TESTING FOR THE EXISTENCE OF AN $I(1)$ VECM REPRESENTATION

MITSUHIRO ODAKI

Abstract

This paper proposes a method to test whether or not a VECM representation for a vector time series consisting of $I(1)$ components, claimed in Granger representation theorem, is derived validly from a VMA as the DGP. The test is also available for the detection of the occurrence of multicointegration etc. Utilizing the idea of the nonparametric test employed by Phillips and Oulialis (1990) and Shintani (2001) to detect the existence of cointegration or the cointegrating rank, we construct a test statistic and establish its limiting properties with some similarity to of those tests. Finite sample performances of the test proposed are also investigated through some Monte Carlo experiment.

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. Particularly, numerous econometric researches have been concentrated on the situation in which individual time series are integrated of order 1 ($I(1)$) and also paid attention to the occurrence cointegration introduced by Granger (1981). *The Granger representation theorem (GRT)* presented in Engle and Granger (1987) is to provide a VAR form coping with such a situation called the VECM representation. However, there are some cointegrated systems of $I(1)$ components resulting in the failure in the valid derivation of the VECM representation by GRT.

Provided that the DGP is a VMA representation in first differences of original data series, the derivation of an $I(1)$ VECM by GRT becomes invalid by the existence of some hidden unit roots in the VMA characteristic equation (see Gregoir and Laroque (1993) e.g.), and it is appeared as the phenomenon that a linear combination either between a linear combination of $I(1)$ series forming a cointegrating relation in the ordinary sense and a first differenced $I(1)$ series or a linear combination between $I(1)$ series itself is weakly stationary but overdifferenced (i.e., $I(-j)$ with a positive integer j), as stated by Lemma 3 in the subsequent section. The former as a relation between series of different orders of integration is associated with the occurrence of *multicointegration* defined and discussed by the literatures such as Granger and Lee (1990) and Engle and Yoo (1991), whereas the latter is interpreted as one included in relations on *higher order cointegration*, in the sense that it exactly corresponds to the case $d=1$ and $b>1$ in Engle and Granger's (1987) definition of cointegration.

As seen in many literatures, such a situation compels the VECM representation/formulation to be far apart from the conventional one. The VECM formulation for time series systems of $I(2)$ components or a condition for the level series forming such a VECM representation to be $I(2)$ with some inferences based on it has been studied vigorously in many literatures such as Granger and Lee (1990), Johansen (1995, p. 57 and p. 132) and Paruolo (1996), and it is pointed out that an $I(2)$ VECM representation is characterized by the error correction mechanism associated with relations on multicointegration or higher order cointegration, although it is limited to a particular one such as $b=2$. On the other hand, Gregoir and Laroque (1993) set up a representation theorem as an extension of GRT under a more general framework without excluding the occurrence of hidden unit roots, any type of multicointegration or higher order cointegration and provided a representation not only in levels and differences but also in their integrated one. In those literatures, we may observe that the generalization of GRT or multicointegration has been discussed under the formulation for $I(d)$ components with a positive integer d greater than 1, particularly for $I(2)$ components. However, it will be noticed in this paper that the analysis of vector time series systems whose individual components are $I(d)$ with an integer d greater than 1 can be converted into that of systems consisting of $I(1)$ components considering the $d-1$-th difference series as the level ones.

The purpose of this paper is to provide a meaningful testing method to determine whether or not the conventional GRT holds, motivated by that it has not checked in many empirical researches for $I(1)$ cointegrated systems. The detection of multicointegration or higher order cointegration will be also aimed at through our test. We therefore adopt the vector time series system of $I(1)$ components as the usual framework in which the conventional GRT is discussed. It will be emphasized that a necessary and sufficient condition for GRT to hold is used for formulating the null hypothesis. Adopting a specific VECM such as in Engsted et al. (1997) is not suitable for our purpose since we are not interested in what a sort of specific relation of multicointegration exists.

The test constructed here is motivated by a nonparametric approach introduced in Phillips and Oulialis (1990) to test the existence of cointegration and then led by Shintani (2001) to testing for determining the cointegrating rank. The essential idea leading to the characteristics of the tests is that some sample covariance matrices or sample long-run covariance matrices of $I(1)$ or $I(0)$ series become degenerate or singular if and only if the null hypotheses are not true. Our test will also utilize the idea, exploiting similar matrices. These matrices are interpreted as the sample covariances or sample long-run covariances of $I(1)$ or $I(0)$ series constructed based on the cointegrating rank and cointegrating matrix and exhibit the degeneracy property exclusively for the alternative hypothesis formulating the situation in which GRT does not hold. It is shown that the test is appreciated by the normalization-free property, limiting distributions in terms of standard Brownian motion and forming consistent tests, as ones in the above-mentioned papers. It is also pointed out that the test requires to obtain a consistent estimator of the cointegrating rank and that this can be actualized by Shintani's (2001) test mentioned above even if GRT does not hold. Moreover, it should be noted that one of Kernel estimators (or lag-windows) with a related band-width parameter is required to estimate the sample long-run covariance matrix of $I(0)$ series, similar to in the papers above. Monte Carlo experiments are executed under some particular DGPs and sample sizes 100, 250 and 500 in order to investigate finite sample performances of the test. Generally, the experimental results reveal that a sample size as many as 250 is needed to secure the asymptotics established theoretically for the test with a careful selection from band-width parameters or significance levels. On the other hand, the sample size of 500 was sufficient for the test to approximate the asymptotics and consequently the performances of the test were satisfactory through all the experiments.

The paper is organized as follows. Section 2 formulates the DGP, suppositions and some preliminary concepts. The test proposed and its related results are presented in Section 3. Section 4 deals with Monte Carlo experiments. The remaining issues including some concluding remarks are discussed in Section 5. The proofs of a lemma and theorem in the text are in Appendix.

2 DGP AND VECM FOR $I(1)$ SERIES

Associated with the Wold representation, suppose that the observable vector time series y_t of k-variates considered is generated by a VMA in the first differences[1]:

$$\begin{aligned}\Delta y_t &= C(B)\,\epsilon_t + C(1)\mu \\ &= C(1)\,\epsilon_t + C^{(1)}(B)\,(1-B)\,\epsilon_t L + C(1)\mu, \qquad (1)\end{aligned}$$

where Δ and B are the difference and backward operators respectively, $C(z)$ and $C^{(1)}(z)$ are the power series $C(z)$ and $C^{(1)}(z)$ given as

$$C(z) = I_k + \sum_{j=1}^{\infty} C_j\, z^j, \quad C^{(1)}(z) = \sum_{j=0}^{\infty}\left(-\sum_{i=j+1}^{\infty} C_i\right) z^j,$$

with $k \times k$ constant matrices C_j such that $\Sigma_{j=1}^{\infty} j^{\bar{\nu}}$, $\|C_j\| < \infty$ for some real number $\bar{\nu} \geq 1$ the row vectors of $C(1)$ are all nonzero, $1 \geq rank\ C(1) \equiv s \geq k-1$ and the notice $C(z)=C(1)+(1-z)C^{(1)}(z)$, $\{\varepsilon_t\}$ is a sequence of unobservable $k \times 1$ random vectors which are iid with $E\varepsilon_t=0$, $E\varepsilon_t\varepsilon_t'=\Omega_{00}>0$ and finite fourth moments, and μ is a k-dimensional constant vector forming the deterministic trends with the supposition that in all the relations removing the stochastic ones $C(1)(\Sigma_{h=1}^{t}\varepsilon_h)$, those are removed similarly. Moreover, we assume that the VMA characteristic equation det $C(z)=0$ has roots either equal to 1 or strictly greater than 1 in absolute value and that $(y_0', y_{-1}', \ldots y_{-q+1}')'$ is either $O(1)$ or $O_p(1)$, with a positive integer q. Note that the former is imposed to exclude noninvertibility caused by a root other than $z=1$ and the latter is put as some initial condition, conventionally for cointegrated systems. Putting

$$\eta_0(1) = y_0 - \sum_{j=0}^{\infty}\left(-\sum_{i=j+1}^{\infty} C_i\right)\epsilon_{-j},$$

$$\eta_0(2) = -\sum_{j=0}^{\infty}\left(-\sum_{i=j+2}^{\infty}(i-1)C_i\right)\epsilon_{-j},$$

with $C^{(2)}(z)=\Sigma_{j=0}^{\infty}(\Sigma_{i=j+2}^{\infty}(i-1)C_i)z^j$, it follows from (1) that for any $t \geq 1$,

$$y_t = C(1)\,(\sum_{h=1}^{t}\epsilon_h) + C^{(1)}(B)\,\epsilon_t + tC(1)\,\mu + \eta_0(1), \qquad (2)$$

$$\begin{aligned}\sum_{h=1}^{t} y_h &= C(1)\,(\sum_{h=1}^{t}(t+1-h)\epsilon_h) + C^{(1)}(1)\,(\sum_{h=1}^{t}\epsilon_h) \\ &\quad + C^{(2)}(B)\,\epsilon_t + \{t(t-1)/2\}C(1)\,\mu + t\eta_0(1) \\ &\quad + \eta_0(2), \qquad (3)\end{aligned}$$

which will be useful for evaluating the initial condition, trends in both stochastic and deterministic sense or relations on cointegration or multicointegration, noting that even a quadratic trend must be considered in deterministic ones for $\Sigma_{h=1}^{t} y_h$.

It is now obvious from the supposition of $rank\ C(1)$ that there exist constant matrices γ of $k \times s$ and β of $k \times r$, with $r=k-s$, such that

$$rank\,\gamma' C(1)=s,\ \beta'\gamma=0,\ \beta' C(1)=0,\ rank\,\beta=r.$$

Without losing generality, suppose that if $\gamma' C(1)\mu \neq 0$, the first component of it is nonzero. As seen easily by (2), all the elements of y_t or all nonzero linear combinations of $\gamma' y_t$ are of $I(1)$ and $\beta' y_t$ becomes weakly stationary if $\eta_0(1)$ as the term on the initial vectors of y_t and ε_t is out of consideration, implying that it forms some coinetgrating relations with the cointegrating rank r and cointegrating matrix β. Letting

$$\Omega = \begin{bmatrix} \beta' C^{(1)}(1) \\ \gamma' C(1) \end{bmatrix} \Omega_{00} \left[C^{(1)}(1)'\beta,\ C(1)'\gamma\right],$$

Ω can be interpreted as the long-run covariance matrix of $(y_t'\beta, \Delta y_t'\gamma)'$, conditionally on the initial vectors, in view of (1) and (2).

Put

$$rank \begin{bmatrix} \beta' C^{(1)}(1) \\ \gamma' C(1) \end{bmatrix} = \bar{s}$$

with the notice $k \geq \bar{s} \geq s$. Now, we provide a necessary and sufficient condition for the VECM representation which GRT claims for $I(1)$ cointegrated systems to exist validly[2]:

Lemma 1 *For y_t generated by (1), there uniquely exists a representation such that*

$$\begin{aligned}\Delta y_t &= \alpha\,\beta' y_{t-1} + \sum_{j=1}^{\infty} H_j \Delta y_{t-j} \\ &\quad + \mu - \alpha\,\beta'\eta_0(1) + \epsilon_t, \qquad \forall\, t \geq 1, \qquad (4)\end{aligned}$$

where $\Delta = 1-B$ and the power series

$$A(z) = -z\,\alpha\beta' + (1-z)\left(I_k - \sum_{j=1}^{\infty} H_j\, z^j\right),$$

satisfies $A(z)C(z)=(1-z)I_k$, with a constant matrix α of $k \times r$ column full rank, $k \times k$ identity matrix I_k, $k \times k$ constant matrices H_j satisfying a convergence condition similar to one for C_j and a k-dimensional *constant vector* μ, *if and only if*

Condition (A) $\bar{s}=k$

is satisfied.

Note that the absence of Condition (A) is equivalent to the singularity of the long-run covariance matrix of $(y_t'\beta,\ \Delta'\gamma)$.[3] It should be also noted in virtue of $A(z)$, $C(z)=(1-z)I_k$ that det $A(z)=0$ has roots either equal to 1 or strictly greater than 1 in absolute value. We now show that the absence of Condition (A) may lead to some relations embodying either multicointegration or higher order cointegration.

Lemma 2 *Suppose that Condition A is not satisfied for y_t generated by* (1). *Then there exists a relation as either* $b_1'\beta'(\Sigma_{h=1}^{t}y_h)+b_2'\gamma' y_t$ *or* $b_1'\beta'(\Sigma_{h=1}^{t}\bar{y}_h)$, *with nonzero constant vectors* b_1 *of* $r\times 1$ *and* b_2 *of* $s\times 1$, *which is weakly stationary except for the terms on the initial vectors of* y_t *and* ε_t *or deterministic trends.*

The VECM for $I(1)$ will be naturally deprived of its validity by the existence of a VECM for $I(2)$, as stated already, and it will be formulated as follows:

Lemma 3 *Suppose that* y_t *is represented as*[4]

$$\Delta y_t=\alpha_1\beta_1'(\sum_{h=1}^{t-1}y_h)+\alpha_2\beta_2' y_{t-1}+\sum_{j=1}^{\infty}H_j\Delta y_{t-j}+\mu+\epsilon_t,\qquad \forall t\geq 1, \tag{5}$$

with suitable constant matrices, provided that α_i *and* β_i, $i=1,2$, *are column full rank, a constant vector* μ *and a suitable initial condition on* y_t. *Then, if* (1) *is derived from* (5), *Condition* (A) *is not satisfied.*

Put

$$A(z)=-z\,\alpha_1\beta_1'-(1-z)z\,\alpha_2\beta'+(1-z)^2\left(I_k-\sum_{j=1}^{\infty}H_j\,z^j\right).$$

The roots of det $A(z)=0$ also satisfies the same condition as for det $C(z)=0$ or det $A(z)=0$ in the Lemma 1, as checked easily in the proof of this lemma.

3 TESTING METHORD

In this section we shall first discuss a method to test whether or not the conventional VECM expressed as (4) is derived validly under the DGP (1), provided that the value of the cointegrating rank r is known. In virtue of Lemma 1, the hypothesis tested can be formulated by the value of $\bar{s}$ associated with Condition (A). This paper will naturally seek to test the null of the situation in which the VECM derivation is valid as

$H_0:\bar{s}=k$ (*Condition* (A) *holds*)
$H_1:\bar{s}<k$ (*Condition* (A) *does not hold*)

The construction of the test statistic proposed in this paper needs some consistent estimates on the cointegrating matrix β with the value of r.

We now propose consistent estimators for β and γ suitably to our situation, although estimating β and its related consistency property have been examined and established in numerous papers including famous Stock and Watson (1993) under several situations. Following the notations in Johansen (1995, pp. 90-91), put $Z_{0t}=\Delta y_t$, $Z_{1t}=y_{t-1}$ and $Z_{2t}=(\Delta y_{t-1}',\cdots,\Delta y_{t-p}',1)'$, with a positive integer p, and based on those,

$$M_{ij}=\sum_{t=2}^{T}Z_{it}Z_{jt}'/T,\quad i,j=0,1,$$

$$M_{i2}=\sum_{t=p+2}^{T}Z_{it}Z_{2t}'/T,\quad i=0,1,2,$$

with $M_{2i}=M_{i2}'$, which in turn construct

$$S_{ij}=M_{ij}-M_{i2}M_{22}^{-1}M_{2j},\quad i,j=0,1.$$

Letting $\hat{\lambda}_1\geq\cdots\geq\hat{\lambda}_k$ and $\hat{\xi}_1,\cdots,\hat{\xi}_k$ be the ordered eigenvalues of S_{11} and the corresponding eigenvectors, the matrices $\hat{\beta}$ of $k\times r$ and $\hat{\gamma}$ of $k\times s$ as

$$\hat{\beta}=[\hat{\xi}_{s+1},\cdots,\hat{xi}_k],\quad \hat{\gamma}=[\hat{\xi}_1,\cdots,\hat{xi}_s].$$

Lemma 4 *Let* D_T^{-1} *be the* $s\times s$ *matrix such that*

$$\begin{aligned} D_T^{-1}&=T^{-1/2}\quad if\ s=1\ and\ \gamma'C(1)\mu\neq 0,\\ &=\begin{bmatrix}T^{-1/2}&0\\0&I_{s-1}\end{bmatrix}\quad if\ s>1\ and\ \gamma'C(1)\mu\neq 0,\\ &=I_s\quad otherwise.\end{aligned}$$

Then, for $\hat{\beta}$ *and* $\hat{\gamma}$ *given above, there exist random* *matrices* $\hat{x}$, $\hat{y}$, $\hat{u}$ *and* $\hat{v}$ *such that*

$$\hat{\beta}=\beta\hat{x}+\gamma\hat{y},\quad \hat{\gamma}=\gamma\hat{u}+\beta\hat{v},$$

with $\hat{x}=O_p(1)$, $\hat{x}^{-1}=O_p(1)$, $D_T\hat{y}=O_p(T^{-1})$, $\hat{u}=O_p(1)$, $\hat{u}^{-1}=O_p(1)$ *and* $D_T\hat{v}=O_p(T^{-1})$.

Similarly to Z_{it} and M_{ij} above, we introduce the notations $Z_{0;t-n}=\Delta y_{t-n}$, $Z_{1;t-n}=y_{t-n}$, $Z_{3;t-n}=\Sigma_{h=1}^{t}y_h$, $Z_{4;t-n}=1$, $Z_{5;t-n}=(1, t)$ and $Z_{6;t-n}=(\Sigma_{h=1}^{t}y_h'\hat{\gamma}, 1, t, t^2)'$, $n=0, 1,\cdots, S_T$, with a sequence of positive integers $\{S_T\}$ such that $\lim_{T\to\infty} S_T=\infty$ and $\lim_{T\to\infty} S_T/T^{1/2}=0$, with the notations for the product moment matrices

$$M_{ij;m;n}=\sum_{t=n+2}^{T} Z_{i;t-m}Z'_{j;t-n}/T,\quad i,j=0,1,3,4,5,6,\quad m=0,n,\quad n=0,1,\cdots S_T.$$

Moreover, let

$$\hat{R}_{ij}(n)=M_{ij;0;n}-M_{i5;0;n}M^{-1}_{55;n;n}M_{5j;n;0},\quad i,j=0,1,\ i=j\neq 0,$$
$$\hat{R}_{00}(n)=M_{00;0;n}-M_{04;0;n}M^{-1}_{44;n;n}M_{41;n;0},$$
$$\tilde{R}_{ij}=M_{ij;0;0}-M_{i6;0;0}M^{-1}_{66;0;0}M_{6j;0;0},\quad i,j=1,3,$$

with $\hat{R}_{ij}(-n)=\hat{R}'_{ij}(n)$. Note that those are constructed so that deterministic terms in Δy_t, y_t or $\Sigma_{h=1}^{t}y_h$ are removed by detrending based on the least squares regression.[5]

Our test statistic is motivated by Shintani's (2001) $P^*(n, s)$ test statistic and is constructed simply by some kernel estimator of the long-run covariance matrix of $(\Sigma_{h=1}^{t}y_h'\hat{\beta}, y_t'\hat{\gamma})'$ (conditioned the initial vectors) and Ω. Those are given by

$$\hat{M}_{ZZ}=\sum_{n=-S_T+1}^{S_T+1}w(n/S_T)\begin{bmatrix}\hat{\beta}'\tilde{R}_{33}(n)\hat{\beta} & \hat{\beta}'\tilde{R}_{31}\hat{\gamma}\\ \hat{\gamma}'\tilde{R}_{13}(n)\hat{\beta} & \hat{\gamma}'\tilde{R}_{11}\hat{\gamma}\end{bmatrix},$$
$$\hat{\Omega}=\sum_{n=-S_T+1}^{S_T+1}w(n/S_T)\begin{bmatrix}\hat{\beta}'\tilde{R}_{11}(n)\hat{\beta} & \hat{\beta}'\tilde{R}_{10}(n)\hat{\gamma}\\ \hat{\gamma}'\tilde{R}_{01}(n)\hat{\beta} & \hat{\gamma}'\tilde{R}_{00}(n)\hat{\gamma}\end{bmatrix},$$

where $w(\cdot)$ is a real-valued kernel function defined by the following assumption:

Assumption KL (*Kernel Condition*): *The kernel function* $w(\cdot)$: $R\to[-1, 1]$ *is a twice continuously differentiable even function with*:

(a) $w(0)=1$, $w'(0)=0$, $w''(0)\neq 0$; *and either*
(b) $w(x)=0$, $|x|\geq 1$, *with either* $\lim_{|x|\to 1}w(x)/(1-|x|)^2=$ *constant or* $\lim_{|x|\to 1}(1-|x|)^3/w(x)=$ *constant, or*
(c) $w(x)=O(x^{-2})$, as $|x|\to 1$.

Note that this assumption is made based on one employed by Phillips (1995) and followed by Shintani (2001), although this is more restrictive. It should be also noted that $\Sigma_{n=-T+2}^{T-2}$ can be replaced with $\Sigma_{n=-S_T}^{S_T}$ and S_T is a band-width parameter, e.g., see Andrews (1991).

The test statitic is now given as

$$\hat{P}=TS_T tr\{\bar{w}\hat{\Omega}(\hat{M}_{ZZ})^{-1}\},$$

with $\bar{w}=\int_{-1}^{1}w(x)\,dx$.

In order to state the asymptotic properties of $\hat{P}$, let us introduce the conventionalized symbols $\Rightarrow$ and $W_n(v)$ standing for weak convergence of probability measures on the unit interval [0, 1] and the n-vector standard Brownian motion on [0, 1] respectively.

Theorem 1 *Suppose that* y_t *is generated by* (1). *Then:*

(i) *Under* H_0,

$$\hat{P}\ \Rightarrow\ tr\{\left(\bar{V}_k-\tilde{V}_{k;s}\breve{U}^{-1}_{s;s}\tilde{V}'_{k;s}\right)^{-1}\},$$

where

$$\bar{V}_k=\int_0^1\begin{bmatrix}W_r(u)\\ W_s(u)\end{bmatrix}[W_r'(u),\ W_s'(u)]\,du,$$
$$\tilde{V}_{k;s}=\int_0^1\begin{bmatrix}W_r(u)\\ W_s(u)\end{bmatrix}\left[\bar{W}_s'(u),\ 1,\ u,\ u^2\right]du,$$
$$\breve{U}_{n;n}=\int_0^1\begin{bmatrix}\int_0^u W_n(v)dv\\ 1\\ u\\ u^2\end{bmatrix}\left[\int_0^u W_n'(v)dv,\ 1,\ u,\ u^2\right]du,\quad n=1,2,\cdots,$$

with the notice that $W_k(u)=(W_r'(u), W_s'(u))'$.

(ii) *Under* H_1,

$$\hat{P}=Ttr\{P_1'\tilde{\Pi}P_1\left(P_1'R_{\bar{u}}P_1\right)^{-1}\}+O_p(S_T^{1/2}),$$

where P_1 *is of* $k\times\bar{r}$ *with* $\bar{r}=k-\bar{s}$ *and constitutes* $[P_1, P_2]$ *of* $k\times k$ *constant matrix such that rank* $[P_1, P_2]=k$, $P_2'P_1=0$,

$$rank\ [P_1,\ P_2]=k,\qquad P_2'P_1=0,$$
$$P_1'\begin{bmatrix}\beta'C^{(1)}(1)\\ \gamma'C(1)\end{bmatrix}=0,\ and\ rank\,P_2'\begin{bmatrix}\beta'C^{(1)}(1)\\ \gamma'C(1)\end{bmatrix}=\bar{s},$$

$\tilde{\Pi}$ *is of* $k \times k$ *such that* $(T^{1/2}/S_T^{1/2})\tilde{\Pi} = O_p(1)$ *and* $\{(T^{1/2}S_T^{1/2})\tilde{\Pi}\}^{-1} = O_p(1)$, *and* $R_{\bar{u}} = E\bar{u}_t\bar{u}_t'$, *putting*

$$\bar{u}_t = \begin{bmatrix} \beta' C^{(2)}(B) \\ \gamma' C^{(1)}(B) \end{bmatrix} \epsilon_t.$$

Theorem 1 (ii) states that test statistic is of $O_p(T^{1/2}S_T^{1/2})$ under the alternative.

Let us denote the limiting distributions of $\hat{P}$ in Theorem 1 as $\Psi_{k;s}$. $\Psi_{k;s}$ is simple, normalization-free and free of nuisance parameters, indicating that to estimate the percentage points using the Monte Carlo method is easy. Table 1 reports the upper percentage points of $\Psi_{k;s}$ as the 5 and 1 percent critical values of the test for several k and s, simulating it with 70000 replications and 5000 samples based on pseudo normal random variables.

This paper had studied the construction and performance of the test based on $\hat{P}$ under the supposition that y_t generated by (1) is cointegrated (i.e., $r \geq 1$) and r is known in spite of the reality that r must be determined through some method. In this section we shall discuss how our test can be dealt with under the situation in which r is unknown. Now, suppose that y_t is generated by (1) in which r is allowed to be 0, i.e., *rank* $C(1)=k$ is allowable. The conventional way to deal with this issue is to determine r before executing our testing procedure. We will take up the test proposed in Shintani (2001) or Phillips and Oulialis (1990), noting that the validity of one based on Johansen's (1988) well-known likelihood ratio test (trace test) is not established under our situation in which the VECM for $I(1)$ is not ensured to be finite-order even if H_0 is true. Let $\hat{\lambda}_1 \geq \cdots \geq \hat{\lambda}_k$ be the ordered eigenvalues of

$$\left(\sum_{n=-T+2}^{T-2} k(n/S_T)\hat{R}_{00}(n)\right)\left(\hat{R}_{11}/T\right)^{-1}\}$$

and put $\tilde{P}(h) = \Sigma_{j=h+1}^{k}\hat{\lambda}_j$ for $h=0,\cdots, k-2$. $\tilde{P}(h)$ is exactly the test statistic $P(n, s)$ in Shintani (2001). Having a keen awareness of the issue arising under the lack of Condition (A), this paper will restate a part of Shintani's (2001) Theorem 3.1 as

Theorem 2 *Suppose that* y_t *is generated by* (1) *in which the case* $r=0$ *is allowed. Then:*

(i) *If* $h=r$,

$$\tilde{P}(h) \Rightarrow tr\{\left(\int_0^1 \tilde{W}_s(u)\tilde{W}_s'(u)du\right)^{-1}\},$$

where

$$\tilde{W}_s(u) = W_s(u)\int_0^1 - \left(\int_0^1 W_s(u)(1,\, u)du\right) \times\left(\int_0^1 (1,\, u)'(1,\, u)du\right)^{-1}(1,\, u)'.$$

(ii) *If* $h \leq r-1$, $\tilde{P}(h) = T\Sigma_{j=h+1}^{r}\check{\lambda}_{j-h}(h) + O_p(S_T^{1/2})$, *where* $\check{\lambda}_{j-h}(h), j=h+1,\cdots, r$, *are the eigenvalues of*

$$\hat{w}_{r-h}'\beta'\check{\Pi}\beta\hat{w}_{r-h}(\hat{w}_{r-h}'\beta' R_{\bar{v}}\beta\hat{w}_{r-h})^{-1},$$

where $\hat{w}_{r-h}'$ *is of* $(r-h)\times r$ *such that* $\hat{w}_{r-h} = O_p(1)$ *and* $(\hat{w}_{r-h}'\hat{w}_{r-h})^{-1} = O_p(1)$, $\check{\Pi}$ *is of* $k \times k$ *such that* $(T^{1/2}/S_T^{1/2})\check{\Pi} = O_p(1)$ *and* $\{(T^{1/2}/S_T^{1/2})\check{\Pi}\}^{-1} = O_p(1)$, *and* $R_{\bar{v}} = E\bar{v}_t\bar{v}_t'$, *putting* $\bar{v}_t = C^{(1)}(B)\varepsilon_t$.

Note in (ii) that r must be positive. It is ensured by this theorem that testing successively $r=0$, $r=1,\cdots$, based on $\tilde{P}(h)$ and Johansen's (1996, p. 167) procedure still leads to a consistent estimator for the true value of r even if Condition (A) is not satisfied. Based on this, it will not be so difficult to see that $\hat{P}$ constructed using the estimate of r possesses the asymptotics close to those for one under known r established by Theorem 1.

4 EXAMPLES

In this section we shall illustrate to what extent the asymptotic properties in the previous section are preserved for finite samples, based on Monte Carlo experiments in different DGPs as special cases of (1). All the DGPs are bivariate systems ($k=2$) with $y_t = (y_{1t}, y_{2t})'$ and $\varepsilon_t = (\varepsilon_1, \varepsilon_2)'$ distributed as Gaussian with mean zero and covariance matrix I_2 ($\Omega_{00} = I_2$), provided that whether or not each DGP possesses a linear trend is decided by the parameter g taking either 0 or 1. The DGPs presented below are classified into groups of three: of VMA(2) models in first differences, VECMs of finite lag-lengths for $I(1)$ and those for $I(2)$. The DGPs included in each group are expressed in a unified form with several parameters. All the DGPs are constructed so that the roots of the VMA/VECM characteristic equations are either equal to 1 or strictly greater than 1 in absolute value, which is imposed for det $C(z)=0$ in (1). $y_{-j}, j \geq 0$, are supposed to be zero.

The unified expression for the first group is given as

$$\begin{pmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{pmatrix} = \begin{bmatrix} 1.2 & 0 \\ -0.8 & 0 \end{bmatrix} \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix} + (1-B) \begin{bmatrix} -0.2 - g_1 B & -g_2 B \\ 0.8 & 1 + g_3 B \end{bmatrix} \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix} + \begin{pmatrix} 0.6g \\ -0.4g \end{pmatrix}.$$

The values of g_i are confined to the cases below: Case (1) as $g_1=g_3=0$, $g_2=0.2$, Case (2) as $g_1=0$, $g_2=0.2$, $g_3=0.4$, Case (3) as $g_1=1$, $g_2=1.8$, $g_3=0.2$, and Case (4) as $g_1=1$, $g_2=1.2$, $g_3=-0.2$ respectively. We put $g=0$ through these cases. For $i=1,\cdots,4$, Case (i*) is defined as one added a deterministic trend to Case (i) (i.e., Case (i) replaced by $g=1$ without changing other parameter values). Through all the cases in this group, it is obvious that the matrix of the first term in the right-hand side of the above relation corresponds to $C(1)$ and that it of the second term is $C^{(1)}(B)$. Choosing (0.8, 1.2) as one of β, it is easy to check that $\beta' C(1)=0$ and H_0 is true for Case (1), (1*), (2) and (2*). On the other hand, $\beta' C(1)=\beta' C^{(1)}(1)=0$ for Case (3), (3*), (4) and (4*), implying that H_0 is false for these cases owing to a relation on higher order cointegration ($d=1$ and $b=2$ in Engle and Granger's (1987) definition).

The expression for the second group is the following one.

$$\begin{pmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{pmatrix} = \begin{bmatrix} 0.8 & 1.2 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & g_4 \end{bmatrix} \begin{pmatrix} \Delta y_{1t-1} \\ \Delta y_{2t-1} \end{pmatrix} + \begin{pmatrix} 0.5g \\ 0.5g \end{pmatrix} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}.$$

The cases considered here are as follows: Case (5) as $g_4=g=0$, Case (6) as $g_4=-0.2$, $g=0$ and Case (i*) corresponding to Case (i) replaced by $g=1$, $i=5, 6$, respectively, similar to in the first group case. All the DGPs in the second group are cointegrated systems with $\beta'=(0.8, 1.2)$ as the cointegrating matrix and can be converted into some VMA representations formulated as (1), letting

$$C(z) = \begin{bmatrix} \frac{1-z}{1-0.2z} & \frac{-1.2z}{(1-0.2z)(1-g_4 z)} \\ 0 & \frac{1}{1-g_4 z} \end{bmatrix}.$$

in view of Lemma 1, it is obvious that H_0 is true for both cases.

The third group is expressed by

$$\begin{pmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{pmatrix} = \begin{bmatrix} 0 & 0 \\ 0.8 & 1.2 \end{bmatrix} \{\sum_{h=1}^{t-1} \begin{pmatrix} y_{1h} \\ y_{2h} \end{pmatrix}\} + \begin{bmatrix} 0.8 & 1.2 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & g_5 \end{bmatrix} \begin{pmatrix} \Delta y_{1t-1} \\ \Delta y_{2t-1} \end{pmatrix} + \begin{pmatrix} 0.5g \\ 0.5g \end{pmatrix} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}.$$

The cases considered are as follows: Case (7) as $g_5=g=0$, Case (8) as $g_5=-0.2$, $g=0$ and Case (i*) added a deterministic trend to Case (i), $i=7, 8$, respectively, as in the first or second group case. All the DGPs in this group are included into ones expressed by (5), and (1) can be also derived with

$$C(z) = \begin{bmatrix} 1 - 0.8z + z^2 - g_5 z(1-z)^2 & -1.2z(1-z) \\ -0.8z & (1-0.2z)(1-z) \end{bmatrix} / f(z)$$

where $f(z)$ is $(1-0.447213iz)(1+0.447213iz)$ for Example (7) and $\{1-(0.134025-0.259804i)z\}\{1+(0.134025-0.259804i)z\}(1+0.468051z)$ for Example (8) with i denoting the imaginary, indicating that H_0 is false under any DGP here.

Throughout all of DGPs, we ran 10000 replications of experiments under 100, 250 and 500 as the sample size T, and pseudo normal random variables for ε_{it}, $i=1, 2$, were adopted to produce $\hat{P}$ in each experiment. $\hat{\beta}$ and $\hat{\gamma}$ (and S_{11} as a basis for those) are obtained as $p=4$, and this paper confines $k(n/S_T)$ to Bartlett Kernel with $S_T=4, 6$.

The aim of the experiments is to obtain the relative frequency (the probability) for the test based on $\hat{P}$ to make a correct decision on H_0 among 10000 replications under each DGP, T, S_T and critical point, using 5% and 1% critical points as the upper percentage points of $\Psi_{k;s}$ reported in Table 1 (the case $k=2$, $s=1$). We will provide it in two ways. One is the relative frequency of the events in which H_0 is rejected as the empirical size or power of the test and follows automatically from $r=1$ given. Another is the relative frequency of the events in which a correct decision on H_0 is made provided that the events that $\tilde{P}$ (h) test as $k=2$ and $h=0$ accepts $r=0$ are excluded.[6] Table 2 provides the empirical sizes and powers of the test $\hat{P}$ using the 5% and 1% critical values in Table 1. On the other hand, the results on the relative frequency of a correct decision in association with $\tilde{P}(0)$ are in Table 3. Note that the figures in Table 2 and 3 are percentiles.

The experimental results in Table 2 show that the performance of the test is not so undesirable under the level of significance set at 5%, T which is as many as 250 and $S_T=4$. However, it may be observed that as $T=250$, the convergence of the empirical size/power to theoretical one is not generaally prominent. In particular, such a sample size yields unsatisfactory results on the power. We will find that such power distortion is noticeably deteriorated as the 1% level of significance is used or $S_T=6$, suggesting that there is great deviation of the empirical distribution of $\hat{P}$ from the limiting one in the part of the tail more. Moreover, we may recognize severer size/power distortion in the VECM cases such as Case (5), (6), (5′) and (6′).

The results for $T=100$ are surprisingly favorable for H_0 through all the cases and consequently it was revealed that the test was extremely powerless under such a sample size. We will note that such powerless performances were recognized similarly in the results on $\tilde{P}(0)$, although not provided in this paper.

On the other hand, the results for T equal to 500 reflect the asymptotics considerably. As $T=500$, the greater value of S_T ($S_T=6$) tends to lead to more favorable results for H_0. However, we will report that the test showed noticeable deterioration in the power in contrast to small improvement in the size as S_T attains 8, although not provided in this paper.

The relative frequency of a correct decision in association with $\tilde{P}(0)$ generally shows a similarity with its corresponding empirical size/power, based on that the performance of $\tilde{P}(0)$ as the test for $r=0$ was similar to of $\hat{P}$. although the difference between both was not necessarily slight for all the cases. In general, the test under the level of significance set at 5% seems to be sufficiently supportable as long as T is as many as 250 and $S_T=4$ is chosen. However, as $T=250$, we may find many cases which yield extremely poor results in the case that the 1% level of significance or $S_T=6$ is used. We will note that the results for $T=100$ excluded in Table 3 are far more undesirable than those for $T=250$, as expected from the results reported above on the power. As T attains about 500, we will see that the relative frequency of a correct decision in association with $\tilde{P}(0)$ sufficiently becomes close to one, reflecting strongly the asymptotics of $\hat{P}$ and the consistency of $\tilde{P}(0)$.

5 CONCLUDING REMARKS

As has been discussed above, the valid derivation of the conventional $I(1)$ VECM representation is not always ensured even if the vector time series system considered is cointegrated with $I(1)$ components. The interest of this paper is put not in formulating concrete relations on multicointegration or higher order cointegration to make the derivation invalid but in finding a method to check whether or not the derivation is possible, stimulated by that it has not been checked empirically. As one of useful methods, this paper proposed the test based on $\hat{P}$.

The null hypothesis of the test formulates the situation in which the derivation is achieved, whereas the details of the situation under the alternative are out of consideration. Theorem 1 (i) in this paper provides the limiting distribution of the test statistic under the null as an available form for the critical values, in the sense that it is normalization-free, free of nuisance parameters and constructed based on standard Brownian motions. We also show that the test is consistent, although the convergence of the power to 1 (as the limit) is rather slow, as verified by the order of $\hat{P}$ in Theorem 1 (ii).

Such slow convergence will be also reflected in the experimental results for $T=100$ or 250 in the previous section, resulting in poor performances of the test in most cases. In particular, the experiments revealed that the sample size as many as 100 was completely insufficient to distinguish H_1 from H_0. However, as long as T is as many as 250, the results under the level of significance of 5% and $S_T = 4$ were not so undesirable through all the cases. It should be also emphasized that T attaining about 500 improved the test performances noticeably.

This paper has not sufficiently dealt with the experiments on how the performance of the test $\hat{P}$ is improved by a selection from various kernel functions or band-width parameters under finite samples. It may be plausible that some kernel exhibits superiority with a value of the band with parameter under some DGPs as $T=250$, following the experiments of this paper. We will only state that in the DGPs adopted for the experiments, a smaller value of the band-width parameter, i.e., $S_T=4$, led to better results.

The test proposed in this paper will be easily extended to one for a DGP taking account not only $tC(1)\mu$ but also higher-order deterministic trends into account, based on replacement of Z_{2t}, $Z_{4;t-n}$, $Z_{5;t-n}$ and $Z_{6;t-n}$ with ones arranged in order to remove such trends, although we have a view that it is not so necessary or crucial for $I(1)$ formulation.

Our test is nonparametric in the sense that $\hat{P}$ is constructed without being dependent on any parametric model, similar to in Phillips and Oulialis (1990) or Shintani (2001) and was not necessarily satisfactory under some specific DGPs and finite samples, as seen in the experimental results. The test will be nevertheless appreciated as a unified and wide-range method, although supposing some parametric model such as the $I(2)$ VECM may lead to more powerful tests.

References

Andrews, W. K., (1991) Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, Vol. 59, pp 817-58.

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

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

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

Engle, R. F. and Yoo, B. S. (1987) Forecasting and testing in co-integrated systems. *Journal of Econometrics*, Vol. 35, pp 143-59.

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

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

Engsted, T., Gonzalo, J. and Haldrup, N. (1997) Testing for multicointegration. *Economic Letters*, Vol. 56, pp 259-66.

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

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

Hamilton, J. D., (1994) *Time Series Analysis*. New Jersey: Princeton University Press.

Hannan, E. J., (1970) *Multiple Time Series*. New York: Wiley.

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

Johansen, S., (1991) Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica*, Vol. 59, pp 1551-80.

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

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

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

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

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

Phillips, P. C. B. and Perron, P. (1988) Testing for a unit root in time series regression. *Biometrika*, Vol. 75, pp 335-46.

Qu, Z. and Perron, P. (2007) A modified information criterion for cointegration tests based on a VAR approximation. *Econometric Theory*, Vol. 23, pp 638-85.

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

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

Sims, C. A., Stock, Stock, J. H. and Watson, M. F. (1990) Inference in linear time series models with some unit roots. *Econometrica*, Vol. 58, pp 113-44.

Stock, J. H. and Watson, M. F. (1993) A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems. *Econometrica*, Vol. 61, pp 783-820.

FOOTNOTES

1 The VMAs in the d-th differences with an integer d greater than 1 are allowable for our consideration, since those can be interpreted as (1) by regarding $\Delta^{d-1}y_t$ as y_t, as mentioned in Introduction.

2 As another one required to derive the VECM representation, we refer to Assumption $B3$ in Banerjee et al. (1993, p. 258), which is equivalent to our Condition (A) as easily checked.

3 Condition (A) holds if and only if any linear combination of ($y_t'\beta$, $\Delta'\gamma$) is invertible, noting the supposition of det $C(z)$.

4 Regarding Δy_t, $\Sigma_{h=1}^{t-1} y_h$ and y_{t-1} as $\Delta^2 y_t$, y_{t-1} and Δy_{t-1} in connection with stated in our introduction or footnote 1, (5) becomes the familiar VECM form for $I(2)$ given in Johansen (1995, p. 57 and p. 132).

5 $\tilde{R}_{ij}$ are constructed in order to remove the stochastic trends in $\Sigma_{h=1}^{t}\hat{\gamma}' y_h$ as well. We note that it is required to make the limiting distribution of the test statistic stated in Theorem 1 later be free of nuisance parameters.

6 $\tilde{P}(h)$ here corresponds to the demeaned and detrended version of Shintani's $P(n, s)$ with the Bartlett Kernel under each of S_T here and therefore Phillips and Ouliaris's (1990) Table IV_c is utilized for the critical values.

Appendix

Proof of Lemma 1 It is trivial that

$$C(z) = \left[\beta(\beta'\beta)^{-1},\ \gamma(\gamma'\gamma)^{-1}\right] \begin{bmatrix} (1-z)\,I_r & 0 \\ 0 & I_s \end{bmatrix} \begin{bmatrix} \beta' C^{(1)}(z) \\ \gamma' C(z) \end{bmatrix}. \quad (A.1)$$

The sufficiency will be proved immediately by putting

$$A(z) = \begin{bmatrix} \beta' C^{(1)}(z) \\ \gamma' C(z) \end{bmatrix}^{-1} \begin{bmatrix} I_r & 0 \\ 0 & (1-z)\,I_s \end{bmatrix} \begin{bmatrix} \beta' \\ \gamma' \end{bmatrix}.$$

and multiplying both sides of (2) by $A(B)$.

Next, suppose that there exists a $A(z)$ satisfying the requirements for (4). Noting that (4) is written as

$$A(B)\, y_t = \varepsilon_t + \mu - A(1)\,\eta_0(1)$$

with $C(z) = I_k + \Sigma_{j=1}^{t-1} C_j z^j + o(1)$ and multiplying both sides of the above relation by $I_k + \Sigma_{j=1}^{t-1} C_j B^j$, we attain to

$$C(B)\,A(B)\,y_t = C(B)\,\varepsilon_t + C(1)\,\mu + o_p(1), \quad \forall t \geq 1. \quad (A.2)$$

Equating the left-hand side of (1) with that of (A.2) as $t \to \infty$, it must be asserted that $C(z)\,A(z) = (1-z),\ I_k$. Consequently,

$$(1-z)^k = (1-z)^r \det\left[\beta(\beta'\beta)^{-1},\ \gamma(\gamma'\gamma)^{-1}\right] \times \det\begin{bmatrix} \beta' C^{(1)}(z) \\ \gamma' C(z) \end{bmatrix} \det A(z). \quad (A.3)$$

On the other hand, by the supposition, det $A(z) = (1-z)^s$, $g(z)$ with a power series $g(z) = 1 + \Sigma_{j=1}^{\infty} g_j z^j$ such that g_j satisfy a convergence condition similar to one for C_j. This, together with (A.3), requires Condition (A).

Proof of Lemma 2 In view of (2) and (3), it is obvious that $((\Sigma_{h=1}^{t} y_h)'\beta, y_t'\gamma)'$ is expressed as

$$\begin{bmatrix} \beta' C^{(1)}(1) \\ \gamma' C(1) \end{bmatrix} \left(\sum_{h=1}^{t} \epsilon_t \right) + \begin{bmatrix} \beta' C^{(2)}(B) \\ \gamma' C^{(1)}(B) \end{bmatrix} \epsilon_t + \begin{bmatrix} t\beta'\eta_0(1) + \beta'\eta_0(2) \\ t\gamma' C(1)\mu + \gamma'\eta_0(1) \end{bmatrix}.$$

Noting that $rank\ \gamma' C(1) = s$, the result desired is derived immediately.

Proof of Lemma 3 Using an argument similar to used in the proof of Lemma 1, it is shown that the derivation of (1) requires $C(z)\,A(z) = (1-z)^2 I_k$, therefore, $A(z)\,C(z) = (1-z)^2 I_k$, where $A(z)$ is in the statement following this lemma and $C(z)$ is on (1). This in turn leads to

$$\alpha_1 \beta_1' C(1) = 0, \quad \alpha_1 \beta_1' C^{(1)}(1) + \alpha_2 \beta_2' C(1) = 0, \quad (A.4)$$

noting that the terms multiplied by $(1-z)^j$, $j=0, 1$, in $A(z)\,C(z)$ must be zero. In view of the definitions of $C(1)$ and β, it is trivial to see that there exists a suitable matrix b_1 such that $\beta_1' = b_1' \beta'$ and $rank\ b_1 = rank \beta_1 = rank\ \alpha_1$. Since $[\beta, \gamma]$ is full rank, we can let

$$\beta_2' = b_{21}' \beta' + b_{22}' \gamma'$$

with suitable matrices b_{2i}, $i=1, 2$, such that $rank\ [b_{21}', b_{22}'] = rank \beta_2 = rank\ \alpha_2$. Then it is obvious that

$$\alpha_2 \beta_2' C(1) = \alpha_2 b_{22}' \gamma' C(1).$$

In view of these results, it follows from the second equation of $(A.4)$ that

$$\bar{b}_1' \beta' C^{(1)}(1) + \bar{b}_2' \gamma' C(1) = 0$$

with a suitable matrix $[\bar{b}_1', \bar{b}_2']$ whose rank is equal to either $rank\ \alpha_1$ or $rank\ \alpha_2$, implying that Condition (A) does not hold.

Proof of Lemma 4 By combining the well-known asymptotic theory for $I(0)$ and $I(1)$ series, which are given in numerous literatures (e.g., Hamilton (1994, p. 548)), and its simple application with the suppositions on β and γ, we see that $D_T^{-1}\gamma' S_{11}\gamma D_T^{-1} = O_p(T)$, $(D_T^{-1}\gamma' S_{11}\gamma / T D_T^{-1})^{-1} = O_p(1)$, $\beta' S_{11}\beta = O_p(1)$ and $D_T^{-1}\gamma' S_{11}\beta = O_p(1)$. It is also obvious from the suppositions of β and γ that there exist random matrices $\hat{x}$, $\hat{y}$, $\hat{u}$ and $\hat{v}$ such that $\hat{\beta} = \beta\hat{x} + \gamma\hat{y}$ and $\hat{\gamma} = \gamma\hat{u} + \beta\hat{v}$. Since $\hat{\beta}$ corresponds to the r smallest eigenvalues of S_{11}, it follows that $\hat{\beta}' S_{11}\hat{\beta} = O_p(1)$, which in turn requires $\hat{x} = O_p(1)$, $\hat{x}^{-1} = O_p(1)$ and $D_T\hat{y} = O_p(\nu_T)$ with ν_T such that $\lim_{T\to\infty} \nu_T = 0$ and $T^{-1/2} \leq \nu_T > 0$. Similarly, $\hat{u} = O_p(1)$ and $\hat{u}^{-1} = O_p(1)$ since $D_T^{-1}\hat{\gamma}' S_{11}\hat{\gamma} D_T^{-1} = O_p(1)$. Then it is derived from $0 = \hat{\beta}'\hat{\gamma}$ and the results derived already that $D_T\hat{v} = O_p(\nu_T)$. Moreover, $0 = \hat{\beta}' S_{11}\hat{\gamma}$ implies

$$\begin{aligned} 0 = {} & \hat{u}' D_T (D_T^{-1}\gamma' S_{11}\gamma / T D_T^{-1})\, D_T T\hat{y} \\ & + \hat{u}' D_T (D_T^{-1}\gamma' S_{11}\beta)\hat{x} + \hat{v}'\beta' S_1 1 \beta\hat{x} \\ & + \hat{v}'(\beta' S_{11}\gamma D_T^{-1})\, D_T\hat{y}. \end{aligned}$$

Evaluating the orders of the terms in the right-hand side of the above equation, we must conclude that $\nu_T = T^{-1}$ to cancel out the order of $\hat{u}' D_T (D_T^{-1}\gamma' S_{11}\gamma / T D_T^{-1})\, D_T T\hat{y}$.

The following lemma states the consistency property of the kernel estimator for weakly stationary series in a general form and is needed to prove Theorem 1.

Lemma A.1 *For ξ_t given as $\xi_t = \Sigma_{j=0}^{\infty} D_j \varepsilon_{t-j}$, where D_j are $n \times k$ constant matrices with $n \leq k$ and $\Sigma_{j=1}^{\infty} j^{\check{\nu}}$, $\| D_j \| < \infty$ for some real number $\check{\nu} \geq 1$, let*

$$\Lambda = D(1)\Omega_{00}D(1)', \quad \hat{\Lambda} = \sum_{n=-T+2}^{T-2} k(n/S_T)\hat{R}_{\xi}(n),$$

where $D(1) = \Sigma_{j=0}^{\prime} D_j$ and $\hat{R}_{\xi}(n) = \Sigma_{t=n+2}^{T} \xi_t \xi_{t-n}' / T$ with $\hat{R}_{\xi}(-n) = \hat{R}_{\xi}(n)'$. Then, $\hat{\Lambda} = \Lambda + \hat{\Pi}$, with $\hat{\Pi}$ such that $(T^{1/2}/S_T^{1/2})\hat{\Pi} = O_p(1)$ and $\{(T^{1/2}/S_T^{1/2})\hat{\Pi}\}^{-1} = O_p(1)$.

This lemma is established by evaluating the covariance matrix of $(T^{1/2}/S_T^{1/2})\hat{\Lambda}$ in the same way as for Theorem 9 in Hannan (1970, p. 280) or Proposition 1 (a) in Andrews (1991). It should be noted that the row full rank property of $D(1)$ is not necessary to be assumed.

Lemma A.2 *Let η_t be $\Sigma_{j=0}^{\infty} D_{j;\eta}\varepsilon_{t-j}$ and suppose that $w(x;\ h-1)$ is defined and $rank\ D_{\xi}(1) = m_{\xi}$, $\xi = \eta, \zeta$. Then, for fixed positive integer h and nonnegtaive integers $\bar{i}$ and $\bar{j}$, we have:*

(i) *Putting*

$$\tilde{\psi}(\eta\zeta) = \sum_{n=-S_T+1+\bar{i}}^{S_T-1-\bar{j}} w((n-\bar{i})/S_T; h-1) \times\left(\hat{R}_{\eta\zeta}(n) - E\,\hat{R}_{\eta\zeta}(n)\right),$$

$\tilde{\Psi}(\eta\,\zeta)=O_p(S_T^{1/2}/T^{1/2})$ and $\{T^{1/2}/S_T^{1/2}\tilde{\Psi}(\eta\,\eta)\}^{-1}=O_p(1)$ or $\{T^{1/2}/S_T^{1/2}\tilde{\Psi}(\zeta\,\zeta)^{-1}=O_p(1)$.

(ii)

$$\sum_{n=-S_T+1+\bar{i}}^{S_T-1-\bar{j}} w((n-\bar{i})/S_T; h-1)E\,\hat{R}_{\eta\zeta}(n) = D_\eta\Omega_{00}D'_\zeta + o(1).$$

Proof (i) can be regarded as exactly the same as Theorem 9 of Hannan (1970, p. 280) or Proposition 1 (a) of Andrews (1991), evaluating a submatrix of the Kernel estimator in the theroem/proposition, regarding $w((n-\bar{i})/S_T;h-1)$ as $w(n/S_T;0)$ and noting that

$\hat{R}_{\eta\zeta}(n)-E\hat{R}_{\eta\zeta}(n)=O_p(T^{-1/2})\quad\forall$ *integer n*,

$$\begin{aligned}&\sum_{n=-S_T+1+\bar{i}}^{S_T-1-\bar{j}} w(n/S_T;0)\left(\hat{R}_{\eta\zeta}(n) - E\,\hat{R}_{\eta\zeta}(n)\right)\\ =&\sum_{n=-S_T+1}^{S_T-1} w(n/S_T;0)\left(\hat{R}_{\eta\zeta}(n) - E\,\hat{R}_{\eta\zeta}(n)\right)\\ &+O_p(T^{-1/2}),\end{aligned}$$

based on the standard theory for $I(0)$ series.

For (ii), it is trivial that

$$\begin{aligned}&\sum_{n=-S_T+1+\bar{i}}^{S_T-1-\bar{j}} w((n-\bar{i})/S_T;h-1)E\,\hat{R}_{\eta\zeta}(n)\\ =&\sum_{n=-S_T+1+\bar{i}}^{S_T-1-\bar{j}} w((n-\bar{i})/S_T;h-1)R_{\eta\zeta}(n)\,(1-|n|/T)\,.\end{aligned}$$

Letting $\{M_T\}$ denote a series of positive integers such that $\lim_{T\to\infty}S_T^{1/\bar{\nu}}/M_T=0$ and $\lim_{T\to\infty}M_T/S_T=0$, it is asserted by the condition on $D_{\bar{j},\xi}$ that

$R_{\eta\eta}(M_T)=O(M_T^{-\bar{\nu}}),\quad R_{\eta\zeta}(-M_T+1)=O(M_T^{-\bar{\nu}})$.

It is also easy to see

$$\begin{aligned}&|\sum_{n=M_T}^{S_T-1-\bar{j}} w((n-\bar{i})/S_T;h-1)R_{\eta\zeta}(n)\,(1-|n|/T)|\\ \le&\ \bar{C}\sum_{n=M_T}^{S_T-1-\bar{j}}|R_{\eta\zeta}(n)| = O(S_TM_T^{-\bar{\nu}})=o(1),\\ &|\sum_{n=-S_T+1+\bar{i}}^{-M_T} w((n-\bar{i})/S_T;h-1)R_{\eta\Delta\zeta}(n)\,(1-|n|/T)|\\ \le&\ \bar{C}\sum_{n=-S_T+1+\bar{i}}^{-M_T}|R_{\eta\zeta}(n)| = O(S_TM_T^{-\bar{\nu}})=o(1).\end{aligned}$$

Moreover, we have

$$\begin{aligned}&\sum_{n=-M_T+1}^{M_T-1} w((n-\bar{i})/S_T;h-1)R_{\eta\zeta}(n)\,(1-|n|/T)\\ =&\ w(0;h-1)\sum_{n=-M_T+1}^{M_T-1}R_{\eta\zeta}(n)\,(1-|n|/T)+o(1)\end{aligned}$$

since $w((n-\bar{i})/S_T;h-1)=w(0;h-1)\ (1+O(M_T/S_T))$ for any $|n|\le M_T-1$.

Putting these results together and noting that

$$\begin{aligned}&\lim_{N\to\infty}\sum_{i=-N}^{N}|n||R_{\eta\zeta}(n)|<\infty,\\ &\lim_{N\to\infty}\sum_{i=-N}^{N}R_{\eta\zeta}(n)=D_\eta\Omega_{00}D'_\zeta,\end{aligned}$$

it is easy to derive the required relation.

We note that Lemma A.2 (ii)corresponds to Theorem 10 of Hannan (1970, p. 283) except that the assumption on the Kernel function is different from ours.

Lemma A.3 *Let η_t be $\Sigma_{h=1}^{t}\varepsilon_h$ and suppose that $w(x;h-1)$ is defined and rank $D_\zeta(1)=m_\zeta$. Then, for h, $\bar{i}$ and $\bar{j}$ in Lemma A.2, we have:*

$$\begin{aligned}&S_T^{-1}\sum_{n=-S_T+1+\bar{i}}^{S_T-1-\bar{j}} w((n-\bar{i})/S_T;h-1)\hat{R}_{\eta\eta}(n)/T\\ \Rightarrow&\ \bar{w}(h-1)\Omega_{00}^{1/2}\left(\int_0^1 W_k(u)W'_k(u)du\right)\Omega_{00}^{1/2},\\ &S_T^{-1}\sum_{n=-S_T+1+\bar{i}}^{S_T-1-\bar{j}} w((n-\bar{i})/S_T;h-1)\hat{R}_{\eta\zeta}(n)=O_p(1)\end{aligned}$$

where $\bar{w}(h-1)==\int_{-1}^{1}w(x;h-1)dx$.

Proof For any $n=1,\cdots,S_T-1$, put

$\tilde{\Psi}_\eta(n)=\hat{R}_{\eta\eta}(n)/T-\hat{R}_{\eta\eta}(0)/T$,

$\tilde{\Psi}_\eta(-n)=\hat{R}_{\eta\eta}(-n)/T-\hat{R}_{\eta\eta}(0)/T$,

$\tilde{\Psi}_{\eta\zeta}(n)=\hat{R}_{\eta\zeta}(n)-\hat{R}_{\eta\zeta}(0)$,

$\tilde{\Psi}_{\eta\zeta}(-n)=\hat{R}_{\eta\zeta}(-n)-\hat{R}_{\eta\zeta}(0)$

with the notice $\tilde{\Psi}_\eta(-n)=\tilde{\Psi}'_\eta(n)$. Then we have

$$\begin{aligned}&S_T^{-1}\sum_{n=-S_T+1+\bar{i}}^{S_T-1-\bar{j}} w((n-\bar{i})/S_T;h-1)\hat{R}_{\eta\eta}(n)/T\\ =&\ \{S_T^{-1}\sum_{n=-S_T+1+\bar{i}}^{S_T-1-\bar{j}} w((n-\bar{i})/S_T;h-1)\}\hat{R}_{\eta\eta}(0)/T\\ &+S_T^{-1}\sum_{n=-S_T+1+\bar{i}}^{S_T-1-\bar{j}} w((n-\bar{i})/S_T;h-1)\tilde{\psi}_\eta(n),\end{aligned}\tag{A.5}$$

$$S_T^{-1}\sum_{n=-S_T+1+\bar{i}}^{S_T-1-\bar{j}} w((n-\bar{i})/S_T;h-1)\hat{R}_{\eta\zeta}(n)$$
$$= \{S_T^{-1}\sum_{n=-S_T+1+\bar{i}}^{S_T-1-\bar{j}} w((n-\bar{i})/S_T;h-1)\}\hat{R}_{\eta\zeta}(0)$$
$$+S_T^{-1}\sum_{n=0}^{S_T-1-\bar{j}} w((n-\bar{i})/S_T;h-1)\tilde{\psi}_{\eta\zeta}(n)$$
$$+S_T^{-1}\sum_{n=1}^{S_T-1-\bar{i}} w((-n-\bar{i})/S_T;h-1)\tilde{\psi}_{\eta\zeta}(-n). \qquad (A.6)$$

The well-known theory on Brownian motions (e.g., Hamilton (1994, p. 548)) states

$$\eta_t/T^{1/2} \;\Rightarrow\; \Omega_{00}^{1/2} W_k(t/T),$$

which, together with the asymptotics for $I(0)$ and $I(1)$ used in the proof of Lemma 4, leads to

$$\{S_T^{-1}\sum_{n=-S_T+1+\bar{i}}^{S_T-1-\bar{j}} w((n-\bar{i})/S_T;h-1)\}\hat{R}_{\eta\eta}(0)/T$$
$$\Rightarrow \quad \bar{w}(h-1)\Omega_{00}^{1/2}\left(\int_0^1 W_k(u)W_k'(u)du\right)\Omega_{00}^{1/2}, \qquad (A.7)$$

$$\{S_T^{-1}\sum_{n=-S_T+1+\bar{i}}^{S_T-1-\bar{j}} w((n-\bar{i})/S_T;h-1)\}\hat{R}_{\eta\zeta}(0)$$
$$= O_p(1). \qquad (A.8)$$

Since $\sum_{j=0}^{S_T-2}\varepsilon_{t-j}=O_p(S_T^{1/2})$, it is also easy to see that

$$\tilde{\psi}_\eta(n) = -\sum_{t=n+2}^{T}(\sum_{h=1}^{t}\epsilon_h)(\sum_{j=0}^{n-1}\epsilon_{t-j}')/T^2$$
$$-\sum_{t=2}^{n+1}(\sum_{h=1}^{t}\epsilon_h)(\sum_{h=1}^{t}\epsilon_h')/T^2$$
$$= O_p(S_T^{1/2}T^{-1/2}),$$
$$\tilde{\psi}_{\eta\zeta}(n) = \hat{R}_{\eta\zeta}(n) - \sum_{t=n+2}^{T}(\sum_{h=1}^{t-n}\epsilon_h)\zeta_{t-n}'/T$$
$$+\sum_{t=n+2}^{T}(\sum_{h=1}^{t-n}\epsilon_h)\zeta_{t-n}'/T - \sum_{t=2}^{T}(\sum_{h=1}^{t}\epsilon_h)\zeta_t'/T$$
$$= \sum_{t=n+2}^{T}(\sum_{j=0}^{n-1}\epsilon_{t-j})\zeta_{t-n}'/T - \sum_{t=T-n+1}^{T}(\sum_{h=1}^{t}\epsilon_h)\zeta_t'/T$$
$$= O_p(S_T^{1/2}T^{-1/2}),$$
$$\tilde{\psi}_{\eta\zeta}(-n) = \hat{R}_{\eta\zeta}(-n) - \sum_{t=n+2}^{T}(\sum_{h=1}^{t}\epsilon_h)\zeta_t'/T$$
$$-\sum_{t=2}^{n+1}(\sum_{h=1}^{t}\epsilon_h)\zeta_t'/T$$
$$= -\sum_{t=n+2}^{T}(\sum_{j=0}^{n-1}\epsilon_{t-j})\zeta_t'/T + O_p(S_T/T) = O_p(1),$$

for any $n=1,\cdots, S_T-1$. In view of these results, we can find suitable positive definite matrices M_i whose components are of $O(1)$ with quantities $O(S_T/T)$ and $O(1)$ so that

$$E\tilde{\theta}_i\,\tilde{\theta}_i' < \bar{w}^2(h-1)\,O(S_T/T)\,M_i, \quad i=1,2,$$
$$E\tilde{\theta}_3\;\tilde{\theta}_3' < \bar{w}^2(h-1)\,O(1)\,M_3, \qquad (A.9)$$

where

$$\tilde{\theta}_1 = vec\{S_T^{-1}\sum_{n=-S_T+1+\bar{i}}^{S_T-1-\bar{j}} w((n-\bar{i})/S_T;h-1)\tilde{\psi}_\eta(n)\},$$
$$\tilde{\theta}_2 = vec\{S_T^{-1}\sum_{n=0}^{S_T-1-\bar{j}} w((n-\bar{i})/S_T;h-1)\tilde{\psi}_{\eta\zeta}(n)\},$$
$$\tilde{\theta}_3 = vec\{S_T^{-1}\sum_{n=1}^{S_T-1-\bar{i}} w((-n-\bar{i})/S_T;h-1)\tilde{\psi}_{\eta\zeta}(-n)\}.$$

Evaluating ($A.7$) or ($A.8$) with ($A.9$) in ($A.5$) or ($A.6$), we obtain the desired result for the lemma.

We note that such asymptotics as in Lemma A.2 are provided as well in Shintani (2001), although our kernel function $w(x;h-1)$ is more general in construction. Now, for any nonnegative integers h, n, h_i, m_2, $\bar{i}$ and $\bar{j}$ such that $h_i \geq n\,(i=1,2)$, put

$$\tilde{\psi}_{1;T}(h;n;h_1;m_2;\bar{i};\bar{j})$$
$$=-\sum_{j=0}^{S_T-1-\bar{j}} w((j-\bar{i})/S_T;h+n-2)\left(\Delta^{h_1-n}\eta_{j+1}\Delta^{m_2}\zeta_2'\right.$$
$$\left.-\Delta^{h_1-n}\eta_T\Delta^{m_2}\zeta_{T-j+1}'\right)/T,$$
$$\tilde{\psi}_{2;T}(h;n;h_1;m_2;\bar{i};\bar{j})$$
$$=-w((-S_T+1)/S_T;h+n-2)\hat{R}_{\Delta^{h_1-n}\eta\Delta^{h_2}\zeta}(-S_T+\bar{i})$$
$$+w((S_T-1-\bar{j})/S_T;h+n-2)\hat{R}_{\Delta^{h_1-n}\eta\Delta^{m_2}\zeta}(S_T-1-\bar{j}),$$
$$\tilde{\psi}_{3;T}(h;n;h_2;\bar{i};\bar{j})$$
$$=-\sum_{j=0}^{S_T-1-\bar{i}} w((-j-\bar{i})/S_T;h+n-2)\left(\eta_2\Delta^{h_2-n}\zeta_{j+1}'\right.$$
$$\left.-\eta_{T-j+1}\Delta^{h_2-n}\zeta_T'\right)/T,$$
$$\tilde{\psi}_{4;T}(h;n;h_2;\bar{i};\bar{j})$$
$$=-w((S_T-1-\bar{j})/S_T;h+n-2)\hat{R}_{\eta\Delta^{h_2-n}\zeta}(S_T-\bar{j})$$
$$+w((-S_T+1)/S_T;h+n-2)\hat{R}_{\eta\Delta^{h_2-n}\zeta}(-S_T+1+\bar{i}).$$

Lemma A.4 *For h, $\bar{i}$ and $\bar{j}$ in Lemma A.2, let η_t be $\sum_{j=0}^{\infty}D_{j;\eta}\varepsilon_{t-j}$ and suppose that $w(x;h+n-1)$ is defined and rank $D_\xi(1)=m_\xi$, $\xi=\eta$, ζ. Then:*

$$\tilde{\psi}_{1;T}(h;n;h_1;m_2;\bar{i};\bar{j}) = O_p(S_T^{1/2-\check{q}}T^{-1}) \quad if\ h_1=n\ and\ m_2=0,$$
$$= O_p(S_T^{-\check{q}}T^{-1}) \quad otherwise,$$
$$\tilde{\psi}_{2;T}(h;n;h_1;m_2;\bar{i};\bar{j}) = O_p(S_T^{-\check{q}})O_p(max\{S_T^{\bar{\nu}},\,T^{-1/2}\}),$$
$$\tilde{\psi}_{3;T}(h;n;h_2;\bar{i};\bar{j}) = O_p(S_T^{1/2-\check{q}}T^{-1}),$$

where $\check{q}$ is 1 if $h+n-2=0$ and Assumption KL(b) is *satisfied and is* 0 *otherwise.*

The following results will be needed to establish asymptotics for the kernel estimator constructed overdifferenced series.

Lemma A.5 *Suppose that* $w(x;h-1)$ *is defined. Then, for h,* $\bar{i}$ *and* $\bar{j}$ *in Lemma A.2, we have:*

(i)

$$\sum_{n=-S_T+1+\bar{i}}^{S_T-1-\bar{j}} w((n-\bar{i})/S_T;h-1)\hat{R}_{\Delta\eta\zeta}(n)$$
$$= -S_T^{-1}\sum_{n=-S_T+1+\bar{i}}^{S_T-2-\bar{j}} w((n-\bar{i})/S_T;h)\hat{R}_{\eta\zeta}(n)$$
$$\times\Big(1+O(S_T^{-1})\Big)+\sum_{i=1}^{2}\tilde{\psi}_{i;T}(h;1;1;0;\bar{i};\bar{j}),$$
$$\sum_{n=-S_T+1+\bar{i}}^{S_T-1-\bar{j}} w((n-\bar{i})/S_T;h-1)\hat{R}_{\eta\Delta\zeta}(n)$$
$$= S_T^{-1}\sum_{n=-S_T+2+\bar{i}}^{S_T-1-\bar{j}} w((n-1-\bar{i})/S_T;h)\hat{R}_{\eta\zeta}(n)$$
$$\times\Big(1+O(S_T^{-1})\Big)+\sum_{i=3}^{4}\tilde{\psi}_{i;T}(h;1;1;\bar{i};\bar{j})$$

if $w(x;h-1)$ *is not constant over* $[-1, 1]$, *and*

$$\sum_{n=-S_T+1+\bar{i}}^{S_T-1-\bar{j}} w((n-\bar{i})/S_T;h-1)\hat{R}_{\Delta\eta\zeta}(n)$$
$$= \sum_{i=1}^{2}\tilde{\psi}_{i;T}(h;1;1;1;\bar{i};\bar{j}),$$
$$\sum_{n=-S_T+1+\bar{i}}^{S_T-1-\bar{j}} w((n-\bar{i})/S_T;h-1)\hat{R}_{\eta\Delta\zeta}(n)$$
$$= \sum_{i=3}^{4}\tilde{\psi}_{i;T}(h;1;1;\bar{i};\bar{j})$$

otherwise,

(ii)

$$\sum_{n=-S_T+1+\bar{i}}^{S_T-1-\bar{j}} w((n-\bar{i})/S_T;h-1)\hat{R}_{\Delta\eta\Delta\zeta}(n)$$
$$= -S_T^{-2}\sum_{n=-S_T+2+\bar{i}}^{S_T-2-\bar{j}} w((n-1-\bar{i})/S_T;h+1)\hat{R}_{\eta\zeta}(n)$$
$$\Big(1+O(S_T^{-1})\Big)^2+\sum_{i=1}^{2}\hat{\psi}_{i;T}(h;1;1;1;\bar{i};\bar{j})+S_T^{-1}(-1)$$
$$\times\sum_{i=3}^{4}\tilde{\psi}_{i;T}(h+1;1;1;\bar{i};\bar{j})\Big(1+O(S_T^{-1})\Big)$$

if $w(x;h)$ *is defined and is not constant over* $[-1, 1]$, *and*

$$\sum_{n=-S_T+1+\bar{i}}^{S_T-1-\bar{j}} w((n-\bar{i})/S_T;h-1)\hat{R}_{\Delta\eta\Delta\zeta}(n)$$
$$= \sum_{i=1}^{2}\tilde{\psi}_{i;T}(h;1;1;1;\bar{i};\bar{j})$$
$$+S_T^{-1}(-1)\sum_{i=3}^{4}\tilde{\psi}_{i;T}(h+1;1;1;\bar{i};\bar{j})\Big(1+O(S_T^{-1})\Big)$$

if $w(x;h)$ *is defined and is constant over* $[-1, 1]$.

Proof Noting that

$$\sum_{t=n+2}^{T}\eta_{t-1}\zeta'_{t-n}/T$$
$$= \sum_{t=n+2}^{T}\eta_t\zeta'_{t-n+1}/T+\Big(\eta_{n+1}\zeta'_2-\eta_T\zeta'_{T-n+1}\Big)/T$$

for *n* such that $S_T-1-\bar{j}\geq n\geq 1$, it follows that

$$\sum_{n=-S_T+1+\bar{i}}^{S_T-1-\bar{j}} w((n-\bar{i})/S_T;h-1)\hat{R}_{\Delta\eta\zeta}(n)$$
$$= \sum_{n=-S_T+1+\bar{i}}^{S_T-1-\bar{j}} w((n-\bar{i})/S_T;h-1)\Big(\hat{R}_{\eta\zeta}(n)$$
$$-\hat{R}_{\eta\zeta}(n-1)\Big)-\sum_{j=0}^{S_T-1-\bar{j}} w((j-\bar{i})/S_T;h-1)$$
$$\Big(\eta_{j+1}\zeta'_2-\eta_T\zeta'_{T-j+1}\Big)/T$$
$$= -\sum_{n=-S_T+1+\bar{i}}^{S_T-2-\bar{j}}\{w((n+1-\bar{i})/S_T;h-1)$$
$$-w((n-\bar{i})/S_T;h-1)\}\hat{R}_{\eta\zeta}(n)+\tilde{\psi}_{1;T}(h;1;1;0;\bar{i};\bar{j})$$
$$-w((-S_T+1)/S_T;h-1)\hat{R}_{\eta\zeta}(-S_T+\bar{i})+w((S_T-$$
$$1-\bar{j})/S_T;h-1)\hat{R}_{\eta\zeta}(S_T-1-\bar{j}).$$

Note that

$$-\sum_{n=-S_T+1+\bar{i}}^{S_T-1-\bar{j}}\{w((n+1-\bar{i})/S_T;h-1)-w((n-\bar{i})/$$
$$S_T;h-1)\}\hat{R}_{\eta\zeta}(n)$$

above is zero if $w(x;h-1)$ is constant over $[-1, 1]$. It is also trivial that

$$w((n+1-\bar{i})/S_T;h-1)-w((n-\bar{i})/S_T;h-1)$$
$$= S_T^{-1}w((n-\bar{i})/S_T;h)\Big(1+O(S_T^{-1}\Big)$$

if it is not so. These results immediately leads to the first relation of (i). Similarly, noting that

$$\sum_{t=n+2}^{T}\eta_{t-n}\zeta'_{t-1}/T$$
$$= \sum_{t=n+2}^{T}\eta_{t-n+1}\zeta'_t/T+\Big(\eta_2\zeta'_{n+1}-\eta_{T-n+1}\zeta'_T\Big)/T$$

for *n* such that $S_T-1-\bar{i}\geq n\geq 1$,

$$\sum_{n=-S_T+1+\bar{i}}^{S_T-1-\bar{j}} w((n-\bar{i})/S_T;h-1)\hat{R}_{\eta\Delta\zeta}(n)$$

$$= \sum_{n=-S_T+1+\bar{i}}^{S_T-1-\bar{j}} w((n-\bar{i})/S_T;h-1)\Big(\hat{R}_{\eta\zeta}(n) - \hat{R}_{\eta\zeta}(n+1)\Big) - \sum_{j=0}^{S_T-1-\bar{i}} w((-j-\bar{i})/S_T;h-1) \Big(\eta_2\zeta'_{j+1}-\eta_{T-j+1}\zeta'_T\Big)/T$$

$$= \sum_{n=-S_T+2+\bar{i}}^{S_T-1-\bar{j}} \{w((n-\bar{i})/S_T;h-1)-w((n-1-\bar{i}) /S_T;h-1)\}\hat{R}_{\eta\zeta}(n) + \tilde{\psi}_{3;T}(h;1;1;\bar{i};\bar{j}) - w((S_T -1-\bar{j})/S_T;h-1)\hat{R}_{\eta\zeta}(S_T-\bar{j}) + w((-S_T+1)/ S_T;h-1)\hat{R}_{\eta\zeta}(-S_T+1+\bar{i}).$$

Thus the second relation of (i) can be derived in the same way as for the first relation of (i).

Turn to the proof for (ii). By the first relation of (i) in which ζ is replaced by $\Delta\zeta$,

$$\sum_{n=-S_T+1+\bar{i}}^{S_T-1-\bar{j}} w((n-\bar{i}/S_T;h-1)\hat{R}_{\Delta\eta\Delta\zeta}(n)$$

$$= -S_T^{-1}\sum_{n=-S_T+1+\bar{i}}^{S_T-2-\bar{j}} w((n-\bar{i})/S_T;h)\hat{R}_{\eta\Delta\zeta}(n) \Big(1+O(S_T^{-1})\Big) + \sum_{i=1}^{2}\tilde{\psi}_{i;T}(h;1;1;1;\bar{i};\bar{j}),$$

noting that $\Sigma_{i=1}^{2}\tilde{\psi}_{i;T}(h;1;1;0;\bar{i};\bar{j})$ must be also replaced by $\Sigma_{i=1}^{2}\tilde{\psi}_{i;T}(h;1;1;1;\bar{i};\bar{j})$. The required result is led to by the second relation of (i), regarding h in the above relation as $h-1$.

We will find the probability order of the kernel estimator constructed based on $I(-1)$ series by combining Lemma A.5 with A.2 and note that the essentials of this lemma is obtained in the course of the proof of Lemma 8.1 of Phillips (1995)
or by some application of the argument used there.

Proof of Theorem 1 First, let M_{ZZ} and $\breve{\Omega}$ denote the quantities obtained by replacing $\hat{\beta}$ and $\hat{\gamma}$ with β and γ in $\hat{M}_{ZZ}$ and $\hat{\Omega}$ respectively, noting that $\hat{\gamma}$ in $\tilde{R}_{ij}$ is also replaced with γ. It follows from Lemmas A.1 to A.5 and the essential results above that

$$\breve{\Omega} = \Omega + \tilde{\Pi}, \tag{A.10}$$

with $\tilde{\Pi}$ in the statement of the theorem. Next, let α denote a $k\times r$ column full rank constant matrix such that $C(1)\alpha=0$.

If $\bar{s}=k$ (i.e., Condition (A) holds), we can let

$$\begin{bmatrix} \beta' C^{(1)}(1) \\ \gamma' C(1) \end{bmatrix} = P \begin{bmatrix} \alpha'\Omega_{00}^{-1} \\ \gamma' C(1) \end{bmatrix},$$

with a $k\times k$ full rank matrix P. Note that $\{\alpha'\Omega_{00}^{-1}\varepsilon_t\}$ and $\{\gamma' C(1)\varepsilon_t\}$ are independent. Based on the well-known results on Brownian motions (e.g., Hamilton (1994, p. 548)), we have

$$\begin{bmatrix} \beta' C^{(1)}(1) \\ \gamma' C(1) \end{bmatrix} (\sum_{h=1}^{t}\epsilon_h/\sqrt{T}) = P \begin{bmatrix} \alpha'\Omega_{00}^{-1} \\ \gamma' C(1) \end{bmatrix} (\sum_{h=1}^{t}\epsilon_h/\sqrt{T}) \Rightarrow \Omega^{1/2}\begin{bmatrix} W_r(t/T) \\ W_s(t/T) \end{bmatrix},$$

with Ω in Section 2, noting that

$$P\begin{bmatrix} \alpha'\Omega_{00}^{-1} \\ \gamma' C(1) \end{bmatrix}\Omega_{00}\Big[\Omega_{00}^{-1}\alpha,\ C(1)'\gamma\Big]P' = \Omega$$

and that $W_r(t/T)$ and $W_s(t/T)$ are independent. Similarly,

$$\begin{bmatrix} \gamma'\{\sum_{h=1}^{t}\left(\sum_{i=1}^{h}\epsilon_i\right)/T^{2/3}\} \\ 1 \\ t/T \\ (t/T)^2 \end{bmatrix} \Rightarrow \bar{\Omega}^{1/2}\begin{bmatrix} \int_0^{t/T} W_s(v)dv \\ 1 \\ t/T \\ (t/T)^2 \end{bmatrix},$$

where

$$\bar{\Omega} = \begin{bmatrix} \gamma' C(1)\Omega_{00}C(1)'\gamma & 0 \\ 0 & I_3 \end{bmatrix},$$

regardless of whether or not $\bar{s}=k$ holds. Note that

$$T^{-2/3}\sum_{h=1}^{t}\left(\sum_{i=1}^{h}\epsilon_i\right) = \sum_{h=1}^{t}(t+1-h)\epsilon_h/T^{2/3}.$$

We shall now state several essential results for $I(2)$, $I(1)$ and $I(0)$ series with deterministic trends, not depended upon whether or not $\bar{s}=k$. Letting $R_{\xi}=E\xi_t\xi_t'$ with ξ_t in Lemma A.1,

$$\bar{D}_T = \begin{bmatrix} I_k & 0 \\ 0 & I_k/T \end{bmatrix},$$

and $\bar{M}_{ij;n;n}=\Sigma_{t=n+2}^{T}\bar{Z}_{i;t}\bar{Z}_{j;t}'/T$, $i, j=1, 2, 3, 4$, $n=0, 1, \cdots, S_T$, with

$$\bar{Z}_{1;t} = \left(\sum_{h=1}^{t}(t+1-h)\epsilon_h' C(1)'\hat{\gamma},\ 1,\ t,\ t^2\right)',$$
$$\bar{Z}_{2;t} = \left(\sum_{h=1}^{t}(t+1-h)\epsilon_h' C(1)'\gamma,\ 1,\ t,\ t^2\right)',$$
$$\bar{Z}_{3;t} = \left(y_t',\ \sum_{h=1}^{t} y_h'\right)',$$
$$\bar{Z}_{4;t} = \left(\sum_{h=1}^{t}\epsilon_h' C^{(1)}(1)',\ \sum_{h=1}^{t}(t+1-h)\epsilon_h' C(1)'\right)',$$

$$\sum_{t=n+2}^{T}\xi_t\xi_t'/T = R_\xi + O_p(T^{-1/2}),$$
$$\left(\sum_{t=n+2}^{T}\xi_t\bar{Z}_{4;t}'/T\right)\bar{D}_T = O_p(1),$$
$$\left(\sum_{t=n+2}^{T}\xi_t\bar{Z}_{2;t}'/T\right)\bar{M}_{22;n;n}^{-1}\left(\sum_{t=n+2}^{T}\bar{Z}_{2;t}\xi_t'/T\right) = O_p(T^{-1}),$$
$$\left(\sum_{t=n+2}^{T}\xi_t\bar{Z}_{2;t}'/T\right)\bar{M}_{22;n;n}^{-1}\bar{M}_{24;n;n}\bar{D}_T = O_p(T^{-1/2}),$$
$$\bar{D}_T\left(\bar{M}_{42;n;n}\bar{M}_{22;n;n}^{-1}\bar{M}_{24;n;n}/T\right)\bar{D}_T = O_p(1),$$

which are included essentially in the asymptotics on $I(1)$ established in past papers or shown by a simple application of those, e.g., Park and Phillips (1988, 1989) or Hamilton (1994, p. 548). It is also obvious in view of Lemma 4 that

$$\begin{aligned}&\hat{B}'\bar{D}_T\left(\bar{M}_{33;0;0}/T - \bar{M}_{31;0;0}\bar{M}_{11;0;0}^{-1}\bar{M}_{13;0;0}/T\right)\bar{D}_T\hat{B}\\ =\ &\bar{B}'\bar{D}_T\left(\bar{M}_{44;0;0}/T - \bar{M}_{42;0;0}\bar{M}_{22;0;0}^{-1}\bar{M}_{24;0;0}/T\right)\bar{D}_T\bar{B}\\ &+O_p(T^{-1}),\end{aligned}$$

putting

$$\hat{B}' = \begin{bmatrix}\hat{\beta}' & 0\\ 0 & \hat{\gamma}'\end{bmatrix},\qquad \bar{B}' = \begin{bmatrix}\beta' & 0\\ 0 & \gamma'\end{bmatrix}.$$

In view of the expression for $((\Sigma_{h=1}^{t} y_h)'\beta, y_t'\gamma)'$ in the proof of Lemma 2, the limiting behaviors based on Brownian motions above, together with the essential results, lead to

$$M_{ZZ}/T \Rightarrow PF^{1/2}\left(\bar{V}_k - \tilde{V}_{k;s}\breve{U}_{s;s}^{-1}\tilde{V}_{k;s}'\right)F^{1/2}P'. \quad (A.11)$$

It is also easy to find $[P_1, P_2]$ claimed in the statement of the theorem for the case $\bar{s} < k$. Then the expression for $((\Sigma_{h=1}^{t} y_h)'\beta, y_t'\gamma)'$ becomes

$$\begin{bmatrix}P_1'\\ P_2'\end{bmatrix}^{-1}\begin{bmatrix}0\\ \bar{P}_2'(\sum_{h=1}^{t}\epsilon_h)\end{bmatrix} + \begin{bmatrix}P_1'\\ P_2'\end{bmatrix}\bar{u}_t,$$

with $\bar{u}_t$ in the statement of Theorem 1 (ii) and

$$\bar{P}_2' = P_2'\begin{bmatrix}\beta' C^{(1)}(1)\\ \gamma' C(1)\end{bmatrix},$$

provided that the terms on the initial vectors or deterministic trends are out of consideration. By a similar argument to one used for the case $\bar{s}=k$,

$$\bar{P}_2'(\sum_{h=1}^{t}\epsilon_h/\sqrt{T}) = P_2'P\begin{bmatrix}\alpha'\Omega_{00}^{-1}\\ \gamma' C(1)\end{bmatrix}(\sum_{h=1}^{t}\epsilon_h/\sqrt{T}) \Rightarrow \breve{\Omega}_2^{1/2}W_{\bar{s}}(t/T),$$

with $\breve{\Omega}_2 = P_2'\Omega P_2$. It follows from these results with the essential ones stated above that

$$\begin{bmatrix}T^{1/2}P_1'\\ P_2'\end{bmatrix}M_{ZZ}/T\left[T^{1/2}P_1,\ P_2\right]$$
$$=\begin{bmatrix}\sum_{t=2}^{T}P_1'\bar{u}_t\bar{u}_t'P_1/T + O_p(T^{-1}) & O_p(T^{-1/2})\\ O_p(T^{-1/2}) & \sum_{t=2}^{T}\bar{P}_2'(\sum_{h=1}^{t}\epsilon_h)(\sum_{h=1}^{t}\epsilon_h')\bar{P}/T^2 + O_p(T^{-1})\end{bmatrix}$$
$$-\{\begin{bmatrix}0\\ \sum_{t=2}^{T}\bar{P}_2'(\sum_{h=1}^{t}\epsilon_h)\end{bmatrix}\bar{Z}_{2;t}'/T^{3/2}\}$$
$$\times\bar{M}_{22;0;0}^{-1}\{\sum_{t=2}^{T}\bar{Z}_{2;t}\left[0, (\sum_{h=1}^{t}\epsilon_h')\bar{P}_2\right]/T^{3/2}\} + O_p(T^{-1/2})$$
$$=\begin{bmatrix}P_1'R_{\bar{u}}P_1 + O_p(T^{-1/2}) & O_p(T^{-1/2})\\ O_p(T^{-1/2}) & \bar{W}_{22} + O_p(1)\end{bmatrix},$$

with

$$\begin{aligned}\bar{W}_{22} =\ &\bar{P}_2'\{\sum_{t=2}^{T}(\sum_{h=1}^{t}\epsilon_h)(\sum_{h=1}^{t}\epsilon_h')/T^2\}\bar{P}_2\\ &-\bar{P}_2'\{\sum_{t=2}^{T}(\sum_{h=1}^{t}\epsilon_h)\bar{Z}_{2;t}'/T^{3/2}\}\\ &\bar{M}_{22;0;0}^{-1}\{\sum_{t=2}^{T}\bar{Z}_{2;t}(\sum_{h=1}^{t}\epsilon_h'/T^{3/2}\}\bar{P}.\end{aligned}$$

Noting that $P_1'\Omega = 0$, it is easy to see from $(A.10)$ that

$$\begin{bmatrix}T^{1/2}P_1'\\ P_2'\end{bmatrix}\breve{\Omega}\left[T^{1/2}P_1,\ P_2\right]$$
$$=\begin{bmatrix}0 & 0\\ 0 & P_2'\Omega P_2\end{bmatrix} + \begin{bmatrix}TP_1'\tilde{\Pi}P_1 & T^{1/2}P_1'\tilde{\Pi}P_2\\ T^{1/2}P_2'\tilde{\Pi}P_1 & P_2'\tilde{\Pi}P_2\end{bmatrix}.$$

Consequently,

$$\begin{bmatrix}T^{1/2}P_1'\\ P_2'\end{bmatrix}\breve{\Omega}\,(M_{ZZ}/T)^{-1}\begin{bmatrix}T^{1/2}P_1'\\ P_2'\end{bmatrix}^{-1}$$
$$=\begin{bmatrix}TP_1'\tilde{\Pi}P_1\,(P_1'R_{\bar{u}}P_1)^{-1} & T^{1/2}P_1'\tilde{\Pi}P_2\left(\bar{W}_{22}\right)^{-1}\\ T^{1/2}P_2'\tilde{\Pi}P_1\,(P_1'R_{\bar{u}}P_1)^{-1} & P_2'\left(\Omega + \tilde{\Pi}P_2\right)\left(\bar{W}_{22}\right)^{-1}\end{bmatrix}$$
$$+O_p(T^{-1/2}).$$

Noting that $\tilde{\Pi} = O_p(S_T^{1/2}/T^{1/2})$ in view of the statement of the thereom, it is established that

$$tr\{\breve{\Omega}\,(M_{ZZ}/T)^{-1}\}$$
$$= Ttr\{P_1'\tilde{\Pi}P_1\,(P_1'R_{\bar{u}}P_1)^{-1}\} + O_p(S_T^{1/2}). \qquad (A.12)$$

On the other hand, using the essential results and Lemma 4 with the expression for $((\Sigma_{h=1}^{t} y_h)'\beta, y_t'\gamma)'$,

$$\hat{M}_{ZZ}/T = M_{ZZ}/T + O_p(T^{-1}), \qquad (A.13)$$

regardless of whether or not $\bar{s} < k$. Similarly,

$$\hat{\Omega} = \breve{\Omega} + O_p(S_T/T), \qquad (A.14)$$

recalling that $\Sigma_{h=-T+2}^{T-2}$ in $\hat{\Omega}$ actually becomes summation over $-S_T \leq t \leq S_T$.

Putting $(A.10)$ to $(A.14)$ together, the results required for the theorem are established.

Proof of Theorem 2 We will chcek that this theorem can be proved essentially along the course of the proof of Shintani's Theorem 3.1 and using its related arguments except for the minor points stated below. First, $\tilde{P}(h)$ is constructed based on demeaning and detrending unlike Shintani's $P(n, s)$.
However, it does not make much difference to the result. A similar matter is applied to our band-width parameter S_T given in a more general form than Shintani's one. We also note in (ii) that the order of $\beta'\breve{\Pi}\beta$ is established by arguments, similar to of $P_1'\tilde{\Pi}P_1$ in Theorem 1 (ii). This term may be expressed in a more specific form if $rank\,\beta' C^{(1)}(1) = r$ holds, which is necessary for Condition (A), as in Shintani (see (5) in p. 341 and p. 356). However, it is not assumed in this paper and it should be emphasized that Condition (A) is not needed for Shintani's Theorem 3.1 except for this matter. The important matter is that the degeneracy property of $\Sigma_{n=-T+2}^{T-2} k(n/S_T)\hat{R}_{00}(n)$ or $\hat{R}_{11}/T$ is caused by only β (or a matrix spanned by β) and that P_1 in Theorem 1 never affect it.

Table 1

5% and 1% Percentage Points of $\Psi_{k;s}$

$k=2, s=1$		$k=3, s=1$		$k=3, s=2$		$k=4, s=1$	
5%	1%	5%	1%	5%	1%	5%	1%
116.79	140.256	179.233	136.544	145.03	160.671	134.521	165.341
$k=4, s=2$		$k=4, s=3$		$k=5, s=1$		$k=5, s=2$	
5%	1%	5%	1%	5%	1%	5%	1%
116.79	140.256	179.233	136.544	145.03	160.671	134.521	165.341
$k=5, s=3$		$k=5, s=4$		$k=6, s=1$		$k=6, s=2$	
5%	1%	5%	1%	5%	1%	5%	1%
116.79	140.256	179.233	136.544	145.03	160.671	134.521	165.341
$k=6, s=3$		$k=6, s=4$		$k=6, s=5$		$k=7, s=1$	
5%	1%	5%	1%	5%	1%	5%	1%
116.79	140.256	179.233	136.544	145.03	160.671	134.521	165.341
$k=7, s=2$		$k=7, s=3$		$k=7, s=4$		$k=7, s=5$	
5%	1%	5%	1%	5%	1%	5%	1%
116.79	140.256	179.233	136.544	145.03	160.671	134.521	165.341
$k=7, s=6$		$k=8, s=1$		$k=8, s=2$		$k=8, s=3$	
5%	1%	5%	1%	5%	1%	5%	1%
116.79	140.256	179.233	136.544	145.03	160.671	134.521	165.341
$k=8, s=4$		$k=8, s=5$		$k=8, s=6$		$k=8, s=7$	
5%	1%	5%	1%	5%	1%	5%	5%
116.79	140.256	179.233	136.544	145.03	160.671	210,344	221.556

Table 2

Empirical Size and Power

Theoretical Size	$T=100$		$T=250$		$T=500$	
	$S_T=4$	$S_T=6$	$S_T=4$	$S_T=6$	$S_T=4$	$S_T=6$
Case 1 (H_0 is true.)						
5%	0.08	0	1.89	0.59	6.36	3.16
1%	0.01	0	0.18	0.04	0.8	0.32
Case 2 (H_0 is true.)						
5%	0.09	0.09	1.33	0.45	3.31	2.1
1%	0.01	0.01	0.06	0.01	0.49	0.15
Case 3 (H_0 is false.)						
5%	0.32	0.17	95.37	25.98	100.0	100.0
1%	0.08	0.09	47.76	1.79	100.0	99.99
Case 4 (H_0 is false.)						
5%	0.22	0.11	90.02	16.25	100.0	100.0
1%	0.04	0.04	30.0	0.83	100.0	99.97
Case 5 (H_0 is true.)						
5%	0.5	0.54	3.49	1.34	10.98	5.46
1%	0.04	0.07	0.24	0.04	1.91	0.69
Case 6 (H_0 is true.)						
5%	0.16	0.14	3.57	0.94	11.19	5.4
1%	0.02	0.03	0.24	0.07	2.14	0.55
Case 7(H_0 is false.)						
5%	0.03	0.01	91.95	19.12	100.0	100.0
1%	0	0	35.9	0.94	100.0	100.0
Case 8 (H_0 is false.)						
5%	0.09	0.01	92.21	17.28	100.0	100.0
1%	0	0	34.16	0.35	100.0	99.96
Case 1* (H_0 is true.)						
5%	1.96	0.03	2.4	0.75	6.42	3.55
1%	0.01	0	0.1	0.01	0.96	0.29
Case 2* (H_0 is true.)						
5%	0.19	0.22	1.36	0.49	3.45	2.12
1%	0	0	0.05	0.01	0.3	0.14
Case 3* (H_0 is false.)						
5%	0.05	0.01	95.68	25.71	100.0	100.0
1%	0	0	46.33	1.42	100.0	100.0
Case 4* (H_0 is false.)						
5%	0.08	0.03	89.8	15.69	100.0	100.0
1%	0.01	0.01	30.38	0.31	100.0	99.97
Case 5* (H_0 is true.)						
5%	0.58	0.76	3.76	1.36	10.93	5.43
1%	0.04	0.08	0.38	0.1	2.09	0.72
Case 6* (H_0 is true.)						
5%	0.26	0.21	3.86	1.3	11.73	6.03
1%	0.01	0	0.39	0.07	2.45	0.87
Case 7* (H_0 is false.)						
5%	0.03	0	91.55	18.82	100.0	100.0
1%	0	0	35.72	0.46	100.0	99.99
Case 8* (H_0 is false.)						
5%	0.1	0.01	92.34	17.93	100.0	100.0
1%	0	0	34.82	0.42	100.0	99.96

Table 3

Relative Frequency of a Correct Decision in association with $\sim\tilde{P}(0)$

Significance Level : 5%				Significance Level : 1%			
T=250		T=500		T=250		T=500	
S_T=4	S_T=6	S_T=4	S_T=6	S_T=4	S_T=6	S_T=4	S_T=6
Case 1 (H_0 is true.)							
99.11	86.71	93.64	96.84	96.83	17.38	99.2	99.68
Case 2 (H_0 is true.)							
78.88	21.59	96.69	99.9	28.51	3.58	99.51	99.85
Case 3 (H_0 is false.)							
95.37	24.81	100.0	100.0	47.41	1.01	100.0	99.99
Case 4 (H_0 is false.)							
90.02	15.11	100.0	100.0	29.52	0.32	100.0	99.98
Case 5 (H_0 is true.)							
96.51	85.58	89.02	94.54	96.08	17.49	98.09	99.31
Case 6 (H_0 is true.)							
96.43	87.1	88.81	94.6	96.88	18.5	97.86	99.45
Case 7 (H_0 is false.)							
91.95	18.2	100.0	100.0	35.69	0.48	100.0	100.0
Case 8 (H_0 is false.)							
92.21	17.28	100.0	100.0	33.8	0.35	100.0	99.96
Case 1* (H_0 is true.)							
97.6	96.71	93.58	96.45	87.03	17.24	99.04	99.71
Case 2* (H_0 is true.)							
78.0	21.46	96.55	97.88	21.46	3.2	99.7	99.86
Case 3* (H_0 is false.)							
95.68	45.99	100.0	100.0	24.78	0.8	100.0	100.0
Case 4* (H_0 is false.)							
89.79	30.0	100.0	100.0	15.69	0.31	100.0	99.97
Case 5* (H_0 is true.)							
96.24	85.87	89.07	94.57	96.09	17.78	97.91	99.28
Case 6* (H_0 is true.)							
96.14	86.89	88.27	93.97	96.62	18.71	97.55	99.13
Case 7* (H_0 is false.)							
91.55	17.87	100.0	100.0	35.72	0.46	100.0	99.99
Case 8* (H_0 is false.)							
92.34	16.83	100.0	100.0	34.56	0.42	100.0	99.96