

# STATISTICAL ANALYSIS AND TESTING OF COINTEGRATED SYSTEMS

**O. E. OLOWOFESO, K. A. KHOGALI and O. A. FASORANBAKU**

(Received 7 November 2002; Revision accepted 5 February 2003)

## ABSTRACT

This paper deals with estimation and testing for cointegration when deterministic trends are present in the data generating process. The study confirmed that to estimate the Vector Error Correction Model (VECM) when there is no cointegration will produce an egregious pitfall. Derivation of the linkage between the residual matrix of VECM and the corresponding eigenvalues of the product moment matrices is provided. The bivariate system designed shows a reversal relationship between the lag-lengths and the values of the likelihood ratio (LR) statistic. Moreover, the values of the (LR) test for different lags at various sample sizes are reported in the simulation. The Monte Carlo experiment shows that the null hypothesis of no cointegration is rejected in favours of cointegration inspite of the deterministic trend in the data. The standard Z-test and t-test prove to be more robust via size properties for a wider range of nuisance parameter than the coefficient based tests.

**KEYWORDS:** Cointegration, Deterministic trend, Vector Autoregressive Estimates, Hypothesis Testing and Data generation process(DGP).

Correspondence address: [olowofeso@yahoo.com](mailto:olowofeso@yahoo.com)

## INTRODUCTION

The concept of cointegration have been studied extensively by some researchers in the last ten years. A critical examination of cointegration from theory point of view says that variables  $x_t$  and  $y_t$  for instance, are both integrated of order 1, denoted by  $I(1)$ , if their changes are stationary, then they are said to be cointegrated if there exists a linear combination, say  $z_t = y_t - \alpha x_t$ , which is integrated of order 0, denoted by  $I(0)$ . Many scholars have defined cointegration as stated above; but the changes in some of the important feature of an economy can interrupt equilibrium type relationship possible for an extended period of time. A lot of statistical explanations for not rejecting the null of no cointegration in empirical works have been examined by Engle and Granger (1987), Johansen (1996), Perron (1989), Ahn et al (1990), Hendry (1995) and Stock and Watson (1989) among others. In particular, since the introduction of cointegration tests by Engle and Granger (1987), the usage of these tests on long-run relationships between non-stationary time series variables have grown in popularity amongst applied statisticians, economists and econometricians. Engle and Smith (1998) examined cointegration very carefully and concluded that a major fallacy of cointegration tests, however, is the need for a considerable span of the data.

In empirical research works, many researchers have used deterministic trending functions like polynomials to represent secular characteristics like growth over time to model non-stationarity. In that case, the time series  $y_t$  is broken into two components, one to capture trend and another to capture stationarity fluctuations. Generally, model of this form can be written as :

$$y_t - h_t - y_t^* = 0, h_t = \gamma^t x_t \quad (1)$$

$(t = 1, 2, \dots, n)$

where  $y_t^*$  is stationary time series,  $x_t$  is an m-vector of deterministic trend  $\gamma$  is a vector of m parameters and  $h_t$  is the deterministic trend. A more general example where the trends are

piecewise higher order polynomial is

$$h_t = \sum_{j=0}^p q_j t^j + \sum_{j=0}^p q_{mj} t_m^j \quad (2)$$

where

$$t_m^j = \begin{cases} (t-m)^j & t \in (m+1, \dots, n) \\ 0 & t \in (1, \dots, m) \end{cases}$$

It must be noted that unsatisfactory feature of trend stationary models is that no random elements appear in the trending mechanism and only the stationary component is subject to stochastic shocks. By adding a stationary component  $v_t$  to the (R.H.S) of the first equation of (1) and allow  $y_t^*$  to be generated by first order process, we have

$$y_t - h_t - y_t^* = v_t, y_t^* = y_{t-1}^* + u_t \quad (3)$$

where  $u_t \sim N(0, \sigma_u^2)$ . Now equation (3) decomposes the time series  $y_t$  into a deterministic trends, a stochastic trend and stationary residual. When  $\sigma_u^2 = 0$ , the stochastic trend in (3) corresponds to a null hypothesis of the trend stationary. By Gaussian assumption, and if the error is identically and independently normally distributed, the hypothesis can be tested in a simple way using the likelihood method principle. This procedure can easily be extended to more general cases where there is serial dependence, by using parametric or semi parametric methods. Let

$$M_t = y_t^* + u_t \quad (4)$$

and writing its difference as

$$\Delta M_t = (1 - \theta L)\eta_t$$

where  $\eta_t$  is stationary. It is clear that  $\sigma_u^2 = 0$ , then Equation (5) corresponds to the null hypothesis of a moving average unit root  $\theta = 1$ . Thus, there is a correspondence between testing for stationarity and testing for moving average unit root (Saikkonen et al (1993)).

In general form, this work attempts to extend the researches done on spurious regression where the misspecification is also in the long memory component and the pitfalls in testing for long run relationships of the Granger and Newbold (1974). In general, this paper is based on the work of Gonzalo and Lee (1998) who recommend the non-singularity of the error covariance matrix and that the cointegrated variables must have a trending and long memory not different from the unit root tests as two prerequisite conditions for the likelihood ratio test not to suffer from pitfalls. This work also extend the work of Olowofeso *et al* (2001) who examined the estimation of cointegration system using a Monte Carlo experiment. The issue of model misspecification and its consequences are investigated by using the simple equation-based tests and the system of equation-based tests of Johansen. More specifically, this paper attempts to link the eigenvalues of the product covariance matrices to the residual matrix of the VECM. The simulation adopted attempt to examine the behaviour of the likelihood ratio with different lag-lengths as well as the testing and estimating of trend stationary models. We estimate the statistics at various sample sizes and fit a response surface and observed the asymptotic distribution. Precisely, the design of  $Q_1$  and  $Q_2$  of Johansen trace statistic and Maximum eigenvalue statistic respectively were carefully conducted. Section 2 focuses on mathematical and computational framework as well as hypothesis testing. Section 3 presents the simulated results obtained. Section 4 shows the concluding remarks.

## Mathematical and Computational Framework

### Trend Stationary models with deterministic component

The objective of this section is to present the framework which systematically analyses and test cointegration systems. We examined how non-cointegrated systems that contains deterministic trend components can be improved with the method of analysing cointegration relationship.

The DGP used here is as defined below when  $I(1)$  variables with deterministic components

$$\Delta y_t - \alpha_1 - \beta_1 t = e_{1t} \tag{6}$$

$$\Delta x_t - \alpha_2 - \beta_2 t = e_{2t} \tag{7}$$

which is an example of a trend stationary model of the general form defined in (1) with  $\gamma^1 = (\alpha_1, \beta_1)$ ,  $x_t = (1, t)'$  and  $i = 1, 2$ .

Gonzalo and Lee (1998) provide the following proposition for  $I(1)$  process with deterministic component of the type defined in (6) and (7):

Proposition 1: If  $(y_t, x_t)'$  are two different  $I(1)$  processes with deterministic components generated from equation (6) and (7).

- (i) If  $\beta_1 \neq 0$  or  $\beta_2 \neq 0$  and if the vector error correction model (VECM) does not include any deterministic components, then the first eigenvalue,  $(\lambda_1)$ , from the product matrices  $\sum_{i=0}^{t-1} \sum_{j=0}^{t-1} \sum_{k=0}^{t-1} \sum_{l=0}^{t-1}$  does not converge to zero in probability.
- (ii) If  $\beta_1 = \beta_2 = 0$ ,  $\alpha_1 \neq 0$  or  $\alpha_2 \neq 0$ , and if the VECM does not include any deterministic components, then  $\lambda$  does not converge to zero in probability.

Proof: See Olowofeso (2000)

In order to capture vector error correction model with  $X_t$  augmented with 1 and/ or  $t$  as considered in Johansen (1995), we use

$X_{t-1}^* = (x_{t-1}^1 1)'$  or  $X_{t-1}^* = (x_{t-1}^1 t)'$  to get the following VEC model:

$$\Delta X_t^* = \Pi X_{t-1}^* + \epsilon_t$$

This model was formulated to capture a cointegration relationship around a common deterministic trend (that is, stochastic cointegration as well as deterministic cointegration)

The two cases examined are:

Case1: When  $\alpha_1 = \alpha_2 = 0$  and  $\beta_1 = \beta_2 = 0.01$  and

Case2: When  $\alpha_1 = \alpha_2 = 0$  and  $\beta_1 = \beta_2 = 0$

For null hypothesis of no cointegration, that is

$H_0: r = 0$  using  $Q_1$  and  $Q_2$  defined as:

$$Q_1 = -T \ln[(1 - \lambda_1)(1 - \lambda_2)]$$

$$Q_2 = -T \ln(1 - \lambda_1)$$

respectively;

where  $\lambda_1$  and  $\lambda_2$  are eigenvalues of  $\sum_{i=1}^T \sum_{j=1}^T \sum_{k=1}^T \sum_{l=1}^T$  and

$\sum_{i=1}^T \sum_{j=1}^T \sum_{k=1}^T \sum_{l=1}^T R_{it} R_{jt}'$  ( $i, j = 0, 1$ ) are the product moment matrices of the residuals,  $R_{0t}$  and  $R_{1t}$ , from the regression of  $\Delta y_t$  and  $y_{t-1}$  on the lagged differences respectively. The  $T$  is the sample size,  $Q_1$  is the Johansen trace statistic and  $Q_2$  is the maximum eigenvalue statistic. Then the trace test statistics for the hypothesis that there are at most  $r$  cointegrating vectors, and hence  $s = (2 - r)$  unit roots or equivalently,  $s$  zero characteristic roots, is

$\xi = -n^2 \sum_{i=r+1}^2 \ln(1 - \lambda_i)$ . The asymptotic distribution of  $Q_1$  depends only on  $s$ , and 5% critical value for  $s = 1$  is 3.84 (see Johansen's Likelihood Ratio Test Statistical table), and that of  $s = 2$  is 12.53. Since the eigenvalues of  $\sum_{i=1}^T \sum_{j=1}^T \sum_{k=1}^T \sum_{l=1}^T$  are the same as the one of  $\sum_{i=1}^T \sum_{j=1}^T \sum_{k=1}^T \sum_{l=1}^T$  and since the vector error correction representation of  $X_t = (x_t, y_t)'$  is

$$\Delta X_t = \Pi X_{t-1} + \sum_{j=0}^{k-1} \Gamma_j \Delta X_{t-j} + \epsilon_t \tag{9}$$

with  $\epsilon_t$  assumed to be a Gaussian i.i.d  $(0, \Omega)$  process and  $\Omega > 0$ , then

$\hat{\Omega} = \sum_{i=1}^T \sum_{j=1}^T \sum_{k=1}^T \sum_{l=1}^T$  it follows that  $|\sum_{i=1}^T \sum_{j=1}^T \sum_{k=1}^T \sum_{l=1}^T \hat{\Omega}| = |I_p - \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} \sum_{l=1}^T \sum_{m=1}^T|$ , which produces the

relationship  $\log|\Omega(r)| = \log\left|\sum_{00} + \sum_{i=1}^r \log(1 - \tilde{\lambda}_i)\right|$ . Little mathematical manipulation produces the maximised likelihood function

$$L_{\max}^{\lambda, r} = |\hat{\Omega}[\Gamma_{(r,r)}^r(1 - \hat{\lambda}_i)]^{-1}$$

It must be noted that if the hypothesis involved just  $I(0)$  variables, we would expect twice the log likelihood ratio, that is,

$$2(L_{01}^* - L_0^*) = -T \sum_{i=h+1}^n \log(1 - \hat{\lambda}_i)$$

(See Hamilton, (1994)), where  $L_{01}^*$  is the log-likelihood function under the absence of constraints,  $L_0^*$  is the log-likelihood function under the constraint and  $\lambda$  is as earlier defined. Another approach is to test the null of hypothesis of  $h$  cointegrating relations against the alternative of  $h+1$  cointegrating relations. Twice the log likelihood ratio for this case is given by

$$= T \log(1 - \hat{\lambda}_{h+1}^2)$$

We also considered the four cases for the first -order autoregressive when the true process is random walk.

- Case1: No constant term or time trend in the regression; true process is a random walk.
- Case2: Constant term but no time trend included in the regression; true process is a random walk.
- Case3: Constant term but no true trend included in the regression; true process is a random walk with drift.
- Case4: Constant term and time trend included in the regression; true process is a random walk with or without drift.

If the true model for case 4 is

$$y_t = \alpha + y_{t-1} + u_t \tag{12}$$

where  $u_t \sim N(0, \sigma^2)$ , then the true value of  $\alpha$  turns out not to matter for the asymptotic distribution. In contrast to the previous cases, we now assumed that a time trend is included in the regression that is actually estimated by OLS:

$$y_t = \alpha + \rho y_{t-1} + \delta t + u_t \tag{13}$$

If  $\alpha \neq 0$ ,  $y_{t-1}$  would be asymptotically equivalent to time trend.

The hypothesis are:  $H_0: y_t = \alpha + y_{t-1} + u_t$ ,  $\alpha > 0$  that is whether this trend arise from the positive drift term of a random walk against  $H_A: y_t = \alpha + \delta t + \rho y_{t-1} + v_t$ ,  $|\rho| < 1$

Gonzalo *et al* (1998) attributed the pitfalls of Johansen's likelihood ratio test to the behaviour of  $\Omega$  and  $\sum_{01}^{-1}$  when they postulated the following proposition:

**Proposition 2:** (a) The eigenvalues of  $\sum_{11}^{-1} \sum_{00} \sum_{00}^{-1} \sum_{01}$  are the same as eigenvalues of  $\hat{L} = I - \sum_{01}^{-1} \hat{\Omega}$ ; where  $\Omega$  is the covariance matrix of the residuals of the VEC model:

$$\Delta X_t = \Pi X_{t-1} + \epsilon_t$$

calculated under no rank constraint and  $\sum_{01}$  is the residual covariance matrix.

**Proof:** See Gonzalo and Lee (1998)

Indeed, Gonzalo *et al* (1998) stressed the need for a deep pre-cointegration analysis for adequate application of likelihood ratio test to avoid pitfalls. Accordingly, we examined the link between the eigenvalues and the coefficient of determination of the residuals of the VECM by postulating the following proposition:

**Proposition 3:** If  $R_i^2$  represent the coefficient of determination of the  $i^{th}$  equation of the VECM, under the null of  $p$  independent random walks, then

$$\sum_{i=1}^p \hat{\lambda}_i = \sum_{i=1}^K R_i^2 + O_p(1) \tag{15}$$

**Proof:**

Since equation (6) and (7) can be written in the following VECM form:

$$\begin{aligned} \Delta X_t &= \begin{pmatrix} 0 & 0 & \alpha & \beta \\ 0 & 0 & \alpha & \beta \end{pmatrix} [X_{t-1} \ 1 \ t]' + \epsilon_t \\ &= \Pi^* X_{t-1}^* + \epsilon_t \end{aligned}$$

By examining the result from bivariate point of view; we have  $\Delta X_t = X_{t-1} + \epsilon_t$  where  $X_t = (x_{1t}, x_{2t})'$  and the errors are assumed to have zero mean and the following covariance matrix

$$\Omega = \begin{pmatrix} \sigma_1^2 & 0\sigma_1\sigma_2 \\ 0\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

the eigenvalues of  $\sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^m$ . It can be prove that these eigenvalues are the same with  $\sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^m$  due to the symmetry of the matrices. In addition,  $\hat{\Omega}_t = \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^m$ , where  $\hat{\Omega}_t$  is the estimated error covariance matrix. This implies that eigenvalues above are the same with eigenvalues  $I - \sum_{i=1}^n \hat{\Omega}_t$ . That is,  $\sum_{i=1}^p \hat{\lambda}_i = \text{Trace}(I - \sum_{i=1}^n \hat{\Omega}_t)$

$$\begin{aligned} &= \text{Trace} \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} \sum \epsilon_{1t}^2 & \sum \epsilon_{1t} \epsilon_{2t} \\ \sum \epsilon_{2t} \epsilon_{1t} & \sum \epsilon_{2t}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum \epsilon_{1t}^2 & \sum \epsilon_{1t} \epsilon_{2t} \\ \sum \epsilon_{2t} \epsilon_{1t} & \sum \epsilon_{2t}^2 \end{pmatrix} \right\} \\ &= R_1^2 + R_2^2 - \frac{2 \sum \epsilon_{1t} \epsilon_{2t} [\sum \hat{\epsilon}_{1t} \epsilon_{2t} - \sum \epsilon_{1t} \epsilon_{2t}]}{\sum \epsilon_{1t}^2 \sum \epsilon_{2t}^2} \\ &\qquad\qquad\qquad 1 - \frac{(\sum \epsilon_{1t} \epsilon_{2t})^2}{\sum \epsilon_{1t}^2 \sum \epsilon_{2t}^2} \end{aligned}$$

where  $R_i^2$ , ( $i = 1, 2$ ) are the coefficients of determination from the first and the second equations of the ECM. The first term in the numerator above converges to zero in probability, by the law of large number. Suppose  $0 = 0$  in  $\Omega$ , the numerator converges to 1, giving us the required result for  $p=2$ , by induction it is also true for  $p > 2$  eigenvalues.

**2.2 Statistical Models for case (SMC) 1 and 2 of Integration of order one I(1) with deterministic component**

DGP:

$$\Delta Y_t = \alpha_1 + \beta_1 t + \epsilon_{1t} \text{ and } \Delta X_t = \alpha_2 + \beta_2 t + \epsilon_{2t}$$

Thus

$$\text{SMC}_{1t}: \Delta X_t = \Pi_1 X_{t-1} + \epsilon_{1t} \tag{18}$$

in matrix form we have

$$\begin{pmatrix} \Delta y_t \\ \Delta X_t \end{pmatrix} = \Delta X_t = \begin{pmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{pmatrix} \begin{pmatrix} \Delta y_{t-1} \\ \Delta X_{t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{11t} \\ \epsilon_{12t} \end{pmatrix}$$

Similarly

$$SMC_{2i}: \Delta X_t = \mu + \Pi_3 X_{t-1} + \epsilon_{2t} \tag{19}$$

$$\begin{pmatrix} \Delta y_t \\ \Delta X_t \end{pmatrix} = \begin{pmatrix} \alpha_0 \\ \alpha_0 \end{pmatrix} + \begin{pmatrix} \alpha_1 & \beta_1 \\ \alpha_1 & \beta_2 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{11t} \\ \epsilon_{12t} \end{pmatrix}$$

$$SMC_{3i}: \Delta X_t = \mu + \Pi_3 X_{t-1} + \epsilon_{2t} \tag{20}$$

$$\begin{pmatrix} \Delta y_t \\ \Delta X_t \end{pmatrix} = \begin{pmatrix} \alpha_0 \\ \beta_0 \end{pmatrix} + \begin{pmatrix} \alpha_{1t} \\ \beta_{1t} \end{pmatrix} t + \begin{pmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ X_{t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{31t} \\ \epsilon_{32t} \end{pmatrix}$$

$$SMC_{4i}: \Delta X_t = \Pi_4 [X_{t-1}^1 \ 1]^1 + \epsilon_{3t} \tag{21}$$

$$\begin{pmatrix} \Delta y_t \\ \Delta X_t \end{pmatrix} = \Delta X_t = \begin{pmatrix} 0 & \alpha_0 & \beta_1 \\ 0 & \alpha_2 & \beta_2 \end{pmatrix} [X_{t-1}^1 \ 1]^1 + \begin{pmatrix} \epsilon_{11t} \\ \epsilon_{12t} \end{pmatrix}$$

$$SMC_{5i}: \Delta X_t = \Pi_5 [X_{t-1}^1 \ 1 \ t]^1 + \epsilon_{5t} \tag{22}$$

$$\begin{pmatrix} \Delta y_t \\ \Delta X_t \end{pmatrix} = \Delta X_t = \begin{pmatrix} 0 & 0 & \alpha_0 & \beta_1 \\ 0 & 0 & \alpha_2 & \beta_2 \end{pmatrix} [X_{t-1}^1 \ t]^1 + \begin{pmatrix} \epsilon_{51t} \\ \epsilon_{52t} \end{pmatrix}$$

with T=100, SMC<sub>j</sub>, (j=1,2,..., 5) and i=1,2 ; equations (18), (19), ..., (22) are statistical models used to estimate the Johansen statistics of Q<sub>1</sub> and Q<sub>2</sub>.

The Engle and Granger (EG) are calculated from the cointegrating ordinary least squares regression without a constant (SMC<sub>1i</sub>), with a constant (SMC<sub>2i</sub>) and with both a constant and trend (SMC<sub>3i</sub>). A total of 1,000 replications are used and reported at 5% level. The critical value for each of the three regression models are simulated from 100,000 replications using the DGP with α<sub>1</sub>= α<sub>2</sub> = 0 and β<sub>1</sub>= β<sub>2</sub> = 0. The simulation results are presented in section 3.

### 2.3 Hypothesis Testing.

In addition, various test of hypotheses for the parameters of the regression model  $y_t = \alpha + \rho y_{t-1} + u_t$  under the assumption that the true  $\alpha = 0$ ,  $\rho = 1$  and  $u_t$  is i.i.d with mean zero and variance  $\sigma^2$ , was conducted. The first one is the Phillips-Perron  $\rho$  statistic defined by

$$T(\hat{\rho} - 1) - (1/2) \left( T^2 \sigma_\rho^2 / s_\rho^2 \right) (\hat{\lambda}^2 - \gamma_0^2) \tag{23}$$

where

T is the sample size,  $\hat{\rho}$  is the estimated coefficient of  $y_{t-1}$  in the autoregressive model,  $\hat{\sigma}_\rho^2$  is the ordinary least squares variance of  $\hat{\rho}$ ,  $s_\rho^2 = (T - k)^{-1} \sum_{i=1}^n u_i^2$  is the usual formula for the residuals from the model, k is number of parameters in the estimated regression, in this case k=2;  $\gamma_0$  is the autocovariance,  $\hat{\lambda}^2$  is the Newey-West estimator defined by

$$\hat{\lambda}^2 = \gamma_0 + 2 \sum_{j=1}^q \left[ 1 - j/(q+1) \right] \hat{\gamma}_j \tag{24}$$

$q$  is the number of autocovariances considered and  $\gamma_j = T^{-1} \sum_{i=j+1}^T \hat{u}_i \hat{u}_{i-j}'$  is the  $j^{\text{th}}$  autocovariances of  $u_t$  (See Hamilton, (1989) and Hamilton (1994) for the derivations of Phillips-Perron  $p$  statistic and Newey-West estimator).

**3. Simulation Results**

The DGP used when Equations (6) and (7) are I(1) with deterministic components and when we vary lag from 2 to 10 in the simulation, then produces LR test for the DGP recorded in Table 1 below. The likelihood test statistics for the bivariate system defined in (6) and (7) shows that as the number of lag increases then the values of the likelihood statistics decreases for a given fixed sample size.

Table 1: Values of LR Test for lag 2 to 10.

Sample	Time	Lag = 2	Lag = 4	Lag = 8	Lag 10
N = 100	R <sub>0</sub>	62.170	61.885	34.887	22.055
N = 100	R <sub>2</sub>	35.158	28.826	12.636	8.715
N = 100	R <sub>3</sub>	16.926	10.929	3.090	2.486
N = 200	R <sub>0</sub>	147.902	112.608	67.886	53.557
N = 200	R <sub>2</sub>	72.138	50.009	36.257	30.604
N = 200	R <sub>3</sub>	24.648	22.630	17.038	13.186
N = 400	R <sub>0</sub>	205.118	196.547	175.294	125.524
N = 400	R <sub>2</sub>	112.658	101.025	95.556	56.707
N = 400	R <sub>3</sub>	52.125	48.194	40.001	10.125

The truncated and untruncated time used are defined as R<sub>0</sub>, R<sub>1</sub>, R<sub>2</sub> of fixed time of 20 seconds, 21-30seconds and computer unrestricted time respectively.

The result of the non-cointegrated systems that contained deterministic trend components when we used the DGP in (6) and (7) are presented in table 2.

Table 2: I(1) with deterministic components.

Case A: When $\alpha_1 = \alpha_2 = 0$ and $\beta_1 = \beta_2 = 0.01$	EG	Q <sub>1</sub> = JTS	Q <sub>2</sub> = MES
SMC <sub>1</sub> : $\Delta X_t = \Pi_1 X_{t-1} + \epsilon_t$	0.047	1.000	1.000
SMC <sub>2</sub> : $\Delta X_t = \mu + \Pi_2 X_{t-1} + \epsilon_t$	0.068	0.625	0.586
SMC <sub>3</sub> : $\Delta X_t = \mu + \theta t + \Pi_3 X_{t-1} + \epsilon_t$	0.054	0.016	0.044
SMC <sub>4</sub> : $\Delta X_t = \Pi_4 \left( X_{t-1}^1 \ 1 \right)' + \epsilon_t$		1.000	1.000
SMC <sub>5</sub> : $\Delta X_t = \Pi_5 \left( X_{t-1}^1 \ t \right)' + \epsilon_t$		1.000	1.000
Case B: When $\alpha_1 = \alpha_2 = 0$ and $\beta_1 = \beta_2 = 0.01$	EG	Q <sub>1</sub> = JTS	Q <sub>2</sub> = MES
SMC <sub>1</sub> : $\Delta X_t = \Pi_1 X_{t-1} + \epsilon_t$	0.073	1.000	1.000
SMC <sub>2</sub> : $\Delta X_t = \mu + \Pi_2 X_{t-1} + \epsilon_t$	0.080	0.024	0.045
SMC <sub>3</sub> : $\Delta X_t = \mu + \theta t + \Pi_3 X_{t-1} + \epsilon_t$	0.067	0.063	0.064
SMC <sub>4</sub> : $\Delta X_t = \Pi_4 \left( X_{t-1}^1 \ 1 \right)' + \epsilon_t$		1.000	1.000
SMC <sub>5</sub> : $\Delta X_t = \Pi_5 \left( X_{t-1}^1 \ t \right)' + \epsilon_t$		1.000	1.000

T= 100, SMC<sub>i</sub> (i = 1,2,...,5) are the statistical models used for estimating the Johansen statistics Q<sub>1</sub> and Q<sub>2</sub>. The dimension are as earlier discussed in equations (18) to (22). EG is computed from SMC<sub>i</sub>, i = 1, 2, 3. The frequency of rejecting the null in 10,000 replications is reported at 5% level. The critical values for each of the three regression models are simulated from 100,000 replications using the DGP with  $\alpha_1 = \alpha_2 = 0$  and  $\beta_1 \beta_2 = 0$ . The integral of order 1 with deterministic components for case A and B of the DGP of table 2 shows that the Engle-Granger (EG) statistic for SMC<sub>2</sub> in both cases has higher values when compared with the EG values of SMC<sub>1</sub> and SMC<sub>3</sub> in cases A and B. The model developed for I(1) processes with deterministic component as shown in table 2 shows that VECM with X<sub>t-1</sub> augmented with 1 and / or t as consistent with Johansen (1995). This model was developed mainly to capture a cointegration relationship around a common deterministic trend. It shows that estimating the VECM with X<sub>t-1</sub> when there is no

cointegration will produce an egregious pitfall. The  $Q_1$  and  $Q_2$  behave very similar under the null of no pitfall but the results are totally different under the alternative of pitfall. We recommend that both tests should be used in research work.

In addition, when the approach of Mackinnon *et al.* (1996) and Nielsen (1997) was used to estimate the statistics at various sample sizes and fit a response surface and observed the asymptotic distribution. Precisely, the design of  $Q_1$  and  $Q_2$  of Johansen trace statistic and Maximum eigenvalue statistic respectively of  $SMC_i$  ( $i=1, 2, \dots, 5$ ) cases with 100,000 replications, and sample size 50, 100, 250, 1000, 2000, 4000 produces the result in Table 3 below.

Table 3: Estimation of the response surface regression.

n	T	$\bar{D}$	SSEM	T	$\bar{D}$	SE
1	500	1.243	0.0045	50,100,...,4000	1.184	0.0041
2	500	5.843	0.009	50,100,...,4000	5.082	0.0011
4	500	42.39	0.021	50,100,...,4000	40.22	0.048
10	1000	188.10	0.050	50,100,...,4000	186.14	0.083

In Table 3,  $\bar{D}$  is the mean and SSEM is the simulated standard error of the mean when  $M=100000$ . Columns five, six and seven are the sample size, mean and the standard error, respectively. These results were obtained with a response surface of average of thirty experiments with  $T = 500$ . The response surface is a regression of  $\mu$  on a constant,  $\frac{1}{T}$  and  $\frac{1}{T_2}$ ;  $\mu$  is the coefficient on the constant term, with corresponding standard error (S.E). The standard error of the constant in the regression coefficient is quite informative on the mean of the asymptotic distribution. The SSEM is informative on the accuracy for that specific size. The estimates of the mean of each test are obtained by automated regression on a constant,  $\frac{1}{T}$  and  $\frac{1}{T_2}$  and dummy for  $T=50$ . Any of the term, that is, the dummy,  $\frac{1}{T}$  and  $\frac{1}{T_2}$  are dropped when we discovered they are not significant. It was also observed that as the  $n$  increases for both the single simulation with  $M=100,000$  and the response surface based on the average of thirty experiments with  $M=500$ . The standard error also displayed this increase characteristic value as dimension  $n$  increases from 1 to 10.

Table 4: Phillips-Perron Tests - Truncation lag =8 for  $x_t$

Null hypothesis	Test Statistics	Asy. Critical Value 10%
Constant, No trend		
$\Lambda(1)=0$ , Z-Test	-41.813	-11.2
$\Lambda(1)=0$ , T-Test	-5.8574	-2.57
$\Lambda(0)=\Lambda(1)=0$	16.963	3.78
Constant, No trend		
$\Lambda(1)=0$ , Z-Test	-41.720	-18.2
$\Lambda(1)=0$ , T-Test	-5.8036	-3.13
$\Lambda(0)=\Lambda(1)=\Lambda(2) = 0$	11.562	4.03
$\Lambda(0)=\Lambda(2) = 0$	16.562	5.34

Table 5: Phillips-Perron Tests - Truncation lag =8 for  $y_t$

Null hypothesis	Test Statistics	Asy. Critical Value 10%
Constant, No trend		
$\Lambda(1)=0$ , Z-Test	-42.460	-11.2
$\Lambda(1)=0$ , T-Test	-6.0181	-2.57
$\Lambda(0)=\Lambda(1)=0$	17.891	3.78
Constant, No trend		
$\Lambda(1)=0$ , Z-Test	-42.125	-18.2
$\Lambda(1)=0$ , T-Test	-5.9735	-3.13
$\Lambda(0)=\Lambda(1)=\Lambda(2) = 0$	11.667	4.03
$\Lambda(0)=\Lambda(2) = 0$	17.498	5.34



Table 6: Johansen Cointegration Test of no deterministic trend in the data generated.

Eigenvalue	Likelihood Ratio	5% Critical Value	1% Critical Value	Hypo. No of CE(s)
0.708	19041.271	12.53	16.31	None * *
0.490	6735.194	3.84	6.51	Atmost 1 * *

(\*\*) denotes rejection of the hypothesis at 1% significance level

Table 7. Unrestricted Vector Autoregression Estimates when constant is not included.

Lag variables(s) and Statistic	Variable (x)	Variable (y)
x(-1)	-0.682115 (0.00992) (-68.7844)	-0.517381 (0.01050) (-49.2669)
x(-2)	-0.341389 (0.01056) (-32.3382)	-0.388649 (0.01118) (-34.7645)
y(-1)	0.019067 (0.00775) (2.45921)	-1.142572 (0.00821) (-139.157)
y(-2)	0.054381 (0.00699) (7.78224)	-0.508019 (0.00740) (-68.6514)
R-squared	0.336278	0.780058
Adj. R-squared	0.336078	0.779992
Sum sq. resides	39396.84	44181.16
S.E. equation	1.985758	2.102878
Log likelihood	-21036.83	-21609.61
Akaike AIC	1.372401	1.487014
Schwarz SC	1.375286	1.489899
Mean dependent	-0.000378	-0.000366
S.D. dependent	2.437069	4.483265
Det. Residual Covariance = 15.70717		
Log Likelihood = -32133.28		
Akaike Information Criteria = 2.754918		
Schwarz Criteria = 2.757803		

Standard error and t- statistics in parentheses from top to down respectively

Table 8. Unrestricted Vector Autoregression Estimates when constant is included.

Lag variables(s) and Statistic	Variable (x)	Variable (y)
x(-1)	-0.682114 (0.00992) (-68.7809)	-0.517381 (0.01050) (-49.2644)
x(-2)	-0.341389 (0.01056) (-32.3366)	-0.388649 (0.01118) (-34.7627)
y(-1)	0.019067 (0.00775) (2.45909)	-1.142572 (0.00821) (-139.150)
y(-2)	0.054381 (0.00699) (7.78185)	-0.508019 (0.00740) (-68.6480)
C	-7.77E-05 (0.01986) (-0.00391)	6.16E-05 (0.02104) (0.00293)
R-squared	0.336278	0.780058
Adj. R-squared	0.336012	0.779969
Sum sq. resides	39396.84	44181.16
S.E. equation	1.985857	2.102983
Log likelihood	-21036.83	-21609.61
Akaike AIC	1.372601	1.487214
Schwarz SC	1.376208	1.490821
Mean dependent	-0.000378	-0.000366
S.D. dependent	2.437069	4.483265
Det. Residual Covariance = 15.70717		
Log Likelihood = -32133.28		
Akaike Information Criteria = 2.755118		
Schwarz Criteria = 2.758725		

Standard error and t- statistics in parentheses from top to down respectively

Table 9. Vector error correction estimates with linear trend involving constant and trend in data.

Lag variable(s) and Statistic	Variable (x)	Variable (y)
Cointegrating Eq.:	1.000000	
x(-1)	6.887494	
y(-1)	(0.27099)	
	(25.4165)	
@TREND(1)	-5.67E-08	
	(1.0E-05)	
	(-0.00561)	
C	-0.002302	
	(1.0E-04)	
	(1.0E-04)	
Error Correction:	D(x)	D(y)
Co-int Eq.1	-0.052844	-0.601402
	(0.00582)	(0.00406)
	(-9.08250)	(-148.214)
D(x(-1))	-0.963761	0.340135
	(0.01265)	(0.00882)
	(-76.2001)	(38.5614)
D(x(-2))	-0.482219	0.100743
	(0.01100)	(0.00767)
	(-43.8548)	(13.1373)
D(y(-1))	0.299435	1.653935
	(0.02879)	(0.02008)
	(10.4016)	(82.3815)
D(y(-2))	0.169372	0.479424
	(0.01149)	(0.00801)
	(14.7449)	(59.8459)
C	-0.000970	-0.001744
	(0.02681)	(0.01870)
	(-0.03617)	(-0.09324)
R-squared	0.598708	0.950362
Adj. R-squared	0.598507	0.950337
Sum sq. residues	71766.85	34905.32
S.E. equation	2.680542	1.869418
Log likelihood	-24032.11	-20430.36
Akaike AIC	1.972638	1.251855
Schwarz SC	1.976967	1.256183
Mean dependent	-0.000766	-0.000850
S.D. dependent	4.230424	8.388651
Determinant Residual Covariance= 21.06042		
Log Likelihood = -33595.58		
Akaike Information Criteria = 3.048896		
Schwarz Criteria = 3.054307		

Standard error and t- statistics in parentheses from top to down respectively.

It was observed that the unrestricted vector autoregression estimates when constant is included and when constant is not included in the model show that the two models exhibit the same characteristics in terms of their robustness and significance of the estimated parameters. For instance, the Akaike Information Criteria values are 2.758725 and 2.757803 respectively (See tables 7 and 8). From table 6, the L.R. test indicates 2 cointegrating equation(s) at 5% significance level.

Log likelihood is -33439.395.

The first line of table 6 shows that the likelihood ratio is greater than the critical value at

The hypothesis is rejected in favour of cointegration in spite of the deterministic trend in the data.

It must be noted that all programmes are written in Microsoft Visual Basic 6.0 Enterprise, Shazam program and subroutines are partially adopted from White Kenneth's SHAZAM Econometric computer program and E-Views computer software

#### 4. CONCLUDING REMARKS

This work shows that cointegration tests for both univariate unit root tests and likelihood ratio test are very important since pretest for individual unit root is not enough to detect spuriousity in trend stationarity models. It was confirmed that the singularity in VAR that comes from error correction matrix gave a misleading results unlike the one obtained from long-run impact matrix. One other picture displayed by our investigation on the case where variables that have long-memory properties and a trending behaviour but they are not pure  $I(1)$  processes are that the VEC estimates gave better estimates in terms of the R-squared and lower standard error for the model when compared with the unrestricted cases. This paper shows the existence of reversal relationship between the values of LR statistic and the lag-lengths. We also observed that the AIC and Schwarz are more robust and the Z-test and T-test are more stable for size properties for a wider range of nuisance parameter than the coefficient based tests. The effect of the sample size was captured from the simulation. This work also derived the asymptotic relationship between the eigenvalues of the product matrices and the coefficient of determination of the VECM residual matrix. We hope the above mentioned reversal relationship between LR statistics and lag-length could be a base for further mathematical investigations that establish the conditions needed to eliminate this type of pitfalls attributed to the LR. The work on the stationary model with both deterministic component and stochastic components of the type described by the general form of equation (3) is in progress.

#### REFERENCES

- Ahn, S. K. and Reinsel, G. C., 1990. Estimation for partially Non-stationary Multivariate Autoregressive Models, *Journal of American Statistical Association*, 85: 813-823.
- Engle, R. F., and Granger, C. W. J., 1987. Cointegration and Error Correction: Representation, Estimation and Testing", *Econometrica*, 55: 251-276.
- Engle, R. F. and Smith, A. D., 1998. Stochastic Permanent Breaks, Discussion paper 98-03, University of California, San Diego.
- Gonzalo, J. and Lee, T., 1998. Pitfalls in testing for long run relationships, *Journal of Econometrics*, 86: 129-154.
- Granger, C.W. J. and Newbold, P., 1974. Spurious regression in econometrics, *Journal of Econometrics*,
- Hamilton, J. D., 1989. A New Approach to the Economic Analysis of Nonstationary Time Series and Business Cycle, *Econometrica*, 57: 357-84.
- Hamilton, J. D., 1994. *Time Series Analysis*. Princeton; Princeton University Press.
- Hendry, D. G., 1995. *Dynamic Econometrics* (Oxford University Press, Oxford).
- Johansen S., 1995. *Likelihood based Inference on Cointegration in the Vector Autoregressive Model*. Oxford University Press.
- Johansen, S., 1996. A Statistical Analysis of Cointegration for  $I(2)$  variables, *Econometric Theory* 11: 25-59.
- Khogali, A. K, Olowofeso, O.E. and Owino J. O., 2002. Duality, Forecasting and Selection of Autoregressive Moving Average Models. *American InterStat Journal-Statistics on Internet*, April, 2002, No.1, page 1-17.
- Mackinnon, J. G. Haug, A. A. and Michelis, L., 1996. Numerical distribution functions of likelihood ratio tests for Cointegration: Mimeo, Queens University, Canada.
- Nielsen, B., 1997. On the distribution of tests for cointegration ranks" Mimeo, Nuffield College, U.K.
- Olowofeso, O. E. and Fagbohun, A. B., 2001. 'Estimation of cointegrating vectors: A Monte Carlo Comparism, *Journal of Research Communication in Statistics* (Forthcoming)
- Olowofeso, O. E., 2000. Estimation and Hypothesis Testing for the Class of Non-stationary Time Series. Unpublished Ph.D. Thesis, Federal University of Technology, Akure.
- Perron, P., 1989. The great crash, the oil price shock and the unit root hypothesis, *Econometrica*, 57: 1361-1401.
- Saikkonen, P. and Lukkonen, R., 1993. Testing for a Moving Average Unit Root in Autoregressive Moving Average Models, *Journal of American Statistical Association*, 88, 596-601.
- Stock, J. H., and Watson, M. W., 1989. Testing for Common Trends, *Journal of American Statistical Association* 83, 1097-1107.
- White, K. J., 1993. *ZHAZAM*, The Econometrics Computer Program Version 7.0, User's Reference Manual, Mc Graw-Hill.