THE RANK OF A SUBMATRIX OF COINTEGRATION

EIJI KUROZUMI
Hitotsubashi University
and
Boston University

This paper proposes a test of the rank of the submatrix of β, where β is a cointegrating matrix. In addition, the submatrix of $\beta_\perp$, an orthogonal complement to β, is investigated. We construct the test statistic by using the eigenvalues of the quadratic form of the submatrix. We show that the test statistic has a limiting chi-square distribution when data are nontrending, whereas for trending data we have to consider a conservative test or other testing procedure that requires the pretest of the structure of the matrix. Finite sample simulations show that, although the simulation settings are limited, the proposed test works well for nontrending data, whereas we have to carefully use the test for trending data because it may become too conservative in some cases.

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 $t = 1, \ldots, T$, where $\{\varepsilon_t\}$ is independently and identically distributed (i.i.d.) with mean zero and a positive definite covariance matrix Σ and $\det(I_n - A_1 z - \cdots - A_m z^m)$ has all 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 rela-

299

tions. The exact condition of the existence of cointegration is given by Johansen (1991, 1992). We also assume that the cointegrating rank r is known or estimated by some testing procedure, such as the likelihood ratio (LR) 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 submatrix of β, and the rank of $\beta_{\perp,1}$, the submatrix 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 $a'\beta$ is of full rank. In such a situation, we would like to know whether the first r rows of β have full rank. The second example is the Granger noncausality test. As shown in Toda and Phillips (1993), when there is a cointegrating relationship, in general the Wald statistic of the Granger noncausality test from the last n_3 variables of x_t to the first n_1 variables has a nonstandard limiting distribution, depending on nuisance parameters. However, if either the last n_3 rows of β or the first n_1 rows of α have full row rank, the Wald statistic is asymptotically χ^2 distributed. Then, the testing procedure in this paper is useful to check the rank of the submatrix of β, whereas the existing testing procedure may be available for the test of the rank of the submatrix of α. The third example is the test of long-run Granger noncausality proposed by Yamamoto and Kurozumi (2001, 2003). In a usual sense, Granger causality is concerned with the one period ahead forecast. This concept can be extended to the predictability of h period ahead horizon, and long-run Granger causality is defined when the forecast horizon h goes to infinity. See, for example, Bruneau and Jondeau (1999) and Dufour and Renault (1998). Yamamoto and Kurozumi (2003) proposed the test for long-run block noncausality, in which it is shown that the ranks of the submatrices of β and $\beta_\perp$ play an important role in constructing the test statistic. See Yamamoto and Kurozumi (2003) for more details.

Tests of the rank of a matrix have been investigated in the literature, and recent econometric developments can be seen in works by Camba-Mendez, Kapetanios, Smith, and Weale (2003), Cragg and Donald (1996, 1997), and Robin and Smith (2000), among others. Although these papers proposed tests of the rank of a matrix, they assumed that the estimator of the matrix is $T^{1/2}$ consistent and has a limiting normal distribution with a nonstochastic variance matrix. However, the estimator of the cointegrating matrix is T (or $T^{3/2}$) consistent and has an asymptotic nonstandard distribution. As a result, we cannot apply existing testing procedures to the cointegrating matrix.

The paper is organized as follows. In Section 2, we propose tests of the rank of β_1 and $\beta_{\perp,1}$ for nontrending data. We will show that the two test statistics proposed have asymptotically a χ^2 distribution and a distribution of the maximum eigenvalue of the product of normal random matrices. Section 3 considers the case of trending data. In this case, the test statistics do not necessarily converge to a χ^2 distribution and a distribution of the maximum eigenvalue. To overcome this situation, we propose two testing procedures. Section 4 investigates the finite sample properties of the tests. Section 5 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, and $\xrightarrow{P}$, $\xrightarrow{d}$, 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 such as $\int_0^1 X(s)\,dY(s)'$ simply as $\int XdY'$ to achieve notational economy, and all integrals are from 0 to 1 except where otherwise noted.

2. TEST OF THE RANK OF THE SUBMATRIX FOR NONTRENDING DATA

2.1. The Model with *d* = 0

In this section we consider a test of rank for nontrending 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 maximum likelihood (ML) method assuming that $\{\varepsilon_t\}$ is Gaussian, although asymptotic properties are preserved under more general assumptions. We denote the ML estimator with ^. 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, 1995) showed that

$$T\beta'_{\perp}(\tilde{\beta} - \beta) \xrightarrow{d} \left(\int G_0 G_0' \, ds\right)^{-1} \int G_0 \, dV', \tag{4}$$

where $\tilde{\beta} = \hat{\beta}(\bar{\beta}'\hat{\beta})^{-1}$, $G_0(\cdot) = \bar{\beta}'_{\perp} CW(\cdot)$ 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 $\hat{\Gamma}_i$ ($i = 1,\ldots,m-1$) 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 ($0 < n_1 < n$). Similarly, we partition

$\beta'_\perp = [\beta'_{\perp,1}, \beta'_{\perp,2}]$ conformably. Note that $\beta'_1 \beta_{\perp,1}$ does not necessarily equal zero, whereas $\beta' \beta_\perp = \beta'_1 \beta_{\perp,1} + \beta'_2 \beta_{\perp,2}$ must be zero. Our interest lies in finding the rank of β_1, and thus we consider the following testing problem:

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

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 nonnegative 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.). Because 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 $T^{1/2}$, whereas $\hat{\beta}_1$ 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

$$\Phi = [\beta_1, \beta_{\perp,1}(\beta'_\perp \beta_\perp)^{-1}] \begin{bmatrix} (\beta'\beta)^{-1} & 0 \\ 0 & \left(\int G_0 G'_0 \, ds\right)^{-1} \end{bmatrix} \begin{bmatrix} \beta'_1 \\ (\beta'_\perp \beta_\perp)^{-1} \beta'_{\perp,1} \end{bmatrix}. \tag{6}$$

These Ψ and Φ are chosen so that the limiting distribution of the test statistic does not depend on nuisance parameters. Other choices of Φ may be possible because, as shown in the Appendix, the test statistic asymptotically does not depend on $\beta_1(\beta'\beta)^{-1}\beta'_1$, which appears when (6) is expanded. For example, we can use a constant multiple of $(\beta'\beta)^{-1}$ in the second term of (6). However, as indicated in the Appendix, Φ has to be invariant to the normalization of β. We use the definition (6) just because it seems simplest among other choices.

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 ML 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

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

$$= \hat{\beta}_1(\hat{\beta}'\hat{\beta})^{-1}\hat{\beta}_1' + \hat{\beta}_{\perp,1}\left(\frac{1}{T}\bar{\beta}_\perp'S_{11}\bar{\beta}_\perp\right)^{-1}\hat{\beta}_{\perp,1}', \tag{10}$$

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 $\triangle x_{t-1}, \ldots, \triangle x_{t-m+1}$, and we denote the ordered eigenvalues of (8) 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.), whereas 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 $n_1 > r$.

We can also consider the null hypothesis of $rk(\beta_1) = f$ against the alternative of $rk(\beta_1) = f + 1$. In this case, the test statistic is defined by

$$\mathcal{M}_T = T^2\hat{\lambda}_{f+1}.$$

To denote the limiting distribution of $\mathcal{M}_T$, we define $\lambda_{max,j,k}^*$ as the maximum eigenvalue of

$$|X^{*'}X^* - \lambda^*I_j| = 0, \tag{11}$$

where $X^{*'}$ is a $j \times k$ matrix with $\text{vec}(X^{*'}) \sim N(0, I_{j\times k})$. The critical points of this distribution are given in Table 1 for the case where $j \geq k$. They are calculated by simulations with 1,000,000 replications. Because the nonzero eigenvalues of $X^*X^{*'}$ are the same as those of $X^{*'}X^*$, we can refer to the percentage points of $\lambda_{max,k,j}$ when $j < k$.

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, $\mathcal{L}_T \xrightarrow{d} \chi^2_{(n_1-f)(r-f)}$ and $\mathcal{M}_T \xrightarrow{d} \lambda_{max,n_1-f,r-f}^*$.*

TABLE 1. Critical values of the $\lambda^*_{min,j,k}$ and $\lambda^*_{max,j,k}$

| | j | | | | | | | | | |
| | 1 | | 2 | | 3 | | 4 | | 5 | |
k	λ^*_{min}	λ^*_{max}	λ^*_{min}	λ^*_{max}	λ^*_{min}	λ^*_{max}	λ^*_{min}	λ^*_{max}	λ^*_{min}	λ^*_{max}
	1.54×10^{-4}	2.70	0.0197	4.61	0.114	6.25	0.294	7.78	0.553	9.23
1	0.00381	3.84	0.103	6.01	0.348	7.82	0.708	9.49	1.14	11.1
	0.0156	6.66	0.210	9.26	0.581	11.4	1.06	13.3	1.60	15.1
	—	—	6.27×10^{-5}	7.01	0.0100	9.00	0.0644	10.8	0.182	12.5
2	—	—	0.00157	8.61	0.0510	10.7	0.197	12.7	0.431	14.5
	—	—	0.00638	12.2	0.105	14.6	0.325	16.8	0.641	18.8
	—	—	—	—	3.97×10^{-5}	11.3	0.00669	13.3	0.0455	15.2
3	—	—	—	—	9.81×10^{-4}	13.1	0.0342	15.3	0.138	17.2
	—	—	—	—	0.00396	17.2	0.0704	19.5	0.230	21.7
	—	—	—	—	—	—	2.81×10^{-5}	15.4	0.00504	17.5
4	—	—	—	—	—	—	7.09×10^{-4}	17.5	0.0254	19.7
	—	—	—	—	—	—	0.00286	22.0	0.0523	24.3
	—	—	—	—	—	—	—	—	2.16×10^{-5}	19.6
5	—	—	—	—	—	—	—	—	5.47×10^{-4}	21.9
	—	—	—	—	—	—	—	—	0.00222	26.6

Note: The upper, middle, and lower entries are 1%, 5%, and 10% critical points for given j and k.

Remark 1. Because 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, $\hat{\lambda}_{f+1} \overset{d}{\to} \lambda_{f+1} > 0$ (a.s.), so that $T^2\hat{\lambda}_{f+1}$ goes to infinity. Therefore, the tests $\mathcal{L}_T$ and $\mathcal{M}_T$ are consistent.

Remark 2. Although the test statistics are constructed using the estimator of $\beta_{\perp,1}$, we do not have to assume that it is of full rank. We can see that the rank of $\beta_{\perp,1}$ is at least $n_1 - f$ under H_0, noting that the column space of $\beta_{\perp,1}$ must contain $n_1 - f$ bases that are orthogonal to $sp(\beta_1)$ because $[\beta_1, \beta_{\perp,1}]$ has full row rank n_1. Because $\beta_1' \beta_{\perp,1}$ is not necessarily equal to zero, it is possible for $sp(\beta_{\perp,1})$ to contain some of the bases that span $sp(\beta_1)$, so that the rank of $\beta_{\perp,1}$ may be greater than $n_1 - f$. It is shown in the Appendix that the limiting distributions of the test statistics depend not on the rank of $\beta_{\perp,1}$ but on the number of the bases orthogonal to $sp(\beta_1)$, $n_1 - f$, unless $f = n_1$. When $f = n_1$, all the eigenvalues are asymptotically greater than zero (a.s.), and then the test statistics will diverge. This case is excluded from the theorem (f is assumed to be less than $p = \min(n_1, r)$). In other words, our tests cannot be applied for the null hypothesis of full rank. If we need to check whether β_1 is of full rank or not, we may test for the null of $f = n_1 - 1$, and if we rejected the null hypothesis, we would conclude that it is a full row rank matrix.

Remark 3. Because the hypothesis about the rank of β_1 can be regarded as a restriction on the cointegrating matrix β, we may consider using the LR test as proposed by, for example, Johansen (1991, 1995) and Johansen and Juselius (1990, 1992). In fact, when $f = 0$ the null hypothesis is equivalent to $\beta_1 = 0$, and this hypothesis can be expressed as a linear restriction on β such as $\beta = H\phi$, where $H = [0, I_{n-n_1}]'$ and ϕ is an $(n - n_1) \times r$ unknown parameter. Then, the LR test is applicable to the test of $f = 0$. However, for $0 < f < p$, the null hypothesis is expressed as $\beta_1 = \beta_{11}\beta_{12}'$ where β_{11} and β_{12} are $n_1 \times f$ matrices with full column rank f. Then, we have to estimate the model with this restriction. Although the LR test might be applicable to the nonlinear hypothesis, it seems tedious to estimate the model with this nonlinear restriction, whereas our test uses only the ML estimator without the restriction. It is beyond our scope to investigate the applicability of the LR test to our case, and we do not discuss this in detail.

We may represent the null hypothesis as proposed by Boswijk (1996) and apply the LR test. According to his paper, the null hypothesis of $rk(\beta_1) = f$ is expressed as $\beta = (H_o\phi, \psi)$ where $H_o = [0, I_{n_2}]'$ and $(\phi, \psi) \in R^{n_2 \times (r-f)} \times R^{n \times f}$. As pointed out by Boswijk (1996, p. 156), the LR test for this hypothesis has an asymptotic χ^2 distribution only when "no linear combination of ψ lies in the column space of" H_o. Because there is no guarantee of this condition, we do not consider his method in this paper.

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

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

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:

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

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 (12) is given by

$$|\hat{\beta}_{\perp,1}\hat{\ddot{\Psi}}\hat{\beta}'_{\perp,1} - \hat{\mu}\hat{\ddot{\Phi}}| = 0, \tag{13}$$

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

$$\hat{\ddot{\Phi}} = [\hat{\beta}_{\perp,1}, \hat{\beta}_1(\hat{\beta}'\hat{\beta})^{-1}] \begin{bmatrix} (\hat{\beta}'_\perp \hat{\beta}_\perp)^{-1} & 0 \\ 0 & (\hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha})^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta}'_{\perp,1} \\ (\hat{\beta}'\hat{\beta})^{-1}\hat{\beta}'_1 \end{bmatrix}. \tag{14}$$

Let $\mu_1 \geq \cdots \geq \mu_{n_1}$ and $\hat{\mu}_1 \geq \cdots \geq \hat{\mu}_{n_1}$ be ordered eigenvalues of (12) and (13), respectively, and we construct the following test statistics, with $q = \min(n_1, n-r)$:

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

THEOREM 2. *Let* $\hat{\ddot{\Psi}} = T^{-1}\hat{\bar{\beta}}'_\perp S_{11}\hat{\bar{\beta}}_\perp$ *and* $\hat{\ddot{\Phi}}$ *be given by (14). If* $g < q$, *under* $H_{0\perp}$, $\mathcal{L}_{\perp T} \xrightarrow{d} \chi^2_{(n_1-g)(n-r-g)}$ *and* $\mathcal{M}_{\perp T} \xrightarrow{d} \lambda^*_{max, n_1-g, n-r-g}$.

Note that the consistency of the tests is shown in a similar way as in Remark 1. We also note that we cannot test the null of $rk(\beta_{\perp,1}) = q$ by a similar reason to that given in Remark 2.

Given the preceding 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 submatrix, as the cointegrating rank is selected sequentially using the test of the cointegrating rank. For example, to decide the rank of β_1, we first 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.

2.2. The Model with $d \neq 0$

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

$$\triangle x_t = \alpha\beta^{+\prime}x^+_{t-1} + \sum_{j=1}^{m-1} \Gamma_j \triangle x_{t-j} + \varepsilon_t, \tag{15}$$

where $\beta^+ = [\beta', \rho_0]'$ and $x^+_{t-1} = [x'_{t-1}, 1]'$. The ML estimator of β^+ can be obtained by the reduced rank regression of $\triangle x_t$ on x^+_{t-1} corrected for $\triangle x_{t-1}, \ldots, \triangle x_{t-m+1}$, and the estimator of the cointegrating matrix is the first n rows of $\hat{\beta}^+$.

To test the rank of the submatrix of β for the model (15), we use $\hat{\Phi}$ defined by

$$\hat{\Phi} = \hat{\beta}_1(\hat{\beta}'\hat{\beta})^{-1}\hat{\beta}'_1 + \hat{\beta}_{\perp,1}(\hat{\beta}'_\perp\hat{\beta}_\perp)^{-1}L'(Y'_T S^+_{11} Y_T)^{-1}L(\hat{\beta}'_\perp\hat{\beta}_\perp)^{-1}\hat{\beta}'_{\perp,1}, \tag{16}$$

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

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

and $S^+_{11} = T^{-1}\sum_{t=1}^{T} R^+_{1t}R^{+\prime}_{1t}$, with R^+_{1t} being the regression residual of x^+_{t-1} on $\triangle x_{t-1}, \ldots, \triangle x_{t-m+1}$.

THEOREM 3. *Consider the model (15) and let $\hat{\Psi} = \hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha}$ and $\hat{\Phi}$ be given by (16). If $f < p$, under H_0, $\mathcal{L}_T \overset{d}{\to} \chi^2_{(n_1-f)(r-f)}$ and $\mathcal{M}_T \overset{d}{\to} \lambda^*_{max, n_1-f, r-f}$.*

THEOREM 4. *Consider the model (15) and let $\hat{\hat{\Psi}} = \{L'(Y'_T S^+_{11} Y_T)^{-1}L\}^{-1}$ and $\hat{\Phi}$ be given by (14). If $g < q$, under $H_{0\perp}$, $\mathcal{L}_{\perp T} \overset{d}{\to} \chi^2_{(n_1-g)(n-r-g)}$ and $\mathcal{M}_{\perp T} \overset{d}{\to} \lambda^*_{max, n_1-g, n-r-g}$.*

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 (1995, 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. Because $sp(M) = sp(M_l)$ and they are orthogonal to $\hat{\beta}$, we can use M_l as $\hat{\beta}_\perp$.

3. THE TEST OF THE RANK OF THE SUBMATRIX FOR TRENDING DATA

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

$$x_t = C \sum_{i=1}^{t} \varepsilon_i + \tau t + C_1(L)(\varepsilon_t + d) + x_0^*, \tag{17}$$

where $C = \beta_\perp (\alpha_\perp' \Gamma \beta_\perp)^{-1} \alpha_\perp'$ as defined in Section 2.1, $\tau = Cd$, $C_1(L) = (C(L) - C(1))/(1 - L)$ 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, 1995) for more details. In this case, $\beta_\perp$ is decomposed to τ, the coefficient of a linear trend in (17), and γ, an $n \times (n - r - 1)$ matrix that is orthogonal to τ. We partition γ and τ into $[\gamma_1', \gamma_2']'$ and $[\tau_1', \tau_2']'$ in the same way as β. As shown in Chapter 13.2 of Johansen (1995), $\tilde{\beta}$ can be expressed as

$$\tilde{\beta} = \beta + \gamma(\gamma'\gamma)^{-1} U_{1T} + \frac{1}{T^{1/2}} \tau(\tau'\tau)^{-1} U_{2T}, \tag{18}$$

where

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

where $G(r) = [G_1'(r), G_2'(r)]'$ with $G_1(r) = G_0(r) - \int G_0 ds$, $G_0(r) = \bar{\gamma}' CW(r)$ and $G_2(r) = r - \frac{1}{2}$. 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 $\triangle x_{t-1}, \ldots, \triangle x_{t-m+1}$. Convergence of $\hat{\Omega}^{11}$ to Ω^{11} is proved in Lemma 2(iii) in the Appendix, whereas the consistency of other ML estimators, such as $\tilde{\alpha}$, $\hat{\Sigma}$, and $\hat{\Gamma}_i$, is shown by Johansen (1991, 1995).

In the following discussion, we will show that the limiting distribution of $\tilde{\beta}_1$ depends on whether the rank of $[\beta_1, \gamma_1]$ is $n_1 - 1$ or n_1, or equivalently, whether $\tau_2 = 0$ or not. We will propose two testing procedures to cope with this problem.

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^* as an orthogonal complement to β_1^*, so that $\delta^{*\prime}\beta_1^* = 0$. 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.

Let us consider the case where $r < n - 1$. Because $\tilde{\beta}_1$ is the first n_1 rows of $\tilde{\beta}$, it is expressed from (18) 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, because the $n \times n$ matrix $[\beta, \gamma, \tau]$ is of full rank, the first n_1 rows of this matrix, $[\beta_1, \gamma_1, \tau_1]$, must be of full row rank, which implies that $a'[\beta_1, \gamma_1, \tau_1] \neq 0$ for any nonzero 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

$$T^{3/2}\tau_1^{*\prime}\tilde{\beta}_1 = \tau_1^{*\prime}\tau_1(\tau'\tau)^{-1}(TU_{2T}) \xrightarrow{d} \tau_1^{*\prime}\tau_1(\tau'\tau)^{-1}U_2 = X_2', \quad \text{say,} \qquad (19)$$

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

$$T\delta_0^{*\prime}\tilde{\beta}_1 = \delta_0^{*\prime}\gamma_1(\gamma'\gamma)^{-1}(TU_{1T}) + O_p(T^{-1/2}) \xrightarrow{d} \delta_0^{*\prime}\gamma_1(\gamma'\gamma)^{-1}U_1 = X_1', \quad \text{say.}$$
$$\qquad (20)$$

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

$$T\delta^{*\prime}\tilde{\beta}_1 = \delta^{*\prime}\gamma_1(\gamma'\gamma)^{-1}(TU_{1T}) + O_p(T^{-1/2}) \xrightarrow{d} \delta^{*\prime}\gamma_1(\gamma'\gamma)^{-1}U_1 = X', \quad \text{say.}$$
$$\qquad (21)$$

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, whereas the nonexistence 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 $n_1 - 1$ when $r < n - 1$.

For further investigation, let us consider the case where the rank of $[\beta_1, \gamma_1]$ equals $n_1 - 1$. In this case, this matrix is expressed as $[\Theta_{11}, 0]$ by some nonsingular transformation from the right-hand side, where Θ_{11} is an $n_1 \times (n_1 - 1)$ matrix with rank $n_1 - 1$. Then, using the same nonsingular transformation, $[\beta, \gamma]$ becomes

$$\begin{bmatrix} \Theta_{11} & 0 \\ \Theta_{21} & \Theta_{22} \end{bmatrix}, \quad \text{say.} \tag{22}$$

Let τ_1^* be the orthogonal complement to the column space of Θ_{11}. Then, because $\tau_1^{*\prime} \Theta_1 = 0$ and using the expression (22), we can see that the $n \times 1$ vector $[\tau_1^{*\prime}, 0]'$ is orthogonal to $[\beta, \gamma]$. Therefore, in this case, the trend parameter τ, which is orthogonal to β and γ, is a constant multiple of $[\tau_1^{*\prime}, 0]'$. In other words, when $rk[\beta_1, \gamma_1] = n_1 - 1$, τ_2 must be equal to zero. Note that, because τ_1^* is orthogonal to $sp(\beta_1)$ and $sp(\gamma_1)$, it is essentially the same as τ_1^*.

On the other hand, when $\tau_2 = 0$, τ is expressed as $[\tau_1', 0]'$ and then $\tau_1'[\beta_1, \gamma_1]$ equals zero because $\tau'[\beta, \gamma] = 0$. This implies that the $n_1 \times (n - 1)$ matrix $[\beta_1, \gamma_1]$ does not have full row rank. Then, we have the following proposition.

PROPOSITION 1. *The rank of $[\beta_1, \gamma_1]$ is $n_1 - 1$ if and only if $\tau_2 = 0$.*

When $r = n - 1$, there is no γ, and in this case, $rk(\beta_1)$ must be $n_1 - 1$ or n_1. Then, under the null hypothesis of $rk(\beta_1) = n_1 - 1$, δ^* becomes an $n_1 \times 1$ vector, and we have

$$T^{3/2} \delta^{*\prime} \tilde{\beta}_1 = \delta^{*\prime} \tau_1 (\tau' \tau)^{-1} (TU_{2T}) \xrightarrow{d} X_3', \quad \text{say.} \tag{23}$$

In this case, the test statistics should be multiplied by T, that is, $T\mathcal{L}_T$ and $T\mathcal{M}_T$ are the appropriate test statistics.

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

$$\hat{\Phi} = [\hat{\beta}_1, \hat{\gamma}_1 (\hat{\gamma}' \hat{\gamma})^{-1}] \begin{bmatrix} (\hat{\beta}' \hat{\beta})^{-1} & 0 \\ 0 & \hat{\Omega}^{11} \end{bmatrix} \begin{bmatrix} \hat{\beta}_1' \\ (\hat{\gamma}' \hat{\gamma})^{-1} \hat{\gamma}_1' \end{bmatrix} \tag{24}$$

or

$$\hat{\Phi} = [\hat{\beta}_1, \hat{\gamma}_1 (\hat{\gamma}' \hat{\gamma})^{-1}, \hat{\tau}_1 (\hat{\tau}' \hat{\tau})^{-1}] \begin{bmatrix} (\hat{\beta}' \hat{\beta})^{-1} & 0 & 0 \\ 0 & \hat{\Omega}^{11} & 0 \\ 0 & 0 & 12 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1' \\ (\hat{\gamma}' \hat{\gamma})^{-1} \hat{\gamma}_1' \\ (\hat{\tau}' \hat{\tau})^{-1} \hat{\tau}_1' \end{bmatrix}. \tag{25}$$

THEOREM 5. *When $r < n - 1$,*

(i.a) Let $\hat{\Psi} = \hat{\alpha}' \hat{\Sigma}^{-1} \hat{\alpha}$ and $\hat{\Phi}$ be given by (24). If $rk([\beta_1, \gamma_1]) = n_1$ and $f < p$, under H_0, $\mathcal{L}_T \xrightarrow{d} \chi^2_{(n_1-f)(r-f)}$ and $\mathcal{M}_T \xrightarrow{d} \lambda^*_{max, n_1-f, r-f}$.

(i.b) *Let* $\hat{\Psi} = \hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha}$ *and* $\hat{\Phi}$ *be given by* (25). *If* $rk([\beta_1, \gamma_1]) = n_1$ *and* $f < p$, *under* H_0, $\mathcal{L}_T$ *and* $\mathcal{M}_T$ *converge in distribution to random variables that are bounded above by* $\chi^2_{(n_1-f)(r-f)}$ *and* $\lambda^*_{max, n_1-f, r-f}$, *respectively.*

(ii) *Let* $\hat{\Psi} = \hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha}$ *and* $\hat{\Phi}$ *be given by* (25). *If* $rk([\beta_1, \gamma_1]) = n_1 - 1$ *and* $f < p$, *under* H_0, $\mathcal{L}_T \overset{d}{\to} \chi^2_{(n_1-f-1)(r-f)}$ *and* $\mathcal{M}_T \overset{d}{\to} \lambda^*_{max, n_1-f-1, r-f}$.

When $r = n - 1$,

(iii) *Let* $\hat{\Psi} = \hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha}$ *and* $\hat{\Phi}$ *be given by* (25). *Under the null hypothesis of* $f = n_1 - 1$,

$$T\mathcal{L}_T = T\mathcal{M}_T = T^3\hat{\lambda}_{n_1} \overset{d}{\to} \chi^2_{n-n_1}.$$

Remark 4. In the case of (i.b), $\mathcal{L}_T$ 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 (25) is conservative if $rk([\beta_1, \gamma_1]) = n_1$.

From Theorem 5, if we know the rank of $[\beta_1, \gamma_1]$ when $r < n - 1$, we can construct the test statistic $\mathcal{L}_T$ that converges to a χ^2 distribution by appropriately using (24) or (25). However, such information is not available in practice. Notice that if $rk[\beta_1, \gamma_1] = n_1 - 1$, $\hat{\Phi}$ given by (24) 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 (24) is not desirable in practice. On the other hand, if we use $\hat{\Phi}$ given by (25), 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 degrees of freedom may change depending on the rank of $[\beta_1, \gamma_1]$. Then, noting that the critical value of $\chi^2_{(n_1-f)(r-f)}$ in Theorem 5(i) is greater than that of $\chi^2_{(n_1-f-1)(r-f)}$ in (ii), we propose to test the null of $rk(\beta_1) = f$ as follows.

1. We construct the test statistic $\mathcal{L}_T$ using (25).
2. If $\mathcal{L}_T$ is greater than the critical value of $\chi^2_{(n_1-f)(r-f)}$, we reject the null hypothesis.
3. If $\mathcal{L}_T$ is less than the critical value of $\chi^2_{(n_1-f-1)(r-f)}$, we accept the null hypothesis.

The test statistic $\mathcal{M}_T$ is used in the same manner. In this procedure, we may encounter the case where the test statistic is greater than the critical value of $\chi^2_{(n_1-f-1)(r-f)}$ but less 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 (25). Suppose that $r < n - 1$ and 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 $\lambda^*_{min, r-f, n_1-f}$, which is the smallest nonzero eigenvalue of (11) with $j = r - f$ and $k = n_1 - f$.*

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

The percentage points of $\lambda^*_{min, r-f, n_1-f}$ are tabulated in Table 1.

Using the preceding 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 less than some percentage (10, 5, or 1%) point of $\lambda^*_{min, r-f, n_1-f}$, 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 greater than the critical point of $\lambda^*_{min, r-f, n_1-f}$, we accept the hypothesis of $rk([\beta_1, \gamma_1]) = n_1$, so that the rank of β_1 is decided to be f. We call this testing procedure TEST1.

The other strategy is to use the result of Proposition 1. From Johansen (1995), $T^{1/2}(\hat{\tau} - \tau)$ converges in distribution to a normal random vector with mean zero and the variance matrix given by $C\Sigma C'$. Although the Wald-type test may not be applicable to the test of $\tau_2 = 0$ because the variance matrix might be degenerate, we can test whether each element of τ_2 is zero or not by the t-test statistic. We call the following testing procedure TEST2.

1. We test each element of τ_2.
2. If some of the elements of τ_2 are significant, we use Theorem 5(i.a).
3. If none of the elements of τ_2 are significant, we use Theorem 5(ii).

Next, we investigate a test of the rank of $\beta_{\perp,1}$. When data are 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 (13) with $\hat{\beta}_{\perp,1}$ replaced by $[\hat{\gamma}_1, \hat{\tau}_1]$ and

$$\hat{\hat{\Psi}}^{-1} = \begin{bmatrix} \hat{\Omega}^{11} & 0 \\ 0 & 12 \end{bmatrix}, \tag{26}$$

$$\hat{\Phi} = [\hat{\gamma}_1, \hat{\tau}_1, \hat{\beta}_1(\hat{\beta}'\hat{\beta})^{-1}] \begin{bmatrix} (\hat{\gamma}'\hat{\gamma})^{-1} & 0 & 0 \\ 0 & (\hat{\tau}'\hat{\tau})^{-2} & 0 \\ 0 & 0 & (\hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha})^{-1} \end{bmatrix} \begin{bmatrix} \hat{\gamma}'_1 \\ \hat{\tau}'_1 \\ (\hat{\beta}'\hat{\beta})^{-1}\hat{\beta}'_1 \end{bmatrix}. \tag{27}$$

We construct the test statistics $\mathcal{L}_{\perp T}$ and $\mathcal{M}_{\perp T}$ in the same way as in the previous section. Similar to Theorem 5, we have to distinguish two cases where $r < n - 1$ and $r = n - 1$. When $r = n - 1$, the rank of $\beta_{1,1}$ $(= \tau_1)$ must be 0 or 1, and in this case, we consider the null hypothesis of $g = 0$.

THEOREM 6. *Let $\hat{\hat{\Psi}}$ and $\hat{\hat{\Phi}}$ be given by (26) and (27). When $r < n - 1$ and $g < q$, under $H_{0\perp}$, $\mathcal{L}_{\perp T}$ and $\mathcal{M}_{\perp,T}$ converge in distribution to random variables that are bounded above by $\chi^2_{(n_1-g)(n-g-r)}$ and $\lambda^*_{max, n_1-g, n-g-r}$, respectively.*
When $r = n - 1$, under the null hypothesis of $g = 0$,

$$T\mathcal{L}_{\perp T} = T\mathcal{M}_{\perp T} = T^3 \hat{\mu}_1 \xrightarrow{d} \chi^2_{n_1}.$$

4. 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 EC model as a data generating process (DGP):

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

where $\{\varepsilon_t\} \sim i.i.d.N(0, I_4)$. Let

$$a_1 = \begin{bmatrix} 0.3 \\ -0.3 \\ -0.8 \\ 0.8 \end{bmatrix}, \quad a_2 = \begin{bmatrix} -0.5 \\ 0 \\ -0.3 \\ -0.5 \end{bmatrix}, \quad a_3 = \begin{bmatrix} 0.8 \\ 1 \\ -0.5 \\ -0.5 \end{bmatrix}, \quad a_4 = \begin{bmatrix} -0.5 \\ -0.8 \\ 0 \\ -0.5 \end{bmatrix},$$

$$b_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -0.5 \end{bmatrix}, \quad b_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \quad b_3 = \begin{bmatrix} 0 \\ 0 \\ 0.5 \\ 1 \end{bmatrix}, \quad b_4 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix},$$

$$d_1 = \begin{bmatrix} -0.5 \\ 1 \\ 0.5 \\ -1 \end{bmatrix}, \quad d_2 = \begin{bmatrix} 0.6 \\ 0.3 \\ -0.4 \\ 0.4 \end{bmatrix},$$

and we consider the following settings of parameters.

Parameter Sets for the Test of $rk(\beta_1)$			Parameter Sets for the Test of $rk(\beta_{1,1})$		
α	β	$\beta_\perp$	α	β	$\beta_\perp$
DGP1 a_1	b_1	$[b_2,b_3,b_4]$	DGP1o a_2	b_2	$[b_3,b_4,b_1]$
DGP2 $[a_1,a_2]$	$[b_1,b_2]$	$[b_3,b_4]$	DGP2o $[a_1,a_2]$	$[b_1,b_2]$	$[b_3,b_4]$
DGP3 $[a_1,a_2,a_3]$	$[b_1,b_2,b_3]$	b_4	DGP3o $[a_1,a_2,a_4]$	$[b_1,b_2,b_4]$	b_3

Here DGP1(1o), 2(2o), and 3(3o) correspond to the cases where the cointegrating rank is 1, 2, and 3, respectively. We set the (2,1) element of β as c_1, which takes values of 0, 0.005, 0.01, 0.025, 0.05, 0.075, and 0.1, and we consider the test of the rank of the first two rows of β. The case of $c_1 = 0$ corresponds to the null hypothesis under which the rank of β_1 is 0, 1, and 1 for DGP1, 2, and 3, whereas it is 1, 2, and 2 when $c_1 \neq 0$, which corresponds to the alternative. For the case of nontrending data, we set $d = 0$ for the zero-mean process, whereas d is defined as $\alpha\rho_0$ for the case of $d \neq 0$, where ρ_0 is set to be 1, $[1,1]'$, and $[1,1,1]'$ for DGP1(1o), 2(2o), and 3(3o), respectively. On the other hand, for the case of trending data, d is set to be d_1 and d_2; the former corresponds to the case where $[\beta_1,\gamma_1]$ is of full rank ($\tau_2 \neq 0$), whereas the rank of $[\beta_1,\gamma_1]$ is $n_1 - 1$ ($\tau_2 = 0$) when $d = d_2$.

Similarly, we set the (2,1) element of $\beta_\perp$ as c_2 and consider the test of the rank of the first two rows of $\beta_\perp$. In this case, $c_2 = 0$ implies that the rank of $\beta_{\perp,1}$ is 1, 1, and 0 for DGPo1, o2, and o3, respectively, whereas it is 2, 2, and 1 under the alternative of $c_2 \neq 0$.

We set $x_0 = 0$ and discard the first 100 observations in all experiments. The number of replication is 5,000, and the level of significance is set equal to 0.05. We only report the results of the test statistics $\mathcal{L}_T$ and $\mathcal{L}_{\perp T}$ because the performances of $\mathcal{M}_T$ and $\mathcal{M}_{\perp T}$ are almost the same as those of $\mathcal{L}_T$ and $\mathcal{L}_{\perp T}$.

Table 2 shows the simulation results of the test of $rk(\beta_1)$. When the cointegrating rank is 1, the empirical size is greater than the nominal size, 0.05, for $T = 100$ when data are nontrending ($d = 0$ or $d = \alpha\rho_0$), whereas it becomes closer to 0.05 for $T = 200$. When data are trending, τ becomes $[-\frac{1}{4},\frac{3}{4},-\frac{1}{6},-\frac{1}{3}]$ for $d = d_1$, whereas it is $[\frac{9}{20},\frac{9}{20},0,0]$ for $d = d_2$. Similar to the case of nontrending data, the testing procedure TEST2 tends to overly reject the null of $c_1 = 0$ for $T = 100$, whereas the testing procedure TEST1 seems to be slightly conservative. Under the alternative of $c_1 \neq 0$, the power increases rapidly around $c_1 = 0.025$ for nontrending data and for trending data with TEST2, whereas the testing procedure TEST1 seems to be less powerful. This is because TEST1 is a conservative test. When data are trending, both TEST1 and TEST2 are more powerful for the model with $rk([\beta_1,\gamma_1]) = n_1$ ($d = d_1$) than the model with $rk([\beta_1,\gamma_1]) = n_1 - 1$ ($d = d_2$).

When the cointegrating rank is 2, the relative performance is preserved for the cases of $d = 0$ and $d = \alpha\rho_0$. For trending data, τ becomes $[\frac{61}{84}, \frac{61}{84}, \frac{43}{126}, \frac{43}{63}]$ and $[\frac{9}{20}, \frac{9}{20}, 0, 0]$ for $d = d_1$ and d_2, respectively. Note that $T^2\hat{\lambda}_p$ is numerically equal to $\mathcal{L}_T$ because the determinant equation (11) with $j = k = 1$ has only one eigenvalue. Then, we can see that $\mathcal{L}_T$ and $T^2\hat{\lambda}_p$ converge in distri-

TABLE 2. Rejection frequencies of the tests of $rk(\beta_1)$

$r = 1$	c_1	$d = 0$	$d = \alpha\rho_0$	TEST1(d_1)	TEST2(d_1)	TEST1(d_2)	TEST2(d_2)
$T = 100$	0.000	0.099	0.118	0.046	0.121	0.040	0.100
	0.005	0.119	0.122	0.047	0.137	0.039	0.104
	0.010	0.179	0.138	0.053	0.171	0.041	0.112
	0.025	0.423	0.222	0.072	0.424	0.065	0.170
	0.050	0.732	0.474	0.222	0.804	0.179	0.359
	0.075	0.886	0.699	0.575	0.952	0.381	0.624
	0.100	0.960	0.849	0.886	0.989	0.653	0.862
$T = 200$	0.000	0.074	0.076	0.028	0.084	0.028	0.080
	0.005	0.141	0.097	0.030	0.143	0.031	0.089
	0.010	0.297	0.156	0.040	0.302	0.041	0.122
	0.025	0.703	0.457	0.169	0.814	0.137	0.329
	0.050	0.950	0.845	0.883	0.992	0.615	0.847
	0.075	0.992	0.971	0.998	1.000	0.966	0.998
	0.100	0.999	0.995	1.000	1.000	0.999	1.000

$r = 2$	c_1	$d = 0$	$d = \alpha\rho_0$	TEST1(d_1)	TEST2(d_1)	TEST1(d_2)	TEST2(d_2)
$T = 100$	0.000	0.088	0.097	0.023	0.095	0.001	0.009
	0.005	0.105	0.100	0.026	0.102	0.000	0.011
	0.010	0.150	0.109	0.030	0.123	0.001	0.014
	0.025	0.338	0.176	0.065	0.251	0.001	0.033
	0.050	0.618	0.353	0.230	0.538	0.005	0.069
	0.075	0.800	0.559	0.533	0.755	0.042	0.101
	0.100	0.899	0.721	0.781	0.873	0.220	0.124
$T = 200$	0.000	0.069	0.070	0.012	0.077	0.000	0.007
	0.005	0.115	0.081	0.018	0.095	0.000	0.012
	0.010	0.227	0.115	0.034	0.161	0.000	0.028
	0.025	0.593	0.348	0.203	0.541	0.000	0.071
	0.050	0.889	0.728	0.799	0.913	0.084	0.095
	0.075	0.976	0.914	0.975	0.989	0.809	0.115
	0.100	0.995	0.972	0.998	0.999	0.990	0.145

$r = 3$	c_1	$d = 0$	$d = \alpha\rho_0$	$d = d_1$	—	$d = d_2$	—
$T = 100$	0.000	0.090	0.094	0.101	—	0.055	—
	0.005	0.680	0.249	0.961	—	0.179	—
	0.010	0.902	0.551	0.998	—	0.399	—
	0.025	0.995	0.938	1.000	—	0.708	—
	0.050	1.000	0.996	1.000	—	0.850	—
	0.075	1.000	1.000	1.000	—	0.904	—
	0.100	1.000	1.000	1.000	—	0.927	—
$T = 200$	0.000	0.064	0.068	0.076	—	0.045	—
	0.005	0.879	0.555	1.000	—	0.474	—
	0.010	0.983	0.901	1.000	—	0.728	—
	0.025	1.000	0.999	1.000	—	0.893	—
	0.050	1.000	1.000	1.000	—	0.944	—
	0.075	1.000	1.000	1.000	—	0.960	—
	0.100	1.000	1.000	1.000	—	0.969	—

bution to χ_1^2 under H_0 when $rk([\beta_1, \gamma_1]) = n_1 = 2$, whereas they converge in probability to zero when $rk([\beta_1, \gamma_1]) = n_1 - 1 = 1$. Then, the testing procedure TEST1 accepts the null hypothesis when it is less than the critical point of χ_1^2. On the other hand, the asymptotic size of TEST1 becomes 0 when $rk([\beta_1, \gamma_1]) = n_1 - 1$ ($d = d_2$) because $\mathcal{L}_T$ converges in probability to zero when $rk([\beta_1, \gamma_1]) = n_1 - 1$. Reflecting this fact, TEST1 is too conservative for $d = d_2$, and it is not powerful when the alternative is close to the null. TEST2 also seems to have no power when the rank of $[\beta_1, \gamma_1] = n_1 - 1 = 1$ ($d = d_2$). This is because τ_2 is very close to zero[1] even under the alternative of $c_1 \neq 0$, so that the pretest of τ_2 cannot reject the null of $\tau_2 = 0$. When τ_2 is judged to be zero, the rank of $[\beta_1, \gamma_1]$ is at most $n_1 - 1 = 1$. Then, because we are testing the null hypothesis of $rk(\beta_1) = f = 1$, we automatically accept the hypothesis when $\tau_2 = 0$ is accepted in this case.

When the cointegrating rank is 3, we can see that the first two variables of x_t are cointegrated, whereas the last two variables are stationary. Note that we cannot generate the process such that the rank of β_1 is 1 while all the variables are nonstationary. Because we want to investigate the property of the test under the null hypothesis, we allowed several variables to be stationary.

In this case, the power property seems to be improved for all the cases compared with the cases where $r = 1$ and 2. For trending data, τ becomes $\left[\frac{1,215}{424}, \frac{1,215}{424}, 0, 0\right]$ for $d = d_1$, whereas it is $\left[\frac{9}{20}, \frac{9}{20}, 0, 0\right]$ for $d = d_2$. Note that in this case the last two rows of the impact matrix C become zero because the corresponding variables are stationary, so that τ_2, the last two rows of Cd, become zero irrespective of the value of d. We also note that the result of Theorem 5(iii) is applied because $r = n - 1 = 3$. That is, we do not have to use the conservative test or the pretest as in the cases where $r < n - 1$. This is the reason why both the size and power properties are improved for trending data compared with the cases where $r < 3$.

Table 3 reports the results of the test of $rk(\beta_{\perp,1})$. From the table, the test tends to overly reject the null hypothesis for several cases when $T = 100$, whereas the size becomes reasonable when $T = 200$, except for the case where $r = 3$ and $d = d_1$. In that case, the test becomes conservative as investigated in Theorem 6. As to the power, we can see that the more complicated the deterministic term becomes, the less powerful is the test.

5. CONCLUSION

In this paper, we proposed tests of the rank of the submatrix of cointegration. We can test the hypothesis straightforwardly when data are nontrending, whereas for trending data, we have to examine whether $[\beta_1, \gamma_1]$ is of full rank or not or we have to use the conservative test. The simulation results show that we have to carefully use the test of $rk(\beta_1)$ when data are trending and $f = n_1 - 1$, because the test might become too conservative to reject the null hypothesis.

TABLE 3. Rejection frequencies of the tests of $rk(\beta_{\perp,1})$

$r = 1$	c_2	$d = 0$	$d = \alpha\rho_0$	$d = d_1$	$d = d_2$
$T = 100$	0.000	0.106	0.128	0.106	0.051
	0.005	0.121	0.135	0.107	0.050
	0.010	0.165	0.140	0.114	0.054
	0.025	0.374	0.200	0.141	0.081
	0.050	0.658	0.389	0.253	0.180
	0.075	0.826	0.599	0.417	0.338
	0.100	0.909	0.750	0.591	0.522
$T = 200$	0.000	0.072	0.082	0.065	0.031
	0.005	0.114	0.089	0.070	0.035
	0.010	0.230	0.126	0.084	0.047
	0.025	0.600	0.351	0.205	0.136
	0.050	0.903	0.752	0.579	0.520
	0.075	0.982	0.923	0.846	0.846
	0.100	0.996	0.978	0.954	0.960

$r = 2$	c_2	$d = 0$	$d = \alpha\rho_0$	$d = d_1$	$d = d_2$
$T = 100$	0.000	0.083	0.091	0.051	0.056
	0.005	0.098	0.097	0.052	0.056
	0.010	0.148	0.109	0.057	0.060
	0.025	0.369	0.176	0.087	0.106
	0.050	0.676	0.397	0.194	0.268
	0.075	0.845	0.624	0.367	0.478
	0.100	0.927	0.777	0.572	0.672
$T = 200$	0.000	0.062	0.065	0.036	0.046
	0.005	0.107	0.079	0.040	0.056
	0.010	0.237	0.119	0.053	0.082
	0.025	0.627	0.377	0.164	0.288
	0.050	0.916	0.772	0.564	0.734
	0.075	0.987	0.940	0.854	0.938
	0.100	0.998	0.985	0.959	0.986

$r = 3$	c_2	$d = 0$	$d = \alpha\rho_0$	$d = d_1$	$d = d_2$
$T = 100$	0.000	0.074	0.094	0.024	0.072
	0.005	0.125	0.102	0.026	0.096
	0.010	0.258	0.119	0.031	0.171
	0.025	0.636	0.234	0.074	0.606
	0.050	0.883	0.539	0.219	0.915
	0.075	0.962	0.760	0.340	0.972
	0.100	0.985	0.885	0.417	0.982
$T = 200$	0.000	0.063	0.064	0.018	0.065
	0.005	0.187	0.085	0.027	0.255
	0.010	0.440	0.146	0.049	0.701
	0.025	0.839	0.497	0.226	0.995
	0.050	0.979	0.879	0.447	1.000
	0.075	0.998	0.976	0.539	1.000
	0.100	1.000	0.996	0.579	1.000

NOTE

1. For example, even when $c_1 = 0.1$, the third and fourth elements of τ are $-\frac{84}{2,150}$ and $\frac{3}{430}$.

REFERENCES

Boswijk, H.P. (1996) Testing identifiability of cointegrating vectors. *Journal of Business & Economic Statistics* 14, 153–160.

Bruneau, C. & 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.

Camba-Mendez, G., G. Kapetanios, R.J. Smith, & M.R. Weale (2003) Tests of rank in reduced rank regression models. *Journal of Business & Economic Statistics* 21, 145–155.

Cragg, J.G. & 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.

Cragg, J.G. & S.G. Donald (1997) Inferring the rank of a matrix. *Journal of Econometrics* 76, 223–250.

Dufour, J.M. & E. Renault (1998) Short run and long run causality in time series: Theory. *Econometrica* 66, 1099–1125.

Engle, R.F. & C.W.J. Granger (1987) Co-integration and error correction: Representation, estimation, and testing. *Econometrica* 55, 251–276.

Hubrich, K., H. Lütkepohl, & P. Saikkonen (2001) A review of systems cointegrating tests. *Econometric Reviews* 20, 247–318.

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

Johansen, S. (1991) Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica* 59, 1551–1580.

Johansen, S. (1992) A representation of vector autoregressive processes integrated of order 2. *Econometric Theory* 8, 188–202.

Johansen, S. (1995) *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press.

Johansen, S. & K. Juselius (1990) Maximum likelihood estimation and inference on cointegration: With applications to the demand for money. *Oxford Bulletin of Economics and Statistics* 52, 169–210.

Johansen, S. & K. Juselius (1992) Testing structural hypotheses in a multivariate cointegration analysis of the PPP and UIP for UK. *Journal of Econometrics* 53, 211–244.

Lütkepohl, H. & P. Saikkonen (2000) Testing for the cointegrating rank of a VAR process with a time trend. *Journal of Econometrics* 95, 177–198.

Magnus, J.R. & H. Neudecker (1988) *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Wiley.

Paruolo, P. (1997) Asymptotic inference on the moving average impact matrix in cointegrated I(1) VAR systems. *Econometric Theory* 13, 79–118.

Robin, J.-M. & R.J. Smith (2000) Test of rank. *Econometric Theory* 16, 151–175.

Saikkonen, P. & H. Lütkepohl (2000) Testing for the cointegrating rank of a VAR process with an intercept. *Econometric Theory* 16, 373–406.

Toda, H.Y. & P.C.B. Phillips (1993) Vector autoregression and causality. *Econometrica* 61, 1367–1393.

Yamamoto, T. & E. Kurozumi (2001) Finite sample properties of the test for long-run Granger non-causality in cointegrated systems. In F. Ghassemi et al. (ed.), *Proceedings of International Congress on Modelling and Simulation 2001*. Modelling and Simulation Society of Australia and New Zealand.

Yamamoto, T. & E. Kurozumi (2003) Tests for Long-Run Granger Non-Causality in Cointegrated Systems. Discussion paper 2003-12, Graduate School of Economics, Hitotsubashi University.

APPENDIX

We use the notation H 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 nonsingular transformation of the columns of $\hat{\beta}$ 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

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

whose columns 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 nonsingular transformation of the columns of $\hat{\beta}_\perp$. Then, we can also replace $\hat{\beta}_\perp$ by $\tilde{\beta}_\perp$ in $\hat{\Phi}$.

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 $\delta^{*'}\beta_1^* = 0$.

LEMMA 1.

(i) $\tilde{\beta} \xrightarrow{P} \beta$, $\tilde{\alpha} \xrightarrow{P} \alpha$, $\tilde{\Sigma} \xrightarrow{P} \Sigma$.
(ii) $T\delta^{*'}\tilde{\beta}_1 = T\delta^{*'}(\tilde{\beta}_1 - \beta_1) \xrightarrow{d} \delta^{*'}\beta_{\perp,1}(\beta_\perp'\beta_\perp)^{-1}(\int G_0 G_0' \, ds)^{-1}\int G_0 \, dV' = X_0'$, say.
(iii) $T(\tilde{\beta}_\perp - \beta_\perp) \xrightarrow{d} -\beta(\beta'\beta)^{-1}\int dV \, G_0'(\int G_0 G_0' \, ds)^{-1}$.
(iv) $T^{-1}\bar{\tilde{\beta}}_\perp' S_{11}\bar{\tilde{\beta}}_\perp \xrightarrow{d} \int G_0 G_0' \, ds$.

Proof.

(i) Proved by Johansen (1988, 1995).
(ii) As shown in Chapter 13.2 of Johansen (1995), $\tilde{\beta}$ can be expressed as $\tilde{\beta} = \beta + \beta_\perp(\beta_\perp'\beta_\perp)^{-1}U_T$ for nontrending data, where TU_T converges in distribution to $(\int G_0 G_0' \, ds)^{-1}\int G_0 \, dV'$. Because $\tilde{\beta}_1$ 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

$$T\delta^{*'}\tilde{\beta}_1 = T\delta^{*'}(\tilde{\beta}_1 - \beta_1) \xrightarrow{d} \delta^{*'}\beta_{\perp,1}(\beta_\perp'\beta_\perp)^{-1}\left(\int G_0 G_0' \, ds\right)^{-1}\int G_0 \, dV'.$$

(iii) 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 (A.1).
(iv) is proved by noting that $T^{-1}\bar{\beta}_\perp' S_{11}\bar{\beta}_\perp \xrightarrow{d} \int G_0 G_0' \, ds$ from Johansen (1988, 1995), $(\bar{\tilde{\beta}}_\perp - \bar{\beta}_\perp) = O_p(T^{-1})$, and $S_{11} = O_p(T)$. ∎

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

$$|H'||\tilde{\beta}_1\tilde{\Psi}\tilde{\beta}_1' - \hat{\lambda}\hat{\Phi}||H| = 0, \tag{A.2}$$

where $H = [\beta_1^*, T\delta^*]$ is an $n \times n$ nonsingular matrix. Using Lemma 1, we have

$$H'\tilde{\beta}_1 \tilde{\Psi} \tilde{\beta}_1' H = \begin{bmatrix} \beta_1^{*'} \tilde{\beta}_1 \tilde{\Psi} \tilde{\beta}_1' \beta_1^* & \beta_1^{*'} \tilde{\beta}_1 \tilde{\Psi} (\tilde{\beta}_1' \delta^* T) \\ (T\delta^{*'} \tilde{\beta}_1) \tilde{\Psi} \tilde{\beta}_1' \beta_1^* & (T\delta^{*'} \tilde{\beta}_1) \tilde{\Psi} (\tilde{\beta}_1' \delta^* T) \end{bmatrix}$$

$$\xrightarrow{d} \begin{bmatrix} \beta_1^{*'} \beta_1 \Psi \beta_1' \beta_1^* & \beta_1^{*'} \beta_1 \Psi X_0 \\ X_0' \Psi \beta_1' \beta_1^* & X_0' \Psi X_0 \end{bmatrix}. \tag{A.3}$$

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

$$H'[\tilde{\beta}_1, \tilde{\beta}_{\perp,1}(\tilde{\beta}_\perp' \tilde{\beta}_\perp)^{-1}] = \begin{bmatrix} O_p(1) & O_p(1) \\ O_p(1) & T\delta^{*'}\beta_{\perp,1}(\beta_\perp' \beta_\perp)^{-1} + o_p(T) \end{bmatrix}$$

because $\beta_1^{*'} \tilde{\beta}_1 \xrightarrow{p} \beta_1^{*'} \beta_1$, $\beta_1^{*'} \tilde{\beta}_{\perp,1} \xrightarrow{p} \beta_1^{*'} \beta_{\perp,1}$, $T\delta^{*'}\tilde{\beta}_1 = O_p(1)$, and $T\delta^{*'}\tilde{\beta}_{\perp,1} = T\delta^{*'}\beta_{\perp,1} + o_p(T)$ by Lemma 1. Then, $\hat{\lambda} H'\tilde{\Phi} H$ is asymptotically equivalent to

$$T^2 \hat{\lambda} \begin{bmatrix} 0 & 0 \\ 0 & \delta^{*'}\beta_{\perp,1}(\beta_\perp' \beta_\perp)^{-1} \left(\int G_0 G_0' \, ds \right)^{-1} (\beta_\perp' \beta_\perp)^{-1} \beta_{\perp,1}' \delta^* \end{bmatrix}.$$

Then, the equation (A.2) is asymptotically equal to

$$\left\| \begin{bmatrix} \beta_1^{*'} \beta_1 \Psi \beta_1' \beta_1^* & \beta_1^{*'} \beta_1 \Psi X_0 \\ X_0' \Psi \beta_1' \beta_1^* & X_0' \Psi X_0 \end{bmatrix} \right.$$

$$\left. - T^2 \hat{\lambda} \begin{bmatrix} 0 & 0 \\ 0 & \delta^{*'}\beta_{\perp,1}(\beta_\perp' \beta_\perp)^{-1} \left(\int G_0 G_0' \, ds \right)^{-1} (\beta_\perp' \beta_\perp)^{-1} \beta_{\perp,1}' \delta^* \end{bmatrix} \right\|$$

$$= |\beta_1^{*'} \beta_1 \Psi \beta_1' \beta_1^*| \times \left| X_0' \{ \Psi - \Psi \beta_1' \beta_1^* (\beta_1^{*'} \beta_1 \Psi \beta_1' \beta_1^*)^{-1} \beta_1^{*'} \beta_1 \Psi \} X_0 \right.$$

$$\left. - T^2 \hat{\lambda} \delta^{*'}\beta_{\perp,1}(\beta_\perp' \beta_\perp)^{-1} \left(\int G_0 G_0' \, ds \right)^{-1} (\beta_\perp' \beta_\perp)^{-1} \beta_{\perp,1}' \delta^* \right| = 0. \tag{A.4}$$

Therefore, the eigenvalues $\hat{\lambda}_{f+1}, \ldots, \hat{\lambda}_p$ 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 an $r \times (r-f)$ matrix J with rank $(r-f)$ such that

$$JJ' = \Psi - \Psi \beta_1' \beta_1^* (\beta_1^{*'} \beta_1 \Psi \beta_1' \beta_1^*)^{-1} \beta_1^{*'} \beta_1 \Psi, \tag{A.5}$$

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, because $|\beta^{*'}_1\beta_1\Psi\beta'_1\beta^*_1| \neq 0$, (A.4) becomes

$$\left| X'_0 JJ'X_0 - T^2\hat{\lambda}\delta^{*'}\beta_{\perp,1}(\beta'_\perp\beta_\perp)^{-1}\left(\int G_0 G'_0 \ ds\right)^{-1}(\beta'_\perp\beta_\perp)^{-1}\beta'_{\perp,1}\delta^* \right| = 0. \qquad \textbf{(A.6)}$$

The variance matrix of X'_0J conditioned on $G_0(\cdot)$ is given by

$$\delta^{*'}\beta_{\perp,1}(\beta'_\perp\beta_\perp)^{-1}\left(\int G_0 G'_0 \ ds\right)^{-1}(\beta'_\perp\beta_\perp)^{-1}\beta'_{\perp,1}\delta^* \otimes I_{r-f}. \qquad \textbf{(A.7)}$$

Noting that $sp(\beta_{\perp,1})$ must contain δ^* because $[\beta_1,\beta_{\perp,1}]$ is of full row rank n_1 and $sp(\beta_1)$ does not contain δ^*, we can see that $\delta^{*'}\beta_{\perp,1}$ has full row rank $n_1 - f$ irrespective of the rank of $\beta_{\perp,1}$, which is greater than $n_1 - f$ as explained in Remark 2. As a result, we can see that the conditional variance matrix of X'_0J is nonsingular (a.s.). Then, by multiplying the square root of the left-hand side of (A.7) from both sides of (A.6), the determinant equation becomes (11) with $j = n_1 - f$ and $k = r - f$, and then $T^2\hat{\lambda}$ converges in distribution to the solution of (11). This proves Theorem 1. ∎

Proof of Theorem 2. The outline of the proof is the same as the proof of Theorem 1, and thus we omit details.

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 $rk(\beta^*_{\perp,1}) = g$, and we denote the orthogonal complement to $\beta^*_{\perp,1}$ by η^*. Consider the following determinant equation:

$$|H'||\tilde{\beta}_{\perp,1}\tilde{\tilde{\Psi}}\tilde{\beta}'_{\perp,1} - \hat{\mu}\tilde{\tilde{\Phi}}||H| = 0, \qquad \textbf{(A.8)}$$

where $H = [\beta^*_{\perp,1}, T\eta^*]$. As in the previous proof, we replaced $\hat{}$ by $\tilde{}$. Because $\tilde{\beta}_{\perp,1}$ is the first n_1 rows of $\tilde{\beta}_\perp$, we obtain, using Lemma 1(iii),

$$T\eta^{*'}\tilde{\beta}_{\perp,1} = T\eta^{*'}(\tilde{\beta}_{\perp,1} - \beta_{\perp,1}) \xrightarrow{d} -\eta^{*'}\beta_1(\beta'\beta)^{-1}\int dV G'_0\left(\int G_0 G'_0 \ ds\right)^{-1}.$$

Then, similar to the previous proof, we can show that $T^2\hat{\mu}$ converges in distribution to a solution of (11) with $j = n_1 - g$ and $k = n - r - g$. 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 (1995), we can show that

$$\text{diag}\{T\beta'_\perp, T^{1/2}\}(\tilde{\beta}^+ - \beta^+) \xrightarrow{d} \left(\int G^+_0 G^{+'}_0 \ ds\right)^{-1}\int G^+_0 \ dV',$$

where $G^+_0 = [G'_0, 1]'$. Then, because $\tilde{\beta}$ is the first n rows of $\tilde{\beta}^+$, we have

$$T\beta'_\perp(\tilde{\beta} - \beta) \xrightarrow{d} L'\left(\int G^+_0 G^{+'}_0 \ ds\right)^{-1}\int G^+_0 \ dV', \qquad \textbf{(A.9)}$$

whose conditional variance is given by $L'(\int G_0^+ G_0^{+'} ds)^{-1} L \otimes (\alpha' \Sigma^{-1} \alpha)^{-1}$. Because $\tilde{\beta} = \beta + \beta_\perp (\beta_\perp' \beta_\perp)^{-1} \beta_\perp' \tilde{\beta}$ as expressed in Johansen (1995, p. 179), we have

$$T \delta^{*'} \tilde{\beta}_1 = \delta^{*'} \beta_{\perp,1} (\beta_\perp' \beta_\perp)^{-1} T \beta_\perp' (\tilde{\beta} - \beta)$$

$$\xrightarrow{d} \delta^{*'} \beta_{\perp,1} (\beta_\perp' \beta_\perp)^{-1} L' \left(\int G_0^+ G_0^{+'} ds \right)^{-1} \int G_0^+ dV'.$$

We also have $\tilde{Y}_T' S_{11}^+ \tilde{Y}_T \xrightarrow{d} \int G_0^+ G_0^{+'} ds$, which is proved as Lemma 1(iv), where $\tilde{Y}_T$ is Y_T with $\hat{\beta}_\perp$ replaced by $\tilde{\beta}_\perp$. Then, the theorems are proved similarly to Theorems 1 and 2. ∎

Proof of Theorem 5. For the case where $r < n - 1$, 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) Because $\tilde{\gamma}_\perp \xrightarrow{P} \gamma_\perp$ and $\gamma_\perp' \gamma = 0$, $\tilde{\gamma}$ converges in probability to γ.
  (ii) Proved by Johansen (1995).
  (iii) Letting $K = [\tilde{\beta}, T^{-1/2} \tilde{\gamma}, T^{-1} \tilde{\tau}]$, we have

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

From Lemma 10.3 in Johansen (1995), $T^{-1}[\bar{\gamma}, T^{-1/2} \bar{\tau}]' S_{11}[\bar{\gamma}, T^{-1/2} \bar{\tau}]$ converges in distribution to Ω whereas $\beta' S_{11} \beta$ 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,

$$K' S_{11} K \xrightarrow{d} \begin{bmatrix} (\beta' \beta)^{-1} \Sigma_\beta (\beta' \beta)^{-1} & 0 \\ 0 & \Omega \end{bmatrix}. \tag{A.10}$$

In addition, we can see that

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

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

$$T^{1/2} \tilde{\gamma}' K = [T^{1/2} \tilde{\gamma}' \beta (\beta' \beta)^{-1}, \tilde{\gamma}' \gamma (\gamma' \gamma)^{-1}, T^{-1/2} \tilde{\gamma}' \tau (\tau' \tau)^{-1}] \xrightarrow{P} [0, I_{n-r-1}, 0]. \tag{A.11}$$

From (A.10) and (A.11), $\tilde{\Omega}^{11}$ converges in distribution to Ω^{11}. ∎

(i.a) Proved in the same way as Theorem 1.

(i.b) In this case, the determinant equation becomes asymptotically equivalent to

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

$$\times (\gamma'\gamma)^{-1}\gamma_1'\delta^* + 12\delta^{*\prime}\tau_1(\tau'\tau)^{-2}\tau_1'\delta^*\}| = 0.$$

Note that, in general, for a given symmetric and positive definite matrix A and a vector b,

$$(A + bb')^{-1} = A^{-1} - A^{-1}bb'A^{-1}/(1 + b'A^{-1}b), \tag{A.12}$$

and then

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

for any nonzero vector c. By substituting $\delta^{*\prime}\gamma_1(\gamma'\gamma)^{-1}\Omega^{11}(\gamma'\gamma)^{-1}\gamma_1'\delta^*$ and $\sqrt{12}\delta^{*\prime}\tau_1(\tau'\tau)^{-1}$ for A and b, we obtain, for a given $G(\cdot)$,

$$\text{tr}(J'X\{\delta^{*\prime}\gamma_1(\gamma'\gamma)^{-1}\Omega^{11}(\gamma'\gamma)^{-1}\gamma_1'\delta^* + 12\delta^{*\prime}\tau_1(\tau'\tau)^{-2}\tau_1'\delta^*\}^{-1}X'J)$$

$$\leq \text{tr}(J'X\{\delta^{*\prime}\gamma_1(\gamma'\gamma)^{-1}\Omega^{11}(\gamma'\gamma)^{-1}\gamma_1'\delta^*\}^{-1}X'J)$$

$$= \text{tr}(X^*X^{*\prime}) = \chi^2_{(r-f)(n_1-f)}, \tag{A.13}$$

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

(ii) Let us consider the determinant equation (A.2) with $H = [\beta_1^*, T\delta_0^*, T\tau_1^*]$. Using Lemma 2 and by some algebra, the determinant equation is shown to be asymptotically equivalent to

$$|\beta_1^{*\prime}\beta_1\Psi\beta_1'\beta_1^*|| - 12T^2\hat{\lambda}(\tau'\tau)^{-2}(\tau_1^{*\prime}\tau_1)^2|$$

$$\times |X_1'JJ'X_1 - T^2\hat{\lambda}\delta_0^{*\prime}\gamma_1(\gamma'\gamma)^{-1}\Omega^{11}(\gamma'\gamma)^{-1}\gamma_1'\delta_0^*| = 0.$$

This determinant equation implies that there are f nonzero eigenvalues, $p - f - 1$ eigenvalues of order T^{-2}, and one eigenvalue of order smaller than T^{-2}. Then, we can see that

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

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

For the case where $r = n - 1$, the limiting distribution is derived similarly using (23). ∎

Proof of Corollary 1.

(i) 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 (A.12). 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 greater than that of $D'(A + bb')^{-1}D$. Then, by substituting $\delta^{*\prime}\gamma_1(\gamma'\gamma)^{-1}\Omega^{11}(\gamma'\gamma)^{-1}\gamma'_1\delta^*$, $\sqrt{12}\delta^{*\prime}\tau_1(\tau'\tau)^{-1}$, and $X'J$ for A, b, and D, the limiting distribution of $T^2\hat{\lambda}_p$ is shown to be bounded above by $\lambda^*_{min, r-f, n_1-f}$ because $D'A^{-1}D = X^*X^{*\prime}$ in this case. Note that $T^2\hat{\lambda}_p \xrightarrow{d} \lambda^*_{min, r-f, n_1-f}$ if and only if $\delta^{*\prime}\tau_1 = 0$.

(ii) is proved in Theorem 5(ii).   ∎

Proof of Theorem 6. Let us define $\beta^*_{\perp,1}$ and η^* as in the proof of Theorem 2.

LEMMA 3.

(i) $T\eta^{*\prime}\tilde{\gamma}_1 \xrightarrow{d} -\eta^{*\prime}\beta_1(\beta'\beta)^{-1}U'_1 = Y'$, say, where

$$\tilde{\gamma}_1 = \gamma_1 - \gamma_{\perp,1}(\tilde{\gamma}'_\perp\gamma_\perp)^{-1}\tilde{\gamma}'_\perp\gamma \quad \text{with } \gamma_{\perp,1} = [\beta_1, \tau_1]. \tag{A.14}$$

(ii) $T\eta^{*\prime}\hat{\tau}_1 \xrightarrow{P} 0$.

Proof.

(i) Because $\eta^{*\prime}\gamma_1 = 0$ and $\eta^{*\prime}\tau_1 = 0$, we have, using (A.14),

$$T\eta^{*\prime}\tilde{\gamma}_1 = T\eta^{*\prime}(\tilde{\gamma}_1 - \gamma_1) = -[\eta^{*\prime}\beta_1, 0](\tilde{\gamma}_\perp\gamma_\perp)^{-1}\begin{bmatrix} (\tilde{\beta} - \beta)'\gamma T \\ (\hat{\tau} - \tau)'\gamma T \end{bmatrix}$$

$$\xrightarrow{d} -\eta^{*\prime}\beta_1(\beta'\beta)^{-1}U'_1.$$

(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 (A.1), we can see that

$$T\eta^{*\prime}\tilde{\beta}_{\perp,1} = -\eta^{*\prime}\beta_1(\tilde{\beta}'\beta)^{-1}[(\tilde{\beta} - \beta)'\gamma T, (\tilde{\beta} - \beta)'\tau T]$$

$$\xrightarrow{d} -\eta^{*\prime}\beta_1(\beta'\beta)^{-1}U'_1[I_{n-r-1}, 0]. \tag{A.15}$$

We also have, from the definition of τ,

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

Because 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

$$[I_{n-r-1}, 0](\tilde{\alpha}'_\perp\hat{\Gamma}\tilde{\beta}_\perp)^{-1}\tilde{\alpha}'_\perp\hat{\mu} \xrightarrow{P} 0. \tag{A.16}$$

Combining (A.15) and (A.16), we obtain $T\eta^{*\prime}\hat{\tau}_1 = (T\eta^{*\prime}\tilde{\beta}_{\perp,1})(\tilde{\alpha}'_\perp\hat{\Gamma}\tilde{\beta}_\perp)^{-1} \times \tilde{\alpha}'_\perp\hat{\mu} \xrightarrow{P} 0$.   ∎

Similar to the proof of Theorem 2, we consider the same determinant equation as (A.8). 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 $[\tilde{\gamma}_1,\hat{\tau}_1] \xrightarrow{P} [\gamma_1,\tau_1] = \beta_{\perp,1}$, (A.8) is expressed as

$$|Y'S_1\underline{J}\,\underline{J}'S_1'Y - T^2\hat{\mu}\eta^{*\prime}\beta_1(\beta'\beta)^{-1}(\alpha'\Sigma^{-1}\alpha)^{-1}(\beta'\beta)^{-1}\beta_1'\eta^*| = 0, \qquad \text{(A.17)}$$

for large values of T, where an $(n-r) \times (n-r-g)$ matrix $\underline{J}$ satisfies $\underline{J}'\ddot{\Psi}^{-1}\underline{J} = I_{n-r-g}$. Noting that the conditional variance of $Y'S_1\underline{J}$ is given by

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

the test statistic $\mathcal{L}_{\perp T}$ conditioned on $G(\cdot)$ converges in distribution to

$$\text{tr}(Y^{*\prime}\underline{J}'S_1'\Omega^{11}S_1\underline{J}Y^*) = \text{tr}(Y^{*\prime}\underline{J}_1'\Omega^{11}\underline{J}_1Y^*), \qquad \text{(A.18)}$$

where $\text{vec}(Y^*) \sim N(0,I_{(n_1-g)(n-r-g)})$ and $\underline{J} = [\underline{J}_1',\underline{J}_2']'$. Because

$$\underline{J}'\ddot{\Psi}^{-1}\underline{J} = \underline{J}_1'\Omega^{11}\underline{J}_1 + 12\underline{J}_2'\underline{J}_2 = I_{n-r-g},$$

the limiting distribution (A.18) is bounded above by

$$\text{tr}(Y^{*\prime}\underline{J}_1'\Omega^{11}\underline{J}_1Y^*) \le \text{tr}(Y^{*\prime}(\underline{J}_1'\Omega^{11}\underline{J}_1 + 12\underline{J}_2'\underline{J}_2)Y^*) = \text{tr}(Y^{*\prime}Y^*). \qquad \blacksquare$$