sub-matrix of $\beta_\perp$, we construct the test statistic $\mathcal{L}_{\perp T}$ using $\hat{\ddot{\Psi}} = \{L'(\Upsilon_T' S_{11}^{+}\Upsilon_T)^{-1}L\}^{-1}$.

Theorem 4 *Consider the model (14) and let $\hat{\ddot{\Psi}} = \{L'(\Upsilon_T' S_{11}^{+}\Upsilon_T)^{-1}L\}^{-1}$ and $\hat{\ddot{\Phi}}$ be given by (13). If $g < q$, under $H_{0\perp}$, $\mathcal{L}_{\perp T} \xrightarrow{d} \chi^2_{(n_1-g)(n-r-g)}$.*

In practical analysis, we will obtain $\hat{\beta}$ by the reduced rank regression and we have to calculate $\hat{\beta}_\perp$ from $\hat{\beta}$. If $d = 0$, $\hat{\beta}_\perp$ can be easily obtained as explained in Johansen (1996, p.95). When $d = \alpha\rho_0$, one of the methods to calculate $\hat{\beta}_\perp$ is as follows: first we calculate the orthogonal projection matrix of $\hat{\beta}$, $M = I_n - \hat{\beta}(\hat{\beta}'\hat{\beta})^{-1}\hat{\beta}'$. Then, by the singular value decomposition, M is expressed as $M_l M_\lambda M_r'$ where M_l and M_r are $n \times (n-r)$ orthogonal matrices and M_λ is an $(n-r)\times(n-r)$ diagonal matrix with positive diagonal elements. Since $sp(M) = sp(M_l)$ and they are orthogonal to $\hat{\beta}$, we can use M_l as $\hat{\beta}_\perp$.

4. The Test of the Rank of the Sub-Matrix for Trending Data

When the data is trending, x_t can be expressed as the sum of the stochastic trend, the deterministic trend, and the I(0) component such that

$$x_t = C\sum_{i=1}^{t}\varepsilon_i + \tau t + C_1(L)(\varepsilon_t + d) + x_0^*, \tag{16}$$

where $\tau = Cd$, $C_1(L) = (C(L) - C(1))/(1-L)$ with $C(L)$ being a lag polynomial when $\triangle x_t$ is represented as the vector moving-average process like $\triangle x_t = C(L)(d + \varepsilon_t)$, and x_0^* is a stochastic component such that $\beta' x_0^* = 0$. See Johansen (1991, 1996) for more details. In this case, $\beta_\perp$ is decomposed to τ, the coefficient of a linear trend in (16), and γ, an $n\times(n-r-1)$ matrix that is orthogonal to τ. As shown in Chapter 13.2 of Johansen (1996), $\tilde{\beta}$ can be expressed as

$$\tilde{\beta} = \beta + \gamma(\gamma'\gamma)^{-1}U_{1T} + \frac{1}{T^{1/2}}\tau(\tau'\tau)^{-1}U_{2T}, \tag{17}$$

where

$$T\begin{bmatrix} U_{1T} \\ U_{2T} \end{bmatrix} \xrightarrow{d} \left(\int GG'ds\right)^{-1}\int GdV' = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}, \quad \text{say,}$$

where $G(r) = [G_1'(r), G_2'(r)]'$ with $G_1(r) = G_0(r) - \int G_0 ds$ and $G_2(r) = r - 1/2$. We denote $\Omega = \int GG'ds$ and partition it into 2×2 blocks conformably with $[U_1', U_2']'$. We express the (i,j) block element of $(\int GG'ds)^{-1}$ as Ω^{ij} for $i, j = 1$ and 2. In this section, we need the estimator of Ω^{11}, which is given by

$$\hat{\Omega}^{11} = T\hat{\gamma}' S_{11}^{-1}\hat{\gamma},$$

and S_{11} is defined in the same way as in the previous section, with R_{1t} being the regression residual of X_{t-1} on a constant and $\triangle x_{t-1}, \cdots, \triangle x_{t-m+1}$. Convergence of $\hat{\Omega}^{11}$ to Ω^{11} is proved in Lemma 2 (iii) in the appendix, while the consistency of other ML estimators, such as $\tilde{\alpha}$, $\hat{\Sigma}$, and $\hat{\Gamma}_i$, is shown by Johansen (1991, 1996).

Let us consider the testing problem (5). Under the null hypothesis, we can find the f linearly independent column vectors in β_1 and we define β_1^* as an $n_1 \times f$ matrix whose columns consist of those f vectors. We also define an $n_1 \times (n_1 - f)$ matrix δ^* whose columns span the space orthogonal to the columns of β_1^* so that $\delta^{*\prime}\beta_1^* = 0$. In the following, we show that the direction of δ^* is important in deciding the convergence rate of $\tilde{\beta}_1$ and it also affects the limiting property of the test statistic.

Since $\tilde{\beta}_1$ is the first n_1 rows of $\tilde{\beta}$, it is expressed from (17) as

$$\tilde{\beta}_1 = \beta_1 + \gamma_1(\gamma'\gamma)^{-1}U_{1T} + \frac{1}{T^{1/2}}\tau_1(\tau'\tau)^{-1}U_{2T}.$$

Suppose that an $n_1 \times 1$ vector τ_1^* exists that is orthogonal to γ_1 ($\tau_1^{*\prime}\gamma_1 = 0$) and belongs to the column space of δ^*. Here, note that, since the $n \times n$ matrix $[\beta, \gamma, \tau]$ has full rank, the first n_1 rows of this matrix, $[\beta_1, \gamma_1, \tau_1]$, must have full row rank, which implies that $a'[\beta_1, \gamma_1, \tau_1] \neq 0$ for any non-zero vector a. Then, because τ_1^* is orthogonal to both β_1 and γ_1 by the assumption, we have $\tau_1^{*\prime}[\beta_1, \gamma_1, \tau_1] = [0, 0, \tau_1^{*\prime}\tau_1] \neq 0$, so that $\tau_1^{*\prime}\tau_1 \neq 0$. This implies that

$$T^{3/2}\tau_1^{*\prime}\tilde{\beta}_1 = \tau_1^{*\prime}\tau_1(\tau'\tau)^{-1}(TU_{2T}) \xrightarrow{d} \tau_1^{*\prime}\tau_1(\tau'\tau)^{-1}U_2 = X_2', \quad \text{say,} \tag{18}$$

while for an $n_1 \times (n-r-1)$ matrix δ_0^* whose columns span the orthogonal complement to τ_1^* in $sp(\delta^*)$,

$$T\delta_0^{*\prime}\tilde{\beta}_1 = \delta_0^{*\prime}\gamma_1(\gamma'\gamma)^{-1}(TU_{1T}) \xrightarrow{d} \delta_0^{*\prime}\gamma_1(\gamma'\gamma)^{-1}U_1 = X_1', \quad \text{say.} \tag{19}$$

On the other hand, if there exists no vector in $sp(\delta^*)$ that is orthogonal to γ_1, we have

$$T\delta^{*\prime}\tilde{\beta}_1 = \delta^{*\prime}\gamma_1(\gamma'\gamma)^{-1}(TU_{1T}) + \delta^{*\prime}\tau_1(\tau'\tau)^{-1}(T^{1/2}U_{2T}) \xrightarrow{d} \delta^{*\prime}\gamma_1(\gamma'\gamma)^{-1}U_1 = X', \text{ say.} \tag{20}$$

Therefore, the convergence rate of $\tilde{\beta}_1$ depends on whether a vector τ_1^* orthogonal to γ_1 exists in $sp(\delta^*)$.

The existence of τ_1^* indicates that the column space of $[\beta_1, \gamma_1]$ does not include τ_1^* because $\tau_1^{*\prime}\beta_1 = 0$ and $\tau_1^{*\prime}\gamma_1 = 0$. We also note that the rank of $[\beta_1, \gamma_1]$ must be $n_1 - 1$ or n_1 because $[\beta_1, \gamma_1, \tau_1]$ has full rank n_1. Then, from another point of view, we can say that the rank of $[\beta_1, \gamma_1]$ is $n_1 - 1$ if a vector τ_1^* exists, while the non-existence of τ_1^* is equivalent to $rk([\beta_1, \gamma_1]) = n_1$. Thus, we have to consider the asymptotic property separately according to the two cases where the rank of $[\beta_1, \gamma_1]$ is n_1 and $n_1 - 1$.

In the following theorem, the test statistic is constructed from the eigenvalues of (8) using the same $\hat{\Psi}$ as in the previous section and either

$$\begin{aligned}\hat{\Phi} &= \left[\hat{\beta}_1, \hat{\gamma}_1(\hat{\gamma}'\hat{\gamma})^{-1}\right] \begin{bmatrix} (\hat{\beta}'\hat{\beta})^{-1} & 0 \\ 0 & \hat{\Omega}^{11} \end{bmatrix} \begin{bmatrix} \hat{\beta}_1' \\ (\hat{\gamma}'\hat{\gamma})^{-1}\hat{\gamma}_1' \end{bmatrix} \\ &= \hat{\beta}_1(\hat{\beta}'\hat{\beta})^{-1}\hat{\beta}_1' + \hat{\gamma}_1(\hat{\gamma}'\hat{\gamma})^{-1}(T\hat{\gamma}'S_{11}^{-1}\hat{\gamma})(\hat{\gamma}'\hat{\gamma})^{-1}\hat{\gamma}_1',\end{aligned} \tag{21}$$

or

$$\begin{aligned}\hat{\Phi} &= \left[\hat{\beta}_1, \hat{\gamma}_1(\hat{\gamma}'\hat{\gamma})^{-1}, \hat{\tau}_1\right] \begin{bmatrix} (\hat{\beta}'\hat{\beta})^{-1} & 0 & 0 \\ 0 & \hat{\Omega}^{11} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1' \\ (\hat{\gamma}'\hat{\gamma})^{-1}\hat{\gamma}_1' \\ \hat{\tau}_1' \end{bmatrix} \\ &= \hat{\beta}_1(\hat{\beta}'\hat{\beta})^{-1}\hat{\beta}_1' + \hat{\gamma}_1(\hat{\gamma}'\hat{\gamma})^{-1}(T\hat{\gamma}'S_{11}^{-1}\hat{\gamma})(\hat{\gamma}'\hat{\gamma})^{-1}\hat{\gamma}_1' + \hat{\tau}_1\hat{\tau}_1'.\end{aligned} \tag{22}$$

Theorem 5 *(i-a) Let $\hat{\Psi} = \hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha}$ and $\hat{\Phi}$ be given by (21). If $rk([\beta_1, \gamma_1]) = n_1$ and $f < p$, under H_0, $\mathcal{L}_T \xrightarrow{d} \chi^2_{(n_1-f)(r-f)}$.*
(i-b) Let $\hat{\Psi} = \hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha}$ and $\hat{\Phi}$ be given by (22). If $rk([\beta_1, \gamma_1]) = n_1$ and $f < p$, under H_0, $\mathcal{L}_T$ converges in distribution to a random variable that is bounded above by $\chi^2_{(n_1-f)(r-f)}$.
(ii) Let $\hat{\Psi} = \hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha}$ and $\hat{\Phi}$ be given by (22). If $rk([\beta_1, \gamma_1]) = n_1 - 1$ and $f < p$, under H_0, $\mathcal{L}_T \xrightarrow{d} \chi^2_{(n_1-f-1)(r-f)}$.

Remark 2: In the case of (i-b), the test statistic converges in distribution to $\chi^2_{(n_1-f)(r-f)}$ if and only if $\delta^{*\prime}\tau_1 = 0$, which is equivalent to the case where $\tau_1 \in sp(\beta_1^*) = sp(\beta_1)$. See the proof in the appendix. In general, the test using (22) is a conservative test if $rk([\beta_1, \gamma_1]) = n_1$.

From Theorem 5, if we know the rank of $[\beta_1, \gamma_1]$, we can construct the test statistic that converges to a χ^2 distribution by appropriately using (21) or (22). However, such information is not available in practice. Notice that if $rk[\beta_1, \gamma_1] = n_1 - 1$, $\hat{\Phi}$ given by (21) may violate the condition that it is a full rank matrix, and in that case, the test statistic converges, not to the same χ^2 distribution as given by Theorem 5 (ii), but to a random variable that depends on a nuisance parameter. Then, the test using (21) is not available in practice. On the other hand, if we use $\hat{\Phi}$ given by (22), we can test the hypothesis by referring to a χ^2 distribution irrespective of the rank of $[\beta_1, \gamma_1]$, although the test may be conservative and the degree of freedom changes depending on the rank of $[\beta_1, \gamma]$. Then, noting that the critical value of $\chi^2_{(n_1-f)(r-f)}$ in Theorem 5 (i-b) is larger than that of $\chi^2_{(n_1-f-1)(r-f)}$ in Theorem 5 (ii), we propose to test the null of $rk(\beta_1) = f$ as follows:

1. We construct the test statistic $\mathcal{L}_T$ using (22).
2. If $\mathcal{L}_T$ is larger than the critical value of $\chi^2_{(n_1-f)(r-f)}$, we reject the null hypothesis.
3. If $\mathcal{L}_T$ is smaller than the critical value of $\chi^2_{(n_1-f-1)(r-f)}$, we accept the null hypothesis.

In this procedure, we may encounter a case where the test statistic is larger than the critical value of $\chi^2_{(n_1-f-1)(r-f)}$ but smaller than that of $\chi^2_{(n_1-f)(r-f)}$, that is, the case where $c_{(n_1-f-1)(r-f)} \leq \mathcal{L}_T \leq c_{(n_1-f)(r-f)}$, where $c_{(n_1-f-1)(r-f)}$ and $c_{(n_1-f-1)(r-f)}$ are corresponding critical values. To cope with such a case, the following corollary is useful.

Corollary 1 *Let $\hat{\Psi} = \hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha}$ and $\hat{\Phi}$ be given by (22). Suppose that the rank of β_1 is f $(< p)$.*
(i) If $rk([\beta_1, \gamma_1]) = n_1$, $T^2\hat{\lambda}_p$ converges in distribution to a random variable that is bounded above by λ^*_{min}, where λ^*_{min} is the smallest non-zero eigenvalue of

$$|X^* X^{*\prime} - \lambda^* I_{n_1-f}| = 0,$$

where X^* *is a* $(r-f)\times(n_1-f)$ *matrix with* $vec(X^*)\sim N(0, I_{(r-f)(n_1-f)})$.

(ii) If $rk([\beta_1,\gamma_1]) = n_1 - 1$, $T^2\hat{\lambda}_p$ *converges in probability to zero.*

Table 1 shows the percentage points of λ^*_{min} for the case when $(n_1-f)\geq(r-f)$. Since the non-zero eigenvalues of $X^*X^{*\prime}$ are the same as those of $X^{*\prime}X^*$, we can refer to the percentage points of $(r-f, n_1-f)$ when $(n_1-f)\leq(r-f)$.

Using the above corollary, we can cope with the situation where $c_{(n_1-f-1)(r-f)}\leq\mathcal{L}_T\leq c_{(n_1-f)(r-f)}$. If $T^2\hat{\lambda}_p$ is smaller than the (10, 5 or 1%) percentage point of λ^*_{min}, we reject the hypothesis of $rk([\beta_1,\gamma_1]) = n_1$. In that case, $c_{(n_1-f-1)(r-f)}$ is an appropriate critical value for $\mathcal{L}_T$, so that the null of $rk(\beta_1) = f$ is rejected. On the other hand, if $T^2\hat{\lambda}_p$ is larger than the critical point of λ^*_{min}, we accept the hypothesis of $rk([\beta_1,\gamma_1]) = n_1$, so that the rank of β_1 is decided to be f.

Next, we investigate a test of the rank of $\beta_{\perp,1}$. When the data is trending, $\beta_{\perp,1}$ can be decomposed into $[\gamma_1,\tau_1]$ where γ_1 and τ_1 are the first n_1 rows of γ and τ, respectively. Then, testing the rank of $\beta_{\perp,1}$ is equivalent to testing the rank of $[\gamma_1,\tau_1]$ and, therefore, we construct a test statistic from $[\hat{\gamma}_1,\hat{\tau}_1]$. Note that $\hat{\beta}_{\perp,1}$ is the first n_1 rows of $\hat{\beta}_\perp$ and is not necessarily numerically equal to $[\hat{\gamma}_1,\hat{\tau}_1]$, although they span the same column space.

Let us consider the same determinant equation as (12) with $\hat{\beta}_{\perp,1}$ replaced by $[\hat{\gamma}_1,\hat{\tau}_1]$ and

$$\hat{\hat{\Psi}} = \begin{bmatrix} (\hat{\Omega}^{11})^{-1} & 0 \\ 0 & 1 \end{bmatrix}, \tag{23}$$

$$\begin{aligned} \hat{\hat{\Phi}} &= [\hat{\gamma}_1,\hat{\tau}_1,\hat{\beta}_1(\hat{\beta}'\hat{\beta})^{-1}]\begin{bmatrix} (\hat{\gamma}'\hat{\gamma})^{-1} & 0 & 0 \\ 0 & (\hat{\tau}'\hat{\tau})^{-2} & 0 \\ 0 & 0 & (\hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha})^{-1} \end{bmatrix}\begin{bmatrix} \hat{\gamma}_1' \\ \hat{\tau}_1' \\ (\hat{\beta}'\hat{\beta})^{-1}\hat{\beta}_1' \end{bmatrix} \\ &= \hat{\gamma}_1(\hat{\gamma}'\hat{\gamma})^{-1}\hat{\gamma}_1' + \hat{\tau}_1(\hat{\tau}'\hat{\tau})^{-2}\hat{\tau}_1' + \hat{\beta}_1(\hat{\beta}'\hat{\beta})^{-1}(\hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha})^{-1}(\hat{\beta}'\hat{\beta})^{-1}\hat{\beta}_1'. \end{aligned}$$

We construct the test statistic $\mathcal{L}_{\perp T}$ in the same way as in the previous section.

Theorem 6 *Let* $\hat{\Psi}$ *and* $\hat{\Phi}$ *be given by (23) and (24). If* $g<q$, *under* $H_{0\perp}$, $\mathcal{L}_{\perp T}$ *converges in distribution to a random variable that is bounded above by* $\chi^2_{(n_1-g)(n-g-r)}$.

In a similar way to the non-trending data case, we can use the results of Theorems 5 and 6 to decide the ranks of β_1 and $\beta_{\perp,1}$ sequentially.

5. Simulation Results

In this section, we investigate the finite sample properties of the tests proposed in the previous sections. We consider the following four-dimensional error-correction model as a data generating process (DGP).

$$\triangle x_t = d_0 + \alpha\beta' x_{t-1} + \varepsilon_t,$$

where $\{\varepsilon_t\} \sim i.i.d.N(0, I_4)$. The following three DGPs are considered.

$$\text{DGP1:} \quad \alpha = \begin{bmatrix} -0.5 & 0.3 \\ 0.0 & -0.3 \\ -0.3 & -0.8 \\ -0.5 & 0.8 \end{bmatrix}, \ \beta = \begin{bmatrix} 1.0 & 0.0 \\ -1.0 & 0.0 \\ 1.0 & 1.0 \\ 0.0 & -0.5 \end{bmatrix}, \ \beta_\perp = \begin{bmatrix} 1.0 & 0.0 \\ 1.0 & 1.0 \\ 0.0 & 1.0 \\ 0.0 & 2.0 \end{bmatrix}, \ d_0 = \begin{bmatrix} -0.5 \\ 1.0 \\ 0.5 \\ -1.0 \end{bmatrix},$$

$$\text{DGP2:} \quad \alpha = \begin{bmatrix} -0.5 & 0.3 \\ 0.0 & -0.3 \\ -0.3 & -0.8 \\ -0.5 & 0.8 \end{bmatrix}, \ \beta = \begin{bmatrix} 1.0 & 0.0 \\ -1.0 & 0.0 \\ 0.0 & 1.0 \\ 0.0 & -0.5 \end{bmatrix}, \ \beta_\perp = \begin{bmatrix} 1.0 & 0.0 \\ 1.0 & 0.0 \\ 0.0 & 0.5 \\ 0.0 & 1.0 \end{bmatrix}, \ d_0 = \begin{bmatrix} -0.5 \\ 1.0 \\ 0.5 \\ -1.0 \end{bmatrix},$$

$$\text{DGP3:} \quad \alpha = \begin{bmatrix} 0.23 & 0.60 \\ 0.34 & 0.40 \\ 0.29 & 0.82 \\ 0.30 & 0.50 \end{bmatrix}, \ \beta = \begin{bmatrix} 1.0 & 0.0 \\ 0.0 & 1.0 \\ 0.5 & -0.5 \\ -1.5 & -1.0 \end{bmatrix}, \ \beta_\perp = \begin{bmatrix} 0 & 5 \\ 5 & 0 \\ 6 & -4 \\ 2 & 2 \end{bmatrix}, \ d_0 = \begin{bmatrix} 0.0 \\ -0.1 \\ 0.7 \\ -0.2 \end{bmatrix},$$

where DGP3 is the same as the simulation example in Reinsel and Ahn (1992). We test the rank of the first 2×2 matrices of β and $\beta_\perp$. Then, for DGP1, DGP2, and DGP3, the true rank of β_1 is 1, 1, and 2 respectively, while that of $\beta_{\perp,1}$ is 2, 1, and 2, respectively. We set $x_0 = 0$ and discard the first 100 observations in all experiments. The number of replication is 1,000, and the level of significance is set equal to 0.05.

Table 2 reports the simulation results of the tests for non-trending data with $d = 0$ and $d = \alpha\rho_0$, and the tests for trending data. For non-trending data, we set $d_0 = 0$. Notice that the rank of β_1 equals 1 for DGP1 and DGP2, and then the corresponding entries in the table are rejection frequencies under the null hypothesis. For the test of non-trending data with $d = 0$, although the null hypothesis is rejected slightly more often than the nominal level, 0.05, the test seems to work well. On the other hand, the other entries for the test of $rk(\beta_1)$ correspond to the power, which is close to 1. Similarly, the test of $rk(\beta_{\perp,1})$ seems to have good finite sample properties both under the null and the alternative hypotheses.

It can be seen in the middle of Table 2 that the tests for non-trending data with $d = \alpha\rho_0$ performs similarly to the case with $d = 0$.

The results of the tests for trending data are reported in the bottom of Table 2. We can see that the test works well both under the null and the alternative hypotheses, although the empirical size of $\mathcal{L}_T$ is slightly lower than the nominal level, 0.05. The reason may be that the test is conservative in these cases.

To further investigate power properties of the tests, we use DGP2 with the $(2, 2)$ element of β replaced by $c_1 > 0$ for the test of $rk(\beta_1)$ and DGP2 with the $(2, 2)$ element of $\beta_\perp$ replaced by $c_2 > 0$ for the tests of $rk(\beta_{\perp,1})$. We test for the null hypothesis of $rk(\beta_1) = 1$ or $rk(\beta_{\perp,1}) = 1$ for each specification of a constant term d. The simulation results are reported in Table 3. From the table, it seems difficult to detect the correct rank for $c_1 < 0.01$, even when the sample size is 200. However, when c_1 becomes larger than 0.01, the power of the test increases and it attains almost 1 when c_1 is 0.1 and $T = 200$.

For the test of $rk(\beta_{\perp,1})$, we have to use different β depending on $\beta_\perp$ to generate a process. In our simulation, we normalized β so that the $(1, 1)$ and $(3, 2)$ elements of β become 1. From the table, we can see that the power tends to increase when c_2 becomes larger than 0.01. We also note that the power of the test for trending data is lower compared with other tests.

We also conducted the simulation using DGP2 with the $(1, 2)$ element of β or $\beta_\perp$ replaced by non-zero value, but the relative performances of the tests are very similar to the results in Table 3 and we do not report to save the space.

6. Conclusion

In this paper, we proposed a test of the rank of the sub-matrix of cointegration. The test statistic is constructed by using the eigenvalues of the quadratic form of the sub-matrix. For non-trending data, the test statistic converges in distribution to a χ^2 distribution under the null hypothesis, while for trending data, the test is conservative in general. Finite sample simulations reveal that, although the simulation settings are limited, the proposed test works well both under the null and the alternative hypotheses with a moderate sample size, $T = 100$ and 200.

Appendix

In this appendix, we use the notation H and J alternately for different definitions if there is no confusion.

Proof of Theorem 1: First, note that we can replace $\hat{\beta}_1$ and $\hat{\beta}$ by $\tilde{\beta}_1$ and $\tilde{\beta}$ in (8), where $\tilde{\beta}_1$ is the first n_1 rows of $\tilde{\beta}$, because $\hat{\beta}\hat{\alpha}' = \tilde{\beta}\tilde{\alpha}'$ and $\hat{\beta}_1(\hat{\beta}'\hat{\beta})^{-1}\hat{\beta}_1' = \tilde{\beta}_1(\tilde{\beta}'\tilde{\beta})^{-1}\tilde{\beta}_1'$. The latter relation is established because $\tilde{\beta}$ is obtained by the non-singular transformation of the columns of $\hat{\beta}$ (see (4)) and $\hat{\beta}_1(\hat{\beta}'\hat{\beta})^{-1}\hat{\beta}_1'$ does not depend on the normalization of $\hat{\beta}$ and $\hat{\beta}_1$. We also define

$$\tilde{\beta}_\perp = \beta_\perp - \beta(\tilde{\beta}'\beta)^{-1}\tilde{\beta}'\beta_\perp. \tag{24}$$

Since $\tilde{\beta}'\tilde{\beta}_\perp = 0$ and $\tilde{\beta}_\perp$ has full column rank, the columns of $\tilde{\beta}_\perp$ span the orthogonal complement to $sp(\tilde{\beta})$, so that $\tilde{\beta}_\perp$ and $\hat{\beta}_\perp$ span the same column space. This implies that $\tilde{\beta}_\perp$ can be obtained by the non-singular transformation of the columns of $\hat{\beta}_\perp$. Then, we can also replace $\hat{\beta}_\perp$ by $\tilde{\beta}_\perp$ in (8).

Under the null hypothesis, $rk(\beta_1)$ is f and then an $n_1 \times f$ matrix β_1^* exists with rank f such that $sp(\beta_1) = sp(\beta_1^*)$. We denote the orthogonal complement to β_1^* by δ^*. That is, δ^* is an $n_1 \times (n_1 - f)$ matrix with rank $(n_1 - f)$ such that $\delta^{*\prime}\beta_1^* = 0$. Note that the $n_1 \times n_1$ square matrix $[\beta_1^*, \delta^*]$ has full rank n_1.

Lemma 1 *: (i)* $\tilde{\beta} \xrightarrow{p} \beta$, $\tilde{\alpha} \xrightarrow{p} \alpha$, $\tilde{\Sigma} \xrightarrow{p} \Sigma$.
(ii) $T\delta^{*\prime}\tilde{\beta}_1 = T\delta^{*\prime}(\tilde{\beta}_1 - \beta_1) \xrightarrow{d} \delta^{*\prime}\beta_{\perp,1}(\beta_\perp'\beta_\perp)^{-1}(\int G_0G_0'ds)^{-1}\int G_0dV' = X_0'$*, say.*
(iii) $T(\tilde{\beta}_\perp - \beta_\perp) \xrightarrow{d} -\beta(\beta'\beta)^{-1}\int dVG_0'(\int G_0G_0'ds)^{-1}$.
(iv) $T^{-1}\bar{\tilde{\beta}}_\perp' S_{11}\bar{\tilde{\beta}}_\perp \xrightarrow{d} \int G_0G_0'ds$.

Proof: (i) is proved by Johansen (1988, 1996).
(ii) As shown in Chapter 13.2 of Johansen (1996), $\tilde{\beta}$ can be expressed as $\tilde{\beta} = \beta + \beta_\perp(\beta_\perp'\beta_\perp)^{-1}U_T$ for non-trending data, where TU_T converges in distribution to $(\int G_0G_0'ds)^{-1}\int G_0dV'$. Since $\tilde{\beta}_1$ is the first n_1 rows of $\tilde{\beta}$, we have $\tilde{\beta}_1 = \beta_1 + \beta_{\perp,1}(\beta_\perp'\beta_\perp)^{-1}U_T$, so that

$$\begin{aligned} T\delta^{*\prime}\tilde{\beta}_1 &= T\delta^{*\prime}(\tilde{\beta}_1 - \beta_1) \\ &= \delta^{*\prime}\beta_{\perp,1}(\beta_\perp'\beta_\perp)^{-1}(TU_T) \end{aligned}$$

$$\xrightarrow{d} \quad \delta^{*\prime}\beta_{\perp,1}(\beta_\perp'\beta_\perp)^{-1}\left(\int G_0G_0'ds\right)^{-1}\int G_0dV',$$

where the first equation holds because $\delta^{*\prime}\beta_1 = 0$.

(iii) This convergence holds because $T(\tilde\beta_\perp - \beta_\perp) = -\beta(\tilde\beta'\beta)^{-1}(\tilde\beta-\beta)'\beta_\perp T = -\beta(\tilde\beta'\beta)^{-1}(TU_T)'$ from (24).

(iv) Note that

$$\begin{aligned}\frac{1}{T}\bar{\tilde\beta}_\perp' S_{11}\bar{\tilde\beta}_\perp &= \frac{1}{T}\bar\beta_\perp' S_{11}\bar\beta_\perp + \frac{1}{T}(\bar{\tilde\beta}_\perp - \bar\beta_\perp)'S_{11}\bar{\tilde\beta}_\perp \\ &\quad + \frac{1}{T}\bar{\tilde\beta}_\perp' S_{11}(\bar{\tilde\beta}_\perp - \bar\beta_\perp) + \frac{1}{T}(\bar{\tilde\beta}_\perp - \bar\beta_\perp)'S_{11}(\bar{\tilde\beta}_\perp - \bar\beta_\perp).\end{aligned}$$

The first term converges in distribution to $\int G_0G_0'ds$ from Johansen (1988, 1996), while the remaining terms converge in probability to zero because $(\bar{\tilde\beta}_\perp - \bar\beta_\perp)$ and S_{11} are of order T^{-1} and T, respectively.□

Now, let us consider the determinant equation (8). Since (8) is equivalent to

$$|H'||\tilde\beta_1\tilde\Psi\tilde\beta_1' - \hat\lambda\tilde\Phi||H| = 0, \tag{25}$$

where H is any $n\times n$ non-singular matrix, we consider (25) with $H = [\beta_1^*, T\delta^*]$. Then, using Lemma 1, we have

$$\begin{aligned}H'\tilde\beta_1\tilde\Psi\tilde\beta_1'H &= \begin{bmatrix} \beta_1^{*\prime}\tilde\beta_1\tilde\Psi\tilde\beta_1'\beta_1^* & \beta_1^{*\prime}\tilde\beta_1\tilde\Psi(\tilde\beta_1'\delta^*T) \\ (T\tilde\delta^{*\prime}\tilde\beta_1)\tilde\Psi\tilde\beta_1'\beta_1^* & (T\tilde\delta^{*\prime}\tilde\beta_1)\tilde\Psi(\tilde\beta_1'\delta^*T)\end{bmatrix} \\ &\xrightarrow{d} \begin{bmatrix} \beta_1^{*\prime}\beta_1\Psi\beta_1'\beta_1^* & \beta_1^{*\prime}\beta_1\Psi X_0 \\ X_0'\Psi\beta_1'\beta_1^* & X_0'\Psi X_0\end{bmatrix}.\end{aligned} \tag{26}$$

To investigate the asymptotic behavior of $H'\tilde\Phi H$, we consider $\tilde\Phi$ with the same expression as (9). Note that

$$\begin{aligned}H'\left[\tilde\beta_1, \tilde\beta_{\perp,1}(\tilde\beta_\perp'\tilde\beta_\perp)^{-1}\right] &= \begin{bmatrix} \beta_1^{*\prime}\tilde\beta_1 & \beta_1^{*\prime}\tilde\beta_{\perp,1}(\tilde\beta_\perp'\tilde\beta_\perp)^{-1} \\ T\delta^{*\prime}\tilde\beta_1 & T\delta^{*\prime}\tilde\beta_{\perp,1}(\tilde\beta_\perp'\tilde\beta_\perp)^{-1}\end{bmatrix} \\ &= \begin{bmatrix} O_p(1) & O_p(1) \\ O_p(1) & T\delta^{*\prime}\beta_{\perp,1}(\beta_\perp'\beta_\perp)^{-1} + o_p(T)\end{bmatrix}\end{aligned}$$

because $\beta_1^{*\prime}\tilde\beta_1 \xrightarrow{p} \beta_1^{*\prime}\beta_1$, $\beta_1^{*\prime}\tilde\beta_{\perp,1} \xrightarrow{p} \beta_1^{*\prime}\beta_{\perp,1}$, $T\delta^{*\prime}\tilde\beta_1 = O_p(1)$ and $T\delta^{*\prime}\tilde\beta_{\perp,1} = T\delta^{*\prime}\beta_{\perp,1} + o_p(T)$ by Lemma 1. Then, $\lambda H'\tilde\Phi H$ is asymptotically equivalent to

$$T^2\hat\lambda\begin{bmatrix} 0 & 0 \\ 0 & \delta^{*\prime}\beta_{\perp,1}(\beta_\perp'\beta_\perp)^{-1}(\int G_0G_0'ds)^{-1}(\beta_\perp'\beta_\perp)^{-1}\beta_{\perp,1}'\delta^*\end{bmatrix}.$$

Then, the equation (25) is asymptotically equal to

$$\left| \begin{bmatrix} \beta_1^{*\prime}\beta_1\Psi\beta_1'\beta_1^* & \beta_1^{*\prime}\beta_1\Psi X_0 \\ X_0'\Psi\beta_1'\beta_1^* & X_0'\Psi X_0 \end{bmatrix} - T^2\hat{\lambda} \begin{bmatrix} 0 & 0 \\ 0 & \delta^{*\prime}\beta_{\perp,1}(\beta_\perp'\beta_\perp)^{-1}(\int G_0G_0'ds)^{-1}(\beta_\perp'\beta_\perp)^{-1}\beta_{\perp,1}'\delta^* \end{bmatrix} \right|$$

$$\begin{aligned} &= |\beta_1^{*\prime}\beta_1\Psi\beta_1'\beta_1^*| \\ &\quad \times \Big| X_0' \left\{ \Psi - \Psi\beta_1'\beta_1^*(\beta_1^{*\prime}\beta_1\Psi\beta_1'\beta_1^*)^{-1}\beta_1^{*\prime}\beta_1\Psi \right\} X_0 \\ &\quad -T^2\hat{\lambda}\delta^{*\prime}\beta_{\perp,1}(\beta_\perp'\beta_\perp)^{-1}\left(\int G_0G_0'ds\right)^{-1}(\beta_\perp'\beta_\perp)^{-1}\beta_{\perp,1}'\delta^* \Big| = 0. \end{aligned} \tag{27}$$

Therefore, the eigenvalues $\hat{\lambda}_{f+1}\cdots\hat{\lambda}_p$ converge in probability to zeros and are of order T^2.

Here, notice that, in the same way as Johansen (1988, p.246), we can find a $r \times (r-f)$ matrix with rank $(r-f)$ such that

$$JJ' = \Psi - \Psi\beta_1'\beta_1^*(\beta_1^{*\prime}\beta_1\Psi\beta_1'\beta_1^*)^{-1}\beta_1^{*\prime}\beta_1\Psi, \tag{28}$$

with $J'(\beta_1'\beta_1^*) = 0$ and $J'\Psi^{-1}J = I_{r-f}$, implying that $J'(\alpha'\Sigma^{-1}\alpha)^{-1}J = I_{r-f}$ because $\Psi = \alpha'\Sigma^{-1}\alpha$. Then, since $|\beta_1^{*\prime}\beta_1\Psi\beta_1'\beta_1^*| \neq 0$, (27) becomes

$$\left| X_0'JJ'X_0 - T^2\hat{\lambda}\delta^{*\prime}\beta_{\perp,1}(\beta_\perp'\beta_\perp)^{-1}\left(\int G_0G_0'ds\right)^{-1}(\beta_\perp'\beta_\perp)^{-1}\beta_{\perp,1}'\delta^* \right| = 0. \tag{29}$$

Since the variance matrix of $X_0'J$ conditioned on $G_0(\cdot)$ is

$$\delta^{*\prime}\beta_{\perp,1}(\beta_\perp'\beta_\perp)^{-1}\left(\int G_0G_0'ds\right)^{-1}(\beta_\perp'\beta_\perp)^{-1}\beta_{\perp,1}'\delta^* \otimes I_{r-f},$$

we can see that $T^2\hat{\lambda}$ converges in distribution to λ^*, which is a solution of

$$|X_0^{*\prime}X_0^* - \lambda^* I_{n_1-f}| = 0, \tag{30}$$

where $X_0^{*\prime}$ is an $(n_1-f)\times(r-f)$ matrix with $vec(X_0^{*\prime}) \sim N(0, I_{n_1-f}\otimes I_{r-f})$. Then, $\mathcal{L}_T$ converges in distribution to the trace of $X_0^{*\prime}X_0^*$, which proves Theorem 1.□

Proof of Theorem 2: In the same way as the proof of Theorem 1, we replace $\hat{\beta}$, $\hat{\beta}_\perp$ and $\hat{\beta}_{\perp,1}$ by $\tilde{\beta}$, $\tilde{\beta}_\perp$ and $\tilde{\beta}_{\perp,1}$.

Under the null hypothesis, an $n_1\times g$ matrix $\beta_{\perp,1}^*$ exists such that $sp(\beta_{\perp,1}^*) = sp(\beta_{\perp,1})$ and $rk(\beta_{\perp,1}^*) = g$. We can also find an $n_1 \times (n_1-g)$ matrix η^* with rank (n_1-g) that is orthogonal to $\beta_{\perp,1}^*$. Here, we consider the following determinant equation.

$$|H'||\hat{\beta}_{\perp,1}\hat{\ddot{\Psi}}\hat{\beta}_{\perp,1}' - \hat{\mu}\hat{\ddot{\Phi}}||H| = 0, \tag{31}$$

where $H = [\beta^*_{\perp,1}, T\eta^*]$. Since $\tilde{\beta}_{\perp,1}$ is the first n_1 rows of $\tilde{\beta}_\perp$, we obtain, using Lemma 1 (iii),

$$T\eta^{*\prime}\tilde{\beta}_{\perp,1} = T\eta^{*\prime}(\tilde{\beta}_{\perp,1} - \beta_{\perp,1}) \xrightarrow{d} -\eta^{*\prime}\beta_1(\beta'\beta)^{-1}\int dVG_0'\left(\int G_0G_0'ds\right)^{-1} = Y_0', \text{ say,}$$

and $\beta^{*\prime}_{\perp,1}\tilde{\beta}_{\perp,1} \xrightarrow{p} \beta^{*\prime}_{\perp,1}\beta_{\perp,1}$. Then,

$$H'\tilde{\beta}_{\perp,1}\tilde{\ddot{\Psi}}\tilde{\beta}'_{\perp,1}H \xrightarrow{d} \begin{bmatrix} \beta^{*\prime}_{\perp,1}\beta_{\perp,1}\ddot{\Psi}\beta'_{\perp,1}\beta^*_{\perp,1} & \beta^{*\prime}_{\perp,1}\beta_{\perp,1}\ddot{\Psi}Y_0 \\ Y_0'\ddot{\Psi}\beta'_{\perp,1}\beta^*_{\perp,1} & Y_0'\ddot{\Psi}Y_0 \end{bmatrix}.$$

In addition, we can see that $\beta^{*\prime}_{\perp,1}\tilde{\beta}_1 \xrightarrow{p} \beta^{*\prime}_{\perp,1}\beta_1$, $T\eta^{*\prime}\tilde{\beta}_{\perp,1} = O_p(1)$ and $T\eta^{*\prime}\tilde{\beta}_1 = T\eta^{*\prime}\beta_1 + o_p(T)$. Then, similar to the previous proof, the determinant equation (31) is asymptotically equivalent to

$$\begin{aligned} &\left|\begin{bmatrix} \beta^{*\prime}_{\perp,1}\beta_{\perp,1}\ddot{\Psi}\beta'_{\perp,1}\beta^*_{\perp,1} & \beta^{*\prime}_{\perp,1}\beta_{\perp,1}\ddot{\Psi}Y_0 \\ Y_0'\ddot{\Psi}\beta'_{\perp,1}\beta^*_{\perp,1} & Y_0'\ddot{\Psi}Y_0 \end{bmatrix} - T^2\hat{\mu}\begin{bmatrix} 0 & 0 \\ 0 & \eta^{*\prime}\beta_1(\beta'\beta)^{-1}(\alpha'\Sigma^{-1}\alpha)^{-1}(\beta'\beta)^{-1}\beta_1'\eta^* \end{bmatrix}\right| \\ &= |\beta^{*\prime}_{\perp,1}\beta_{\perp,1}\ddot{\Psi}\beta'_{\perp,1}\beta^*_{\perp,1}| \\ &\quad\times\Big|Y_0'(\ddot{\Psi} - \ddot{\Psi}\beta'_{\perp,1}\beta^*_{\perp,1}(\beta^{*\prime}_{\perp,1}\beta_{\perp,1}\ddot{\Psi}\beta'_{\perp,1}\beta^*_{\perp,1})^{-1}\beta^{*\prime}_{\perp,1}\beta_{\perp,1}\ddot{\Psi})Y_0 \\ &\quad -T^2\hat{\mu}\eta^{*\prime}\beta_1(\beta'\beta)^{-1}(\alpha'\Sigma^{-1}\alpha)^{-1}(\beta'\beta)^{-1}\beta_1'\eta^*\Big| = 0. \end{aligned} \tag{32}$$

Then, $\hat{\mu}_{g+1}, \cdots, \hat{\mu}_q$ are of order T^{-2}. For a given $G_0(\cdot)$, we can find an $(n-r)\times(n-r-g)$ matrix with rank $(n-r-g)$ such that

$$JJ' = \ddot{\Psi} - \ddot{\Psi}\beta'_{\perp,1}\beta^*_{\perp,1}(\beta^{*\prime}_{\perp,1}\beta_{\perp,1}\ddot{\Psi}\beta'_{\perp,1}\beta^*_{\perp,1})^{-1}\beta^{*\prime}_{\perp,1}\beta_{\perp,1}\ddot{\Psi}$$

with $J'(\beta'_{\perp,1}\beta^*_{\perp,1}) = 0$ and $J'(\int G_0G_0'ds)^{-1}J = I_{n-r-g}$ because $\ddot{\Psi} = \int G_0G_0'ds$. Then, (32) becomes

$$\left|Y_0'JJ'Y_0 - T^2\hat{\mu}\eta^{*\prime}\beta_1(\beta'\beta)^{-1}(\alpha'\Sigma^{-1}\alpha)^{-1}(\beta'\beta)^{-1}\beta_1'\eta^*\right| = 0. \tag{33}$$

Since the variance matrix of $Y_0'J$ conditioned on $G_0(\cdot)$ is given by

$$\eta^{*\prime}\beta_1(\beta'\beta)^{-1}(\alpha'\Sigma^{-1}\alpha)^{-1}(\beta'\beta)^{-1}\beta_1'\eta^* \otimes I_{n-r-g},$$

we can see that $T^2\hat{\mu}$ converges in distribution to μ^*, which is a solution of

$$|Y_0^{*\prime}Y_0^* - \mu^* I_{n_1-g}| = 0,$$

where $Y_0^{*\prime}$ is an $(n_1-g)\times(n-r-g)$ matrix with $vec(Y_0^{*\prime}) \sim N(0, I_{n_1-g}\otimes I_{n-r-g})$. This proves Theorem 2.□

Proof of Theorems 3 and 4: Let $\hat{\beta}^+ = [\hat{\beta}', \hat{\rho}]'$ and $\tilde{\beta}^+ = [\tilde{\beta}', \hat{\rho}]'$. Exactly in the same way as the proof of Lemma 13.2 in Johansen (1996), we can show that

$$\begin{bmatrix} T\beta_\perp' & 0 \\ 0 & T^{1/2} \end{bmatrix} (\tilde{\beta}^+ - \beta^+) \xrightarrow{d} \left(\int G_0^+ G_0^{+\prime} ds\right)^{-1} \int G_0^+ dV',$$

where $G_0^+ = [G_0', 1]'$. Then, since $\tilde{\beta}$ is the first n rows of $\tilde{\beta}^+$, we have

$$T\beta_\perp'(\tilde{\beta}-\beta) \xrightarrow{d} L'\left(\int G_0^+ G_0^{+\prime} ds\right)^{-1} \int G_0^+ dV', \tag{34}$$

whose conditional variance is given by $L'\left(\int G_0^+ G_0^{+\prime} ds\right)^{-1} L$. Since $\tilde{\beta} = \beta + \beta_\perp(\beta_\perp'\beta_\perp)^{-1}\beta_\perp'\tilde{\beta}$ as expressed in Johansen (1996, p. 179), we have

$$\begin{aligned} T\delta^{*\prime}\tilde{\beta}_1 &= \delta^{*\prime}\beta_{\perp,1}(\beta_\perp'\beta_\perp)^{-1}T\beta_\perp'(\tilde{\beta}-\beta) \\ &\xrightarrow{d} \delta^{*\prime}\beta_{\perp,1}(\beta_\perp'\beta_\perp)^{-1}L'\left(\int G_0^+ G_0^{+\prime} ds\right)^{-1} \int G_0^+ dV'. \end{aligned}$$

We also have

$$\tilde{\Upsilon}_T' S_{11}^+ \tilde{\Upsilon}_T \xrightarrow{d} \int G_0^+ G_0^{+\prime} ds,$$

which is proved as Lemma 1(iv), where $\tilde{\Upsilon}_T$ is Υ_T with $\hat{\beta}_\perp$ replaced by $\tilde{\beta}_\perp$. Then, Theorems 3 and 4 is proved in the same way as Theorems 1 and 2.□

Proof of Theorem 5: First, we give the following lemma.

Lemma 2 *(i)* $\tilde{\gamma} \xrightarrow{p} \gamma$ *and* $\tilde{\gamma}_1 \xrightarrow{p} \gamma_1$*, where*

$$\tilde{\gamma} = \gamma - \gamma_\perp(\tilde{\gamma}_\perp'\gamma_\perp)^{-1}\tilde{\gamma}_\perp'\gamma,$$

and $\tilde{\gamma}_1$ *is the first* n_1 *rows of* $\tilde{\gamma}$ *with* $\gamma_\perp = [\beta, \tau]$ *and* $\tilde{\gamma}_\perp = [\tilde{\beta}, \hat{\tau}]$.
(ii) $T^{1/2}(\hat{\tau}-\tau) \xrightarrow{d} CW(1)$.
(iii) $\tilde{\Omega}^{11} \xrightarrow{d} \Omega^{11}$*, where* $\tilde{\Omega}^{11}$ *is defined as* $\hat{\Omega}^{11}$ *with* $\hat{\gamma}$ *replaced by* $\tilde{\gamma}$.

Proof: (i) Since $\tilde{\gamma}_\perp \xrightarrow{p} \gamma_\perp$ and $\gamma_\perp'\gamma = 0$, $\tilde{\gamma}$ converges in probability to γ.
(ii) This is proved by Johansen (1991, 1996).

(iii) Letting $K = [\bar{\beta}, T^{-1/2}\bar{\gamma}, T^{-1}\bar{\tau}]$, we can see that

$$T\tilde{\gamma}' S_{11}^{-1}\tilde{\gamma} = (T^{1/2}\tilde{\gamma}' K)(K' S_{11} K)^{-1}(T^{1/2}K'\tilde{\gamma}).$$

From Lemma 10.3 in Johansen (1996), $T^{-1}[\bar{\gamma}, T^{-1/2}\bar{\tau}]' S_{11}[\bar{\gamma}, T^{-1/2}\bar{\tau}]$ converges in distribution to Ω while $\beta' S_{11}\beta$ converges in probability to a positive definite matrix, Σ_β, and $[\bar{\gamma}, T^{-1/2}\bar{\tau}]' S_{11}\beta = O_p(1)$. Then,

$$K' S_{11} K \xrightarrow{d} \begin{bmatrix} (\beta'\beta)^{-1}\Sigma_\beta(\beta'\beta)^{-1} & 0 \\ 0 & \Omega \end{bmatrix}. \tag{35}$$

In addition, we can see that

$$\begin{aligned} T\beta'\tilde{\gamma} &= -T\beta'\gamma_\perp(\tilde{\gamma}_\perp\gamma_\perp)^{-1}\tilde{\gamma}_\perp'\gamma \\ &= -[\beta'\beta, 0](\tilde{\gamma}_\perp'\gamma_\perp)^{-1}\begin{bmatrix} (\tilde{\beta}-\beta)'\gamma T \\ (\hat{\tau}-\tau)'\gamma T \end{bmatrix} \\ &\xrightarrow{d} -U_1' \end{aligned}$$

because $\beta'\tau = 0$ and $\hat{\tau}'\gamma = O_p(T^{1/2})$. Using this result, we have

$$T^{1/2}K'\tilde{\gamma} = \begin{bmatrix} T^{1/2}(\beta'\beta)^{-1}\beta'\tilde{\gamma} \\ (\gamma'\gamma)^{-1}\gamma'\tilde{\gamma} \\ T^{-1/2}(\tau'\tau)^{-1}\tau'\tilde{\gamma} \end{bmatrix} \xrightarrow{p} \begin{bmatrix} 0 \\ I_{m-r-1} \\ 0 \end{bmatrix}. \tag{36}$$

From (35) and (36), $\tilde{\Omega}^{11}$ converges in distribution to Ω^{11}. $\square$

For the same reason as the previous proofs, we replace $\hat{\beta}$, $\hat{\beta}_1$, $\hat{\gamma}$ and $\hat{\gamma}_1$ by $\tilde{\beta}$, $\tilde{\beta}_1$, $\tilde{\gamma}$ and $\tilde{\gamma}_1$.

(i-a) We consider the same determinant equation as (25) with $H = [\beta_1^*, T\delta^*]$. Using (20), we can see that $H'\tilde{\beta}_1\tilde{\Psi}\tilde{\beta}_1' H$ converges to the same limit as (26), replacing X_0 by X. On the other hand, because $\beta_1^{*\prime}\tilde{\beta}_1 \xrightarrow{p} \beta_1^{*\prime}\beta_1$, $\beta_1^{*\prime}\tilde{\gamma}_1 \xrightarrow{p} \beta_1^{*\prime}\gamma_1$, $T\delta^{*\prime}\tilde{\beta}_1 = O_p(1)$, $T\delta^{*\prime}\tilde{\gamma}_1 = T\delta^{*\prime}\gamma_1 + o_p(T)$ and $\tilde{\Omega}^{11} \xrightarrow{d} \Omega^{11}$ by Lemma 2, $\hat{\lambda}H'\tilde{\Phi}H$ is asymptotically equivalent to

$$T^2\hat{\lambda}\begin{bmatrix} 0 & 0 \\ 0 & \delta^{*\prime}\gamma_1(\gamma'\gamma)^{-1}\Omega^{11}(\gamma'\gamma)^{-1}\gamma_1'\delta^* \end{bmatrix}.$$

Then, similar to (27), as far as the limiting distribution is concerned, it is sufficient to consider,

$$|\beta_1^{*\prime}\beta_1\Psi\beta_1'\beta_1^*|\left|X'JJ'X - T^2\hat{\lambda}\delta^{*\prime}\gamma_1(\gamma'\gamma)^{-1}\Omega^{11}(\gamma'\gamma)^{-1}\gamma_1'\delta^*\right| = 0,$$

where J is the same $r \times (r-f)$ matrix as in (28). Since $vec(X'J)$ conditioned on $G(\cdot)$ is normally distributed with a variance matrix $\delta^{*\prime}\gamma_1(\gamma'\gamma)^{-1}\Omega^{11}(\gamma'\gamma)^{-1}\gamma_1'\delta^* \otimes I_{r-f}$, $\mathcal{L}_T$ converges in distribution to $\chi^2_{(n_1-f)(r-f)}$.

(i-b) We consider the same determinant equation as in the proof of (i-a). We can easily see that $H'\tilde{\beta}_1\tilde{\Psi}\tilde{\beta}_1'H$ converges to the same limit as (i-a), while $\hat{\lambda}H'\tilde{\Phi}H$ is asymptotically equivalent to

$$T^2\hat{\lambda}\begin{bmatrix} 0 & 0 \\ 0 & \delta^{*\prime}\gamma_1(\gamma'\gamma)^{-1}\Omega^{11}(\gamma'\gamma)^{-1}\gamma_1'\delta^* + \delta^{*\prime}\tau_1\tau_1'\delta^* \end{bmatrix}.$$

because $\beta_1^{*\prime}\hat{\tau}_1 \xrightarrow{p} \beta_1^{*\prime}\tau_1$ and $T\delta^{*\prime}\hat{\tau}_1 = T\delta^{*\prime}\tau_1 + o_p(T)$ by Lemma 2. Then, similar to (27), as far as the limiting distribution is concerned, it is sufficient to consider,

$$\left|\beta_1^{*\prime}\beta_1\Psi\beta_1'\beta_1^*\right|\left|X'JJ'X - T^2\hat{\lambda}\left\{\delta^{*\prime}\gamma_1(\gamma'\gamma)^{-1}\Omega^{11}(\gamma'\gamma)^{-1}\gamma_1'\delta^* + \delta^{*\prime}\tau_1\tau_1'\delta^*\right\}\right| = 0,$$

where J is the same $r \times (r-f)$ matrix as (28). Here, note that, in general, for a given symmetric and positive definite matrix A and a vector b,

$$(A + bb')^{-1} = A^{-1} - A^{-1}bb'A^{-1}/(1 + b'A^{-1}b), \tag{37}$$

and then,

$$c'(A + bb')^{-1}c \leq c'A^{-1}c$$

for any non-zero vector c. By substituting $\delta^{*\prime}\gamma_1(\gamma'\gamma)^{-1}\Omega^{11}(\gamma'\gamma)^{-1}\gamma_1'\delta^*$ and $\delta^{*\prime}\tau_1$ for A and b, we obtain, for a given $G(\cdot)$,

$$\begin{aligned} & tr\left(J'X\{\delta^{*\prime}\gamma_1(\gamma'\gamma)^{-1}\Omega^{11}(\gamma'\gamma)^{-1}\gamma_1'\delta^* + \delta^{*\prime}\tau_1\tau_1'\delta^*\}^{-1}X'J\right) \\ \leq \ & tr\left(J'X\{\delta^{*\prime}\gamma_1(\gamma'\gamma)^{-1}\Omega^{11}(\gamma'\gamma)^{-1}\gamma_1'\delta^*\}^{-1}X'J\right) \\ = \ & tr(X^*X^{*\prime}) = \chi^2_{(r-f)(n_1-f)}, \end{aligned} \tag{38}$$

where X^* is a $(r-f) \times (n_1-f)$ matrix with $vec(X^*) \sim N(0, I_{(r-f)(n_1-f)})$. The equality is established if and only if $\delta^{*\prime}\tau_1 = 0$.

(ii) Let us consider the determinant equation (25) with $H = [\beta_1^*, T\delta_0^*, T\tau_1^*]$. Using (18) and (19), we have

$$H'\tilde{\beta}_1\tilde{\Psi}\tilde{\beta}_1'H \xrightarrow{d} \begin{bmatrix} \beta_1^{*\prime}\beta_1\Psi\beta_1'\beta_1^* & \beta_1^{*\prime}\beta_1\Psi X_1 & 0 \\ X_1'\Psi\beta_1'\beta_1^* & X_1'\Psi X_1 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{39}$$

On the other hand, because $\beta_1^{*\prime}\tilde{\beta}_1 \xrightarrow{p} \beta_1^{*\prime}\beta_1$, $\beta_1^{*\prime}\tilde{\gamma}_1 \xrightarrow{p} \beta_1^{*\prime}\gamma_1$, $\beta_1^{*\prime}\hat{\tau}_1 \xrightarrow{p} \beta_1^{*\prime}\tau_1$, $T\delta_0^{*\prime}\tilde{\beta}_1 = O_p(1)$, $T\delta_0^{*\prime}\tilde{\gamma}_1 = T\delta_0^{*\prime}\gamma_1 + o_p(T)$, $T\delta_0^{*\prime}\hat{\tau}_1 = T\delta^{*\prime}\tau_1 + o_p(T)$, $T\tau_1^{*\prime}\tilde{\beta}_1 = O_p(T^{-1/2})$, $T\tau_1^{*\prime}\tilde{\gamma}_1 = O_p(T^{1/2})$, $T\tau_1^{*\prime}\hat{\tau}_1 = T\tau_1^{*\prime}\tau_1 + o_p(T)$ and $\tilde{\Omega}^{11} \xrightarrow{d} \Omega^{11}$ by Lemma 2, $\hat{\lambda}H'\tilde{\Phi}H$ is asymptotically equivalent to

$$T^2\hat{\lambda}\begin{bmatrix} 0 & 0 & 0 \\ 0 & \delta_0^{*\prime}\gamma_1(\gamma'\gamma)^{-1}\Omega^{11}(\gamma'\gamma)^{-1}\gamma_1'\delta_0^* + \delta_0^{*\prime}\tau_1\tau_1'\delta_0^* & \delta_0^{*\prime}\tau_1\tau_1'\tau_1^* \\ 0 & \tau_1^{*\prime}\tau_1\tau_1'\delta_0^* & \tau_1^{*\prime}\tau_1\tau_1'\tau_1^* \end{bmatrix}.$$

Then, after some algebra, we can see that (25) is asymptotically equal to

$$\left|\beta_1^{*\prime}\beta_1\Psi\beta_1'\beta_1^*\right|\left|-T^2\hat{\lambda}(\tau_1^{*\prime}\tau_1)^2\right|\left|X_1'JJ'X_1 - T^2\hat{\lambda}\delta_0^{*\prime}\gamma_1(\gamma'\gamma)^{-1}\Omega^{11}(\gamma'\gamma)^{-1}\gamma_1'\delta_0^*\right| = 0,$$

where J is the same $r \times (r-f)$ matrix as (28). This determinant equation implies that there are f non-zero eigenvalues, $p-f-1$ eigenvalues of order T^{-2}, and one eigenvalue of order smaller than T^{-2}. In the same way as in the previous proofs of the theorems, we can see that

$$T^2\sum_{i=f+1}^{p-1}\hat{\lambda}_i \xrightarrow{d} \chi^2_{(n_1-f-1)(r-f)}.$$

Since $T^2\hat{\lambda}_p \xrightarrow{p} 0$, we have

$$\mathcal{L}_T = T^2\sum_{i=f+1}^{p}\hat{\lambda}_i = T^2\sum_{i=f+1}^{p-1}\hat{\lambda}_i + o_p(1) \xrightarrow{d} \chi^2_{(n_1-f-1)(r-f)}.$$

We can also see that $\hat{\lambda}_p$ is of order T^3 if we choose $H = [\beta_1^*, T\delta_0^*, T^{3/2}\tau_1^*]$.□

Proof of Corollary 1: (i) When $rk[\beta_1, \gamma_1] = n_1$, from (38) in the proof of Theorem 5, $T^2\hat{\lambda}_p$ converges in distribution to the $p-f$-th eigenvalue (the smallest non-zero eigenvalue) of

$$J'X\left\{\delta^{*\prime}\gamma_1(\gamma'\gamma)^{-1}\Omega^{11}(\gamma'\gamma)^{-1}\gamma_1'\delta^* + \delta^{*\prime}\tau_1\tau_1'\delta^*\right\}^{-1}X'J.$$

Here, note that, in general, for a given positive definite matrix A, a vector b and a matrix D,

$$D'A^{-1}D = D'(A+bb')^{-1}D + D'A^{-1}bb'A^{-1}D/(1+b'A^{-1}b),$$

where we used the relation (37). By Theorem 9 of Magnus and Neudecker (1988, p.208), we can see that the $p-f$-th eigenvalue of $D'A^{-1}D$ is larger than that of $D'(A+bb')^{-1}D$. Then, by substituting $\delta^{*\prime}\gamma_1(\gamma'\gamma)^{-1}\Omega^{11}(\gamma'\gamma)^{-1}\gamma_1'\delta^*$, $\delta^{*\prime}\tau_1$ and $X'J$ for A, b and D, respectively, we can see that the limiting distribution of $T^2\hat{\lambda}_p$ is bounded above by λ^*_{min}, the smallest non-zero eigenvalue of

$$J'X\left\{\delta^{*\prime}\gamma_1(\gamma'\gamma)^{-1}\Omega^{11}(\gamma'\gamma)^{-1}\gamma_1'\delta^*\right\}^{-1}X'J = X^*X^{*\prime},$$

where X^* is an $(r-f)\times(n_1-f)$ matrix with $vec(X^*)\sim N(0, I_{(n-r)(n_1-f)})$. Note that $T^2\hat{\lambda}_p \xrightarrow{d} \lambda^*_{min}$ if and only if $\delta^{*\prime}\tau_1 = 0$.
(ii) This is proved in Theorem 5 (ii).□

Proof of Theorem 6: Similar to the proof of Theorem 2, under the null hypothesis there exists an $n_1\times g$ matrix $\beta^*_{\perp,1}$ with rank g whose columns span the same space as $sp(\beta_{\perp,1})$, and an $n_1\times(n_1-g)$ matrix η^* with rank (n_1-g) that is orthogonal to $\beta^*_{\perp,1}$. For the same reason as before, we replace $\hat{\beta}$, $\hat{\gamma}$ and $\hat{\gamma}_1$ by $\tilde{\beta}$, $\tilde{\gamma}$ and $\tilde{\gamma}_1$.

First, we give the convergence result of $\tilde{\gamma}_1$ and $\hat{\tau}_1$, where

$$\tilde{\gamma}_1 = \gamma_1 - \gamma_{\perp,1}(\tilde{\gamma}_\perp'\gamma_\perp)^{-1}\tilde{\gamma}_\perp'\gamma, \tag{40}$$

with $\gamma_{\perp,1} = [\beta_1, \tau_1]$.

Lemma 3 *(i)* $T\eta^{*\prime}\tilde{\gamma}_1 \xrightarrow{d} -\eta^{*\prime}\beta_1(\beta'\beta)^{-1}U_1' = Y'$, *say.*
(ii) $T\eta^{*\prime}\hat{\tau}_1 \xrightarrow{p} 0$.

Proof: (i) Since $\eta^{*\prime}\gamma_1 = 0$ and $\eta^{*\prime}\tau_1 = 0$, we have, using (40),

$$\begin{aligned} T\eta^{*\prime}\tilde{\gamma}_1 &= T\eta^{*\prime}(\tilde{\gamma}_1-\gamma_1) \\ &= -[\eta^{*\prime}\beta_1, 0](\tilde{\gamma}_\perp\gamma_\perp)^{-1}\left[\begin{array}{c}(\tilde{\beta}-\beta)'\gamma T \\ (\hat{\tau}-\tau)'\gamma T\end{array}\right] \\ &\xrightarrow{d} -\eta^{*\prime}\beta_1(\beta'\beta)^{-1}U_1', \end{aligned}$$

where the last convergence is established because $(\hat{\tau}-\tau)$ is $O_p(T^{-1/2})$ and $\tilde{\beta}'\tau$ is $O_p(T^{-3/2})$.
(ii) First, note that, because $\hat{\tau}_1 = \hat{\beta}_{\perp,1}(\hat{\alpha}_\perp'\hat{\Gamma}\hat{\beta}_\perp)^{-1}\hat{\alpha}_\perp'\hat{\mu}$, $\hat{\tau}_1$ is invariant to each normalization of $\hat{\alpha}_\perp$ and $\hat{\beta}_\perp$. Then, we can express $\hat{\tau}_1$ as

$$\hat{\tau}_1 = \tilde{\beta}_{\perp,1}(\tilde{\alpha}_\perp'\hat{\Gamma}\tilde{\beta}_\perp)^{-1}\tilde{\alpha}_\perp'\hat{\mu}.$$

From the expression (24), we can see that

$$\begin{aligned} T\eta^{*\prime}\tilde{\beta}_{\perp,1} &= T\eta^{*\prime}(\tilde{\beta}_{\perp,1}-\beta_{\perp,1}) \\ &= -\eta^{*\prime}\beta_1(\tilde{\beta}'\beta)^{-1}\left[(\tilde{\beta}-\beta)'\gamma T, (\tilde{\beta}-\beta)'\tau T\right] \\ &\xrightarrow{d} -\eta^{*\prime}\beta_1(\beta'\beta)^{-1}U_1'[I_{n-r-1},0]. \end{aligned} \quad (41)$$

Next, from the definition of τ, we can see that

$$\bar{\gamma}'\tau = [I_{n-r-1},0](\alpha_\perp'\Gamma\beta_\perp)^{-1}\alpha_\perp'\mu.$$

Since the left-hand side is zero from the orthogonality between γ and τ, the first $n-r-1$ rows of $(\alpha_\perp'\Gamma\beta_\perp)^{-1}\alpha_\perp'\mu$ are zero. Then, because each estimator is consistent, we have

$$[I_{n-r-1},0](\tilde{\alpha}_\perp'\hat{\Gamma}\tilde{\beta}_\perp)^{-1}\tilde{\alpha}_\perp'\hat{\mu} \xrightarrow{p} 0. \quad (42)$$

Combining (41) and (42), we obtain

$$T\eta^{*\prime}\hat{\tau}_1 = (T\eta^{*\prime}\tilde{\beta}_{\perp,1})(\tilde{\alpha}_\perp'\hat{\Gamma}\tilde{\beta}_\perp)^{-1}\tilde{\alpha}_\perp'\hat{\mu} \xrightarrow{p} 0. \quad \square$$

Similar to the proof of Theorem 2, we consider the same determinant equation as (31) with $H=[\beta^*_{\perp,1}, T\eta^*]$. Using Lemma 3, we have

$$T\eta^{*\prime}[\tilde{\gamma}_1,\hat{\tau}_1] \xrightarrow{d} [Y',0] = Y'S_1,$$

where $S_1=[I_{n-r-1},0]$, and then, using $[\tilde{\gamma}_1,\hat{\tau}_1] \xrightarrow{p} [\gamma_1,\tau_1]=\beta_{\perp,1}$,

$$H'[\tilde{\gamma}_1,\hat{\tau}_1]\tilde{\ddot{\Psi}}[\tilde{\gamma}_1,\hat{\tau}_1]'H \xrightarrow{d} \begin{bmatrix} \beta^{*\prime}_{\perp,1}\beta_{\perp,1}\ddot{\Psi}\beta'_{\perp,1}\beta^*_{\perp,1} & \beta^{*\prime}_{\perp,1}\beta_{\perp,1}\ddot{\Psi}S_1'Y \\ Y'S_1\ddot{\Psi}\beta'_{\perp,1}\beta^*_{\perp,1} & Y'S_1\ddot{\Psi}S_1'Y \end{bmatrix}.$$

On the other hand, because $\beta^{*\prime}_{\perp,1}\tilde{\gamma}_1 \xrightarrow{p} \beta^{*\prime}_{\perp,1}\gamma_1$, $\beta^{*\prime}_{\perp,1}\hat{\tau}_1 \xrightarrow{p} \beta^{*\prime}_{\perp,1}\tau_1$, $\beta^{*\prime}_{\perp,1}\tilde{\beta}_1 \xrightarrow{p} \beta^{*\prime}_{\perp,1}\beta_1$, $T\eta^{*\prime}\tilde{\gamma}_1 = O_p(1)$, $T\eta^{*\prime}\hat{\tau}_1 = o_p(1)$, and $T\eta^{*\prime}\tilde{\beta}_1 = T\eta^{*\prime}\beta_1 + o_p(T)$ by Lemma 3, $\hat{\mu}H'\hat{\ddot{\Phi}}H$ is asymptotically equivalent to

$$T^2\hat{\mu}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \eta^{*\prime}\beta_1(\beta'\beta)^{-1}(\alpha'\Sigma^{-1}\alpha)^{-1}(\beta'\beta)^{-1}\beta_1'\eta^* \end{bmatrix}.$$

Then, for large values of T, the determinant equation (31) can be seen as

$$\begin{aligned}
& \left|\beta^{*\prime}_{\perp,1}\beta_{\perp,1}\ddot{\Psi}\beta'_{\perp,1}\beta^{*}_{\perp,1}\right| \\
\times & \left|Y'S_1\{\ddot{\Psi} - \ddot{\Psi}\beta'_{\perp,1}\beta^{*}_{\perp,1}(\beta^{*\prime}_{\perp,1}\beta_{\perp,1}\ddot{\Psi}\beta'_{\perp,1}\beta^{*}_{\perp,1})^{-1}\beta^{*\prime}_{\perp,1}\beta_{\perp,1}\ddot{\Psi}\}S_1'Y\right. \\
& \left. -T^2\hat{\mu}\eta^{*\prime}\beta_1(\beta'\beta)^{-1}(\alpha'\Sigma^{-1}\alpha)^{-1}(\beta'\beta)^{-1}\beta_1'\eta^*\right| \\
\propto & \left|Y'S_1JJ'S_1'Y - T^2\hat{\mu}\eta^{*\prime}\beta_1(\beta'\beta)^{-1}(\alpha'\Sigma^{-1}\alpha)^{-1}(\beta'\beta)^{-1}\beta_1'\eta^*\right| \\
= & \ 0,
\end{aligned} \tag{43}$$

where an $(n-r)\times(n-r-g)$ matrix J satisfies $J'\ddot{\Psi}^{-1}J = I_{n-r-g}$. Noting that the conditional variance of $Y'S_1J$ is given by

$$\eta^{*\prime}\beta_1(\beta'\beta)^{-1}(\alpha'\Sigma^{-1}\alpha)^{-1}(\beta'\beta)^{-1}\beta_1'\eta^* \otimes J'S_1'\Omega^{11}S_1J,$$

the test statistic conditioned on $G(\cdot)$ converges in distribution to

$$tr(Y^{*\prime}J'S_1'\Omega^{11}S_1JY^*) = tr(Y^{*\prime}J_1'\Omega^{11}J_1Y^*), \tag{44}$$

where $vec(Y^*) \sim N(0, I_{n_1-g} \otimes I_{n-r-g})$ and $J = [J_1', J_2']'$. Since

$$J'\ddot{\Psi}^{-1}J = J_1'\Omega^{11}J_1 + J_2'J_2 = I_{n-r-g},$$

the limiting distribution (44) is bounded above by

$$\begin{aligned}
tr(Y^{*\prime}J_1'\Omega^{11}J_1Y^*) &\leq tr(Y^{*\prime}(J_1'\Omega^{11}J_1 + J_2'J_2)Y^*) \\
&= tr(Y^{*\prime}Y^*) \sim \chi^2_{(n_1-g)(n-r-g)}.
\end{aligned}$$

This proves the statement of Theorem 6.□

References

[1] Bruneau, C. and E. Jondeau (1999) Long-Run Causality, with Application to International Links between Long-Term Interest Rates. *Oxford Bulletin of Economics and Statistics* 61, 545-568.

[2] Cragg, J. G. and S. G. Donald (1996) On the Asymptotic Properties of LDU-Based Tests of the Rank of a Matrix. *Journal of the American Statistical Association* 91, 1301-1309.

[3] Cragg, J. G. and S. G. Donald (1997) Inferring the Rank of a Matrix. *Journal of Econometrics* 76, 223-250.

[4] Dolado, J. J. and H. Lütkepohl (1996) Making Wald Tests Work for Cointegrated VAR Systems. *Econometric Reviews* 15, 369-386.

[5] Engle, R. F. and C. W. J. Granger (1987) Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica* 55, 251-276.

[6] Hubrich, K., H. Lütkepohl and P. Saikkonen (2001) A Review of Systems Cointegrating Tests. *Econometric Reviews* 20, 247-318.

[7] Johansen, S. (1988) Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamics and Control* 12, 231-254.

[8] Johansen, S. (1991) Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica* 59, 1551-1580.

[9] Johansen, S. (1992) A Representation of Vector Autoregressive Processes Integrated of Order 2. *Econometric Theory* 8, 188-202.

[10] Johansen, S. (1996) *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford: Oxford University Press.

[11] Kurozumi, E. and T. Yamamoto (2000) Modified Lag Augmented Vector Autoregressions. *Econometric Reviews* 19, 207-231.

[12] Lütkepohl, H. and P. Saikkonen (2000) Testing for the Cointegrating Rank of a VAR Process with a Time Trend. *Journal of Econometrics* 95, 177-198.

[13] Magnus, J. R. and H. Neudecker (1988) *Matrix Differential Calculus with Applications in Statistics and Econometrics*. New York: Wiley.

[14] Paruolo, P. (1997) Asymptotic Inference on the Moving Average Impact Matrix in Cointegrated I(1) VAR Systems. *Econometric Theory* 13, 79-118.

[15] Phillips, P. C. B. (1995) Fully Modified Least Squares and Vector Autoregression. *Econometrica* 63, 1023-1078.

[16] Phillips, P. C. B., and B. E. Hansen (1990) Statistical Inference in Instrumental Variables Regression with I(1) Processes. *Review of Economic Studies* 57, 99-125.

[17] Reinsel, G. C. and S. K. Ahn (1992) Vector Autoregressive Models with Unit Roots and Reduced Rank Structure: Estimation, Likelihood Ratio Test, and Forecasting. *Journal of Time Series Analysis* 13, 353-375.

[18] Robin, J-M. and R. J. Smith (2000) Test of Rank. *Econometric Theory* 16, 151-175.

[19] Saikkonen, P. and H. Lütkepohl (2000) Testing for the Cointegrating Rank of a VAR Process with an Intercept. *Econometric Theory* 16, 373-406.

[20] Toda, H. Y., and P. C. B. Phillips (1993) Vector Autoregression and Causality. *Econometrica* 61, 1367-1393.

[21] Toda, H. Y., and T. Yamamoto (1995) Statistical Inference in Vector Autoregressions with Possibly Integrated Processes. *Journal of Econometrics* 66, 225-250.

[22] Yamada, H., and H. Y. Toda (1997) A Note on Hypothesis Testing Based on the Fully Modified Vector Autoregression. *Economic Letters* 56, 27-39.

[23] Yamada, H., and H. Y. Toda (1998) Inference in Possibly Integrated Vector Autoregressive Models: Finite Sample Evidence. *Journal of Econometrics* 86, 55-95.

[24] Yamamoto, T. (2002) Finite Sample Properties of the Granger Non-Causality Test in Possibly Cointegrated Systems. manuscript.

[25] Yamamoto, T. and E. Kurozumi (2001) Finite Sample Properties of the Test for Long-Run Granger Non-Causality in Cointegrated Systems. *Proceedings of International Congress on Modelling and Simulation 2001*, Modelling and Simulation Society of Australia and New Zealand Inc.

[26] Yamamoto, T. and E. Kurozumi (2002) A Test for Long-Run Granger Non-Causality in Cointegrated Systems. manuscript.

Table 1. Critical values of the $T^2\hat{\lambda}_p$ statistic

	$r-f$				
$n_1 - f$	1	2	3	4	5
1	1.54×10^{-4}				
	0.00381				
	0.0156				
2	0.0197	6.27×10^{-5}			
	0.103	0.00157			
	0.210	0.00638			
3	0.114	0.0100	3.97×10^{-5}		
	0.348	0.0510	9.81×10^{-4}		
	0.581	0.105	0.00396		
4	0.294	0.0644	0.00669	2.81×10^{-5}	
	0.708	0.197	0.0342	7.09×10^{-4}	
	1.06	0.325	0.0704	0.00286	
5	0.553	0.182	0.0455	0.00504	2.16×10^{-5}
	1.14	0.431	0.138	0.0254	5.47×10^{-4}
	1.60	0.641	0.230	0.0523	0.00222

Table 2. Rejection frequencies of the tests

		$rk(\beta_1) = 0$	$rk(\beta_1) = 1$	$rk(\beta_{\perp,1}) = 0$	$rk(\beta\perp, 1) = 1$
Non-trending data with $d = 0$					
DGP1	$T = 100$	1.000	0.088	1.000	1.000
	$T = 200$	1.000	0.068	1.000	1.000
DGP2	$T = 100$	1.000	0.092	1.000	0.077
	$T = 200$	1.000	0.072	1.000	0.056
DGP3	$T = 100$	1.000	0.998	1.000	0.994
	$T = 200$	1.000	1.000	1.000	1.000
Non-trending data with $d = \alpha\rho_0$					
DGP1	$T = 100$	1.000	0.098	1.000	1.000
	$T = 200$	1.000	0.074	1.000	1.000
DGP2	$T = 100$	1.000	0.097	1.000	0.083
	$T = 200$	1.000	0.074	1.000	0.071
DGP3	$T = 100$	1.000	0.983	1.000	0.981
	$T = 200$	1.000	0.999	1.000	1.000
Trending data					
DGP1	$T = 100$	1.000	0.065	1.000	1.000
	$T = 200$	1.000	0.047	1.000	1.000
DGP2	$T = 100$	1.000	0.063	1.000	0.084
	$T = 200$	1.000	0.041	1.000	0.061
GDP3	$T = 100$	1.000	0.998	1.000	0.995
	$T = 200$	1.000	1.000	1.000	1.000

Table 3. Powers of the tests

	$rk(\beta_1) = 1$		$rk(\beta_{\perp,1}) = 1$	
$T =$	100	200	100	200
c_1 or c_2	Non-trending data with $d = 0$			
0.001	0.090	0.071	0.077	0.055
0.0025	0.095	0.084	0.074	0.068
0.005	0.107	0.122	0.084	0.103
0.0075	0.130	0.186	0.110	0.163
0.01	0.151	0.253	0.145	0.241
0.025	0.333	0.626	0.351	0.631
0.05	0.615	0.885	0.672	0.916
0.075	0.798	0.973	0.849	0.986
0.1	0.898	0.994	0.927	1.000
c_1 or c_2	Non-trending data with $d = \alpha\rho_0$			
0.001	0.099	0.067	0.086	0.069
0.0025	0.109	0.074	0.085	0.067
0.005	0.111	0.088	0.093	0.073
0.0075	0.114	0.100	0.097	0.093
0.01	0.115	0.121	0.100	0.111
0.025	0.182	0.348	0.167	0.352
0.05	0.361	0.728	0.383	0.743
0.075	0.558	0.912	0.588	0.926
0.1	0.747	0.975	0.765	0.980
c_1 or c_2	Trending data			
0.001	0.058	0.043	0.082	0.061
0.0025	0.061	0.044	0.082	0.060
0.005	0.062	0.053	0.087	0.058
0.0075	0.070	0.083	0.085	0.064
0.01	0.077	0.101	0.082	0.078
0.025	0.185	0.473	0.124	0.248
0.05	0.488	0.888	0.254	0.630
0.075	0.749	0.980	0.452	0.886
0.1	0.888	0.997	0.647	0.961