FINITE LAG ORDER VECTOR AUTOREGRESSIONS AND COINTEGRATING RANK DETECTION

MITSUHIRO ODAKI

Abstract

This paper discusses on how the number of independent cointegrating relations known as the cointegrating rank can be formulated and detected when some finite lag order vector autoregressive (VAR) schemes are fitted without imposing the assumptions which make the Granger representation theorem (GRT) hold. Adopting a generalized framework on the data generation processes (DGPs) and theoretically formulating each of the VAR schemes as a linear least-square predictor, we show that it precisely captures the cointegrating rank even if the existence of the VAR representation in GRT is not ensured. It is also established that estimating the rank through direct application of one of the information criteria under any finite lag order VAR scheme leads to some asymptotic desirability such as the conventional consistency. For finite sample performances of the estimation procedure proposed, some Monte Carlo experiments are executed, and it is observed that those are not so far from the asymptotics established theoretically, although affected by the selection of the scheme fitted or its lag order. We also point out that under finite sample sizes, the schemes specified by comparatively small lags such as 1 to 3 tend to produce desirable estimation results.

1 Introduction

Among model formulations for a system of multivariate economic time series, the *vector autoregression* (VAR) has been considered to be the handiest one and applied widely in a large amount of econometric researches. In most of such VAR formulations, embodying the empirical belief that many of time series considered are integrated of some orders need to be cared. Particularly, numerous econometric researches have been concentrated on the situation in which individual time series are integrated of order 1. Organizing VARs through differencing the data series considered was the approach adopted mainly until early 1980's along the prosperity of the Box and Jenkins methodology.

On the other hand, the concept of *cointegration* has been playing an important role in both theoretical and empirical econometrics since it was formulated by Engle and Granger (1987) and others. It brought about a significant change in the VAR model formulation along with recent development of the inference theory for integrated processes, emphasizing that the framework of VARs in differences is not always valid and that some linear combinations of individual series may be weakly stationary.

Following *Granger's representation theorem* (GRT) in Engle and Granger (1987), a cointegrated system whose individual data series is integrated of order 1 is expressed as a VAR in levels of the data series or an vector error correction model (VECM) as an equivalent form, in which such linear combinations referred as *the cointegrating relations* as well as first differences are included with serially uncorrelated error vectors.[1] The number of independent cointegrating relations, called *the cointegrating rank*, is essential and indispensable for the model formulation, parametrization and inferences under the occurrence of cointegration, particularly under the VAR approach in which the system consisting of more than 3 series tends to be considered and consequently the cointegrating rank may be greater than 1.

Conventionally, the issue of detecting the cointegrating rank has been dealt with based on the fitting of data series to a finite lag order (or lag length) VAR in levels and Johansen's rank test (trace test) (see Johansen (1988, 1992b)) for estimation of the true value of the rank. It is also conventional that the VAR lag order is automatically determined through one of the information criteria used in the model selection prior to 'estimation' of the rank. It is widely accepted that the Akaike information criterion introduced in Akaike (1973), whose asymptotic evaluation is established by Shibata (1976) etc., the Bayesian information criterion originated by Schwartz (1978) and the criterion in Hannan and Quinn (1979), referred as AIC, SIC and HQ respectively following the conventional manner, are examples of such criteria.

Among many researches utilizing such an information criterion for the lag order selection, Aznar and Salvador (2002) proposed to apply not Johansen's rank test but it to estimation of the cointegrating rank itself. They show that using the criterion which is established to possess the consistency property in the conventional statistical analyses such as SIC or HQ, simultaneous determination of both the rank and the VAR lag order achieves consistent estimation. However, those methods including Aznar and Salvador's one are not available unless the lag order of the VAR in GRT is finite. As a way to make up this defect, Shintani (2001) proposed a nonparametric test which is less powerful than Johansen's rank test without formulating any VAR scheme. On the other hand, Saikkonen (1992) considered a VAR whose lag order is finite but 'large' as an approximation of the infinite lag order VAR, in the sense that it increases at a 'slower' rate as the sample size goes to infinity, and derived such asymptotics as in the Johansen's rank test. Qu and Perron (2007) discussed the determination of an optimal lag order based on such a VAR approximation and one of information criteria, although the rank estimation using it was not dealt with. It should be noted that all the approaches stated above are based on the supposition that GRT holds.

We should not overlook that there are some cointegrated systems in which the VAR approximation is insufficient or GRT itself does not hold. As mentioned in Introduction of Qu and Perron (2007), such a situation occurs if the data generation process (DGP) expressed as a vector moving average (VMA) possesses a root close to one in the VMA characteristic equation, and a condition/restriction to rule out the occurrence of *polynomial cointegration or multicointegration*, discussed by the literatures such as Granger and Lee (1990), Engle and Yoo (1991), Gregoir and Laroque (1993) and Stock and watson (1993), or *noninvertibility/overdifferencing* in some time series system is indispensable for GRT/the VAR derivation itself, as mentioned in later section. It should be also recognized that polynomial cointegration requires another type of VAR representation, formed in not only in levels but also in their integrated ones, which is not suitable for formulating the cointegrating rank (see Theorem 2.1 of Gregoir and Laroque (1993)). Besides, the exclusion of noninvertibility following from a factor except overdifferencing must be assumed for GRT, and recall that even if GRT holds, the VAR is not always described by a finite lag. A similar matter rises in the case in which GRT holds and the VAR lag order is finite as well: it is on whether fitting a VAR scheme of a lag order smaller than the true one can lead to effective detection of the true rank.

We can fit our data series to any finite lag order VAR scheme/model even if the difficulty on GRT stated above occurs, although formulating a VAR with serially uncorrelated errors is not expected. Actually, many empirical researches on cointegration have been based on such a finite lag order VAR fitting and the use of one of the information criteria without verifying whether GRT or the matter raised above is realized or not. Under such a background, this paper is aroused by the question whether some meaningful detection for the cointegrating rank based on a finite lag order VAR scheme can be achieved or not, provided that such difficulty on GRT occurs. It may be necessary for some resolution to consider the theoretical formulation/implication of a VAR scheme subject to such a matter, and seeking a procedure for meaningful and effective estimation of the true rank value under such a scheme must be examined as well.

The purpose of this paper is to provide a clear resolution of the matters above, motivated by the belief that those have been rarely considered in empirical researches. Supposing the DGP as a VMA, we seek a theoretical representation for a finite lag order VAR scheme fitted under the situation where the matters including the failure of GRT itself may occur. The usual concept of cointegration will be extended to one such that such matters are dealt with well. It is pointed out that the concept of linear least-square (l.l.s.) prediction or projection (see Whittle (1983, p. 9) e.g.) provides exact formulation for our purpose: VAR schemes characterized by finite lags are interpreted as a l.l.s. predictors. It is shown that the cointegrating rank is precisely reflected on some matrix parameter in each of such schemes, accompanied with the derivation of related theoretical properties. For estimation of the true rank value, we propose to adopt one of such information criteria as the above-mentioned ones as a 'method'. In such a method we construct the related statistics based on residual matrices from reduced rank regression on each of the VAR schemes as in the Johansen's rank test. What should be emphasized is that unlike the Johansen's test, the procedure proposed in this paper only pursues to estimate the rank through the direct application of one of the information criteria under an arbitrary finite lag order VAR scheme, whereas determining an 'optimal' lag order is not needed and the conventional asymptotics for the Johansen's test are not established here since the error vectors in the VAR scheme may be serially correlated. It is established that using the information criterion under such a VAR scheme leads to the conventional asymptotics such as the consistency, similar to those used in the conventional statistical analyses or the approach of Aznar and Salvador (2002), emphasizing that the asymptotic validity holds whatever the VAR scheme fitted is. It is also noticed that the properties on the VAR scheme stated above are indispensable for those asymptotics. Monte Carlo experiments are executed in some particular examples/DGPs and sample sizes 100, 200 and 500 in order to investigate finite sample performances of the criteria mentioned above. The experimental results reveal that each criterion strongly depends on the VAR lag order unlike one claimed by the asymptotics, and it will be recognized that the results close to the asymptotics are mostly realized under the schemes by small lags such as 1 to 3, particularly for some information criterion such as SIC. Generally, the results brought about by the information criteria are not unsatisfactory compared with those through application of Johansen's test.

The paper is organized as follows. Section 2 formulates the DGP and some preliminary concepts. The results on the VAR formulation stated in the above paragraph are in Section 3. Section 4 is used for presenting the rank estimation procedure and related information criteria. Asymptotics for the procedure in Section 4 are established in Section 5. Section 6 deals with Monte Carlo experiments. The remained issues including some concluding remarks are discussed in Section 7. The proofs of lemmas, theorems and a corollary in the text, together with some preliminary results, are provided in Appendix.

2 The DGP and some preliminaries

Let us begin our discussion by conventionalizing some notations appeared in the text. The symbols L and Δ are the lag and difference operators defined as $L^j u_t = u_{t-j}$ and $\Delta^j u_t = (1-L)^j u_t$ for any positive integer j, with a time series u_t. The determinant of a square matrix F is denoted as det F, I_m denotes the $m \times m$ identity matrix and $\|F\|$ denotes the Euclidean distance of F.[2] In connection with $F(z)$ denoting a power series of a complex variable z with matrix coefficients $F_j, j \geq 0$:

$$F(z) = \sum_{j=0}^{\infty} F_j z^j,$$

$F(L)$ and $F(1)$ are defined as

$$F(L) = \sum_{j=0}^{\infty} F_j L^j, \qquad F(1) = \sum_{j=0}^{\infty} F_j.$$

All power series in the text are defined over the complex plane and all notations in the text except z are interpreted as real numbers or vectors/matrices which consist of components of real numbers.

Next, without losing the natures of the usual definitions of $I(0)$ or $I(1)$ time series (see Banerjee et.al (1993, p. 84) e.g.) and cointegration (see Engle and Granger (1987) or Banerjee et.al (1993, p. 145) e.g.), let us extend those to:

Definition 1 A scalar time series η_t with mean zero and no deterministic component is said to be $I(0)$ if η_t is weakly stationary with a moving average (MA) representation following the Wold decomposition and $O_p(1)$ property and its partial $\sum_{h=1}^{t}\eta_h$ is of $O_p(t^{1/2})$, and η_t is said to be $I(1)$ if its first difference $\Delta\eta_t$ is of $I(0)$ for any $t \geq 1$ and η_0 is of $O_p(1)$.

Definition 2 A $\bar{n}$-dimensional vector time series $\bar{\eta}_t$ is said to be *cointegrated* if all the elements of $\bar{\eta}_t$ are of $I(1)$ in stochastic parts and there exists a column full rank constant (nonrandom) matrix $\bar{b}$ of $\bar{n}\times\bar{m}$ such that the stochastic part of $\bar{b}'\bar{\eta}_t$ plus a $\bar{m}$-dimensional random vector of $O_p(1)$ which does not depend upon t is weakly stationary and of $O_p(1)$ with a $\bar{m}$-dimensional VMA representation following the vector version of the Wold decomposition and an integer $\bar{m}$ satisfying $\bar{n}-1 \geq \bar{m} \geq 1$, and then $\bar{m}$ and $\bar{b}$ are called *cointegrating rank* and *cointegrating matrix* respectively.

Following Definition 1, MA processes are regarded as $I(0)$ unless overdifferenced, and other type of noninvertibility, caused by some root other than 1 in the MA characteristic equation, is acceptable.[3] Similarly, Definition 2 does not ensure that any linear combination of $\bar{b}'\bar{\eta}_t$ is of $I(0)$ stochastically, unlike the usual definition of cointegration. In other words, the situation in which some of the linear combinations of $\bar{b}'\bar{\eta}_t$ are overdifferenced, referred as *higher-order cointegration*, is allowable.[4] It should be also noted that for the case in which the random vector added is constant, the stochastic part of $\bar{b}'\bar{\eta}_t$ is weakly stationary and that VMA representations are accompanied with purely nondeterministic series or their covariance matrix which are positive definite.

Consider a $k-$variates vector time series y_t whose components are of $I(1)$ in stochastic parts. Without losing generality, the DGP is formulated as a VMA representation: based on the power series $C(z)$ and $C^{(1)}(z)$ given as

$$C(z)=I_k+\sum_{i=1}^{\infty}C_i z^i, \qquad C^{(1)}(z)=\sum_{i=1}^{\infty}\Big(-\sum_{h=i+1}^{\infty}C_h\Big)z^i,$$

with $k\times k$ constant matrices C_j such that $\sum_{i=1}^{\infty}i^{\bar{\nu}}\|C_i\|<\infty$ for some real number $\bar{\nu}\geq 1$ and the row vectors of $C(1)$ are all nonzero, $\bar{y}_t$ as $\bar{y}_t=y_t-Ey_t$ and $\{\varepsilon_t;\ t=\cdots,-1,0,1,\cdots,\}$ as a sequence of unobservable $k-$variates random vectors such that $E\varepsilon_t=0$, $E\varepsilon_t\varepsilon_t'=\Lambda$ with a positive definite matrix Λ and $E\varepsilon_t\varepsilon_{t'}'=0$ for any integers $t\neq t'$,

$$\Delta\bar{y}_t=C(L)\,\varepsilon_t=C(1)\,\varepsilon_t+C^{(1)}(L)(1-L)\,\varepsilon_t, \qquad t=1,2,\cdots, \quad (1)$$

noting that

$$C(z)=C(1)+(1-z)C^{(1)}(z). \quad (2)$$

From (1) we derive

$$\bar{y}_t=C(1)\Big(\sum_{h=1}^{t}\varepsilon_h\Big)+v_t+\xi_0, \qquad t=1,2,\cdots, \quad (3)$$

where $v_t=C^{(1)}(L)\,\varepsilon_t$ and $\xi_0=y_0-Ey_0-C^{(1)}(L)\,\varepsilon_0$.

Now, put $rank\ C(1)=s$ and $r=k-s$ with an integer s such that $1\leq s\leq k$. We can find column full rank constant matrices γ, τ and δ such that

$$C(1)=\gamma\tau\delta', \quad \gamma: k\times s, \quad \tau: s\times s, \quad \delta; k\times s.$$

Hereafter we impose $y_0=O_p(1)$ as some suitable initial condition on y_t. It is obvious from (3) that all the elements of $\bar{y}_t$ or all nonzero linear combinations of $\gamma'\bar{y}_t$ are of $I(1)$. If $s<k$ (or equivalently $r\geq 1$), there exits a column full rank constant matrix β of $k\times r$ such that $\beta'\gamma=0$, and $\beta'\bar{y}_t$ is of $O_p(1)$, as clarified by

$$\beta'\bar{y}_t=\beta' v_t+\beta'\xi_0, \qquad t=1,2,\cdots, \quad (4)$$

following from (3). It is seen from (4) that $\beta'\bar{y}_t$ is weakly stationary if either $\beta'\xi_0$ is out of consideration or stronger initial conditions of $\bar{y}_t$ and ε_t such as $y_0=Ey_0$ and $\varepsilon_{-j}=E\varepsilon_{-j}$, $j=0,1,\cdots$, are imposed. Thus, if $s<k$, we can consider $\bar{y}_t$ (or y_t) cointegrated with the cointegrating rank r and cointegrating matrix β, whereas $\bar{y}_t$ is not so if $s=k$. For discussion in the following section, we also provide the following relation here:

$$\begin{bmatrix} \beta' v_t \\ \gamma' \Delta \bar{y}_t \end{bmatrix} = \begin{bmatrix} \beta' C^{(1)}(L) \\ \gamma' C(L) \end{bmatrix} \epsilon_t, \qquad t=1, 2, \cdots. \quad (5)$$

In general, y_t may possibly possess some deterministic trends and drift formed as a q−th order polynomial of time t, expressed as

$$Ey_t = \sum_{j=0}^{q} \check{\mu}_j t^j, \qquad t=1, 2, \cdots, \quad (6)$$

with k−dimensional constant vectors $\check{\mu}_j$.[5] It is in turn derived from (6) that

$$E\Delta y_t = \sum_{j=0}^{q-1} \mu_j t^j, \qquad t=2, 3, \cdots, \quad (7)$$

with μ_j following from the relation

$$\sum_{j=0}^{q-1} \mu_j t^j = \sum_{j=0}^{q} \check{\mu}_j \{t^j - (t-1)^j\}.$$

Notice that $E\Delta y_1 = \sum_{j=0}^{q} \check{\mu}_j - Ey_0 = O(1)$.

For later discussion, for the case $s \geq 2$, partition γ constituting $C(1)$ as

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

with γ_1 of $s \times 1$ and γ_2 of $s \times (s-1)$. Then we can suppose $\gamma_2' \check{\mu}_q = 0$, since there exist a column full rank matrix $\bar{\gamma}_2$ of $s \times (s-1)$ and a nonzero s-dimensional vector $\bar{\gamma}_1$ such that $\bar{\gamma}_2' \gamma' \check{\mu}_q = 0$, and note that $\begin{bmatrix} \bar{\gamma}_1' \\ \bar{\gamma}_2' \end{bmatrix} \gamma'$ can be regarded as γ'.

3 The VAR Formulation

In this section we shall provide some theoretical formulation of finite lag order VAR schemes fitted for the data series considered with properties on the cointegrating rank. We first mention the VAR derivation by GRT and the conditions which make it valid in order to make our VAR formulation be more noticeable. Under the DGP (1), its derivation requires

Condition Ⅰ If $s < k$, $\det \begin{bmatrix} \beta' C^{(1)}(1) \\ \gamma' C(1) \end{bmatrix} \neq 0$.

Condition Ⅱ All the roots of $\det C(z) = 0$ are greater than 1 in absolute values except $z=1$.

Both conditions are related on the invertibility of $\begin{bmatrix} \beta' v_t \\ \gamma' \Delta \bar{y}_t \end{bmatrix}$ or $I(0)$ property of any linear combination of it. Condition I is put to exclude the existence of relations of polynomial cointegration as well as higher-order one. We note that if this is not satisfied, there exists a weak stationary series as either $b_1'\beta'(\sum_{h=1}^{t} \bar{y}_h) + b_2'\gamma'\bar{y}_t$ or $b_1'\beta'(\sum_{h=1}^{t} \bar{y}_h)$ with nonzero vectors b_1 of $r \times 1$ and b_2 of $s \times 1$, provided that ξ_0 is suitably dealt with as stated already. It can be easily checked that Condition I is equivalent to Assumption $B3$ in Banerjee et.al (1993, p. 258). Condition Ⅱ is imposed to ensure the invertibility on roots other than $z=1$ in $\det C(z) = 0$ as the VMA characteristic equation of (1). Notice that cointegration under Definition 2 becomes the usual one if $rank\, \beta' C^{(1)}(1) = r$ as well as Condition Ⅱ holds. For the case $s=k$, Condition Ⅱ implies that all the roots are greater than 1 in absolute values.

We now note that neither Condition I nor Ⅱ is necessary for most of the results provided in this paper (except Theorem 1 (iv), (v) and (vi)), as clarified later. If Conditions Ⅰ and Ⅱ are imposed, GRT leads to a VAR representation (as a VECM form) from (1) (see Engle and Granger (1987) or Banerjee et.al (1993, pp. 258-260) etc.): for $t=1, 2, \cdots$,

$$\Delta \bar{y}_t = \alpha \beta' v_{t-1} + \sum_{i=1}^{\infty} H_i \Delta \bar{y}_{t-i} + \varepsilon_t \qquad if\ s < k, \quad (8)$$

$$\Delta \bar{y}_t = \sum_{i=1}^{\infty} H_i \Delta \bar{y}_{t-i} + \varepsilon_t \qquad if\ s = k, \quad (9)$$

with α as a column full rank constant matrix of $k \times r$ such that $\delta' \alpha = 0$, defined only for the case $s < k$, and H_i of $k \times k$ constant matrices.[6] It should be noticed that H_i satisfy such a condition on the Euclidean distance as for C_i and that (8)/(9) is generally characterized by the infinite lag order.

Apart from formulating the 'pure' VAR such as (8)/(9), consider the formulation of 'VAR-like/VECM-like' representations of some finite lag orders under the case in which neither Condition I nor II is imposed. Let

$$P(\bar{w}_t | \bar{z}_{t;i};\ i=1, \cdots, \check{n}), \quad P(\bar{w}_t | \bar{z}_{t;0}, \bar{z}_{t;i};\ i=1, \cdots, \check{n}),$$
$$P(\bar{w}_t | \bar{z}_{t;-1}, \bar{z}_{t;0}; \bar{z}_{t;i};\ i=1, \cdots, \check{n}), \quad P(\bar{w}_t | \bar{z}_{t;-1}, \bar{z}_{t;0})$$
$$or\ P(\bar{w}_t | \bar{z}_{t;-1})$$

stand for the l.l.s. predictor of a vector time series $\bar{w}_t$ onto $\{\bar{z}_{t;i};\ i=\bar{m}, \cdots, \bar{n}\}$ as the (Hilbert) space spanned by vector time series $\bar{z}_{t;i}$, $i=\bar{m}, \cdots, \bar{n}$, with the inclusion of the case in which $\bar{z}_{t;-1}=1$ for all t, $\bar{m}$ as one of -1, 0 or 1 and $\bar{n}$ as one of -1, 0 or $\check{n}$ such that $\bar{n} \geq \bar{m}$ and $\check{n}$ is a positive integer, formulated as

$$(E\bar{w}_t \bar{Z}'_{t;\bar{m};\bar{n}})\,(E\bar{Z}_{t;\bar{m};\bar{n}}\,\bar{Z}'_{t;\bar{m};\bar{n}})^{-1}\,\bar{Z}_{t;\bar{m};\bar{n}},$$

with $\bar{Z}_{t;\bar{m};\bar{n}}$ standing for $(\bar{z}'_{t;\bar{m}},\cdots,\bar{z}'_{t;\bar{n}})'$. Now, let us p be a nonnegative integer, fixed in the sense that it does not depend upon the sample size T, unlike in Saikkonen (1992) or Qu and Perron (2007). For $p=1, 2,\cdots$, and $t=p+2, p+3,\cdots$,
put

$$\begin{aligned}\varepsilon_t(p) &= \Delta\bar{y}_t - P(\Delta\bar{y}_t \mid \beta' v_{t-1}, \Delta\bar{y}_{t-i};\ i=1,\cdots,p) \quad \text{if } s<k,\\ &= \Delta\bar{y}_t - P(\Delta\bar{y}_t \mid \Delta\bar{y}_{t-i};\ i=1,\cdots,p) \quad \text{if } s=k,\end{aligned}$$

Following the definition of the l.l.s. predictors, for p and t given above we have:

$$\Delta\bar{y}_t = \alpha(p)\beta' v_{t-1} + \sum_{i=1}^{p} H_i(p)\Delta\bar{y}_{t-i} + \varepsilon_t(p) \qquad \text{if } s<k, \quad (10)$$

$$\Delta\bar{y}_t = \sum_{i=1}^{p} H_i(p)\Delta\bar{y}_{t-i} + \varepsilon_t(p) \qquad \text{if } s=k, \quad (11)$$

with $\alpha(p)$ as a constant matrix of $k\times r$, defined only for the case $s<k$, and $H_i(p)$ of $k\times k$ constant matrices. Similarly, for $t=2, 3,\cdots$,

$$\Delta\bar{y}_t = \alpha(p)\beta' v_{t-1} + \varepsilon_t(0) \qquad \text{if } s<k, \quad (12)$$

letting $\varepsilon_t(0) = \Delta\bar{y}_t - P(\Delta\bar{y}_t \mid \beta' v_{t-1})$, and put

$$\Delta\bar{y}_t = \varepsilon_t(0) \qquad \text{if } s=k, \quad (13)$$

Replacing $\sum_{i=1}^{p} H_i(p)$ with $\sum_{i=1}^{\max\{p,1\}} H_i(p)$ in (10)/(11) and defining $H_1(0)=0$, (12)/(13) can be incorporated into (10)/(11) as the case $p=0$.

For the purpose of statistical inferences, representations using Δy_{t-i} and $\beta' y_{t-1}$ may be more preferable than those in $\Delta\bar{y}_{t-i}$ and $\beta' v_{t-1}$. The following lemma states how such a representation is obtained in connection with (10)/(11) above.

Lemma 1 *Suppose that y_t (or $\bar{y}_t$) is generated by* (1) *accompanied with* (6). *Then, for $t=p+2, p+3,\cdots$, we have*

$$\begin{aligned}&\Delta y_t - P(\Delta y_t \mid 1, \beta' y_{t-1} - \beta'\xi_0, \Delta y_{t-i};\ i=1,\cdots,p) = \varepsilon_t(p),\\ &\quad \Delta y_t - P(\Delta y_t \mid 1, \beta' y_{t-1} - \beta'\xi_0) = \varepsilon_t(0) \qquad \text{if } s<k, \quad (14)\end{aligned}$$

$$\begin{aligned}&\Delta y_t - P(\Delta y_t \mid 1, \Delta y_{t-i};\ i=1,\cdots,p) = \varepsilon_t(p),\\ &\quad \Delta y_t - P(\Delta y_t \mid 1) = \varepsilon_t(0) \qquad \text{if } s=k, \quad (15)\end{aligned}$$

$$\Delta y_t = \alpha(p)\beta' y_{t-1} + \sum_{i=1}^{\max\{p,1\}} H_i(p)\Delta y_{t-i} + \sum_{j=0}^{q} \bar{\mu}_j t^j + \varepsilon_t(p) \qquad \text{if } s<k, \quad (16)$$

$$\Delta y_t = \sum_{i=1}^{\max\{p,1\}} H_i(p)\Delta y_{t-i} + \sum_{j=0}^{q-1} \bar{\mu}_j t^j + \varepsilon_t(p) \qquad \text{if } s=k, \quad (17)$$

with the notations introduced on (1) *to* (7) *and* (10) *to* (13) *and k-dimensional vectors $\bar{\mu}_j$ satisfying*

$$\sum_{j=0}^{q} \bar{\mu}_j t^j = \alpha(p)\beta' \sum_{j=0}^{q} \check{\mu}_j (t-1)^j - \sum_{i=1}^{\max\{p,1\}} H_i(p) \sum_{j=0}^{q-1} \mu_j (t-i)^j - \alpha(p)\beta'\xi_0 + \sum_{j=0}^{q-1} \mu_j t^j$$

if $s<k$, and

$$\sum_{j=0}^{q-1} \bar{\mu}_j t^j = -\sum_{i=1}^{\max\{p,1\}} H_i(p) \sum_{j=0}^{q-1} \mu_j (t-i)^j + \sum_{j=0}^{q-1} \mu_j t^j$$

if $s=k$.

Now, turn our interest to the characterization of $\alpha(p)$ and $\varepsilon_t(p)$, particularly of the rank value of $\alpha(p)$ and the invertibility on $\varepsilon_t(p)$, and those are summarized in:

Theorem 1 *Suppose that y_t is generated by* (1). *Then, with the notations on* (1) *and* (10)/(11), *we have the following results.*

(i) *For the case $s<k$, rank $\alpha(p)=r$.*

(ii) *$\varepsilon_t(p)$ in* (10)/(11) *possesses the following representation*

$$\varepsilon_t(p) = B(L;p)\,\varepsilon_t, \quad t=p+2, p+3,\cdots, \qquad (18)$$

where the power series $B(z;p)$ is given as

$$B(z;p) = I_k + \sum_{i=1}^{\infty} B_i(p) z^i,$$

with constant matrice $B_i(p)$ of $k\times k$ such that $\sum_{i=1}^{\infty} \|B_i(p)\| < \infty$ for some $\nu \geq 1$, and for the case $s=k$, rank $B(1;p)=k$.

(iii) *For any column full rank constant matrix $\delta(p)$ of $k\times s$ such that $\delta'(p)\alpha(p)=0$ is satisfied if and only if $s<k$, there exists a full rank constant matrix $\tilde{\tau}(p)$ of $s\times s$ such that*

$$\delta'(p)\,B(1;p) = \tilde{\tau}(p)\,\delta',$$

with $B(1;p)$ in (*ii*) *and δ on $C(1)$ in Section 2.*

(iv) *For the case $s<k$, suppose that Condition I holds. Then, for any column full rank constant matrix ψ of $k\times r$ such that $\delta'\psi=0$, there exists a full rank constant matrix $\bar{\tau}(p)$ of $r\times r$ such that*

$$\psi=B^{-1}(1;p)\ \alpha(p)\,\bar{\tau}(p),$$

with rank $B(1;p)=k$ for $B(1;p)$ in (ii).

(v) *Suppose that Condition Ⅱ as well as Ⅰ holds for the case $s<k$ and only Ⅱ holds for the case $s=k$. Then, for $B(z;p)$ in (ii), all the roots of* $\det B(z;p)=0$ *are greater than* 1 *in absolute values.*

(vi) *For the case $s<k$, suppose that Conditions Ⅰ and Ⅱ hold. Then, for α in* (8) *and $B(1;p)$ in (ii),*

$$\alpha=B^{-1}(1;p)\ \alpha(p).$$

Theorem 1 (i) implies that (10) or (16) is regarded as an acceptable model to formulate cointegration in the sense that the cointegrating rank is precisely captured by the parameter of the model, and as presented in Section 5, it also plays an important role for estimation of the rank and the asymptotic evaluation. That $E\varepsilon_t(p)=0$, $E\varepsilon_t(p)\ v'_{t-1}\beta=0$ and $E\varepsilon_t(p)\Delta\bar{y}'_{t-i}=0$, $i=1,\cdots,p$, may be another favorable factor, although $\varepsilon_t(p)$ are not ensured to be serially uncorrelated unlike ε_t. Based on these matters, (10)/(11) or (16)/(17) is regarded as a theoretical representation for the VAR scheme of p-th order. Notice that neither Conditions Ⅰ nor Ⅱ is needed to establish *rank* $\alpha(p)=r$. Similarly, without these conditions, it is ensured by (iii) that any linear combination of $\delta'(p)\ \varepsilon_t(p)$ is of $I(0)$. On the other hand, (iv) states that Condition I rules out overdiffencing in the MA representation of $\varepsilon_t(p)$: any linear combination of $\varepsilon_t(p)$ is of $I(0)$. Moreover, it should be noted from (v) that the invertibility of (any linear combination of) $\varepsilon_t(p)$ itself is ensured by the combination of Conditions Ⅰ and Ⅱ.

Before completing this section, we state some relations or properties on $\alpha(p)$, some l.l.s. predictors and their innovations which contribute to the derivation of some of the asymptotics and are similar to ones in Johansen's (1988) Lemma 2:

Corollary 1 *Suppose that y_t is generated by* (1) *with $s<k$, put*

$$\begin{aligned}
&u_t(p)=\Delta\bar{y}_t-P(\Delta\bar{y}_t\,|\,\Delta\bar{y}_{t-i};\ i=1,\cdots,p),\\
&\zeta_{t-1}(p)=v_{t-1}-P(v_{t-1}|\,\Delta\bar{y}_{t-i};\ i=1,\cdots,p),\\
&\Omega(p)=E\,\varepsilon_t(p)\,\varepsilon'_t(p),\quad \textstyle\sum_{00}(p)=Eu_t(p)\ u'_t(p),\\
&\textstyle\sum_{01}(p)=Eu_t(p)\ \zeta'_{t-1}(p),\quad \sum_{10}(p)=\sum'_{01}(p),\\
&\textstyle\sum_{11}(p)=E\ \zeta_{t-1}(p)\ \zeta'_{t-1}(p),
\end{aligned}$$

and let $\lambda_1(p)\geq\cdots\geq\lambda_r(p)$ be the ordered eigenvalues of

$$(\beta\textstyle\sum_{11}(p)\beta)^{1/2}\ \alpha'(p)\sum_{00}^{-1}(p)\,\alpha(p)\,(\beta\sum_{11}(p)\beta)^{1/2},$$

with β and $\alpha(p)$ on (1) *and* (10). *Then we have:*

$$\begin{aligned}
&\textstyle\sum_{00}(p)=\sum_{j=0}^{\infty}\bar{K}_j(p;0)\,\Lambda\,\bar{K}'_j(p;0),\\
&\textstyle\sum_{i1}(p)=\sum_{j=1}^{\infty}\bar{K}_j(p;i)\,\Lambda\,\bar{K}'_j(p;1)\qquad i=0,1,
\end{aligned}\tag{19}$$

with constant matrices $\bar{K}_j(p;i)$ of $k\times k$ in

$$u_t(p)=\sum_{j=0}^{\infty}\bar{K}_j(p;0)\ \varepsilon_{t-j},\quad \zeta_{t-1}(p)=\sum_{j=1}^{\infty}\bar{K}_j(p;1)\ \varepsilon_{t-j},$$

$$\Omega(p)=\textstyle\sum_{00}(p)-\sum_{01}(p)\,\beta\,(\beta'\sum_{11}(p)\,\beta)^{-1}\beta'\sum_{10}(p),\tag{20}$$

$$\alpha(p)=\textstyle\sum_{01}(p)\,\beta\,(\beta'\sum_{11}(p)\,\beta)^{-1},\tag{21}$$

$$1>\lambda_1(p),\qquad \lambda_r(p)>0.\tag{22}$$

In connection with (20), notice that $\Omega(p)=\Lambda+\sum_{i=1}^{\infty}B_i(p)\,\Lambda\,B'_i(p)$, followed directly from (18) in Theorem 1.

4 Information Criteria

Given T observations $y_1,\cdots,y_T$ in the DGP (1) accompanied with (6), we shall discuss a statistical procedure to estimate the cointegrating rank r. It is constructed under each of the VAR schemes fitted, expressed as (10)/(11) or (16)/(17) in the previous section. For each p, we define the matrices/vectors Y_{-1}, ΔY_{-i} of $\check{T}\times k$, with $\check{T}=T-p-1$ and $-0=0$, $\check{\tau}_j$ of $\check{T}\times 1$, $\hat{\tau}(\bar{q})$ of $\check{T}\times\bar{q}$, with $\bar{q}=q,\ q+1$, $Z_{-1}(p)$ of $\check{T}\times(kp+q)$, $\check{Z}_{-1}(p)$ of $\check{T}\times(kp+q+1)$, $\Delta Z_{-1}(p)$ of $\check{T}\times kp$, $M_Z(p)$, $M_{\check{Z}}(p)$ and $M_{\Delta Z}(p)$ of $\check{T}\times\check{T}$ as

$$\begin{aligned}
&Y'_{-1}=[y_{p+1},\,y_{p+2},\cdots,\,y_{T-1}],\\
&\Delta Y'_{-i}=[\Delta y_{p+2-i},\,\Delta y_{p+3-i},\cdots,\Delta y_{T-i}]\qquad i=0,1,\cdots,p,\\
&\check{\tau}'_j=\big((p+2)^j,\,(p+3)^j,\cdots,\,T^j\big)\qquad j=0,1,\cdots,q,
\end{aligned}$$

$$\check{\tau}(\overline{q})=[\check{\tau}_0,\ \check{\tau}_1,\cdots,\ \check{\tau}_{q-1}],$$
$$Z_{-1}(0)=\hat{\tau}(q),\qquad Z_{-1}(p)=[\Delta Y_{-1},\cdots,\Delta Y_{-p},\hat{\tau}(q)]\quad \text{if } p\geq 1,$$
$$\check{Z}_{-1}(0)=\hat{\tau}(q+1),\ \check{Z}_{-1}(p)=[\Delta Y_{-1},\cdots,\Delta Y_{-p},\hat{\tau}(q+1)]\quad \text{if } p\geq 1,$$
$$\Delta Z_{-1}(p)=[\Delta Y_{-1},\cdots,\Delta Y_{-p}]\qquad \text{if } p\geq 1,$$
$$M_Z(p)=I_{\check{T}}-Z_{-1}(p)\left(Z'_{-1}(p)\,Z_{-1}(p)\right)^{-1}Z'_{-1}(p),$$
$$M_{\check{Z}}(p)=I_{\check{T}}-\check{Z}_{-1}(p)\left(\check{Z}'_{-1}(p)\,\check{Z}_{-1}(p)\right)^{-1}\check{Z}'_{-1}(p),$$
$$M_{\Delta Z}(p)=I_{\check{T}}\qquad \text{if } p=0,$$
$$=I_{\check{T}}-\Delta Z_{-1}(p)\left(\Delta Z'_{-1}(p)\Delta Z_{-1}(p)\right)^{-1}\Delta Z'_{-1}(p)\quad \text{if } p\geq 1.$$

We also let $\tilde{M}(p)$ denote one of $M_Z(p)$, $M_{\check{Z}}(p)$ or $M_{\Delta Z}(p)$, provided that it is not permissible for $\tilde{M}(p)$ to be $M_Z(p)$ unless $\check{\mu}_q=C(1)\check{\mu}\neq 0$ holds with a k-dimensinal constant vector $\check{\mu}$ and that the choice of $M_{\Delta Z}(p)$ is allowed if and only if $\check{\mu}_0=\check{\mu}_1=\cdots\check{\mu}_q=0$ holds. Then, following the notations S_{ij} used in Johansen (1988, 1992b, 1996), let us define $S_{ij}(p)$ as

$$S_{00}(p)=\Delta Y'_0\tilde{M}(p)\Delta Y_0/T,$$
$$S_{01}(p)=\Delta Y'_0\tilde{M}(p)Y_{-1}/\mathrm{T},\qquad S_{10}(p)=S'_{01}(p),$$
$$S_{11}(p)=Y'_{-1}\tilde{M}(p)Y_{-1}/T.$$

Moreover, let $\hat{\lambda}_1(p)\geq\cdots\geq\hat{\lambda}_k(p)$ and $\hat{\psi}_1(p),\cdots,\ \hat{\psi}_k(p)$ be the ordered eigenvalues of the equation

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

and the corresponding eigenvectors, and with $\hat{\rho}_1(p)\leq\cdots\leq\hat{\rho}_k(p)$ as the ordered eigenvalues of $S_{11}(p)$, $diag\{\hat{\rho}_1^{-1/2}(p),\cdots\hat{\rho}_k^{-1/2}(p)\}$ denoting the $k\times k$ diagonal matrix and $\hat{\xi}_1(p),\cdots,\ \hat{\xi}_k(p)$ as the corresponding eigenvectors. Then, as seen easily, $\hat{\lambda}_1(p),\cdots,\hat{\lambda}_k(p)$ are calculated actually as the (ordered) eigenvalues of

$$S_{11}^{-1/2}(p)S_{10}(p)S_{00}^{-1}(p)S_{01}(p)S_{11}^{-1/2}(p).$$

It should be noted that the above matrix and its eigenvalues do not depend upon the scale on which $y_1,\cdots,y_T$ are measured.

The information criteria adopted in this paper and related expressions are described in a unified form:

$$I(j;p)=T\log\det\hat{\Omega}(j;p)+\left(2jk+k^2p+\tfrac{k^2}{2}+\tfrac{k}{2}\right)C_T,\tag{23}$$

where

$$\hat{\Omega}(0;p)=S_{00}(p),$$
$$\hat{\Omega}(j;p)=S_{00}(p)-S_{01}(p)\hat{\beta}(j;p)\hat{\beta}'(j;p)S_{10}(p),\quad j=1,\cdots,k-1,$$

with

$$\hat{\beta}(j;p)=S_{11}^{-1/2}(p)[\hat{\psi}_1(p),\cdots,\ \hat{\psi}_k(p)],$$

and $\{C_T\}$ is a sequence such that $\lim_{T\to\infty}C_T>0$ and $\lim_{T\to\infty}\frac{C_T}{T}=0$. Notice that the first term $T\log\det\hat{\Omega}(j;p)$ of the right-hand side of (23) corresponds to a quantity on the residual moment matrix from reduced rank regression or the concentrated log-likelihood, regarding (10)/(11) as a VAR/VECM of lag order p and cointegrating rank j and that $2jk+k^2p+\frac{k^2}{2}+\frac{k}{2}$ in the second term corresponds to the number of parameters $\alpha(p)$, β, $H_i(p)$, $i=1,\cdots,p$, $\Omega(p)$. Each of the information criteria yields an estimator of r through minimization of $I(j;p)$ with respect to j for each fixed p and C_T, and any of such estimators is denoted as $\hat{r}(p)$ in a unified form, noting that $\hat{r}(p)$ is realized as an integer producing the minimum of $I(j;p)$ over the set $J=\{0,1,\cdots,k-1\}$:

$$I(\hat{r}(p);p)=\min_{j\in J}I(j;p).\tag{24}$$

Noting that

$$\log\det\hat{\Omega}(j;p)$$
$$=\log\{I_j-\det\hat{\beta}'(j;p)S_{10}(p)S_{00}^{-1}(p)S_{01}(p)\hat{\beta}(j;p)\}+\log\det S_{00}(p)$$
$$=\sum_{i=1}^{j}\log\{1-\hat{\lambda}_i(p)\}+\log\det S_{00}(p)$$

and adding the quantity not dependent on j

$$-T\sum_{i=1}^{k}\log\{1-\hat{\lambda}_i(p)\}-T\log\det S_{00}(p)-\left(k^2p+\tfrac{k^2}{2}+\tfrac{k}{2}\right)C_T$$

to $I(j;p)$, we also derive a simpler form:

$$\overline{I}(j;p)=-T\sum_{i=j+1}^{k}\log\{1--\hat{\lambda}_i(p)\}+2jkC_T.\tag{25}$$

Since obviously minimization of $\overline{I}(j;p)$ with respect to j provides the identical conclusion as that of $I(j;p)$, we have another definition of $\hat{r}(p)$:

$$\overline{I}(\hat{r}(p);p)=\min_{j\in J}\overline{I}(j;p).\tag{26}$$

In (23) or (25), each information criterion is characterized by C_T. It should be noted that $C_T=2$ for AIC, $C_T=\log T$ for SIC and $C_T=2\log\log T$ for HQ.

5 Asymptotics

In order to establish some asymptotic desirability of $\hat{r}(p)$ in the previous section, ε_t in (1) are assumed to be iid with finite fourth moments hereafter in addition to the supposition put already. We also provide notations on the Brownian motion: let the symbols $\Rightarrow$ and $W_m(u)$ stand for weak convergence of probability measures on the unit interval [0, 1] and a m-dimensional standard Brownian motion of on [0, 1] respectively, noting that $W_m(u)$ is distributed pointwisely for each u as m-variate Gaussian with mean zero and covariance matrix $u\, I_m$ (for the detailed definition, see Johansen (1996, p. 241) or Davidson (1994, pp. 418, 442-443) e.g.), and with q given on (6), let us regard $\bar{W}_m(u)$ and $\psi_u(\bar{q})$ on [0, 1] as

$$\bar{W}_m(u) = u^q \qquad \text{if } m=1,$$
$$= \left(u^q,\, W'_{m-1}(u)\right)' \qquad \text{if } m>1$$

and $\psi_u(\bar{q}) = (1, u, \cdots, u^{\bar{q}-1})'$, $\bar{q} = q, q+1$.

Lemma 2: *Suppose that y_t is generated by* (1) *accompanied with* (6), *the notations on* (1) *and* (10)/(11) *and the assumption stated above. Then we have the following asymptotics on $S_{ij}(p)$ in Section 4:*

$$S_{00}(p) = \Sigma_{00}(p) + O_p(T^{-1/2}), \tag{27}$$

$$\beta' S_{11}(p)\beta = \beta' \Sigma_{11}(p)\beta + O_p(T^{-1/2}), \qquad \text{if } s<k, \tag{28}$$

$$S_{01}(p)\beta = \alpha(p)\beta' \Sigma_{11}(p)\beta + O_p(T^{-1/2}) \quad \text{if } s<k, \tag{29}$$

$$\delta'(p) S_{01}(p)\gamma D_T^{-1}) \Rightarrow \tilde{\tau}(p)\check{G}\left(\int_0^1 dW_s(u)\tilde{W}'_s(u)\right)\tilde{G} + \tilde{\tau}(p)\,\delta' Q(p)\bar{\gamma} \quad \text{as } T\to\infty \tag{30}$$

$$S_{01}(p)\gamma D_T^{-1}) \Rightarrow \bar{K}(1;p;0)\,\check{F}\bar{P}\left(\int_0^1 \binom{dW_s(u)}{dW_r(u)}\tilde{W}'_s(u)\right)\tilde{G} + \bar{Q}(p;0)\bar{\gamma} \quad \text{as } T\to\infty \tag{31}$$

$$\beta' S_{11}(p)\gamma D_T^{-1}) \Rightarrow \beta'\bar{K}(1;p;1)\,\check{F}\bar{P}\left(\int_0^1 \binom{dW_s(u)}{dW_r(u)}\tilde{W}'_s(u)\right)\tilde{G} + \beta'\bar{Q}(p;1)\bar{\gamma} \quad \text{as } T\to\infty \quad \text{if } s<k, \tag{32}$$

$$D_T^{-1}(\gamma' S_{11}(p)\gamma/T)D_T^{-1}) \Rightarrow \tilde{G}\left(\int_0^1 \tilde{W}_s(u)\tilde{W}'_s(u)du\right)\tilde{G} \quad \text{as } T\to\infty \tag{33}$$

where $\delta(p)$ is as in Theorem 1 (*iii*) *with the corresponding $\tilde{\tau}(p)$, $\Sigma_{ij}(p)$ are as in Corollary 1, and the matrices D_T^{-1} of $s\times s$, $\bar{\gamma}$ of $k\times s$, $Q(p)$, $\check{G}$, $\tilde{G}$, $\bar{K}$(1, p, i), $\check{F}$, $\bar{P}$ and $\bar{Q}(p;i)$, of $k\times k$, are defined as*

$$D_T^{-1} = T^{-q+1/2} \quad \text{if } \tilde{M}(p) = M_Z(p),\ s=1 \text{ and } \gamma'_1\check{\mu}_q \neq 0,$$
$$= \begin{bmatrix} T^{-q+1/2} & 0 \\ 0 & I_{s-1} \end{bmatrix} \quad \text{if } \tilde{M}(p) = M_Z(p),\ s>1 \text{ and } \gamma'_1\check{\mu}_q \neq 0,$$
$$= I_s \quad \text{otherwise,}$$

$$Q(p) = \left(\sum_{j=1}^{\infty} B_j(p)\right)\Lambda C'(1) + \sum_{j=1}^{\infty} B_j(p)\Lambda\left(-\sum_{h=j}^{\infty} C'_h\right),$$

$$\bar{\gamma} = [0,\ \gamma_2] \quad \text{if } \tilde{M}(p) = M_Z(p) \text{ and } \gamma'_1\check{\mu}_q \neq 0,$$
$$= \gamma \quad \text{otherwise,}$$

$$\check{G} = (\delta'\Lambda\delta)^{1/2},$$

$$\tilde{G} = \gamma'_1\check{\mu}_q \quad \text{if } \tilde{M}(p) = M_Z(p),\ s=1 \text{ and } \gamma'_1\check{\mu}_q \neq 0,$$
$$= \begin{bmatrix} \gamma'_1\check{\mu}_q & 0 \\ 0 & \gamma'_2\gamma\tau\check{G} \end{bmatrix} \quad \text{if } \tilde{M}(p) = M_Z(p),\ s>1 \text{ and } \gamma'_1\check{\mu}_q \neq 0,$$
$$= \gamma'\gamma\,\tau\check{G} \text{ otherwise,}$$

$$\bar{K}(1;p;0) = \sum_{j=0}^{\infty}\bar{K}_j(p;0),\ \bar{K}(1;p;1) = \sum_{j=1}^{\infty}\bar{K}_j(p;1),$$

$$\check{F} = [\delta(\delta'\delta)^{-1},\ \psi(\psi'\psi)^{-1}] \quad \text{if } s<k,$$
$$= \delta(\delta'\delta)^{-1} \quad \text{if } s=k,$$

$$\bar{P} = \begin{bmatrix} G & 0 \\ 0 & (\psi'\Lambda\psi)^{1/2} \end{bmatrix} \quad \text{if } s<k,$$
$$= \check{G} \quad \text{if } s=k,$$

$$\bar{Q}(p;0) = \left(\sum_{j=1}^{\infty}\bar{K}_j(p;0)\right)\Lambda C'(1) + \sum_{j=1}^{\infty}\bar{K}_j(p;0)\Lambda\left(-\sum_{h=j}^{\infty}C'_h\right),$$

$$\bar{Q}(p;1) = \left(\sum_{j=1}^{\infty}\bar{K}_j(p;1)\right)\Lambda C'(1) + \sum_{j=1}^{\infty}\bar{K}_j(p;1)\Lambda\left(-\sum_{h=j}^{\infty}C'_h\right),$$

where ψ is a column full rank constant matrix of $k\times r$ such that $\psi'\delta = 0$ and $\psi'\Lambda\delta = 0$, defined for the case $s<k$, $W_s(u)$ is formulated above, $W_r(u)$ is a standard Brownian motion of r-dimension independent of $W_s(u)$, and $\tilde{W}_s(u)$ is defined as

$$\tilde{W}_s(u) = W_s(u) \quad \text{if } \tilde{M}(p) = M_{\Delta Z}(p),$$

$$=\bar{W}_s(u)-\Big(\int_0^1 \bar{W}_s(u)\,\psi_u'(q)du\Big)\cdot\Big(\int_0^1 \psi_u(q)\,\psi_u'(q)du\Big)^{-1}\psi_u(q)$$
if $\tilde{M}(p)=M_Z(p)$ *and* $\gamma_1'\check{\mu}_q\neq 0$,

$$=W_s(u)-\Big(\int_0^1 W_s(u)\,\psi_u'(q)du\Big)\cdot\Big(\int_0^1 \psi_u(q)\,\psi_u'(q)du\Big)^{-1}\psi_u(q)$$
if $\tilde{M}(p)=M_Z(p)$ *and* $\gamma_1'\check{\mu}_q=0$,

$$=W_s(u)-\Big(\int_0^1 W_s(u)\,\psi_u'(q+1)du\Big)\cdot\Big(\int_0^1 \psi_u(q+1)\,\psi_u'(q+1)du\Big)^{-1}\psi_u(q+1)$$
if $\tilde{M}(p)=M_{\dot{Z}}(p)$,

with $\bar{W}_s(u)$ *and* $\psi_u(\bar{q})$, $\bar{q}=q, q+1$, *formulated above.*

Note in (30) to (32) above that if $\tilde{M}(p)=M_Z(p)$, $s>1$ and $\gamma_1'\check{\mu}_q\neq 0$, the first column vectors of $\tilde{\tau}(p)\,\delta'\bar{Q}(p)\bar{\gamma}$, $\bar{Q}(p;0)\bar{\gamma}$ and $\beta'\bar{Q}(p;1)\bar{\gamma}$ are zero. We also notice in Lemma 2 that if $\varepsilon_t(p)=\varepsilon_t$(i.e., $H_i=0$ for $\forall_i\geq p+1$ in (8)), $Q(p)=0$ holds since $B(z;p)=I_k$ (i.e., $B_i(p)=0$ for $\forall_i\geq 1$), and it is seen that $\tilde{\tau}(p)=I_s$ for $\delta(p)$ such that $\delta'(p)\,\delta(p)=\delta'\delta$. Then the limiting distribution of the trace of $T\delta'(p)S_{01}(p)\gamma(\gamma'S_{11}(p)\gamma)^{-1}\gamma'S_{10}(p)\,\delta(p)$ is equal to one for Johansen's rank test (under the null), diversified by $\tilde{M}(p)$ or $\check{\mu}_j$.

Lemma 3: *Suppose that* y_t *is generated by* (1) *with the same supposition as in Lemma* 2. *Then, for* $\hat{\lambda}_j(p)$ *given in Section 4, we have:*

(i) *For the case* $s<k$ *and* $j=1,\cdots,r$,

$$-\log\{1-\hat{\lambda}_j(p)\}=O_p(1),$$
$$\big(-\log\{1-\hat{\lambda}_j(p)\}\big)^{-1}=O_p(1).$$

(ii) *For* $j=r+1,\cdots,k$,

$$-T\sum_{h=r+1}^{j}\log\{1-\hat{\lambda}_h(p)\}=O_p(1),$$
$$\Big(-T\sum_{h=r+1}^{j}\log\{1-\hat{\lambda}_h(p)\}\Big)^{-1}=O_p(1).$$

(iii) *Putting*

$$-\log\{1-\hat{\lambda}_j(p)\}=f_j(\Lambda)\qquad j=1,\cdots,r,$$
$$-T\log\{1-\hat{\lambda}_{r+h}(p)\}=f_{r+h}\big((\gamma_1'\check{\mu}_q)^2,\ \Lambda\big)\quad \textit{if } \tilde{M}(p)=M_Z(p) \textit{ and } \gamma_1'\check{\mu}_q\neq 0,$$
$$=f_{r+h}(\Lambda)\qquad \textit{otherwise},\quad h=1,\cdots,s,$$

as some functions whose inputs are either elements of Λ *or* $(\gamma_1'\check{\mu}_q)^2$ *as well as those, the functions are asymptotically scale invariant to all the inputs in the sense that for any nonzero real number* $\bar{c}$, *the asymptotics of* $f_j(\bar{c}\Lambda)$ or $f_j(\bar{c}(\gamma_1'\check{\mu}_q)^2, \bar{c}\Lambda)$ *formulated by convergence in probability or weak convergence of probability measure are equal to those of* $f_j(\Lambda)$ *or* $f_j\big((\gamma_1'\check{\mu}_q)^2,\ \Lambda\big)$ *respectively*, $1\geq j\geq k$.

Notice on Lemma 3 (i) that for sufficient large T,

$$\infty>-log\{1-\hat{\lambda}_1(p)\}\geq\cdots\geq-log\{1-\hat{\lambda}_r(p)\}>0,$$

which follows from (22) of Corollary 1, (A.36) and (A.37) in the proof of Lemma 3. We also note that *rank* $\alpha(p)=r$ of Theorem 1 (i) is indispensable for the derivation of $\big(-log\{1-\hat{\lambda}_j(p)\}\big)^{-1}=O_p(1)$, $j=1,\cdots,r$, as clarified in the proof of Lemma 3. Similarly, *rank* $\delta'(p)B(1;p)=s$ of Theorem 1 (iii) is needed for the derivation of (ii), although $\big(-T\sum_{h=r+1}^{j}log\{1-\hat{\lambda}_h(p)\}\big)^{-1}=O_p(1)$ is unnecessary for the main results stated below. Moreover, we may expect (iii) to have effects as some boundary to the first term of (25) expressed as $-T\sum_{h=j+1}^{k}log\{1-\hat{\lambda}_h(p)\}$, although neither ensured to be free of all the nuisance parameters nor required directly for the main results below. We now attain to:

Theorem 2: *Suppose* y_t *is generated by* (1) *with the same supposition as in Lemma* 2. *Then, for* $\hat{r}(p)$ *chosen by* (26) *in Section* 4, *we have:*

$$\lim_{T\to\infty}Pr(\hat{r}(p)=r)=1\qquad \textit{if}\lim_{T\to\infty}C_T=\infty,\tag{34}$$

$$\lim_{T\to\infty}Pr(\hat{r}(p)\geq r)=1\qquad \textit{if}\lim_{T\to\infty}C_T<\infty,\tag{35}$$

with the notation $P_r(\cdot)$ *denoting the probability.*

Following the above theorem, $\hat{r}(p)$ chosen by an information criterion satisfying $\lim_{T\to\infty}C_T=\infty$ such as SIC or HQ converges to r with probability one, and for $\hat{r}(p)$ under a criterion characterized by $\lim_{T\to\infty}C_T<\infty$ such as AIC, the probability of underestimating r tends to zero as T increases, although overestimation of r possibly occurs with nonzero probability. These properties of consistency and overestimability are completely conformed to the conventionality including Aznor and Salvador (2002) on the determination of cointegrating rank.

6 Monte Carlo Experiments

In this section we execute Monte Carlo experiments on the cointegrating rank estimation based on the methods such as AIC, SIC and HQ under each of several finite lag order VAR schemes and the DGPs as special cases of (1). The main purpose of the experiments is to observe to what extent the asymptotics established theoretically in the previous section are preserved for finite samples. The DGPs in Examples 1-3 below are of 4-variates systems ($k=4$ in (1)) with ε_t as Gaussian with mean zero and covariance matrix I_4 (i.e., $\Lambda=I_4$), and it is assumed that $y_{-j}=\varepsilon_{-j}=0$, $j\geq 0$, $\breve{\mu}_i=0$, $i\geq 2$, and $\breve{\mu}_0=0$, with the supposition that either $\breve{\mu}_1=0$ or $\breve{\mu}_1=C(1)\ \breve{\mu}\neq 0$ holds, implying that $\beta'\bar{y}_{t-1}=\beta' v_{t-1}$ in (4) and that the only allowable deterministic trend is one for $q=1$ in (6). Each example consists of four DGPs identified by f and g as scalar parameters, provided that whether the DGP possesses a deterministic trend or not is decided by the value of g. It will be also explained that the DGP can be converted to a special cases of (1), although not provided in a direct form as (1). On the other hand, we suppose that p as the VAR lag order takes 8 as the value at its maximum under each estimation method: p possibly takes integers from 0 to 8. For each DGP, p and estimation method, an estimate of the cointegrating rank r as a realized value of $\hat{r}(p)$ is produced. Calculating $S_{ij}(p)$ in Section 4, we adopt $\tilde{M}(p)=M_Z(p)$ for the case $\breve{\mu}_q\neq 0$ with $q=1$ (or $\mu_0\neq 0$), provided that

$$Z_{-1}(0)=\hat{\tau}(1),\ Z_{-1}(p)=[\Delta Y_{-1},\cdots,\Delta Y_{-p},\ \hat{\tau}(1)] \quad \text{if } p\geq 1,$$

and $\tilde{M}(p)=M_{\Delta Z}(p)$ for the case $\mu_0=0$. Throughout all of examples/DGPs, we ran 10, 000 replication of experiments, and pseudo normal random variables were adopted as elements of ε_t for actual calculation of the estimates under 100, 200 and 500 of the sample size T in each experiment. The method of estimating r based on a consecutive application of Johansen's rank tests (see Johansen (1996, p. 71) e.g.), simply denoted as JT or JT*, was adopted as well as the information criteria, For the critical values under the cases $\tilde{M}(p)=M_{\Delta Z}(p)$ and $\tilde{M}(p)=M_{\Delta Z}(p)$, we follow Johansen's (1996) Table 15.1 and 15.3 respectively.[7] All of the estimators including those based on Johansen's tests are denoted as $\hat{r}(p)$.

Example 1: The DGP is: for $t=1, 2,\cdots$,

$$\Delta y_t=C(1)\ \varepsilon_t+\frac{(1-L)}{(1-0.8L)}(I_4-C(1)-0.8I_4L)\ \varepsilon_t+\mu_0, \tag{35}$$

where

$$C(1)=\begin{pmatrix} 1 & -2 & 0 & 0\\ 0.3 & 1 & 0 & 0\\ -1 & 2 & f & 0.4f\\ -0.15 & -0.5 & 0.5f & 0.2f\end{pmatrix},$$

$$\mu_0=C(1)\begin{pmatrix}0.8g\\0\\0\\0\end{pmatrix}=\begin{pmatrix}0.8g\\0.24g\\-0.8g\\-0.12g\end{pmatrix},$$

with $f=0.8$, 1.6 and $g=0$, 1. It is obvious that (35) is converted to a special case of (1). For any f, the VMA characteristic equation of (35) possesses a pair of complex-conjugate roots less than 1 in absolute values, indicating that Condition Ⅱ is not satisfied and noting that other roots either greater than 1 in absolute values or equal to 1. It is obvious that $C(1)$ is of rank 3, in other words, the cointegrating rank is 1. We can also see that Condition I holds, constructing β and γ based on the eigenvectors of $C(1)C(1)'$ and noting that $C^{(1)}(1)$ is given as

$$\left\{\begin{pmatrix} 0 & 2 & 0 & 0\\ -0.3 & 0 & 0 & 0\\ 1 & -2 & 1-f & -0.4f\\ 0.15 & 0.5 & -0.5f & 1-0.2f\end{pmatrix}-0.8I_4\right\}/0.2$$

in view of (35).

Example 2: The DGP is: for $t=1, 2,\cdots$,

$$\Delta y_t=\{f\gamma\delta'\varepsilon_t+\frac{(1-L)}{(1-0.7L)}\{I_4-f\gamma\delta' -\left(I_4-0.3\beta(\beta'\beta)^{-1}\gamma'\gamma\delta'\right)L\}\ \varepsilon_t+\mu_0, \tag{36}$$

where

$$\gamma' = \begin{pmatrix} 1 & -0.1 & -1 & 0.05 \\ -2 & 1 & 2 & -0.5 \end{pmatrix},$$

$$\delta' = \begin{pmatrix} 1 & 0 & 0 & -0.4 \\ 0 & 0.3 & 0.3 & -0.7 \end{pmatrix},$$

$$\beta' = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 2 \end{pmatrix},$$

$$\mu_0 = f\gamma\delta' \begin{pmatrix} 0.8g \\ 0 \\ 0.6g \\ 1.4g \end{pmatrix} = \begin{pmatrix} 1.84fg \\ -0.824fg \\ -1.84fg \\ 0.412fg \end{pmatrix},$$

with $f=1.6$, 2.4 and $g=0, 1$. We easily see that (36) is a special case of (1), similar to the case of (35). Since $C(1)=f\gamma\delta'$, it is obvious that the cointegrating rank is of 2 with β as the cointegrating matrix. The VMA characteristic equation of (36) does not satisfy Condition Ⅱ owing to the existence of a pair of complexconjugate roots. On the other hand, noting that

$$C^{(1)}(z)=\{I_4-f\gamma\delta'-\big(I_4-0.3\beta(\beta'\beta)^{-1}\gamma'\gamma\delta'\big)z\}/(1-0.7z),$$

$$C^{(2)}(z)=\{I_4-\beta(\beta'\beta)^{-1}\gamma'\gamma\delta'+(0.7/0.3)f\gamma\delta'\}/(1-0.7z)$$

in (2) and $C^{(1)}(z)=C^{(1)}(1)+(1-z)C^{(2)}(z)$, it follows that $(1/f)\gamma'C(1)-\beta'C^{(1)}(1)=0$ and $rank\ \{(1/f)\gamma'C^{(1)}(1)-\beta'C^{(2)}(1)\}=2$. These results indicates the occurrence of polynomial cointegration in the sense that any linear combination of $(1/f)\gamma'\bar{y}_t-\beta'(\sum_{h=1}^{t}\bar{y}_h)$ is of $I(0)$, i.e., the situation in which Condition I is unsatisfied.

Example 3: The DGP is: for $t=1, 2, \cdots$,

$$\Delta y_t=f\bar{\alpha}\beta' y_{t-1}+\sum_{i=1}^{3}H_i\Delta y_{t-i}+\breve{\mu}+\varepsilon_t, \qquad (37)$$

where

$$\bar{\alpha} = (-0.2,\ -0.2,\ -0.5,\ -0.2)',$$

$$\beta' = (1,\ 1,\ 1,\ 1),$$

$$H_1 = \begin{pmatrix} 0 & 0.2 & 0.2 & 0.2 \\ 0 & 0 & 0.2 & 0.2 \\ 0 & 0.5 & 0 & 0.5 \\ 0.2 & 0.2 & 0.2 & 0 \end{pmatrix},$$

$$H_2 = \begin{pmatrix} 0 & 0 & 0.2 & 0.2 \\ 0 & 0 & 0 & 0.2 \\ 0 & 0 & 0 & 0 \\ 0.2 & 0.2 & 0 & 0 \end{pmatrix},$$

$$H_3 = \begin{pmatrix} 0 & 0 & 0 & 0.2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0.2 & 0 & 0 & 0 \end{pmatrix},$$

$$\breve{\mu} = (0.8g,\ 0,\ 0.6g,\ 1.4g)',$$

with $f=0.8$, 1.6 and $g=0, 1$. (37) is a finite lag order VAR. Notice that the constant vector $\breve{\mu}$ as $g=1$ is linearly independent of $f\bar{\alpha}$. We see that the roots of the VAR characteristic equation

$$\det\{-f\bar{\alpha}\beta' z+(1-z)\big(I_4-\sum_{i=1}^{3}H_i z^i\big)\}=0$$

are either equal to 1 having multiplicity 3 or greater than 1 in absolute values. It can be also checked that Assumption $A3$ in Banerjee et.al (1993, p. 147), which is analogous to their Assumption $B3$ referred in Section 3, is satisfied, constructing δ and α in a manner similar to for β and γ in Example 1. These results ensures that (37) can be converted to a special case of (1) with the cointegrating rank 1 and that (37) is the VAR in GRT (see Banerjee et.al (1993, pp. 148-150) e.g.). On the other hand, in view of $(1-L)\breve{\mu}=0$ and $\beta'\mu_0=0$ etc., it is derived that $(I_4-\sum_{i=1}^{3}H_iL^i)\mu_0=\breve{\mu}$. Based on this, (37) is converted to a special case of (8) as $\alpha=f\bar{\alpha}$.

The tables below show relative frequency (or probability) distributions for the estimators produced from JT, JT*, AIC, SIC and HQ as the methods, provided that JT and JT* correspond to 0.05 and 0.01 significant levels respectively. For each method, lag order (VAR scheme) and sample size, we tabulate the relative frequencies classified into three events, corresponding to the occurrence of underestimation, correct or consistent estimation, and overestimation on the true rank value, denoted by the notations 'U', **'C'** and 'O' respectively. The numerical values in each row are the relative frequencies for one method, one scheme and three sample sizes. The first column of each table lines up 0, 1,$\cdots$, 8 for p, and in Tables 5 and 6, 3 as the true VAR lag order is suffixed by $^{+}$.

Now, let us survey finite sample performances of these estimation methods through the tables comparatively. The frequency that each method selects the true value strongly depends on the VAR scheme/the value of p under $T=100$ or 200, unlike the large sample ones established in the previous section. It is recognized that the increase of p tends to cause underestimation more frequently although to what extent it appears depends on the method/DGP. The case in which a finite lag order VAR is exactly one in GRT such as in Example 3 seems to produce relatively good results throughout all the methods. The performances of SIC and HQ are better than that of JT or JT* throughout all the DGPs if any of suitable lag orders in the sense that those are able to exhibit the ability is adopted. In most of the methods and DGPs except SIC in Example 3, the results for $p=0$ are far from being satisfactory mainly owing to the occurrence of overestimation. AIC tends to overestimate the true rank than SIC or HQ, in concord with the conventional on the information criteria. SIC achieves accurate estimation with high frequencies under $p=1$ to 3, but its performance is noticeably poor under $p=6$ to 8 and $T=100$ or 200, as clarified by the occurrence of underestimation with high frequencies. HQ seems to show the best performance among all the methods on the whole if T is as many as 100 or 200. On the other hand, as T attains to about 500, SIQ shows remarkably accurate estimation with relative frequencies close to one and it is robust for the selection of p.

7 Discussion

We have discussed the issue of detecting the cointegrating rank based on finite lag order VAR schemes which are not derived from GRT and the information criteria. It was established that based on the result that the rank of $\alpha(p)$ in (10)/(11) or (16)/(17) as the coefficient matrix associated with the cointegrating relations is equal to the cointegrating rank, estimating the rank by the direct application of each of the information criteria can achieve the conventional asymptotic desirability such as the consistency under any VAR scheme even if the lag order p is arbitrarily given.

Monte Carlo experiments in the previous section generally show that the results on estimation brought about by the information criteria are better than those of Johansen's rank test, particularly under DGPs in which GRT does not hold and the sample size T attaining to 200. Observing the experiments wholly, HQ seems to be favorable under T less than 200. On the other hand, the results for T equal to 500 reflect the asymptotics considerably, although those for $p=0$ are not necessarily so in most cases. The accuracy of the rank test is controlled by the significance level constraint even if it exhibits its asymptotics. The superiority of SIC or HQ under $T=500$ is far more noticeable compared with one for $T=100$ or 200. In particular, SIC sufficiently shows the consistency property, and as a result, under T as many as 500, SIC may be strongly recommended.

However, as observed in Monte Carlo experiments, several theoretical conclusions for large samples are not necessarily tenable under such finite sample size as $T=100$ or 200. Then the performance of each information criterion is different according to the VAR scheme or its lag order denoted as p: some lag order schemes lead to the true rank with high frequencies, whereas others do not. We can also read that any method possesses a tendency to select of a smaller rank value as p increases, and it is guessed that the effect of $\beta' y_{t-1}$ on the behavior of Δy_t is absorbed by that of Δy_{t-i}, $i=1,\cdots,p$, as p is not so small and T is not large, resulting in the weakness of the effect. Similarly, the unsatisfactory results for $p=0$ may be caused by that $-\log\{1-\hat{\lambda}_{j+1}(0)\}$ is considerably large in comparison with $2kC_T$ even for $j \geq r$. Note that this is not peculiar to the situation in which GRT does not hold such as Example 1 or 2. We should not overlook that even when GRT holds and the VAR lag order is finite, the asymptotics do not sufficiently appear under such finite samples, as observed for Example 3.

The selection of an 'optimal' VAR lag order (including the true lag order for such cases as Example 3) may be significant, particularly under T as many as 100 or 200, as stated above. It is well-established based on the information criteria if GRT holds and T is large, particularly for the case in which the lag order of the VAR in GRT is finite. For the infinite lag order case, Qu and Perron (2007) adopted not fixed p but p_T such that $\lim_{T\to\infty} p_T = \infty$ and $\lim_{T\to\infty} p_T / T^{1/3} = 0$, i.e., a lag order which increases at a slower rate than $T^{1/3}$ as T goes to ∞, as the upper boundary for the lag orders considered. In view of this, it may be reasonable to select p either equal or less than such a number as the boundary: the number is guessed to be at most 4 for $T=100$, 5 for $T=200$, 7 for $T=500$, noting that $100^{1/3}=4.64159$, $200^{1/3}=5.848$, $500^{1/3}=7.9372$. Generally, the asymptotics in Section 5 are not established under p which is 'large' in the sense that it is not fixed but increases as T goes to ∞, and the undesirable results of the experiments under such lag orders as 4 to 8 seem to be owing to this. Under the general situation in which GRT is not ensured to hold, finding the optimal one or establishing a valid procedure to achieve it is not easy even if T is large and left to the future research.

The issue of inferring other parameters in (10)/(11) or (16)/(17) after the rank is determined was not discussed in this paper. It will not be so difficult to show that under each scheme, using reduced rank regression substituted by the estimated rank value leads to some consistent estimation, since the rank can be consistently estimated as discussed above. However, constructing some practical hypothesis tests on those may not be easy unlike in Johansen (1992a), since it seems to be difficult to derive asymptotic distributions which are free of nuisance parameters in consequence of the existence of serial correlation of the error vectors. We will leave the formal discussion to future research.

FOOTNOTES

1 Even some models/schemes with serially correlated error vectors may be referred to as those using the term *VAR* afterwards in this paper.

2 Following Davidson (1994, p. 23), $\|F\| = \{\sum_{j=1}^{k} f_{jj}^2\}^{1/2}$ with c_{jj} as the j−th diagonal element of F of $k \times k$.

3 For example, both $\overline{\eta}_t = e_t + e_{t-1}$ and $\check{\eta}_t = e_t - 1.2e_{t-1}$ are of $I(0)$ in spite of their noninvertibility, where $\{e_t\}$ is a sequence of serially uncorrelated random variables with $Ee_t = 0$ and $Ee_t^2 = \sigma < \infty$.

4 The term *higher-order* cointegration is used for the case that $b > 1$ and $d = 1$ in Engle and Granger's (1987) definition of cointegration.

5 For the coefficient vector of $\check{\mu}_q t^q$ as the deterministic trend of the highestorder, it is often supposed that $\check{\mu}_q = C(1)\check{\mu}$, with a k-dimensional constant vector $\check{\mu}$. This implies that as y_t is cointegrated, $\check{\mu}_q t^q$ as well as the stochastic ones $C(1)\left(\sum_{h=1}^{t} \varepsilon_h\right)$ is removed in $\beta' y_t$, as seen by (4) and (6).

6 We use the common notation for H_i in both (8) and (9), although not necessarily identical in values. We may interpret (8) and (9) to be nested models within the framework of hypothesis testing for the cointegrating rank like Johansen's test. A similar matter is applied to other notations presented later such as $H_i(p)$ in (10) and (11).

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

References

Akaike, H. (1973) Information theory and an extension of the maximum likelihood principle, in B.N. Petrov & F. Csáki (eds.) *Information Theory and an Extension of the Maximum Likelihood Principle*, 2nd International Symposium on Information Theory, pp. 267-281. Budapest: Academic Kiadó.

Aznar, A. & M. Salvador (2002) Selecting the rank of the cointegrating space and the form of the intercept using an information criterion, *Econometric Theory* **18**, 926-947.

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

Davidson, J. (1994) *Stochastic Limit Theory*, Oxford: Oxford University Press.

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

Engle, R.F., & B.S. Yoo (1987) Forecasting and testing in co-integrated systems, *Journal of Econometrics* **35**, 143-159.

Engle, R.F., & B.S. Yoo (1991) Cointegrated economic time series: a survey with new results, in

Long-Run Economic Relationships: Readings in Cointegration, ed. by R. F. Engle & C. W. J. Granger. Oxford: Oxford University Press, pp. 237-260.

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

Gregoir S. & G. Laroque (1993) Multivariate time series: a polynomial error correction representation theorem, *Econometric Theory* **9**, 329-342.

Hannan E. J. & B.G. Quinn (1979) The determination of the order of an autoregression, *Journal of the Royal Statistical Society, Series B* **41**, 190-195.

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

Johansen, S. (1992b) Determination of cointegration rank in the presence of a linear trend, *Oxford Bulletin of Economics and Statistics* **54**, 383-397.

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

MacKinnon, J.G., A.A. Haug & L. Michelis (1999) Numerical distribution functions of likelihood ratio tests for cointegration, *Journal of Applied Econometrics* **14**, 563-577.

Park, J.Y. & P.C.B. Phillips (1988) Statistical inference in regressions with integrated processes: Part Ⅰ, *Econometric Theory* **4**, 468-497.

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

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

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

Schwarz, G. (1978) Estimating the dimension of a model, *Annals of Statistics* **6**, 461-464.

Shibata, R. (1976) Selection of the order of an autoregressive model by Akaike's information criterion, *Biometrika* **63**, 117-126.

Shintani M. (2001) A simple cointegrating rank test without vector autoregression, *Journal of Econometrics* **105**, 337-362.

Sims, C.A., J.H. Stock, & M.F. Watson (1990) Inference in linear time series models with some unit roots, *Econometrica* **58**, 113-144.

Whittle, P. (1983) *Prediction and Regulation by Linear Least-Square Methods*, 2nd ed. Oxford: Basil Blackwell.

Appendix

Proof of Lemma 1: Suppose $s<k$. It is trivial by the definition of l.l.s. prediction that

$$P(E\Delta y_{t-i}|1)=E\Delta y_{t-i},\quad \forall_i\geq 0,$$
$$P(E\beta' y_{t-1}|1)=E\beta' y_{t-1}, \tag{A.1}$$

and that

$$P(\Delta\bar{y}_{t-i}|1)=0\qquad \forall_i\geq 0,\quad P(\beta' v_{t-1}|1)=0. \tag{A.2}$$

Noting

$$P(\Delta y_{t-i}|1)=P(\Delta\bar{y}_{t-i}|1)+P(E\Delta y_{t-i}|1)\quad \forall_i\geq 0,$$
$$P(\beta' y_{t-1}-\beta'\xi_0|1)=P(\beta' v_{t-1}|1)+P(E\beta' y_{t-1}|1),$$

$(A.1)$ and $(A.2)$ lead to

$$\Delta y_{t-i}-P(\Delta y_{t-i}|1)=\Delta\bar{y}_{t-i},\qquad \forall_i\geq 0,$$
$$\beta' y_{t-1}-\beta'\xi_0-P(\beta' y_{t-1}-\beta'\xi_0|1)=\beta' v_{t-1}. \tag{A.3}$$

It follows from the first relation of $(A.2)$ as $i=0$ and $(A.3)$ that

$$P(\Delta\bar{y}_t|1,\beta' y_{t-1}-\beta'\xi_0,\Delta y_{t-i};i=1,\cdots,p)$$
$$=P(\Delta\bar{y}_t|\beta' v_{t-1},\Delta\bar{y}_{t-i};i=1,\cdots,p). \tag{A.4}$$

Since obviously

$$P(\Delta y_t|1,\beta' y_{t-1}-\beta'\xi_0,\Delta y_{t-i};i=1,\cdots,p)$$
$$=P(\Delta\bar{y}_t|1,\beta' y_{t-1}-\beta'\xi_0,\Delta y_{t-i};i=1,\cdots,p)$$
$$+P(E\Delta y_t|1,\beta' y_{t-1}-\beta'\xi_0,\Delta y_{t-i};i=1,\cdots,p)$$

and

$$P(E\Delta y_t|\,1, \beta' y_{t-1}-\beta'\xi_0, \Delta y_{t-i};\, i=1,\cdots,p)=E\Delta y_t,$$

from (A.4) we derive

$$P(\Delta y_t|\,1, \beta' y_{t-1}-\beta'\xi_0, \Delta y_{t-i};\, i=1,\cdots,p)-E\Delta y_t$$
$$=P(\Delta\bar{y}_t|\beta' v_{t-1}, \Delta\bar{y}_{t-i},\, i=1,\cdots,p).$$

If $s=k$, by a similar manner

$$P(\Delta y_t|\,1, \Delta y_{t-i};\, i=1,\cdots,p)-E\Delta y_t$$
$$=P(\Delta\bar{y}_t|\,\Delta\bar{y}_{t-i};\, i=1,\cdots,p).$$

(14) or (15) of the lemma is only a direct consequence of the above relation for each case of s, and (16) and (17) are derived straightforwardly by substituting $\Delta y_{t-i}-\sum_{j=0}^{q-1}\mu_j(t-i)^j$, $i=0, 1,\cdots, p$, and $\beta' y_{t-1}-\beta'\sum_{j=0}^{q}\check{\mu}_j(t-1)^j-\beta'\xi_0$ for $\Delta\bar{y}_{t-i}$, $i=0, 1,\cdots, p$, and $\beta' v_{t-1}$ in (10) and (11) respectively.

Next, with the notations on (1), define the k-dimensional series W_{t-i}, $i=0, 1,\cdots$, and $\bar{\eta}_t(p)$ as

$$W_{t-i}=\begin{bmatrix}\beta' v_{t-i}\\ \gamma'\Delta\bar{y}_{t-i}\end{bmatrix},\quad \bar{\eta}_t(0)=W_t-P(W_t|\beta' v_{t-1}),$$

$$\bar{\eta}_t(p)=W_t-P(W_t|\beta' v_{t-p-1}, W_{t-i};\, i=1,\cdots,p),\ \ \textit{if}\ p\geq 1$$

if $s<k$ and

$$W_{t-i}=\Delta\bar{y}_{t-i},\ \bar{\eta}_t(0)=W_t,$$
$$\bar{\eta}_t(p)=W_t-P(W_t|\,W_{t-i};\, i=1,\cdots,p),\ \textit{if}\ p\geq 1$$

if $s=k$. Note that $\bar{\eta}_t(p)=\varepsilon_t(p)$ if $s=k$.

The following two lemmas are used in the proof of Theorem 1:

Lemma A.1 *For the case* $s<k$, *with* β, γ, $\alpha(p)$ *and* $H_i(p)$ *on* (1) *or* (10), *let* $\bar{H}_i(p)$ *be* $k\times k$ *matrices constructed as*

$$\bar{H}_1(p)=\begin{bmatrix}\beta'\\ \gamma'\end{bmatrix}\Big([\beta(\beta'\beta)^{-1}+\alpha(p), 0]$$
$$+H_1(p)[\beta(\beta'\beta)^{-1},\ \gamma(\gamma'\gamma)^{-1}]\Big),$$

$$\bar{H}_i(p)=\begin{bmatrix}\beta'\\ \gamma'\end{bmatrix}\Big(H_i(p)[\beta(\beta'\beta)^{-1},\ \gamma(\gamma'\gamma)^{-1}]$$
$$-H_{i-1}(p)\,[\beta(\beta'\beta)^{-1}, 0]\Big),$$
$$i=2,\cdots,p,\qquad p\geq 2,$$

$$\bar{H}_{p+1}(p)=\begin{bmatrix}\beta'\\ \gamma'\end{bmatrix}H_p(p)[\beta(\beta'\beta)^{-1}, 0]\qquad p\geq 1.$$

Then, with W_t *and* $\bar{\eta}_t(p)$ *given above and* $\varepsilon_t(p)$ *in* (10), *we have:*

$$W_t=\sum_{i=1}^{p+1}\bar{H}_i(p)W_{t-i}+\bar{\eta}_t(p), \tag{A.5}$$

$$I_k-\sum_{i=1}^{p+1}\bar{H}_i(p)=\begin{bmatrix}\beta'\\ \gamma'\end{bmatrix}\Big[-\alpha(p),$$
$$\Big(I_k-\sum_{i=1}^{p}H_i(p)\Big)\gamma(\gamma'\gamma)^{-1}\Big], \tag{A.6}$$

$$\alpha(p)=-\begin{bmatrix}\beta'\\ \gamma'\end{bmatrix}^{-1}\Big(I_k-\sum_{i=1}^{p+1}\bar{H}_i(p)\Big)\begin{bmatrix}I_r\\ 0\end{bmatrix}, \tag{A.7}$$

$$\bar{\eta}_t(p)=\begin{bmatrix}\beta'\\ \gamma'\end{bmatrix}\varepsilon_t(p). \tag{A.8}$$

Proof: It suffices to show only the case $p\geq 1$ since the case $p=0$ are trivial.
Using

$$\beta'\Delta\bar{y}_{t-i}=\beta'\Delta v_{t-i},\qquad I_k=\beta(\beta'\beta)^{-1}\beta'+\gamma(\gamma'\gamma)^{-1}\gamma',$$

and multiplying $[\beta,\ \gamma]'$ to both sides of (10), we obtain

$$\begin{bmatrix}\beta'\\ \gamma'\end{bmatrix}\Delta\bar{y}_t=\begin{bmatrix}\beta'\\ \gamma'\end{bmatrix}\Big(\alpha(p)+H_1(p)\beta(\beta'\beta)^{-1}\Big)\beta' v_{t-1}$$
$$+\sum_{i=2}^{p}\begin{bmatrix}\beta'\\ \gamma'\end{bmatrix}(H_i(p)-H_{i-1}(p))\beta(\beta'\beta)^{-1}\beta' v_{t-i}$$
$$+\sum_{i=1}^{p}\begin{bmatrix}\beta'\\ \gamma'\end{bmatrix}H_i(p)\gamma(\gamma'\gamma)^{-1}\gamma\Delta\bar{y}_{t-i}$$
$$-\begin{bmatrix}\beta'\\ \gamma'\end{bmatrix}H_p(p)\beta(\beta'\beta)^{-1}\beta' v_{t-p-1}+\begin{bmatrix}\beta'\\ \gamma'\end{bmatrix}\varepsilon_t(p).$$

Using $\beta'\Delta\bar{y}_{t-1}=\beta'\Delta v_{t-1}$ again in the above relation with some arrangements leads to

$$W_t=\sum_{i=1}^{p+1}\bar{H}_i(p)W_{t-i}+\begin{bmatrix}\beta'\\ \gamma'\end{bmatrix}\varepsilon_t(p). \tag{A.9}$$

It is easy to see

$$P(W_t|\beta' v_{t-p-1}, W_{t-i};\, i=1,\cdots,p)$$
$$=P(W_t|\beta' v_{t-1}, \Delta\bar{y}_{t-i};\, i=1,\cdots,p),$$

which, together with $P(\beta' v_{t-1}|\beta' v_{t-1}, \Delta\bar{y}_{t-i};\, i=1,\cdots,p)=\beta' v_{t-1}$ as the nature of the l.l.s. prediction, $\Delta v_t=\Delta\bar{y}_t$ and the definition of $\varepsilon_t(p)$, derives (A.8). (A.5) follows immediately after (A.8) is used in (A.9).

For (A.6), using the definitions of $\bar{H}_i(p)$ and noting

$$[-\alpha(p)z, 0]$$

$$=-\alpha(p)\beta' L/(1-L)[\beta(\beta'\beta)^{-1}(1-L),\ \gamma(\gamma'\gamma)^{-1}],$$

we obtain

$$\Big(I_k-\sum_{i=1}^{p+1}\bar{H}_i(p)L^i\Big)W_t$$

$$=\begin{bmatrix}\beta'\\ \gamma'\end{bmatrix}\Big(-\alpha(p)\beta' L/(1-L)+I_k-\sum_{i=1}^{p}H_i(p)L^i$$

$$\cdot[\beta(\beta'\beta)^{-1}(1-L),\ \gamma(\gamma'\gamma)^{-1}]W_t,$$

which requires

$$I_k-\sum_{i=1}^{p+1}\bar{H}_i(p)z^i$$

$$=\begin{bmatrix}\beta'\\ \gamma'\end{bmatrix}[-\alpha(p)z,\ 0]+\begin{bmatrix}\beta'\\ \gamma'\end{bmatrix}\Big(I_k-\sum_{i=1}^{p}H_i(p)z^i\Big)$$

$$\cdot[\beta(\beta'\beta)^{-1}(1-z),\ \gamma(\gamma'\gamma)^{-1}].$$

Substituting 1 for z in the above equation is followed immediately by $(A.6)$. $(A.7)$ is only a direct consequence of $(A.6)$.

Putting

$$\tilde{H}_i(p)=H_i(p),\quad i=1,\cdots,p,\quad p\geq 1,\quad \tilde{H}_{p+1}(p)=0$$

for the case $s=k$, it is obvious that $(A.5)$ holds for any s.

Lemma A.2 *For $\bar{H}_i(p)$ given in Lemma A.1 and subsequent statement, all the roots of*

$$\det\Big(I_k-\sum_{i=1}^{p+1}\bar{H}_i(p)z^i\Big)=0$$

are greater than 1 in absolute values.

Proof: For $p\geq 1$ and $i=0,\ 1$, define $W_{t-i}(p)$ *of* $(kp+r)\times 1$, $\check{\eta}_t(p)$ of $(kp+r)\times 1$, $\bar{H}(p)$ of $(kp+r)\times(kp+r)$ and $\check{H}(p)$ of $k(p+1)\times k(p+1)$ as

$$W_{t-i}(p)=\begin{bmatrix}W_{t-i}\\ W_{t-i-1}\\ \vdots\\ W_{t-i-p+1}\\ \beta' v_{t-i-p}\end{bmatrix}\quad if\ s<k,$$

$$W_{t-i}(p)=\begin{bmatrix}W_{t-i}\\ W_{t-i-1}\\ \vdots\\ W_{t-i-p+1}\end{bmatrix}\quad if\ s=k,$$

$$\check{\eta}_t(p)=[\bar{\eta}_t',\ 0,\cdots,\ 0]',$$

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

$$if\ s<k,$$

with

$$\bar{H}_{p+1;1}(p)=\bar{H}_{p+1}\begin{bmatrix}I_r\\ 0\end{bmatrix}=-\begin{bmatrix}\beta'\\ \gamma'\end{bmatrix}H_p(p)\beta(\beta'\beta)^{-1},$$

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

$$if\ s=k,$$

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

Then it is easy to see

$$\det\Big(I_k-\sum_{i=1}^{p+1}\bar{H}_i(p)z^i\Big)=\det\Big(I_{k(p+1)}+\check{H}(p)z\Big).\tag{A.10}$$

It also follows that

$$\det\Big(I_{k(p+1)}-\check{H}(p)z\Big)=\det\Big(I_{kp+r}-\bar{H}(p)z\Big).\tag{A.11}$$

On the other hand, put

$$W_{t-i}(0)=\beta' v_{t-i}\quad i=0,\ 1,\quad \check{\eta}_t(0)=[I_r,\ 0]\,\bar{\eta}_t(0),$$
$$\bar{H}(0)=I_r+\beta'\alpha(0)$$

if $s<k$ and

$$W_{t-i}(0)=W_{t-i}=\Delta\bar{y}_{t-i}\quad i=0,\ 1,$$
$$\check{\eta}_t(0)=\bar{\eta}_t(0)=\varepsilon_t(0),\quad \bar{H}(0)=0$$

if $s=k$ in order to incorporate the case $p=0$ into the framework using $\bar{W}_t(p)$. If $s<k$, note

$$\det\Big(I_k-\bar{H}_1(0)z\Big)=\det\Big(I_r-\bar{H}(0)z\Big)$$

in view of Lemma A.1.

Using the notations above, $(A.5)$ is rewritten as

$$W_t(p)=\bar{H}(p)W_{t-1}(p)+\check{\eta}_t(p).\tag{A.12}$$

It is easy to see in view of (1) to (5) that $W_{t-i}(p)$ is weakly stationary, purely nondeterministic and ergodic with mean zero. Therefore, we can let

$$R_W(0;p)=EW_{t-1}(p)W'_{t-1}(p),\ R_W(1;p)=EW_t(p)W'_{t-1}(p)$$

with the existence of the inverse of $R_W(0;p)$. Since $E\check{\eta}_t(p)W'_{t-1}(p)=0$ by definition, it follows from (A.12) that

$$\bar{H}(p)=R_W(1;p)R_W^{-1}(0;p),$$

which, together with (A.10) and (A.11), implies that the equation

$$\det\left(I_k-\sum_{i=1}^{p+1}\bar{H}_i(p)z^i\right)=0$$

is equivalent to

$$\det\left(R_W(0;p)-R_W(1;p)z\right)=0. \qquad (A.13)$$

Now, consider z satisfying (A.13). Suppose that z is nonzero and real, noting that (A.13) does not hold for $z=0$. Then there exists a $(kp+r)\times 1$ real vector $b\neq 0$ satisfying

$$b'(R_W(0;p)-R_W(1;p)z)=0,$$

which leads to

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

Since z^{-1} is exactly the first-order autocorrelation coefficient of $b'W_t(p)$ which is is weakly stationary, purely nondeterministic and 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.13). Then we can find $(kp+r)\times 1$ complex vector b and $\bar{b}$ satisfying

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

With i denoting the imaginary, $(kp+r)\times 1$ 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

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

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

$$b_1'R_W(0;p)=(z_1b_1'-z_2b_2')R_W(1;p),$$
$$b_2'R_W(0;p)=(z_1b_2'+z_2b_1')R_W(1;p),$$

which requires that either $b_1\neq 0$ and $z_1b_1-z_2b_2\neq 0$ or $b_2\neq 0$ and $z_1b_2+z_2b_1$ holds. Noting that $(z_1b_1'-z_2b_2')R_W(0;\ p)\ (z_1b_1-z_2b_2)$ and $b_1'R_W(0;\ p)b_1$ are the variances of $(z_1b_1'-z_2b_2')W_t(p)$ and $b_1'W_{t-1}(p)$ respectively and that $(z_1b_1'-z_2b_2')R_W(1;\ p)b_1$ is the covariance of those series, we have

$$|(z_1b_1'-z_2b_2')R_W(1;p)b_1|^2$$
$$\leq\left(z_1^2b_1'R_W(0;p)b_1+z_2^2b_2'R_W(0;p)b_2-2z_1z_2b_1R_W(0;p)b_2\right)(b_1'R_W(0;p)b_1).$$

Similarly,

$$|(z_1b_2'+z_2b_1')R_W(1;p)b_2|^2$$
$$\leq\left(z_1^2b_2'R_W(0;p)b_2+z_2^2b_1'R_W(0;p)b_1+2z_1z_2b_1R_W(0;p)b_2\right)(b_2'R_W(0;p)b_2).$$

In view of the restriction on b_i and the properties of the series stated above, we see that at least one of the above two inequalities holds strictly (i.e., one of the sign $\leq$ can be replaced by $<$). Consequently,

$$(b_1'R_W(0;p)b_1+b_2'R_W(0;p)b_2)$$
$$<(z_1^2+z_2^2)(b_1'R_W(0;p)b_1+b_2'R_W(0;p)b_2),$$

which requires $|z|>1$.

Proof of Theorem 1: (A.7), together with $rank\left(I_k-\sum_{i=1}^{p+1}\bar{H}_i(p)\right)=k$ following from Lemma A.2, immediately leads to (i).

For (ii), put $\bar{H}(z;p)=I_k-\sum_{i=1}^{p+1}\bar{H}_i(p)z^i$. Then ($A$.5) is written as

$$\bar{H}(L;p)W_t=\bar{\eta}_t(p),$$

and it follows from (5)/(1) and (A.8) that

$$\varepsilon_t(p)=\begin{bmatrix}\beta'\\ \gamma'\end{bmatrix}^{-1}\bar{H}(L;p)\begin{bmatrix}\beta'C^{(1)}(L)\\ \gamma'C(L)\end{bmatrix}\varepsilon_t, \quad \text{if } s<k, \qquad (A.14)$$

$$\varepsilon_t(p)=\bar{H}(L;p)C(L)\,\varepsilon_t. \quad \text{if } s=k. \qquad (A.15)$$

Putting

$$B(z;p)=\begin{bmatrix}\beta'\\ \gamma'\end{bmatrix}^{-1}\bar{H}(z;p)\begin{bmatrix}\beta'C^{(1)}(z)\\ \gamma'C(z)\end{bmatrix} \quad \text{if } s<k,$$
$$=\bar{H}(z;p)C(z) \quad \text{if } s=k$$

and noting that $\bar{H}(0;p)=C(0)=I_k$, $C^{(1)}(0)=Ik-\gamma\tau\delta'$, the supposition on C_i in (1) and $rank\ \bar{H}(1;p)=k$ by Lemma A.2, it is easy to see that all the requirements for (ii) are satisfied.

For (iii), to obtain the result for the case $s=k$ is trivial since $B(1;p)=\bar{H}(1;p)C(1)$ in the proof of (ii).

For the case $s<k$, with regard to (10), define $A(z;p)$ as

$$\begin{aligned} A(z;p) &= -\alpha(p)\beta' z+(1-z)\Big(I_k-\sum_{i=1}^{p}H_i(p)z^i\Big) \\ &= -\alpha(p)\beta'+(1-z)A^{(1)}(z;p), \end{aligned}$$

with $A^{(1)}(z;p)=I_k+\alpha(p)\beta'-\sum_{i=1}^{p}H_i(p)z^i$. Recalling that $\beta' v_t=\beta'(\bar{y}_t-\xi_0)$, noting that $(1-L)\ \xi_0=\xi_0-\xi_0=0$ and using (18), (10) is written as for $t=p+2, p+3, \cdots$,

$$A(L;p)(\bar{y}_t-\xi_0)=B(L;p)\varepsilon_t. \qquad (A.16)$$

On the other hand, based on both sides of (3) multiplied by $A(L;p)$,

$$\begin{aligned} &A(L;p)(\bar{y}_t-\xi_0) \\ &=A(L;p)\Big(C(1)+(1-L)C^{(1)}(L)\Big)\Big(\sum_{h=1}^{t}\varepsilon_h\Big). \end{aligned}$$

Recalling that $\beta' C(1)=0$, the above relation is converted to

$$\begin{aligned} &A(L;p)(\bar{y}_t-\xi_0) \\ &=\Big(A^{(1)}(L;p)C(1)-\alpha(p)\beta' C^{(1)}(L)\Big)\varepsilon_t \\ &\quad +A^{(1)}(L;p)C^{(1)}(L)(1-L)\varepsilon_t, \quad (A.17) \end{aligned}$$

Equating the right-hand sides of $(A.16)$ and $(A.17)$, we obtain

$$\begin{aligned} B(z;p)=&A^{(1)}(z;p)C(1)-\alpha(p)\beta' C^{(1)}(z) \\ &+(1-z)A^{(1)}(z;p)C^{(1)}(z), \end{aligned}$$

which is followed by

$$B(1;p)=A^{(1)}(1;p)\gamma\tau\delta'-\alpha(p)\beta' C^{(1)}(1).$$

In view of the definition of $A^{(1)}(1;p)$, we see from the above relation that for any $\delta(p)$ satisfying the requirements for (iii),

$$\delta'(p)B(1;p)=\delta'(p)\Big(I_k-\sum_{i=1}^{p}H_i(p)\Big)\gamma\tau\delta'. \qquad (A.18)$$

On the other hand, $(A.6)$, together with Lemma A.2, requires that

$$rank\Big[\alpha(p), \Big(I_k-\sum_{i=1}^{p}H_i(p)\Big)\gamma\Big]=k,$$

which ensures that there exist P_{11}^{-1} and P_{12} as $s\times s$ full rank and $r\times s$ matrices respectively such that

$$\Big(I_k-\sum_{i=1}^{p}H_i(p)\Big)\gamma=\delta(p)P_{11}^{-1}+\alpha(p)P_{12}.$$

This in turn implies that

$$\delta'(p)\Big(I_k-\sum_{i=1}^{p}H_i(p)\Big)\gamma$$

is full rank. Since τ is also full rank by definition, it is established that

$$rank\ \delta'(p)\Big(I_k-\sum_{i=1}^{p}H_i(p)\Big)\gamma\tau=s.$$

Putting

$$\tilde{\tau}(p)=\delta'(p)\Big(I_k-\sum_{i=1}^{p}H_i(p)\Big)\gamma\tau,$$

$(A.18)$, together with the rank value of $\tilde{\tau}(p)$ derived above, completes the proof of (iii).

Next, move to the proof of (iv). The construction of $B(z;p)$ in the proof of (ii), together with Lemma A.2, implies that $rank\ B(1;p)=k$ if and only if Condition Ⅰ is satisfied. This, together with the form of $\delta(p)$ satisfying the requirements for (iii), establishes that ψ takes the form required for (iv).

It is easy to establish (v) in view of the construction of $B(z;p)$ in the proof of (ii) and Lemma A.2 as well as Conditions Ⅰ and Ⅱ.

For (vi), with regard to (8), define $A(z)$ as

$$A(z)=-\alpha\beta' z+(1-z)\Big(I_k-\sum_{i=1}^{\infty}H_i z^i\Big).$$

Noting $A(z)$ above, $A(z;p)$ in the proof of (iii) and $B(z;p)$ in (ii), for $t=p+2, p+3, \cdots$, (8) and (10) are written as

$$A(L)\bar{y}_t=\varepsilon_t, \qquad (A.19)$$

$$A(L,p)\bar{y}_t=B(L;p)\varepsilon_t. \qquad (A.20)$$

Based on $(A.20)$ and the validity of $B^{-1}(z;p)$ established in (v), we have the following representation:

$$B^{-1}(L;p)A(L;p)\bar{y}_t=\varepsilon_t. \qquad (A.21)$$

It follows from $(A.19)$ and $(A.21)$ that

$$A(1)=B^{-1}(1;p)A(1;p). \qquad (A.22)$$

Since $A(1)=-\alpha\beta'$ and $A(1;p)=-\alpha(p)\beta'$ in virtue of (8), $A(z)$ above, and $A(z;p)$ in the proof of (iii), multiplying both sides of $(A.22)$ from right by $-\beta(\beta'\beta)^{-1}$ provides the result for (vi).

Proof of Corollary 1: For the purpose of simplicity, we shall write $\sum_{ij}(p)$ as $\sum_{ij}$ hereaafter. Based on the definitions of $u_t(p)$ and $\zeta_{t-1}(p)$, we easily find $k\times k$ constant matrices $\bar{K}_j(p;\ i)$ stated, which immediately leads to (19). In view of the definition of l.l.s. prediction,

$$u_t(p)=\varepsilon_t(p)+P(\Delta\bar{y}_t|\beta'\zeta_{t-1}(p)), \qquad (A.23)$$
$$\beta'\zeta_{t-1}(p)=\beta'\check{\zeta}_{t-1}(p)+P(\beta'v_{t-1}|u_t(p)), \qquad (A.24)$$

where

$$\check{\zeta}_{t-1}(p)=v_{t-1}-P(v_{t-1}|\Delta\bar{y}_{t-i};\ i=0,1,\cdots,p),$$

with the notice that $E\check{\zeta}_{t-1}(p)u_t'(p)=0$. Noting that $E\Delta\bar{y}_{t-i}\ \zeta_{t-1}'(p)\ \beta=0$ and $E\Delta\bar{y}_{t-i}\ u_t'(p)=0,\ i=1,\cdots,p$, we also see

$$P(\Delta\bar{y}_t|\beta'\zeta_{t-1}(p))=P(u_t(p)|\beta'\zeta_{t-1}(p))$$
$$=\textstyle\sum_{01}\beta(\beta'\sum_{11}\beta)^{-1}\beta'\zeta_{t-1}(p), \quad (A.25)$$
$$P(\beta'v_{t-1}|u_t(p))=P(\beta'\zeta_{t-1}(p)|u_t(p))$$
$$=\beta'\textstyle\sum_{10}\sum_{00}^{-1}u_t(p). \qquad (A.26)$$

Then (20) follows immediately from $(A.23)$ and $(A.25)$. Evaluating the predictions of all the terms of (10) onto $\{\Delta\bar{y}_{t-i};\ i=1,\cdots,p\}$, (21) can be also derived. For the remainder of the proof, put

$$\check{\textstyle\sum}_{11}=E\check{\zeta}_{t-1}(p)\check{\zeta}_{t-1}'(p).$$

Evaluating the covariance matrices of both sides of $(A.24)$ and using (21) in $(A.26)$, we have

$$\beta'\textstyle\sum_{11}\beta=\beta'\check{\sum}_{11}\beta+\beta'\sum_{11}\beta\alpha'(p)\sum_{00}^{-1}\alpha(p)\beta'\sum_{11}\beta. \qquad (A.27)$$

Since $u_t(p)$, $\beta'\zeta_{t-1}(p)$ and $\beta'\check{\zeta}_{t-1}(p)$ are purely nondeterministic in terms of Wold decomposition, $\sum_{00}$, $\sum_{00}^{-1}$, $\beta'\sum_{11}\beta$ and $\beta'\check{\sum}_{11}\beta$ are all positive definite. Putting this with $(A.27)$ and $rank\ \alpha(p)=r$ together, we see that $\beta'\sum_{11}\beta-\beta'\check{\sum}_{11}\beta$ is positive definite as well, and thus it is established that both

$$I_r-(\beta'\textstyle\sum_{11}\beta)^{-1/2}\beta'\check{\sum}_{11}\beta(\beta'\sum_{11}\beta)^{-1/2}$$

and

$$(\beta'\textstyle\sum_{11}\beta)^{-1/2}\beta'\check{\sum}_{11}\beta(\beta'\sum_{11}\beta)^{-1/2}$$

are positive definite, which implies (22).

Proof of Lemma 2: We first note that this lemma can be proved in the same manner as in the counterparts of Johansen (1988, 1995) essentially, although not applicable directly to some of the results. In the proof we shall state only the 0outline under the case $\tilde{M}(p)=M_Z(p)$ which is expected to be more complicated. For the purpose of simplicity, let us write $S_{ij}(p)$ and $M_Z(p)$ as S_{ij} and M_Z respectively, and suppose that i below can be any integer in $\{0,1,\cdots,p\}$ unless specified newly. Let $\Delta\bar{Y}_{-i}$, $\bar{E}_0$ and S_{-1} be $\check{T}\times k$ matrices, given as

$$\Delta\bar{Y}_{-i}^{1}=[\Delta\bar{y}_{p+2-i},\Delta\bar{y}_{p+3-i},\cdots,\Delta\bar{y}_{T-i}],$$
$$\bar{E}_0'=[\varepsilon_{p+2},\ \varepsilon_{p+3},\cdots,\ \varepsilon_T],$$

$$S_{-1}'=\Big[\sum_{h=1}^{p+1}\varepsilon_h,\ \sum_{h=1}^{p+2}\varepsilon_h,\cdots,\sum_{h=1}^{T-1}\varepsilon_h\Big],$$

let $\Delta\bar{Z}_{-1}(p)$ and $M_{\hat{\tau}}$ be $\check{T}\times pk$ and $\check{T}\times\check{T}$ matrices respectively, given as

$$\Delta\bar{Z}_{-1}(p)=[\Delta\bar{Y}_{-1},\cdots,\Delta\bar{Y}_{-p}],$$
$$M_{\hat{\tau}}=I_{\check{T}}-\hat{\tau}(q)D_T^{-1}(q)\big(D_T^{-1}(q)\hat{\tau}'(q)\hat{\tau}(q)D_T^{-1}(q)\big)^{-1}D_T^{-1}(q)\hat{\tau}'(q),$$

with the $(q+1)\times(q+1)$ matrix

$$D_T^{-1}(q)=diag\{T^{-1/2},T^{-3/2},\cdots,T^{-q+1/2}\},$$

and put

$$\tilde{Y}_{-1}'=C(1)\big(S_{-1}'+\check{\mu}\check{\tau}_q'\big).$$

Next, for i, $i'=0,1,\cdots,p$, let $\check{w}_{t-i}$ and $\check{u}_{t-i'}$ denote any linear combination of $\Delta\bar{y}_{t-i}$, v_{t-i}, $v_{t-i}(p)$, ε_{t-i}, $\varepsilon_t(p)$ and $(\Delta\bar{y}_{t-1}',\cdots,\Delta\bar{y}_{t-p}')$. Note that any of those series is weakly stationary and ergodic with mean zero and that therefore there exist constant matrices $\tilde{K}_{j;w}$ such that $\check{w}_{t-i}=\sum_{j=0}^{\infty}\tilde{K}_{j;w}\varepsilon_{t-i-j}$. Letting

$$\check{W}_{-i}'=[\check{w}_{p+2-i},\check{w}_{p+3-i},\cdots,\check{w}_{T-i}],$$
$$\check{U}_{-i'}'=[\check{u}_{p+2-i'},\check{u}_{p+3-i'},\cdots,\check{u}_{T-i'}]$$

and $\tilde{Y}_{-1}=(S_{-1}+\check{\tau}_q\check{\mu}')C'(1)$ and noting

$$M_Z = M_{\hat\tau} \qquad \text{if } p=0,$$
$$= M_{\hat\tau} - M_{\hat\tau}\Delta\bar Z_{-1}(p)\big(\Delta\bar Z'_{-1}(p)\, M_{\hat\tau}\Delta\bar Z_{-1}(p)\big)^{-1}\Delta\bar Z'_{-1}(p)\, M_{\hat\tau} \qquad \text{if } p\ge 1,$$

we can see from (1), (3), (16) and the constructions of S_{ij} that evaluating the asymptotics of the quantities

$$\check W'_{-i} M_{\hat\tau}\, \check U_{-i'}/T, \quad \check W'_{-i} M_{\hat\tau}\, \check\tau_q/T^{q+1/2},$$
$$\big(\Delta\bar Z'_{-1}(p)\, M_{\hat\tau}\Delta\bar Z_{-1}(p)/T\big)^{-1}, \quad \bar E'_0 M_{\hat\tau}\tilde Y_{-1}\gamma D_T^{-1}/T,$$
$$D_T^{-1}\gamma'\tilde Y'_{-1} M_{\hat\tau}\,\tilde Y_{-1}\gamma D_T^{-1}/T^2,$$
$$\big(\check W'_0 - \check W'_{-1}\big) M_{\hat\tau}\,\tilde Y_{-1}\gamma D_T^{-1}/T,$$

suffices for the required results. It is not so difficult to achieve it since these quantities are constructed only by weakly stationary and ergodic time series, deterministic trends and partial sums and the well-known statistical theory can be easily utilized for those. In fact, deriving the results

$$\check W'_{-i}\,\check\tau_j/T^{j+1/2} = O_p(1),$$
$$\big(D_T^{-1}(q)\,\hat\tau'(q)\,\hat\tau(q) D_T^{-1}(q)\big)^{-1} = O(1),$$
$$S'_{-1}\,\check\tau_j/T^{j+3/2} = O_p(1),$$
$$\big(\check W'_0 - \check W'_{-1}\big)\,\check\tau_j/T^{j+1/2} = O_p(T^{-1/2}), \quad \forall j\ge 0, \qquad (A.28)$$

related on the deterministic terms is trivial and the standard statistical theory for weakly stationary and ergodic processes, together with (A.28), shows that

$$\check W'_{-i} M_{\hat\tau}\,\check U_{-i'}/T = \check W'_{-i}\check U_{-i'}/T + O_p(T^{-1})$$
$$= R_{\check w\check u}(i'-i) + O_p(T^{-1/2}), \qquad (A.29)$$

where $R_{\check w\check u}(i'-i) = E\check w_{t-i}\,\check u'_{t-i'}$. It also follows from ($A$.29) that

$$\big(\Delta\bar Z'_{-1}(p)\, M_{\hat\tau}\,\Delta\bar Z_{-1}(p)/T\big)^{-1} = O_p(1).$$

However, there are some points it should be explained particularly for the framework in this paper. One notice should be turned to that not conventional ε_t but $\varepsilon_t(p)$ are used to construct some quantities. As a related matter, we will pay attention to the asymptotic property of $(\check W'_0 - \check W'_{-1})\, M_{\hat\tau}\,\tilde Y_{-1}\gamma D_T^{-1}/T$: we see that $E\,\check w_{t-i}\,\varepsilon'_t$ becomes $\check K_{0;w}\Lambda$ if $i=0$ and 0 if $i\ge 1$. It is also obvious that

$$\big(\check W'_0 - \check W'_{-1}\big)\, S_{-1} C'(1)\bar\gamma/T$$
$$= \{\check w_T\big(\sum_{h=1}^{T-1}\varepsilon'_h\big)/T\}C'(1)\bar\gamma - \{\check w_{p+1}\big(\sum_{h=1}^{p}\varepsilon'_h\big)/T\}C'(1)\bar\gamma$$
$$-\big(\sum_{t=p+1}^{T-1}\check w_t\,\varepsilon'_t/T\big)C'(1)\bar\gamma.$$

Thus it follows from (A.28), (A.29) and the definitions of $M_{\hat\tau}$, $\tilde Y_{-1}$, $\bar\gamma$ and D_T^{-1} that

$$\big(\check W'_0 - \check W'_{-1}\big)\, M_{\hat\tau}\,\tilde Y_{-1}\gamma D_T^{-1}/T$$
$$= -\check K_{0;w}\Lambda C'(1)\bar\gamma + O_p(T^{-1/2}). \qquad (A.30)$$

On the other hand, it is established by Park and Phillips (1988, 1989) etc. that as T increases,

$$\tilde\tau(p)\delta'\varepsilon_t/\sqrt{T} \Rightarrow \tilde\tau(p)\check G\, dW_s(t/T),$$
$$C(1)\varepsilon_t/\sqrt{T} \Rightarrow \gamma\, dW_s(t/T),$$
$$\varepsilon_t/\sqrt{T} \Rightarrow \check F\bar P\begin{pmatrix} dW_s(t/T)\\ dW_r(t/T)\end{pmatrix},$$
$$\gamma' C(1)\big(\sum_{h=1}^{t-1}\varepsilon_h=\big)/\sqrt{T} \Rightarrow \gamma'\gamma\,\check G\, W_s(t/T),$$

For the case $\gamma'_1 C(1)\,\check\mu \ne 0$, it is also shown that as T increases,

$$D_T^{-1}\gamma' C(1)\{\big(\sum_{h=1}^{t-1}\varepsilon_h\big) + t^q\check\mu\}/\sqrt{T}$$
$$= \bar\gamma' C(1)\big(\sum_{h=1}^{t-1}\varepsilon_h\big)/\sqrt{T} + O_p(T^{-q+1/2})$$
$$+ (\gamma'_1 C(1)\,\check\mu)\,(t^q/T^q, 0)' \Rightarrow \tilde G\,\bar W_s(t/T).$$

Based on the above results, we obtain

$$\tilde\tau(p)(\delta'\bar E'_0 M_{\hat\tau}\,\tilde Y_{-1}\gamma D_T^{-1}/T) \Rightarrow$$
$$\tilde\tau(p)\check G\big(\int_0^1 dW_s(u)\,\tilde W'_s(u)\big)\,\tilde G$$
$$\text{as } T\to\infty, \qquad (A.31)$$

$$C(1)\,(\bar E'_0 M_{\hat\tau}\,\tilde Y_{-1}\gamma D_T^{-1}/T) \Rightarrow$$
$$\gamma\,\tau\check G\big(\int_0^1 dW_s(u)\,\tilde W'_s(u)\big)\,\tilde G$$
$$\text{as } T\to\infty, \qquad (A.32)$$

$$\bar E'_0 M_{\hat\tau}\,\tilde Y_{-1}\gamma D_T^{-1}/T \Rightarrow$$
$$\check F\bar P\Big(\int_0^1\begin{pmatrix} dW_s(u)\\ dW_r(u)\end{pmatrix}\tilde W'_s(u)\Big)\,\tilde G$$
$$\text{as } T\to\infty, \qquad (A.33)$$

$$D_T^{-1}\gamma'\,\tilde Y'_{-1} M_{\hat\tau}\,\tilde Y_{-1}\gamma D_T^{-1}/T^2 \Rightarrow$$
$$\tilde G\big(\int_0^1\tilde W_s(u)\,\tilde W'_s(u)du\big)\,\tilde G$$
$$\text{as } T\to\infty, \qquad (A.34)$$

All the results required for the lemma follow from (A.28) to (A.34).

Proof of Lemma 3: Similar to Lemma 2, the results claimed by (i) and (ii) are essentially the same as in Johansen's (1988) Lemma 4 and 6 except that this lemma is under more general suppositions and shall be proved using arguments similar to in such lemmas. First, notice that the equation $\det\{\lambda S_{11}-S_{10}S_{00}^{-1}S_{01}\}=0$ is equivalent to

$$\det\{\lambda\begin{bmatrix}\beta'S_{11}\beta & \beta'S_{11}\gamma D_T^{-1}/T^{1/2}\\ D_T^{-1}\gamma'S_{11}\beta/T^{1/2} & D_T^{-1}\gamma'S_{11}\gamma D_T^{-1}/T\end{bmatrix} - \begin{bmatrix}\beta'S_{10}S_{00}^{-1}S_{01}\beta & \beta'S_{10}S_{00}^{-1}S_{01}\gamma D_T^{-1}/T^{1/2}\\ D_T^{-1}\gamma'S_{10}S_{00}^{-1}S_{01}\beta/T^{1/2} & D_T^{-1}\gamma'S_{10}S_{00}^{-1}S_{01}\gamma D_T^{-1}/T\end{bmatrix}\}=0. \tag{A.35}$$

In view of (27) and (31) to (33), λ satisfying (A.35) must be a root of either

$$\det\{\lambda D_T^{-1}\gamma'S_{11}\gamma D_T^{-1}/T+O_p(T^{-1})\}=0$$

or

$$\det\{\lambda\beta'S_{11}\beta-\beta'S_{10}S_{00}^{-1}S_{01}\beta+T^{-1}\tilde G_T(\lambda)\}=0,$$

with

$$\tilde G_T(\lambda)=T\{\sum_{j=1}^{\infty}\lambda^{2-j}O_p(T^{-j}) + \sum_{j=1}^{\infty}\lambda^{1-j}O_p(T^{-j})+\sum_{j=1}^{\infty}\lambda^{-j}O_p(T^{-j})\}.$$

This implies that $\hat\lambda_{r+h}(p)=O_p(T^{-1})$, $h=1,\cdots,s$. Letting $\tilde\lambda_j(p)$, $j=1,\cdots,r$, be the roots of detf$\{\lambda\beta'S_{11}\beta-\beta'S_{10}S_{00}^{-1}S_{01}\beta\}=0$, it is not so difficult to show that

$$\hat\lambda_j(p)=\tilde\lambda_j(p)+O_p(T^{-1})\quad j=1,\cdots,r. \tag{A.36}$$

Using (20), (21) and (27) to (29) and recalling $\lambda_j(p)$ in Corollary 1, it is also established that

$$\tilde\lambda_j(p)=\lambda_j(p)+O_p(T^{-1/2})\quad j=1,\cdots,r. \tag{A.37}$$

(A.36) and (A.37), together with (22), ensures that (i) holds.

For (ii), notice that

$$\{T\hat\lambda_k(p)\}^{-1}\ge\{T\hat\lambda_{k-1}(p)\}^{-1}\ge\cdots\ge\{T\hat\lambda_1(p)\}^{-1}$$

are the ordered eigenvalues of the equation

$$\det\{\begin{bmatrix}\beta'S_{11}\beta/T & \beta'S_{11}\gamma D_T^{-1}/T\\ D_T^{-1}\gamma'S_{11}\beta/T & D_T^{-1}\gamma'S_{11}\gamma D_T^{-1}/T\end{bmatrix} -\mu\begin{bmatrix}\beta'S_{10}S_{00}^{-1}S_{01}\beta & \beta'S_{10}S_{00}^{-1}S_{01}\gamma D_T^{-1}\\ D_T^{-1}\gamma'S_{10}S_{00}^{-1}S_{01}\beta & D_T^{-1}\gamma'S_{10}S_{00}^{-1}S_{01}\gamma D_T^{-1}\end{bmatrix}\}=0 \tag{A.38}$$

Following an argument similar to used in the proof of (i) and putting

$$\tilde B_T(\mu)=D_T^{-1}\gamma'S_{10}S_{00}^{-1}S_{01}\beta(\beta'S_{10}S_{00}^{-1}S_{01}\beta)^{-1}\beta'S_{10}S_{00}^{-1}S_{01}\gamma D_T^{-1}+\sum_{j=1}^{\infty}\mu^{-j}O_p(T^{-j}),$$

(A.38) accompanied by (28) and (32) establishes that $\{T\hat\lambda_{r+h}(p)\}^{-1}$, $h=1,\cdots,s$, are the roots of the

$$\det\{D_T^{-1}\gamma'S_{11}\gamma D_T^{-1}/T-\mu D_T^{-1}\gamma'S_{10}S_{00}^{-1}S_{01}\gamma D_T^{-1}+\mu\tilde B_T(\mu)\}=0.$$

Putting

$$\hat Q=(D_T^{-1}\gamma'S_{11}\gamma D_T^{-1}/T)^{-1/2}D_T^{-1}\gamma'S_{10}S_{00}^{-1}\{I_k-S_{01}\beta(\beta'S_{10}S_{00}^{-1}S_{01}\beta)^{-1}\beta'S_{10}S_{00}^{-1}\}\cdot S_{01}\gamma D_T^{-1}(D_T^{-1}\gamma'S_{11}\gamma D_T^{-1}/T)^{-1/2}$$

and letting $\tilde\nu_{r+1;T}(p),\cdots,\tilde\nu_{r+s;T}(p)$ denote the eigenvalues of $\hat Q$, it follows that

$$\{T\hat\lambda_{r+h}(p)\}^{-1}=\{\tilde\nu_{r+h;T}(p)\}^{-1}+O_p(T^{-1})\quad h=1,\cdots,s, \tag{A.39}$$

similar to (A.36). It is ensured from (27) to (29), (30) and (33) that that $\hat Q$ and $\hat Q^{-1}$ are of $O_p(1)$. This, together with (A.39), completes the proof of (ii).

For (iii), put

$$\check Q=(\beta'S_{11}\beta)^{-1/2}\beta'S_{10}S_{00}^{-1}S_{01}\beta(\beta'S_{11}\beta)^{-1/2}.$$

It is obvious that $\tilde\lambda_j(p)$, $j=1,\cdots,r$, are the eigenvalues of $\check Q$. Since $\check Q$ is asymptotically scale invariant to Λ and $\hat Q$ is in a similar condition in view of (19) and Lemma 2, we attain to the required result via (A.36) and (A.39).

Proof of Theorem 2: By definition we have

$$\begin{aligned}Pr(\hat r(p)<r)&=\sum_{j=0}^{r-1}Pr(\hat r(p)=j)\\ &\le\sum_{j=0}^{r-1}Pr\big(\bar I(j;p)\le\bar I(r;p)\big)\\ &=\sum_{j=0}^{r-1}Pr\Big(\frac{2(r-j)kC_T}{T}\ge -\sum_{i=j+1}^{r}\log\{1-\hat\lambda_i(p)\}\Big).\end{aligned}$$

Recalling that $-\log\{1-\hat\lambda_1(p)\}\ge\cdots\ge-\log\{1-\hat\lambda_r(p)\}$

>0, we see

$$Pr\Big(\frac{2(r-j)kC_T}{T} \geq -\sum_{i=j+1}^{r} \log\{1-\hat{\lambda}_i(p)\}\Big)$$

$$\leq Pr\Big(\frac{2kC_T}{T} \geq -\log\{1-\hat{\lambda}_r(p)\}.$$

Since $\lim_{T\to\infty}\frac{CT}{T}=0$ by definition, it must be required by Lemma 3 (i) that

$$\lim_{T\to\infty} Pr\Big(\frac{2(r-j)kC_T}{T} \geq -\sum_{i=j+1}^{r} \log\{1-\hat{\lambda}_i(p)\}\Big)=0$$

$$j=0,\cdots,r-1$$

Hence

$$\lim_{T\to\infty} Pr\,(\hat{r}(p)<r)=0, \tag{A.40}$$

which implies (35).

On the other hand,

$$Pr(\hat{r}(p)>r)=\sum_{j=r+1}^{k-1} Pr\,(\hat{r}(p)=j)$$

$$\leq \sum_{j=r+1}^{k-1} Pr\big(\bar{I}(j;p) \leq \bar{I}(r;p)\big)$$

$$=\sum_{j=r+1}^{k-1} Pr\Big(-\sum_{i=r+1}^{j} T\log\{1-\hat{\lambda}_i(p)\} \geq 2(j-r)k\,C_T\Big).$$

In view of Lemma 3 (ii),

$$\lim_{T\to\infty} \sum_{j=r+1}^{k-1} Pr\Big(-\sum_{i=r+1}^{j} T\log\{1-\hat{\lambda}_i(p)\} \geq 2(j-r)k\,C_T\Big)=0 \tag{A.41}$$

must hold if $\lim_{T\to\infty} C_T=\infty$. ($A.41$), together with ($A.40$), implies (34).

Table 1

DGP : (35) with $g=0$ in Example 1; $\tilde{M}(p)=M_{\Delta Z}(p)$ in $S_{ij}(p)$

$f=0.8$		$T=100$			$T=200$			$T=500$		
p	$\hat{r}$(p)	U	C	O	U	C	O	U	C	O
0	JT	0	0.1399	0.8601	0	0.1777	0.8223	0	0.2199	0.7801
	JT*	0	0.2282	0.7718	0	0.2884	0.7156	0	0.3295	0.6705
	AIC	0	0.1042	0.8958	0	0.1377	0.8623	0	0.1685	0.8315
	SIC	0	0.5994	0.4006	0	0.7408	0.2592	0	0.8585	0.1415
	HQ	0	0.2978	0.7022	0	0.4058	0.5942	0	0.5241	0.4759
1	JT	0.0002	0.3964	0.6034	0	0.4745	0.5255	0	0.5357	0.4643
	JT*	0.0023	0.5447	0.453	0	0.6254	0.3746	0	0.6771	0.3229
	AIC	0	0.3293	0.6707	0	0.3961	0.6039	0	0.4611	0.5389
	SIC	0.0199	0.8858	0.0943	0	0.9422	0.0578	0	0.9757	0.0243
	HQ	0	0.6362	0.3638	0	0.7467	0.2533	0	0.8252	0.1748
2	JT	0.0093	0.5718	0.4189	0	0.6659	0.3341	0	0.7166	0.2834
	JT*	0.0452	0.709	0.2458	0	0.8104	0.1896	0	0.8483	0.1517
	AIC	0	0.511	0.489	0	0.5941	0.4059	0	0.647	0.353
	SIC	0.2479	0.7385	0.0136	0.0002	0.9918	0.008	0	0.9987	0.0013
	HQ	0.0038	0.8381	0.1581	0	0.9065	0.0935	0	0.949	0.051
3	JT	0.0208	0.5912	0.388	0	0.6906	0.3094	0	0.731	0.269
	JT*	0.0954	0.7068	0.1978	0.0003	0.8321	0.1676	0	0.8608	0.1392
	AIC	0	0.5528	0.4472	0	0.6232	0.3768	0	0.6587	0.3413
	SIC	0.4614	0.535	0.0036	0.0069	0.9898	0.0033	0	0.9993	0.0007
	HQ	0.0223	0.877	0.1007	0	0.9293	0.0707	0	0.955	0.045
4	JT	0.0341	0.5355	0.4304	0.0001	0.6499	0.35	0	0.698	0.302
	JT*	0.1265	0.6494	0.2241	0.0013	0.8084	0.1903	0	0.8401	0.1599
	AIC	0.0003	0.5193	0.4804	0	0.5909	0.4091	0	0.6261	0.3739
	SIC	0.5377	0.4604	0.0019	0.0494	0.9482	0.0024	0	0.9989	0.0011
	HQ	0.0528	0.8446	0.1026	0	0.9219	0.0781	0	0.9485	0.0515
5	JT	0.0452	0.4993	0.4555	0.0004	0.6422	0.3574	0	0.7021	0.2979
	JT*	0.1391	0.6172	0.2437	0.0064	0.7992	0.1944	0	0.8402	0.1598
	AIC	0.0017	0.5003	0.498	0	0.5819	0.4181	0	0.6296	0.3704
	SIC	0.5713	0.426	0.0027	0.1552	0.8431	0.0017	0	0.9989	0.0011
	HQ	0.0826	0.8104	0.107	0.0005	0.9174	0.0821	0	0.9534	0.0466
6	JT	0.0556	0.4827	0.4617	0.0029	0.6575	0.3396	0	0.7396	0.2604
	JT*	0.1584	0.6001	0.2415	0.0238	0.7996	0.1766	0	0.8675	0.1325
	AIC	0.0048	0.5031	0.4921	0	0.598	0.402	0	0.6699	0.3301
	SIC	0.6008	0.3966	0.0026	0.3185	0.6805	0.001	0	0.9996	0.0004
	HQ	0.1115	0.7887	0.0998	0.0048	0.9315	0.0637	0	0.966	0.034
7	JT	0.0565	0.4645	0.479	0.0116	0.676	0.3124	0	0.7883	0.2117
	JT*	0.1594	0.5912	0.2494	0.0557	0.7946	0.1497	0	0.9054	0.0946
	AIC	0.0057	0.4958	0.4985	0.0001	0.6343	0.3656	0	0.723	0.277
	SIC	0.5916	0.405	0.0034	0.4727	0.5267	0.0006	0.0004	0.9993	0.0003
	HQ	0.1176	0.7778	0.1046	0.019	0.9334	0.0476	0	0.9764	0.0236

Table 1 (Continued)

8	JT	0.0523	0.4399	0.5078	0.0245	0.6813	0.2942	0	0.822	0.178
	JT*	0.1476	0.57	0.2824	0.0967	0.7702	0.1331	0	0.9263	0.0737
	AIC	0.0051	0.4636	0.5313	0.0002	0.6544	0.3454	0	0.7636	0.2364
	SIC	0.5541	0.4421	0.0038	0.5779	0.4219	0.0002	0.0107	0.9893	0
	HQ	0.1122	0.7668	0.121	0.0504	0.9146	0.035	0	0.9869	0.0131
$f=1.6$			$T=100$			$T=200$			$T=500$	
p	$\hat{r}(p)$	U	C	O	U	C	O	U	C	O
0	JT	0	0.2372	0.7628	0	0.2852	0.7148	0	0.3322	0.6678
	JT*	0	0.3567	0.6433	0	0.4125	0.5875	0	0.4593	0.5407
	AIC	0	0.1836	0.8164	0	0.2248	0.7752	0	0.265	0.735
	SIC	0	0.7073	0.2927	0	0.8241	0.1759	0	0.9119	0.0881
	HQ	0	0.4247	0.5753	0	0.5312	0.4688	0	0.6354	0.3646
1	JT	0.0002	0.3831	0.6167	0	0.432	0.568	0	0.4777	0.5223
	JT*	0.0025	0.526	0.4715	0	0.5785	0.4215	0	0.619	0.381
	AIC	0	0.3208	0.6792	0	0.3637	0.6363	0	0.407	0.593
	SIC	0.0199	0.8856	0.0945	0	0.9335	0.0665	0	0.9695	0.0305
	HQ	0	0.6221	0.3779	0	0.7132	0.2868	0	0.7885	0.2115
2	JT	0.0075	0.4988	0.4937	0	0.5829	0.4171	0	0.6277	0.3723
	JT*	0.0354	0.6512	0.3134	0	0.7273	0.2727	0	0.7647	0.2353
	AIC	0	0.4424	0.5576	0	0.5064	0.4936	0	0.5483	0.4517
	SIC	0.2212	0.7591	0.0197	0.0001	0.9836	0.0163	0	0.9945	0.0055
	HQ	0.0026	0.7862	0.2112	0	0.8475	0.1525	0	0.9031	0.0969
3	JT	0.0155	0.5543	0.4302	0	0.6526	0.3474	0	0.7029	0.2971
	JT*	0.0824	0.6749	0.2427	0.0001	0.8013	0.1986	0	0.8371	0.1629
	AIC	0	0.5136	0.4864	0	0.5834	0.4166	0	0.6279	0.3721
	SIC	0.4204	0.5729	0.0067	0.0056	0.9882	0.0062	0	0.9984	0.0016
	HQ	0.0173	0.8537	0.129	0	0.9091	0.0909	0	0.9469	0.0531
4	JT	0.0316	0.5298	0.4386	0	0.6598	0.3402	0	0.7233	0.2767
	JT*	0.112	0.6561	0.2319	0.0009	0.8223	0.1768	0	0.8556	0.1444
	AIC	0.0004	0.511	0.4886	0	0.5992	0.4008	0	0.6488	0.3512
	SIC	0.51	0.4869	0.0031	0.0426	0.9549	0.0025	0	0.9991	0.0009
	HQ	0.0482	0.8407	0.1111	0.0001	0.9286	0.0713	0	0.9586	0.0414
5	JT	0.0405	0.5081	0.4514	0.0005	0.6709	0.3286	0	0.7414	0.2586
	JT*	0.1318	0.6302	0.238	0.0057	0.8261	0.1682	0	0.871	0.129
	AIC	0.0012	0.5084	0.4904	0	0.6088	0.3912	0	0.6694	0.3306
	SIC	0.557	0.4391	0.0039	0.1506	0.848	0.0014	0	0.9992	0.0008
	HQ	0.0771	0.8205	0.1024	0.0003	0.9307	0.069	0	0.965	0.035
6	JT	0.0508	0.4925	0.4567	0.0034	0.6808	0.3158	0	0.7745	0.2255
	JT*	0.1546	0.6092	0.2362	0.0259	0.8229	0.1512	0	0.8963	0.1037
	AIC	0.0047	0.5062	0.4891	0	0.6228	0.3872	0	0.7057	0.2943
	SIC	0.5902	0.4077	0.0021	0.3157	0.6834	0.0009	0	0.9999	0.0001
	HQ	0.1039	0.7961	0.1	0.0056	0.9426	0.0518	0	0.9749	0.0251

Table 1 (Continued)

7	JT	0.0539	0.4609	0.4852	0.0109	0.6912	0.2979	0	0.8068	0.1932
	JT*	0.1525	0.5872	0.2603	0.055	0.8073	0.1377	0	0.9191	0.0809
	AIC	0.0044	0.4852	0.5104	0.0001	0.6436	0.3563	0	0.7443	0.2557
	SIC	0.5756	0.4209	0.0035	0.4611	0.5384	0.0005	0.0004	0.9994	0.0002
	HQ	0.1114	0.78	0.1086	0.0194	0.9402	0.0404	0	0.9832	0.0168
8	JT	0.0507	0.4344	0.5149	0.0244	0.6855	0.2901	0	0.824	0.176
	JT*	0.1405	0.5723	0.2872	0.0935	0.7795	0.127	0	0.9317	0.0683
	AIC	0.0045	0.4576	0.5379	0.0002	0.6561	0.3437	0	0.7651	0.2349
	SIC	0.5425	0.4516	0.0059	0.5689	0.4308	0.0003	0.0106	0.9894	0
	HQ	0.1059	0.7664	0.1277	0.0484	0.9165	0.0351	0	0.9889	0.0111

Table 2

DGP : (35) with $g=1$ in Example 1; $\tilde{M}(p)=M_Z(p)$ as $q=1$ in $S_{ij}(p)$

$f=0.8$		$T=100$			$T=200$			$T=500$		
p	$\hat{r}(p)$	U	C	O	U	C	O	U	C	O
0	JT	0	0.1852	0.8148	0	0.231	0.769	0	0.2893	0.7107
	JT*	0	0.3166	0.6834	0	0.3755	0.6245	0	0.4444	0.5556
	AIC	0	0.0723	0.9277	0	0.0954	0.9046	0	0.1224	0.8776
	SIC	0	0.6085	0.3915	0	0.7655	0.2345	0	0.8897	0.1103
	HQ	0	0.2761	0.7239	0	0.388	0.612	0	0.5232	0.4768
1	JT	0.0005	0.5092	0.4903	0	0.6171	0.3829	0	0.6904	0.3096
	JT*	0.009	0.6737	0.3173	0	0.7639	0.2361	0	0.8116	0.1884
	AIC	0	0.2749	0.7251	0	0.3707	0.6293	0	0.4533	0.5467
	SIC	0.0168	0.9003	0.0829	0	0.959	0.041	0	0.9833	0.0167
	HQ	0	0.6322	0.3678	0	0.7673	0.2327	0	0.8562	0.1438
2	JT	0.0232	0.6798	0.297	0	0.7822	0.2178	0	0.83	0.17
	JT*	0.1009	0.7533	0.1458	0	0.895	0.105	0	0.9179	0.0821
	AIC	0	0.4531	0.5469	0	0.5566	0.4434	0	0.6265	0.3735
	SIC	0.2097	0.7802	0.0101	0	0.9948	0.0052	0	0.9994	0.0006
	HQ	0.0031	0.8309	0.166	0	0.9166	0.0834	0	0.9528	0.0472
3	JT	0.0516	0.6633	0.2851	0	0.7699	0.2301	0	0.8127	0.1873
	JT*	0.1774	0.6973	0.1253	0.0006	0.8924	0.107	0	0.9109	0.0891
	AIC	0.0001	0.4777	0.5222	0	0.5534	0.4466	0	0.5995	0.4005
	SIC	0.398	0.599	0.003	0.0047	0.9928	0.0025	0	0.9994	0.0006
	HQ	0.0168	0.8695	0.1137	0	0.9252	0.0748	0	0.9537	0.0463
4	JT	0.0631	0.5735	0.3634	0.0004	0.6855	0.3141	0	0.741	0.259
	JT*	0.1865	0.639	0.1745	0.0048	0.8383	0.1569	0	0.8683	0.1317
	AIC	0.0002	0.4121	0.5877	0	0.4619	0.5381	0	0.5089	0.4911
	SIC	0.4785	0.5191	0.0024	0.0422	0.9535	0.0043	0	0.9985	0.0015
	HQ	0.0383	0.8286	0.1331	0	0.894	0.106	0	0.9352	0.0648

Table 2 (Continued)

5	JT	0.0625	0.4968	0.4407	0.0024	0.6212	0.3764	0	0.6965	0.3035
	JT*	0.1713	0.6146	0.2141	0.0128	0.7872	0.2	0	0.8467	0.1533
	AIC	0.0009	0.348	0.6511	0	0.3977	0.6023	0	0.4645	0.5355
	SIC	0.492	0.5026	0.0054	0.1247	0.8713	0.004	0	0.9984	0.0016
	HQ	0.0535	0.7866	0.1599	0.0004	0.8746	0.125	0	0.9264	0.0736
6	JT	0.0623	0.463	0.4747	0.006	0.5955	0.3985	0	0.6968	0.3032
	JT*	0.1732	0.5783	0.2485	0.0307	0.755	0.2143	0	0.8483	0.1517
	AIC	0.0012	0.3158	0.683	0	0.3772	0.6228	0	0.4632	0.5368
	SIC	0.4903	0.5021	0.0076	0.2681	0.7298	0.0021	0	0.9987	0.0013
	HQ	0.0651	0.7503	0.1846	0.0043	0.866	0.1297	0	0.9331	0.0669
7	JT	0.0634	0.4337	0.5029	0.0126	0.5947	0.3927	0	0.7226	0.2774
	JT*	0.1565	0.56	0.2835	0.058	0.7423	0.1997	0	0.8717	0.1283
	AIC	0.0011	0.2849	0.714	0	0.3814	0.6186	0	0.4926	0.5074
	SIC	0.4659	0.5218	0.0123	0.3941	0.6035	0.0024	0.0005	0.9985	0.001
	HQ	0.064	0.7225	0.2135	0.0143	0.8717	0.114	0	0.9466	0.0534
8	JT	0.0565	0.3956	0.5479	0.0262	0.5899	0.3839	0	0.7592	0.2408
	JT*	0.1388	0.5388	0.3224	0.0971	0.7167	0.1862	0.0001	0.897	0.1029
	AIC	0.0007	0.2586	0.7407	0.0001	0.3943	0.6056	0	0.5313	0.4687
	SIC	0.413	0.5685	0.0185	0.4987	0.4998	0.0015	0.0075	0.9921	0.0004
	HQ	0.0558	0.6844	0.2598	0.0303	0.8702	0.0995	0	0.9612	0.0388
f=1.6			T=100			T=200			T=500	
p	$\hat{r}(p)$	U	C	O	U	C	O	U	C	O
0	JT	0	0.5573	0.4427	0	0.6572	0.3428	0	0.7095	0.2905
	JT*	0	0.7191	0.2809	0	0.7917	0.2083	0	0.8283	0.1717
	AIC	0	0.3346	0.6654	0	0.4235	0.5765	0	0.5024	0.4976
	SIC	0	0.921	0.079	0	0.9658	0.0342	0	0.9811	0.0189
	HQ	0	0.6824	0.3176	0	0.8044	0.1956	0	0.8701	0.1299
1	JT	0.0018	0.6352	0.363	0	0.6906	0.3094	0	0.7133	0.2867
	JT*	0.0164	0.7654	0.2182	0	0.7988	0.2012	0	0.8112	0.1888
	AIC	0	0.4161	0.5839	0	0.4733	0.5267	0	0.5124	0.4876
	SIC	0.0176	0.9322	0.0502	0	0.9655	0.0345	0	0.9812	0.0188
	HQ	0	0.7483	0.2517	0	0.8093	0.1907	0	0.8548	0.1452
2	JT	0.0329	0.6804	0.2867	0	0.7547	0.2453	0	0.7809	0.2191
	JT*	0.1256	0.7266	0.1478	0	0.8601	0.1399	0	0.8757	0.1243
	AIC	0	0.5039	0.4961	0	0.5515	0.4485	0	0.5862	0.4138
	SIC	0.192	0.7964	0.0116	0	0.9901	0.0099	0	0.9967	0.0033
	HQ	0.002	0.8324	0.1656	0	0.8833	0.1167	0	0.916	0.084
3	JT	0.0563	0.6667	0.277	0	0.7544	0.2456	0	0.7935	0.2065
	JT*	0.1804	0.6878	0.1318	0.0007	0.874	0.1253	0	0.8939	0.1061
	AIC	0	0.4992	0.5008	0	0.5418	0.4582	0	0.5795	0.4205
	SIC	0.3604	0.6356	0.004	0.0035	0.9916	0.0049	0	0.9984	0.0016
	HQ	0.0125	0.8645	0.123	0	0.9057	0.0943	0	0.9421	0.0579

Table 2 (Continued)

4	JT	0.0592	0.5824	0.3584	0.0004	0.6827	0.3169	0	0.7288	0.2712
	JT*	0.1799	0.6503	0.1698	0.003	0.8356	0.1614	0	0.8701	0.1299
	AIC	0.0002	0.4246	0.5752	0	0.4572	0.5428	0	0.5077	0.4923
	SIC	0.4497	0.5481	0.0022	0.0372	0.9591	0.0037	0	0.9987	0.0013
	HQ	0.033	0.8314	0.1356	0	0.8903	0.1097	0	0.941	0.059
5	JT	0.0563	0.5017	0.442	0.0018	0.6033	0.3949	0	0.6753	0.3247
	JT*	0.1604	0.6215	0.2181	0.009	0.784	0.207	0	0.8426	0.1574
	AIC	0.0009	0.3516	0.6475	0	0.3861	0.6139	0	0.4463	0.5537
	SIC	0.482	0.514	0.004	0.1187	0.8769	0.0044	0	0.9987	0.0013
	HQ	0.0473	0.7922	0.1605	0.0003	0.8726	0.1271	0	0.9312	0.0688
6	JT	0.0596	0.4519	0.4885	0.0041	0.565	0.4309	0	0.6649	0.3351
	JT*	0.1614	0.5844	0.2542	0.0225	0.7526	0.2249	0	0.8349	0.1651
	AIC	0.0008	0.304	0.6952	0	0.3572	0.6428	0	0.4389	0.5611
	SIC	0.4798	0.5126	0.0076	0.2611	0.7362	0.0027	0	0.999	0.001
	HQ	0.0622	0.7493	0.1885	0.0043	0.8679	0.1278	0	0.9364	0.0636
7	JT	0.0578	0.421	0.5212	0.0108	0.556	0.4332	0	0.6758	0.3242
	JT*	0.1441	0.5593	0.2966	0.0488	0.7278	0.2234	0	0.8475	0.1525
	AIC	0.001	0.2704	0.7286	0	0.3511	0.6489	0	0.4486	0.5514
	SIC	0.45	0.5371	0.0129	0.3773	0.6203	0.0024	0.0005	0.9988	0.0007
	HQ	0.0568	0.7202	0.223	0.013	0.8681	0.1189	0	0.9479	0.0521
8	JT	0.0535	0.3977	0.5488	0.0222	0.5696	0.4082	0	0.7144	0.2856
	JT*	0.13	0.5418	0.3282	0.0841	0.7133	0.2026	0	0.875	0.125
	AIC	0.0009	0.2475	0.7516	0	0.3625	0.6375	0	0.4797	0.5203
	SIC	0.3923	0.5876	0.0201	0.4855	0.5124	0.0021	0.0071	0.9925	0.0004
	HQ	0.0505	0.6745	0.275	0.0274	0.8647	0.1079	0	0.9594	0.0406

Table 3

DGP : (36) with $g=0$ in Example 2; $\tilde{M}(p) = M_{\Delta Z}(p)$ in $S_{ij}(p)$

$f=1.6$		$T=100$			$T=200$			$T=500$		
p	$\hat{r}(p)$	U	C	O	U	C	O	U	C	O
0	JT	0	0.3912	0.6088	0	0.4226	0.5774	0	0.4412	0.5588
	JT*	0	0.5408	0.4592	0	0.5669	0.4331	0	0.5881	0.4119
	AIC	0	0.5591	0.4409	0	0.5864	0.4136	0	0.608	0.392
	SIC	0	0.9242	0.0758	0	0.9505	0.0495	0	0.9772	0.0288
	HQ	0	0.7635	0.2365	0	0.822	0.178	0	0.8628	0.1372
1	JT	0	0.7626	0.2374	0	0.815	0.185	0	0.8503	0.1497
	JT*	0	0.8866	0.1134	0	0.922	0.078	0	0.943	0.057
	AIC	0	0.8986	0.1014	0	0.929	0.071	0	0.9496	0.0504
	SIC	0.0011	0.9972	0.0017	0	0.9997	0.0003	0	0.9999	0.0001
	HQ	0	0.9807	0.0193	0	0.9916	0.0084	0	0.9985	0.0015

Table 3 (Continued)

3	JT	0.0004	0.8603	0.1393	0	0.8891	0.1109	0	0.9047	0.0953
	JT*	0.0058	0.9521	0.0421	0	0.9699	0.0301	0	0.972	0.028
	AIC	0	0.9674	0.0326	0	0.9769	0.0231	0	0.9782	0.0218
	SIC	0.3228	0.6772	0	0	1	0	0	1	0
	HQ	0.0068	0.9917	0.0015	0	0.999	0.001	0	0.9997	0.0003
4	JT	0.017	0.8225	0.1605	0	0.8615	0.1385	0	0.8784	0.1216
	JT*	0.0803	0.8704	0.0493	0	0.9567	0.0433	0	0.9586	0.0414
	AIC	0.0051	0.9566	0.0383	0	0.9651	0.0349	0	0.9661	0.0339
	SIC	0.7369	0.2631	0	0.0077	0.9923	0	0	1	0
	HQ	0.1143	0.8842	0.0015	0	0.9985	0.0015	0	0.9986	0.0014
5	JT	0.1094	0.734	0.1566	0.0002	0.8522	0.1476	0	0.8667	0.1333
	JT*	0.2822	0.6706	0.0472	0.002	0.9468	0.0512	0	0.95	0.05
	AIC	0.0517	0.9111	0.0372	0	0.959	0.041	0	0.9597	0.0403
	SIC	0.9252	0.0748	0	0.1965	0.8035	0	0	0.9999	0.0001
	HQ	0.395	0.6035	0.0015	0.0049	0.9926	0.0025	0	0.9977	0.0023
6	JT	0.1997	0.6527	0.1476	0.0067	0.8481	0.1452	0	0.8674	0.1326
	JT*	0.4261	0.5301	0.0438	0.0276	0.9228	0.0496	0	0.9506	0.0494
	AIC	0.1172	0.8509	0.0319	0.0012	0.9575	0.0413	0	0.9582	0.0418
	SIC	0.9572	0.0427	0.0001	0.5648	0.4351	0.0001	0.0006	0.9993	0.0001
	HQ	0.5661	0.4325	0.0014	0.0546	0.9432	0.0022	0	0.9977	0.0023
7	JT	0.2218	0.6357	0.1425	0.0161	0.8561	0.1278	0	0.8811	0.1189
	JT*	0.438	0.5226	0.0394	0.0659	0.8935	0.0406	0	0.9589	0.0411
	AIC	0.1433	0.8276	0.0291	0.0046	0.9621	0.0333	0	0.9677	0.0323
	SIC	0.9451	0.0549	0	0.7391	0.2609	0	0.0036	0.9964	0
	HQ	0.0167	0.5595	0.4238	0.133	0.8654	0.0016	0	0.9988	0.0012
8	JT	0.2179	0.6407	0.1414	0.0309	0.8651	0.104	0	0.9085	0.0915
	JT*	0.424	0.5352	0.0408	0.1122	0.8585	0.0293	0	0.9751	0.0249
	AIC	0.1418	0.8251	0.0331	0.0102	0.9672	0.0226	0	0.9811	0.0189
	SIC	0.9299	0.0701	0	0.8196	0.1804	0	0.0218	0.9782	0
	HQ	0.5545	0.4446	0.0009	0.2102	0.7894	0.0004	0.0001	0.9994	0.0005
$f=2.4$			$T=100$			$T=200$			$T=500$	
p	$\hat{r}(p)$	U	C	O	U	C	O	U	C	O
0	JT	0	0.6121	0.3879	0	0.6548	0.3452	0	0.6816	0.3184
	JT*	0	0.7536	0.2464	0	0.7932	0.2068	0	0.8198	0.1802
	AIC	0	0.7774	0.2226	0	0.8153	0.1847	0	0.8386	0.1614
	SIC	0	0.9866	0.0134	0	0.9961	0.0039	0	0.9994	0.0006
	HQ	0	0.9251	0.0749	0	0.9569	0.0431	0	0.978	0.022
1	JT	0	0.8794	0.1206	0	0.9126	0.0874	0	0.9268	0.0732
	JT*	0	0.9659	0.0341	0	0.9756	0.0244	0	0.9833	0.0167
	AIC	0	0.9723	0.0277	0	0.98	0.02	0	0.9876	0.0124
	SIC	0.001	0.999	0	0	1	0	0	1	0
	HQ	0	0.9974	0.0026	0	0.9993	0.0007	0	1	0

Table 3 (Continued)

2	JT	0	0.8493	0.1507	0	0.8732	0.1268	0	0.8839	0.1161
	JT*	0.0003	0.948	0.0517	0	0.9568	0.0432	0	0.9595	0.0405
	AIC	0	0.9596	0.0404	0	0.9659	0.0341	0	0.9666	0.0334
	SIC	0.0709	0.9291	0	0	1	0	0	1	0
	HQ	0.0003	0.9967	0.003	0	0.9978	0.0022	0	0.9989	0.0011
3	JT	0.0022	0.8175	0.1803	0	0.8359	0.1641	0	0.8461	0.1539
	JT*	0.0133	0.9192	0.0675	0	0.9341	0.0659	0	0.942	0.058
	AIC	0.0003	0.9458	0.0539	0	0.9459	0.0541	0	0.9506	0.0494
	SIC	0.4211	0.5788	0.0001	0.0001	0.9999	0	0	0.9999	0.0001
	HQ	0.017	0.9783	0.0047	0	0.9958	0.0042	0	0.9968	0.0032
4	JT	0.0442	0.7816	0.1742	0	0.829	0.171	0	0.8489	0.1511
	JT*	0.1525	0.7869	0.0606	0.0003	0.9315	0.0682	0	0.9374	0.0626
	AIC	0.0146	0.9359	0.0495	0	0.9432	0.0568	0	0.9479	0.0521
	SIC	0.843	0.157	0	0.0508	0.9491	0.0001	0	0.9998	0.0002
	HQ	0.1987	0.7979	0.0034	0.0006	0.9947	0.0047	0	0.9954	0.0046
5	JT	0.1153	0.7299	0.1548	0.0001	0.8495	0.1504	0	0.8619	0.1381
	JT*	0.2835	0.663	0.0535	0.0038	0.9403	0.0559	0	0.9464	0.0536
	AIC	0.0493	0.9086	0.0421	0.0001	0.9539	0.046	0	0.9559	0.0441
	SIC	0.9063	0.0937	0	0.2703	0.7296	0.0001	0	0.9999	0.0001
	HQ	0.3726	0.6254	0.002	0.008	0.9878	0.0042	0	0.9973	0.0027
6	JT	0.1473	0.7271	0.1256	0.0031	0.8921	0.1048	0	0.9032	0.0968
	JT*	0.3421	0.6231	0.0348	0.0153	0.954	0.0307	0	0.9706	0.0294
	AIC	0.0741	0.8976	0.0283	0.0002	0.9755	0.0243	0	0.9775	0.0225
	SIC	0.9014	0.0986	0	0.4701	0.5298	0.0001	0	1	0
	HQ	0.4375	0.5615	0.001	0.0288	0.9701	0.0011	0	0.9995	0.0005
7	JT	0.1544	0.7342	0.1114	0.0067	0.9093	0.084	0	0.9295	0.0705
	JT*	0.3379	0.6336	0.0285	0.0308	0.9504	0.0188	0	0.9859	0.0141
	AIC	0.0816	0.8959	0.0225	0.0014	0.9837	0.0149	0	0.9891	0.0109
	SIC	0.8741	0.1259	0	0.5926	0.4074	0	0.0004	0.9996	0
	HQ	0.4315	0.5679	0.0006	0.0622	0.9375	0.0003	0	1	0
8	JT	0.146	0.7253	0.1287	0.0131	0.8987	0.0882	0	0.9254	0.0746
	JT*	0.3213	0.6418	0.0369	0.0584	0.9195	0.0221	0	0.9853	0.0147
	AIC	0.0838	0.8845	0.0317	0.0041	0.9769	0.019	0	0.9882	0.0118
	SIC	0.8505	0.1495	0	0.698	0.302	0	0.0077	0.9923	0
	HQ	0.4089	0.5896	0.0015	0.1198	0.8801	0.0001	0	1	0

Table 4

DGP : (36) with $g=1$ in Example 2; $\tilde{M}(p)=M_Z(p)$ as $q=1$ in $S_{ij}(p)$

$f=1.6$		$T=100$			$T=200$			$T=500$		
p	$\hat{r}(p)$	U	C	O	U	C	O	U	C	O
0	JT	0	0.0896	0.9104	0	0.0964	0.9036	0	0.0929	0.9071
	JT*	0	0.1929	0.8071	0	0.1962	0.8038	0	0.194	0.806
	AIC	0	0.1455	0.8545	0	0.1499	0.8501	0	0.1478	0.8522
	SIC	0	0.7299	0.2701	0	0.7871	0.2129	0	0.8706	0.1294
	HQ	0	0.3989	0.6011	0	0.4552	0.5448	0	0.5324	0.4676
1	JT	0	0.5343	0.4657	0	0.6149	0.3851	0	0.6806	0.3194
	JT*	0	0.7258	0.2742	0	0.7867	0.2133	0	0.8396	0.1604
	AIC	0	0.649	0.351	0	0.7182	0.2818	0	0.774	0.226
	SIC	0.0017	0.9854	0.0129	0	0.9966	0.0034	0	0.9996	0.0004
	HQ	0	0.8889	0.1111	0	0.9493	0.0507	0	0.9815	0.0185
2	JT	0.0001	0.8123	0.1876	0	0.883	0.117	0	0.9246	0.0754
	JT*	0.0019	0.9278	0.0703	0	0.9641	0.0359	0	0.9802	0.0198
	AIC	0	0.8806	0.1194	0	0.9351	0.0649	0	0.9605	0.0395
	SIC	0.0849	0.914	0.0011	0	0.9999	0.0001	0	1	0
	HQ	0.0004	0.9834	0.0162	0	0.9975	0.0025	0	0.9993	0.0007
3	JT	0.0114	0.877	0.1116	0	0.9379	0.0621	0	0.9578	0.0422
	JT*	0.0489	0.9144	0.0367	0	0.9827	0.0173	0	0.9884	0.0116
	AIC	0.0012	0.9295	0.0693	0	0.9657	0.0343	0	0.9766	0.0234
	SIC	0.3272	0.6725	0.0003	0.0002	0.9998	0	0	1	0
	HQ	0.0212	0.9721	0.0067	0	0.9992	0.0008	0	0.9993	0.0007
4	JT	0.0615	0.8328	0.1057	0.0003	0.9359	0.0638	0	0.9526	0.0474
	JT*	0.1785	0.7871	0.0344	0.0014	0.9783	0.0203	0	0.9844	0.0156
	AIC	0.0103	0.921	0.0687	0	0.9603	0.0397	0	0.9706	0.0294
	SIC	0.5972	0.4025	0.0003	0.0243	0.9757	0	0	1	0
	HQ	0.0955	0.898	0.0065	0.0007	0.9969	0.0024	0	0.9987	0.0013
5	JT	0.1393	0.7384	0.1223	0.0059	0.9179	0.0762	0	0.941	0.059
	JT*	0.3191	0.6392	0.0417	0.0142	0.9588	0.027	0	0.978	0.022
	AIC	0.0286	0.8879	0.0835	0.0012	0.9484	0.0504	0	0.9621	0.0379
	SIC	0.783	0.2168	0.0002	0.1407	0.8593	0	0.0001	0.9999	0
	HQ	0.2272	0.7642	0.0086	0.0107	0.9866	0.0027	0	0.9982	0.0018
6	JT	0.2097	0.6385	0.1518	0.0236	0.8882	0.0882	0	0.9331	0.0669
	JT*	0.3992	0.5448	0.056	0.0599	0.9053	0.0348	0.0001	0.9728	0.0271
	AIC	0.0569	0.8387	0.1044	0.0051	0.935	0.0599	0	0.955	0.045
	SIC	0.8361	0.1634	0.0005	0.3996	0.6004	0	0.0047	0.9953	0
	HQ	0.3564	0.6341	0.0095	0.0491	0.9475	0.0034	0.0001	0.9978	0.0021
7	JT	0.2118	0.6021	0.1861	0.0548	0.8441	0.1011	0.0004	0.9298	0.0698
	JT*	0.3917	0.5369	0.0714	0.1262	0.8389	0.0349	0.0007	0.975	0.0243
	AIC	0.0766	0.7972	0.1262	0.0148	0.9186	0.0666	0.0001	0.9557	0.0442
	SIC	0.823	0.1769	0.0001	0.5798	0.4202	0	0.0203	0.9796	0.0001
	HQ	0.3803	0.6051	0.0146	0.1206	0.8762	0.0032	0.0012	0.9968	0.002

Table 4 (Continued)

8	JT	0.196	0.578	0.226	0.0876	0.799	0.1134	0.0013	0.9339	0.0648
	JT*	0.3674	0.5431	0.0895	0.1834	0.7849	0.0317	0.0034	0.977	0.0196
	AIC	0.0794	0.7641	0.1565	0.0241	0.9045	0.0714	0.0002	0.9602	0.0396
	SIC	0.795	0.2044	0.0006	0.6752	0.3248	0	0.0619	0.9381	0
	HQ	0.3667	0.6119	0.0214	0.1927	0.8034	0.0039	0.0049	0.9942	0.0009
f=2.4		T=100			T=200			T=500		
p	$\hat{r}(p)$	U	C	O	U	C	O	U	C	O
0	JT	0	0.3329	0.6671	0	0.3691	0.6309	0	0.398	0.602
	JT*	0	0.5182	0.4818	0	0.5508	0.4492	0	0.5934	0.4066
	AIC	0	0.4568	0.5432	0	0.4878	0.5122	0	0.5254	0.4746
	SIC	0	0.9417	0.0583	0	0.9769	0.0231	0	0.9935	0.0065
	HQ	0	0.7512	0.2488	0	0.8329	0.1671	0	0.8959	0.1041
1	JT	0	0.8252	0.1748	0	0.8901	0.1099	0	0.9127	0.0828
	JT*	0	0.9353	0.0647	0	0.9683	0.0317	0	0.9764	0.0236
	AIC	0	0.8972	0.1028	0	0.943	0.057	0	0.9574	0.0426
	SIC	0.0022	0.9972	0.0006	0	1	0	0	1	0
	HQ	0	0.9861	0.0139	0	0.9979	0.0021	0	0.9996	0.0004
2	JT	0.0008	0.9223	0.0769	0	0.9464	0.0536	0	0.9539	0.0461
	JT*	0.0082	0.9697	0.0221	0	0.9849	0.0151	0	0.9846	0.0154
	AIC	0	0.9552	0.0448	0	0.968	0.032	0	0.9728	0.0272
	SIC	0.1006	0.8991	0.0003	0	1	0	0	1	0
	HQ	0.0011	0.9947	0.0042	0	0.999	0.001	0	0.9988	0.0012
3	JT	0.0177	0.9066	0.0757	0	0.9394	0.0606	0	0.9419	0.0581
	JT*	0.0744	0.9004	0.0252	0	0.9784	0.0216	0	0.9784	0.0216
	AIC	0.0014	0.9463	0.0523	0	0.9589	0.0411	0	0.9616	0.0384
	SIC	0.3586	0.6413	0.0001	0.0011	0.9989	0	0	1	0
	HQ	0.0246	0.9698	0.0056	0	0.9968	0.0032	0	0.9976	0.0024
4	JT	0.0823	0.8191	0.0986	0.0017	0.9228	0.0755	0	0.9332	0.0668
	JT*	0.2291	0.737	0.0339	0.0033	0.964	0.0327	0	0.9723	0.0277
	AIC	0.0108	0.9202	0.069	0.0001	0.945	0.0549	0	0.9543	0.0457
	SIC	0.6625	0.3373	0.0002	0.0396	0.9604	0	0	0.9999	0.0001
	HQ	0.1157	0.876	0.0083	0.0024	0.9921	0.0055	0	0.9968	0.0032
5	JT	0.1486	0.7286	0.1228	0.0068	0.9093	0.0839	0	0.933	0.067
	JT*	0.3234	0.6329	0.0437	0.0171	0.9499	0.033	0	0.9724	0.0276
	AIC	0.0295	0.8834	0.0871	0.0019	0.9383	0.0598	0	0.9546	0.0454
	SIC	0.7725	0.2272	0.0003	0.1808	0.8189	0.0003	0.0007	0.9992	0.0001
	HQ	0.2304	0.759	0.0106	0.0136	0.9816	0.0048	0	0.9972	0.0028
6	JT	0.1868	0.6708	0.1424	0.0207	0.8984	0.0809	0	0.946	0.054
	JT*	0.3652	0.5833	0.0515	0.0505	0.9241	0.0254	0.0001	0.9831	0.0168
	AIC	0.0525	0.8486	0.0989	0.0047	0.9419	0.0534	0	0.9676	0.0324
	SIC	0.7887	0.2111	0.0002	0.37	0.63	0	0.0048	0.9952	0
	HQ	0.3148	0.6749	0.0103	0.043	0.9547	0.0023	0.0001	0.999	0.0009

Table 4 (Continued)

7	JT	0.1839	0.6383	0.1778	0.038	0.863	0.099	0.0003	0.9357	0.064
	JT*	0.3479	0.5907	0.0614	0.0919	0.8801	0.028	0.0009	0.9842	0.0149
	AIC	0.0642	0.8106	0.1252	0.0114	0.9211	0.0675	0.0001	0.9643	0.0356
	SIC	0.7728	0.2272	0	0.5009	0.499	0.0001	0.0173	0.9827	0
	HQ	0.3257	0.6614	0.0129	0.0883	0.9079	0.0038	0.0014	0.9983	0.0003
8	JT	0.1656	0.5977	0.2367	0.0568	0.8026	0.1406	0.0008	0.8952	0.104
	JT*	0.3106	0.5915	0.0979	0.1297	0.8219	0.0484	0.003	0.9656	0.0314
	AIC	0.0606	0.7733	0.1661	0.0161	0.8801	0.1038	0	0.9288	0.0712
	SIC	0.7432	0.2557	0.0011	0.599	0.4008	0.0002	0.0476	0.9524	0
	HQ	0.3149	0.6624	0.0227	0.1455	0.8487	0.0058	0.0045	0.9943	0.0012

Table 5

DGP : (37) with $g=0$ in Example 3; $\tilde{M}(p)=M_{\Delta z}(p)$ in $S_{ij}(p)$

$f=0.8$		$T=100$			$T=200$			$T=500$		
p	$\hat{r}(p)$	U	C	O	U	C	O	U	C	O
0	JT	0.0002	0.5676	0.4322	0	0.5236	0.4764	0	0.493	0.507
	JT*	0.0026	0.7717	0.2257	0	0.7147	0.2853	0	0.683	0.317
	AIC	0	0.4769	0.5231	0	0.4432	0.5568	0	0.4121	0.5879
	SIC	0.0441	0.9457	0.0102	0	0.9895	0.0105	0	0.9922	0.0078
	HQ	0	0.8548	0.1452	0	0.8499	0.1501	0	0.867	0.133
1	JT	0	0.7865	0.2135	0	0.7922	0.2078	0	0.7807	0.2193
	JT*	0	0.9148	0.0852	0	0.9096	0.0904	0	0.9057	0.0943
	AIC	0	0.7027	0.2973	0	0.7052	0.2948	0	0.696	0.304
	SIC	0	0.9969	0.0031	0	0.9983	0.0017	0	0.9993	0.0007
	HQ	0	0.9462	0.0538	0	0.964	0.036	0	0.9733	0.0267
2	JT	0.0006	0.8763	0.1231	0	0.9004	0.0996	0	0.9062	0.0938
	JT*	0.0063	0.9572	0.0365	0	0.9725	0.0275	0	0.9737	0.0263
	AIC	0	0.8215	0.1785	0	0.8472	0.1528	0	0.8505	0.1495
	SIC	0.0241	0.9754	0.0005	0	0.9998	0.0002	0	1	0
	HQ	0	0.9838	0.0162	0	0.9937	0.0063	0	0.9965	0.0035
3^{+}	JT	0.0159	0.8686	0.1155	0	0.9181	0.0819	0	0.9341	0.0659
	JT*	0.0892	0.8789	0.0319	0	0.9807	0.0193	0	0.986	0.014
	AIC	0.0001	0.8346	0.1653	0	0.8775	0.1225	0	0.8944	0.1056
	SIC	0.2787	0.7211	0.0002	0.0004	0.9995	0.0001	0	1	0
	HQ	0.0071	0.9805	0.0124	0	0.9965	0.0035	0	0.9986	0.0014
4	JT	0.107	0.7538	0.1392	0	0.9133	0.0867	0	0.9324	0.0676
	JT*	0.2984	0.6615	0.0401	0.0008	0.9747	0.0245	0	0.9842	0.0158
	AIC	0.0017	0.8067	0.1916	0	0.8651	0.1349	0	0.893	0.107
	SIC	0.6717	0.3283	0	0.0349	0.965	0.0001	0	1	0
	HQ	0.0983	0.8893	0.0124	0	0.9958	0.0042	0	0.9984	0.0016

Table 5 (Continued)

5	JT	0.1855	0.6624	0.1521	0.0025	0.9019	0.0956	0	0.9296	0.0704
	JT*	0.4206	0.5383	0.0411	0.02	0.9548	0.0252	0	0.9825	0.0175
	AIC	0.0113	0.7862	0.2025	0	0.8533	0.1467	0	0.8877	0.1123
	SIC	0.8116	0.1884	0	0.2423	0.7577	0	0	1	0
	HQ	0.2253	0.7632	0.0115	0.0016	0.9939	0.0045	0	0.999	0.001
6	JT	0.2391	0.5893	0.1716	0.0152	0.8809	0.1039	0	0.9266	0.0734
	JT*	0.4869	0.465	0.0481	0.0797	0.8925	0.0278	0	0.9812	0.0188
	AIC	0.0228	0.7593	0.2179	0	0.8449	0.1551	0	0.8839	0.1161
	SIC	0.8651	0.1349	0	0.5115	0.4885	0	0	1	0
	HQ	0.3212	0.6659	0.0129	0.0161	0.9791	0.0048	0	0.9985	0.0015
7	JT	0.2488	0.5495	0.2017	0.0389	0.8472	0.1139	0	0.9222	0.0778
	JT*	0.4841	0.4588	0.0571	0.1624	0.8058	0.0318	0	0.9818	0.0182
	AIC	0.0316	0.7284	0.24	0.0003	0.8354	0.1643	0	0.8789	0.1211
	SIC	0.8748	0.1251	0.0001	0.6934	0.3066	0	0.0003	0.9997	0
	HQ	0.3614	0.6225	0.0161	0.0509	0.9451	0.004	0	0.9982	0.0018
8	JT	0.2352	0.5364	0.2284	0.0738	0.8072	0.119	0	0.9203	0.0797
	JT*	0.4625	0.4601	0.0774	0.2428	0.7253	0.0319	0	0.981	0.019
	AIC	0.0331	0.7017	0.2652	0.0004	0.8279	0.1717	0	0.8785	0.1215
	SIC	0.8503	0.1496	0.0001	0.7974	0.2026	0	0.0069	0.9931	0
	HQ	0.357	0.6236	0.0194	0.1139	0.8823	0.0038	0	0.9985	0.0015
$f=1.6$			$T=100$			$T=200$			$T=500$	
p	$\hat{r}(p)$	U	C	O	U	C	O	U	C	O
0	JT	0	0.6388	0.3612	0	0.6033	0.3967	0	0.5695	0.4305
	JT*	0	0.8274	0.1726	0	0.7812	0.2188	0	0.7512	0.2488
	AIC	0	0.547	0.453	0	0.5151	0.4849	0	0.4801	0.5199
	SIC	0	0.993	0.007	0	0.9944	0.0056	0	0.9946	0.0054
	HQ	0	0.8887	0.1113	0	0.8907	0.1093	0	0.9009	0.0991
1	JT	0	0.791	0.209	0	0.7985	0.2015	0	0.7882	0.2118
	JT*	0	0.9118	0.0882	0	0.912	0.088	0	0.9069	0.0931
	AIC	0	0.7089	0.2911	0	0.7104	0.2896	0	0.7023	0.2977
	SIC	0	0.9974	0.0026	0	0.9986	0.0014	0	0.9993	0.0007
	HQ	0	0.9429	0.0571	0	0.9655	0.0345	0	0.973	0.027
2	JT	0	0.8746	0.1254	0	0.8995	0.1005	0	0.9075	0.0925
	JT*	0.0007	0.962	0.0373	0	0.9714	0.0286	0	0.9737	0.0263
	AIC	0	0.8207	0.1793	0	0.8471	0.1529	0	0.8507	0.1493
	SIC	0.004	0.9958	0.0002	0	0.9999	0.0001	0	1	0
	HQ	0	0.9835	0.0165	0	0.9935	0.0065	0	0.9962	0.0038
3^{+}	JT	0.0053	0.8804	0.1143	0	0.919	0.081	0	0.9349	0.0651
	JT*	0.0377	0.9324	0.0299	0	0.9788	0.0212	0	0.9858	0.0142
	AIC	0	0.8289	0.1711	0	0.8772	0.1228	0	0.8944	0.1056
	SIC	0.1398	0.8601	0.0001	0	0.9998	0.0002	0	1	0
	HQ	0.0016	0.986	0.0124	0	0.9964	0.0036	0	0.9986	0.0014

Table 5 (Continued)

4	JT	0.0469	0.8131	0.14	0	0.9157	0.0843	0	0.9328	0.0672
	JT*	0.1703	0.7878	0.0419	0.0005	0.9747	0.0248	0	0.9838	0.0162
	AIC	0.0004	0.8084	0.1912	0	0.8689	0.1311	0	0.8945	0.1055
	SIC	0.4523	0.5477	0	0.0103	0.9896	0.0001	0	1	0
	HQ	0.037	0.9506	0.0124	0	0.9959	0.0041	0	0.9987	0.0013
5	JT	0.089	0.7538	0.1572	0.0007	0.9055	0.0938	0	0.9296	0.0704
	JT*	0.2426	0.714	0.0434	0.0075	0.9676	0.0249	0	0.982	0.018
	AIC	0.0028	0.7909	0.2063	0	0.8566	0.1434	0	0.8886	0.1114
	SIC	0.5699	0.43	0.0001	0.1006	0.8994	0	0	1	0
	HQ	0.0949	0.8912	0.0139	0.0005	0.9947	0.0048	0	0.9992	0.0008
6	JT	0.1026	0.7155	0.1819	0.0047	0.8935	0.1018	0	0.9261	0.0739
	JT*	0.2528	0.6899	0.0573	0.0297	0.9437	0.0266	0	0.9813	0.0187
	AIC	0.0063	0.7588	0.2349	0	0.8475	0.1525	0	0.8839	0.1161
	SIC	0.5761	0.4238	0.0001	0.251	0.749	0	0	1	0
	HQ	0.1221	0.8598	0.0181	0.0032	0.992	0.0048	0	0.9987	0.0013
7	JT	0.091	0.69	0.219	0.0118	0.8781	0.1101	0	0.9245	0.0755
	JT*	0.2227	0.7011	0.0762	0.0555	0.9136	0.0309	0	0.9816	0.0184
	AIC	0.0061	0.7201	0.2738	0	0.8376	0.1624	0	0.8807	0.1193
	SIC	0.5434	0.4563	0.0003	0.3547	0.6453	0	0	1	0
	HQ	0.115	0.8583	0.0267	0.0125	0.9826	0.0049	0	0.9985	0.0015
8	JT	0.0715	0.6701	0.2584	0.0169	0.8643	0.1188	0	0.9207	0.0793
	JT*	0.1783	0.7213	0.1004	0.075	0.8937	0.0313	0	0.9825	0.0175
	AIC	0.0056	0.6883	0.3061	0	0.8305	0.1695	0	0.8777	0.1223
	SIC	0.4776	0.5215	0.0009	0.4121	0.5879	0	0.0007	0.9993	0
	HQ	0.0965	0.8658	0.0377	0.024	0.9703	0.0057	0	0.9986	0.0014

Table 6 Relative frequency Distribution for $\hat{r}(p)$ under Example 3

DGP : (37) with $g=1$ in Example 3; $\tilde{M}(p)=M_Z(p)$ as $q=1$ in $S_{ij}(p)$

$f=0.8$		$T=100$			$T=200$			$T=500$		
p	$\hat{r}(p)$	U	C	O	U	C	O	U	C	O
0	JT	0.0015	0.5397	0.4588	0	0.45	0.55	0	0.3871	0.6129
	JT*	0.0088	0.7517	0.2395	0	0.6583	0.3417	0	0.5933	0.4067
	AIC	0	0.2741	0.7259	0	0.2158	0.7842	0	0.1745	0.8255
	SIC	0.0382	0.9325	0.0293	0	0.9646	0.0354	0	0.9764	0.0236
	HQ	0	0.7267	0.2733	0	0.7038	0.2962	0	0.7237	0.2763
1	JT	0	0.7199	0.2801	0	0.7152	0.2848	0	0.7161	0.2839
	JT*	0	0.882	0.118	0	0.8706	0.1294	0	0.8716	0.1284
	AIC	0	0.4461	0.5539	0	0.4425	0.5575	0	0.4308	0.5692
	SIC	0	0.9881	0.0119	0	0.9951	0.0049	0	0.9989	0.0011
	HQ	0	0.8503	0.1497	0	0.8836	0.1164	0	0.9228	0.0772

Table 6 (Continued)

2	JT	0.0091	0.8398	0.1511	0	0.8739	0.1261	0	0.8919	0.1081
	JT*	0.0396	0.9161	0.0443	0	0.9625	0.0375	0	0.9722	0.0278
	AIC	0	0.6188	0.3812	0	0.6507	0.3493	0	0.6736	0.3264
	SIC	0.0698	0.9292	0.001	0	0.9998	0.0002	0	1	0
	HQ	0.001	0.9465	0.0525	0	0.9748	0.0252	0	0.9911	0.0089
3^{+}	JT	0.0861	0.7904	0.1235	0	0.9095	0.0905	0	0.9333	0.0667
	JT*	0.2385	0.7238	0.0377	0.0006	0.975	0.0244	0	0.9853	0.0147
	AIC	0.0008	0.6546	0.3446	0	0.7124	0.2876	0	0.7559	0.2441
	SIC	0.3625	0.6371	0.0004	0.0028	0.9972	0	0	1	0
	HQ	0.0283	0.9347	0.037	0	0.9844	0.0156	0	0.9961	0.0039
4	JT	0.1943	0.6716	0.1341	0.0044	0.894	0.1016	0	0.9296	0.0704
	JT*	0.407	0.5561	0.0369	0.0198	0.954	0.0262	0	0.9835	0.0165
	AIC	0.003	0.6308	0.3662	0	0.6941	0.3059	0	0.7456	0.2544
	SIC	0.609	0.3906	0.0004	0.0905	0.9094	0.0001	0	1	0
	HQ	0.1007	0.8591	0.0402	0.0007	0.9813	0.018	0	0.9964	0.0036
5	JT	0.2288	0.6125	0.1587	0.0223	0.8711	0.1066	0	0.9279	0.0721
	JT*	0.4493	0.5047	0.046	0.0918	0.8778	0.0304	0	0.9816	0.0184
	AIC	0.0054	0.595	0.3996	0	0.68	0.32	0	0.7366	0.2634
	SIC	0.6769	0.323	0.0001	0.326	0.6739	0.0001	0	1	0
	HQ	0.1476	0.803	0.0494	0.0066	0.9746	0.0188	0	0.9954	0.0046
6	JT	0.228	0.5903	0.1817	0.0623	0.8249	0.1128	0	0.9245	0.0755
	JT*	0.4337	0.5051	0.0612	0.1952	0.7744	0.0304	0	0.9812	0.0188
	AIC	0.0074	0.5534	0.4392	0.0002	0.6622	0.3376	0	0.7386	0.2614
	SIC	0.6795	0.3195	0.001	0.5446	0.4554	0	0.0001	0.9999	0
	HQ	0.1597	0.7801	0.0602	0.0336	0.9471	0.0193	0	0.9946	0.0054
7	JT	0.1852	0.5842	0.2306	0.1056	0.7778	0.1166	0	0.9224	0.0776
	JT*	0.3729	0.54	0.0871	0.2838	0.6844	0.0318	0	0.98	0.02
	AIC	0.0069	0.4989	0.4942	0.0006	0.6546	0.3448	0	0.7317	0.2683
	SIC	0.6201	0.3781	0.0018	0.6499	0.35	0.0001	0.0044	0.9956	0
	HQ	0.1402	0.7798	0.08	0.0744	0.9078	0.0178	0	0.9952	0.0048
8	JT	0.1341	0.5709	0.295	0.1413	0.7357	0.123	0	0.9217	0.0783
	JT*	0.2797	0.5998	0.1205	0.322	0.6404	0.0376	0.0008	0.9797	0.0195
	AIC	0.0045	0.4338	0.5617	0.0014	0.6488	0.3498	0	0.7351	0.2649
	SIC	0.5253	0.4698	0.0049	0.6786	0.3214	0	0.04	0.96	0
	HQ	0.1028	0.7818	0.1154	0.1143	0.8661	0.0196	0	0.9952	0.0048
$f=1.6$			$T=100$			$T=200$			$T=500$	
p	$\hat{r}(p)$	U	C	O	U	C	O	U	C	O
0	JT	0	0.5931	0.4069	0	0.4995	0.5005	0	0.4351	0.5649
	JT*	0	0.7954	0.2046	0	0.7072	0.2928	0	0.6477	0.3523
	AIC	0	0.3211	0.6789	0	0.2521	0.7479	0	0.2085	0.7915
	SIC	0	0.9822	0.0178	0	0.9828	0.0172	0	0.988	0.012
	HQ	0	0.7763	0.2237	0	0.7618	0.2382	0	0.7808	0.2192

Table 6 (Continued)

1	JT	0	0.7075	0.2925	0	0.7022	0.2978	0	0.702	0.298
	JT*	0	0.8693	0.1307	0	0.8631	0.1369	0	0.8642	0.1358
	AIC	0	0.4384	0.5616	0	0.4327	0.5673	0	0.4221	0.5779
	SIC	0	0.9874	0.0126	0	0.9943	0.0057	0	0.9986	0.0014
	HQ	0	0.8456	0.1544	0	0.8799	0.1201	0	0.9192	0.0808
2	JT	0.0007	0.8336	0.1657	0	0.865	0.135	0	0.8854	0.1146
	JT*	0.005	0.9443	0.0507	0	0.9575	0.0425	0	0.9685	0.0315
	AIC	0	0.6031	0.3969	0	0.6374	0.3626	0	0.6635	0.3365
	SIC	0.0056	0.993	0.0014	0	0.9998	0.0002	0	1	0
	HQ	0.0001	0.9436	0.0563	0	0.9716	0.0284	0	0.9906	0.0094
3^{+}	JT	0.0317	0.8408	0.1275	0	0.9087	0.0913	0	0.9338	0.0662
	JT*	0.1195	0.8424	0.0381	0	0.9752	0.0248	0	0.9847	0.0153
	AIC	0.0001	0.6526	0.3473	0	0.712	0.288	0	0.754	0.246
	SIC	0.1901	0.8094	0.0005	0.0001	0.9998	0.0001	0	1	0
	HQ	0.0051	0.9568	0.0381	0	0.9856	0.0144	0	0.9962	0.0038
4	JT	0.1263	0.7355	0.1382	0.001	0.8995	0.0995	0	0.9317	0.0683
	JT*	0.3013	0.6571	0.0416	0.0069	0.9678	0.0253	0	0.983	0.017
	AIC	0.0012	0.6359	0.3629	0	0.7005	0.2995	0	0.7474	0.2526
	SIC	0.4712	0.5284	0.0004	0.028	0.9718	0.0002	0	1	0
	HQ	0.0514	0.906	0.0426	0	0.9845	0.0155	0	0.9959	0.0041
5	JT	0.1578	0.681	0.1612	0.0089	0.8852	0.1059	0	0.9285	0.0715
	JT*	0.3414	0.6083	0.0503	0.0423	0.9278	0.0299	0	0.9818	0.0182
	AIC	0.0022	0.5977	0.4001	0	0.6849	0.3151	0	0.7383	0.2617
	SIC	0.5289	0.4705	0.0006	0.1812	0.8187	0.0001	0	1	0
	HQ	0.0916	0.8535	0.0549	0.0022	0.9772	0.0206	0	0.9955	0.0045
6	JT	0.1411	0.66	0.1989	0.0268	0.8589	0.1143	0	0.9255	0.0745
	JT*	0.2988	0.6271	0.0741	0.1072	0.8585	0.0343	0	0.9816	0.0184
	AIC	0.0039	0.548	0.4481	0	0.6632	0.3368	0	0.7404	0.2596
	SIC	0.499	0.4989	0.0021	0.3317	0.6683	0	0	1	0
	HQ	0.089	0.8393	0.0717	0.011	0.9683	0.0207	0	0.9945	0.0055
7	JT	0.1017	0.6331	0.2652	0.0406	0.8331	0.1263	0	0.9237	0.0763
	JT*	0.2236	0.6699	0.1065	0.1414	0.8199	0.0387	0	0.9801	0.0199
	AIC	0.0019	0.472	0.5261	0.0002	0.648	0.3518	0	0.7306	0.2694
	SIC	0.4224	0.5747	0.0029	0.3891	0.6109	0	0.0005	0.9995	0
	HQ	0.0677	0.8297	0.1026	0.0233	0.9548	0.0219	0	0.9954	0.0046
8	JT	0.0664	0.5926	0.341	0.0546	0.8035	0.1419	0	0.9207	0.0793
	JT*	0.1541	0.6908	0.1551	0.1541	0.8005	0.0454	0.0003	0.9798	0.0199
	AIC	0.0017	0.4049	0.5934	0.0001	0.6281	0.3718	0	0.732	0.268
	SIC	0.3344	0.6591	0.0065	0.4111	0.5889	0	0.0072	0.9928	0
	HQ	0.0418	0.8076	0.1506	0.0344	0.9398	0.0258	0	0.9954	0.0046