A TESTING METHOD FOR GRANGER CAUSALITY IN COINTEGRATED TIME SERIES SYSTEMS

Mitsuhiro Odaki

1 Introduction

Since Granger's (1969) definition on causality between economic time series proposed, the so-called Granger causality test has been attracting many researchers in both theoretical and empirical aspects. In 1970's, the causality test was established mainly in the framework of stationary[1)] autoregressive (VAR) processes, based on the approach of Box and Jenkins (1976), and as a result, the standard asymptotic theory[2)] based on asymptotic normality was applied to usual test statistics known as Wald tests. This approach which uses a stationary VAR was verified by the empirical belief that many economic time series are not stationay in levels of the original time series but can be transformed to stationary ones by differencing or detrending.

The remarkble development of statistical inferences for nonstationary time series from the late 1970's suggested the possibility that even the

---

1) It may be appropriate that the term 'stationarity' we use here is refered as 'weak one' than 'strong one'.

2) The most essential requirement to make this theory hold is not 'stationarity' but 'regurality'. See Grenander and Rosenblatt (1957) etc. for example.

Granger causality may be tested using level series under VAR modeling in which some unit toots are postulated to exist. On the other hand, the concept of cointegration formulated by Engle and Granger (1987) etc. revealed the situations in which individual time series considered are integrated of an order (i.e. have nonstationarity caused by one unit root) but some linear combinations of those are stationay. The so-called the Granger Representation Theorem (GRT) of Engle and Granger (1987) states that modelling under such a situation should be done based on either a VAR representation in levels of the orginal series or the error correction model (ECM) derived equivalently from it. In other words, any cointegrated system of time series cannot be transformed to a stationary VAR by differencing. The ECM consists of stationay series only; that is, such linear combinations called cointegarating relations and first differences. Numerous literatures have studied the inferences on cointegration theoretically and empirically, and as a result those on the VAR or ECM framework have got the reputation of being most general and desirable: Sims et al. (1990)'s OLS in the VAR (often referred as the unrestricred VAR) and Johansen (1988)'s ML in the ECM.

Those developments based on the inference theory for nonstationary time series and the concept of cointegration or error correction suggested the possibility that the Granger causality may be tested using level seies under VAR modeling in which some unit roots are postulated to exist, and the consideration of this matter drives us into a new testing procedure for Granger causality. Toda and Phillips (1993) tried to discuss Granger causality tests in the general framework of VAR's in which the presence of unit roots and cointegration is considered. More concretely, they replaced stationay VAR's used in 1970's by VAR's with a unit root (the unrestricted levels VAR's in terms of Sims et al. (1990)) and ECM's, and investigated statistical properties of the Wald statistics to test null hypotheses of the absence of Granger causality under both models. In the paper, the vector time sries system which are the target of the analysis is partitioned into three subsystems and Granger causality between two subsystems of them is dealt with. In levels VAR's whose individual series are supposed to be stationary after differencing once and are thought to be models for level series themselves, the Wald test are simply constructed based on OLS estimation for coefficient parameters. On the other hand, in ECM's the Wald tests are more complex in structure because the null hypothesis is not expressed simply by linear restrictions of the coefficients and it may be redundant in some cases. Estimation for coefficients are made by not OLS but Johansen's (1988) style of ML method, as stated already.

The main purpose of Toda and Phillips (1993) was to evaluate their Wald tests asymptotically. They conclude that under cointegrated VAR system, the Wald tests asymptotically follow χ^2 distribution under the null of the absence of Granger causality if some conditions on the parameters are imposed. That is, their results are restrictive in the sense that there exist many situations which are not valid to use those tests based on χ^2 critical values. Particularly, in the unrestricted levels VAR, the deviation from χ^2 was more serious.

The purpose of the present paper is to establish a valid Granger causality test under the same framework of Toda and Phillips (1993) without imposing any condition on the parameters. Our arguments are developed in ECM's, taking account of that VAR's are redundant on the parameters concerning to the cointegration. Therefore, our testing method will be constructed based on Johansen's (1988) ML estimation. Unlike Toda and Phillips (1993), the system considered is always to be cointegrated since the cases which are not cointegrated corresponds not to ECM's but to stationary VAR's for differenced series and the analysis for such a series is conventional within the framework of Box and Jenkins (1976). The ECM's parameter we must noticeably pay attention to in both Johansen's (1988) ML estimation and our testing method is the cointegrating matrix which consists of all linearly independent cointegrating vectors. Before we get to estimating the cointegrating matrix, the rank of the cointegrating matrix, named *the coitegrating rank*, need to be first decided via some statistical procedure or pretested usually by Johansen's (1988) likelihood ratio test. As some elaboration to actualize our purpose, we introduce another form of the cointegrating matrix derived by an orthogonal transformation of its original one, so that the null hypothesis is expressed by some linear restrictions on the parameters. Along that line, as the estimator of the cointegrating matrix we use not Johansen's (1988) one but an orthogonal transformation of it. One of the characteristics of our testing method is that in order to decide the rank of a submatrix of the matrix transformed from the Johansen's (1988) estimator of the cointegrating matrix, another pretest need to be executed. Given an estimator of the rank, our testing method can be constructed using it as a usual Wald test for a linear hypothesis. We prove that both the pretest to decide the rank and our Granger causality test following it are asymptotically valid as χ^2 criterion. An emphasized matter is that those results are established without imposing any condition on the ECM's parameters.

The remained part of the present paper is as follows. Sections 2 and 3 play the role of some preliminaries, in which the models, notations and assumptions for the paper are introduced to formulate our hypothesis and some fundamental statistics and asymptotic results are also presented. Section 4 discusses the above-mentioned pretests. Our Granger causality test, based on the pretests, is provided in section 5. Asymptotic properties of those are presented in each of sections 4 and 5. In section 6 we give some remarks and implications on the testing method proposed in this paper. Featuring different special cases explains why the asymptotic validity of our method holds unconditionally unlike Toda and Phillips' (1993) one. Section 7 concludes the paper. Some of the proofs for theorems and lemmas presented in this paper are placed in the appendix.

2 The model and formulation of Granger noncausality

In what follows, we shall consider the k-dimensional vector time series $\{y_t\}$ generated by the p-th order VAR model, which is supposed to be the same one as that of Toda and Phillips (1993) except for some differences in notation, i. e.

$$y_t = \sum_{j=1}^{p} A_j y_{t-j} + \varepsilon_t \quad \forall t \geq 1, \tag{1}$$

where $\{\varepsilon_t\}$ is an iid sequence of k-dimensional random vectors with mean zero and nonsingular covariance matrix Λ, such that $E \mid \varepsilon_{it} \mid^{2+\delta} < \infty$ for some $\delta > 0$ with the i-th element of ε_t, ε_{it}, $i = 1, ..., k$. It is also assumed that Δy_{it} is stationary (in weak sense) for any $i = 1, ..., k$, where y_{it} is the i-th element of y_t and Δ denotes the first differenced operator, i. e. $\Delta y_{it} = y_{it-1}$. Defining

$$A(\lambda) = I_k - \sum_{j=1}^{p} A_j \lambda^j,$$

with the notation I_n denoting the $n \times n$ identity matrix for any positive integer n and a complex number λ, this implies that the root of det A $(\lambda) = 0$ is confined to one such that either $\mid \lambda \mid > 1$ or $\lambda = 1$, and requires some initial condition on y_t. We shall impose an assumption on the intial values $\{y_0, y_{-1}, ..., y_{-p+1}\}$ so that those do not affect asymptotic results; for example, those initial values are supposed to be constant vectors. Furthermore, the present paper assumes that y_t is cointegrated in the sense of Engle and Granger's (1987) formulation. (1) is the data-generating process for this paper.

As shown in Engle and Granger (1987) or Johansen (1988), from the VAR representation as the data-generating process (1), we can derive the equivalent ECM

$$\Delta y_t = \sum_{j-1}^{p-1} \Pi_j (\Delta y_{t-j}) - A(1) y_{t-1} + \varepsilon_t, \quad \forall t \geq 1, \qquad (2)$$

where $\Pi_j = -\sum_{m=j+1}^{p} A_m, j = 1, ..., p-1$, and $A(1) = I_k - \sum_{j=1}^{p} A_j$. With (2), we make other assumptions embodying that y_t is cointegrated, which are also put in Toda and Phillips (1993):

$$A(1) = -\alpha \beta' \qquad (3)$$

where α and β[3] are $k \times r$ matrices of full column rank r, with $1 \leq r \leq k-1$.

$$rank\ \delta' \Pi(1) \gamma = s \equiv k - r, \qquad (4)$$

where r is given as the rank of α or β in (3) and δ and γ are $k \times s$ matrix of full column rank s such that

$$\delta' \alpha = 0 = \gamma', \beta, \quad \Pi(1) = I_k - \sum_{j-1}^{p-1} \Pi_j.$$

(3) and (4) just imply that the cointegrating rank and the cointegrating matrix are r and β respectively. Throughout the paper, r (or s) is assumed to known. Notice that a statistical method 'to estimate' r is a consecutive application of Johansen's (1988) likelihood ratio tests we mentioned already.

Now, consider how Granger noncausality can be expressed by the parameters in (2). Suppose that our interest is in Granger causality from the first k_1 elements of y_t to the last k_3 elements of y_t, with positive integers k_1, k_3 and a nonnegative integer k_2 such that $k_1 + k_2 + k_3 = k$. For convenience, we introduce some notations to partition the system; that is,

---

3) The decomposition of $A(1)$ into α and β is not unique. The expression of β can be freely chosen in later discussion as long as it satisfies the conditions stated in (3).

letting $k_{**} = k_1 + k_2$ and $k_{++} = k_2 + k_3$, write

$$y_t = \begin{bmatrix} y_t(1) \\ y_t(2) \\ y_t(3) \end{bmatrix} = \begin{bmatrix} y_t(**) \\ y_t(3) \end{bmatrix}, \varepsilon_t = \begin{bmatrix} \varepsilon_t(1) \\ \varepsilon_t(2) \\ \varepsilon_t(3) \end{bmatrix},$$

with $k_i \times 1$ vectors $y_t(i)$ and $\varepsilon_t(i)$, $i = 1, 2, 3$, and a $k_{**} \times 1$ vector $y_t(**)$,

$$\beta' = [\beta'_{**} \beta'_3], \quad \alpha = [\alpha'_1, \alpha'_{++}]', \quad \gamma' = [\gamma'_{**}, \gamma'_3],$$

where β'_{**} is $r \times k_{**}$, β'_3 is $r \times k_3$, α_1 is $k_1 \times r$, α_{++} is $k_{++} \times r$, γ'_{**} is $r \times k_{**}$ and γ'_3 is $r \times k_3$,

$$\Pi_j = \begin{bmatrix} \Pi_j(1, **) & \Pi_j(1, 3) \\ \Pi_j(++, **) & \Pi_j(++, 3) \end{bmatrix}, \ j = 1, \ldots, p\text{-}1,$$

where $\Pi_j(1,**)$ is $k_1 \times k_{**}$, $\Pi_j(1, 3)$ is $k_1 \times k_3$, $\Pi_j(++,**)$ is $k_{++} \times k_{**}$ and $\Pi_j(++, 3)$ is $k_{++} \times k_3$. We must note here that the partition of y_t into two subsystems, i. e. the case in which $y_{**} = y_{1t}$ (therefore $k_{**} = k_1$), is also included in the above partition of y_t into three systems. Using parts of the above notations, the first k_1 equations of (2) are written as

$$\Delta y_t(1) = \sum_{j=1}^{p-1} \Pi_j(1, **), \Delta y_{t-j}(**) + \alpha_1 \beta'_{**} y_{t-1}(**)$$

$$+ \sum_{j-1}^{p-1} \Pi_j(1, 3) \Delta y_{t-j}(3) + \alpha_1 \beta'_3 y_{t-1}(3) + \varepsilon_t(1),$$

for all $t \geq 1$. Moreover, this is rewritten as

$$\Delta y_t(1) = G' x_t(**) + F' x_t(3) + \varepsilon_t(1), \quad \forall t \geq 1, \tag{5}$$

where

$G' = [\Pi_1(1, **), \ldots, \Pi_{p-1}(1, **), \alpha_1 \beta'_{**}]$, $x_t(**) = [\Delta y'_{t-1}(**), \ldots, \Delta y'_{t-p+1}(**) y'_{t-1}(**)]$,

$F' = [\Pi_1(1, 3), \ldots, \Pi_{p-1}(1, 3), \alpha_1 \beta'_3]$, $x_t(3) = [\Delta y'_{t-1}(3), \ldots, \Delta y'_{t-p+1}(3) y'_{t-1}(3)]$.

Following the notion proposed by Granger (1969), the null hypothesis of the absence of causality from y_t (1) to y_t (3) is formulated as

$$H_0 : F = 0. \tag{6}$$

Notice that $F = 0$ is a nonlinear hypothesis on parameters because it contains $\alpha_1\ \beta_3' = 0$ and such 'nonlinearity' causes difficulty for Granger causelity test which Toda and Phillips (1993) has just faced.

Now, we shall discuss another form of the cointegrating matrix and another formulation for the null based on it. Hereafter, let the notation $diag\ \{\nu_1, ..., \nu_m\}$ denote a $m \times m$ diagonal matrix which possesses ν_i as the i-th diagonal element. Let $\bar{L}$ and λ_i, $i = 1, ..., r$, be an orthogonal matrix of $r \times r$ and real numbers such that

$$\bar{L}'\beta_3'\beta_3\bar{L} = diag\ \{\lambda_1, ..., \lambda_r\},\ 0 \leq \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_r. \tag{7}$$

It is ensured that there exist such $\bar{L}$ and λ_i, $i = 1, ..., r$, because $\lambda_j, j = 1, ..., i$, are the i smallest eigenvalues of $\beta_3'\beta_3$, $1 \leq i \leq r$, and the columns of $\bar{L}$ are the corresponding eigenvectors. Let us write the rank of β_3 as r_3. Then it is confirmed that $0 \geq r_3 \geq m_1$, with $m_1 = min\ \{r, k_3\}$. Also, let

$$r_{**} = r - r_3,\ \bar{\beta} = \beta\bar{L},\ \bar{\beta}_{**} = \beta_{**}\bar{L},$$
$$\bar{\beta}_3 = \beta_3\bar{L},\ \ \bar{\alpha} = \alpha\bar{L},\ \bar{\alpha}_1 = \alpha_1\bar{L}.$$

Furthermore, note that $r \geq r_{**} \geq m_2$, with $m_2 = max\ \{0, k_3 - r\}$. Note that the cointegrating rank r is not only the rank of β but also that of $\bar{\beta}$, i. e. $rank\,\bar{\beta} = r$. For the case in which $0 < r_3 < r$, note that $\lambda_i = 0$, $i = 1, ...\ r_{**}$ and $\lambda_i > 0$, $i = r_{**} + 1, ..., r_{**} + r_3 = r$. Therefore, for this case, we can partition $\bar{\beta}_{**}$, $\bar{\beta}$ and $\bar{\alpha}_1$ as

$$\bar{\beta}_{**} = [\bar{\beta}_{**},\ _{**},\ \bar{\beta}_{**,3}],\ \ \bar{\beta}_3 = [0, \bar{\beta}_{3,3}],\ \ \bar{a}_1 = [\bar{a}_{1,**}\},\ \bar{a}_{1,3}], \tag{8}$$

with $rank, \bar{\beta}_{3,3} = r_3$, where $\bar{\beta}_{**,**}$ is $k_{**} \times r_{**}$, $\bar{\beta}_{**,3}$ is $k_{**} \times r_3$, $\bar{\beta}_{3,3}$ is $K_3 \times r_3$, $\bar{a}_{1,**}$

is $k_1 \times r_{**}$ and $\bar{a}_{1,3}$ is $k_1 \times r_3$. $rank\bar{\beta}_{3,3} = r_3$ requires $r_3 \leq k_3$.

Some careful consideration for r_{**} need be done. Since $rank\bar{\beta} = rank\ \beta = r$, all the column vectors are linearly independent of $\bar{\beta}$, the first r_{**} column vectors of $\bar{\beta}$ for the case in which $0 < r_3 < r$ are also linearly independent. In other words, the $k \times r_{**}$ matrix $[\bar{\beta}'_{**,**}, 0]'$ constructed by those column vectors is the column-full rank. This requires that $rank\ \beta_{**,**} = r_{**}$, therefore $r_{**} \leq k_{**}$. The thing noted in this case is that for any $k_{**} \times 1$ vector b which is linearly independent from all the column vectors in $\bar{\beta}_{**,**}$, $b'\ y_t\ (**)$ is not stationary, since all the cointegrating relations are expressed as

$$\bar{\beta}' y_t = [\bar{\beta}'_{**}, \bar{\beta}'_3]\ y_t = \begin{bmatrix} \bar{\beta}_{**,**} & 0 \\ \bar{\beta}_{**,3} & \bar{\beta}_{3,3} \end{bmatrix} \begin{bmatrix} y_t(**) \\ y_t(3) \end{bmatrix}.$$

It is obvious that the case in which $r_3 = 0$ corresponds to the case in which $\bar{\beta}_3 = 0$. Under this case, let us define $y_t = y_t\ (**)$ and $\bar{\beta} = \bar{\beta}_{**,**}$. This case confines the cointegrating relations to one spanned by $\bar{\beta}'_{**,**}\ y_t\ (**)$. From the above notices, we can assert that r_{**} is the cointegrating rank of the subsystem $\{y_t\ (**)\}$ and therefore $r_{**} \leq k_{**} - 1$. Hereafter, let us define $\bar{\beta}_{**,3}$, $\bar{\beta}_{3,3}$ and $\bar{\alpha}_{1,3}$ for the case in which $r_3 = r$ (in other words, $r_{**} = 0$) as $\bar{\beta}_{**,3} = \bar{\beta}_{**}$, $\bar{\beta}_{3,3} = \bar{\beta}_3$ and $\bar{\alpha}_{1,3} = \bar{\alpha}_1$.

It is easy to check that

$$\alpha_1\ \bar{\beta}'_3 = \alpha_1\ \bar{L}\bar{L}'\ \beta'_3 = \bar{\alpha}_1\bar{\beta}'_3 = \bar{\alpha}_{1,3}\bar{\beta}'_{3,3},$$

which in turn says that $\alpha_1\ \beta'_3 = 0$ is equivalent to $\bar{\alpha}_{1,3} = 0$ unless $r_3 = 0$, recalling that $rank\bar{\beta}'_{3,3} = r_3 \geq k_3$. Now define A and B as

$$A' = [\Pi_1\ (1, 3), ..., \Pi_{p\text{-}1}\ (1, 3),\ \bar{\alpha}_{1,3}]\quad if\ r_3 > 0,$$

$$A' = [\Pi_1\ (1, 3), ..., \Pi_{p\text{-}1}\ (1, 3)\quad if\ r_3 = 0,$$

$$B' = [\Pi_1\ (1,), ..., \Pi_{p\text{-}1}(1,**),\ \bar{\alpha}_1],$$

$$\bar{x}_t\ (**) = [\Delta y'_{t\text{-}1}\ (**), ..., \Delta y'_{t\text{-}p+1}\ (**),\ y'_{t\text{-}1}\ (**)\bar{\beta}_{**}]',$$

$$\bar{x}_t(3) = [\Delta y'_{t-1}(3), ..., \Delta y'_{t-p+1}(3), y'_{t-1}(3)\bar{\beta}_{3,3}]',$$

using $\Pi_j(1, 3)$, $\Pi_j(1,**)$, $y_{t-j}(**)$ and $y_{t-j}(3)$ given already. Then (5) is rewritten as

$$\Delta_{yt}(1) = B'\bar{x}_t(**) ++ A'\bar{x}_t(3) + \varepsilon_t(1), \quad \forall_t \geq 1. \tag{9}$$

Under (9) the formulation for the null of Granger noncausality is naturally expressed as

$$H_0 : A = 0. \tag{10}$$

Notice that the hopothesis $A = 0$ is 'linear' unlike $F = 0$. It should be also noticed that in the case in which $r_3 = 0$, we do not need to care cointegration or more concretely the error correction terms to formulate/test Granger noncausality. Thus we see that to check whether $r_3 = 0$ holds or not before testing Granger causlity is particularly important.

3 Some statistics and fundamental asymptotic results

This section discusses some statistics and related asymptotic properties needed for both the pretest and Granger causality proposed in this paper. First, suppose that the symbols $\rightarrow_p$ and $\rightarrow_d$ denote convergence in probability and convergence in distribution respectively as the sample size T goes to ∞. Also, let $\hat{\beta}$ drenote the Johansen's (1988) type of ML estimator[4)] of the cointegrating matrix. That is, suppose that $[y_1, ..., y_T]$ is given as the observations of y_t with the sample size T. Using the notations on observable matrices expressed as

4) The ECM in this paper (2) is slightly different from that of Johansen (1988) and therefore $\hat{\beta}$ is also so. However, the difference can be negligible and we can consider that our estimator is essentially the same as that in Johansen (1988).

$$\Delta Y = \begin{bmatrix} \Delta y'_{p+1} \\ \vdots \\ \Delta y'_T \end{bmatrix}, \; Y_{-1} = \begin{bmatrix} y'_p \\ \vdots \\ y'_{T-1} \end{bmatrix}, \; W_{-1} = \begin{bmatrix} \Delta y'_p & \dots & \Delta y'_2 \\ \vdots & \ddots & \vdots \\ \Delta y'_{T-1} & \dots & \Delta y'_{T-p+1} \end{bmatrix},$$

$$M_W = I_{T-p} - W_{-1}(W'_{-1}W_{-1})^{-1}W'_{-1}, \; S_{11} = Y'_{-1}M_W Y_{-1}/T, \; S_{01} = (\Delta Y)'M_W Y_{-1},$$

$$S_{00} = (\Delta Y)'M_W(\Delta Y), \; S_{10} = S'_{01},$$

$\hat{\beta}$ is defined as a $k \times r$ matrix minimizing

$$\det\{\beta'(S_{11} - S_{10}S_{00}^{-1}S_{01}\})\beta\} \quad subject\ to\ \beta' S_{11}\beta = I_r,$$

with respect to any $k \times r$ matrix β. Let us denote eigenvalues of $S_{11}^{-1/2}S_{10}S_{00}^{-1}S_{01}S_{11}^{-1/2}$ as $\lambda_j, j = 1, \dots, \lambda_k$, and the eigenvector associated with λ_j as $S_{11}^{1/2}\hat{\beta}_j$. Then it can be shown that $\hat{\beta} = [\hat{\beta}_1, \dots, \hat{\beta}_r]$. Also, put

$$\hat{\beta}' = [\hat{\beta}'_{**}, \hat{\beta}'_3]$$

with $\hat{\beta}'_{**}$ of $r \times k_{**}$ and $\hat{\beta}'_3$ of $r \times k_3$. In our pretest, not $\hat{\beta}$ itself but its linear transfoemation is used in order to infer the structure of $\hat{\beta}$. From the eigenvectors and eigenvalues of $\hat{\beta}'_3\hat{\beta}_3$, we can find an $r \times r$ orthogonal matrix $\hat{L}$ and real numbers $\hat{\lambda}_i$, $i = 1, \dots, r$ such that

$$\hat{L}'\hat{\beta}'_3\hat{\beta}_3\hat{L} = diag\{\hat{\lambda}_1, \dots, \hat{\lambda}_r\}, \tag{11}$$

with the supposition of $0 \leq \hat{\lambda}_1 \leq \cdots \leq \hat{\lambda}_r$. Write

$$\tilde{\beta} = \hat{\beta}\hat{L}, \; \tilde{\beta}_{**} = \hat{\beta}_{**}\hat{L}, \; \sim\hat{\beta}_3 = \hat{\beta}_3\hat{L}.$$

Also, for the case in which $r > r_3 > 0$, let

$$\tilde{\beta}_3 = [\tilde{\beta}_{3,**}, \tilde{\beta}_{3,3}],$$

with $\tilde{\beta}'_{**}$ of $r \times k_{**}$ and $\tilde{\beta}'_3$ of $r \times k_3$. Furthermore, define $\tilde{\beta}_3 = \tilde{\beta}_{3,3}$ if $r = r_3$. We shall first state some asymptotic properties on S_{ij}, $i, j = 0, 1$, with respect to $\bar{\beta}$, γ and δ.

Let us turn to some fundamental asymptotic properties. First, we give a fundamental esult on $\bar{\beta}$ and γ.

Lemma 1

For $0 \geq i_1, i_2, i_3 \geq p$, *define*

$$z_{1t}(i_1, i_2) = b_1' \bar{\beta} y_{t-i_1} + b_2' \Delta y_{t-i_2},$$

$$z_{2t}(i_3) = b_3' \bar{\beta} y_{t-i_3} + b_4' \gamma' \Delta y_{t-i_3}$$

where b_h, $h = 1, 2, 3, 4$, *are fixed vectors of* $r \times 1$, $k \times 1$, $r \times 1$ *and* $s \times 1$ *respectively such that* $b_1 \neq 0$ *or* $b_2 \neq 0$ *and* $b_4 \neq 0$. *Then*

$$\sum_{p+1}^{T} z_{1t}(i_1, i_2)\, z_{2t}(i_3) / T = O_p(1).$$

We can see that this lemma is a direct application of well-known results established in such a paper as Phillips (1987) etc. By this, the proof of this lemma is omitted. Next, we shall present a series of lemmas needed to establish some asymptotic justification of our pretest and Granger causality test. The following lemma states some asymptotic properties on S_{ij}, $i, j = 0, 1$, with respect to $\bar{\beta}$, γ and δ given above.

Lemma 2

(*i*) $S_{00} \rightarrow_p \Sigma_{00}$.

(*ii*) $S_{00} - S_{01} S_{11}^{-1} S_{10} \rightarrow_p \Lambda$.

(*iii*) $\bar{\beta}' S_{10} \rightarrow_p \bar{\beta}' \Sigma_{10}$.

(*iii*) $\bar{\beta}' S_{11} \bar{\beta} \rightarrow_p \bar{\beta}' \Sigma_{11} \bar{\beta}$

(*iv*) $\gamma' S_{11} \gamma / T \rightarrow_d B_1 = \int_0^1 U_s(t)\, U_s(t)'\, dt$.

where Σ_{ij} *are* $k \times k$ *matrices given in Johansen (1988) such that rank* $\Sigma_{00} = k$, *rank* $\bar{\beta}' \Sigma_{10} = r$ *and rank* $\bar{\beta}' S_{11} \bar{\beta} = r$ *and* $U_s(t)$ *is a s-dimensional motion with a covariance matrix* $\Omega_1 = \gamma' \gamma \tau \delta' \Lambda \delta \tau' \gamma' \gamma$, *with a stimess matrix* τ *of full rank.*

Essentially, this lemma is the same as Lemma 3 of Johansen (1988) except (ii), and only (ii) of this lemma will be proved in the Appendix.

Recalling that $\hat{\beta} = \hat{\beta}\,\hat{L}$ and $\tilde{\beta} = \beta\,\bar{L}$, the asymptotic results of $\hat{\beta}$ presented in Lemmas 5 and 8 of Johansen (1988), a part of which is refered to Lemma 5 (i) in Toda and Phillips (1988), can be easily rewritten as those of $\tilde{\beta}$. That is,

Lemma 3 *Letting*

$$x = (\bar{\beta}'\,\bar{\beta})^{-1}\bar{\beta}'\,\tilde{\beta},\quad y = (\gamma'\,\gamma)^{-1}\gamma'\,\tilde{\beta},$$

we have:

(*i*) $\tilde{\beta} = \bar{\beta}x + \gamma y.$

(*ii*) $x = O_p(1),\ x^{-1} = O_p(1).$

(*iii*) $Tyx^{-1} \rightarrow_d \beta_1^{-1} B_2 \Phi$

where B_1 *is given in Lemma* 2,

$$\Phi = (\bar{\beta}'\,\Sigma_{10}\,\Sigma_{00}^{-1}\,\Sigma_{01}\,\bar{\beta})^{-1}\,\bar{\beta}'\,\Sigma_{11}\,\bar{\beta},\ B_2 = \int_0^1 U_s(t)\,d\,V_r(t),$$

$U_s(t)$ *is given in Lemma* 2, $V_r(t)$ *is an r-dimensional Brownian motion with a covariance matrix* $\Omega_2 = \bar{\beta}'\,\Sigma_{10}\,\Sigma_{00}^{-1}\,\Lambda\,\Sigma_{00}^{-1}\,\Sigma_{01}\,\bar{\beta}$ *and* $U_s(t)$ *and* $V_r(t)$ *are independent.*

Using the notations $\tilde{\beta}_3$, $\bar{\beta}_3$ and γ_3 which come from the partitioning of $\tilde{\beta}$, $\bar{\beta}$ and γ, the lower r_3 relations in Lemma 3 (i) is written as

$$\tilde{\beta}_3 = \bar{\beta}_3 x + \gamma_3 y. \tag{12}$$

For the case in which $r > r_3 > 0$, partition x and y into submatrices as

$$x = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix},\ y = [y_{\cdot **},\ y_{\cdot 3}],$$

with x_{11} of $r_{**} \times r_{**}$, x_{12} of $r_{**} \times r_3$, x_{21} of $r_3 \times r_{**}$, x_{22} of $r_3 \times r_3$, $y_{\cdot *}$ of $s \times r_{**}$ and $y_{\cdot 3}$ of $s \times r_3$. Also, recalling that

$$\tilde{\beta}_3 = [\tilde{\beta}_{3,**},\ \tilde{\beta}_{3,3}],$$

with $\tilde{\beta}_{**}'$ of $r \times k_{**}$ and $\tilde{\beta}_3'$ of $r \times k_3$, under the case in which $r > r_3 > 0$, by the form of $\tilde{\beta}_3$, (12) can be rewritten as

$$\tilde{\beta}_{3,**} = \bar{\beta}_{3,3} x_{21} + \gamma_3 y_{\cdot **}, \tag{13}$$

$$\tilde{\beta}_{3,3} = \bar{\beta}_{3,3} x_{22} + \gamma_3 y_{\cdot 3}, \tag{14}$$

where $\gamma = [\gamma'_{**}, \gamma'_3]'$ with γ_{**} of $k_{**} \times s$ and γ_3 of $k_3 \times s$. Notice that partitioning γ into γ_{**} and γ_3 is very closely done to that of $\bar{\beta}$ into $\bar{\beta}_{**}$ and $\bar{\beta}_3$. Our next lemma presents other asymptotic results for this case.

Lemma 4

Suppose that $r > r_3 > 0$. Then:

(*i*) $x_{ii} = O_p(1)$, $x_{ii}^{-1} = O_p(1)$, $i = 1, 2$, $x_{21} = O_p(1)$,

with x_{ij} introduced with respect to (12).

(*ii*) $T\sim\tilde{\beta}_{3,**} = M_3 \gamma_3 T y_{\cdot **} + O_p(T^{-1})$,

where $M_3 = I_{k3} - \bar{\beta}_{3,3}(\bar{\beta}'_{3,3}\bar{\beta})^{-1}\bar{\beta}'_{3,3}$, with $\bar{\beta}_{3,3}$, r_3 and $\bar{\beta}_{3,*}$ given already.

What Lemma 4 (ii) indicates is that $\tilde{\beta}_3$ converges in probability to a $k_3 \times r$ matrix which has the same structure as $\bar{\beta}_3$ with respect to zero restrictions. Hereafter, let $A \otimes D$ and $vec\, A$ denote the Kronecker product of matrices A and D and the $m_1 m_2 \times 1$ vector formed by the column vectors of an $m_1 \times m_2$ matrix $A = [a_1, a_2, \ldots, a_{m_2}]$ as

$$vec\, A = [a'_1, a'_2, \ldots, a'_{m_2}]',$$

respectively. Now, we conclude this section by stating an important result on two independent Brownian motions:

Lemma 5

Let C_1 and C_2 be matrices of $m_1 \times s$ and $m_2 \times r$ such that rank $C_1 = m_1$ and rank $C_2 = m_2$, where m_1 and m_2 are integers such that $m_1 \geq s$ and $m_2 \geq r$, with r and s given already. Also, define the $m_1 m_2 \times 1$ vector b as

$$b = \{(C_2 \Omega_2 C'_2)^{-1/2} \otimes (C_1 B_1^{-1} C'_1)^{-1}\, vec\, C_1 B_1^{-1} B_2 C'_2,$$

where B_1 is given in Lemma 2 and Ω_2 and B_2 are given in Lemma 3. Then, b is distributed as m_1 m_2-dimensional multivariate normal with mean zero and the covariance matrix $I_{m_1 m_2}$, i.e. $N(0, I_{m_1 m_2})$.

Noting that the conditional distribution of $vec\, B_2$ given $\{U_s(t)\}$ is $N(0, I_{m_1 m_2})$, this lemma can be proved by the same arguments as those used to prove Lemma 5.1 of Park and Phillips (1988), as suggested in page 1984 of Toda and Phillips (1993). It should be also noted that the distribution of $vec\, B_2$ is named as *mixed normal.*

4 The pretest procedure

As a quite natural thing, the formulation (10) in section 2 is not available as the null hypothesis to test Granger causality unless the number of elements of A is given a priori. It is impossible unless the rank of β_3 or $\bar{\beta}_3$ (i. e. r_3) is known. However, this may be improbable since it is unnatural to suppose that r_3 is known. Thus, we need a valid estimaor of r_3 or equivalently r_{**} (recall that $r = r_{**} + r_3$ and r is supposed to be known.) in order to proceed to some test for Granger causality based on (10). In this section we shall discuss a method[5)] to infer the value of r_3 or r_{**} as a pretest which precedes a Granger causality test.

In order to construct the pretest statistic, several estimtes on γ given in (4) and a particular form of γ are dispensable, though the construction is mainly based on $\tilde{\beta}_3$ introduced in the previous section. Let $\tilde{N} = \tilde{\beta}\,(\tilde{\beta}'\,\tilde{\beta})^{-1}\tilde{\beta}'$. Based on the eigenvectors and eigenvalues of $\tilde{N}$, we can find a $k \times k$ orthogonal matrix $\tilde{L}$ and $\tilde{\lambda}_i$, $i = 1, ..., k$ such that

$$\tilde{N} = \tilde{L}\, diag\,\{\tilde{\lambda}_1, ..., \tilde{\lambda}_k\},\ \tilde{\lambda}_i = 1\ for\ i = 1, ..., r,$$
$$\tilde{\lambda}_i = 0\ \ for\ i = r + 1, ..., k. \tag{15}$$

5) One method is the well-known Johansen's (1988) likelihood ratio test. In this paper, we will propose another one as an information criterion like AIC or BIC rather than a hypothesis test

Letting $\tilde{L} = [\tilde{L}\,(1),\ \tilde{L}\,(2)]$ with matrices $\tilde{L}\,(1)$ of $k \times r$ and $\tilde{L}\,(2)$ of $k \times r$ and recalling that $\tilde{\lambda}_i = 0$ for $i = r + 1, ..., k$, it is easily checked that

$$\tilde{L}\, diag\,\{\tilde{\lambda}_1, ..., \tilde{\lambda}_k\} = \tilde{L}\,(1)\,\tilde{L}'\,(1).$$

Define $\tilde{\gamma} = \tilde{L}\,(2)$.

Now, consider the form of γ we can get under several situations. For r_3 and r_{**} defined in section 2, put $s_3 = k_3 - r_3$ and $s_{**} = k_{**} - r_{**}$. Recalling

$$k - 1 = k_{**} - 1 + k_3 \geq r = r_{**} + r_3, \quad 0 \leq r_3 \leq k_3,$$

we see that $s_{**} + s_3 = k - r = s$. Also, recall that under the case in which $0 < r_3 < r$, either $r_3 \leq k_3 - 1$ and $r_{**} \leq k_{**}$ or $r_3 \leq k_3$ and $r_{**} \leq k_{**} - 1$ must hold since

$$\bar{\beta}' = \begin{bmatrix} \bar{\beta}'_{**,**} & 0 \\ \bar{\beta}'_{**,3} & \bar{\beta}'_{3,3} \end{bmatrix},$$

$\bar{\beta}_{**,**}$ is $r_{**} \times k_{**}$, $\bar{\beta}_{3,3}$ is $r_3 \times k_3$ and $rank\ \bar{\beta}' = r_{**} + r_3 = r$ under this case, as explained already in section 2. It implies that for the case in which $r_3 = k_3$, $r_{**} \leq k_{**} - 1$, therefore $s_{**} \geq 1$ and that for the case in which $r_3 < k_3$, $s_3 \geq 1$. If $r_{**} < k_{**}$, it is easy to find a $k_{**} \times s_{**}$ matrix $\gamma_{**,**}$ such that

$$\gamma'_{**,**}\,\bar{\beta}_{**,**} = 0, \quad rank\ \gamma_{**,**} = s_{**}.$$

Similarly, if $r_3 < k_3$, we can find a $k_3 \times r_3$ matrix $\gamma_{3,3}$ such that

$$\gamma'_{3,3}\,\bar{\beta}_{3,3} = 0, \quad rank\ \gamma_{3,3} = s_3.$$

For the case in which $0 < r_{**} < k_{**}$ and $0 < r_3 < k_3$, define the $k \times s$ matrix $\bar{\gamma}$ as

$$\bar{\gamma} = \begin{bmatrix} \gamma_{**,**} & 0 \\ \gamma_{**,3} & \gamma_{3,3} \end{bmatrix},$$

with $\gamma_{**,3}$ and $\gamma_{3,3}$ given above and

$$\gamma_{**,3} = -\bar{\beta}_{3,3}(\bar{\beta}'_{3,3}\beta_{3,3})^{-1}\bar{\beta}'_{**,3}\,\gamma_{****}.$$

Also, for the case in which $0 < r_3 = k_3$ (therefore $s_3 = 0$ and $0 < r_{**} < k_{**}$), define $s_{**} = s$ and the $k \times s_{**}$ matrix $\bar{\gamma}$ as

$$\bar{\gamma} = [\gamma'_{**,**}, \gamma'_{**,3}]',$$

with $\gamma'_{**,**}$ and $\gamma'_{**,3}$ given above. Furthermore, for the case in which $0 < r_{**} = k_{**}$ (therefore $s_{**} = 0$ and $0 < r_3 < k_3$), define $s_3 = s$ and the $k \times s_3$ matrix $\bar{\gamma}$ as

$$\bar{\gamma} = [0, \gamma'_{3,3}]',$$

with $\gamma_{3,3}$ given above. Recalling that γ given in (4) can be freely defined as long as it satisfies the conditions defined by (4), $\bar{\gamma}$ defined above can be regarded as one particular form of γ. Therefore, without being generality we can suppose hereafter that γ appeared in the above lemmas is one which has the form $\bar{\gamma}$. Then, note that

$$\gamma = \bar{\gamma} = \begin{bmatrix} \gamma_{**} \\ \gamma_3 \end{bmatrix} = \begin{bmatrix} \gamma_{**,**} & 0 \\ \gamma_{**,3} & \gamma_{3,3} \end{bmatrix}$$

and

$$\gamma_{**} = [\gamma_{**,**}, 0], \quad \gamma_3 = [\gamma_{**,3}, \ \gamma_{3,3}].$$

Let us turn to the construction of a pretest statistic. Extend the definition of x_{22} introduced with respect to (14) for the case in which $0 < r_3 < r$ to the case in which $r_3 = r$ by putting $x_{22} = x$, where x is introduced in Lemma 3 (i). Then, with positive integers n_1 and n_2 such that $n_1 + n_2 = r$, we partition $\tilde{\beta}_3$, x and x_{22} regardless of the value of r_3 as follows:

$$\tilde{\beta}_3 = [\tilde{\beta}_3(1), \tilde{\beta}_3(2)], \tag{16}$$

where $\tilde{\beta}_3(1)$ is $k_3 \times n_1$ and $\tilde{\beta}_3(2)$ is $k_3 \times n_2$,

$$x = \begin{bmatrix} x(1,1) & x(1,2) \\ x(2,1) & x(2,2) \end{bmatrix}, \tag{17}$$

where x (1, 1) is $n_1 \times n_1$, x (1, 2) is $n_1 \times n_2$, x (2, 1) is $n_2 \times n_1$ and x (2, 2) is $n_2 \times n_2$, and

$$x_{22} = [x_{22}(1), x_{22}(2)], \tag{18}$$

where $x_{22}(1)$ is $r_3 \times (r_3 - n_2)$ and $x_{22}(2)$ is $r_3 \times n_2$. Also, letting $\tilde{N} = \tilde{\beta}_3(2)$ $(\tilde{\beta}_3'(2)\tilde{\beta}_3(2))^{-1}$ $\tilde{\beta}_3(2)$ for n_1 and n_2 chosen such that $k_3 > n_2$ is satisfied, by the eigenvectors and eigenvalues of $\tilde{N}_3$, we find a $k_3 \times k_3$ orthogonal matrix $\tilde{L}_3$ and $\tilde{\mathrm{v}}_i$, i = 1, ..., k_3, such that

$$\tilde{N}_3 = \tilde{L}_3\, diag\,\{\tilde{\mathrm{v}}_1, ..., \tilde{\mathrm{v}}_{K_3}\}\, \tilde{L}_3' = \tilde{L}_3(1)\, \tilde{L}_3(1), \tag{19}$$

where $\tilde{L}_3 = [\tilde{L}_3(1), \tilde{L}_3(2)]$ with $\tilde{L}_3(1)$ of $k_3 \times n_2$ and $\tilde{L}_3(2)$ of $k_3 \times \bar{n}_2$, where $\bar{n}_2 = k_3 - n_2$. Note that ~ $\tilde{\mathrm{v}}_i$, i = 1, ..., k_3 are chosen such that b $\tilde{\mathrm{v}}_i$ = 1 for i = 1, ..., n_2 and $\tilde{\mathrm{v}}_i$ = 0 for $i = n_2 + 1$, ..., k_3.

Now, define $\tilde{\gamma}_{3,3} = \tilde{L}_3(2)$. The pretest statistics proposed are constructed based on the asymptotic behaviors of $T\tilde{\beta}_3$ and $T\tilde{\gamma}_{3,3}'\tilde{\beta}_3(1)$ for several n_1. Those are stated by the following lemma:

Lemma 6

(*i*) $T\tilde{\beta}_3 x^{-1} \to_d \gamma_3 B_1^{-1} B_2 \Phi$ *if* $r_3 = 0$,

where γ_3 *is given with respect to* (13) *and* (14) *and* x, B_1, B_2 *and* Φ *are defined in Lemmas* 2 *and* 3,

(*ii*) $w^{-1} T\, \tilde{\gamma}_{3,3}'\tilde{\beta}_3(1)\, x_{11}^{-1} \to_d S_3, B_1^{-1} B_2\, \Phi\, S_{**}'$ *if* $1 \le r_3 = n_2 \ge m_1$,

where $m_1 = min\ \{r, k_3\}$, $w = (\tilde{\gamma}_{3,3}'\gamma_{3,3})^{-1}$, $\gamma_{3,3}'\tilde{\gamma}_{3,3}$ *with* $\gamma_{3,3}$ *defined above and the matrices* S_3 *of* $s_3 \times s$ *and* S_{**} *of* $r_{**} \times r$ *are defined as*

$$S_3 = [0, \tilde{\gamma}_{3,3}'\gamma_{3,3}], \quad S_{**} = [I_{r^{**}}, 0].$$

Putting $F_1 = \Phi'\,\Omega_2\,\Phi$, $F_2 = \gamma_3 B_1^{-1}\gamma_3'$, $G_1 = S_{**} F_1 S_{**}'$ and $G_2 = S_3 B_1' S_3'$ with B_1,

Φ and Ω_2 given in Lemmas 2 and 3 and S_3 and S_{**} given in Lemma 6, Lemma 6 suggests that the limiting distributions in this lemma can be transformed into $N(0, I_{\bar{n}_2 n_1})$ if several statistics which converge to F_i and G_i, $i = 1, 2$, are provided and combined with $T\tilde{\beta}_3 x^{-1}$ and $w'^{-1} T\ \tilde{\gamma}'_{3,3}\ \tilde{\beta}_3(1)\ x_{11}^{-1}$. For the $k \times s$ matrix $\tilde{\gamma} = \tilde{L}\,(2)$ defined in the early part of this section as an estimator of $\tilde{\gamma}$, let

$$\tilde{\gamma} = [\tilde{\gamma}'_{**}, \tilde{\gamma}'_3]',$$

with $\tilde{\gamma}_{**}$ of $k_{**} \times s$ and $\tilde{\gamma}_3$ of $k_3 \times s$. Introducing the statistics defined as

$$\tilde{F}_1 = \tilde{\Phi}'\,\tilde{\Omega}_2\,\tilde{\Phi}, \quad \tilde{F}_2 = \tilde{F}_2 = \tilde{\gamma}_3(\tilde{\gamma}'\,S_{11}\,\tilde{\gamma})^{-1}\,\tilde{\gamma}'_3,$$
$$\tilde{G}_1 = S_{**}\tilde{F}_1 S'_{**}, \quad \tilde{G}_2 = \tilde{\gamma}'_{3,3}\,\tilde{F}_2\,\tilde{\gamma}_{3,3},$$

where $\tilde{\Phi} = (\tilde{\beta}'\,S_{10}S_{00}^{-1}\,S_{01}\tilde{\beta})^{-1}\,\tilde{\beta}'\,S_{11}\,\tilde{\beta}, \tilde{\Omega}_2 = \tilde{\beta}'\,S_{10}\,S_{00}^{-1}\,(S_{00} - S_{01}\,S_{11}^{-1}\,S_{10})\,S_{00}^{-1}\,S_{01}\,\tilde{\beta}$, S_{**} is given in Lemma 6 and S_{ij}, $i, j = 0, 1$, and $\tilde{\beta}$ are given in section 3, we have:

Lemma 7

(*i*) $(x \otimes I_k)\,(\tilde{F}^{-1} \otimes \tilde{F}^{-1})\,(x' \otimes I_{k_3}) \to_d F_1^{-1} \otimes F_2^{-1}$ *if* $r \leq k_{**}$.

(*ii*) $(x_{11} \otimes w)\,(\tilde{G}_1^{-1} \otimes \tilde{G}_2^{-1})\,(x'_{11} \otimes w') \to_d \tilde{G}_1^{-1} \otimes \tilde{G}_2^{-1}$ *if* $1 \leq r_3 = n_2 \geq m_1 - 1$, *where* x, x_{11} *and* w *are the same as those in Lemma* 6.

Now, define

$$\tilde{C}\,(1) = T^2\,tr\,\tilde{F}_1^{-1}\,\tilde{\beta}'_3\,\tilde{F}_2^{-1}\,\tilde{\beta}_3, \tag{20}$$

$$\tilde{C}(n_2 + 1) = T^2\,tr\,\tilde{G}_1^{-1}\,\tilde{\beta}'_3(1)\,\tilde{\gamma}_{3,3}\,\tilde{G}_2^{-1}\,\tilde{\gamma}_{3,3}\,\tilde{\beta}'_3(1),\ n_2 = 1, ..., m_1 - 1, \tag{21}$$

under the case in which $r \geq k_{**}$, recalling $m_1 = min\,\{r, k_3\}$.

The pretest procedure proposed is composed by several hypothesis tests which are numbered from $m_4 + 1$ *to* m_1 conveniently, with $m_4 = max\,\{k_3 - s, 0\}$. For $j = m_4 + 1, ..., m_1$, the null and alternative hypotheses in the j-th test are stated as follows:

$$H_0: r_3 = j-1,\ equivalently,\ r_{**} = r-j+1,\ or\ s_3 = k_3-j+1, \qquad (22)$$

$$H_1: r_3 > j-1,\ equivalently,\ r_{**} < r-j+1,\ or\ s_3 = k_3-j+1. \qquad (23)$$

Also, the $\tilde{C}(j)$ defined by (20) and (21) in which n_2 is replaced by $j+1, j = m_4+1, \ldots, m_1$, are the test statistics numbered correspondingly. Recall $m_2 = max\{0, r-k_3\}$ and $m_3 = max\{k_3-r, 0\}$ and noting that

$$r_3 = k_3 - s_3 \leq k_3 - s,\ \ r_{**} = k_{**} - s_{**} = k - k_3 - s + s_3$$
$$= r - r_3 \leq r - k_3,\ \ s_3 = k_3 - r_3 \leq k_3 - r,$$

it is obvious that the theoretical values of r_3, r_{**} and s_3 are subject to the restrictions such that

$$m_4 \leq r_3 \leq m_1,\ \ m_5 \geq r_{**} \geq m_2,\ \ m_6 \geq s_3 \geq m_3,$$

with $m_5 = min\ \{k_{**}, r\}$ and $m_6 = min\ \{s, k_3\}$. However, through the procedure, the hypothesized values of r_3, r_{**} and s_3 are confined to those such that

$$m_4 \leq r_3 \leq m_1 - 1,\ \ m_5 \geq r_{**} \geq m_2 + 1,\ \ m_6 \geq s_3 \geq m_3 + 1.$$

We should pay attention to that this procedure is not necessarily started from the 1st test for which the null hypothesis is formulated as $r_3 = 0$. The values of k_{**} (or k_3) and r (or s) must be considred carefully as some useful information to shorten the procedure before executing the pretests. Letting $h = r - k_{**}$, by $r - k_{**} = k_3 - s = r_3 - s_{**}$ we see the information $r_3 = h + s_{**} \geq h$ without entering any pretest. As a result, the case in which $r > k_{**}$ (or equivalently $k_3 > s$) drives us to the start of the procedure by the $(h+1)$ th test (such that $h \geq 1$) for which the null hypothesis is $r_3 = h$. However, if $r \leq k_{**}$ (or equivalently $k_3 \leq s$), we start the 1st test to decide whether r3 is zero or not.

The pretest procedure is carried out as follows. If the null hypothesis in the j th test is accepted, our procedure is terminated at this stage and then the decision is made as $r_3 = j - 1$, $r_{**} = r - j + 1$ and $s_3 = k_3 - j + 1$, provided that $j = m_4 + 1, \ldots, m_1$. For the case of rejection of the null hypothesis in the j th test, we proceed to the $(j + 1)$ th test based on the decision that $r_3 \geq j$, $r_{**} \leq r - j$ and $s_3 \leq k_3 - j$, provided that $j = m_4 + 1, \ldots, m_1 - 1$. The procedure is continued up to the m_1th test as long as the null hypothesis is rejected. The m_1th test always terminates the procedure. In other words, the maximum number of tests executed must be m_1. The rejection in the m1th test leads to the decision that $r_3 = m_1$, $r_{**} = m_2$ and $s_3 = m_3$ since $r_3 \leq m_1$, $r_{**} \geq m_2$ and $s_3 \geq m_3$.

The results in Lemma 6 and Lemma 7 (ii) are used to justify our pretests under the null hypotheses. For the alternative hypotheses, we need other results as presented in the following two lemmas.

Lemma 8

Let n_1 and n_2 be positive intergers such that $n_1 + n_2 = r$. For the case in which $1 \leq n_2 < r_3$ $(\leq m_1)$, let

$$\bar{x}_{22}(1) = x_{22}(1) - x_{22}(2)\,\{x'_{22}(2)\bar{\beta}'_{3,3}\ \bar{\beta}_{3,3}x_{22}(2)\}^{-1}x'_{22}(2)\,\bar{\beta}_{3,3}{}'\bar{\beta}_{3,3}x_{22}(1),$$
$$\Psi_{11} = \bar{x}'_{22}(1)\ \bar{\beta}'_{3,3}\ \bar{\beta}_{3,3}\,\bar{x}_{22}(1)$$

with $x_{22}(1)$ and $x_{22}(2)$ defined by (18). Notice that Ψ_{11} and $\bar{x}_{22}(1)$ are $(r_3 - n_2) \times (r_3 - n_2)$. Also, recalling that $k_3 - n_2 = s_3 + (r_3 - n_2)$, $k_3 \geq r_3$, $n_1 = r - n_2 = r_{**} + (r_3 - n_2)$ and $r \geq r_3$ (or $r_{**} \geq 0$), define the matrices δ *of* $k_3 \times (k_3 - n_2)$ *and* Ψ *of* $(k_3 - n_2) \times n_1$ *as follows:*

$$\delta = \bar{\beta}_{3,3}\,\bar{x}_{22}(1) \quad \text{if } r_3 = k_3,$$
$$= [\gamma_{3,3},\ \bar{\beta}_{3,3}\bar{x}_{22}(1)] \quad \text{if } r_3 < k_3,$$
$$\Psi = \begin{bmatrix} 0 & 0 \\ 0 & \Psi_{11} \end{bmatrix} \text{ if } r_3 < m_1 (= \min\{r, k_3\}),$$

$$= [0,\ \Psi_{11}] \ \ if\ r_3 = k_3 < r,$$

$$= \begin{bmatrix} 0 \\ \Psi_{11} \end{bmatrix} if\ r_3 = r < k_3,$$

$$= \Psi\ if\ r_3 = r = k_3.$$

Then, for $\tilde{\beta}_3(1)$ *defined by* (16) *and* $\tilde{\gamma}_{3,3}$ *defined with respect to* (19), *we have*

$$\tilde{\gamma}'_{3,3}\,\tilde{\beta}_3(1) = \bar{w}'\,\Psi + O_p(T^{-1}),$$

with

$$\bar{w} = (\delta'\,\delta)^{-1}\,\delta'\,\tilde{\gamma}_{3,3}\},$$

of $(k_3 - n_2) \times (k_3 - n_2)$, *and*

$$\bar{w} = O_p(1),\ \ \bar{w}^{-1} = O_p(1),\ \ \Psi_{11} = O_p(1),\ \ \Psi_{11}^{-1} = O_p(1).$$

Lemma 9

For positive integers n_1 *and* n_2 *such that* $n_1 + n_2 = r$, *define the* $n_1 \times r$ *matrix* $\bar{S}_{**}$ *and* $(k_3 - n_2) \times s$ *matrix* $\bar{S}_3$ *as*

$$\bar{S}_{**} = [I_{n1},\ 0],$$

$$\bar{S}_3 = \begin{bmatrix} 0 & \gamma'_{3,3}\,\gamma_{3,3} \\ -\bar{x}'_{22}(1)\,\bar{\beta}'_{**,3}\,\gamma_{**,**} & 0 \end{bmatrix} if\ r_3 < k_3,$$

$$= -x'_{22}(1)\,\bar{\beta}'_{**,3}\,\gamma_{**,**}\ if\ r_3 = k_3.$$

Also, let

$$\bar{G}_1 = \bar{S}_{**}\,F_1\,\bar{S}'_{**},\quad \bar{G}_2 = \bar{S}_3\,B_1^{-1}\,\bar{S}'_3,$$

where F_1 *is already defined as* $F_1 = \Phi'\,\Omega_2\,\Phi$, *with* Φ *and* Ω_2 *given in Lemma* 2. *Then, for the same case as in Lemma* 7,

$$\tilde{G}_1^{-1} \otimes \tilde{G}_2^{-1} = \{x^{-1}(1,1) \otimes \bar{w}^{-1}\}\,(\bar{G}_1^{-1} \otimes \bar{G}_2^{-1})\,\{x'^{-1}(1,1) \otimes \bar{w}'^{-1}\} + O_p(T^{-1}),$$

where $\tilde{G}_i$, *together with* $\tilde{F}_i$, *are introduced in the text,* x (1, 1) *is given in* (17) *and* $\bar{w}$ *is defined in Lemma* 8.

Note that $n_1 - r_{**} = r_3 - n_2 = k_3 - n_2 - s_3$ and $s - 1 \geq k_3 - n_2$ for the case in which $n_2 > m_4 = max\ \{k_3 - s, 0\}$. Now, through the above lemmas, we attain to a theorem for justifying the pretest procedure proposed:

Theorem 1

$\tilde{C}\ (j)$, $j = m_4 + 1, \ldots, m_1$, *which are defined by* (20) *and* (21), *are asymptotically distributed as chi-square with* $(k_3 - j + 1)\ (r - j + 1)$ *degreees of freedom under the null hypothesis* (21), *i. e.*

$$\tilde{C}\ (j) \rightarrow_d \chi^2_{(k_3 - j+1)\ (r - j+1)} \quad if\ r_3 = j - 1,$$

and are expressed as $\tilde{C}\ (j) = T^2 C\ (j)$ *with random variaqbles* $C\ (j)$ *such that* $C\ (j) = O_p\ (1)$ *and* $C^{-1}\ (j) = O_p\ (1)$ *under the alternative hypothesis* (22), *i. e.*

$$T^{-2}\tilde{C}\ (j) = O_p\ (1),\ \ \{T^{-2}\ \tilde{C}\ (j)\}^{-1} = O_p\ (1) \quad if\ r_3 \geq j.$$

This theorem states that the pretests to determine the value of r_3 are asymptotically valid as χ^2 criteria are used. The results for the alternative hypotheses assert that the pretests are consistent with the advantage that those divergences are remarkably faster as compared with the rate of T in the usual consistent tests. It should be also noted that this theorem holds without imposing such a condition as that in Theorem 3 of Toda and Phillips (1993).

5 The Granger causality test

Suppose that the value of r_3 is known or is correctly infered through the pretests in the previous section. Conveniently, let us introduce the following matrices constructed by the notations in section 2:

$$X_1 = \begin{bmatrix} \Delta y_p'(**) & \dots & \Delta y_2(**) \\ \vdots & \ddots & \vdots \\ \Delta y_{T-1}'(**) & \cdots & \Delta y_{T-p+1}'(**) \end{bmatrix}, X_2 = \begin{bmatrix} \Delta y'_p(3) & \cdots & \Delta y_2(3) \\ \vdots & \ddots & \vdots \\ \Delta y_{T-1}'(3) & \cdots & \Delta y_{T-p+1}'(3) \end{bmatrix},$$

$$\Delta Y_1 = \begin{bmatrix} \Delta y_{p+1}'(1) \\ \vdots \\ \Delta y_T' \end{bmatrix}, \bar{\varepsilon}_1 = \begin{bmatrix} \varepsilon_{p+1}'(1) \\ \vdots \\ \varepsilon_T' \end{bmatrix},$$

$$Y_{**} = \begin{bmatrix} y_p'(**) \\ \vdots \\ y_{T-1}'(**) \end{bmatrix}, Y_3 = \begin{bmatrix} y_p'(3) \\ \vdots \\ y_{T-1}' \end{bmatrix},$$

$$X = [X_1, Y_{**}\bar{\beta}_{**}],\ Z = [X_2, Y_3\ \bar{\beta}_{3,3}]\ \ if\ r_3 > 0,\ Z = X_2\ \ if\ r_3 = 0.$$

Using those matrix notations, (9) is written as

$$\Delta Y_1 = XB + ZA + \bar{\varepsilon}_1, \tag{24}$$

with B and A defined with respect to (9). For the construction of a statistic to test the null hypothesis formulated by (10), define $\tilde{X}$ and $\tilde{Z}$ as

$$\tilde{X} = [X_1, Y_{**}\tilde{\beta}_{**}],\ \tilde{Z} = [X_2, Y_3\tilde{\beta}_{3,3}]\ \ if\ r_3 > 0,\ \tilde{Z} = X_2\ \ if\ r_3 = 0,$$

with $\tilde{\beta}_{**}$ and $\tilde{\beta}_3$ defined in section 3. Note that both $\tilde{X}$ and $\tilde{Z}$ are available as some observation. Also, notice that $\tilde{Z}$ is just Z as $r_3 = 0$. Based on X and $\tilde{Z}$, we shall construct the following statistic:

$$\tilde{A} = (\tilde{Z}' M_{\tilde{X}} \tilde{Z})^{-1} \tilde{Z}'\ M_{\tilde{X}}\ \Delta Y_1, \tag{25}$$

where $M_{\tilde{X}} = I_{T-p} - \tilde{X}\ (\tilde{X}'\ \tilde{X})^{-1}\ \tilde{X}'$. Letting $\Lambda_{11} = E\ \{\varepsilon_t\ (1)\ \varepsilon_t'(1)\}$, we need an estimator for Λ_{11}. Note that

$$\Lambda_{11} = [I_{k1}, 0]\ \Lambda \begin{bmatrix} I_{k1} \\ 0 \end{bmatrix},$$

recalling $\Lambda = E\ \{\varepsilon_t\ \varepsilon_t'\}$. In view of Lemma 2 (ii),

$$\tilde{\Lambda}_{11} = [I_{k_1}, 0]\,(S_{00} - S_{01}\,S_{11}^{-1}\,S_{10})\begin{bmatrix} I_{k_1} \\ 0 \end{bmatrix}$$

is provided as such an estimator. Now, our test statistic is constructed as

$$\tilde{C} = T\,tr\,\tilde{\Lambda}_{11}^{-1}\tilde{A}'\,(\tilde{Z}'\,M_{\tilde{X}}\tilde{Z}\,/\,T)\,\tilde{A}. \tag{26}$$

With respect to (24), $\tilde{A}$ may be interpreted as the OLS estimator for A if $\tilde{\beta}_{**}$ and $\tilde{\beta}_{3,3}$ are dealt with as $\tilde{\beta}_{**}$ and $\tilde{\beta}_{3,3}$ respectively. Also, we can say that $\tilde{C}$ is the usual Wald statistic to test (10) if $\tilde{\beta}_{**}$ and $\tilde{\beta}_{3,3}$ are replaced with $\bar{\beta}_{**}$ and $\bar{\beta}_{3,3}$ respectively. The next theorem states some asymptotic validity of the test based on $\tilde{C}$.

Theorem 2

$\tilde{C}$ defined by (26) *is asymptotically distributed as chi-square with* $k_1\,\{k_3\,(p-1) + r_3\}$ *degrees of freedom when the null hypothesis* (10) *holds, i. e.*

$$\tilde{C} \to_d \chi^2_{k_1\bar{m}} \text{ if } A = 0,$$

with $\bar{m} = k_3\,(p-1) + r_3$. *Also, if* (10) *does not hold, there exists a random variable* $\bar{C}$ *such that*

$$\tilde{C} = T\,\bar{C},\;\; \bar{C} = O_p\,(1),\;\; \bar{C}^{-1} = O_p\,(1).$$

This theorem also concludes, like Theorem 1, that the Granger causality test based on $\tilde{C}$ is aymptotically valid in the sense that the use of *chi*^{2} critical values is ensured with the consistency of the test. We note again that such a condition as that in Theorem 3 of Toda and Phillips (1993) is not required for the results stated in this theorem.

6 Some remarks and implications

As discussed in sections 4 and 5, our Granger causality test strongly depends upon the formulation (10) for the null hypothesis of Granger noncausality and such statitical inferences as the pretests in section 4 to make (10) be available. The purpose of this section is to state some remarks and implications on the methods proposed from the practical viewpoint, comparing with one proposed in Toda and Phillips (1993) and featuring different special cases.

The results in Toda and Phillips (1993) are derived based on the formulation (6) for the null hypothesis. What should be noted is that (6) may contain some 'nonlinear' relations $\alpha_1 \beta_3' = 0$. Also, note that (6) may possess some redundancy according to the rank of $\alpha_1 \beta_3'$. The difficulty / complexity in their testing method comes from those. This is the reason why Toda and Phillips (1993) imposed several conditions, such as $rank\ \alpha_1 = k_1$ or $rank\ \beta_3' = k_3$, to avoid such redundancy and derive clear results. Our method, which uses (10) after executing the pretests in section 4 instead of the direct use of (6), is proposed to overcome such a limitation. Any condition, such as in Toda and Phillips (1993), is not imposed in Theorem 2 of ours.

In the pretest procedure in section 4, consider the situation to make the statistic for the 1st test construct and its asymptotic property hold. As seen in (20) and Lemma 7 (i), $r \geq k_{**}$ is supposed to ensure the inverse matrices of $\tilde{F}_2$ and F_2. However, the supposition never be restrictive since the 1st test is not needed unless $r < k_{**}$. The procedure starts from the 2nd test if $r = k_{**}$. The cases in which $r = k_{**} + 1$ and $r = k_{**} + 2$ set the starting points to the 3rd test and 4th test respectively. The case in which $r = 1$ confines the possible value of r_3 to either 0 or 1. Then only the 1st test is required to attain to a conclusion. Generally, if k is not so large, the procedure often becomes simple and requires a few tests to attain to a conclusion since r is not also so large. The case in which $k_1 = k_2 = 1$, $k_3 = 2$ and $r = 3$ may possibly need only the 3rd test after starting from the 2nd test since $h = r - k_{**} = 1$. Also, we can sometimes find the value of r_3 without executing any pretest. For example, let us take up the case in which $k_1 = k_3 = 1$, $k_2 = 0$ and $r = 2$. As seen already, the fact h $= r - k_{**} = r - k_1 = 1$ gives us $r_3 \geq h = 1$. On the other time, recall that $k_3 = 1 \geq r_3$. Those information drives us to the conclusion $r_3 = 1$ without entering the pretest procedure.

Now, let us turn to the issue of testing for Granger causality. If $p = 1$ and we attained to the conclusion $r_3 = 0$ after executing the pretests, further test to decide to decide whether Granger causality from y_t (1) to y_t (3) exists or not is not required since Granger noncausality is implies by $r_3 = 0$ as $p = 1$. In other words, the pretest for $r_3 = 0$ is interpreted as a test for Granger causality under this case. On the other hand, if we make the decision that $\{y_t\ (**)\}$ is not cointegrated, i. e. $r_{**} = 0$ (or equivalently $r_3 = r$), the relation $\bar{\alpha}_1 = 0$ must be tested for the null hypothesis of Granger noncausality.

The condition imposed by Theorem 3 in Toda and Phillips (1993), i. e. either $rank\ \alpha_1 = k_1$ or $rank\ \beta_3 = k_3$, requires either $k_1 \leq r$ or $k_3 = r_3$ since α_1 is $k_1 \times r$ and β_3 is $k_3 \times r_3$ with the fact rank $\beta_3 = r_3$. Under the situation such that either $k_1 = 1$ or $k_3 = 1$, it may be easy for this condition to be satisfied. However, even if it is so, this condition rules out many situations. The case in which $k_{**} = k_1 = r_3 = r = 1$ and $k_3 = 2$, together with the suppostion of Granger noncausality, implies that

$$rank\ \alpha_1 = rank\ \bar{\alpha}_1 = 0 < k_1 = 1,\ \ rank\ \bar{\beta}_3 = rank\ \bar{\beta} = 1 < k_3 = 2,$$

since $r_3 = r$ implies that $\bar{\alpha}_1 = 0$ is contained in the formulation (10) for Granger noncausality, as stated already. Then it is obvious that the condition in Toda and Phillips (1993) is not satisfied. Also, if $k_{**} = k_1 = 2$, $k_3 = r = 1$ and $r_3 = 0$, the condition is not satisfied because of

$$rank\ \alpha_1 = rank\ \bar{\alpha} \leq r = 1 < k_1 = 2,\ \beta_3 = \bar{\beta}_3 = 0.$$

These two cases explicitly illustrates that Toda and Phillips' (1993) condition is not practical.

Finally, recalling the matrix polynomial $A(\lambda)$ defined with respect to the VAR representation as the data-generating process, i.e.

$$A(\lambda) = I_k - \sum_{j=1}^{p} A_j \lambda_j,$$

we embody the special cases in which the conclusion in Toda and Phillips (1993) is not satisfied under the null hypothesis of Granger noncausality by showing the following two examples of $A(\lambda)$.

Example 1: $p = 1, k_{**} = k_1 = r_3 = r = 1, k_3 = 2$ and

$$A(\lambda) = \begin{bmatrix} 1-\lambda & 0 & 0 \\ -2\lambda & 1-0.5\lambda & \lambda \\ -2\lambda & 0.5\lambda & 1 \end{bmatrix}.$$

Then, it is obvious that $\det A(\lambda) = (1-\lambda)^2 (1 + 0.5\lambda)$. Also, since

$$A(1) = \begin{bmatrix} 0 & 0 & 0 \\ -2 & 0.5 & 1 \\ -2 & 0.5 & 1 \end{bmatrix},$$

we can let

$$\bar{\alpha} = [0, 1, 1]', \ \bar{\beta}' = [2, -0.5, -1],$$

which in turn make us deine

$$\delta = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 0 \end{bmatrix}', \ \gamma' = \begin{bmatrix} 0 & -2 & 1 \\ 1 & 2 & 0 \end{bmatrix}.$$

Furthermore, by $A(\lambda) = (1-\lambda) I_3 + A(1)\lambda$, $\Pi(1) = I_3$. Therefore,

$$\delta' \Pi(1) \gamma = \begin{bmatrix} -3 & 2 \\ 0 & 1 \end{bmatrix}.$$

Thus, all the requirements in the present paper are satisfied. On the other hand, $\bar{\alpha}_1 = 0$ and $\bar{\beta}' = [-0.5, -1]$.

Example 2: $p = 1$, $k_{**} = k_1 = 2$, $k_3 = r = 1$, $r_3 = 0$ and

$$A(\lambda) = \begin{bmatrix} 1 - 0.5\lambda & \lambda & 0 \\ 0.5\lambda & 1 - 0.5\lambda & 0a \\ -\lambda & -2\lambda & 1 - \lambda \end{bmatrix}.$$

Therefore, we can let

$$\bar{\alpha} = [1, 1, -2]', \bar{\beta}' = [-0.5, -1, 0],$$

$$\delta = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 2 & 1 \end{bmatrix}', \gamma' = \begin{bmatrix} 0 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix}.$$

It is also easy to see that $\det A(\lambda) = (1 - \lambda)^2 (1 + 0.5\lambda)$ and $A(\lambda) = (1 - \lambda) I_3 + A(1)\lambda$. Thus, we derive all the requirements imposed. On the other hand, $\bar{\alpha}_1 = [1, 1]'$ and $\bar{\beta}' = 0$.

7 Conclusion

In this paper, we have discussed how a valid test for Granger causality in cointegrated systems can be constructed, and proposed a conclusive procedure summarized as follows:

At the first stage, several pretests to make a decision on the rank of a submatrix of the cointegrating matrix (or the cointegrating rank of a subsystem) are executed. Next, based on the rank value determined by those pretests, a test for the null hypothesis of the absense of Granger causality is constructed. All these tests proposed are asymptotically distributed χ^2 under the null hypotheses and consistent as indicated by the results under the alternative hypotheses.

The basic idea of the testing procedure proposed is formalize the null hypothesis of Granger noncausality without any redundancy and nonlinearity with respect to parametrization of the cointegrating matrix. As a result, all the asymptotic results hold even if such a oondition as those in Toda and Phillips (1993) is not imposed. In all the tests proposed, the cointegrating matrix need to be estimated. We adopt not the ML estimator provided by Johansen (1988) but its orthogonal transformation. The pretest procedure proposed may bge seemingly observed as complicated one. However, it is rather simple practically as seen in section 6 and requires only a few number of tests as long as the dimension of the subsystem is not so large. Even no test is required under some cases. For the test for the null hypothesis of Granger noncausality, we can consider it to be a Wald test, provided that the replacement of the cointegrating matrix with its estimate is tolerated.

As mentioned in many literatures (see Phillips and Toda (1993, p. 1369), e. g.), the existence of constant terms in the VAR's drives us into the consideration of some deterministic trends. Intentionally, we have avoided to include a constant term in our VAR based on the reason below:

(i) The concept of cointegration defined by Engle and Granger (1987) does not cope with such deterministic trends sufficiently; therefore, the ECM's may have different implication and require another interpretation as threr exist deterministic trends.

(ii) If we start our method for the cointegrated system from a multivariate moving average representation with drift based on the Wold decomposion and such a specification as that in Engle and Yoo (1987), any constant term does not appear in the VAR derived from it under the suitable initial condition.

As empahsized already, the validity of our testing procedure for Granger causality is described as χ^2 criteria asymptotically and unconditionally. Both the procedure and its validity may be robust under several extension of the model such as the inclusion of a constant term. However, apart from technical derivations, the implications of those extensions must be carefully considered.

References

Box, G. E. P., and G. M. Jenkins (1976): *Time Series Analysis: Forecasting and Control*, 2nd edn., Holden-Day, San Francisco.

Engle, R. F., and C. W. J. Granger (1987): "Co-Integration and error-correction: representation, estimation and testing," *Econometrica*, 55, 251-276.

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

Granger, C. W. J. (1969): "Investigating causal relations by econometric models and cross spectral methods," *Econometrica*, 37, 424-438.

Grenander, U., and M. Rosenflat (1957): *Statistical Analysis of Stationary Time Series*, Wiley, New York.

Johansen, S. (1988): "Statistical analysis of cointegration vectors," *Jornal of Economic Dynamics and Control*, 12, 231-254.

Loeve, M. (1977): *Probablity Theorey*, Vol. 1, New York: Springer-Verlag.

Park, J., and P. C. B. Phillips 81988): "Statistical inference in regressions with integrated processes: part I," *Econometric Theory*, 4, 468-497.

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

Toda, H. Y., and P. C. B. Phillips (1993): "Vector autoregressions and causality," *Econometrica*, 61, 1367-1393.

Appendix

Proof of Lemma 2 (ii): Defining $B_T = [\bar{\beta}, \gamma T^{-1/2}]$ and by Lemma 1 and other parts of Lemma 2, we have

$$S_{00} - S_{01} S_{11} S_{10} = S_{00} - S_{01} B_T (B_T' S_{11} B_T)^{-1} B_T' S_{10}$$
$$= S_{00} - S_{01} \bar{\beta} (\bar{\beta}' S_{11} \bar{\beta})^{-1} \bar{\beta}' S_{10} + O_p(T^{-1}),$$

that is,

$$S_{00} - S_{01} S_{11} S_{10} = S_{00} - S_{01} \bar{\beta} (\bar{\beta}' S_{11} \bar{\beta})^{-1} \bar{\beta}' S_{10} + O_p(T^{-1}), \quad (27)$$

and

$$S_{01}\bar{\beta}(\bar{\beta}' S_{11}\bar{\beta})^{-1}\bar{\beta}' S_{10} \to_p \Sigma_{00} - \Sigma_{01}\bar{\beta}(\bar{\beta}' \Sigma_{11}\bar{\beta})^{-1}\bar{\beta}' \Sigma_{10}. \tag{28}$$

Combining these results, the conclusion follows.

Proof of Lemma 4: (i) It is obvious from Lemma 3 (ii) that $x_{ii} = O_p(1)$, i = 1, 2. Since $y = O_p(T^{-1})$ by Lemma 3 (iii), (13) and (14) are collected into

$$\tilde{\beta}_3' \tilde{\beta}_3 = [x_{21}, x_{22}]' \bar{\beta}_{3,3}' \bar{\beta}_{3,3} [x_{21}, x_{22}] + O_p(T^{-1}). \tag{29}$$

Recalling that

$$\tilde{\beta}_3' \tilde{\beta}_3 = diag\{\hat{\lambda}_1, ..., \hat{\lambda}_r\}$$

and

$$x_{21}' \bar{\beta}_{3,3}' \bar{\beta}_{3,3} x_{22} = O_p(T^{-1}), \tag{30}$$

in view of (29). Since Lemma 3 (ii) ensures that

$$rank\ \bar{\beta}_{3,3}[x_{21}, x_{22}] = rank\ \bar{\beta}_3 x = rank\ \bar{\beta}_3 = r_3$$

with probability one, the definition $\hat{\lambda}_1 \leq \hat{\lambda}_2 \leq \cdots \leq \hat{\lambda}_r$, together with (29) and (30), yields

$$(x_{22}' \bar{\beta}_{3,3}' \bar{\beta}_{3,3} x_{22})^{-1} = O_p(1), \tag{31}$$

which in turn implies $x_{22}^{-1} = O_p(1)$. In view of (30), the assertion $x_{21} = O_p(T^{-1})$ follows immediately from this result. Also, $x_{11}^{-1} = O_p(1)$ is derived by Lemma 3 (ii). Thus, we complete the proof for (i).

(ii) Inserting the relations (13) and (14) into $\tilde{\beta}_{3,3}'\ \tilde{\beta}_{3,**} = 0$ and using that $y_{.**} = O_p(T^{-1})$ and $x_{21} = O_p(T^{-1})$,

$$x_{22}' \bar{\beta}_{3,3}' \bar{\beta}_{3,3} x_{21} + x_{22}' \bar{\beta}_{3,3}' \gamma_3 y_{.**} + O_p(T^{-2}) = 0, \tag{32}$$

which gives

$$x_{21} = -(\beta_{3,3}'\bar{\beta}_{3,3})^{-1}\bar{\beta}_{3,3}'\gamma_3 y_{\cdot **} + O_p(T^{-2}), \tag{33}$$

noting that $x_{22}^{-1} = O_p(1)$ by (i). Inserting (33) into (13), we have the desired result for (ii).

Proof of Lemma 6: (i) The desired result follows immediately from Lemma 3 (iii) and (12) as $\bar{\beta}_3 = 0$.

(ii) The supposition

$$\tilde{\beta}_3(1) = \tilde{\beta}_{3,**},\ \tilde{\beta}_3(2) = \tilde{\beta}_{3,3} \tag{34}$$

and the definition of $\gamma_{3,3}$ gives

$$\gamma_{3,3}' M_3 = \gamma_{3,3}', \tag{35}$$

$$\tilde{\gamma}_{3,3} = \bar{\beta}_{3,3} v + \gamma_{3,3} w, \tag{36}$$

where $v = (\bar{\beta}'\bar{\beta})^{-1}\bar{\beta}_{3,3}'\tilde{\gamma}_{3,3}$ and recall $w = (\gamma_{3,3}'\gamma_{3,3})^{-1}\gamma_{3,3}'\gamma_{3,3}$. Since $y_{\cdot 3} = O_p(T^{-1})$ by Lemma 3 (iii), from (14), (34), (36) and the definition of $\gamma_{3,3}$, we have

$$\begin{aligned}\gamma_{3,3}'\tilde{\beta}_3(2)\{\tilde{\beta}_3'(2)\tilde{\beta}(2)\}^{-1} &= \tilde{\gamma}_{3,3}'\bar{\beta}_{3,3}(\bar{\beta}_{3,3}'\bar{\beta}_{3,3})^{-1}x_{22}'^{-1} + O_p(T^{-1})\\ &= v\,x_{22}'^{-1} + O_p(T^{-1}),\end{aligned}$$

which implies that $v = O_p(T^{-1})$ since $x_{22}^{-1} = O_p(1)$ by Lemma 4 (i) and $\tilde{\gamma}_{3,3}'\tilde{\beta}_3(2) = 0$ by the definition. Also, it follows that $w = O_p(1)$ and $w^{-1} = O_p(1)$ since $\tilde{\gamma}_{3,3}'\tilde{\gamma}_{3,3} = I_{s3}$. Therefore, by Lemma 4 (ii), (34) and (35),

$$\mathrm{w}'^{-1}\, T\, \tilde{\gamma}_{3,3}'\tilde{\beta}_3(1) = \gamma_{3,3}'\gamma_3 T y x^{-1} x S_{**}' + O_p(T^{-1}). \tag{37}$$

Recalling that $\gamma_3 = [\gamma_{**,3}, \gamma_{3,3}]$ with $\gamma_{**,3}$ defined in (16),

$$\gamma_{3,3}'\gamma_3 = S_3. \tag{38}$$

Noting that

$$x S_{**}' x_{11}^{-1} = S_{**}' + O_p(T^{-1})$$

by Lemma 4 (i), (37) and (28), together with Lemma 3 (iii), lead to the conclusion.

Proof of Lemma 7: (i) Letting

$$W(1) = (\tilde{\beta}'\tilde{\beta})^{-1}\tilde{\beta}'\,\tilde{\gamma},\ \ W(2) = (\gamma'\gamma)^{-1}\gamma'\,\tilde{\gamma}$$

by the same argument as used for $\tilde{\gamma}_{3,3}$ in the proof of Lemma 6,

$$\tilde{\gamma} = \gamma\, W(2) + O_p\,(T^{-1}),$$

therefore,

$$\tilde{\gamma}_3 = \gamma_3\, W(2) + O_p\,(T^{-1}),$$

and

$$W(2) = O_p\,(1),\ \ W^{-1}(2) = O_p\,(1).$$

Using the results in Lemma 2, together with the results for $\tilde{\beta}$ and the above results, we have

$$x'\,\tilde{F}x^{-1} \rightarrow_p F_1,\ \ \tilde{F}_2 \rightarrow_d F_2. \tag{39}$$

Since $k_3 \leq s$ by supposition, the existences of $\tilde{F}_2^{-1}$ and F_2^{-1} are ensured. Hence, (39) completes the proof of (i). Notice that (39) hilds even if $r > k_{**}$.

(ii) Since $x\,S'_{**}\,x'_{11} = S'_{**} + O_p\,(T^{-1})$ as noted in the proof of Lemma 6,

$$\begin{aligned} x'_{11}\tilde{G}_1 x_{11}^{-1} &= x'_{11} S_{**}\, x'\, x'^{-1}\,\tilde{F}_1\, x^{-1}\, x\, S_{**}'\, x_{11}^{-1} \\ &= S_{**}\, x'^{-1}\,\tilde{F}_1\, x^{-1}\, S'_{**} + O_p\,(T^{-1}), \end{aligned}$$

that is,

$$x'_{11}\,\tilde{G}_1 x_{11}^{-1} = S_{**}\, x'^{-1}\,\tilde{F}_1\, x^{-1}\, S'_{**} + O_p\,(T^{-1}). \tag{40}$$

On the other hand, the structure of γ gives

$$G_2 = \gamma'_{3,3} F_2 \gamma_{3,3}. \tag{41}$$

Also, by the result for $\gamma_{3,3}$ in the proof of Lemma 6 (ii),

$$\tilde{G}_2 = w' \gamma'_{3,3} \tilde{F}_2 \tilde{\gamma}_{3,3} w + O_p(T^{-1}). \tag{42}$$

(39), together with (40), (41) and (42), leads to

$$x'^{-1}_{11} \tilde{G} x'_{11} \rightarrow_p G_1, \ w'^{-1} \tilde{G}_2 w^{-1} \rightarrow_d G_2. \tag{43}$$

The desired result for (ii) follows immediately from (43).

Proof of Lemma 8: Define the $k_3 \times n_1$ matrix R as

$$R = [0, \bar{\beta}_{3,3} x_{22}(1)] \ \ if \, r_3 < r,$$
$$= \bar{\beta}_{3,3} x_{22}(1) \ \ if \, r_3 = r.$$

By supposition, (12) is expressed as

$$\tilde{\beta}_3(1) = R + O_p(T^{-1}), \ \tilde{\beta}_3(2) = \bar{\beta}_{3,3} x_{22}(2) + O_p(T^{-1}), \tag{44}$$

since $y = O_p(T^{-1})$ by Lemma 3 (iii). Also, letting

$$\bar{v} = \{x'_{22}(2) \bar{\beta}'_{3,3} \bar{\beta}_{3,3} x_{22}(2)\}^{-1} x'_{22}(2) \bar{\beta}'_{3,3} \tilde{\gamma}_{3,3}$$

and noting

$$x'_{22}(2) \bar{\beta}_{3,3} \bar{\beta}'_{3,3} \bar{\beta}_{3,3} \bar{x}_{22}(1) = 0$$

in view of (44), the same argument as used in the proof of Lemma 6 yields

$$\tilde{\gamma}_{3,3} = \delta \bar{w} + \bar{\beta}_{3,3} x_{22}(2) \bar{v}, \tag{45}$$

with $\bar{w} = O_p(1)$, $\bar{w}^{-1} = O_p(1)$ and $\bar{v} = O_p(T^{-1})$. Since it is easy to check that

$$\Psi_{11} = \bar{x}'_{22}(1) \bar{\beta}'_{3,3} \bar{\beta}_{3,3} x_{22}(1) = O_p(1), \ \ \delta' R = \Psi,$$

the first relation in (44), together with (45), the desired result for the lemma.

Proof of Lemma 9: First, note that $x\,(1, 1) = O_p\,(1)$ and $x^{1}\,(1, 1) = O_p\,(1)$ by either Lemma 3 (ii) or Lemma 4 (i). It is also obvious that

$$x\,\bar{S}'_{**}\;x^{-1}\,(1, 1) = \bar{S}'_{**} + O_p\,(T^{-1}).$$

Recalling that

$$\begin{aligned}\gamma_{**,3} = &- b\,(1)\,\{b'\,(1)\,b\,(1)\}^{-1}\,\bar{x}'_{22}\,(1)\,\bar{\beta}_{**,3}\,\gamma_{**,**}\\ &- b\,(2)\,\{b'\,(2)\,b\,(2)\}^{-1}\,x'_{22}\,(2)\;\bar{\beta}_{**,3}\,\gamma_{**,**},\end{aligned}$$

where $b\,(1) = \bar{\beta}_{3,3}\,\bar{x}_{22}\,(1)$, $b\,(2) = \bar{\beta}_{3,3}\,x_{22}\,(1)$, and $b'\,(1)\,b\,(2) = 0$, for δ defined in Lemma 8 and γ_3 in either (16) or (17), we have

$$\delta'\,\gamma_3 = \bar{S}_3.$$

Noting these results and using the same argument as used in the proof of Lemma 7 (ii) and (45) instead of (36), this lemma can be easily shown.

Proof of Theorem 1: For the case in which $1 \leq r_3 = n_2 \leq m_1 - 1$, let

$$\tilde{b}_{n_2} = (G_1^{-1/2} \otimes G_2^{-1/2})\;vec\;w'^{-1}\,T\;\tilde{\gamma}'_{3.3}\,\tilde{\beta}_3\,(1)\,x_{11}^{-1}.$$

Then we have

$$\begin{aligned}\tilde{C}\,(n_2 + 1) &= \{vec\;T\;\tilde{\gamma}'_{3,3}\;\tilde{\beta}_3\,(1)\}'\;(\tilde{G}_1^{-1} \otimes \tilde{G}_2^{-1})\;\{vec\;T\;\tilde{\gamma}'_{3,3}\;\tilde{\beta}_3(1)\}\\ &= \tilde{b}'_{n_2}(G_1^{1/2} \otimes G_2^{-1/2})\;(x_{11} \otimes w)\;(\tilde{G}_1^{-1} \otimes \tilde{G}_2^{-1})\;(x'_{11} \otimes w')\\ &\quad (G_1^{1/2} \otimes G_2^{1/2})\;\tilde{b}_{n_2}.\end{aligned}$$

Letting

$$\bar{b} = (G_1^{-1/2} \otimes G_2^{-1/2})\;vec\;S_3\,B_1^{-1}\,B_2\;\Phi\;S'_{**}$$

and applying Lemma 5 to b_{r_3}, we see that $\bar{b}$ is ditributed as $N\,(0, I_{s_3 r_{**}})$. Also, Lemma 6 (ii) and Lemma 7 (ii) assert

$$b_{n_2} \to_d \bar{b},$$

$$(G_1^{1/2} \otimes G_2^{1/2})(x_{11} \otimes w)(\tilde{G}_1^{-1} \otimes \tilde{G}_2^{-1})(x_{11}' \otimes w')(G_1^{1/2} \otimes G_2^{1/2}) \to_p I_{s_3 r_{**}},$$

which implies

$$\tilde{C}(n_2 + 1) \to_d \bar{b}' \bar{b}.$$

It is obvious that the distribution of $\bar{b}' \bar{b}$ is $\chi^2_{s_3 r_{**}}$. On the other hand, for the case in which $1 \leq n_2 < r_3 \leq m_1$, define

$$C(n_2 + 1) = \{vec\ \tilde{\gamma}_{3,3}' \tilde{\beta}_3(1)\}' (\tilde{G}_1^{-1} \otimes \tilde{G}_2^{-1}) \{vec\ \tilde{\gamma}_{3,3}' \tilde{\beta}_3(1)\}.$$

Then C $(n_2 + 1) = O_p(1)$ and $C^{-1}(n_2 + 1) = O_p(1)$ are ensured by Lemmas 8 and 9. Thus, we derive the desired result for either $j \geq 2$ as $k_3 \leq s$ or $j \geq k_3 - s + 2$ as $k_3 > s$.

The remaining case, i. e. the result for $j = 1$ as $k_3 < s$, is also established in a similar manner, using Lemma 6 (i) under the null hypothesis and the result $rank\ \tilde{\beta}_3 \geq 1$ with probability one, which is obvious by Lemma 2 and (12), as well as Lemma 7 (i) under the alternative hypothesis.

Proof of Theorem 2: First, define the $(k_{**}p - k_{**} + r) \times (k_{**}p - k_{**} + r)$ matrix D_1 and $(k_3p - k_3 + r_3) \times (k_3p - k_3 + r_3)$ matrix D_2 as

$$D_1 = \begin{bmatrix} I_{k_{**}(p-1)} & 0 \\ 0 & x \end{bmatrix},$$

$$D_2 = \begin{bmatrix} I_{k_3(p-1)} & 0 \\ 0 & x_{22} \end{bmatrix} \quad if\ r_3 > 0,$$

$$= I_{k_3(p-1)} \quad if\ r_3 = 0.$$

Then, in view of Lemma 1, Lemma 3 or 4, together with (14), yields

$$\tilde{Z}'\tilde{X}/T = D_2' Z' X D_1 / T + O_p(T^{-1}),$$

$$\tilde{Z}'\tilde{Z}/T = D_2' Z' Z D_2 / T + O_p(T^{-1}),$$

$$\tilde{Z}'\tilde{X}/T = D_2' Z' X D_1 / T + O_p(T^{-1}),$$
$$\tilde{X}'\tilde{X}/T = D_1' X' X D_1 / T + O_p(T^{-1}),$$
$$\tilde{Z}'\tilde{X}/T = D_2' Z' X D_1 / T + O_p(T^{-1}),$$
$$\tilde{Z}'\,\Delta Y_1/T = D_2' Z'\,\Delta Y_1 / T + O_p(T^{-1}),$$
$$\tilde{Z}'\,\Delta Y_1/T = D_1' X'\,\Delta Y_1 / T + O_p(T^{-1}),$$

with

$$D_i = O_p(1),\ D_i^{-1} = O_p(1),\ i = 1, 2,$$

which in turn leads to

$$\tilde{A} = D_2^{-1}(Z' M_X Z / T)^{-1}(Z' M_X\,\Delta Y_1 / T) + O_p(T^{-1}), \tag{46}$$

where $M_X = I_{T\text{-}p} - X(X'X)^{-1}X'$. Substituting the right-hand side of (24) for ΔY_1 in (47),

$$vec\, T^{1/2}(\tilde{A} - D_2^{-1}A) = \{\mathrm{I}_{k_1} \otimes D_2^{-1}\}\,(Z' M_X Z / T)^{-1}\}$$
$$vec\, Z' M_X\ \bar{\varepsilon}_1 / T^{1/2} + O_p(\mathrm{T}^{-1/2}). \tag{47}$$

On the other hand, because of Lemma 2 (i), (iii) and (iv), we can find a $(k_{**}p - k_{**} + r) \times (k_{**}p - k_{**} + r)$ positive definite matrix Λ_Z such that

$$Z' M_X Z / T \rightarrow_p \Lambda_Z. \tag{48}$$

Also, by Lemma 2 (ii),

$$\hat{\Lambda}_{11} \rightarrow_p \Lambda_{11}. \tag{49}$$

Combining (46) with (49), we obtain

$$\Lambda^{1/2}(Z' M_X Z / T)^{-1} D_2^{-1}(\tilde{Z}' M_{\tilde{X}} \tilde{Z} / T)$$
$$D_2^{-1}(Z' M_X Z / T)^{-1}\Lambda^{1/2} \rightarrow_p I_{k_{**}p - k_{**} + r}. \tag{50}$$

Noting

$$\tilde{C} = (vec\, T^{1/2}\tilde{A})' \{\Lambda_{11}^{-1} \otimes (\tilde{Z}' M_{\tilde{X}} \tilde{Z})\}(vec\, T^{1/2}\tilde{A})$$

and using (47), under the null hypothesis (10), we have

$$\tilde{C} = \eta' (\Lambda_{11}^{1/2}\, \Lambda_{11}^{-1} \Lambda_{11}^{1/2} \otimes \tilde{\textstyle\sum}_Z)\, \eta + O_p\,(T^{-1/2}), \tag{51}$$

where

$$\eta = (\Lambda_{11}^{-1/2} \otimes \Lambda_{11}^{-1/2})\; vec\, Z'\, M_X\, \bar{\varepsilon}_1 / T^{1/2},$$

$$\tilde{\textstyle\sum}_Z = \Lambda_Z^{1/2} (Z'\, M_X\, Z \,/\, T)^{-1}\; D_2'^{-1} (\tilde{Z}' M_{\tilde{X}} \tilde{Z} \,/\, T)\, D_2^{-1} (Z'\, M_X\, Z \,/\, T)^{-1}\, \Lambda_Z^{1/2}.$$

Since both Z and X are constructed by stationary series which possess some valid Wold representations using ε_{it}, by the assumptions on ε_{it}, (48) and (49), the Liapounov's central limit theorem (see Loeve (1977, p. 287), e. g.) is applicable to η, that is,

$$\eta \rightarrow_d N\,(0, I_{k_1\,(k_{**}p\,-\,k_{**}\,+\,r)}). \tag{52}$$

Note that Lemma 1 of Toda and Phillips (1993) can also derive (52). Thus, through the equations from (49) to (52), we establish that the asymptotic distribution of $\tilde{C}$ is $\chi^2_{\bar{n}}$ under (10), where $\bar{n} = k_1\,(k_{**}p - k_{**} + r)$.

Note that the results from (46) to (50) still hold under the alternative hypothesis. Define

$$C = (vec\, D_2\tilde{A})' \{\hat{\Lambda}_{11}^{-1} \otimes D_2'^{-1}\, (\tilde{Z}' M_{\tilde{X}} \tilde{Z} \,/\, T)\, D_2^{-1}\}\, (vec\, D_2\tilde{A}).$$

If (10) does not hold,

$$D_2\tilde{A} = A + O_p\,(T^{-1/2}),\;\; \mathrm{A} \neq 0 \tag{53}$$

must be satisfied by (47). (53) immediately gives the desired result for the alternative hypothesis.