Generalized VECM formulation and cointegrating rank determination by Johansen method

MITSUHIRO ODAKI

Abstract

This paper discusses the issues on what representation can be formulated based on vector autoregressions and how the cointegrating rank can be determined by the well-known Johansen method when the valid derivation of the conventional VECM is not possible. A generalized VECM representation is derived in connection with the cointegrating rank and determining the cointegrating rank by using the Johansen method is proposed under this representation. It is shown that the asymptotics and the procedure used for the rank determination under the conventional VECM are essentially applicable to the present situation.

Keywords: VECM, cointegrating rank, Johansen method, GRT.

1 Introduction

The concept of cointegration defined by Engle and Granger (1987) and others brought about a significant change in the model formulation based on vector autoregressions by picking up the situations in which individual time series considered are integrated of order one ($I(1)$) but some linear combinations of those are both weakly stationary and invertible ($I(0)$). Engle and Granger (1987) derived a *vector error correction model (VECM)* for such a situation in their *Granger representation theorem (GRT)*, emphasizing the existence of the error correcting term characterized by *the cointegrating rank* defined as the number of all the independent cointegrating relations considered. Detecting the cointegrating rank is the elementary among statistical analyses over cointegration, and the method proposed by Johansen (1988, 1992b) (Johansen method) is reputed to provide an handy and desirable estimator for it in the VECM.

On the other hand, the concept of multicointegration (or polynomial cointegaration) introduced by Granger and Lee (1990) compels us to reconsider such VECM formulation under cointegrated systems of $I(1)$ components. Engle and Yoo (1991) derived a VECM representation for systems in which a simple relation on multicointegration or a relation corresponding to the case of $b=2$ in the definition of cointegration by Engle and Granger (1987) is realized as an extension of GRT, and it was clarified that it results from a vector autoregeression for $I(2)$ components. Johansen (1992a), Paruolo (1996) and Engsted and Johansen (1999) have also studied the VECM for $I(2)$ components with the derivation of the condition for the level series forming the VECM to be be $I(2)$. Among those papers dealing with the VECMs for $I(2)$ components, Gregoir and Laroque (1993) set up a representation theorem for a general situation coping with more complicated multicointegration or various cases ($b>1$) in the definition of cointegration and derived a generalized VECM representation. The representation seems to be far from the form to which the Johansen method can be applied in appearance. In addition, the explanation on how it is related to the cointegrating rank was not given.

The purpose of this paper is to establish another VECM formulation for the general situation where GRT does not necessarily hold so that the cointegrating rank is described through some parameters and to show that determining the cointegrating rank by the Johansen method is still valid. We first derive some generalized VECM representation for this purpose, adopting a vector moving average (VMA) representation as the data generating process (DGP) and following a manner that is different from Gregoir and Laroque (1993). We also study the situation where the generalized VECM obtained above itself is adopted as the DGP, and conversely derive a VMA representation to evaluate the orders of integration for the data series considered. It is shown that the cointegrating rank is formulated explicitly in connection with some parameter of the representation, unlike the one in the representation derived by Gregoir and Laroque (1993). Also, based on this, we establish applicability of the Johansen method to the general situation, provided that our representation is formaulated by some finite lag-order in differences. As shown later, the limiting distribution of the statistic of Johansen's log-likelihood ratio test under the null is essentially the same as that for the conventional case, and the conventional critical points and the procedure to determine the rank value are also not altered. Several Monte Carlo experiments also are executed to illustrate the results obtained theoretically, and we will observe that the asymptotic desirability of the method is recognized even under finite samples as many as 200.

The paper is organized as follows: Section 2 formulates the DGP and the generalized VECM representation in connection with the cointegrating rank. The asymptotics of the method as the main results are presented in Section 3. Section 4 deals with Monte Carlo experiments. Some concluding remarks are put in Section 5. The Appendix contains proofs of theorems in the text.

2 DGP, VECM formulation and cointegrating rank description

In this paper we will adopt two generalized DGPs to discuss VECM formulation for cointegrated systems.

First, consider the observable k-dimensional vector time series y_t generated by

$$\Delta^d y_t = \sum_{j=0}^{\infty} C_j \epsilon_{t-j} + \sum_{n=0}^{q} t^n \mu_n, \qquad \forall t \geq 1, \quad (1)$$

where d is a positive integer, $\Delta = 1 - B$ with B denoting the backward operator C_j are $k \times k$ constant matrices satisfying

$$\sum_{j=0}^{\infty} (j+1)^{\rho} \parallel C_j \parallel < \infty,$$

for some $\rho > 1$, where $\|C_j\| = tr(C_j' C_j)^{1/2}$, $C_0 = I_k$ with I_k denoting the $k \times k$ identity matrix I_k, $\{\varepsilon_t\}$ is a sequence of unobservable k-dimensional random vectors which are mutually uncorrelated such that $E(\varepsilon_t) = 0$, $E(\varepsilon_t \varepsilon_t') = \Omega$, Ω is positive definite, and the fourth moments of components of ε_t are all finite, q is a nonnegative integer, and μ_n are k-dimensional constant vectors. We note that (1) is a VMA process with some deterministic terms and is expressed based on the Wold decomposition. Defining the power series $C(z)$ as

$$C(z) = \sum_{j=0}^{\infty} C_j z^j,$$

we make the following assumptions: det $C(z) = 0$ has roots either equal to 1 or strictly greater than 1 in absolute value, neither of the row vectors of $C(1)$ is zero vector, $rank\ C(1) = k - r$ with an integer r satisfying $k - 1 \geq r \geq 1$, there exist some positive integer $\bar{m}$ and some power series $g(z) = 1 + \sum_{j=0}^{\infty} g_j z^j$, where g_j are real numbers, such that det $C(z) = (1 - z)^{\bar{m}} g(z)$ and $g(1) \neq 0$, and with some positive integer q_0, $(y_0', \ldots y_{-q_0+1}')'$ is either $O(1)$ or $O_p(1)$, noting that these are thought to be conventionally imposed for cointegrated systems. As clarified by (1) and by the assumption that neither of row vectors of $C(1)$ is zero vector, the components of $\Delta^d y_t$ are $I(0)$ in the stochastic parts.[1] Also, the assumption that $rank\ C(1) = k - r$ is directly connected with the occurrence of cointegration in the definition of cointegration by Engle and Granger (1987) or Banerjee et *al*. (1993, p.

[1] It may suffice to adopt the framework of $d = 1$ for our purpose, since our study below is automatically applicable to one subject to this restriction by regarding $\Delta^{d-1} y_t$ as y_t.

145): There exist γ of size $k\times(k-r)$ and β of size $k\times r$ such that

$$rank\ \gamma' C(1)=k-r, \qquad \beta' C(1)=0, \qquad rank\ \beta=r,$$

and at least two of the components of each column of β are not zero. Following the above-mentioned definition, this implies that r corresponds to the cointegrating rank in the sense that $\beta'\Delta^{d-1}y_t$ is weakly stationary but $\gamma'\Delta^{d-1}y_t$ is not so.

For the derivation of a generalized VECM representation under (1), we need to construct some integrated series of y_t by extending the definition of Δ^j to the case $j\le 0$:

$$\Delta^0 y_t = y_t, \qquad \Delta^{-n-1}y_t = \sum_{h=1}^{t}\Delta^{-n}y_h, \qquad \forall n\ge 0, \qquad \forall t\ge 1.$$

We now establish:

Theorem 1: *For y_t generated by (1), we have the following representation:*

$$\Delta^d y_t = \bar{\alpha}\left(\sum_{n=0}^{m-1}\bar{\beta}_n'\,\Delta^{d-m+n}y_{t-1}\right) + \sum_{j=1}^{\infty}H_j\,\Delta^d y_{t-j} + \epsilon_t + \sum_{n=m}^{m+q}t^n\,\check{\mu}_n + \sum_{n=0}^{m-1}t^n\,\hat{\mu}_n, \quad \forall t\ge 1, \tag{2}$$

where $\bar{\alpha}$ is a column full rank matrix of $k\times r$, m is a positive integer, $\bar{\beta}_n$ are $k\times r$ matrices such that rank $[\bar{\beta}_{m-1}',\cdots,\bar{\beta}_1',\bar{\beta}_0']=r$, H_j are $k\times k$ constant matrices, $\check{\mu}_n$ for $m+q\ge n\ge m$ costant vectors and are expressed as linear combinations of μ_h, and $\hat{\mu}_n$ for $m-1\ge n\ge 1$ are given as some linear combinations of $\tilde{\mu}_h$, with

$$\tilde{\mu}_h = \Delta^{d-h}y_0 - \sum_{j=0}^{\infty}C_j^{(h)}\epsilon_{-j}, \quad h=1,\cdots,m.$$

Also, H_j satisfy the same convergence condition as that for C_j if det *$C(z)$ is rational, $\check{\mu}_{m+q}$ vanishes if μ_q is spanned by the columns of $C(1)$. Moreover, at least two of the components of each column of*

$$\bar{A}^{-1}\,\bar{\alpha}\left[\bar{\beta}_{m-1}',\cdots,\bar{\beta}_1',\ \bar{\beta}_0'\right]'$$

are not zero for any nonsingular $k\times k$ matrix $\bar{A}^{-1}$.

(2) is different from the representation derived by Gregoir and Laroque (1993), and therefore may be interpreted as another general VECM formaulation for the general situation where GRT does not necessarily hold. Note that (2) for the case $d=m=1$ corresponds to the conventional VECM (and is the same as Gregoir and Laroque's one under this case). Also, recalling that the components of $\Delta^d y_t$ are $I(0)$ in the stochastic parts, we see from (2) that $\sum_{n=0}^{m-1}\bar{\beta}_n'\Delta^{d-1-n}y_{t-1}$ consists of some relations on multicointegration or cointegrating relations for the case $b>1$ in Engle and Granger's (1987) definition except for the terms constructed from the initial vectors.

Next, suppose that the observable k-dimensional vector time series yt considered is generated by

$$\Delta^{d_1}y_t = \bar{\alpha}_1\left(\sum_{n=0}^{m_1-1}\bar{\beta}_{n;1}'\,\Delta^{d_1-m_1+n}y_{t-1}\right) + \sum_{j=1}^{p_1}H_{j;1}\,\Delta^d y_{t-j} + \epsilon_t + \sum_{n=0}^{q}t^n\,\bar{\mu}_n, \quad \forall t\ge 1, \tag{3}$$

where d_1 is a positive integer, $\bar{\alpha}_1$ is a column full rank matrix of $k\times r_1$, r_1 is a positive integer less than k, m_1 and p_1 are positive integer, $\bar{\beta}_{1;n}$ are $k\times r_1$ constant matrices, $H_{j;1}$ are $k\times k$ constant matrices, $\{\varepsilon_t\}$ is the one as given in (1), q is a nonnegative integer and $\bar{\mu}_n$ are k-dimensional constant vectors. Defining the matrix polynomial $A(z)$ as

$$A(z) = -\bar{\alpha}_1\left(\sum_{n=0}^{m_1-1}\bar{\beta}_{n;1}'\,(1-z)^n\,z\right) + (1-z)^{m_1}\left(I_k - \sum_{j=1}^{p_1}H_{j;1}\,z^j\right),$$

we assume that det $A(z)=0$ has roots either equal to 1 or strictly greater than 1 in absolute value, and also assume that with some nonnegative integer q_0, $(y_0',\ldots y_{-q_0+1}')'$ is either $O(1)$ or $O_p(1)$. Moreover, it is assumed that the frequency of the differencing operations needed to transform the components of y_t into $I(0)$ in the stochastic parts is identical amnog all the components. (3) is seemingly similar to (2), although (3) is a finite lag-order model and possesses no term on the initial vectors unlike (2). We may convert (3) to another suitable form in order to relate cointegration with the parameters appeared in the representation.

We also need to derive a VMA representation for this purpose. These will be provided in the following theorem:

Theorem 2: *For y_t generated by (3), we have the following representations:*

$$\Delta^d y_t = \bar{\alpha}\left(\sum_{n=0}^{m-1} \bar{\beta}_n' \Delta^{d-m+n} y_{t-1}\right) + \sum_{j=1}^{p} H_j \Delta^d y_{t-j} + \epsilon_t + \sum_{n=0}^{q} t^n \bar{\mu}_n, \quad \forall t \geq 1, \quad (4)$$

where $\bar{\alpha}$ and $\bar{\beta}_n$ are the ones as given in Theorem 1 with r as a positive integer less than k, d is a positive integer that is not less than d_1, $m=m_1+d-d_1$, p is a positive integer, H_j are $k \times k$ constant matrices such that

$$A(z) = -\bar{\alpha}\left(\sum_{n=0}^{m-1} \bar{\beta}_n' (1-z)^{m-1-n} z\right) + (1-z)^m \left(I_k - \sum_{j=1}^{p} H_j z^j\right)$$

and *rank* $\bar{\delta}'(I_k - \sum_{j=1}^{p} H_j) = k-r$ *for any* $k \times (k-r)$ *column full rank matrix* $\bar{\delta}$ *such that* $\bar{\delta}'\bar{\alpha}=0$,

$$\Delta^d y_t = \sum_{j=0}^{\infty} C_j \epsilon_t + t^q C(1) \bar{\mu}_q + \sum_{n=0}^{q-1} t^n \check{\mu}_n, \quad \forall t \geq 1, \quad (5)$$

where C_j are $k \times k$ constant matrices, the term $\sum_{n=0}^{q-1} t^n \check{\mu}_n$ is defined for the case $q \geq 1$ and is removed if $q=0$, and $\check{\mu}_n$ are k-dimensional constant vectors. Also, C_j satisfy the convergence condition as given in (1), *rank* $C(1)=k-r$ *and* $C(1)\bar{\mu}_q$ *vanishes if $\bar{\mu}_q$ is spanned by the columns of $\bar{\alpha}$ in* (4).

(5) shows that the components of $\Delta^d y_t$ are $I(0)$ in the stochastic parts. It is also seen from the result *rank* $C(1)=k-r$ that r corresponds to the cointegrating rank for $\Delta^{d-1} y_t$. Moreover, note that if the deterministic part is out of consideration, the VECM for $I(2)$ components in Johansen (1992a) or Johansen (1996, p. 57) corresponds to the case $d=m=2$ in (4) with $\bar{\alpha}, \bar{\beta}_0$ and $\bar{\beta}_1$ specified as

$$\bar{\alpha} = [\bar{\alpha}_0,\ \bar{\alpha}_1], \quad \bar{\beta}_0 = \left[\bar{\beta}_{00},\ 0\right], \quad \bar{\beta}_1 = \left[0,\ \bar{\beta}_{11}\right],$$

where $\bar{\alpha}_i$ and $\bar{\beta}_{ii}$ are $k \times r_i$ column full matrices for $i=0, 1$ and $r=r_0+r_1$.

3 Johansen method

Given T observations $y_1, \cdots, y_T$ in (2) or (4), we will discuss the Johansen method to determine the cointegrating rank r and its related log-likelihood ratio test (trace test) statistics. For this purpose, we assume that $H_j=0$ for all $j \geq p+1$ for the case where (1) is adopted as the DGP. (2) and (4) are then expressed as a unified form:

$$\Delta^d y_t = \bar{\alpha}\left(\sum_{n=0}^{m-1} \bar{\beta}_n' \Delta^{d-m+n} y_{t-1}\right) + \sum_{j=1}^{p} H_j \Delta^d y_{t-j} + \epsilon_t + t^q \bar{\mu}_q + \sum_{n=0}^{q-1} t^n \hat{\mu}_n, \quad \forall t \geq 1, \quad (6)$$

where q, $\bar{\mu}_q$ is a k-dimensional constant vector and $\hat{\mu}_n$ are redefined suitably. Note again that the term $\sum_{n=0}^{q-1} t^n \hat{\mu}_n$ is defined unless $q=0$. Furthermore, we incorporate the case $p=0$ into (6) by considering that H_j are all zero matrices under this case and assume that $H_p \neq 0$ for the case $p \geq 1$.

Now, put $\bar{T}=T-(p+d)$ and define the matrices/vectors Y_{-1} of size $\bar{T} \times k$, ΔY_{-j} of size $\bar{T} \times k$, $\check{\tau}_n$ of dimension $\bar{T}$, $\hat{\tau}(n)$ of size $\bar{T} \times (q+1)$ as

$$Y_{-1} = [\Delta^{d-1} y_{p+d}, \Delta^{d-1} y_{p+d+1}, \cdots, \Delta^{d-1} y_{T-1}]',$$
$$\Delta Y_{-j} = [\Delta^d y_{p+d+1-j}, \Delta^d y_{p+d+2-j}, \cdots, \Delta^d y_{T-j}]', \quad j=0, 1, \cdots, p,$$
$$\check{\tau}_n = ((p+d+1)^n, (p+d+2)^n, \cdots, T^n)', \quad n=0, \cdots, q,$$
$$\hat{\tau}(n) = [\check{\tau}_0, \cdots, \check{\tau}_n], \quad n=0, 1, \cdots, q.$$

Also, define $\check{Z}_{-1}$ and $\check{M}$ as

$$\check{Z}_{-1} = [\Delta Y_{-1}, \cdots, \Delta Y_{-p}, \hat{\tau}(q)] \quad \text{if } p \geq 1,$$
$$\check{M} = I_{\bar{T}} - \check{Z}_{-1}(\check{Z}_{-1}'\check{Z}_{-1})^{-1}\check{Z}_{-1}' \quad \text{if } p \geq 1,$$
$$\check{M} = I_{\bar{T}} \quad \text{if } p=0,$$

for the case where $q=0$ and $\bar{\mu}_0=0$ and define those as

$$\check{Z}_{-1} = \hat{\tau}(q) \quad \text{if } p=0,$$
$$\check{Z}_{-1} = [\Delta Y_{-1}, \cdots, \Delta Y_{-p}, \hat{\tau}(q)] \quad \text{if } p \geq 1,$$
$$\check{M} = I_{\bar{T}} - \check{Z}_{-1}(\check{Z}_{-1}'\check{Z}_{-1})^{-1}\check{Z}_{-1}'$$

for the other cases. Following the notations S_{ij} used in Johansen (1988, 1992b, 1996), let us introduce S_{ij}

given as

$$S_{00}=\Delta Y_0' \check{M}\Delta Y_0/T, \quad S_{11}=Y_{-1}' \check{M}Y_{-1}/T,$$
$$S_{01}=\Delta Y_0' \check{M}Y_{-1}/T, \quad S_{10}=S_{01}'.$$

Furthermore, let $\hat{\lambda}_1 \geq \cdots \geq \hat{\lambda}_k$ be the ordered eigenvalues of the equation $\det\{\lambda S_{11}-S_{10}S_{00}^{-1}S_{01}\}=0$. A test statistic for the null $r=j$ and the alternative $r \geq j+1$ is given as

$$-T\sum_{h=j+1}^{k}\log\left(1-\hat{\lambda}_h\right).$$

It should be noted that the value of m is not used in the construction of the statistic.

In order to derive the asymptotics for the test statistic, let the symbols $\Rightarrow$ and $W_s(u)$ stand for weak convergence of probability measures on the unit interval [0, 1] and a s-dimensional standard Brownian motion of on [0, 1] respectively, and put $\psi_q(u)=(1,\cdots, u^q)'$,

$$\bar{W}_s(u)=u^{q+1} \quad if\ s=1,$$
$$\bar{W}_s(u)=\left(u^{q+1},\ W_{s-1}'(u)\right)' \quad if\ s>1.$$

Furthermore, define $\tilde{W}(u)$ as $W_{k-r}(u)$ for the case where $q=0$ and $\bar{\mu}_0=0$, define it as

$$W_{k-r}(u)-\left(\int_0^1 W_{k-r}(u)\,\psi_q'(u)du\right)$$
$$\cdot\left(\int_0^1 \psi_q(u)\psi_q'(u)du\right)^{-1}\psi_q(u)$$

for the case where $\bar{\mu}_q\neq 0$ is spanned by the columns of $\bar{\alpha}$ (including the case $\bar{\mu}_q=0$), and define it as

$$\bar{W}_{k-r}(u)-\left(\int_0^1 \bar{W}_{k-r}(u)\,\psi_q'(u)du\right)$$
$$\cdot\left(\int_0^1 \psi_q(u)\psi_q'(u)du\right)^{-1}\psi_q(u)$$

for the other cases. Next, define u_t and v_t as $u_t=\sum_{j=0}^{\infty}C_j\varepsilon_t$ and $v_t=\sum_{j=0}^{\infty}(-\sum_{i=j+1}^{\infty}C_i)\,\varepsilon_t$, Also, for the case $p\geq 1$, let $\check{u}_t$ and $\check{v}_{t-1}$ denotes the errors from the linear least-square predictors of u_t and v_{t-1} onto the (Hilbert) space spanned by u_{t-j}, $j=1,\cdots,p$, respectively, and for the case $p=0$, put $\check{u}_t=u_t$ and $\check{v}_{t-1}=v_{t-1}$. Moreover, put

$$R_u=E\left(\check{u}_t\check{u}_t'\right), \quad R_v=E\left(\check{v}_t\check{v}_t'\right),$$
$$R_{uv}=E\left(\check{u}_t\check{v}_{t-1}'\right), \quad R_{vu}=R_{uv}'.$$

We now set up the following theorem for the test statistics:

Theorem 3: *Suppose y_t is generated by* (6). *Then, for $\hat{\lambda}_h$ given above, we have:*

(*i*) $$-T\sum_{h=r+1}^{k}\log\left(1-\hat{\lambda}_h\right)$$
$$\Rightarrow tr\left\{\left(\int_0^1 dW_{k-r}(u)\tilde{W}'(u)\right)\cdot\left(\int_0^1\tilde{W}(u)\tilde{W}'(u)du\right)^{-1}\left(\int_0^1\tilde{W}(u)dW_{k-r}(u)'\right)\right\}.$$

(*ii*) $$-\sum_{h=j}^{k}\log\left(1-\hat{\lambda}_h\right)$$
$$=-\sum_{h=j}^{r}\log\left(1-\nu_h\right)+O_p(T^{-1/2}),$$
$$1>\nu_h>0, \qquad j=1,\cdots,r,$$

where $\nu_1\geq\cdots\geq\nu_r$ are the ordered eigenvalues of

$$\left(\beta' R_v\,\beta\right)^{-1/2}\beta' R_{vu}R_u^{-1}R_{uv}\beta\left(\beta' R_v\,\beta\right)^{-1/2}.$$

Theorem 3 ensures that the asymptotics of the test statistics obtained under the conventional VECM are valid even if (6) is adopted as the DGP and therefore indicates that the procedure to determine the value of r in the Johansen method, which is presented in Johansen (1992b) or Johansen (1996, pp. 98-100), should be used as it is. Also, for the limiting distribution given by Theorem 3 (i), Johansen (1996, p. 94) describes several special cases: (6.20) for the case where $q=0$ and $\bar{\mu}_0=0$; (6.21) for the case where $q=0$ and $\bar{\mu}_0\neq 0$ is not spanned by the columns of $\bar{\alpha}$; and (6.22) for the case where $q=1$ and $\bar{\mu}_1\neq 0$ is not spanned by the columns of $\bar{\alpha}$. We note that these are used for the examples presented in the following section.

4 Examples

In this section, we execute Monte Carlo experiments on the cointegrating rank determination based on the Johansen method in several DGPs as special cases of (6). The purpose of the experiments is to observe the extent to which the asymptotics established theoretically in the previous section are preserved for finite samples. The DGPs presented below are of 4-variates systems ($k=4$) with $d=1$, $m=2$, ε_t as Gaussian with mean zero and covariance matrix I_4 ($\Omega=I_4$) and are classified into two groups according to the value of r: $r=1$ for the first group and $r=2$ for the second one. All the DGPs are constructed so that $y_{-j}=0$ for any $j \geq 0$, $q=0$ or 1, $\hat{\mu}_0$ is constant, both $\bar{\mu}_1$ and $\hat{\mu}_0$ are not spanned by the columns of $\bar{\alpha}$, and the roots of det $A(z)=0$ are either greater than 1 in absolute values or equal to 1 for each DGPs. Also, for any $t \geq 1$, the DGPs included in each group are expressed in a unified VECM form:

$$\Delta y_t = \bar{\alpha}\,\{\bar{\beta}_0'(\sum_{h=1}^{t-1} y_h) + \bar{\beta}_1' y_{t-1}\} + \sum_{j=1}^{3} H_j\,\Delta y_{t-j} + t\,g_2\,\bar{\mu} + g_1\bar{\mu} + \epsilon_t,$$

where g_1 and g_2 are parameters to administer the existence of deterministic terms and are set to be either 1 or 0 and $\bar{\mu}=(1,\,0,\,0.5,\,1)'$. We note that y_t possesses no deterministic term ($q=0$ and $\bar{\mu}_0=0$) if $g_1=g_2=0$, it possesses only a linear trend ($q=0$ and $\bar{\mu}_0 \neq 0$) if $g_1=1$ and $g_2=0$, it possesses only a quadratic trend ($q=1$, $\bar{\mu}_1 \neq 0$ and $\hat{\mu}_0=0$) if $g_1=0$ and $g_2=1$, and the case $g_1=g_2=1$ is not dealt with.

The matrices $\bar{\alpha}, \bar{\beta}_0, \bar{\beta}_0$ and H_j are also described by the parameters g_h for $h=3, 4, 5, 6$ below, and those for the first group are given as

$$\bar{\alpha} = (-0.2,\ -0.2,\ -0.5,\ -0.2)',$$
$$\beta_0' = (1,\ 1,\ 1,\ 1),$$
$$\beta_1' = (g_3,\ g_4,\ 1,\ g_5),$$
$$H_1 = \begin{pmatrix} -0.4 & 0.2 & 0.2 & -0.2 \\ -0.4 & 0 & 0.2 & 0 \\ 0 & 0.5 & 0 & -0.5 \\ -0.2 & 0.2 & 0.2 & -0.5 \end{pmatrix}, \quad H_2 = H_3 = 0.$$

The values of g_h are confined to the ones in the following cases: Case (1) $g_3=1$, $g_4=-1$, $g_5=g_6=0$; Case (2) $g_3=1$, $g_4=1$, $g_5=g_6=0$; Case (3) $g_3=0$, $g_4=-1$, $g_5=1$, $g_6=0$; and Case (4) $g_3=-1$, $g_4=-1$, $g_5=g_6=1$.

Those for the second group are

$$\bar{\alpha} = \begin{pmatrix} -0.5 & -0.5 & 0 & 0 \\ 0 & 0 & -0.1 & 0 \end{pmatrix}',$$

$$\bar{\beta}_0' = \begin{pmatrix} g_3 & g_3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\bar{\beta}_1' = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 2.5 & 2.5 & 6 & -6 \end{pmatrix},$$

$$H_1 = g_4 \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0.25 & 0.25 & 0 & 0 \\ 0 & -2 & -0.2 & -0.3 \end{pmatrix},$$

$$H_2 = g_5 \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0.25 & 0.25 & 0 & 0 \\ -1 & 1 & 0 & 0 \end{pmatrix},$$

$$H_3 = g_5 \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -0.25 & 0.25 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The values of g_h are confined to the ones in the following cases: Case (5) $g_3=1$, $g_4=g_5=1$; Case (6) $g_3=3/5$, $g_4=g_5=1$; Case (7) $g_3=1$, $g_4=1$, $g_5=0$; and Case (8) $g_3=3/5$, $g_4=1.5$, $g_5=0$.

For each of the above DGPs, we ran 10000 simulations using a sample size of 200 ($T=200$), adopting pseudo normal random variables for the components of ϵ_t to produce the test statistics needed for the Johansen method. In each DGP, the test statistics are constructed based on the true value of p, which is 1, 2 or 3. The aim of the experiments is to obtain the relative frequency for the Johansen method to make a correct decision on the value of r over 10000 simulations for each DGP. Two values for it are calculated for each DGP according to the tests executed consecutively at both 5% and 1% significance levels. For the critical values, we follow Johansen's (1996, pp. 214-216) Tables 15.1, 15.3 and 15.5, which correspond to the cases $g_1=g_2=0$, $g_1=1$, $g_2=0$ and $g_1=0$, $g_2=1$ respectively.[2] The experimental results are reported in Table 1, and the figures in the table are expressed as percentiles.

As observed in the table, the frequency that the method detects the true value of r is sufficiently close to the theoretical one (1 minus the significance level) for all the DGPs, indicating that the asymptotics established in the previous section are sufficiently tenable for the sample size of $T=200$ as far as those DGPs are concerned.

[2] This paper does not adopt more accurate critical values in MacKinnon et *al*. (1999) since 1% critical values are not available.

5 Concluding remarks

We have been discussed the valid derivation of some VECM representation for the general situation in which the conventional VECM is not derived and some applicability of the Johansen method to determine the cointegrating rank to the situation. In the VECM representation derived in this paper, the cointegrating rank was explicitly described as the rank (or the column size) of some coefficient matrix for the generalized error correcting term. It was also established theoretically that the the use of Johansen method in the conventional manner leads to essentially the same asymptotics for the rank determination as the ones for the conventional VECM due to the form of the representation. It is enforced by the performance of the method for several examples and a sample size of 200, based on the satisfactory results of the experiments in the previous section.

Some semiparametric and nonparametric approaches to the rank determination may be sometimes considered as measures against some limitation of the Johansen method. For the case in which the valid derivation of the conventional VECM is ensured but its lag-order is infinite, Saikkonen (1992) and Qu and Perron (2007) have discussed the applicability of the Johansen method based on a finite lag-order approximation of the infinite lag-order or the determination of an optimal lag-order. On the other hand, Shintani (2001) proposed some nonparametric tests for the rank determination without formulating any vector autoregression scheme. The issues on how such these methods can be applied to the general situation and what alterations should be made need to be discussed formally. We will leave these issues to future research and only state that the Johansen method is always useful for the cointegrating rank detection within the framework of finite lag-order vector autoregressions.

References

Banerjee, A., Dolado, J. J., Galbraith, J. W. and Hendry, D. F. (1993) *Co-Integration, Error-Correction and the Econometric Analysis of Non-stationary Data*. Oxford: Oxford University Press.

Engle, R. F. and Granger, C.W.J. (1987) Co-integration and error-correction: Representation, estimation and testing. *Econometrica*, Vol. 55, pp. 251-76.

Engle, R. F. and Yoo, B. S. (1991) Cointegrated economic time series: A survey with new results. in Granger, C.W.J. and Engle, R.F. (eds.). *Long-Run Economic Relationships: Readings in Cointegration*. Oxford: Oxford University Press.

Engsted, T., and Johansen, S. (1999) Granger's representation theorem and multicointegration. in Engle, R.F. and White, H. (eds.). *Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive Granger*. Oxford: Oxford University Press.

Granger, C. W. J. and Lee, T-H. (1990) Multicointegration. *Advances in Econometrics*. Vol. 8, pp. 71-84.

Gregoir, S. and Laroque, G. (1993) Multivariate time series: a polynomial errorcorrection representation theorem. *Econometric Theory*, Vol. 9, pp. 329-42.

Johansen, S. (1988) Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, Vol. 12, pp. 231-54.

Johansen, S. (1992a) A representation of vector autoregressive processes inte grated of order 2. *Econometric Theory*, Vol. 8, pp. 188-202.

Johansen, S. (1992b) Determination of cointegration rank in the presence of a linear trend, *Oxford Bulletin of Economics and Statistics*, Vol. 54, pp. 383-97.

Johansen, S. (1996) *Likelihood-Based Inference in Cointegrated Vector Auto-regressive Models*. (2nd printing), Oxford: Oxford University Press.

MacKinnon, J.G., Haug A. A. and Michelis, L. (1999) Numerical distribution functions of likelihood ratio tests for cointegration, *Journal of Applied Econometrics*, Vol. 14, pp. 563-77.

Paruolo, P. (1996) On the determination of integration indices in *I*(2) systems. *Journal of Econometrics*, Vol. 72, pp. 313-56.

Park, J. Y. and Phillips, P. C. B. (1988) Statistical inference in regressions with integrated processes: Part Ⅰ. *Econometric Theory*, Vol. 4, pp. 468-97.

Park, J. Y. and Phillips, P. C. B. (1989) Statistical inference in regressions with integrated processes: Part Ⅱ. *Econometric Theory*, Vol. 5, pp. 95-131.

Qu, Z. and Perron P. (2007) A modified information criterion for cointegration tests based on a VAR approximation, *Econometric Theory*, Vol. 23, pp. 638-685.

Saikkonen, P. (1992) Estimation and testing of cointegrated systems by an autoregressive approximation, *Econometric Theory*, Vol. 9, pp. 1-27.

Shintani M. (2001) A simple cointegrating rank test without vector autoregression, *Journal of Econometrics*, Vol. 105, pp. 337-62.

Appendix

Proof of Theorem 1 By the assumptions for $C(z)$, we see that there uniquely exists a combination of a power series $G(z)=I_k+\sum_{j=1}^{\infty}G_j z^h$ of $k\times k$ matrices G_j and a nonnegative integer $\bar{n}$ such that

$$adj\, C(z) = (1-z)^{\bar{n}}\, G(z), \quad G(1)\neq 0, \quad \bar{n}<\bar{m},$$

where $adj\ C(z)$ denotes the adjoint matrix of $C(z)$. Putting $A(z)=G(z)/g(z)$ and $m=\bar{m}-\bar{n}$, it is obvious that $A(z)C(z)=(1-z)^m I_k$ holds and that $A(z)$ is expressed as

$$A(z) = I_k + \sum_{j=1}^{\infty} A_j\, z^j,$$

where A_j are $k\times k$ matrices satisfying the same convergence condition as for C_j.

Next, for $n=0, 1, \cdots, m$, put

$$C_0 = I_k, \quad C_j^{(0)} = C_j, \quad \forall j\geq 1, \quad C^{(0)}(z) = C(z),$$

$$C_j^{(n+1)} = -\sum_{i=j+1}^{\infty} C_i^{(n)}, \qquad \forall j\geq 0,$$

$$C^{(n)}(z) = \sum_{j=0}^{\infty} C_j^{(n)}\, z^j,$$

$$A_0 = I_k, \quad A_j^{(0)} = A_j, \quad \forall j\geq 1, \quad A^{(0)}(z) = A(z),$$

$$A_j^{(n+1)} = -\sum_{i=j+1}^{\infty} A_i^{(n)}, \qquad \forall j\geq 0,$$

$$A^{(n)}(z) = \sum_{j=0}^{\infty} A_j^{(n)}\, z^j,$$

which induce

$$C(z) = \sum_{n=0}^{b-1}(1-z)^n\, C^{(n)}(1) + (1-z)^b\, C^{(b)}(z),$$

$$A(z) = \sum_{n=0}^{b-1}(1-z)^n\, A^{(n)}(1) + (1-z)^b\, A^{(b)}(z), \qquad b=1,\cdots,m, \tag{A.1}$$

Using (A.1) when $b=m$ in $A(z)\ C(z)=(1-z)^m I_k$, it follows that

$$\left[A^{(m-1)}(1),\cdots,A^{(1)}(1),\, A^{(0)}(1)\right]\begin{bmatrix} C^{(1)}(1)\\ \vdots \\ C^{(m-1)}(1)\\ C^{(m)}(1)\end{bmatrix} + A^{(m)}(1)\, C^{(0)}(1) = I_k, \tag{A.2}$$

$$\left[A^{(m)}(1),\ A^{(m-1)}(1),\cdots,\ A^{(1)}(1),\ A^{(0)}(1)\right]\bar{C}_{0;m-1} = 0, \tag{A.3}$$

where

$$\bar{C}_{0;m-1} = \begin{bmatrix} C^{(0)}(1) & & & \mathbf{0} \\ C^{(1)}(0) & \ddots & & \\ \vdots & \vdots & C^{(0)}(1) & \\ C^{(m-1)}(1) & \cdots & C^{(1)}(1) & C^{(0)}(1)\end{bmatrix}.$$

Setting $rank\ [A^{(m-1)}(1),\cdots, A^{(1)}(1), A^{(0)}(1)]=\bar{r}$, (A.2) requires that $\bar{r}\geq r$, since

$$rank\, C^{(0)}(1)\, A^{(m)}(1) \leq rank\, C^{(0)}(1) = k-r.$$

Noting that $[\beta,\ \gamma]$ has full rank and $A^{(0)}(1)\ C^{(0)}(1)=0$, we can easily find matrices F_n and G_n such that

$$\bar{B}'\, A^{(n)}(1) = F_n\,\beta' + G_n\,\gamma', \qquad G_0 = 0, \qquad n=0,\cdots,m-1,$$

for some column full rank matrix $\bar{B}$ of size $k\times\bar{r}$. (A.3) can be then expressed as

$$\sum_{n=0}^{m-h} F_n\,\beta' C^{(m-h-n)}(1) + \sum_{n=0}^{m-h} G_n\,\gamma' C^{(m-h-n)}(1) = 0, \qquad h=1,\cdots,m. \tag{A.4}$$

It is not difficult to see that for $h=2,\cdots,m$, the relation

$$\sum_{n=0}^{m-h+1} F_n\, \beta' C^{(m-h+1-n)}(1) + \sum_{n=0}^{m-h+1} G_n\, \gamma' C^{(m-h+1-n)}(1) = 0$$

results from

$$\sum_{n=0}^{m-h} F_n\, \beta' C^{(m-h-n)}(1) + \sum_{n=0}^{m-h} G_n\, \gamma' C^{(m-h-n)}(1) = 0,$$

noting that

$$\sum_{n=0}^{m-h} F_n\, \beta' C^{(m-h-n)}(z) \;+\; \sum_{n=0}^{m-h} G_n\, \gamma' C^{(m-h-n)}(z)$$

$$= \;(1-z) \sum_{n=0}^{m-h} F_n\, \beta' C^{(m-h+1-n)}(z) + \sum_{n=0}^{m-h} G_n\, \gamma' C^{(m-h+1-n)}(z).$$

In other words, F_n and G_n satisfying the former necessarily satisfy the latter. As a result, we can see that all the relations in (A.4) result from $F_0 \beta' C^{(0)}(1) = 0$, which requires that $\bar{r} = r$.

Next, define the $k \times k$ matrices $\bar{A}^{(m-n)}(1)$ and H_j for $n = 1, \cdots, m$ and for $j \geq 1$ and the power series $H(z)$ as those satisfying

$$\left[\bar{A}^{(m-1)}(1), \cdots, \bar{A}^{(1)}(1),\; \bar{A}^{(0)}(1)\right]$$

$$= \left[A^{(m-1)}(1), \cdots, A^{(1)}(1),\; A^{(0)}(1)\right] \begin{bmatrix} I_k & & & 0 \\ \vdots & \ddots & & \\ I_k & \cdots & I_k & \\ I_k & \cdots & I_k & I_k \end{bmatrix},$$

$$H(z) = A^{(m)}(z) \;+\; \sum_{n=0}^{m} A^{(n)}(1) = I_k \;-\; \sum_{j=0}^{\infty} H_j\, z^j.$$

It is now easy to see that there exist some column full rank matrices $\bar{\alpha}$ and $\bar{\beta}_n$ of size $k \times r$ such that

$$\left[\bar{A}^{(m-1)}(1), \cdots, \bar{A}^{(1)}(1),\; \bar{A}^{(0)}(1)\right] = -\,\breve{\alpha}\, \left[\bar{\beta}'_{m-1}, \cdots, \bar{\beta}'_1,\; \bar{\beta}'_0\right]. \qquad (A.5)$$

It can be also shown that

$$A(z) = -\,\bar{\alpha}\, \left(\sum_{n=0}^{m-1} \bar{\beta}'_n\, (1-z)^n\, z\right) + (1-z)^m\, H(z). \qquad (A.6)$$

Note that the matrices $\bar{\alpha}$, $\bar{\beta}_{m-1-n}$ and H_j satisfy the condition as given in the statement of the lemma.

On the other hand, using the expansion for $C(z)$ when $b = 1$ in (A.1), (1) is written as

$$\Delta^d y_t = C^{(0)}(1)\, \epsilon_t(1) \;+\; C^{(1)}(B)\, \epsilon_t \;+\; \sum_{n=0}^{q} t^n\, \mu_n, \qquad \forall\, t \geq 1.$$

Noting that $\sum_{h=1}^{t} \Delta^d y_h = \Delta^{d-1} y_t - \Delta^{d-1} y_0$, it is straightforward from this that

$$\Delta^{d-1}\, y_t = C^{(0)}(1)\, \epsilon_t(1) \;+\; C^{(1)}(B)\, \epsilon_t + \sum_{n=1}^{q+1} t^n\, \breve{\mu}_{n;1} \;+\; \tilde{\mu}_1, \qquad \forall\, t \geq 1, \qquad (A.7)$$

where

$$\epsilon_t(1) = \sum_{h=1}^{t} \epsilon_h, \qquad \tilde{\mu}_1 = \Delta^{d-1}\, y_0 \;-\; \sum_{j=0}^{\infty} C_j\, \epsilon_{-j},$$

$$\sum_{n=1}^{q+1} t^n\, \breve{\mu}_{n;1} = \sum_{n=0}^{q} \Big(\sum_{h=1}^{t} h^n\Big)\, \mu_n.$$

Note that $\breve{\mu}_{q+1} = \mu_q$. By induction on b, we can derive

$$\Delta^{d-b} y_t \;=\; \sum_{n=0}^{b-1} C^{(n)}(1)\, \epsilon_t(b-n) \;+\; C^{(b)}(B)\, \epsilon_t + \sum_{n=b}^{q+b} t^n\, \breve{\mu}_{n;b} \;+\; \sum_{n=0}^{b-1} d_t(n)\, \tilde{\mu}_{b-n}, \qquad \forall\, t \geq 1, \quad b = 1, \cdots, m, \qquad (A.8)$$

$$\epsilon_t(n+1) = \sum_{h=1}^{t} \epsilon_h(n), \qquad n = 1, \cdots, m-1,$$

$$d_t(n+1) = \sum_{h=1}^{t} d_h(n), \quad d_t(0) = 1, \quad n = 0, 1, \cdots, m,$$

$$\sum_{n=b}^{q+b} t^n\, \breve{\mu}_{n;b} = \sum_{n=b-1}^{q+b-1} \Big(\sum_{h=1}^{t} h^n\Big)\, \breve{\mu}_{n;b-1}, \qquad b = 1, \cdots, m, \qquad \breve{\mu}_{n;0} = \mu_n,$$

noting that $\breve{\mu}_{q+b;b} = \mu_q$. Putting $\varepsilon_t(0) = \varepsilon_t$, noting that

$\varepsilon_t(n) = (1-B)^{m-n}\, \varepsilon_t(m), \quad n = 0, 1, \cdots, m-1,$

and using the expansion for $C(z)$ when $b = m$ in (A.1), (A.8) when $b = m$ is rewritten as

$$\Delta^{d-m} y_t = C(B)\, \epsilon_t(m) \;+\; \sum_{n=m}^{q+m} t^n\, \breve{\mu}_{n;m} + \sum_{n=0}^{m-1} d_t(n)\, \tilde{\mu}_{m-n}, \qquad \forall\, t, \geq 1. \qquad (A.9)$$

Note that $\breve{\mu}_{q+m;m} = \mu_q$. Premultiplying both sides of (A.8) by $A(B)$, we can obtain

$$A(B)\, \Delta^{d-m} y_t = \epsilon_t + \sum_{n=m}^{m+q} t^n\, \breve{\mu}_n + \sum_{n=0}^{m-1} d_t(n)\, \Big\{\sum_{h=0}^{m-n} A^{(h)}(1)\, \tilde{\mu}_{m-n-h}\Big\}, \qquad \forall\, t \geq 1, \qquad (A.10)$$

where $\check{\mu}_n$ are constructed suitably from $\check{\mu}_{n;m}$. t Substituting the right-hand side of (A.6) in which z is replaced by B for $A(B)$ in (A.10) and recalling the natures of $d_t(n)$ and $\tilde{\mu}_n$, it is easy to obtain (2). If det $C(z)$ is rational, $g(z)$ given above must be so. Noting this, the condition of the roots of $g(z)=0$, together with this and the convergence condition for C_j, ensures that H_j satisfy the same convergence condition as that for C_j. It is also obvious that $\check{\mu}_{m+q}=A(1)\ \mu_q$, which implies that $\check{\mu}_{m+q}$ vanishes if μ_q is spanned by the columns of $C(1)$. Moreover, if the components of one row of

$$\bar{A}^{-1}\,\bar{\alpha}\,\left[\bar{\beta}'_{m-1},\cdots,\bar{\beta}'_1,\ \bar{\beta}'_0\right]'$$

for some nonsingular $k\times k$ matrix $\bar{A}^{-1}$ are zero except one, it contradicts the fact that the components of $\Delta^d y_t$ are $I(0)$ in the stochastic parts. Thus the result required for the lemma is derived.

Proof of Theorem 2 First, note that there exists some positive integer $\check{m}$ and some polynomial $\bar{g}(z)$ such that det $A(z)=(1-z)^{\check{m}}\ \bar{g}(z)$ and $\bar{g}(1)\neq 0$. Also, define $A^{(n)}(z)$ as given in the proof of Theorem 1. By the same manner as that used for establishing the equivalence between (A.10) and (2) in the proof of Theorem 1, we see that (3) is written as

$$A(B)\,\Delta^{d_1-m_1}y_t=\epsilon_t+\sum_{n=0}^{q}t^n\,\bar{\mu}_n,\quad \forall\, t\geq 1. \qquad (A.11)$$

If it is satisfied that $\bar{\delta}'_1\ (I_k-\sum_{j=1}^{p}H_j)=k-r_1$ for any $k\times(k-r_1)$ column full rank matrix $\bar{\delta}_1$ such that $\bar{\delta}'_1\,\bar{\alpha}_1=0$, (3) itself can be regarded as (4). If it is not so, it can be shown that $rank\ [A^{(m_1)}(1),\cdots,A^{(1)}(1),A^{(0)}(1)]<k$. However, since det $A(z)=(1-z)^{\check{m}}\bar{g}(z)$ and $\bar{g}(1)\neq 0$, there must exist some integer m greater than m_1 such that

$$rank\ [A^{(m-1)}(1),\cdots,A^{(1)}(1),A^{(0)}(1)]\equiv r<k,$$
$$rank\ [A^{(m)}(1),\cdots,A^{(1)}(1),A^{(0)}(1)]\equiv k, \qquad (A.12)$$

which ensures that there exist $\bar{\alpha}$ and $\bar{\beta}_n$ as given in Theorem 1. (4) can be then obtained by using the argument used for the equivalence between (A.10) and (2) again and (A.4) and by putting $d=d_1-m_1+m$ and

$$A^{(m)}(z)+\sum_{n=0}^{m}A^{(n)}(1)=I_k-\sum_{j=1}^{p}H_j\,z^j=H(z)$$

with some positive integer p. We then have the expression required for $A(z)$.

For the derivation of (5), recall the nature of det $A(z)$. It ensures the existence of $G(z)$ such that

$$adj\ A(z)=(1-z)^{\check{n}}\,\bar{G}(z),$$
$$G(z)=I_k+\sum_{j=1}^{\infty}G_j\,z^h,\quad \bar{G}(1)\neq 0,\quad \check{m}>\check{n},$$

where G_j are $k\times k$ matrices and $\check{n}$ is a nonnegative integer. Put $C(z)=\bar{G}(z)/\bar{g}(z)$, and define $C^{(n)}$ and $\bar{C}_{0;\,m-1}$ as given in the proof of Theorem 1. It is now trivial to see that

$$A(z)\,C(z)=C(z)\,A(z)=(1-z)^{\check{m}-\check{n}}\,I_k,$$

from which follows that (A.2) and (A.3), in which m is replaced by $\check{m}-\check{n}$, hold. This, together with (A.11), requires that $\check{m}-\check{n}=m$. Also, let $C(z)=\sum_{j=0}^{\infty}C_j$ with some $k\times k$ constant matrices C_j. $C(z)$ given above is rational due to the condition of the roots of $\bar{g}(z)=0$. This ensures that C_j satisfy the convergence condition as given in (1). Moreover, it is asserted by (A.2) that $rank\ C^{(0)}(1)\ A^{(m)}(1)\geq k-r$, therefore, $rank\ C(1)\geq k-r$. Now, suppose that $rank\ C(1)>k-r$. This implies that there exist some integer $\check{r}$ less than r, $\check{\gamma}$ of size $k\times(k-\check{r})$ and $\check{\beta}$ of size $k\times\check{r}$ such that

$$rank\,\check{\gamma}'\,C(1)=rank\,C(1)=k-\check{r},$$
$$\check{\beta}'\,C(1)=0,\qquad rank\,\check{\beta}=\check{r}.$$

By using argument similar to those used for the derivation of (A.5) in the proof of Theorem 1 and $[\check{\beta},\ \check{\gamma}]$ instead of $[\beta,\ \gamma]$, it can be led to that

$$rank\ \left[A^{(m-1)}(1),\cdots,A^{(1)}(1),\ A^{(0)}(1)\right]=\check{r},$$

which contradicts the definition of r in (A.12). Thus it is established that $rank\ C(1)=k-r$. Premultiplying both sides of (A.10) by $C(B)$, using that $C(B)\ A(B)=(1-B)^m\ I_k$ and that $d_1-m_1=d-m$, and defining $\check{\mu}_n$ as the ones satisfying

$$t^q\,C(1)\,\bar{\mu}_q+\sum_{n=0}^{q-1}t^n\,\check{\mu}_n=\sum_{n=0}^{q}\{C(B)t^n\}\,\bar{\mu}_n,$$

(5) follows. It is also obvious from the assumption for

the components of y_t that neither of the row vectors of $C(1)$ is zero vector. Moreover, we see from $C(1)A(1)=0$ that $C(1)\bar{\alpha}=0$, which implies that $C(1)\bar{\mu}_q$ vanishes if $\bar{\mu}_q$ is spanned by the columns of $\bar{\alpha}$. Thus the proof is completed.

To prove Theorem 3, we set up the following three lemmas:

Lemma A.1 *Suppose that the same conditions as given in Theorem 3 hold. Also, for the case $p \geq 1$, let the linear least-square predictor of u_t onto the space spanned by $\beta' v_{t-1}$, u_{t-j}, $j=1,\cdots,p$,*

$$\alpha(p)\,\beta' v_{t-1} + \sum_{j=1}^{p} H_i(p)\,u_{t-i},$$

and for the case $p=0$, let the linear least-square predictor of u_t onto the space spanned by only $\beta' v_{t-1}$ be $\alpha(0)\beta' v_{t-1}$, where u_t and v_t are defined in the text, $\alpha(p)$ is a $k\times r$ constant matrix and $H_i(p)$ are $k\times k$ constant matrices. Then:

$$S_{01}\,\beta = \alpha(p)\,\beta' R_v\,\beta + O_p(T^{-1/2}) = \bar{\alpha}\,\bar{\beta}'\,S_{21}\,\beta + O_p(T^{-1/2}),$$

where R_v is given in the text,

$$\bar{\beta} = \left[\beta_{m-1}', \cdots, \beta_1', \beta_0'\right]' \quad and \quad S_{21} = [Y_{-1;m-1}, \cdots, Y_{-1;1}, Y_{-1}]'\,\breve{M}\,Y_{-1}/T,$$

with

$$Y_{-1;n} = \left[\Delta^{d-1-n} y_{p+d}, \Delta^{d-1-n} y_{p+d+1}, \cdots, \Delta^{d-1-n} y_{T-1}\right], \qquad n = 1, \cdots, m-1.$$

Proof It suffices to give the proof for the case $p \geq 1$, since it is trivial to prove the lemma for the case $p=0$. Note that $v_t = C^{(1)}(B)\,\epsilon_t$ for $C^{(1)}(z)$ given in the proof of Theorem 1, that u_t and $\beta' v_t$ correspond to the ones obtained by removing the deterministic parts and the terms on the initial vectors from $\Delta^d y_t$ and $\beta'\Delta^{d-1} y_t$ and that

$$u_t = \alpha(p)\,\beta' v_{t-1} + \sum_{j=1}^{p} H_i(p)\,u_{t-i} + \bar{\epsilon}_{t;p},$$

where $\bar{\epsilon}_{t;p}$ is the error from the prediction stated in the lemma. Also, by the standard statistics for weakly stationary, ergodic time series and deterministic trends (Banerjee *et al*. (1993) and Johansen (1988, 1996) e.g.),

$$\sum_{t=p+d+1}^{T} t^n \left(v_{t-1}', u_t', u_{t-1}', \cdots, u_{t-p}'\right)/T^{n+1/2} = O_p(1),$$

$$\sum_{t=p+d+1}^{T} (\epsilon_t', \epsilon_t'(p))' \left(u_{t-1}', \cdots, u_{t-p}'\right)/T = O_p(T^{-1/2}),$$

$$\sum_{t=p+d+1}^{T} \beta' \breve{v}_{t-1} \breve{v}_{t-1}' \beta/T = \beta' R_v\,\beta + O_p(T^{-1/2}).$$

We note the following facts: It is not required that the weakly stationary, ergodic series in the above results are $I(0)$, and some linear combinations of v_t and u_t are overdifferenced. By arranging these results suitably or from (6), it is not difficult to obtain the desired results.

Lemma A.2 *Suppose that the same conditions as given in Lemma A.1 hold with $p \geq 1$. Also, define the $k \times k$ matrices $\bar{H}_j(p)$ as the ones satisfying*

$$I_k - \sum_{j=1}^{p+1} \bar{H}_j(p)\,z^j = \begin{bmatrix} \beta' \\ \gamma' \end{bmatrix} \left\{ \begin{bmatrix} (1-z)I_r & 0 \\ 0 & I_{k-r} \end{bmatrix} - \bar{\alpha}(p)\,[I_r, 0]\,z - \sum_{j=1}^{p} H_j(p)\left[(1-z)\,\beta\,(\beta'\beta)^{-1}, \gamma\,(\gamma'\gamma)^{-1}\right] z^j \right\}$$

and then define the k-dimensional series W_t and $\bar{\eta}_t(p)$ as

$$W_t = \begin{bmatrix} \beta' v_t \\ \gamma' u_t \end{bmatrix}, \quad \bar{\eta}_t(p) = \begin{bmatrix} \beta' \\ \gamma' \end{bmatrix} \bar{\epsilon}_{t;p}.$$

Then we have

$$\alpha(p) = -\begin{bmatrix} \beta' \\ \gamma' \end{bmatrix}^{-1} \{I_k - \sum_{i=1}^{p+1} \bar{H}_i(p)\} \begin{bmatrix} I_r \\ 0 \end{bmatrix}.$$

Proof The lemma is easily established by putting $z=1$ in the definitions for $\bar{H}_j(p)$.

Lemma A.3 *Suppose that the same conditions as given in Lemma A.2 hold. Then we have rank $\alpha(p) = r$.*

Proof Consider the case $p \geq 1$. Since $\beta' u_t = \beta' \Delta v_t$, it is also obvious that the space spanned by $[v_{t-j}'\beta, u_{t-j}'\gamma]$, $j=1,\cdots,p$, and $\beta' v_{t-p-1}$ is equal to the one spanned by $\beta' v_{t-1}$, u_{t-j}, $j=1,\cdots,p$. Based on this, it is easy to check that

$$W_t = \sum_{j=1}^{p+1} \bar{H}_j(p) W_{t-j} + \bar{\eta}_t(p),$$
$$E\{\bar{\eta}_t(p) W'_{t-j}\} = 0, \quad j = 1, \cdots, p. \qquad (A.13)$$

Next, define $W_{t-j}(p)$ of size $(kp+r) \times 1$, $\check{\eta}_t(p)$ of dimension $(kp+r)$, $\bar{H}(p)$ of size $(kp+r) \times (kp+r)$ and $\check{H}(p)$ of size $k(p+1) \times k(p+1)$ as

$$W_{t-j}(p) = \begin{bmatrix} W_{t-j} \\ W_{t-j-1} \\ \vdots \\ W_{t-j-p+1} \\ \beta' v_{t-i-p} \end{bmatrix}, \quad j = 0, 1, \quad \check{\eta}_t(p) = \begin{bmatrix} \bar{\eta}_t \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

$$\bar{H}(p) = \begin{bmatrix} \bar{H}_1(p) & \cdots & \bar{H}_{p-1}(p) & \bar{H}_p(p) & \bar{H}_{p+1;1}(p) \\ I_k & & & & \\ & \ddots & & & 0 \\ 0 & & I_k & & \\ & & & I_r & 0 \end{bmatrix},$$

$$\check{H}(p) = \begin{bmatrix} \bar{H}_1(p) & \cdots & \bar{H}_{p-1}(p) & \bar{H}_p(p) & \bar{H}_{p+1}(p) \\ I_k & & & & \\ & \ddots & & & 0 \\ 0 & & I_k & & \\ & & & I_k & \end{bmatrix},$$

where

$$\bar{H}_{p+1;1}(p) = \bar{H}_{p+1} \begin{bmatrix} I_r \\ 0 \end{bmatrix}.$$

We then see that

$$\det\left(I_{kp+r} - \bar{H}(p) z\right) = \det\left(I_{k(p+1)} - \check{H}(p)\, z\right)$$
$$= \det\left(I_k - \sum_{j=1}^{p+1} \bar{H}_j(p)\, z^j\right) \qquad (A.14)$$

since

$$\begin{bmatrix} I_k - \bar{H}_1(p)\, z & -\bar{H}_2(p)\, z & \cdots & -\bar{H}_p(p)\, z & -\bar{H}_{p+1}(p)\, z \\ -I_k\, z & I_k & & & \\ & -I_k\, z & \ddots & & 0 \\ 0 & & \ddots & I_k & \\ & & & -I_k\, z & I_k \end{bmatrix}$$

$$\times \begin{bmatrix} I_k & & & & \\ I_k\, z & I_k & & & 0 \\ \vdots & \vdots & \ddots & & \\ I_k\, z^{p-1} & I_k\, z^{p-2} & \ddots & I_k & \\ I_k\, z^p & I_k\, z^{p-1} & \cdots & I_k\, z & I_k \end{bmatrix}$$

$$= \begin{bmatrix} I_k - \sum_{j=1}^{p+1} \bar{H}_j(p)\, z^j & -\sum_{j=2}^{p+1} \bar{H}_j(p)\, z^{j-1} & \cdots & -\bar{H}_{p+1}(p)\, z \\ & I_k & & \\ & & \ddots & 0 \\ 0 & & & I_k \end{bmatrix}.$$

Also, the first equation of (A.12) is rewritten as

$$W_t(p) = \bar{H}(p)\, W_{t-1}(p) + \check{\eta}_t(p). \qquad (A.15)$$

Noting that $W_t(p)$ is weakly stationary, purely nondeterministic and ergodic with mean zero and putting

$$R_W(0) = E\{W_{t-1}(p)\, W'_{t-1}(p)\},$$
$$R_W(1) = E\{W_t(p)\, W'_{t-1}(p)\},$$

using the property of $\check{\eta}_t(p)$ stated in (A.13), (A.14) and (A.15), it can be shown that $\det\{I_k - \sum_{j=1}^{p+1} \bar{H}_j(p)\, z^j\} = 0$ is equivalent to

$$\det\{R_W(0) - R_W(1)\, z\} = 0. \qquad (A.16)$$

Now, consider z satisfying (A.16). Suppose that z is nonzero and real, noting that (A.16) does not hold for $z = 0$. Therefore, there must exist some $(kp+r)$-dimensional real vector $b \neq 0$ satisfying

$$b'\{R_W(0) - R_W(1)\, z\} = 0,$$

which leads to

$$z^{-1} = \frac{b'\, R_W(1)\, b}{b'\, R_W(0)\, b}.$$

Since z^{-1} is exactly the first-order autocorrelation coefficient of $b' W_t(p)$ that is weakly stationary, ergodic with mean zero, it must be satisfied that $|z| > 1$.

Next, suppose that z and $\bar{z}$ are a pair of complex-conjugate roots of (A.16). Then we can find $(kp+r) \times 1$ complex vector b and $\bar{b}$ satisfying

$$b'\{R_W(0) - R_W(1)\, z\} = 0, \bar{b}'\{R_W(0) - R_W(1)\, \bar{z}\} = 0.$$

With i denoting the imaginary, $(kp+r)$-dimensional real vectors b_j and real numbers z_j ($j = 1, 2$) such that $b_j \neq 0$ for at least one j and $z_2 \neq 0$, we can let

$$b = b_1 + ib_2, \quad z = z_1 + iz_2.$$

Since both real and imaginary parts of $b'\{R_W(0) - R_W(1)\, z\}$ must be zero, we have

$$\{b_1' R_W(0)\, b_1\}^2 = |(z_1 b_1' - z_2 b_2')\, R_W(1)\, b_1|^2,$$
$$\{b_2' R_W(0)\, b_2\}^2 = |(z_1 b_2' + z_2 b_1')\, R_W(1)\, b_2|^2. \qquad (A.17)$$

This requires that either $b_1 \neq 0$, $z_1 b_1 - z_2 b_2 \neq 0$ or $b_2 \neq 0$, $z_1 b_2 + z_2 b_1$ holds, recalling that $b_j \neq 0$ for at least one j. Noting that $(z_1 b_1' - z_2 b_2')\, R_W(0)\, (z_1 b_1 - z_2 b_2)$ and $b_1' R_W(0)\, b_1$ are the variances of $(z_1 b_1' - z_2 b_2')\, W_t(p)$ and $b_1' W_{t-1}(p)$ respectively and that $(z_1 b_1' - z_2 b_2')\, R_W(1)\, b_1$ is the covariance of those series, we have

$$\begin{aligned} &|(z_1 b_1' - z_2 b_2')\, R_W(1)\, b_1|^2 \\ &\leq \{(z_1 b_1' - z_2 b_2')\, R_W(0)\, (z_1 b_1 - z_2 b_2)\}\, (b_1' R_W(0)\, b_1) \\ &= \{z_1^2\, b_1' R_W(0)\, b_1 + z_2^2\, b_2' R_W(0)\, b_2 \\ &\quad - 2 z_1 z_2\, b_1\, R_W(0)\, b_2\}\, (b_1' R_W(0)\, b_1). \end{aligned}$$

Similarly,

$$\begin{aligned} &|(z_1 b_2' + z_2 b_1')\, R_W(1)\, b_2|^2 \\ &\leq \{z_1^2\, b_2' R_W(0)\, b_2 + z_2^2\, b_1' R_W(0)\, b_1 \\ &\quad + 2 z_1 z_2\, b_1\, R_W(0)\, b_2\}\, (b_2' R_W(0)\, b_2). \end{aligned}$$

By combining these with (A.17), it is led to that

$$\begin{aligned} &b_1' R_W(0)\, b_1 \leq z_1^2\, b_1' R_W(0)\, b_1 + z_2^2\, b_2' R_W(0)\, b_2 \\ &- 2 z_1 z_2\, b_1\, R_W(0)\, b_2, \end{aligned}$$

$$\begin{aligned} &b_2' R_W(0)\, b_2 \leq z_1^2\, b_2' R_W(0)\, b_2 + z_2^2\, b_1' R_W(0)\, b_1 \\ &+ 2 z_1 z_2\, b_1\, R_W(0)\, b_2. \end{aligned}$$

By the restriction on b_i and the properties of the series stated above, at least one of the above two inequalities must hold strictly. By adding each of both sides of the first inequality to those of the second one, we attain to

$$\begin{aligned} &\{b_1' R_W(0)\, b_1 + b_2' R_W(0)\, b_2\} \\ &< (z_1^2 + z_2^2)\{b_1' R_W(0)\, b_1 + b_2' R_W(0)\, b_2\}, \end{aligned}$$

which requires that $|z| > 1$.

Thus all the roots of $\det\{I_k - \sum_{j=1}^{p+1} \bar{H}_j(p)\, z^j\} = 0$ must be greater than 1 in absolute values, which, together with Lemma A.2, implies that the result required for the lemma holds for the case $p \geq 1$. Since it is trivial to obtain the result for the case $p = 0$, we can now finish the proof.

Proof of Theorem 3 We first note that this theorem is essentially the same as the counterparts of Johansen (1988, 1996) and many part of the proof can be proved by the same manner as used for such ones. However, several matters need to be proved here.

Now, put

$$\check{E}_0 = [\varepsilon_{p+d+1},\ \varepsilon_{p+d+2}, \cdots,\ \varepsilon_T]',$$

let $C(1)$ be $\gamma\delta'$ for some $k \times (k-r)$ column full rank matrices γ and δ. Also, without loss of generality, for the case where $\check{\mu}_q$ is not spanned by the columns of $\bar{\alpha}$ and $k - r > 1$, let us partition γ as

$$\gamma' = \begin{bmatrix} \gamma_1' \\ \gamma_2' \end{bmatrix}$$

for some k-dimensional vector γ_1 and $k \times (k-r-1)$ matrix γ_2 such that $\gamma_1' C(1)\, \check{\mu}_q \neq 0$ and $\gamma_2' C(1)\, \check{\mu}_q = 0$. Moreover, note that S_{ij} are obtained by removing the deterministic parts and the terms on the initial vectors. By (1)/(5) and (A.8) in the proof of Lemma A.1 and using the elementary of Brownian motion and the asymptotics for $I(1)$ series (Park and Phillips (1988, 1889), Banerjee *et al*. (1993) or Johansen (1988, 1996) e.g.) in addition to the statistics for weakly stationary, ergodic time series and deterministic trends mentioned in the proof of Lemma A.1, it can be established that

$$\check{D}_T^{-1}\, (\gamma' S_{11}\, \gamma / T)\, \check{D}_T^{-1} \Rightarrow \tilde{G}\, \left(\int_0^1 \tilde{W}(u) \tilde{W}'(u) du\right)\, \tilde{G} \qquad (A.18)$$

$$\beta' S_{11} \beta = \beta' R_v \beta + O_p(T^{-1/2}),$$
$$(\beta' S_{11} \beta)^{-1} = (\beta' R_v \beta)^{-1} + O_p(T^{-1/2}), \qquad (A.19)$$

$$S_{00} = R_u + O_p(T^{-1/2}), \quad S_{00}^{-1} = R_u^{-1} + O_p(T^{-1/2}),$$
$$\check{E}_0'\, \check{M} \check{E}_0 = \Omega + O_p(T^{-1/2}), \qquad (A.20)$$

$$(\beta' S_{11}\, \gamma)\, \check{D}_T^{-1} = O_p(1), \quad (S_{01}\, \gamma)\, \check{D}_T^{-1} = O_p(1). \qquad (A.21)$$

where $\check{D}_T$ and $\tilde{G}$ are defined as

$$I_{k-r} \quad \textit{and} \quad (\gamma' C(1)\, \Omega\, C(1)'\, \gamma)^{1/2}$$

respectively if $\check{\mu}_q$ is spanned by the columns of $\bar{\alpha}$, those are defined as

$$T^{q+3/2} \quad \textit{and} \quad |\gamma' C(1)\, \check{\mu}_q|$$

respectively if $\check{\mu}_q$ is not spanned by the columns of $\bar{\alpha}$ and $k - r = 1$, and those are defined as

$$\begin{bmatrix} T^{q+3/2} & 0 \\ 0 & I_{k-r-1} \end{bmatrix}$$
$$\textit{and} \quad \begin{bmatrix} |\gamma_1' C(1)\, \check{\mu}_q| & 0 \\ 0 & (\gamma_2' C(1)\, \Omega\, C(1)'\, \gamma_2)^{1/2} \end{bmatrix}$$

respectively otherwise, and Ω, R_v and R_u are given in the text. It is also led to automatically from (6) that

$$\bar{\delta}'\,S_{00}\,\bar{\delta} = \bar{\delta}'\,\breve{E}_0'\,\breve{M}\,\breve{E}_0\,\bar{\delta}, \tag{A.22}$$

$$(S_{01}\,\gamma)\,\breve{D}_T^{-1} = \bar{\alpha}\,(\bar{\beta}'\,S_{21}\,\gamma)\,\breve{D}_T^{-1} + (\breve{E}_0'\,\breve{M}\,Y_{-1}\,\gamma/T)\,\breve{D}_T^{-1}, \tag{A.23}$$

where $\bar{\delta}$ is a $k\times(k-r)$ column full rank matrix such that $\bar{\delta}'\,\bar{\alpha}=0$ and $\bar{\beta}$ and S_{21} are given in Lemma A.1. Moreover, from (A.23) and using the elementary of Brownian motion and the asymptotics for $I(1)$ mentioned above again, we have

$$(\bar{\delta}'\,S_{01}\,\gamma)\,\breve{D}_T^{-1} \Rightarrow \left(\bar{\delta}'\,\Omega\,\bar{\delta}\right)^{1/2}\left(\int_0^1 dW(u)\tilde{W}'(u)du\right)\tilde{G}. \tag{A.24}$$

For the derivation of (i), notice that $(T\,\hat{\lambda}_k)^{-1}\geq\cdots\geq(T\,\hat{\lambda}_1)^{-1}$ are the ordered eigenvalues of the equation

$$\det\{\begin{bmatrix} \beta' S_{11}\beta/T & \beta' S_{11}\gamma/T \\ \breve{D}_T^{-1}\,(\gamma' S_{11}\beta/T) & \breve{D}_T^{-1}\,(\gamma' S_{11}\gamma/T)\,\breve{D}_T^{-1} \end{bmatrix} - \mu\begin{bmatrix} \beta' S_{10}S_{00}^{-1}S_{01}\beta & \beta' S_{10}S_{00}^{-1}(S_{01}\gamma)\,\breve{D}_T^{-1} \\ \breve{D}_T^{-1}\,(\gamma' S_{10})S_{00}^{-1}S_{01}\beta & \breve{D}_T^{-1}\,(\gamma' S_{10})S_{00}^{-1}(S_{01}\gamma)\,\breve{D}_T^{-1} \end{bmatrix}\} = 0. \tag{A.25}$$

Now, put

$$\hat{M} = S_{00}^{-1}\left[I_k - S_{01}\,\beta\left(\beta'\,S_{01}\,S_{00}^{-1}\,S_{01}\,\beta\right)^{-1}\beta'\,S_{01}\,S_{00}^{-1}\right],$$

and let $\hat{\nu}_{k-r}\geq\cdots\geq\hat{\nu}_1$ denote the ordered eigenvalues of the equation

$$\det\{\breve{D}_T^{-1}\,(\gamma'\,S_{11}\,\gamma/T)\breve{D}_T^{-1} - \mu\,\breve{D}_T^{-1}\,(\gamma'\,S_{10}\,\hat{M}\,S_{01}\,\gamma)\,\breve{D}_T^{-1}\} = 0.$$

By using (A.18) to (A.21) in (A.25), it can be shown that

$$T\,\hat{\lambda}_{r+h} = \hat{\nu}_h + O_p(T^{-1}), \qquad h = 1,\cdots,k-r. \tag{A.26}$$

On the other hand, Lemma A.1 ensures that

$$\bar{\beta}'\,S_{21}\,\beta = O_p(1), \qquad (\bar{\beta}'\,S_{21}\,\beta)^{-1} = O_p(1).$$

It is also easy to check that$=\hat{M}'=\hat{M}$, $\hat{M}S_{01}\,\beta=0$ and $\beta' S_{10}\,\hat{M}=0$. These results, together with Lemma A.1, imply that $\bar{\alpha}'\hat{M}=O_p(T^{-1/2})$ and $\hat{M}\,\bar{\alpha}=O_p(T^{-1/2})$. Hence,

$$\hat{M} = \bar{\delta}\,\bar{\delta}' + O_p(T^{-1/2}), \tag{A.27}$$

for some $\bar{\delta}$ as given in (A.22) or (A.24). Moreover, it is easily checked that $M\,S_{00}\,\hat{M}=\hat{M}$, in other words,

$$\bar{\delta}\,\bar{\delta}'\,S_{00}\,\bar{\delta}\,\bar{\delta}' = \bar{\delta}\,\bar{\delta}' + O_p(T^{-1/2}),$$

which leads to

$$\bar{\delta}'\,S_{00}\,\bar{\delta} = I_{k-r} + O_p(T^{-1/2}).$$

It follows from this, (A.20) and (A.22) that

$$\bar{\delta}'\,\Omega\,\bar{\delta} = I_{k-r}. \tag{A.28}$$

Putting (A.18), (A.24), (A.27) and (A.28) together, we see that the limiting distribution of $\sum_{h=1}^{k-r}\hat{\nu}_h$ becomes the one stated by (ii). It is also obvious from this and (A.26) that $\hat{\lambda}_{r+h}=O_p(T^{-1})$ for $h=1,\cdots,k-r$, which leads to

$$-T\log\{1-\hat{\lambda}_{r+h}\}=T\,\hat{\lambda}_{r+h}+O_p(T^{-1}) \quad h=1,\cdots,k-\text{r}.$$

The result required for (i) follows from this, (A.26) and the result on the limiting distribution of $\sum_{h=1}^{k-r}\hat{\nu}_h$ established above.

For (ii), note that $\beta' R_v\,\beta$ has full rank, since $\beta'\breve{u}_t$ is not degenerated. Also, from the nature of the linear least-square prediction, we have

$$\alpha\,(p)=R_{uv}\,\beta\,(\beta' R_v\,\beta)^{-1}, \tag{A.29}$$

which, together with Lemmas A.1 and A.3, ensures that $1>\nu_1\geq\cdots\geq\nu_r>0$. Moreover, letting $\tilde{\nu}_1\geq\cdots\geq\tilde{\nu}_r$ be the ordered eigenvalues of the equation

$$\det\{\lambda\,\beta' S_{11}\,\beta-\beta' S_{10}\,S_{00}^{-1}\,S_{01}\,\beta\},$$

it follows from Lemma A.1, (A.19), (A.20) and (A.29) that

$$\tilde{\nu}_j=\nu_j+O_p(T^{-1/2}), \quad j=1,\cdots,r. \tag{A.30}$$

On the other hand, it is trivial to see that the roots of the equation $\det\{\lambda\,S_{11}-S_{10}S_{00}^{-1}\,S_{01}\}=0$ are equivalent to those of

$$\det\{\lambda\begin{bmatrix} \beta' S_{11}\beta & (\beta' S_{11}\gamma/T^{1/2})\,\breve{D}_T^{-1} \\ \breve{D}_T^{-1}\,(\gamma' S_{11}\beta/T^{1/2}) & \breve{D}_T^{-1}\,(\gamma' S_{11}\gamma/T)\,\breve{D}_T^{-1} \end{bmatrix}$$

$$
-\left[\begin{matrix} \beta' S_{10} S_{00}^{-1} S_{01}\beta & \left(\beta' S_{10} S_{00}^{-1} S_{01}\gamma/T^{1/2}\right)\breve{D}_T^{-1} \\ \breve{D}_T^{-1}\left(\gamma' S_{10} S_{00}^{-1} S_{01}\beta/T^{1/2}\right) & \breve{D}_T^{-1}\left(\gamma' S_{10} S_{00}^{-1} S_{01}\gamma/T\right)\breve{D}_T^{-1} \end{matrix}\right]\} = 0.
$$

It follows from this, Lemma A.1 and (A.18) to (A.21) that

$$
\hat{\lambda}_j = \tilde{\nu}_j + O_p(T^{-1/2}) \qquad j=1, \cdots, r. \tag{A.31}
$$

By combining (A.31) with $1 > \nu_1 \geq \cdots \geq \nu_r > 0$ and (A.30), (ii) is established.

TABLE 1

Relative Frequency of a Correct Determination of r

DGP	Significance Level: 5%			Significance Level: 1%		
	$g_1=0$	$g_1=1$	$g_1=0$	$g_1=0$	$g_1=1$	$g_1=0$
	$g_2=0$	$g_2=0$	$g_2=1$	$g_2=0$	$g_2=0$	$g_2=1$
Case (1)	93.72	93.16	92.69	98.43	98.46	98.52
Case (2)	93.55	93.48	92.66	98.6	98.49	98.32
Case (3)	93.96	93.22	92.75	98.68	98.46	98.46
Case (4)	93.41	92.89	92.96	98.48	98.57	98.44
Case (5)	93.2	93.14	92.35	98.61	98.35	98.2
Case (6)	93.23	92.83	93.64	98.45	98.34	97.4
Case (7)	94.18	94.01	93.59	98.9	98.66	98.5
Case (8)	93.71	93.85	93.67	98.6	98.64	98.32

TABLE 2

Relative Frequency of detecting the true value of r: The 1st Group

DGP＼Test	$\hat{Q}_j$	$\hat{Q}_{*;j}$	LR	$P(k,j)$		$P_*(k,j)$	
K_T				4	8	4	8
VMA: $c_1=c_2=c_3=c_4=1.5$, $c_5=0$ and $c_6=c_7=c_8=c_9=0.56$							
				$T=200$			
$r=0$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
$r=1$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
$r=2$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
				$T=500$			
$r=0$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
$r=1$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
$r=2$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
VMA: $c_1=c_2=c_3=c_4=0.2$, $c_5=0$ and $c_6=c_7=c_8=c_9=-0.48$							
				$T=200$			
$r=0$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
$r=1$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
$r=2$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
				$T=500$			
$r=0$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
$r=1$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
$r=2$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
VMA: $c_1=1.5$, $c_2=c_4=0.9$, $c_3=0.4$, $c_5=0.5$, $c_6=0.56$, $c_7=c_9=0.2$ and $c_8=0.2$							
				$T=200$			
$r=0$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
$r=1$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
$r=2$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
				$T=500$			
$r=0$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
$r=1$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
$r=2$	93.72	93.16	92.69	98.43	98.46	96.4	98.52

In all the tables, figures indicate percentiles, and K_T is required only for $P(k,j)$ and $P_*(k,j)$.

TABLE 3

Relative Frequency of detecting the true value of r: The 3rd Group

DGP＼Test	$\hat{Q}_j$	$\hat{Q}_{*;j}$	L_R	$P(k,j)$		$P_*(k,j)$	
K_T				4	8	4	8
$q=2$: $c_1=c_2=c_3=c_4=0.6$, $c_5=0$ and $c_6=c_7=c_8=c_9=0$							
			$T=200$				
$r=2$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
			$T=500$				
	93.72	93.16	92.69	98.43	98.46	96.4	98.52
$q=1$: $c_1=c_2=c_3=c_4=-0.6$, $c_5=0$ and $c_6=c_7=c_8=c_9=0$							
			$T=200$				
$r=2$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
			$T=500$				
	93.72	93.16	92.69	98.43	98.46	96.4	98.52
$q=2$: $c_1=c_2=c_3=c_4=1.0$, $c_5=0$ and $c_6=c_7=c_8=c_9=0.24$							
			$T=200$				
$r=2$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
			$T=500$				
	93.72	93.16	92.69	98.43	98.46	96.4	98.52
$q=2$: $c_1=c_2=c_3=c_4=1.5$, $c_5=0$ and $c_6=c_7=c_8=c_9=0.56$,							
			$T=200$				
$r=2$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
			$T=500$				
	93.72	93.16	92.69	98.43	98.46	96.4	98.52
$q=2$: $c_1=1.5$, $c_2=c_4=0.9$, $C_3=0.4$, $c_5=0.5$, $c_6=0.56$, $c_7=c_9=0.2$ and $c_8=0.2$							
			$T=200$				
$r=2$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
			$T=500$				
	93.72	93.16	92.69	98.43	98.46	96.4	98.52

TABLE 3 (Continued)

Relative Frequency of detecting the true value of r: The 3rd Group

DGP＼Test	$\hat{Q}_j$	$\hat{Q}_{*;j}$	LR	$P(k,j)$		$P_*(k,j)$	
K_T				4	8	4	8
$q=2$: $c_1=1.5$, $c_2=0.9$, $c_3=0.4$, $c_4=-0.4$, $c_5=0.5$, $c_6=0.56$, $c_7=c_9=0.2$ and $c_8=0$							
			$T=200$				
$r=2$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
			$T=500$				
	93.72	93.16	92.69	98.43	98.46	96.4	98.52

TABLE 4

Relative Frequency of detecting the true value of r: The 4th Group

DGP＼Test	$\hat{Q}_j$	$\hat{Q}_{*;j}$	LR	$P(k,j)$		$P_*(k,j)$	
K_T				4	8	4	8
$q=1$: $c_1=c_2=c_3=c_4=0.6$, $c_5=0$ and $c_6=c_7=c_8=c_9=0$							
			$T=200$				
$r=1$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
			$T=500$				
	93.72	93.16	92.69	98.43	98.46	96.4	98.52
$q=1$: $c_1=c_2=c_3=c_4=-0.6$, $c_5=0$ and $c_6=c_7=c_8=c_9=0$							
			$T=200$				
$r=1$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
			$T=500$				
	93.72	93.16	92.69	98.43	98.46	96.4	98.52
$q=2$: $c_1=c_2=c_3=c_4=1.0$, $c_5=0$ and $c_6=c_7=c_8=c_9=0.24$							
			$T=200$				
$r=1$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
			$T=500$				
	93.72	93.16	92.69	98.43	98.46	96.4	98.52
$q=2$: $c_1=c_2=c_3=c_4=1.5$, $c_5=0$ and $c_6=c_7=c_8=c_9=0.56$,							
			$T=200$				
$r=1$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
			$T=500$				
	93.72	93.16	92.69	98.43	98.46	96.4	98.52
$q=2$: $c_1=1.5$, $c_2=c_4=0.9$, $c_3=0.4$, $c_5=0.5$, $c_6=0.56$, $c_7=c_9=0.2$ and $c_8=0.2$							
			$T=200$				
$r=1$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
			$T=500$				
	93.72	93.16	92.69	98.43	98.46	96.4	98.52

TABLE 4 (Continued)

Relative Frequency of detecting the true value of r: The 4th Group

DGP＼Test	$\hat{Q}_j$	$\hat{Q}_{*;j}$	LR	$P(k,j)$		$P_*(k,j)$	
K_T				4	8	4	8
$q=2$: $c_1=1.5$, $c_2=0.9$, $c_3=0.4$, $c_4=-0.4$, $c_5=0.5$, $c_6=0.56$, $c_7=c_9=0.2$ and $c_8=0$							
			$T=200$				
$r=1$	93.72	93.16	92.69	98.43	98.46	96.4	98.52
			$T=500$				
	93.72	93.16	92.69	98.43	98.46	96.4	98.52