Advances in Social Sciences Research Journal – Vol.3, No.4
Publication Date: April. 25, 2016
DoI:10.14738/assrj.34.1961.
Nkoro, E., & Uko, A. K. (2016). The Johansen-Juselius Multivariate Cointeration Technique: Application and Interpretation. *Advances in Social Sciences Research Journal, 3(4)* 248-267.

The Johansen-Juselius Multivariate Cointegration Technique: Application and Interpretation.

Dr. Emeka Nkoro
Department of Economics
University of Port Harcourt
Port Harcourt, Nigeria

Dr. Aham Kelvin Uko
Department of Economics
University of Port Harcourt
Port Harcourt, Nigeria

Abstract

Cointegration is the simplest way of detrending series whose mean, variance as well as autocorrelation functions changes over time due to the presence of unit roots. However, the issue has shifted to the application of the appropriate cointegration technique as various cointegration techniques abound. Hence, this study reviews the issues surrounding the application and interpretation cointegration techniques within the context of Johansen-Juselius multivariate cointegration framework. The study shows that the adoption of the Johansen-Juselius multivariate cointegration technique rests on the pretests for unit roots. The study reveals that Johansen-Juselius multivariate cointegration technique is preferable when dealing with more than two variables that are integrated of the same and as well different order, I(d) .However, the Johansen-Juselius multivariate cointegration technique is robust when dealing with variables of the same order of integration. The number of cointegrating vectors is detected through the two likelihood ratio test statistics (trace and maximum test). Although the major difficulty lies in the identification of the cointegrating vectors where there are multiple cointegrating vectors. In this approach, cointegration is said to be established when there is at least one cointegrating vector. Based on forecast and policy implications, this paper explores the conditions that necessitate the application of the Johansen and Juselius cointegration technique. This is to avoid its wrongful application, which may in turn lead to model misspecification, inconsistent and unrealistic estimates. However, this paper cannot claim to have treated the underlying issues in their greatest details, but have endeavoured to provide sufficient insight into the issues surrounding Johansen and Juselius cointegration technique to practitioners to enable them apply and interpret and also, follow discussions of the issues in some more advanced texts.

Keywords: Cointegration, Unit Roots, the Johansen-Juselius multivariate technique, Error Correction mechanism.

INTRODUCTION

Recent advances in econometrics have shown that many macroeconometric time series data are Differenced Stationary Process (DSP), rather than a Trend-Stationary Process (TSP) as was initially thought by econometricians. It is clearly possible to see that most time series exhibit the characteristics of non-stationarity (DSP), that is, diverging away from their mean with the passage of time (non-mean reverting) rather than exhibiting the characteristic of stationarity(converging to the mean overtime). Therefore, when such non-stationary time series (DSP) are used in estimation of an econometric model, the traditional diagnostic

statistics such as, Goodness-of-fit (R2), Fisher's Ratio (F-Statistic), Durbin-Watson (DW Stat), t-statistic etc. use in evaluating the validity of the model becomes highly unreliable. In such series, the mean, variance, covariance and autocorrelation functions change overtime. Ordinary Least Square (OLS) method which was the standard regression method which is based on the assumptions that the means and variances of the variables being tested are constant over the time becomes inefficient and biased in estimating non-stationary variables. To overcome this problem of modeling of DSP series, econometric analysis of time series data has increasingly moved towards the issue of unit roots, cointegration and error- correction modeling. The testing for and the analysis of unit roots clearly lead to the cointegration theory. The reason being that cointegration basically deals with the method of modeling Difference stationary time series.

In applied econometrics, the Granger (1981) and Engle and Granger (EG) (1987), Autoregressive Distributed Lag(ARDL) cointegration technique or bound test of cointegration(Pesaran and Shin 1999 and Pesaran et al. 2001) and Johansen and Juselius(1990) cointegration techniques have become the solution to determining the long run relationship between series that are non-stationary and as well as reparameterizing them to the Error Correction Model (ECM). The reparameterized result gives the short-run dynamics and long run relationship of the underlying variables. Despite the versatility of cointegration technique in estimating relationship between non-stationary variables(DSP) and reconciling the short run dynamics with long run equilibrium, most researchers still adopt the conventional way of estimation even when it is glaring to test for cointegration among the variables under consideration. That is most of the researchers are not conversant with the conditions that necessitate the application of cointegration test and as well the interpretation of the results therein, hence presenting misleading inferences.

This paper is examines the conditions that necessitate the application and interpretation of Johansen and Juselius cointegration technique. Accordingly, this paper is divided into six sections. Section one, which is the introduction. Section two, examines the concept of stationary and non-stationary series, section three looks at the unit roots stochastic processes, section four deals on cointegration test in the context of Johansen and Juselius approach(JJA), section five examines causality tests while section six focuses on summary and conclusions.

STATIONARY AND NON-STATIONARY SERIES CONCEPT

A non-stationary time series is a stochastic process with unit root, hence the presence of a unit root implies that a time series under consideration is non-stationary while the absence of it entails that a time series is stationary. This depicts that unit root is one of the sources of non-stationarity. A non-stationary stochastic process could be Trend Stationary (deterministic) Process (TSP) or Difference Stationary Process (DSP). A time series is said to be trend stationary process if the trend is completely predictable and not variable, whereas if it is not predictable, we call it difference or integrated stochastic trend or difference stationary process. In the case of deterministic trend, the divergence from the initial value(represents non-stationary mean) is purely random and they die out quickly; they do not contribute or affect the long run development of the time series. However, in the case of integrated stochastic trend, the random component (Ut) or divergence affects the long run development of the series. To utilizing time series with these features in any meaningful empirical analysis, the series must be purged of this trend. This is referred as detrending of the series. This could be carried out in two ways, depending on whether the series is difference stationary process or deterministic stationary process. If a series is DSP, it means it has a unit root; hence the

differencing of such series is stationary. Therefore, the solution to the non-stationary series is to difference the series. Also, if a series is TSP, it means it exhibits a deterministic trend, while a trend stationary variable with non-constant mean may be I(0) after removal of a deterministic trend. That is, regressing such series on time(t) and the residuals from this regression will be stationary(Yt =βt + Ut). Hence, cointegration cannot be seen as a means to an end but restricted. It should be made clear that if a time series is TSP but it is treated as DSP, this is called over-differencing. On the other hand, if a time series is DSP, but treated as TSP, this is referred as under-differencing. The implications of these types of specification error can be serious, depending on how the serial correlation properties of the resulting error terms are handled. However, it has been observed that most time series are DSP rather than TSP.

UNIT ROOTS STOCHASTIC PROCESS

The unit roots test is a key pre-condition in the study of a time series model and cointegration within the context of JJA. Knowing the order of integration is crucial in terms of setting up an econometric model and doing inference. The second motive is that economic theory suggests that certain variables should be integrated since they exhibit a random walk (Bo Sj¨o, 2008), with the divergence of its mean and variance over time. According to Maddala(1992), unit root testing is formalization of the Box-Jenkins method of differencing the time series after a visual inspection of the correlogram. A stochastic process Yt is assumed to have a unit root problem if its first difference, (Yt-Yt-1), is stationary. In practice, the presence of unit root shows that the time series under consideration is non-stationary while the reverse is the case. On the other hand, non-stationary series have no tendency to return to long-run deterministic path and the variance of the series is time dependent. Non-stationary series suffer permanent effects from random shocks and thus the series follow a random walk. That is, when all (dependent and independent) time series are non-stationary, the classical regression results may be misleading.

Non-stationarity can be illustrated using a random walk model (RWM) with or without drift:

$Yt = \beta Yt\text{-}1 + \mu t$ (3.1) Random walk model without drift

$Yt = \alpha + \beta Yt\text{-}1 + \mu t$ (3.2)) Random walk model with drift(α)

$(-1 \leq \beta \leq 1)$

Where μt is assumed to satisfy the usual Ordinary Least Square(OLS) assumptions. We then run the regression (equation 3.1 and 3.2) and examine the behavior of β, that is, the coefficient of Yt-1. If β in absolute term is greater or equal to one(I(1)) we say that the series is non-stationary but if the β is less than one (I(0)) we say that the series is stationary. If β =1, in this case, the stochastic element in the series may be a pure random walk. If β>1, then the series actually exhibit explosive characteristics. Testing for stationarity is also referred to as unit root test. If β =1, we face the problem of unit root.

A stochastic process Yt is assumed to have a unit root problem if its first difference, (Yt-Yt-1) is stationary. In practice, the presence of unit root shows that the time series under consideration is non-stationary while the reverse is the case. On the other hand, a series with unit root have no tendency to return to long-run deterministic path and the variance of the series is time dependent. A series with unit root suffers permanent effects from random shocks and thus the series follow a random walk. That is, using (dependent and independent) time series that contain unit root in regression analysis, the classical results of the regression may be misleading. However, I(1) variables that exhibit a random walk without drift may have a mean that is constant over time, expected value of zero and, with trending variance; hence making the series with unit root to having the tendency to return to long-run path after removing

deterministic trend. This reveals that cointegration cannot be seen as a means to an end but restricted. However, this paper focuses on series with unit root, I(1) (no constant mean and variance) that have no tendency to returning to the long-run path.

There are various methods of testing for unit roots. They include; Durbin-Watson (DW) test, Dickey-Fuller test(1979)(DF), Augmented Dickey-Fuller(1981)(ADF) test, Philip-Perron (1988) (PP) test and among others. It is of the views that before pursuing a formal test, to plot the time series under consideration to see the likely features of the series and also run the classical regression. If the series is trending upward it shows that the mean of the series has been changing with time. In the case of the classical regression, if Durbin – Watson statistics is very low and a high R2 (Granger–Newbold, 1974), this perhaps reveals that the series is not stationary. These are the step for a starting point of a more formal test of stationarity. The most popular strategy for testing of the stationarity property of a single time series involves using the Dickey-Fuller or Augmented Dickey Fuller test respectively. The choice of the right tests depends on the set up of the problem which is of interest to the practitioner. It is difficult to follow the latest advances or to understand the problems by employing various tests. This should not be understood as a motive for not performing other types of unit root tests. Comparing different results from different test methods is a good way of testing the sensitivity of the conclusions. The understanding of how these tests work, and their limitations, brings about the understanding of when to use any of the tests. The advantage is that it enables us to understand the meaning and purpose of each of the test. However, when a test results is inconclusive, the usual way is to continue the analysis with a warning note, or simply assume one of the alternatives. Thus, the unit roots test is basically required to ascertain the number of times a variable/series has to be differenced to achieve stationarity, and which is the first step of Johansen and Juselius cointegration analysis. From this comes the definition of integration (Engle and Granger, 1987). For instance, A variable Y, is said to be integrated of order d, I(d)] if it attained stationarity after differencing d times.

The Durbin-Watson Test.

This test is simple but unreliable test for unit root. To understand how this test works, recollect that the DW-value is calculated as DW =2(1−ˆρ)(Harvey, 1981), where ρ = ˆρ is the estimated first order autocorrelation. Thus, if Yt is a random walk, ρ will equal unity and the DW value is zero. Under the null that Yt is a random walk, the DW statistic calculated from the first order autocorrelation of the series Yt = Yt−1 +Vt, will approach one. The DW value approaches 0 under the null of a random walk. A DW value significantly different from zero rejects the hypothesis that Yt is a random walk and I(1), in favor of the alternative that Yt is not I(1), and perhaps I(0). The test is limited by the assumption that Yt is a random walk variable. The test is not good for integrated variables in general. The critical value at the 5% level for the maintained hypothesis of I(1) versus I(0) is 0.17. A higher value rejects I(1)(Bo Sj¨o, 2008).

Dickey-Fuller (DF) (1979) Test for Unit Roots

Assume that Yt is random walk process, Yt = Yt−1 + ut, then the regression model becomes Yt =ρYt−1 + ut. Subtract Yt−1 from both sides of the equation:

$$Yt-Yt-1 = \alpha1Yt-1 -Yt-1 + ut \quad (3.3)$$

$$\Delta Yt = (\alpha-1)Yt-1 + ut \quad (3.4)$$

$$Yt-Yt-1 = \alpha1Yt-1 -Yt-1 + \alpha2T + ut \quad (3.5)$$

$$\Delta Yt = (\alpha-1)Yt-1+ \alpha2T + ut \quad (3.6)$$

Where α-1= p1, Δ is change in Yt or first difference operator and T is the trend factor. μt is assumed to be white noise residual(Ut ~ iiDN(0,σ2).

$$\Delta Yt = p1Yt\text{-}1 + \alpha 1T + \mu t \quad (3.7)$$

With a drift we have;

$$\Delta Yt = \alpha 0 + p1Yt\text{-}1 + \mu t \quad (3.8)$$

$$\Delta Yt = p1Yt\text{-}1 + \alpha 1T + \mu t \quad (3.9)$$

In practice, we test the hypothesis that p=0. If p=0, "α" in equation 3.4 or 3.6 will be equal to 1, meaning that we have a unit root and infer that the series under consideration is non-stationary and the reverse is the case of ρ < 0.

Alternatively, due to case of autocorrelation in the observed series, the ADF test is resorted to since it adjusts the DF test to take care of possible autocorrelation in the error terms by adding the lagged difference term of the dependent variable.

The Augmented Dickey-Fuller (ADF) (1981) Tests

$$\text{Restrictive ADF Model: } \Delta Yt = p1Yt\text{-}1 + \sum_{i=1}^{k} \alpha i\, \Delta Yt\text{-}i + ut \quad (3.10)$$

$$\text{Restrictive ADF Model: } \Delta Yt = p1Yt\text{-}1 + \alpha 2T + \sum_{i=1}^{k} \alpha i\, \Delta Yt\text{-}i + ut \quad (3.11)$$

This model is a restricted model, because it does not allows a constant and but a deterministic trend.

$$\text{General ADF Model: } \Delta Yt = \alpha 0 + p1Yt\text{-}1 + \sum_{i=1}^{k} \alpha i\, \Delta Yt\text{-}i + ut \quad (3.12)$$

$$\text{General ADF Model: } \Delta Yt = \alpha 0 + p1Yt\text{-}1 + \alpha 2T + \sum_{i=1}^{k} \alpha i\, \Delta Yt\text{-}i + ut \quad (3.13)$$

This model is a non restricted model, because it allows a constant and a deterministic trend. k is the lagged values of Δ*Y*(ΔYt-1 - ΔYt-2), to control for higher-order correlation assuming that the series follows an AR(p).

H0: p1 =0 (p1 ~ I(1)), against Ha : p1 < 0(p1~ I(0)).

In practice, DF or ADF value less than its critical value shows that the underlying series is non-stationary and the reverse is the case. However, the null cannot be rejected about non-stationarity based on ADF test since its power is not strong as such. This decision can be verified using other related tests, such as Kwiatkowski-Phillips-Schmidt-Shin (1992)(KPSS) or Philips-Perron (PP) test. PP test has the same null hypothesis as ADF, and its asymptotic distribution is the same as the ADF test statistic. But in the case of KPSS test, the null hypothesis is different. It assumes stationarity of the variable of interest. The results from ADF test differ from KPSS as KPSS does not provide a p-value, showing different critical values instead. In this case, the test statistic value is compared with the critical value on desired significance level. If the test statistic is higher than the critical value, we reject the null hypothesis and when test statistic is lower than the critical value, we cannot reject the null hypothesis. However, when there are conflicting results of the tests, it all depends on the researchers aim and objective. . In general, the null hypothesis for ADF reads that the series is non- stationary while KPSS reads that the series is stationary. For the treatment of serial

correlation, PP reads that there is serial correlation (non-parametric) while ADF reads that there is serial correlation (parametric).

The test can also be performed on variables in first differences as a test for I(2). Under the null ˆp1 will be negatively biased in a limited sample, thus unless yt is explosive. A significant positive value implies an explosive process, which can be a very difficult alternative hypothesis to handle. Contrarily, When testing for I(2) or differencing twice, a trend term is not a possible alternative. The two interesting models here are the ones with and without a constant term. Furthermore, lag length in the augmentation can also be assumed to be shorter.

However, It is a good strategy to start with the model containing both a constant and a trend (equ 3.12), because this model is the least restricted. If a unit root is rejected here, due to a significant p1 there is no need to continue testing. If p1 =0 cannot be rejected, the improved efficiency in a model without a time trend might be better. There is also the lag length in the augmentation to consider (Bo Sj¨o, 2008).

A substantial weakness of the original Dickey-Fuller test (equation 3.8) as earlier stated is that it does not take account of possible autocorrelation in the error process ut. If ut is autocorrelated (that is, it is not white noise) then the OLS estimates of the equations and of its variants are inefficient. Therefore the simple solution is to apply ADF by using the difference lagged dependent variable as explanatory variables to take care of the autocorrelation.

The choice of the number of lags (p) to be included in the unit root test is based on the significant lag of the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) plots of the correlogram and partial correlogram. The value of p is taken to be the number of lags at which the ACF cuts off or the number of lags of the PACF that are significantly different from zero. By rule of thumb, is to compute ACF up to one-third to one-quarter of the length of the time series. The ACF and PACF show different lags that are correlated and compared with the confidence bounds, mostly at 95 percent level. This will lead to AR process in cognizance of the properties of the residual(Uko and Nkoro, 2012). The characteristic of a time series has a far reaching implication for economic and policy formulation and implementation. When a series has a unit root (p1 =0), any shock to the data series is long lasting. Hence, there will be a cumulative divergence from the mean/trend of the series. The instability exhibited by this series will tend to render any policy formulated and implemented on the basis of a model estimated using such data series inefficient. This is because what drives any policy formulation and implementation is the clear assumption of the stability of the series.

COINTEGRATION TEST

Modeling time series in order to keep their long-run information intact can be done through cointegration. Granger (1981) and, Engle and Granger(1987) were the first to formalize the idea of cointegration, providing tests and estimation procedure to evaluate the existence of long-run relationship between set of variables as postulated by the economic theory. Cointegration involves a certain stationary linear combination of variables which are individually non-stationary but integrated, I(d). That is, cointegration analysis involves how to test whether the combination of two or more variables which are individually non stationary are cointegrated or spurious, and how to estimate the cointegrating parameters. It shows the existence of error correction representation of the relevant variables. Thus, cointegration establishes a stronger statistical and economic basis for empirical error correction representation, which brings together short and long-run information in modeling variables.

Testing for cointegration is a necessary step to establish if a model empirically exhibits meaningful long run relationships. If failed to establish the cointegration among underlying variables in their level, it becomes imperative to continue to work with variables in differences instead. But however, the long run information will be missing. There are several tests of cointegration, other than Engle and Granger(1987) procedure, among them is; Johansen-Juselius(1990) cointegration technique. This becomes the focal point of this paper.

Johansen-Juselius (1990) Cointegration Procedure

The residual-based cointegration tests(suitable for bivariate analysis) are inefficient and can lead to contradictory results, especially when there are more than two I(d) variables under consideration and when it seems they have a feedback effect. This test for cointegration, as proposed by Engle and Granger (1987) do not distinguish between the existence of one or two cointegrating vectors. Importantly, Engle and Granger (1987) test relies on super-convergence result. In practice, applying Ordinary Least Square (OLS), its estimates will differ with the arbitrary normalization implicit in the selection of the left-hand-side variables for the regression equation. These different arbitrary normalizations can alter Engle and Granger (1987) test results. Also ARDL cointegration approach (Pesaran and Shin, 1995 and Pesaran et al, 1996b) cannot be applied when there are multiple long-run relations. Therefore, a more satisfactory approach would be to employ Johansen-Juselius(1990) cointegration procedures. Not only does this approach yield maximum likelihood estimators of unconstrained cointegrating vectors, but it allows one to explicitly test for the number of cointegrating vectors and does not rely on an arbitrary normalization and non-pretesting of unit roots.

The Johansen (1988) and, Johansen and Juselius(1990) methods of testing for the existence of cointegrating relationships has become a standard in the econometrics literature. The framework developed by Johansen and Juselius provides a multivariate maximum likelihood technique that leads to the determination of the number of cointegration vectors in an equation. However, in most cases this technique has been wrongly applied by most practitioners or econometricians. This wrong application can lead to significantly misleading results. Also, the associated error correction model estimates will be misleading and unrealistic in relation to forecast and policy implication.

The Johansen- Juselius approach is an extension of the single-equation error correction model to a multivariate one. Suppose we have n variables which can all be endogenous (using matrix notation for Zt = [Y1t, Y2t, Y3t]

$$Zt = \delta 0 + \delta 1t + \prod 1Yt\text{-}1 + \prod 2Yt\text{-}1 + \ldots\ldots + \prod pYp\text{-}1 + ut \qquad (4.1)$$

Where Zt is a vector containing n variables, I(d). The subscript t denotes the time period. δ is an ($nx1$) vector of constants, δ1 is trend, ∏1 through ∏p is an ($n \times n$) matrix of coefficients where ρ is the maximum lag included in the model, and μt is an ($n\ x1$) vector of error terms(i.e. vector of impulses or innovation or shock in the language of VAR, which represent the unanticipated movements in Yt) (μt ∼iidN(0,δ2)).

Reparameterising equation 3.1, that is, subtracting Zt−1 on both sides, leads to

$$\Delta Zt = \delta 0 + \delta 1t + (\prod 1 - I)\, Yt{-}1 + \prod 2\, Zt{-}2 + \cdots + \prod p\, Zt{-}p + \varepsilon t, \qquad (4.2)$$

a final broad form,

$$\Delta Zt = \delta 0 + \delta 1t + \Gamma 1\, \Delta Zt{-}1 + \Gamma 2\, \Delta Zt{-}2 + \cdots + \Gamma\, p{-}1\, \Delta Zt{-}p{+}1 + \prod p\, Zt{-}p + \varepsilon t \qquad (4.3)$$

Where Γi = (I − ∏1 − ∏2 − ∏3 −.. − ∏p) (i = 1, 2, ..., P-1) and ∏ = − (I − ∏1 − ∏2 − ∏3 − .. − ∏p).

Γi is the first row of Γ, and ∏p is the first row of ∏. ΔZt are stationary, i=1,2,…P-1 are all I(d). $\prod p\ Zt-$ must be stationary, I(d).

The only difference between a standard first difference version of a VAR model and equation 4.3 is the term $\prod p\ Zt-p$. Γj represents the dynamics of the model in the short run, that is, the estimable parameters while ∏ rank represents the long run relationship among the variables included in the vector Zt, and I is the identity vector. Cointegration can be detected by examining the ∏ matrix. That is, the matrix ∏ determines the extent to which the system is cointegrated (the number of independent cointegrating vectors) and is called the impact matrix. In other words, it determines how many error correction terms belong in the model.

Conventionally, the rank of the matrix Π is determined by the number of eigenvalues Π that are significantly different from zero. That is, the cointegrating relationship is determined using the rank of the matrix ∏. In doing this, Johansen-Juselius(1990) approach distinguishes three possibilities depending on the values of $\prod(r)$. r = rank of ∏, is the number of linearly independent cointegrating vectors.

Case one: when the rank equals n, all variables in Zt are I(0) (stationary). That is, when $r = n$, where n the number of variables in the system, shows that the vector process Zt is stationary, and there is no stochastic trend in the underlying series. In other words, when Π is of full rank, it implies that the initial assumption that all variables included in the Zt vector are $I\sim(1)$ is no longer valid (Johansen and Juselius, 1990). That is, no common stochastic trends. In this case there is no problem of spurious regression and the simple VAR in levels model can be used to model this case

Case two: when the rank is 0, $r = 0$. That is, n matrix is an n x n matrix of zeros. This implies that the variables included in the model are not cointegrated. This shows that, there is no linear combinations of the variables exist in the vector Zt , and

Case three: when the rank(r) is lower than n. This implies that there exist a maximum of n-1 cointegrating relationships the form β'Zt-l ~ 1(0). However, the issue is to find r greater than zero but less than (0<r<p) or r ≤ (n-1), which implies that a stationary number of linear combinations exist among the vector process Zt (Enders, 2004, and Quang, 2007).

The existence of cointegration between underlying variables in VAR model means that it can be represented in a form of error correction mechanism, in that case, equation 4.3 can be reparameterized to obtain a Vector error correction mechanism;

$$\Delta Zt = \delta 0 + \delta 1t + \Gamma 1\ \Delta Zt-1 + \Gamma 2\ \Delta Zt-2 + \cdots + \Gamma\ p-1\ \Delta Zt-p+1 + \prod p\ Zt-p + \varepsilon t \qquad (4.4)$$

Cointegration is a system property. For simplicity we assume that k = 2, so that we have only two lagged terms, and the model is as following:

$$\begin{Bmatrix} \Delta Y_{1t} \\ \Delta Y_{2t} \\ \Delta Y_{3t} \end{Bmatrix} = \Gamma \begin{Bmatrix} \Delta Y_{1t-1} \\ \Delta Y_{2t-1} \\ \Delta Y_{3t-1} \end{Bmatrix} + \prod \begin{Bmatrix} Y_{t-1} \\ Y_{t-1} \\ Y_{t-1} \end{Bmatrix} + \varepsilon_t \qquad \textbf{(4.5)}$$

The ∏ matrix is 3 x 3 due to the fact that we assume three variables in Zt = [Y1t, Y2t, Y3t]. The ∏ matrix contains information regarding the long-run relationships. We can decompose ∏ =

αβ', where α represents the speed of adjustment to equilibrium coefficients while β' the long-run matrix of coefficients.

Or
We may write the model as the so-called vector error correction model

$$\begin{Bmatrix}\Delta Y_{1t}\\ \Delta Y_{2t}\\ \Delta Y_{3t}\end{Bmatrix} = \begin{Bmatrix}\delta_1\\ \delta_2\\ \delta_3\end{Bmatrix} + \begin{Bmatrix}\Gamma_{11} & \Gamma_{22}\\ \Gamma_{21} & \Gamma_{22}\\ \Gamma_{31} & \Gamma_{32}\end{Bmatrix}\begin{Bmatrix}Y_{1t-1}\\ Y_{2t-1}\\ Y_{3t-1}\end{Bmatrix} + \begin{Bmatrix}\delta_1\\ \delta_2\\ \delta_3\end{Bmatrix}\{\beta' Y_{t-1}\} + \begin{Bmatrix}\varepsilon 1t\\ \varepsilon 2t\\ \varepsilon 3t\end{Bmatrix} \quad \textbf{(4.6)}$$

Or simply

$$\Delta Zt = \delta + \Gamma 1\ \Delta Zt{-}1 + \Gamma 2\ \Delta Zt{-}2 + \cdots + \alpha\beta' Zt{-}p + \varepsilon t, \qquad (4.7)$$

Where $\prod = \alpha\beta'$, is equivalent to the error-correction term (Yt-1 – βo - β1Xt-1) in the single-equation case, except that now β'Zt-1 contains up to (n- 1) vectors in a multivariate framework.. More precisely, in a three variable setting where Y1, Y2 and Y3 are integrated of order one or $I{\sim}(1)$, the Vector Error Correction Model (VECM) can be formulated.

The first part of the error-correction of the first equation (i.e. for $\Delta Y1t$ on the left-hand side) is as follows:

$$\prod_1 Z_{1t-1} = (\{\alpha_{11}\beta_{11} + \alpha_{12}\beta_{12}\}\{\alpha_{11}\beta_{21} + \alpha_{12}\beta_{22}\}\{\alpha_{11}\beta_{31} + \alpha_{12}\beta_{32}\})\begin{Bmatrix}Y_{1t-1}\\ Y_{2t-1}\\ Y_{3t-1}\end{Bmatrix} \qquad \textbf{(4.8)}$$

Where ∏1 is the first row of the n matrix. Equation (4.8) can be rewritten as:

$$\prod 1Zt\text{-}1 = \alpha 11\ (\beta 11\ Y1t\text{-}1 + 21Y2t\text{-}1 + \beta 31Y3t\text{-}1) + \alpha 12(\beta 12Y1t\text{-}1 + \beta 22Y2t\text{-}1 + \beta 32Y3t\text{-}1) \quad (4.9)$$

Equation 4.9 shows two cointegrating vectors with their respective speed of adjustment terms α11 and α12.

Requirements for the Application of Johansen and Juselius Cointegration Technique.

- When there are more than two I(d) variables in the system, the Johansen and Juselius maximum likelihood cointegration approach has the advantage over residual-based approach of Engle and Granger.
- When the variables are integrated of the same order or of different order Johansen and Juselius procedure can handle variables of the same and different orders of integration. The most desirable case is when all the variables are integrated of the same order and then to proceed with the cointegration test. However, this is not always the case, as there could be cases where there is a mix of I(0), I(1) and I(2) variables present in the model, cointegrating relationships might well exist. The inclusion of these variables in our analysis, nevertheless, will massively affect the results. Caution should be applied in such cases as differing orders produce complicated results. Although the positive thing is that it is quite common in econometric analysis to have I(l) variables (Asteriou and Hall, 2007). Johansen and Juselius cointegration technique provides a unified system framework for estimation and testing of cointegration relations in the context of Vector Autoregression (VAR) model. This approach is based on a VAR model of order p to examine the long run relationships that may exist among representative variables and

overcomes some drawbacks associated with the two-step Engle-Granger (1987) method.

- When there are more than one (multiple) cointegrating vectors (relations) among the underlying variables, n>1, the Engle and Granger (1987) or ARDL single-equation approach cannot solve this problem. Therefore, a system analysis such as Johansen-Juselius(1990) cointegration approach is relied upon to solve this problem. In this situation, a single equation analysis cannot provide information about the number of cointegrating relations and, cannot test for several cointegration relations jointly. However, Johansen-Juselius(1990) cointegration approach can handle

The result of a multiple cointegrating regression with k regressors can be expressed as:

$Yt = \varphi 1Xt + ut$
Where $Xt = (X1t; X2t, \ldots\ldots Xkt)'$ is a k-dimensional I(1) regressors given by,
$Xt = Xt\text{-}1 + et$

Where $et=(e1t; e2t,\ldots\ldots ekt)'$ are k-dimensional stationary disturbances, and Independently distributed of ut. In this case we need one additional condition.

Hypothesis: Xt's not cointegrated among themselves.
The Violation of the hypothesis means that there may be more than one cointegrating relations. Such a case cannot be analyzed in a single equation approach, and will be taken care of in the JJ system VAR approach to cointegration.

- If a single long-run relation (one cointegrating vector) exists in a multiple equation approach and the series are integrated of the same order: In this case, the multiple equation method reduces to the single equation approach. Johansen and Juselius approach can be applied in this situation. Here, the variables that come after the dependent variable are equivalent of being weakly exogenous. However, it is only when all the independent variables in a single equation relation are weakly exogenous does the multivariate equation approach provide the same result as a single-equation approach. Also, ARDL and the Engle-Granger (1987) approach can be applied if there is one cointegrating vector of the underlying series and the variables integrated of the same order, I(1). However, the ARDL is robust when the data size is n< 30 and the integration order differs while Johansen and Juselius approach can still be applied but the result should be handled with care.
- Where there are more than one cointegrating vectors, r < n -1 and assuming that only one cointegrating relationship exists, is a very serious problem that cannot be resolved by the Engel and Granger single-equation approach, except johansen-Juselius approach for multiple equations. Only when there is one cointegrating vector does the single-equation approach provide the same result as a multivariate equation approach.

DISADVANTAGE

- Johansen and Juselius cointegration procedure does not require the choice of dependent and independent variables. All variables entering the VAR models are treated as endogenous variables. Also, it is a one step calculation approach. The VAR specification is used to model each variable as a function of all the lagged endogenous variables in the system. Johansen-Juselius(1990) considers that the process Yt is defined by an unrestricted VAR system of order (p).

- Another major difficulty of Johansen and Juselius cointegration procedure is the identification of the cointegrating vectors where there are multiple cointegrating vectors. It assumes that the cointegrating vector remains constant during the period of review. However, after establishing the number of cointegrating relations, it ignores identification of line of direction of the relationship established. This can be handled by ARDL cointegration approach.

The Steps of the Johansen and Juselius Cointegration Approach

This sub-section explores how one determines whether the above requirements are met.

Step 1: Testing for Stationarity of the Variables and its Order of Integration.

In order to avoid the problem of spurious regressions, unit roots test is a key pre-condition in the study of a time series model and cointegration within the context of JJA. If the variables are integrated of same order or of different order of integrations, then we proceed with Johansen and Juselius cointegration test. The most desirable case is when all the variables are integrated of the same order and then to proceed with the cointegration test. However, this is not always the case, as there could be cases where there is a mix of I(0), I(1) and I(2) variables present in the model, cointegrating relationships might well exist. The inclusion of these variables in our analysis, nevertheless, will massively affect the results and caution should be applied in such cases as differing orders are complicated. See section three for types of unit root tests and how they are carried out.

Step 2: Optimum Lag Length Selection Process

It is necessary to determine the dynamic specification of VAR model of equation 4.3 above before the cointegration tests are carried out. Hence, the selection of appropriate lags length (k) using proper information criteria. The determination of the appropriate lag length (p) starts by estimating a VAR model including all the variables in level (non-differenced data). The VAR model should be estimated with variable set with p arbitrary set equal n-1 lags, (i.e. the number of variables under consideration minus one) and this restriction continues until the restriction is rejected(then reducing down by re-estimating the model for one lag less until we reach the appropriate lags). As we estimating and re-estimating the models we inspect the values of lag selection criteria, as well as the diagnosing autocorrelation, heteroskedasticity, possible ARCH effects and normality of the residuals. The specification of a VAR model starts with establishing the order of integration and then appropriate lag-length of its variables. That is, the lag-length is being determined after testing for unit roots. The most common optimum lag selection criteria used are; LR (Log-Likelihood Ratio Criterion), AIC (Akaike Information Criterion), SIC (Schwarz Information Criterion), FPE (Final Prediction Error), HQ (Hannan-Quinn Information Criterion). The model with lowest value of these criteria is chosen as the one with the optimal lag length. This model should also pass all the diagnostic checks.

The values of AIC, SBC and LP for model 4.3 are given by;

$$AICp = -n/2(1+\log 2\pi)-n/2\log\delta 2-P$$
$$SBCp = \log(\delta 2) +(\log n/n)P$$
$$HQC =\log \delta +(2\log\log n/n)P$$
$$LRp,p = n(\log[\sum p]-\log[\hat{}\sum p])$$

Where $\delta 2$ is Maximum Likelihood(ML) estimator of the variance of the regression disturbances, $\hat{}\sum p$ is the estimated sum of squared residuals, and n is the number of estimated parameters, p=0,1,2……P, where P is the maximum order of the VAR model chosen. Having

identifies the lag length, p, the p is now used to test for no autocorrelation in the VAR model using the Godfrey(1988) Lagrange Multiplier(LM).

In small sample, the use of LR statistic tends to result in over-rejection of the hypothesis. In an attempt to take account of this small sample problem, in practice the following degree of freedom adjusted LR statistics are also computed as follows;

LRpp,p= (n-q-2-mP)(log[ˆ∑p]-log[ˆ∑p]), for p =0,1,2….P-1.

These adjusted LR statistics have the same asymptotic distribution as the unadjusted statistics given above.

Step 3: Choosing the Appropriate Model Regarding the Deterministic Components in the Multivariate System

Before carrying out the Johansen cointegration test it is necessary to examine whether an intercept and/or a trend should enter either the short-run or the long-run model, or both models.

$$\Delta Zt = \delta 0 + \delta 1t + \Gamma 1\, \Delta Zt{-}1 + \Gamma 2\, \Delta Zt{-}2 + \cdots + \Gamma\, p{-}1\, ZYt{-}p{+}1 + \prod p\, Zt{-}p + \varepsilon t \qquad (4.10)$$

Equation 4.10 contains an intercept and/or a trend in the long-run model (the cointegrating equation (CE)). Johansen (1995) distinguishes between five models depending on whether equation 4.10 contains intercepts and/or time trends, and whether the intercepts and the trend coefficients are restricted. Although the first and the fifth model are not that realistic, however we present all of them for easy clarification.

Ordering these models

1. No intercepts and no deterministic trends in CE or VAR: (δ0y = δ1y = 0) In this model there are no deterministic components in the data or in the cointegrating relations. However, this is quite unlikely to occur in practice, especially as the intercept is generally needed in order to account for adjustments in the units of measurements of the variables.
2. Restricted intercepts and no deterministic trends or intercept (no trend) in CE, no intercept in VAR:(δ0y =∏yuy, δ1y=0) In this model there are no linear trends in the data, and therefore the first differenced series have a zero mean. In this case the intercept is restricted to the long-run model (i.e. the cointegrating equation) to account for the unit of measurement of the variables
3. Unrestricted intercepts and no deterministic trends or intercept in CE and VAR, no trend in CE and VAR: (δ0y ≠0, δ1y =0). In this type of model there are no linear trends in the levels of the data, but allows both specifications to drift around an intercept. In this model the intercept in the CE is assumed to cancell out by the intercept in the VAR, leaving just one intercept in the short-run model.
4. Unrestricted intercepts and restricted deterministic trends or intercept and linear trend in CE and VAR, no intercept in VAR: (δ0y ≠0, δ1y ==∏yΓy). In this model a trend is included in the CE as a trend-stationary variable in order to take into account exogenous growth (i.e. technical progress). The model also allows for intercepts.
5. unrestricted intercepts and deterministic trends or intercept and quadratic trend in CE, intercept in VAR: δ0y ≠0, δ1y ≠0) .This model allows for linear trends in the short-run model and thus quadratic trends in the CE. Thus, in this model intercept and trend are

unrestricted. However, this model will be very difficult to interpret from an economics point of view.

The importance of distinguishing between the five cases above is that, in the case where intercept and trend are both unrestricted (i.e. case V), Zt will be trend-stationary when the rank of ∏z is full. But when ∏z is rank deficient (n-1), the solution of Zt will contain quadratic trends, unless the trend is restricted as in a case IV. Similarly, in case III, when ∏z is rank deficient then Zt will contain a linear deterministic trend, unless intercept is restricted as in case II.

The problem is which of the three model (1, 2 and 4) is appropriate in testing for cointegration, given that model one and five which are unlikely to happen. Johansen (1992) suggested Pantula principle which is a joint test hypothesis of both the rank order and the deterministic components. The Pantula principle involves the estimation of all three models and the presentation of the results from the most restrictive hypothesis (i.e. r = 0) through the least restrictive hypothesis (i.e. r = n- 1). The model-selection procedure then comprises moving from the most restrictive model, at each stage comparing the trace test statistic to its critical value of the three models, stopping only when, for the first time the null hypothesis of no cointegration is not rejected. To apply the Pantula principle, first specify the appropriate number of lags and estimate the three models. The trace statistics from these models are needed collectively for the Pantula principle to choose the appropriate model. In selecting the appropriate model we start with the smaller number of cointegrating vectors r = 0, and check whether the trace statistic rejects the null in each of the models, if yes we proceed to checking whether other models rejects the null, and so on. We consider the results from model that have the highest cointegrating vectors relatively. However, as far as the specification of intercept and trend in the VAR is concerned, Johansen-Juselius, (1990) opted for VAR model that does not contain deterministic trends, but contains unrestricted intercepts; case III .

Step 4: Determining the Number of Cointegrating Vectors

According to Johansen (1988) and Johansen and Juselius (1990) developed two likelihood ratio test statistics for determining the number of cointegrating relationships, and both involve estimation of the matrix ∏. This is a k x k matrix with rank r. The procedures are based on propositions about eigenvalues. The two likelihood ratio test statistics are the trace and maximum eigenvalues tests.

The Trace Test

The trace test examines the null hypothesis of cointegrating vectors is less than or equal to r against the alternative hypothesis that Π(r) is of the full rank, $r = n$ cointegrating vectors. Lütkepohl and Saikkonen (2000) formulated the hypothesis as follows:

H0 (r0): rk(Π)=r0 with the alternative H1 (r1): rk(Π)>r1. Its test statistic is given by;

$$LR_{trace}(r) = -T\sum_{i=r+1}^{n} \ln(1-\hat{\lambda}_i)$$

for $r = 0,1,2..,\ p-2,\ p-1.$

The Maximum Eigenvalue Test

On the other hand, the maximum eigenvalue test examines the null hypothesis of $r = 1$ cointegrating vectors against the alternative hypothesis of $r + 1$ or $r > 1$ cointegrating vectors. H0 (r0): rk(Π)=r0 with the alternative H1 (r1): rk(Π)=r1+1. Its test statistic is given by:

LRMax(r, r+1)= -TIn(1-ˆλ r+1) = LRtrace(r)- LRtrace(r+1)
The maximum eigenvalue statistic, r =0, 1, 2…. n-1

Where T is the sample size and ˆλi ,ˆλr+1 are the P-r eigenvalues, or characteristic roots, which have been obtained from the matrix Π. Notably, the both tests of cointegration are distributed asymptotically X2 with degrees of freedom $p-r$.

The computed log-likelihood values and eigenvalues are compared to the critical values, by this the exact number of cointegrating equations is determined. After cointegration relations have been established, the resulted long-run equation, the cointegration equation is viewed and analyzed base on the objective of the researcher.

Tests based on log-likelihood values have non-standard distributions. The critical values depend upon which deterministic terms are included, and whether they are restricted or unrestricted. However, in practice, it is advisable to use the unrestricted constant in testing for cointegrating rank since most macroeconomic variables are trended (Johansen-Juselius,1990).

One of the weaknesses of these Johansen and Juselius trace and maximum tests is discrepancy of the results of both tests. However, the choice of r, the number of cointegrating vector should be based on the test that has the least cointegrating vectors perhaps from economic theory. In most cases, the maximum eigenvalue statistic test is recommended based on its low r cointegrating vector (Enders, 2004 and Banerjee et al, 1993), since Johansen and Juselius(1990) noted that the power of trace test is lower. Thus, the Johansen and Juselius analysis allows for one cointegrating vector.

The Johansen-Juselius(1990) approach of testing cointegrating rank is very sensitive to the lag length and upon which deterministic terms are included in the VAR system. Therefore, it is important to determine the appropriate lag length and deterministic terms in order to prevent hypothesis testing that may be misleading (Enders, 2004). This is done before carrying out the cointegration tests.

Step 5: Testing for Exogeneity

After determining the number of cointegrating vectors, next is the tests for exogeneity. Remember that the ∏ matrix contains information about the long-run relationships. From this, when there are r ≤ n-1 cointegrating vectors in β, then this automatically means that at least (n - r) columns of α are equal to zero. Thus, once we have determined the number of cointegrating vectors r ≤ n-1, we should proceed with testing which of the variables are exogenous. Categorizing variables into endogenous and exogenous depends mainly on the theory. In this regards, Wu-Hausman(1973) T2 statistic provided useful approach to testing for exogeneity. The Wu-Hausman statistic (Wu's T2 statistic) is equal the value of the F statistic. This approach enables us to test if explanatory/exogenous variables are really exogenous variables. Given this equation:

Zt = δ0 + δ1t + ∏1Yt-1 + ∏2Yt-1 +……. + ∏pYp-1 + ut (4.11)

In this equation we can test exogeneity with respect to the long run parameters which equivalent to testing which of the rows of α are equal to zero.

Computation of Wu-Hausman T2 statistic can be carried out in the following manner:

Run a regression of the variables on the other variables' lags respectively and save the residuals therein. After which run another regression of the variables on the other variables' lags and the residuals generated from the equations. Use the F test (or the t test for one regression coefficient) to test the significance of the coefficients of the residuals. If the test shows significant, reject the null hypothesis that the variable(s) is exogenous in the underlying equation. That is, accept that the corresponding variable(s) is endogenous. If the test reveals not significant, accept the null hypothesis, that is accept that the corresponding variable(s) is exogenous.

If a variable is found to be exogenous it can drop as an endogenous part of the system. This means that we can drop the whole equation for that variable although it will continue to feature on the right-hand side of the other equations. This can be carried out using Eview VAR specification.

Step 6: Vector Error Correction Model (VECM)

After establishing the number of cointegrating vectors, the next is to estimate the Error Correct Model. In estimating the VECM we need to specify which model we are estimating, based on Pantula principle result and numbers of cointegrating vectors as previously determined from the step.

We may write the model as the so-called vector error correction model

$$\begin{Bmatrix}\Delta Y_{1t}\\ \Delta Y_{2t}\\ \Delta Y_{3t}\end{Bmatrix} = \begin{Bmatrix}\delta_1\\ \delta_2\\ \delta_3\end{Bmatrix} + \begin{Bmatrix}\Gamma_{11} & \Gamma_{22}\\ \Gamma_{21} & \Gamma_{22}\\ \Gamma_{31} & \Gamma_{32}\end{Bmatrix}\begin{Bmatrix}Y_{1t-1}\\ Y_{2t-1}\\ Y_{3t-1}\end{Bmatrix} + \begin{Bmatrix}\delta_1\\ \delta_2\\ \delta_3\end{Bmatrix}\{\beta' Y_{t-1}\} + \begin{Bmatrix}\varepsilon_{1t}\\ \varepsilon_{2t}\\ \varepsilon_{3t}\end{Bmatrix} \quad \textbf{(4.12)}$$

Or simply

$$\Delta Zt = \delta + \Gamma 1\ \Delta Zt-1 + \Gamma 2\ \Delta Zt-2 + \cdots + \alpha\beta' Zt-p + \varepsilon t, \qquad (4.13)$$

Where $\prod = \alpha\beta'$, is equivalent to the error-correction term (Yt-1 – βo - β1Xt-1) in the single-equation case, except that now β'Zt-1 contains up to (n- 1) vectors in a multivariate framework. More precisely, in a three variable setting where Y1, Y2 and Y3 are integrated of order one or $I\sim(1)$.

Most times we might be testing restrictions regarding the long-run proportionality between endogenous variable and exogenous variable(s). To show how a proportionate change in exogenous variable(s) result in the same proportionate change in endogenous variable(s). This leads to impose restrictions on the elements of the α and β matrices. The restrictions are entered as b(1, 1) = 0 for the β11 = 0 restriction. More than one restriction can be entered and they should he separated by commas. For example, doubling xt leads also to the doubling of yt..

CAUSALITY TESTS

It is essential to consider the relationship among the variables under consideration using the Causality tests. Causality establishes the ability of one variable to predict the other(s). It is a statistical measure that provides the extent to which lagged values of a set of variables (say xt) are important in predicting another set of variables (say yt) once lagged values of the latter set are included in the model. The causality between the underlying variables can be captured by a VAR model. The causality tests utilize the concept of VAR models. VAR model allows for the test of the direction of causality. Causality tests can be conducted in two different ways depending on the results of the long run analysis of the underlying variables. There are appropriate tests that detect the cause and effect relationship among the variables. These tests

include Granger causality (1969) test, Sims causality (1972) test and Geweke et al (1983) causality test; however, we focus on Granger causality test.

Granger Causality Test

The Granger causality test (Granger, 1969) is the most well known test causality. It is suitable for analyzing the short-run relationship if no cointegration exists among the variables which are integrated. The Granger causality test shows how one variable can predict the other with or without a feedback effect. A variable Yt is said to Granger-cause Xt, if Xt can be predicted with greater accuracy by using past values of the Yt variable rather than not using such past values, all other terms remaining unchanged. The conventional Granger Causality test is given by the following;

$$Z_{1t} = \alpha_{10} + \alpha_{11t} + \sum_{i=1}^{P} \Phi_{i11} Z_{1t-1} + \sum_{i=1}^{p} \Phi_{i,12} Z_{2,t-1} + \psi_1 W_t + \mu_{1t} \quad (5.1)$$

$$Z_{2t} = \alpha_{20} + \alpha_{21t} + \sum_{i=1}^{P} \Phi_{i21} Z_{1t-1} + \sum_{i=1}^{p} \Phi_{i,22} Z_{2,t-1} + \psi_2 Wt + \mu_{2t} \quad (5.2)$$

where Zt = (Z'1t, Z'2t), where Z1t and Z2t are m1x 1 and m2 x1 subset of Zt and m=m1+m2. The hypothesis that the subset Z2t do not Granger Cause Z1t is defined by;

HG: Φ12= 0, where Φ12=(Φ1,21, Φ2,21,……. Φp,21).

It is assumed that both μ1t and μ2t are uncorrelated white-noise error terms. In this model we can have the following different cases:

Procedure in Carrying out Causality Test

Stage 1: Regress current Z1t on all lagged Z1t terms as in the following model:

$$Z_{1t} = \alpha_{10} + \alpha_{11t} + \sum_{i=1}^{P} \Phi_{i11} Z_{1t-1} + \sum_{i=1}^{p} \Phi_{i,12} Z_{2,t-1} + \mu_{1t} \quad (5.3)$$

From this regression obtain the restricted Residual Sum Square (RSSR)·

Stage 2: Regress current Z1t on lagged Z1t terms and Z2t lagged terms as in the following model:

$$Z_{1t} = \sum_{i=1}^{P} \Phi_{i11} Z_{1t-1} + \sum_{i=1}^{p} \Phi_{i,12} Z_{2,t-1} + \mu_{1t} \quad (5.4)$$

From this regression obtain the unrestricted Residual Sum Square (RSSUR)·

Stage 3: Set the hypotheses as follows:
$\sum_{j=1}^{n} \Phi i\, j = 0$ and $\sum_{j=1}^{n} \Phi i\, j \neq 0$,

unidirectional Granger causality exists from *X* to Y, but not vice versa. In other words, changes in *X* can help to predict future values of *Y*, but *Y* cannot help to predict future values of *X*.

$\sum_{j=1}^{n} \Phi i\, j \neq 0$ and $\sum_{j=1}^{n} \Phi i\, j = 0$,
changes in *Y* can help to predict future values of *X* but not vice versa.
$\sum_{j=1}^{n} \Phi i\, j = 0$ and $\sum_{j=1}^{n} \Phi i\, j = 0$,

it can be concluded that *X* and *Y* do not help to predict one another.
$\sum n_{j=1} \Phi i\, j$ and $\sum n_{j=1} \Phi i\, j$,

If there is a feedback relationship between the two variables *X* and *Y*, then there is bi-directional Granger causality among the variables. $\sum n_{j=1} \Phi ij$ and $\sum n_{j=1} \Phi ij$, are both significantly different from zero.

Stage 4: Hypotheses Testing

$$F_{cal} = \frac{[RSS_R - RSS_{UR}/P]}{[RSS_{UR}/n-k-1]}$$

Where *RSSR* and *RSSUR* are residual sum of squares of the restricted and unrestricted models, respectively, *p* is the number of Z2 lagged terms, *k* is the number of parameters estimated in the unrestricted model, is the number of observations. The restricted model occurs when the above model's parameters are restricted by the null hypotheses conditions mentioned above.

Stage 5: If the Fcal value exceeds the critical F value at the given level of significant, reject the null hypothesis and conclude that Z2t causes Z1t·

Stage 1-5 can be repeated to test equation 5.1b, that Z1t causes Z2t
Notably, the concept of causality in the Granger test does not mean that changes in one variable cause changes in another variable, as the term is used in the context of policy discuss. The Granger test provides a statistical measure of the extent to which predictability exist among the variables of interest. That is, the extent to which lagged values of a set of variables(say Z2t) are important in predicting another set of variables(say Z1t) once lagged values of the latter set are included in the model.

Disadvantage of Granger Causality Test

- The conventional Granger causality test becomes misleading when the series are integrated and long run relationship exists among the series (cointegrated). In such a case, it becomes imperative to reparameterize the model into VECM form (see Hendry et al., 1984; Johansen, 1988) in order to test for causality among the cointegrating variables. If the underlying variables are integrated,I(d) and no long run relationship exist, in such case, VAR models in differences, I(d) are used. The standard F- test is still used to test the hypothesis. The reparameterized models are as follows:

$$\Delta Z_{1t} = \alpha_{10} + \Phi_{2i}\sum_{i=1}^{P}\Delta Z_{2,t-1} + \psi_1\sum_{i=1}^{P}\Delta\gamma_{t-1} + \Phi_{1i}\varepsilon_{1t-1,} + \mu_{1t} \qquad (5.5)$$

$$\Delta Z_{2t} = \alpha_{20} + \Phi_{1i}\sum_{i=1}^{P}\Delta Z_{1t-1} + \psi_2\sum_{i=1}^{P}\Delta\gamma_{t-1} + \Phi_{2i}\varepsilon_{2t-1,} + \mu_{2t} \qquad (5.6)$$

where $\varepsilon 1t\text{-}1 = Z1t\text{-}1 - \Phi 2iZ2t\text{-}1 - \psi 1\gamma t\text{-}1$ and $\varepsilon 2t\text{-}1 = Z2t\text{-}1 - \Phi 1iZ2t\text{-}1 - \psi 1\gamma t\text{-}1$ are the residuals of the cointegration equations.

The null hypothesis now reads: Z2 does not Granger-cause Z1, given γ, is H0 ($\Phi 1 = \Phi 2 = 0$). This means that there are two sources of causation for Z1, either through the lagged terms ΔZ2 through the lagged cointegrating vector. If the coefficients cointegrating vectors, ε1t-1 and ε2t-1 are statistically significant and the F- tests reject the hypothesis that the coefficients are equal to zero, suggesting that in both cases there is a long bi-directional causality relationship

among the underlying variables. Causality in the long run exists only when the coefficient of the cointegrating vector, ε1t-1 is statistically significant and different from zero (Granger and Lin, 1995). However, due difficulties associated with the interpretation of VAR models.

- When the underlying variables are integrated but not cointegrated the conventional Granger causality test becomes invalid. Integrated series cannot Granger cause each other in the long run unless they are cointegrated. Therefore test for causality among cointegrated series can only be carried out using the VECM representations (Asteriou and Hall, 2007).
- The direction of the causality to an extent depends on the number of lagged terms included.
- The Granger causality test assumes that the error terms entering the causality test are uncorrelated. If this is not the case, the Granger causality test gives an unrealistic estimate.

SUMMARY AND CONCLUSIONS

Cointegration theory is definitely an innovation in applied econometrics that has created the most interest among economists in the last decade in trying to solve the problem of spurious regression. Like any other area in economics, there have been methodological issues regarding the application of the right cointegration techniques. Based on this, the study reviewed Johansen and Juselius Multivariate (1990) cointegration technique in terms of its application and interpretation. This motivated the following findings: that pretests of the unit roots is the necessary condition for the adoption of the Johansen and Juselius cointegration technique. Therefore, Johansen-Juselius multivariate technique is adopted when the underlying variables are integrated of the same order, I(1) and of different order I(d). The Johansen-Juselius multivariate cointegration technique is robust when dealing with variables of the same order of integration than differing order. The number of cointegrating vectors are determined through the two likelihood ratio statistics (trace and maximum test), and the long run relationship is established when there is at least one cointegrating vector among the underlying variables. The Johansen-Juselius multivariate technique provides a unified framework for testing and estimation of cointegration relations in the context of Vector Autoregression (VAR) model or system equations.

Re-parameterising the VAR model into the ECM have become the solution to determining the long run relationships between integrated series. The reparameterized result gives the short-run dynamics and long run relationships between the underlying variables.

This review is an important starting point for future reliable research. Johansen and Juselius cointegration technique is one of the greatest discoveries of the 20th century as it helps in solving the problem of spurious relationships. Therefore, this paper explores the conditions that necessitate the application of the Johansen and Juselius cointegration technique. This is to avoid its wrongful application and interpretation which may in turn lead to model misspecification and unrealistic estimates. However, this paper cannot claim to have treated the underlying issues in their greatest details, but have endeavoured to provide sufficient insight into the issues surrounding Johansen and Juselius cointegration technique to young practitioners to enable them apply and interpret and, also, follow discussions of the issues in some more advanced texts.

References

Asteriou, Dimitrios and Hall, Stephen G. (2007). Applied Econometrics: A Modern Approach, PALGRAVE MACMILLAN, New York.

Banerjee, A., J. Dolado, J. W. Galbraith, and H. David (1993): Co-integration, Error Correction

and the Econometric Analysis of non-Stationary Data. Oxford University Press, USA.

Bo Sj¨o (2008): Testing for Unit Roots and Cointegration, Memo.

Brandt, P. T. and J.T. Williams (2006): Multiple Time Series Models. Series: Quantitative Applications in the Social Sciences. 148, Sage Publications Ltd, London.

Dickey, D. and W. Fuller (1979): Distribution of the Estimators for Autoregressive Time Series with a Unit Root, Journal of the American Statistical Association 74, 427-431.

----------(1981): Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root, Econometrica 49, 1057-1072.

Enders, W. (2004): Applied Econometric Time Series, 2nd Ed, John Wiley & Sons Inc, New York.

Engle, R. F. (1982): Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of UK Inflation, Econometrica, 50, 987-1007.

Engle, R. and G. Granger (1987): Cointegration and Error Correction: Representation, Estimation and Testing, Econometrica 55, 251-276.

Geweke,J., Meese, R. and Dent, W. (1983): Comparing Alternative Tests of Causality in Temporal System, Journal of Econometrics, 77, Pp 161-194.

Godfery, L.G (1988): Misspecification Test in Econometrics: The Lagrange Multiplier Principle and Other Approaches, Cambridge University press, Cambridge.

Granger, C. W. J. (1969): Investigating Causal Relations by Econometric Models and Cross Spectral Methods, Econometrica, 37, 428-438.

Granger, C. (1981): Some Properties of Time Series Data and Their Use in Econometric Model Specification. Journal of Econometrics 28, 121-130.

Granger, C.W.J. (1983): CoIntegrated Variables and Error-Correcting Models, UCSD Discussion, Paper 83-13.

----------(1986): Developments in the Study of Cointegrated Economic Variables, Oxford Bulletin of Economics and Statistics 48, 213-228.

----------(1988): Some Recent Developments in a Concept of Causality, Journal of Econometrics 39, 199-211.

Granger, C. and P. Newbold,(1974): Spurious Regressions in Econometrics, Journal of Econometrics 26, 1045-1066.

Granger, C.W.J. and Weiss, A.A (1983): "Time Series Analysis of Error-Correcting Models", in: Studies in Econometrics, Time Series and Multivariate Statistics. Academic Press: New York, 255-78.

Gujarati, D.N. (2004): Basic Econometrics, Fourth Edtion, Tata McGraw-Hill Publishing Company Limited, New Delhi

Johansen, S. (1988): Statistical Analysis of Cointegration Vectors, Journal of Economic Dynamics and Control 12, 231-254.

Johansen, S. (1995): Likelihood Based Inference in Cointegrated Vector Autoregressive Models, Oxford University Press: Oxford.

Johansen, S and K. Juselius(1992): Testing Structural Hypotheses in Multivariate Cointegration Analysis of the PPP and UIP for Uk, Journal of Econometrica, 53, Pp 211-44

Johansen, S. and K. Juselius (1988) : Hypothesis Testing for Cointegration Vectors with an Application to the Demand for Money in Denmark and Finland, Working Paper No. 88-05, University of Copenhagen.

Johansen, S. and K. Juselius (1990): Maximum Likelihood Estimation and Inference on Cointegration-With Applications to the Demand for Money, Oxford Bulletin of Economics and Statistics, 52(2), 169-210.

Lütkepohl,H. and , P. Saikkonen (200):Testing for the Cointegrating Rank of a VAR Process with a Time Trend, Journal of Econometrics 95, 177-198.

Maddalas, G.S(1992): Introduction to Econometrics, 2nd Ed, Englewood Cliffs, Prentice Hall.

Pesaran, M.H. and B. Pesaran (1996a). Microfit 4.0, Cambridge London, Windows Version, CAMFIT DATA LIMITED

Phillips, P, and P. Perron (1988): "Testing for a unit root in time series regression." Bimetrika, 75: 335-346.

Quang, T.V. (2007) :Testing Cointegration for Czech Stock Market, Center for Research in Economic Dynamics and Econometrics.

Ruxanda, G and A. Botezatu (2008): Spurious Regression and Cointegration. Numerical Example: Romania's M2 Money Demand, Romanian Journal of Economic Forecast, 3.

Seddighi, H.R, K.A. Lawler and A.V. Katos (2000): Econometrics: A Practical Approach, Routledge, London EC4P4EE.

Uko, A.K and E. Nkoro (2012): Inflation Forecast with ARIMA, Vector Autoregressive and Error Correction Models in Nigeria, EJEFAS, Issue 50, July.

