92-F-7

Tests of Unit Roots Hypotheses in Econometric Models

by

Naoto Kunitomo
The University of Tokyo

August 1992

Tests of Unit Roots Hypotheses in Econometric Models *

Naoto Kunitomo †

July 1992

Abstract

A number of statistical procedures for testing the unit roots hypotheses has been proposed by statisticians and econometricians. This paper is unifying many of the previous studies on unit roots tests in the framework of a multivariate regression model and developing some new test statistics.

We give a convenient quadratic representation of the limiting distributions of test statistics using stochastic integrals with respect to the Brownian Motion. The test procedures in this paper include the statistics for testing the unit root, the double unit roots, the seasonal unit roots, and the cointegrating relations as special cases.

1 Introduction

An important underlying assumption in many of the traditional econometric analyses has been the condition that the stochastic part of econometric model is stationary in some statistical sense. Although this assumption were often made a priori ground in practice, it may be advisable to examine this condition from a statistical view. In this respect, a number of statistical testing procedures for this problem has been proposed by statistician

*This research was supported by Grant-in-Aid 04301071 of the Ministry of Education, Science, and Culture at the Faculty of Economics, University of Tokyo

†Faculty of Economics, University of Tokyo, JAPAN

and econometricians in the past decade. The problem of testing this condition has been often called the unit root test and the cointegration test. For instance, the test procedures of Dickey and Fuller (1979), Hasza and Fuller (1979), Dickey and Fuller (1981), Hasza and Fuller (1982), Phillips (1987), Phillips and Perron (1988), Perron (1988) among many others have been drawn attention and have been applied in empirical works under the the name of unit root test. Also the test procedures proposed by Engle and Granger (1987), Phillips and Ouliaris (1990), and Johansen (1991) have been also drawn special attention among econometricians under the name of cointegration tests. There have been many econometric empirical studies using the statistical testing procedures of the unit roots hypotheses mainly in the areas of macroeconomics and financial economics. However, since many different testing procedures have been introduced, it may be difficult to understand the meaning of tests in some cases and how to choose one from many tests in many cases.

The main purpose of this paper is to derive systematically several test procedures for the condition of unit roots and to obtaine the relationships among many test statistics. We start our discussion from the general reduced form equations in the traditional econometric framwork and formulate them in a multivariate regression model. Then we consider the hypotheses on regression coeffiicients which includes the unit root hypothesis, the double unit roots hypothesis, the seasonal unit roots hypothesis, and the cointegration hypothesis as special cases. We shall introduce a general class of test statistics, which includes the likelihood ratio (LR) test, the Lagrange Multiplier (LM) test, and the Wald test for these hypotheses as special cases. In this framework the test statistics we shall derive include many of the test procedures mentioned above and also explore the possibility of constructing some new test procedures.

Secondly, we shall derive the asymptotic distributions of the test criteria under a set of general conditions for martingale difference sequences using a certain type of invariance principle. We shall derive a general form of limiting distributions of test statsitics under the null hupotheses. It gives some insight to the problem of unit roots tests and may be useful for obtaining the percentage points of ststistics based on simulations. Besides, we shall show that they include many previous results as special cases by the use of Ito's formula in the theory of stochastic integration.

In Section 2, we shall formulate the unit roots hypotheses in a multivariate regression model. In Section 3, we introduce a general class of test statistics for these hypotheses. In Section 4, we derive the general forms of limiting distributions of test statistics. In Section5, some concluding remarks on the distributions of tests statistics will be given. The proofs of Theorems are given in Section 6.

2 Unit Roots Hypotheses in a Multivariate Regression Model

We first consider a simple univariate AR(1) model

$$y_t = a_0 + a_1 y_{t-1} + v_t, \tag{2.1}$$

where v_t are independently, identically, and normally distributed disturbance terms with $E(v_t) = 0$ and $E(v_t v_t') = \sigma^2$. Then the unit root hypothesis without drift in this model is given by

$$H_0 : a_0 = 0, \ a_1 = 1. \tag{2.2}$$

This testing problem was first investigated by Dickey and Fuller (1979). Subsequently, a number of studies have been done mainly in order to deal with more general unit roots problems for applications. Especially, a number of applied works has been conducted using more general tests than those developed by Dickey and Fuller (1979) in some sense because of many empirical problems encountered by macroeconomists and financial economists.

From a statistical point of view, it seems there can be three directions in the generalizations of this testing problem. The first one is to consider more general serial correlation structures on v_t. One interesting line of reserach in this derection may be a series of papers since Phillips (1987). The second one is to allow some deterministic parts (or exogenous variables in some econometric sense) as well as stochastic parts in the statistical models. The third one is to consider the multivariate versions of the testing problems of unit roots hypotheses. In this section we shall try to unify these possible extensions in a systematic way.

For this purpose, we consider a system of linear structural equations in econometric models implicitly. The reduced form equations for the G-dimensional dependent vector y_t can be often written as

$$(2.3) \qquad y_t = \Gamma z_t^* + A_1 y_{t-1} + \cdots + A_p y_{t-p} + v_t,$$

where z_t^* is a $K^* \times 1$ vector of strictly exogenous variables, Γ is a $G \times K^*$ coefficient matrix, $A_1, \ldots, A_p$ are $G \times G$ coefficient matrices, and v_t are $G \times 1$ vector of disturbances. Re-arranging the explanatory variables in the right hand side of equations, we can re-write these equations as

$$(2.4) \qquad y_t = \boldsymbol{\beta} z_t + v_t,$$

where z_t is a $K \times 1$ vector of predetermined variables including $z_t^*, y_{t-1}, \ldots, y_{t-p}$ (or their linear combinations), and $\boldsymbol{\beta}$ is a $G \times K$ regression coefficient matrix. This is the standard multivariate regression model.

Let $\mathcal{F}_t$ be an increasing sequence of σ-fields generated by $z_1, v_1, \cdots, z_t, v_t, z_{t+1}$. Then z_t is $\mathcal{F}_{t-1}$ -measurable and v_t is $\mathcal{F}_t$-measurable. We assume that the disturbance terms v_t are a sequence of martingale differences with

$$(2.5) \qquad E(v_t|\mathcal{F}_{t-1}) = 0 \quad \text{a.s.},$$

$$(2.6) \qquad E(v_t v_t'|\mathcal{F}_{t-1}) = \Omega_t \quad \text{a.s.}.$$

In this paper we shall consider two types of hypotheses on the coefficient matrix $\boldsymbol{\beta}$ in (2.4). In roder to state these hypotheses, we partition $z_t' = (z_t^{(1)'}, z_t^{(2)'})'$ into K_1 and K_2 vectors of predetermined variables and $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)$ into $G \times (K_1 + K_2)$ matrices. Then the general linear hypothesis on the coefficient matrix $\boldsymbol{\beta}$ in the multivariate analysis (Chapter 8 of Anderson (1984), for example) is given by

$$(2.7) \qquad H_1 : \boldsymbol{\beta}_2 = \boldsymbol{\beta}_2^*,$$

where $\boldsymbol{\beta}_2^*$ is a fixed $G \times K_2$ matrix. Another important hypothesis in the multivariate analysis is the rank condition on the coefficient matrix $\boldsymbol{\beta}_2$, which is given by

$$H_2 : \mathrm{rank}(\boldsymbol{\beta}_2) = r, \tag{2.8}$$

where $r < G$. This hypothesis, originally considered by Anderson (1951), is mathematically equivalent to the block identification hypothesis for a system of simultaneous structural equations in econometrics. The latter testing problem has been systematically investigated by Anderson and Kunitomo (1990), (1992).

We are now in the position to relate two hypotheses in the multivariate analysis with the testing problems of unit roots hypotheses in econometrics. In the following discussions we mainly consider four examples, namely, the simple unit roots tests, the double unit roots tests, the seasonal unit roots tests, and the cointegration tests, although there can be some other interesting examples for applications. In order to state these hypotheses, we also partition $K^* \times 1$ vector of exogenous variables $z_t^* = (z_{1t}^{*'}, z_{2t}^{*'})'$ into K_1^* and K_2^* vectors, and $\Gamma = (\Gamma_1, \Gamma_2)$ into $G \times (K_1^* + K_2^*)$ matrices.

(i)*Simple Unit Roots Case*

We rewrite (2.3) as

(2.9)

$$\begin{aligned} y_t &= \Gamma z_t^* + A_1 y_{t-1} + \cdots + (A_{p-1} + A_p) y_{t-(p-1)} - A_p \Delta y_{t-(p-1)} + v_t \\ &= \Gamma z_t^* + B_1 y_{t-1} + \sum_{i=2}^{p} B_i \Delta y_{t-(i-1)} + v_t, \end{aligned}$$

where the difference operator Δ is defined by $\Delta y_t = y_t - y_{t-1}$, and

$$B_1 = \sum_{i=1}^{p} A_i, \quad B_j = -\sum_{i=j}^{p} A_i \ (j = 2, \ldots, p). \tag{2.10}$$

Then the unit roots hypothesis can be rewritten as

$$H_1^{(1)} : \Gamma_2 = O, B_1 = I_G. \tag{2.11}$$

For testing this hypothesis, we assume that the abosolute values of all roots of

$$|\lambda^{p-1} I_G - \sum_{i=2}^{G} \lambda^{p-i} B_i| = 0 \tag{2.12}$$

are less than one.

When G=1, the hypothesis given by (2.11) has been extensively investigated as the univariate unit root problem. In Dickey and Fuller (1979), (1981) they used two exogenous variables of constants and time trend, and hence $z_t^* = (1, t - T/2)', t = 1, 2, \ldots, T$. Also Perron (1988) used some exogenous variables of dummuy variables. One example in his paper can be expressed in our notation as $z_t^* = (1, DU_t, t)$, where

$$DU_t = \begin{cases} 0 & \text{if } t < \lambda T \\ 1 & \text{if } \lambda T \leq t \leq T. \end{cases}$$

We note that all unit roots hypotheses discussed by Perron (1988) are special cases of (2.11) when G=1.

(ii) Double Unit Roots Case

The second example is the hypothesis of double unit roots. From (2.3), it can be further rewritten as

$$\begin{aligned} y_t &= \Gamma z_t^* + B_1 y_{t-1} + \cdots - B_p \Delta^2 y_{t-(p-2)} + v_t \\ &= \Gamma z_t^* + C_1 y_{t-1} + C_2 \Delta y_{t-1} + \sum_{i=3}^{p} C_i \Delta^2 y_{t-(i-2)} + v_t, \end{aligned} \tag{2.13}$$

where the double difference operator Δ^2 is defined by $\Delta^2 y_t = \Delta y_t - \Delta y_{t-1}$, and

$$C_1 = B_1, \quad C_2 = \sum_{i=2}^{p} B_i, \quad C_j = -\sum_{i=j}^{p} B_i \quad (j = 3, \ldots, p). \tag{2.14}$$

Then the hypothesis of double unit roots is defined by

$$H_1^{(2)} : \Gamma_2 = O, \quad C_1 = C_2 = I_G. \tag{2.15}$$

For testing this hypothesis, we assume that the abosolute values of all roots of

$$|\lambda^{p-2} I_G - \sum_{i=3}^{G} \lambda^{p-i} C_i| = 0 \tag{2.16}$$

are less than one.

When G=1, the hypothesis of double unit roots given by (2.14) has been investigated. Hasza and Fuller (1979) considered the case when there are no exogenous variables and also the case with $z_t^* = (1, t)'$.

(iii) Seasonal Unit Roots Case

The third example is the hypothesis of seasonal unit roots. Let d be the seasonal lag, which includes 2, 4, and 12 as special cases. The seasonal difference operator Δ_d is defined by $\Delta_d y_t = y_t - y_{t-d}$. For convenience, in this paper we assume that

$$p \geq d + 1. \tag{2.17}$$

Using the relation that $\Delta y_t = -\Delta_d \Delta y_{t+d} + \Delta y_{t+d}$, we rewrite (2.3) as

(2.18)
$$\begin{aligned} y_t &= \Gamma z_t^* + B_1 y_{t-1} + B_2 \Delta y_{t-1} + \cdots + B_{d+1} \Delta y_{t-d} + \cdots + B_p \Delta y_{t-(p-1)} + v_t \\ &= \Gamma z_t^* + D_1^* y_{t-1} + D_2^* \Delta y_{t-1} + \cdots + D_{d+1}^* \Delta y_{t-d} \\ &\quad + D_{d+2}^* \Delta_d \Delta y_{t-(d+1)} + \cdots + D_p^* \Delta_d \Delta_{t-(p-1-d)} + v_t, \end{aligned}$$

where

$$\begin{aligned} D_1^* &= B_1, \\ D_j^* &= \sum_{i=0}^{[p/d]} B_{[j+di]}, \quad (j = 2, \ldots, d+1), \\ D_j^* &= -\sum_{i=0}^{[p/d]} B_{[j+di]}, \quad (j = d+2, \ldots, p). \end{aligned} \tag{2.19}$$

Further, it can be rewritten as

$$y_t = \Gamma z_t^* + D_1 y_{t-1} + \sum_{i=2}^{d+1} D_i \Delta_{d+2-i} y_{t-(i-1)} + \sum_{i=d+2}^{p} D_i \Delta_d \Delta y_{t-(i-1-d)} + v_t, \tag{2.20}$$

where

$$D_1 = D_1^*, \quad D_2 = D_2^*, \quad D_j = D_j^* - D_{j-1}^*, \quad (j = 3, \ldots, d+1). \tag{2.21}$$

Then the hypothesis of seasonal unit roots is given by

$$H_1^{(3)} : D_1 = D_{d+1} = I_G, \quad D_2 = \cdots = D_d = O, \quad \Gamma_2 = O. \tag{2.22}$$

For testing this hypothesis, we assume that the abosolute values of all roots of

$$|\lambda^{p-d-1} I_G - \sum_{i=d+2}^{G} \lambda^{p-i} D_i| = 0 \tag{2.23}$$

are less than one.

When G=1, Hasza and Fuller (1981) has considered this testing problem. However, it should be noted that they further assumed that $D_2 = \cdots = D_d = O$, $\Gamma_2 = O$ under both the null as well as the alternative hypotheses. The double unit roots case can be considered as a special case of the seasonal unit roots model when d=1.

(iv) Cointegration Case

The last example is the test of cointegration relations. From (2.9), we also can rewrite the model as

$$\Delta y_t = \Gamma z_t^* + B_1^* y_{t-1} + \sum_{i=2}^{p} B_i \Delta y_{t-(i-1)} + v_t, \tag{2.24}$$

where

$$B_1^* = B_1 - I_G. \tag{2.25}$$

Then by the Granger Representation Theorem (see Engle and Granger (1987)), the cointegration relation of $I(1)$ among the variables in y_t implies the condition

$$H_2^{(1)} : \text{rank}(B_1^*) = r < G. \tag{2.26}$$

For testing this hypothesis, we assume that the abosolute values of all roots of

$$(2.27) \qquad |\lambda^p I_G - \sum_{i=1}^{G} \lambda^{p-i} A_i| = 0$$

are less than one except $G_0(= G - r)$ unit roots.

When $G \geq 1$, Johansen (1991) has developed a test procedure for the hypothesis, which is similar to $H_2^{(1)}$. In his model, there is an exogenous constant term and $z_{1t}^* = 1$.

When $r = 0$, the hypothesis $H_2^{(1)}$ is equivalent to the first condition of the hypothesis $H_1^{(1)}$. We need the second condition of $H_1^{(1)}$ if we allow to have some exogenous variables in the multivariate regression model. This consideration brings us a more general hypothesis given by

$$(2.28) \qquad H_2^{(2)} : \text{rank}(B_1^*) = r < G, \quad \Gamma_2 = O.$$

We note that the second condition of $\Gamma_2 = O$ can be regarded as a special case of the hypothesis H_1. Then the hypothesis in (2.26) can be interpreted as a combination of two hypotheses H_1 and H_2.

3 Test Statistics for Hypotheses

We shall introduce some test statistics for the general linear hypothesis and the rank hypothesis in the multivariate regression model. Let the least squares estimator of the coefficent matrix $\boldsymbol{\beta}$ be

$$(3.1) \qquad \hat{\boldsymbol{\beta}} = \sum_{t=1}^{T} y_t z_t' \left(\sum_{t=1}^{T} z_t z_t' \right)^{-1} = Y'Z(Z'Z)^{-1},$$

where T is the number of observations, Y is a $T \times G$ matrix of observations on the variables y_t, $Z = (Z_1, Z_2)$ is a $T \times K$ matrix of observations on the $K(= K_1 + K_2)$ predetermined variables $z_t = (z_t^{(1)}, z_t^{(2)})$. We partition the least squares estimator

$$\hat{\boldsymbol{\beta}} = (\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2), \tag{3.2}$$

which corresponds to $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)$. We apply the standard principle of invariant tests in the multivariate statistical analysis to the general linear hypothesis and the rank hypothesis. (See Section 8.6 of Anderson (1984).) By the standard method, the invariant test statistics should depend on

$$H = (\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta}_2^*)' A_{22.1} (\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta}_2^*), \tag{3.3}$$

and

$$\begin{aligned} G &= T\hat{\Omega} = \sum_{t=1}^{T} (y_t - \hat{\boldsymbol{\beta}} z_t)(y_t - \hat{\boldsymbol{\beta}} z_t)' \\ &= Y' \bar{P}_Z Y, \end{aligned} \tag{3.4}$$

where

$$A_{22.1} = Z_2' \bar{P}_{Z_1} Z_2, \tag{3.5}$$

and $P_Z = Z(Z'Z)^{-1}Z'$ denotes the projection operator onto the space spanned by the column vector of Z, and $\bar{P}_Z = I_T - P_Z$ for any (full column) matrix Z. Furthermore, the only function of the sufficent statistics invariant under the orthogonal transformation from the left has the form of

$$T_1 = T \cdot f(\lambda_1, \ldots, \lambda_G), \tag{3.6}$$

which defines a general class of statistics for testing the hypothesis H_1, where $0 \leq \lambda_1 \leq \cdots \leq \lambda_G$ are the roots of the characteristic equation

$$|H - \lambda G| = 0. \tag{3.7}$$

In this paper we impose the conditions that $f(\cdot)$ is a smooth function, satisfying (i) $f(0, \ldots, 0) = 0$, (ii) $f(\lambda_1, \ldots, \lambda_G)$ is totally differentialble at $(\lambda_1, \ldots, \lambda_G) = (0, \ldots, 0)$, and (iii)

$$\left.\frac{\partial f}{\partial \lambda_i}\right|_{\lambda_1=\cdots,\lambda_G=0} = 1, \quad (i = 1, \ldots, G). \tag{3.8}$$

We note that this restriction is mainly for the simplicity of the following results on test statistics. It is certainly possible to consider more general class of statistics. However, there are many examples of statistics within this class for practical purposes. For instance, Anderson and Kunitomo (1989) has shown that it includes the Likelihood Ratio (LR) tests, the Lagrange Multiplier (LM) tests, and the Wald tests. Under the assumptions that the disturbance terms v_t are normally distributed, and the covariance matrices $\Omega_t = \Omega$ for all t, the log likelihood ratio (LR) times -2 is given by

$$LR_1 = T\sum_{i=1}^{G} \log(1 + \lambda_i). \tag{3.9}$$

The Lagrange Multiplier (LM) statistic is given by

$$LM_1 = T\sum_{i=1}^{G} \frac{\lambda_i}{(1 + \lambda_i)}, \tag{3.10}$$

which has been known as the Bartlett=Nanda=Pillai trace criterion in multivariate analysis. Similarly, a form of the Wald test is given by

$$W_1 = T\sum_{i=1}^{G} \lambda_i, \tag{3.11}$$

which has also been known as the Lawley=Hotelling trace criterion in multivariate analysis.

Under the null hypothesis of H_1, we notice that

$$\hat{\boldsymbol{\beta}}_2' - \boldsymbol{\beta}_2^{*'} = (Z_2'\bar{P}_{Z_1}Z_2)^{-1}Z_2'\bar{P}_{Z_1}V, \tag{3.12}$$

where V is a $T \times G$ matrix of disturbance terms whose t-th row is v_t'. Then the two matrices H and G can be rewritten as

$$G = V'\bar{P}_Z V, \tag{3.13}$$

and

$$\begin{aligned} H &= V'\bar{P}_{Z_1} Z_2 (Z_2'\bar{P}_{Z_1} Z_2)^{-1} Z_2' \bar{P}_{Z_1} V \\ &= V'(P_Z - P_{Z_1})V. \end{aligned} \tag{3.14}$$

Now we consider the second hypothesis of the rank condition given by H_2. Let $0 \le \nu_1 \le \ldots \le \nu_G$ be the roots of the characteristic equation

$$\left| \frac{1}{T}\Theta_T - \nu\Omega \right| = 0, \tag{3.15}$$

where

$$\Theta_T = \boldsymbol{\beta}_2 A_{22.1} \boldsymbol{\beta}_2'. \tag{3.16}$$

The hypothesis of rank condition H_2 is mathematically equivalent to the hypothesis on the characteristic roots $H_\nu : \nu_1 = \cdots = \nu_{G_0} = 0$ and $\nu_{G_0+1} > 0$ where $G_0 = G - r$. The sample analogue of the characteristic equation (3.15) is given by

$$|Y'(P_Z - P_{Z_1})Y - \lambda^* Y'\bar{P}_Z Y| = 0. \tag{3.17}$$

In the above derivation, we have used the relation

$$\hat{\boldsymbol{\beta}}_2 A_{22.1} \hat{\boldsymbol{\beta}}_2' = Y'(P_Z - P_{Z_1})Y. \tag{3.18}$$

Then by the same arguments of the invariance principle in the multivariate analysis we may introduce a general class of test statistics T_2, which is given by

$$T_2 = f(\lambda_1^*, \cdots, \lambda_{G_0}^*), \tag{3.19}$$

where $\lambda_1^*, \cdots, \lambda_{G_0}^*$ are the roots of characteristic equation (3.17). There are also many examples in this class of test statistics. Anderson and Kunitomo (1989) has shown that it includes the LR test, LM test, and a type of Wald test for the null-hypothesis H_2 under the assumption that the disturbance terms $\{v_t\}$ are normally distributed and the homoscedasticity of covariance matrices $\Omega_t = \Omega$, $t = 1, \ldots, T$. As we may expect from the discussion on the hypothesis H_1, the log likelihood ratio times -2 for H_2 is given by

$$LR_2 = T \sum_{i=1}^{G_0} \log(1 + \lambda_i^*). \tag{3.20}$$

Similarly, the LM statitic and the Wald statistic are given by

$$LM_2 = T \sum_{i=1}^{G_0} \frac{\lambda_i^*}{(1 + \lambda_i^*)}, \tag{3.21}$$

and

$$W_2 = T \sum_{i=1}^{G_0} \lambda_i^*, \tag{3.22}$$

respectively.

Johansen (1991) has developed the likelihood ratio statistic for testing the hypothesis of cointegration, which is similar to a special case of LR_2.

The last hypothesis developed in Section 2 is a combination of two hypotheses H_1 and H_2. In order to test the null hypothesis $H_2^{(2)}$, it may be natural to use

$$T_3 = T \left[f(\lambda_1^*, \cdots, \lambda_{G_0}^*) - f(\lambda_1, \cdots, \lambda_{G_0}) \right]. \tag{3.23}$$

A simple form of this type of test statistics is the LR test statistic, which is given by

$$LR_3 = T \left[\sum_{i=1}^{G_0} \log(1 + \lambda_i^*) - \sum_{i=1}^{G_0} \log(1 + \lambda_i) \right]. \tag{3.24}$$

We notice that the first term is the LR test for the hypothesis $H_2^{(1)}$ and the second term is the LR test statistic for the hypothesis of $\Gamma_2 = O$.

4 Asymptotic Distributions of Test Statistics

In this section we shall investigate the asymptotic distributions of the test statistics derived in Section 3 under the null-hypotheses. We assume that the disturbance terms v_t are a sequence of martingale differences given by (2.5) and (2.6) and the conditional covariance matrix $\Omega_t = E(v_t v_t' | \mathcal{F}_{t-1})$ can be a function of $z_1, v_1, \cdots, z_{t-1}, v_{t-1}, z_t$. In the conditional expectation operator, $\mathcal{F}_{t-1}$ is the information set available at time $t-1$. The predetermined variables z_t may includes a finite number of past dependent variables $y_{t-1}, y_{t-2}, \cdots, y_{t-p}$ or their differences. Theoretically, the maximum lag p in the autoregression may depend on the sample size T such that $p/T \to 0$ in the following discussion. However, we treat p as if it was fixed for the simplicity. In order to investigate the limiting distributions of statistics, we prepare the following theorem.

Theorem 1 *Let $z_t, v_t, t = 1, \cdots,$ be a sequence of pairs of random vectors, and let $\mathcal{F}_t$ be an increasing sequence of σ-field such that z_t is $\mathcal{F}_{t-1}-$ measurable and v_t is $\mathcal{F}_t-$ measurable. (i) Let the matrix D_T be $\mathcal{F}_0-$ measurable such that*

$$(4.1) \qquad D_T^{-1} \sum_{s=1}^{[Tt]} z_s z_s' (D_T')^{-1} \xrightarrow{p} \int_0^t m(s) m(s)' ds \equiv M(t),$$

say, where $m(s)$ is a deterministic vector-valued function of s, and

$$(4.2) \qquad \max_{t=1,\cdots,T} z_t' (D_T D_T')^{-1} z_t \xrightarrow{p} 0.$$

Suppose further that $E(v_t | \mathcal{F}_{t-1}) = 0$ a.s. , and $E(v_t v_t' | \mathcal{F}_{t-1}) = \Omega_t$ a.s. ,

$$(4.3) \qquad \sum_{s=1}^{[Tt]} [\Omega_s \otimes D_T^{-1} z_s z_s' (D_T')^{-1}] \xrightarrow{p} \Omega \otimes M(t),$$

where Ω is a constant matrix, and

$$(4.4) \qquad \sup_{s=1,2,\cdots} E[v_s' v_s I(v_s' v_s > a) | \mathcal{F}_{s-1}] \xrightarrow{p} 0$$

as $a \to \infty$. *Also let*

$$(4.5) \qquad S_T(t) = \frac{1}{\sqrt{T}} \sum_{s=1}^{[Tt]} v_s.$$

Then

$$(4.6) \qquad \text{vec}\left(\sum_{s=1}^{[Tt]} \begin{bmatrix} D_T^{-1} z_s \\ \frac{1}{\sqrt{T}} S_T(s) \end{bmatrix} v_s' \right) \xrightarrow{\mathcal{L}} \text{vec}\left(\int_0^t \begin{bmatrix} m(s) \\ B(s) \end{bmatrix} dB(s)' \right),$$

where $B(s)$ *is the vector Brownian motion on* $[0,1] \times \cdots \times [0,1]$ *with* $E(B(1)B(1)') = \Omega$. *(ii) Suppose further*

$$(4.7) \qquad \frac{1}{T} \sum_{t=1}^{T} \Omega_t \xrightarrow{p} \Omega.$$

Then

$$(4.8) \qquad \frac{1}{T} \sum_{t=1}^{T} v_t v_t' \xrightarrow{p} \Omega.$$

The proof given in the Appendix is similar to that of Theorem 1 in Anderson and Kunitomo (1992). The first part of this theorem is a functional central limit theorem or an invariance principle. The proof is based on a functional central limit theorem given by Helland (1982), which was originally derived from a very general central limit theorem by Dvoretzky (1972). The upper part of (4.6) means that

$$(4.9) \qquad \text{vec}\left(\sum_{s=1}^{[Tt]} D_T^{-1} z_s v_s' \right) \xrightarrow{\mathcal{L}} N\left(O, \Omega \otimes \int_0^t m(s) m(s)' ds \right),$$

which is a standard central limit theorem on the discrete time series. Further if we take $z_t = 1$, we simply have

$$\text{(4.10)} \qquad \frac{1}{\sqrt{T}} \sum_{s=1}^{[Tt]} v_s \xrightarrow{\mathcal{L}} B(t).$$

This type of functional central limit theorem was systematically discussed by Billingsley (1968) and has been extensively used in the unit roots problems. The second part of Theorem 1 is a general convergence theorem for martingale difference sequences. The conditions implied in (4.4) are known to be too strong for the desired result. (See Theorem 2 in Anderson and Kunitomo (1992).) But we shall use them in the following for their simplicity. Our conditions in the above theorem are on the conditional second order moments based on the Lindeberge conditions. One important feature is that we allow some types of conditional heteroscedasticities on the covariance matrices of disturbance terms.

We shall make an additional assumption on the exogenous variables z_t^* in (2.3). As $T \to \infty$,

$$\text{(4.11)} \qquad \Psi_T = \frac{1}{T} \text{tr} \left(Z_{-s}^{*'} \bar{P}_{Z^*} Z_{-s}^{*} \right)$$

is bounded uniformly in s, where Z_{-s}^* is a $T \times K^*$ matrix whose t-th row is z_{t-s}^*. It should be noticed that $\Psi_T = O$ and the above condition is automatically satisfied when we have time trends such as $z_t^* = (1, t)'$. In order to obtain the limiting distributions of test statistics under the null hypotheses, we partition the matrix M(t) into $(K_1^* + K_2^*) \times (K_1^* + K_2^*)$ sub-matrices

$$\text{(4.12)} \qquad M(t) = \begin{pmatrix} M_{11}(t) & M_{12}(t) \\ M_{21}(t) & M_{22}(t) \end{pmatrix}.$$

Theorem 2 *Suppose (4.1) to (4.4), (4.7) and (4.11) hold for $\{v_t, z_t^*\}$. Also suppose Ω and $M(1)$ are nonsingular. Then under the null hypothesis either $H_1^{(1)}, H_1^{(2)}$ or $H_1^{(3)}$ the statistic T_1 has the limiting distribution of the form*

$$\text{(4.13)} \qquad T_1^* = tr\left(\tilde{N}' \left[\tilde{M}^{-1} - \begin{pmatrix} M_{11}(1)^{-1} & O \\ O & O \end{pmatrix} \right] \tilde{N} \right),$$

where $B(t)$ is the G-dimensional standard Brownian motion, and the vector of exogenous variables z_t^* and two matrices $\tilde{N}$ and $\tilde{M}$ are defined as in the following three cases. (i) In the simple unit roots case ($H_1^{(1)}$),

$$\tilde{M} = \int_0^1 \begin{pmatrix} m(s) \\ B(t) \end{pmatrix} \begin{pmatrix} m(s) \\ B(t) \end{pmatrix}' dt, \tag{4.14}$$

$$\tilde{N} = \int_0^1 \begin{pmatrix} m(s) \\ B(t) \end{pmatrix} dB(t)' \tag{4.15}$$

and $z_{1t}^* = \Delta z_{2t}^*$. (ii) In the double unit roots case ($H_1^{(2)}$),

$$\tilde{M} = \int_0^1 \begin{pmatrix} m(s) \\ \int_0^t B(s)ds \\ B(t) \end{pmatrix} \begin{pmatrix} m(s) \\ \int_0^t B(s)ds \\ B(t) \end{pmatrix}' dt, \tag{4.16}$$

$$\tilde{N} = \int_0^1 \begin{pmatrix} m(s) \\ \int_0^t B(s)ds \\ B(t) \end{pmatrix} dB(t)' \tag{4.17}$$

and $z_{1t}^* = \Delta^2 z_{2t}^*$. (iii) In the seasonal unit roots case ($H_1^{(3)}$),

(4.18)

$$\tilde{M} = J \int_0^1 \sum_{i=1}^{d} \begin{pmatrix} m(t) \\ \sum_{j=1}^{d} \int_0^t B_{(i-(j-1))}(s)ds \\ \sum_{j=1}^{d} B_{(i-1-(j-1))}(t) \\ \sum_{j=1}^{d-1} B_{(i-2-(j-1))}(t) \\ \vdots \\ \sum_{j=1}^{1} B_{(i-d-(j-1))}(t) \end{pmatrix} \begin{pmatrix} m(t) \\ \sum_{j=1}^{d} \int_0^t B_{(i-(j-1))}(s)ds \\ \sum_{j=1}^{d} B_{(i-1-(j-1))}(t) \\ \sum_{j=1}^{d-1} B_{(i-2-(j-1))}(t) \\ \vdots \\ \sum_{j=1}^{1} B_{(i-d-(j-1))}(t) \end{pmatrix}' dt J',$$

$$
(4.19)\qquad \tilde{N} = \int_0^1 \sum_{i=1}^{d} \begin{pmatrix} m(t) \\ \sum_{j=1}^{d} \int_0^t B_{(i-(j-1))}(s)ds \\ \sum_{j=1}^{d} B_{(i-1-(j-1))}(t) \\ \sum_{j=1}^{d-1} B_{(i-2-(j-1))}(t) \\ \vdots \\ \sum_{j=1}^{1} B_{(i-d-(j-1))}(t) \end{pmatrix} dB_i(t)',
$$

where $z_{1t}^* = \Delta\Delta_d z_{2t}^*$, J *is a* $K^* + G(d+1)$ *diagonal matrix*

$$
(4.20).\qquad J = \begin{pmatrix} I_{K^*} & O & O & O \\ O & d^{-2}I_G & O & O \\ O & O & d^{-1}I_G & O \\ O & O & \ddots & O \\ O & O & O & d^{-1}I_G \end{pmatrix},
$$

and

$$
(4.21)\qquad B_{(i)} = \begin{cases} B_{i+d} & \text{if } i < 1 \\ B_i & \text{if } 1 \le i \le d \ . \end{cases}
$$

When there is not any unit root term in the standard case (i), the statistic T_1 has the limiting distribution of χ^2. If we drop $B(t)$ from $\tilde{M}$ and $\tilde{N}$, the distribution of T_1^* is a χ^2. On the other hand, when there is no exogenous variables in the model, the limiting distribution of T_1 has a simpler form. For instance, in the unit roots case

$$
(4.22)\qquad T_1^* = \mathrm{tr}\left(\int_0^1 B(t)dB(t)'\right)' \left[\int_0^1 B(t)B(t)'dt\right]^{-1} \left(\int_0^1 B(t)dB(t)'\right).
$$

This representation of the limiting distribution has been obtained in the units roots case by Phillips and Durlauf (1986).

When G=1, it has been quite common to use the least squares estimator of the coefficient and its t-ratio statistic for testing the hypothesis of a unit root in (2.9). In this case we have the following simple representation of the limiting distribution.

Theorem 3 *Suppose (4.1) to (4.4), (4.7) and (4.11) hold for* $\{v_t, z_t^*\}$. *Also suppose* Ω *and* $M(1)$ *are nonsingular. Then unedr the null hypothesis* $H_1^{(1)}$ $T(\hat{B}_1 - 1)$ *has the limiting distribution of the form*

$$t_1^* = \frac{\left| \int_0^1 \begin{pmatrix} m(t) \\ B(t) \end{pmatrix} \begin{pmatrix} m(t)dt \\ dB(t) \end{pmatrix}' \right|}{\left| \int_0^1 \begin{pmatrix} m(t) \\ B(t) \end{pmatrix} \begin{pmatrix} m(t) \\ B(t) \end{pmatrix}' dt \right|}. \tag{4.23}$$

Now we consider some special cases we have mentioned in Section 2. The following examples may illustrate the usefulness of representations of the limiting distributions in Theorems 2 and 3.

Example 1: When $G = 1$ and there is no exogenous variables in the model,

$$t_1^* = \frac{\int_0^1 B(t)dB(t)}{\int_0^1 B(t)^2 dt} = \frac{\frac{1}{2}(B(t)^2 - 1)}{\int_0^1 B(t)^2 dt}, \tag{4.24}$$

which has been a well-known representation of the limiting distribution of the least squares estimator since the classical papers by White (1958) and Anderson (1959). We notice that the last equality is a simple consequence of Ito's Lemma in the theory of stochastic integration. This observation is a key fact to get many representations in the following examples. When $G = 1$ and we have linear trend, then $z_t^* = (1, t - T/2)'$ and

$$
(4.25) \qquad t_1^* = \frac{\begin{vmatrix} 1 & 0 & T \\ 0 & \frac{1}{12} & \frac{T-2W}{2} \\ W & \frac{V}{2} & \frac{T^2-1}{2} \end{vmatrix}}{\begin{vmatrix} 1 & 0 & W \\ 0 & \frac{1}{12} & \frac{V}{2} \\ W & \frac{V}{2} & \Gamma \end{vmatrix}} = \frac{\frac{1}{2}((T-2W)(T-6V)-1)}{\Gamma - W^2 - 3V^2},
$$

where we have used the notations of Dikey and Fuller (1979). Since we take exogenous variables mutually orthogonal in this case, the final expression has a simple structure. Similarly, when G=1 and $z_t^* = 1$, (4.13) becomes

$$
(4.26) \qquad \begin{aligned} T_1^* &= \frac{\left| \int_0^1 \begin{pmatrix} 1 \\ B(t) \end{pmatrix} \begin{pmatrix} dt \\ dB(t) \end{pmatrix}' \right|^2}{\left| \int_0^1 \begin{pmatrix} 1 \\ B(t) \end{pmatrix} \begin{pmatrix} 1 \\ B(t) \end{pmatrix}' dt \right|} - \left(\int_0^1 dB(t) \right)^2 \\ &= \frac{\begin{vmatrix} 1 & T \\ W & \frac{T^2-1}{2} \end{vmatrix}}{\begin{vmatrix} 1 & W \\ W & \Gamma \end{vmatrix}} - T^2, \end{aligned}
$$

by using the notations of Dickey and Fuller (1979).

Example 2: When G=1, Perron (1989) used some interesting dummy variables DU_t and DT_t^* as exogenous variables. The final representations of test statistics in Perron (1988) appeared to be quite complicated at first glance. When we let $z_t^* = (1, DU_t, t)$, then the limiting distribution of $T(\hat{\alpha}_A - 1)$ in his paper becomes

$$
(4.27) \qquad t_1^* = \frac{\begin{vmatrix} 1 & 1-\lambda & \frac{1}{2} & \int_0^1 dB \\ 1-\lambda & 1-\lambda & \frac{1-\lambda^2}{2} & \int_\lambda^1 dB \\ \frac{1}{2} & \frac{1-\lambda^2}{2} & \frac{1}{3} & \int_0^1 t dB \\ \int_0^1 B dt & \int_\lambda^1 B dt & \int_0^1 tB dt & \int_0^1 B dB \end{vmatrix}}{\begin{vmatrix} 1 & 1-\lambda & \frac{1}{2} & \int_0^1 B dt \\ 1-\lambda & 1-\lambda & \frac{1-\lambda^2}{2} & \int_\lambda^1 B dt \\ \frac{1}{2} & \frac{1-\lambda^2}{2} & \frac{1}{3} & \int_0^1 t B dt \\ \int_0^1 B dt & \int_\lambda^1 B dt & \int_0^1 tB dt & \int_0^1 B(t)^2 dt \end{vmatrix}}.
$$

Similarly, it is possible to derive the simple representations of the limiting distributions of $T(\hat{\alpha}_B - 1)$ and $T(\hat{\alpha}_C - 1)$ in Perron (1989). Our representations are much simpler than his final expressions.

Example 3: When G=1, Hasza and Fuller (1979) developed several test procedures for testing the double unit roots hypothesis. If we let $z_t^* = (1, t)$, the limiting distribution of $T_1 = \tilde{N}'\tilde{M}^{-1}\tilde{N}$ has a form T_1^* when we take

$$
(4.28) \qquad \tilde{M} = \int_0^1 \begin{pmatrix} 1 \\ t \\ \int_0^t B(s)ds \\ B(t) \end{pmatrix} \begin{pmatrix} 1 \\ t \\ \int_0^t B(s)ds \\ B(t) \end{pmatrix}' dt,
$$

$$
(4.29) \qquad \tilde{N} = \int_0^1 \begin{pmatrix} 1 \\ t \\ \int_0^t B(s)ds \\ B(t) \end{pmatrix} dB'.
$$

By using the notations in Hasza and Fuller (1979), they are

$$
(4.30) \qquad \tilde{M} = \begin{pmatrix} 1 & \frac{1}{2} & W_3 & W_2 \\ \frac{1}{2} & \frac{1}{3} & W_4 & W_2 - W_3 \\ W_3 & W_4 & W_6 - W_2^2 + 2W_2W_3 & \frac{W_2^2}{2} \\ W_2 & W_2 - W_3 & \frac{W_2^2}{2} & W_5 \end{pmatrix},
$$

and

$$\tilde{N} = \begin{pmatrix} W_1 \\ W_1 - W_2 \\ W_1 W_2 - W_5 \\ \dfrac{W_1^2 - 1}{2} \end{pmatrix}. \tag{4.31}$$

Then we notice that T_1^* is mathematically equivalent to $h_3' H_3^{-1} h_3 = 4\Psi_3(4)$ in Hasza and Fuller (1979). To evidence this point, we have

$$W_1 = B(1), \tag{4.32}$$

$$\int_0^1 t dB(t) = B(1) - \int_0^1 B(t)dt = W_1 - W_2, \tag{4.33}$$

$$\int_0^1 \left(\int_0^t B(s)ds \right) dB(t) = B(1) \int_0^1 B(t)dt - \int_0^1 B(t)^2 dt = W_1 W_2 - W_5, \tag{4.34}$$

$$\int_0^1 B(t)dB(t) = \frac{B(1)^2 - 1}{2}, \tag{4.35}$$

for instance. These relations and others needed for verifying (4.30) and (4.31) can be directly obtained from the Ito's formula and the partial integration fourmula in the theory of stochastic integration. Then we confirm that each element of $\tilde{M}$ and $\tilde{N}$ corresponds to that of h_3 and H_3, respectively.

Example 4: When G=1, Hasza and Fuller (1982) assumed $D_2 = \cdots = D_d = 0$ under both the null hypothesis of seasonal unit root and the alternative hypothesis. When there is no exogenous variables, the limiting distribution of T_1 has a simpler form $T_1^* = \tilde{N}' \tilde{M}^{-1} \tilde{N}$, *where*

$$\tilde{M} = J \int_0^1 \sum_{i=1}^d \begin{pmatrix} \sum_{j=1}^d \int_0^t B_j(s)ds \\ \sum_{j=1}^d B_j(t) \\ B_i \end{pmatrix} \begin{pmatrix} \sum_{j=1}^d \int_0^t B_j(s)ds \\ \sum_{j=1}^d B_j(t) \\ B_i \end{pmatrix}' dt J', \tag{4.36}$$

$$(4.37) \qquad \tilde{N} = \int_0^1 \sum_{i=1}^{d} \begin{pmatrix} \sum_{j=1}^{d} \int_0^t B_j(s)ds \\ \sum_{j=1}^{d} B_j(t) \\ B_i \end{pmatrix} dB_i(t)',$$

and

$$J = \begin{pmatrix} d^{-2} & O & O \\ O & d^{-1} & O \\ O & O & d^{-1} \end{pmatrix}.$$

Then it is easily seen that $\tilde{M}$ and $\tilde{N}$ in this paper correspond to H and h in Hasza and Fuller (1982), respectively. Consequently, T_1^* is equivalent to $h'H^{-1}h$ in Hasza and Fuller (1982).

Now we turn to consider the problem of testing cointegration hypotheses. For this purpose, we partition the G_0 -dimensional Brownian motion $B(t) = (B(t)_1', B(t)_2')'$ into $[(G - K_1^* - r) + K_1^*]$ -dimensional Brownian motions. The next two theorems are on the testing a simple hypothesis of cointegration and a composite hypothesis of cointegration and zero restrictions in the multivariate regression model.

Theorem 4 *Suppose (4.1) to (4.4), (4.7) and (4.11) hold for $\{v_t, z_t^*\}$. Also suppose Ω and $M(1)$ are nonsingular, $z_{1t}^* = \Delta z_{2t}^*$, and $G - K_1^* - r \geq 0$. Then under the null hypothesis of $H_2^{(2)}$ the statistic T_2 has the limiting distribution of the form*

$$(4.38) \qquad T_2^* = \operatorname{tr}\left(\tilde{N}' \left[\tilde{M}^{-1} - \begin{pmatrix} M_{11}(1)^{-1} & O \\ O & O \end{pmatrix}\right] \tilde{N}\right),$$

where

$$(4.39) \qquad \tilde{M} = \int_0^1 \begin{pmatrix} m(s) \\ B_1(t) \end{pmatrix} \begin{pmatrix} m(s) \\ B_1(t) \end{pmatrix}' dt,$$

$$\tilde{N} = \int_0^1 \begin{pmatrix} m(s) \\ B_1(t) \end{pmatrix} dB(t)'. \tag{4.40}$$

Notice that the dimension of Brownian motion B(t) is $G_0(= G-r)$ in the cointegration case, while it is G in the unit roots case in Theorem 2. This is because there are r cointegrating relations among the variables in y_t so that the number of independent random walk should be $G-r$. Since we allow some exogenous variables in the model, they may make some effects with the cointegration relations. This is the reason why we partitioned the Brownian motion $B(t)$ in the present situation.

When $G-K_1^*-r<0$, the statistic T_2 may have the limiting distribution of χ^2 with G_0^2 degrees of freedom under the condition that the stochastic order of z_{1t}^* is greater than $\sqrt{T}$. This condition is satisfied if there are a sufficient number of deterministic trend terms.

Example 5: When $r = G-1$ and there is no exogenous variables under both the null and alternative hypotheses, the second term in (4.38) disappears. Then we have

$$T_2^* = \frac{\left(\int_0^1 B(t)dB(t)\right)^2}{\int_0^1 B(t)^2 dt}. \tag{4.41}$$

When we use a different normalization factor, we should change the denominator slightly and we have

$$T_2^\dagger = \frac{\int_0^1 B(t)dB(t)}{\int_0^1 B(t)^2 dt}, \tag{4.42}$$

for instance. Again by using the relation (4.34), $T_2^\dagger$ is identical to the expression of the limiting distribution obtained by Fountis and Dickey (1989).

Example 6: When $G \geq 1$, Johansen (1991) assumed $z_{1t}^* = 1, z_{2t}^* = t$, and the order of seasonal dummy variables is negligible. In this case the limiting distribution T_2^* can be further re-written as

$$(4.43) \quad T_2^* = \mathrm{tr}\left(\int_0^1 F(t)dB(t)'\right)'\left(\int_0^1 F(t)F(t)'dt\right)^{-1}\left(\int_0^1 F(t)dB(t)'\right),$$

where F is the G_0 vector given by

$$(4.44) \qquad F = \begin{pmatrix} B_1(t) - \int_0^1 B_1(t)dt \\ t - \frac{1}{2} \end{pmatrix},$$

which is the same as the limiting distribution of the statistic obtained by Johansen (1991).

Theorem 5 *Suppose (4.1) to (4.4), (4.7) and (4.11) hold for* $\{v_t, z_t^*\}$. *Also suppose* Ω *and* $M(1)$ *are nonsingular,* $z_{1t}^* = \Delta z_{2t}^*$, *and* $G - K_1^* - r \geq 0$. *Then under the null hypothesis of* $H_2^{(2)}$ *the statistic* T_3 *has the limiting distribution of the form*

$$(4.45) \qquad \begin{aligned} T_3^* &= \mathrm{tr}\left(\tilde{N}'\left[\tilde{M}^{-1} - \begin{pmatrix} M_{11}(1)^{-1} & O \\ O & O \end{pmatrix}\right]\tilde{N}\right) \\ &\quad + \mathrm{tr}\left(\tilde{L}'\left[\tilde{M}^{-1} - \begin{pmatrix} \tilde{M}_{11}(1)^{-1} & O \\ O & O \end{pmatrix}\right]\tilde{L}\right), \end{aligned}$$

where $\tilde{M}$ *and* $\tilde{N}$ *are defined by (4.14) and (4.15),*

$$(4.46) \qquad \tilde{M}_{11}(1) = \int_0^1 \begin{pmatrix} m_1(s) \\ B_1(t) \end{pmatrix}\begin{pmatrix} m_1(s) \\ B_1(t) \end{pmatrix}' dt,$$

$$(4.47) \qquad \tilde{L} = \int_0^1 \begin{pmatrix} m(s) \\ B(t) \end{pmatrix} dW(t)',$$

and $B(t)$ *and* $W(t)$ *are* $G_0(= G - r)$*- and* $r-$ *demensional Brownian motions, which are mutually independent.*

The representation T_3^* is the most general one in this paper. The first two terms correspond to the cointegration hypothesis. In addition to these terms, we have two more terms because of the zero restrictions in the model. Since $\tilde{L}$ is independent of $\tilde{M}, \tilde{N}, \tilde{M}_{11}$, the second part has the distribution of χ^2 and its degrees of freedom is

$$\text{df} = \text{r} \cdot \min(K_1^*, G - r). \tag{4.48}$$

When $G \geq 1$ and $r = 0$, the last two terms disappear and we have (4.13)-(4.15) in Theorem 2. In this sense, the first part of Theorem 2 is a special case of Theorem 5. When $G = r$, the stochastic part of y_t is stationary. Then the lower part of the first and second terms disappear, and T_3^* has the limiting distribution of χ^2 with $G \cdot K_1^*$ degrees of freedom. This is the situation when we can use the standard asymptotic theory. Thus Theorem 5 includes both the non-standard unit roots case and the standard stationary case as special cases.

5 Concluding Remarks

First, it may be important to note that the numerical calculation of the limiting distributions of the test statistics is not easy even under the null hypotheses of unit roots. The limiting distributions could be numerically obtained only in some simple cases (see Tanaka (1991), for instance). The percentage points of some test statistics have been tabulated based on simulations. In this respect, the simple representations of the limiting distributions of statistics could be useful for efficient simulations. Chan (1988) has noted that it is quite efficient to use a stochastic integral representation of a test statistic in the univariate case. In the multivariate case, we may expect that this aspect is more evident and the representation of the limiting distributions obtained in Section 4 may be potentially useful for practical purposes.

Second, we have investigated the limiting distributions of test statistics only under the null-hypotheses in Section 4. It is possible to obtain the power functions of test statistics under a sequence of local alternatives. For instance, in the unit roots case we may consider

$$B_1 = exp(C(T)), \tag{5.1}$$

where $C(T) \to 0$ as $T \to \infty$. This type of local alternatives has been called the near-unit root case and investigated by Phillips (1987), (1988), and Chan and Wei (1988) when there is not any exogenous variable. For the cointegtation case, it is also possible to consider a sequence of local alternative hypotheses. Since the rank of B_1^* is r, we may take

$$B_1^*\delta = C(T), \tag{5.2}$$

where δ is a $G \times G_0$ matrix and $G_0 \times G_0$ matrix $C(T) \to 0$ as $T \to \infty$. This type of local alternatives has been investigated as the near-overidentification condition by Kunitomo (1988), and Anderson and Kunitomo (1990) because δ is a coefficient matrix in the system of structural equations. It may be straightforward to extend their results to the present near-cointegration case where δ is usually interpreted as a coefficient matrix of long-term relationships among variables.

Finally, from the arguments discussed in this section, it may be possible to consider more complicated unit roots hypotheses. For instance, we can think of some combinations of the unit roots and seasonal unit roots in a variety of ways. By using the methods similar to Theorems in Section 4, it may be possible to construct test statistics and obtain their limiting distributions in general situations.

6 Proofs of Theorems

Proof of Theorem 1: First we shall show that

$$\operatorname{tr}\left(\sum_{s=1}^{[Tt]} D_T^{-1} z_s v_s' C\right) = \sum_{s=1}^{[Tt]} v_s' D_T^{-1\prime} C z_s' \xrightarrow{\mathcal{L}} \int_0^t dB(s)' C m(s) \tag{6.1}$$

for every C. Let $u_{[Ts]} = CD_T^{-1} z_s$, $s = 1, \ldots, T$. Then the conditions (4.1) and (4.3) imply that

$$\sum_{s=1}^{[Tt]} u_{Ts} u_{Ts}' \xrightarrow{p} C \int_0^t m(s) m(s)' ds C' \equiv F(t), \tag{6.2}$$

say, and

$$\max_{1\le s\le T} \|u_{Ts}\|^2 \xrightarrow{p} 0. \tag{6.3}$$

Let $w_{Ts} = u_{Ts} I(\|u_{Ts}\| \le 1), s = 1,\cdots,T, T = 1,\cdots$. Then $\|u_{Ts}\| \le 1$, and $\Pr(w_{Ts} = u_{Ts}, s = 1,\cdots,T) \to 1$, and

$$\sum_{s=1}^{[Tt]} E\left[(w'_{Ts}v_s)^2 | \mathcal{F}_{s-1}\right] = \sum_{s=1}^{[Tt]} w'_{Ts}\Omega_t w_{Ts} \xrightarrow{p} \mathrm{tr}(\Omega F(t)). \tag{6.4}$$

Also by (4.4), we have

$$\sum_{s=1}^{[Tt]} E\left\{(w'_{Ts}v_s)^2 I\left[\left(w'_{Ts}v_s^2\right) > \delta | \mathcal{F}_{s-1}\right]\right\} \xrightarrow{p} 0. \tag{6.5}$$

Hence the result of (6.1) is a direct consequence of Theorem 3.1 of Helland (1982). Next, we shall show

$$\mathrm{tr}\left\{\frac{1}{\sqrt{T}}\sum_{s=1}^{[Tt]} S_T(s)v'_s C\right\} \xrightarrow{\mathcal{L}} \mathrm{tr}\left\{\int_0^t B(s)dB(s)'C\right\}, \tag{6.6}$$

for every C. From what we have just shown in (6.1),

$$S_T(t) \xrightarrow{\mathcal{L}} B(t). \tag{6.7}$$

Then we can use the proof of Theorem 2.4 in Chan and Wei (1988). The conditions in their theorem are automatically satisfied except one. The only modification we need is to replace the uniform integration condition (4.4) on the conditional covariance matrix for their uniform bounded condition. The last convergence result in Theorem 1 is a simple consequence of Theorem 2.23 in Hall and Heyde (1980). (See Theorem 2 in Anderson and Kunitomo (1992).) □.

Lemma 1 *Let y'_{t-1} and $y^{*'}_{t-1}$ be the t-th row of matrices Y_{-1} and $\bar{P}_{Z^*}Y_{-1}$, respectively. Then under the assumptions of Theorem 2,*

(6.8)
$$\frac{1}{T}\sum_{s=1}^{[Tt]} y^*_{s-1}v'_s \xrightarrow{\mathcal{L}} C\left[\int_0^t B(s)dB(s)' - \int_0^t B(s)m(s)'ds(\int_0^t m(s)m(s)'ds)^{-1}\int_0^t m(s)dB(s)'\right],$$

and

(6.9)
$$\frac{1}{T^2}\sum_{s=1}^{[Tt]} y^*_{s-1}y^{*'}_{s-1} \xrightarrow{\mathcal{L}} C\left[\int_0^t B(s)B(s)'ds - \int_0^t B(s)m(s)'ds(\int_0^t m(s)m(s)'ds)^{-1}\int_0^t m(s)B(s)'ds\right]C',$$

where $C = \sum_{s=0}^{\infty} W_s$ and W_s are the coefficients in the moving average representation of the stationary vector AR process Δy_{t-1}.

Proof of Lemma 1: We shall show (6.8) and (6.9) only when p=2. The method below can be extended easily to the general case with some complex notations. Under the null hypothesis we rewrite

$$y_t = \psi_t + \xi_t + (I_G - B_2^{t+1})y_0, \tag{6.10}$$

where

$$\psi_t = \sum_{s=0}^{t} B_2^s S_1(t-s), \tag{6.11}$$

and

$$\xi_t = \sum_{s=0}^{t} B_2^s \Gamma_1 z^*_{2,t-s}. \tag{6.12}$$

Let also Ψ_{-1} and Ξ_{-1} be $T \times G$ matrices whose t-th rows are ψ_{t-1} and ξ_{t-1}, respectively. Then

$$\text{(6.13)} \qquad \frac{1}{T}\sum_{t=1}^{T} y_{t-1}^{*} v_t' = \frac{1}{T}\Psi_{-1}' \bar{P}_{Z^*} V + \frac{1}{T}\Xi_{-1}' \bar{P}_{Z^*} V.$$

By the condition (4.11) and v_t being a sequence of martingale differences, we have

$$\text{(6.14)} \qquad \frac{1}{T}\Xi_{-1}' \bar{P}_{Z^*} V \xrightarrow{p} 0.$$

We notice that

$$\text{(6.15)} \qquad \begin{aligned} \frac{1}{\sqrt{T}}\psi_{Tt} &= \frac{1}{\sqrt{T}}\sum_{s=0}^{\infty} B_2^s \sum_{j=1}^{[Tt]-s} v_j \\ &= \frac{1}{\sqrt{T}}\sum_{s=0}^{\infty} B_2^s \sum_{j=1}^{[Tt]} v_j - \frac{1}{\sqrt{T}}\sum_{s=0}^{\infty} B_2^s (\sum_{j=[Tt]-s+1}^{[Tt]} v_j). \end{aligned}$$

Since the second term converges to zero, we have

$$\text{(6.16)} \qquad \frac{1}{\sqrt{T}}\psi_{Tt} \xrightarrow{p} CB(t),$$

where $C = \sum_{s=0}^{\infty} B_2^s$ in this case. Then by Theorem 1

$$\text{(6.17)} \qquad \begin{aligned} &\frac{1}{T}\sum_{s=1}^{[Tt]} y_{s-1}^{*} v_s' = \frac{1}{T}\sum_{s=1}^{[Tt]} \psi_{s-1}^{*} v_s' \\ &- (\sum_{s=1}^{[Tt]} \frac{1}{T}\psi_{s-1} z_s^{*'} D_T^{-1})(D_T^{-1}\sum_{s=1}^{[Tt]} z_s^{*} z_s^{*'} D_T^{-1'})^{-1}(D_T^{-1}\sum_{s=1}^{[Tt]} z_s^{*} v_s') \end{aligned}$$

converges to (6.8) as $T \to \infty$. Similarly

(6.18)

$$\begin{aligned} \frac{1}{T^2}\sum_{s=1}^{[Tt]} y_{s-1}^{*} y_{s-1}^{*'} &= \frac{1}{T^2}\Psi_{-1}' \bar{P}_{Z^*} \Psi_{-1} \\ &+ \frac{1}{T^2}\Psi_{-1}' \bar{P}_{Z^*} \Xi_{-1} + \frac{1}{T^2}\Xi_{-1}' \bar{P}_{Z^*} \Psi_{-1} + \frac{1}{T^2}\Xi_{-1}' \bar{P}_{Z^*} \Xi_{-1}. \end{aligned}$$

The last three terms converges to zero as $T \to \infty$. By using Theorem 1, the first term converges to (6.9) as $T \to \infty$. □

Proof of Theorem 2: (i) Let $\Delta y^*_{t-1} = (\Delta y'_{t-1}, \ldots, \Delta y'_{t-(p-1)})'$, $z_{1t} = (z^{*'}_{1t}, \Delta y^{*'}_{t-1})'$, and $z_{2t} = (z^{*'}_{2t}, y'_{t-1})'$. Let also ΔY^*_{-1} be a matrix whose t-th row is Δy^*_{t-1}. Then

(6.19)

$$V'\left(\bar{P}_{Z_1} - \bar{P}_Z\right)V = V'\left\{\bar{P}_{Z^*_1} - \bar{P}_{Z^*_1}\Delta Y^*_{-1}\left(\Delta Y^{*'}_{-1}\bar{P}_{Z^*_1}\Delta Y^*_{-1}\right)^{-1}\Delta Y^{*'}_{-1}\bar{P}_{Z^*_1}\right\}V$$
$$+ V'\left\{\bar{P}_{Z^*} - \bar{P}_{Z^*}(\Delta Y^*_{-1}, Y_{-1})\left[\left(\Delta Y^*_{-1}, Y_{-1}\right)'\bar{P}_{Z^*}\left(\Delta Y^*_{-1}, Y_{-1}\right)\right]^{-1}\left(\Delta Y^*_{-1}, Y_{-1}\right)'\bar{P}_{Z^*}\right\}V.$$

Since $\frac{1}{T\sqrt{T}}Y'_{-1}\bar{P}_{Z^*}\Delta Y_{-1} \xrightarrow{p} 0$, the second parenthethis of (6.19) is asymptotically equivalent to

$$\text{(6.20)}\qquad \begin{aligned} V'\bar{P}_{Z^*}V &- V'\bar{P}_{Z^*}\Delta Y^*_{-1}\left(\Delta Y^{*'}_{-1}\bar{P}_{Z^*}\Delta Y^*_{-1}\right)^{-1}\Delta Y^{*'}_{-1}\bar{P}_{Z^*}V \\ &- V'\bar{P}_{Z^*}Y^*_{-1}\left(Y^{*'}_{-1}\bar{P}_{Z^*}Y^*_{-1}\right)^{-1}Y^{*'}_{-1}\bar{P}_{Z^*}V. \end{aligned}$$

Because Δy^*_t has a stationary autoregressive representation and Ω is a positive definite matrix,

$$\text{(6.21)}\qquad \begin{aligned} \operatorname*{plim}_{T\to\infty}\frac{1}{T}\Delta Y^{*'}_{-1}\bar{P}_{Z^*}\Delta Y^*_{-1} &= \operatorname*{plim}_{T\to\infty}\frac{1}{T}\Delta Y^{*'}_{-1}\bar{P}_{Z^*_1}\Delta Y^*_{-1} \\ &= \Gamma, \end{aligned}$$

which is a positive definite matrix. (See Lemma 2 of Anderson and Kunitomo (1992), for instance.) Then the left-hand side of (6.19) is asymptotically equivalent to

$$\text{(6.22)}\qquad V'\left\{\bar{P}_{Z^*_1} - \bar{P}_{Z^*} + \bar{P}_{Z^*}Y_{-1}\left(Y'_{-1}\bar{P}_{Z^*}Y_{-1}\right)^{-1}Y'_{-1}\bar{P}_{Z^*}\right\}V.$$

(We note that (6.22) is (6.19) when p=1.) We apply Theorem 1 to the first two terms and Lemma 1 to the last term of (6.22), respectively. Then the limiting distribution of $\Omega^{-1/2}V'\left(\bar{P}_{Z_1}-\bar{P}_Z\right)V\Omega^{-1/2}$ is T_1^* in (4.13). In the above derivation, we have used that $\Omega^{-1/2}\times$ (the last term in (6.22)) $\times\Omega^{-1/2}$ converges to

(6.23)

$$\left[\int_0^t dB(s)B(s)' - \int_0^t dB(s)m(s)'\left(\int_0^t m(s)m(s)'ds\right)^{-1}\int_0^t m(s)B(s)'dt\right]$$

$$\times\left[\int_0^t B(s)B(s)'ds - \int_0^t B(s)m(s)'ds\left(\int_0^t m(s)m(s)'ds\right)^{-1}\int_0^t m(s)B(s)'ds\right]^{-1}$$

$$\times\left[\int_0^t B(s)dB(s)' - \int_0^t B(s)m(s)'ds\left(\int_0^t m(s)m(s)'ds\right)^{-1}\int_0^t m(s)dB(s)'\right],$$

$$= \tilde{N}'\tilde{M}^{-1}\tilde{N} - \left[\int_0^t dB(s)m(s)'\left(\int_0^t m(s)m(s)'ds\right)^{-1}\int_0^t m(s)dB(s)'\right].$$

From this representation, it is obvious that the second term of (6.23) is cancelled with the second term of (6.22). Next, we can take a $K\times K$ normalization matrix D_T^* such that

$$\frac{1}{T}V'\bar{P}_Z V = \frac{1}{T}V'V - \frac{1}{T}V'ZD_T^{*-1}\left(D_T^{*-1'}Z'ZD_T^{*-1}\right)^{-1}D_T^{*-1'}Z'V. \tag{6.24}$$

Hence from the last part of Theorem 1, the second term of (6.24) converges to zero and we have

$$\frac{1}{T}V'\bar{P}_Z V \xrightarrow{p} \Omega. \tag{6.25}$$

Hence we obtain the result for the unit roots case.

(ii) The proof of the double unit roots case is included in that of the seasonal unit roots case when $d=1$ below.

(iii) We consider the case of seasonal unit roots hypothesis when $p=d+1$. Under the null hypothesis, we define the stochastic process

$$x_t = \Delta y_t. \tag{6.26}$$

Then we notice that

$$x_t = x_{t-d} + v_t = x_0 + \sum_{j=0}^{[t/d]} v_{t-dj}, \tag{6.27}$$

and

$$\Delta_d y_t = \Delta_d y_{t-1} + v_t = \Delta y_{-1} + \sum_{j=0}^{t} v_{t-d}. \tag{6.28}$$

Hence for $d > i \geq 0$,

$$\Delta_{d-i} y_t = \sum_{j=0}^{d-i} \Delta y_{t-j} = \sum_{j=0}^{d-i-1} \Delta y_{t-j} + x_{t-j}. \tag{6.29}$$

For the simplicity of notations, we assume $T = nd$. The initial conditions x_0 and $\Delta y_{-j}, j = 0, \cdots, d$ can be ignored asymptotically. Then

$$\begin{aligned}
&\frac{1}{T^2}\sum_{t=1}^{T} \Delta_k y_t \Delta_l y'_{t-s} \\
&= \frac{1}{d^2}\cdot\frac{1}{n^2}\sum_{i=1}^{d}\sum_{j=1}^{n} \Delta_k y_{(j-1)d+i}\Delta_l y'_{(j-1)d+i-s} \\
&\cong \frac{1}{d^2}\sum_{i=1}^{d}\left\{\frac{1}{n^2}\sum_{j=1}^{n}\Big(\sum_{m=0}^{k-1} x_{(j-1)d+i-m}\Big)\Big(\sum_{m'=0}^{l-1} x_{(j-1)d+i-s-m'}\Big)'\right\} \\
&= \frac{1}{d^2}\sum_{i=1}^{d}\sum_{m=0}^{k-1}\sum_{m'=0}^{l-1}\left\{\frac{1}{n^2}\sum_{j=1}^{n} x_{(j-1)d+i-m} x_{(j-1)d+i-s-m'}\right\} \\
&\xrightarrow{\mathcal{L}} \frac{1}{d^2}\sum_{i=1}^{d}\sum_{m=0}^{k-1}\sum_{m'=0}^{l-1}\left\{\Omega^{1/2}\int_0^1 B_{(i-m)}(t)B_{(i-s-m')}(t)'dt\Omega^{1/2}\right\},
\end{aligned} \tag{6.30}$$

where $B_i(t), i = 1, \ldots, d$ are mutually independent G-dimensional standard Brownian Motions. Similarly, we have

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T} \Delta_k y_{t-s} v'_t &= \frac{1}{d}\frac{1}{n}\sum_{i=1}^{d}\sum_{j=1}^{n}\left\{\sum_{m=0}^{k} x_{(j-1)d+i-s-m} v'_{(j-1)d+i}\right\} \\
&\xrightarrow{\mathcal{L}} \frac{1}{d}\sum_{i=1}^{d}\sum_{m=0}^{k}\left\{\Omega^{1/2}\int_0^1 B_{(i-s-m)}(t)dB'_i(t)\Omega^{1/2}\right\}.
\end{aligned} \tag{6.31}$$

The rest of the proof in the general case is similar to that of (i) and so it is omitted. □.

Proof of Theorem 3: Define z_{1t} as in the proof of Theorem 2 and $B_\Delta = (B_2, \ldots, B_p)$. Then under the null hypothesis, we write

$$\left(\begin{array}{c} \hat{\Gamma}' - \Gamma' \\ \hat{B}'_\Delta - B'_\Delta \\ \hat{B}_1 - 1 \end{array} \right)' = \sum_{t=1}^{T} v_t \left(\begin{array}{c} z_{1t} \\ z^*_{2t} \\ y_{t-1} \end{array} \right)' \left[\sum_{t=1}^{T} \left(\begin{array}{c} z_{1t} \\ z^*_{2t} \\ y_{t-1} \end{array} \right) \left(\begin{array}{c} z_{1t} \\ z^*_{2t} \\ y_{t-1} \end{array} \right)' \right]^{-1}. \tag{6.32}$$

Then we have

$$T\left(\hat{B}_1 - 1\right) = TV'\bar{P}_{Z_1,Z_2^*}Y_{-1}\left(Y'_{-1}\bar{P}_{Z_1,Z_2^*}Y_{-1}\right)^{-1} \tag{6.33}$$

We note that

(6.34)
$$\begin{aligned} \tfrac{1}{T^2}Y'_{-1}\bar{P}_{Z_1,Z_2^*}Y_{-1} &= \tfrac{1}{T^2}Y'_{-1}\bar{P}_{Z^*}Y_{-1} \\ &- \left(\tfrac{1}{T\sqrt{T}}Y'_{-1}\bar{P}_{Z^*}\Delta Y_{-1}\right)\left(\tfrac{1}{T}\Delta Y'_{-1}\bar{P}_{Z^*}\Delta Y_{-1}\right)^{-1}\left(\tfrac{1}{T\sqrt{T}}\Delta Y'_{-1}\bar{P}_{Z^*}Y_{-1}\right). \end{aligned}$$

The second term of (6.20) converges to zero. Hence the limiting distribution of $T(\hat{B}_1 - 1)$ is asymptotically equivalent to that of $TV'\bar{P}_{Z^*}Y_{-1}$. Then by using Lemma 1, we have

$$T\left(\hat{B}_1 - 1\right) \xrightarrow{\mathcal{L}} \frac{\left[\int_0^1 \left(\begin{array}{c} m(t) \\ B(t) \end{array} \right) \left(\begin{array}{c} m(t)dt \\ dB(t) \end{array} \right)'\right]_{22.1}}{\left[\int_0^1 \left(\begin{array}{c} m(t) \\ B(t) \end{array} \right) \left(\begin{array}{c} m(t) \\ B(t) \end{array} \right)' dt\right]_{22.1}}. \tag{6.35}$$

By multiplying

$$\left|\int_0^1 m(t)m(t)'dt\right|$$

to the numerator and the denominator of (6.35), we have (4.23).

□

Proof of Theorem 4: We give the proof only for the case when $p = 1$ and we take $z_{1t} = z_{1t}^*$ and $z_{2t} = y_{t-1}$. The general case when $p > 1$ can be dealt by a similar method as the proofs of Theorems 2 and 3. We use the results and notations in Anderson and Kunitomo (1990) extensively. From the same arguments as the proof of their Lemma 3, the distribution of T_3 is asymptotically equivalent to

$$T_3' = \mathrm{tr}\left\{U'\left(\bar{P}_{(Z_1^*,Y_{-1})D} - \bar{P}_{Z_1^*,Y_{-1}}\right)U\Sigma^{-1}\right\}, \tag{6.36}$$

where

$$U = VC^*, \Sigma = C^{*\prime}\Omega C^*, \tag{6.37}$$

and C^* is a $G \times G_0$ matrix such that

$$B_1^{*\prime}C^* = O, \quad D = \left\{\Pi_*, \begin{bmatrix} I_{K_1^*} \\ O \end{bmatrix}\right\}, \quad \Pi_* = \begin{bmatrix} \Gamma_1 & I_{K_1^*} \\ B_1^* & O \end{bmatrix}\begin{bmatrix} O \\ I_{K-G_0} \end{bmatrix}. \tag{6.38}$$

Define a $G \times r$ matrix β $(rank(\beta) = r)$ and a $r \times r$ matrix α such that $\bar{P}_{Z_1^*}Z\Pi_* = \bar{P}_{Z_1^*}Y_{-1}\beta\alpha$. Then

$$\begin{aligned} T_3' &= \mathrm{tr}U'\left(\left[\bar{P}_{Z_1^*} - \bar{P}_{Z_1^*}Z\Pi_*\left(\Pi_*'Z'\bar{P}_{Z_1^*}Z\Pi_*\right)^{-}\Pi_*'Z'\bar{P}_{Z_1^*}\right]\right. \\ &\left. - \left[\bar{P}_{Z_1^*} - \bar{P}_{Z_1^*}Y_{-1}\left(Y_{-1}\bar{P}_{Z_1^*}Y_{-1}\right)^{-1}Y_{-1}'\bar{P}_{Z_1^*}\right]\right)U\Sigma^{-1}. \end{aligned} \tag{6.39}$$

In this case we should use a $G \times G$ normalization matrix,

$$D_T^{*\prime} = \left[\frac{1}{\sqrt{T}}\beta, D_{2T}^{-1}\right]', \tag{6.40}$$

where we can take each vector of β and D_{2T}^{-1} being orthogonal. Then we have

$$T_3' = \mathrm{tr}\left\{U'\bar{P}_{Z_1^*}Y_{-1}D_{2T}^{-1}\left[D_{2T}^{-1\prime}Y_{-1}'\bar{P}_{Z_1^*}Y_{-1}D_{2T}^{-1\prime}\right]^{-1}D_{2T}^{-1\prime}Y_{-1}'\bar{P}_{Z_1^*}U\Sigma^{-1}\right\}. \tag{6.41}$$

By the Granger Representation Theorem, under the null hypothesis we have

$$(6.42) \qquad y_{t-1} \cong y_0 + C\left(S_1(t-1) + \Gamma_1 \sum_{s=1}^{t-1} z_{1s}^*\right),$$

where a $G \times G$ matrix C is of rank r and $z_{2t}^* = \sum_{s=1}^{t} z_{1s}^*$. Since $K_1^* \le G-r$, we can take a $G \times (G-K_1^*-r)$ matrix $\gamma_1(T)$ and a $G \times K_1^*$ matrix $\gamma_2(T)$ consisting of linearly independent vectors such that $\Gamma_1' C \gamma_1(T) = O$ and $\Gamma_1' C \gamma_2(T) \neq O$. We partition

$$D_{2T}^{-1} = (\gamma_1(T), \gamma_2(T))'.$$

Then we have

$$(6.43) \qquad T_2'' = tr\left\{U'\left[\bar{P}_{Z_1^*} - \bar{P}_{Z_1^*, Z_2^*\gamma_1(T), S_{-1}\gamma_2(T)}\right] U\Sigma^{-1}\right\},$$

where S_{-1} is a $T \times G$ matrix whose t-th row is $S_1(t-1)$. By using Theorem 1, T_2'' converges to (4.38) as $T \to \infty$. □

Proof of Theorem 5: We give the proof only for the case when p=1. The general case when $p > 1$ can be can be dealt with by a similar method as the proofs of Theorems 2 and 3. Let ΔY be a $T \times G$ matrix whose t-th row is $\Delta y_t = y_t - y_{t-1}$. Then the test statistic T_3 is asymptotically equivalent to

$$(6.44) \qquad T_3' = T \cdot \log \frac{|\Delta Y' \bar{P}_{Z_1^*, Y_{-1}} \Delta Y|}{|\Delta Y' \bar{P}_{Z^*, Y_{-1}} \Delta Y|} + T \sum_{i=1}^{G_0} \lambda_i^*.$$

The first term is the LR statistic for testing the null hypothesis $H_1' : \Gamma_2 = O$ against the alternative hypothesis $H_A : \Gamma_2 \neq O$. The second term is to the LR statistic for testing $H_1'' :$ rank$(B_1^*) = r, \Gamma_2 = O$ against the alternative hypothesis $H_{A2} : rank\,(B_1^*) = G, \Gamma_2 \neq O$. From the proof of Theorem 4, the second term of T_3' is asymptotically equivalent to

$$(6.45) \qquad T_3^{(1)} = \mathrm{tr}\left\{U'\left[\bar{P}_{(Z_1^*, Y_{-1})D} - \bar{P}_{Z_1^*, Y_{-1}}\right] U\Sigma^{-1}\right\}.$$

We take a $G \times G$ matrix $E = (C^*, J)$ such that

$$(6.46) \qquad \Omega_* = E'\Omega E = \begin{pmatrix} \Sigma & O \\ O & \Sigma_* \end{pmatrix}.$$

Then under the null hypothesis we rewrite the first term as

(6.47)

$$\begin{aligned}
T_3' &= T \cdot \log \frac{\left| E'V'\frac{1}{T}\bar{P}_{Z_1^*,Y_{-1}} VE \right|}{\left| E'V'\frac{1}{T}\bar{P}_{Z^*,Y_{-1}} VE \right|} \\
&= T \cdot \log \frac{\left| E'V'\frac{1}{T}(\bar{P}_{Z_1^*,Y_{-1}} - \bar{P}_{Z^*,Y_{-1}} + \bar{P}_{Z^*,Y_{-1}})VE \right|}{\left| E'V'\frac{1}{T}\bar{P}_{Z^*,Y_{-1}} VE \right|} \\
&\cong \operatorname{tr}\left\{ E'V'\left(\bar{P}_{Z_1^*,Y_{-1}} - \bar{P}_{Z^*,Y_{-1}}\right) VE\Omega_*^{-1} \right\} \\
&= \operatorname{tr}\left\{ U'\left(\bar{P}_{Z_1^*,Y_{-1}} - \bar{P}_{Z^*,Y_{-1}}\right) U\Sigma^{-1} \right\} + \operatorname{tr}\left\{ V_{*'}\left(\bar{P}_{Z_1^*,Y_{-1}} - \bar{P}_{Z^*,Y_{-1}}\right) V_*\Sigma_*^{-1} \right\} \\
&= T_3^{(2)} + T_3^{(3)},
\end{aligned}$$

where $V_* = VJ$. Then from (6.45) and (6.47), we have

$$(6.48) \qquad T_3^{(1)} + T_3^{(2)} = \operatorname{tr}\left\{ U'\left[\bar{P}_{(Z_1^*,Y_{-1})D} - \bar{P}_{Z^*,Y_{-1}} \right] U\Sigma^{-1} \right\}.$$

Let $D_T^{-1'} = \left(\beta/\sqrt{T}, D_{2T}^{-1}\right)'$ be a $G \times [r + (G-r)]$ nonsingular matrix. We note $\bar{P}_{Z^*}Y_{-1} \cong \bar{P}_{Z^*}S_{-1}C$ under the null hypothesis of cointegration and $\left(Y_{-1}\beta/\sqrt{T}\right)' S_{-1}D_{2T}^{-1} \xrightarrow{p} 0$, then $T_3^{(1)} + T_3^{(2)}$ is asymptotically equivalent to

$$(6.49) \qquad T_3^{*'} = \operatorname{tr}\left\{ U'\left[\bar{P}_{Z_1^*} - \bar{P}_{Z^*,S_{-1}D_{2T}^{-1}} \right] U\Sigma^{-1} \right\},$$

which converges to the first term of (4.45). Next, we obtain the limiting distribution of $T_3^{(3)}$. Under the assumption $\Delta z_{2t}^* = z_{1t}^*$, we take $[r + (G - r -$

$K_1^*) + K_1^*] \times G$ matrix $D_T^{-1'} = \left(\beta/\sqrt{T}, \gamma/T, D_{2T}^{-1}\right)'$ as a normalization factor and

$$\bar{P}_{Z_1^*,Y_{-1}} = \bar{P}_{Z_1^*} - \bar{P}_{Z_1^*}\left(Y_{-1}'\bar{P}_{Z_1^*}Y_{-1}\right)^{-1}Y_{-1}\bar{P}_{Z_1^*} \cong \bar{P}_{Z_1^*,Y_{-1}\beta,S_{-1}\gamma,Z_2^*} \tag{6.50}$$

because of the cointegration relations under the null hypothesis. Similarly, we take a $[r+(G-r)]\times G$ matrix $D_T^{*-1'} = \left(\beta/\sqrt{T}, \gamma^*/T\right)$ for a normalization factor. Then

$$\bar{P}_{Z^*,Y_{-1}} \cong \bar{P}_{Z^*,Y_{-1}\beta,S_{-1}\gamma^*} \tag{6.51}$$

under the null hypothesis. Hence

$$\bar{P}_{Z_1^*,Y_{-1}} - \bar{P}_{Z^*,Y_{-1}} \cong P_{Z^*,Y_{-1}\beta,S_{-1}\gamma^*} - P_{Z_1^*,Y_{-1}\beta,S_{-1}\gamma}\ . \tag{6.52}$$

Let the t-th row of V_* be v_{*t}, which is uncorrelated with the t-th row of U conditional on $\mathcal{F}_{t-1}$. Then by applying Theorem 1 to $(u_t', v_{*t}')'$, we have

$$\frac{1}{\sqrt{T}}\sum_{s=1}^{[Tt]}\begin{pmatrix} u_t \\ v_{*t}\end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} \Sigma^{1/2}B(t) \\ \Sigma_*^{1/2}W(t)\end{pmatrix}, \tag{6.53}$$

where $(B(t)', W(t)')'$ is the $G(=[(G-r)+r])$ dimensional standard Brownian motion. Again the invariance principle, the limiting distribution of $T_3^{(3)}$ has the second form in (4.45), provided that $G - K_1^* - r \geq 0$. □

References

[1] Anderson, T.W. (1951). Estimating linear restrictions on regression coefficients for multivariate normal distributions, *Annals of Mathematical Statistics*, 22, 327-351.

[2] Anderson, T.W. (1959). On asymptotic distributions of estimates of parameters of stochastic difference equations, *Annals of Mathematical Statistics*, 30, 637-687.

[3] Anderson, T.W. (1984). *An Introduction to Multivariate Statistical Analysis*, John-Wiley & Sons.

[4] Anderson, T.W. and Kunitomo, N. (1989). Tests of overidentification and predetermindness in simultaneous equation models, To appear in *Journal of Econometrics.*

[5] Anderson, T.W. and Kunitomo, N. (1990). Asymptotic robustness of tests of overidentification and exogeneity, Discussion Paper No. 90-F-19, Faculty of Economics, University of Tokyo. To appear in *Journal of Econometrics.*

[6] Anderson, T.W. and Kunitomo, N. (1992). Asymptotic distributions of estimators in regression and autoregression based on Lindeberge conditions, *Journal of Multivariate analysis*, 40, 221-243.

[7] Billingsley, P. (1968). *Convergence of Probability Measure*, John-Wiley & Sons.

[8] Chan, N.H. and Wei, C.Z. (1987). Asymptotic inference for nearly non-stationary AR(1) processes, *The Annals of Statistics*, 15, 1050-63.

[9] Chan, N.H. and Wei, C.Z. (1988). Limiting distributions of least squares estimates of unstable autoregressive process, *The Annals of Statistics*, 16 367-401.

[10] Dickey, D.A. and W.A.Fuller (1979). Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association*, 74, 427-431.

[11] Dickey, D.A. and W.A.Fuller (1981). Likelihood ratio statistics for autoregressive time series with a unit root, *Econometrica*, 49, 4, 1057-1072.

[12] Dvoretzky, Aryeh (1972). Asymptotic normality for sum of dependent random variables, *Proceedings of the sixth Berkeley symposium on mathematical statistics and probability, Volume 2*, University of California, Berkeley, 513-135.

[13] Engle, R. and Granger, C.W.J. (1987). Cointegration and error correction: representation, estimation, and testing, *Econometrica*, 55, 2, 251-276.

[14] Fountis, N.G. and Dickey, D.A. (1989). Testing for a unit root nonstationarity in multivariate autoregressive time series," *The Annals of Statistics*, 17, 1, 419-428.

[15] Fuller, W.A. (1976). *Introduction to Statistical Time Series*, John-Wiley.

[16] Hasza,D.P. and W.A. Fuller (1979). Estimation for autoregressive processes with unit roots, *The Annals of Statistics*, 7, 5, 1106-1120.

[17] Hasza,D.P. and W.A. Fuller (1982). Testing for nonstationary parameter specifications in seasonal time series models, *The Annals of Statistics*, 10, 4, 1209-1216.

[18] Helland, I.S. (1982). Central limit theorems for martingales with discrete or continuous time, *Scandinavian Journal of Statistics*, 9, 79-94.

[19] Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models, *Econometrica*, 59, 6, 1551-80.

[20] Kunitomo, N. (1988). Approximate distributions and power of the test statistics for overidentifying restrictions in a system of simultaneous equations, *Econometric Theory*, 4, 248-274.

[21] Perron, P. (1989). The Great crash, the oil price shock, and the unit root hypothesis, *Econometrica*, 57, 6, 1361-1401.

[22] Phillips, P.C.B. (1987). Time series regression with a unit root, *Econometrica*, 55,2, 277-301.

[23] Phillips, P.C.B. (1987),"Towards a unified asymptotic theory for autoregression, *Biometrika*, 74, 3, 535-47.

[24] Phillips, P.C.B. and Ouliaris, S. (1990). Asymptotic properties of residual based tests for cointegration, *Econometrica*, 58, 1, 165-193.

[25] Phillips, P.C.B. and S. N. Durlauf (1986). Multiple time series regression with integrated processes, *Review of Economic Studies*, 53, 473-495.

[26] Phillips, P.C.B. and P. Perron (1988). Testing for a unit root in time series regression, *Biometrika*, 72,2,335-46.

[27] Tanaka, K. (1991). An alternative approach to the asymptotic theory of spurious regression, cointegration and near cointegration, unpublished manuscript.

[28] White, J.S.(1958). The limiting distribution of the serial correlation coefficient in the explosive case, *Annals of Mathematical Statistics*, 29,1188-1197.