A Simple Integration Order Test: An Alternative to Unit **Root Testing**

C. G. Amaefula

Abstract — The paper introduces order of integration test (OIT) which serves as a simple alternative to unit root test built generally using auxiliary autoregressive AAR(3) model. The parametric boundary conditions necessary and sufficient for testing the null hypothesis that the non-stationary variable under test is integrated order zero I(0) were estimated via generalized least squares (GLS). The decision on the hypothesis is evaluated using t-statistic. The test procedure was applied to a simulated non-stationary series (y1) of sample size n = 2000 and a known non-stationary time series data (y_2) with two unit roots. The results showed that y1 is integrated order one (I(1)) and y₂ is I(2). These results were confirmed by Augmented Dickey Fuller (ADF); Phillips-Perron (PP); Kwiatkowski, Phillips, Schmidt, and Shin (KPSS); Elliot, Rothenberg, and Stock Point Optimal (ERS) and Ng and Perron (NP) unit root tests. For logarithm transformed variable, the divergent opinions of other unit root tests in clear-cut solution of the integrated order of such variable makes the new test procedure a better alternative. Nevertheless, the simplicity and aptness of the integration order test give it leverage over conventional methods of unit root test.

Key words — AAR(3), GLS, Integration order test, unit roots, parametric boundary conditions.

I. INTRODUCTION

The order of integration of a series is very important in Econometrics and Time series modelling. Researches on unit root testing have attracted attention for the past three decades. Detecting the order of integration of a data series becomes somewhat difficult except a unit root test is carried out. If a data series is non-stationary, and a unit root test has confirmed that the level series of such data is not stationary then, after the first differencing and the data series becomes stationary, such data series is said to be integrated order one I(1). But if the second differencing is needed to achieve stationarity, such data series is said to be integrated order two I(2). This means that not less than a double unit root tests is needed before confirming that a non-stationary series is I(1) or I(2).

Generally, the integration order of a variable is denoted by I(d), where d is the number of differencing needed to make the non-stationary variable stationary. If a variable is stationary, then, it is said to be integrated order zero I(0). The series could be non-stationary because of random walk, drift, or trend. A series is said to be stationary if its mean and autocovariances do not depend on time. Any series that is not stationary is said to be non-stationary. According to

Published on July 24, 2021. C. G. Amaefula, Federal University Otuoke, Nigeria. (corresponding e-mail: wordwithflame @gmail.com)

[5], Regressing a non-stationary variable on a deterministic trend generally does not yield a stationary variable (instead the series needs to be differenced prior to processing). Thus, using standard regression techniques with non-stationary data can lead to the problem of spurious regressions involving invalid inferences based on t- and F-tests.

A non-stationary variable becomes stationary after it is differenced (although not necessarily just by firstdifferencing — it will be shown that the number of times a variable needs to be differenced in order to induce stationarity depends on the number of units roots it contains). According to [9], differencing will handle the transformations to stationarity. The question of whether a variable is stationary depends on whether it has a unit root.

The Augmented Dickey-Fuller (ADF) test build a parametric correction for higher-order correlation, and it assumed that the v series follows an AR(p) process and adding p lagged difference terms of the dependent variable y to the right-hand side of the test regression. Moreover, while the assumption that y follows an autoregressive (AR) process may seem restrictive, [12] showed that the (ADF) test is asymptotically valid in the presence of a moving average (MA) component, provided that sufficient lagged difference terms are included in the test regression.

[4] propose a simple modification of the (ADF) tests in which the data are detrended so that explanatory variables are "taken out" of the data prior to running the test regression. [11] propose an alternative (nonparametric) method of controlling for serial correlation when testing for a unit root. The PP method estimates the non-augmented DF test equation and modifies the t-ratio of the α coefficient so that serial correlation does not affect the asymptotic distribution of the test statistic. According to [13], Philips type of test has poor size properties (i.e., the tendency to over reject null hypothesis when it is true) when the underlying data generating process has negative moving average (MA) components and MA terms are present in many macroeconomic time series.

The [6] introduced a test that differs from the other unit root tests called the KPSS test, which has a null of stationarity of a series around either mean or a linear trend; and the alternative assumes that a series is non-stationary due to presence of a unit root. In this respect it is innovative in comparison with earlier Dickey-Fuller test, or Perron type tests, in which null hypothesis assumes presence of a unit root. [10] construct four test statistics that are based upon the GLS detrended data y_t^d . These test statistics are modified forms of Phillips and Perron Z_{α} and Z_{t} statistics, the [1] R_1 statistic, and the ERS Point Optimal statistic.

The study presents a simple analytical method of testing for integration order using AAR(3) process with optional intercept term. The remaining part of the paper is arranged as follows: section II presents the materials and methods, section III presents data analysis and results and section IV deals with summary and conclusion.

II. MATERIALS AND METHODS

The method in this study is evidence-based approach relies on direct observation and experimentation in the acquisition of new knowledge. However, this section presents the methodology of the proposed order of integration test mechanism that can be used as a simple alternative to unit root test.

A. A Non-stationary Series

A series is said to be non-stationary if it is integrated order one I(1) or order two I(2). The order of integration is the number of unit roots contained in the series, or the number of differencing needed to make the series stationary. Consider a non-stationary autoregressive (1) process of the form:

$$y_t = \rho y_{t-1} + e_t \tag{1}$$

where $\rho = 1$, (1) becomes a random walk. e_t is a stationary random disturbance term, since:

$$y_{t} - y_{t-1} = e_{t}$$

$$(1-L) y_{t} = e_{t}$$

$$\nabla y_{t} = e_{t}$$

$$(2)$$

The first difference of y denoted by ∇y_t at time t in (2) is stationary, that is I(0) and $e_t \sim (0, \sigma_t^2)$. Sometimes, unit root might still be present in a differenced series ∇y , this means the variable y is integrated order two I(2), hence, second differencing is required to make it stationary. That

$$\nabla y_{t} = \rho^{\bullet} \nabla y_{t-1} + e_{t1}$$

$$\nabla y_{t} - \nabla y_{t-1} = e_{t1}$$

$$(1 - L') y_{t} = e_{t1}$$

$$\nabla^{d} y_{t} = e_{t1}$$
(3)

where $\rho^{\bullet} = 1$ and d = 2. In (3), y_t is said to be integrated order two I(2) and e_{t1} is a stationary random disturbance term. The null hypothesis $H_0: \rho = 1$ or $\rho^{\bullet} = 1$ can be tested against the alternative H_a : $\rho < 1$ using the critical values of t-statistic. If the null hypothesis cannot be rejected, the data generating process is said to have unit root and to be nonstationary. Therefore, the two-sided significance test performed is for the statistical significance of ρ -1 or ρ^{\bullet} -1 as the case may be. The test:

$$t = \frac{p-1}{se(a_i)} \tag{4}$$

The test resembles a t-test. The null hypothesis $H_0: \rho = 1$ is rejected if $|t| > a_1$, where the value of a_1 depends on the sample size and which other parameters are in the equation.

According to [3] under the null hypothesis of a unit root, this statistic does not follow the conventional t-distribution, and they derived asymptotic results and simulated critical values for various test and sample sizes. Thereafter, ([7], [8]) used larger set of simulations than those tabulated by Dickey and Fuller.

B. Detecting Integration Order I(d) - Differencing *Frequency*

One common way of detecting the number of times a non-stationary series $\{y_t\}$ can be differenced to become stationary most often is allowing d = 1, then test for unit root using any one of the several tests available in literature such as the Augmented Dickey-Fuller (ADF) test, Phillips-Perron (PP) test, Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) test and so on. If after first difference, { y, } is still non-stationary, the second differencing (d = 2) is expected to render the underlying series I(0). However, it is conventional practice that d cannot exceed 2. It has been aforementioned that the number of differencing needed to make a non-stationary series stationary is known as the integration order.

C. Detecting Integrated Order one I(1) – Presence of One Unit Root

However, the differencing frequency to render a nonstationary series I(0) can be visualized using auxiliary autoregressive (AAR) model. Consider a non-stationary series with lagged endogenous regression of order two with a constant term of the form:

$$y_{t} = \beta_{0} + \varphi t r + \beta_{1} y_{t-1} + \beta_{2} y_{t-2} + e_{t}$$
 (5)

where, φ is the coefficient of the trend parameter tr, β_0 is the constant term (intercept) and it is optional to include it. The e_t is the random error term. β_1 and β_2 are the autoregression coefficients. But if the level series $\{y_i\}$ is integrated order one I(1), then it is conditionally sufficient that the absolute value of β_1 must be greater than or equal to unity, that is, $|\beta_1| \ge 1$, β_2 must be strictly less than one, that is $|\beta_2| < 1$ and it is necessary that the ratio of $\frac{|\beta_1|}{|\beta_2|} > 1$. For a differenced $\{y_t\}$, $y_t - y_{t-1} = (1-L)y_t = \nabla y_t$, (5) is rewritten as:

$$\nabla y_{t} = \beta_{0} + \varphi t r + \beta_{1} \nabla y_{t-1} + \beta_{2} \nabla y_{t-2} + e_{t}$$
 (6)

If the level series is integrated order one I(1) after the first difference, the absolute values of β_1 and β_2 in (6) should be strictly less than unity. $|\beta_2| < 1$ and $|\beta_2| < 1$, respectively. And the random disturbance term e_t should be stationary.

D. Detecting Integrated Order One I(2) - Presence of 2 Unit Roots

Consider (4) again, if the level series $\{y_i\}$ is integrated order two I(2), then it is expected that $|\beta_1| > 1$ and the value of $|\beta_2| \ge 1$. Situation could arise where $|\beta_2| \ge 1$ or the value of $|\beta_2|$ is approaching unity. If such situation arises, increase the lag order of (5) to be of the form:

$$y_{t} = \beta_{0} + \varphi t r + \beta_{1} y_{t-1} + \beta_{2} y_{t-2} + \beta_{3} y_{t-3} + e_{t}$$
 (7)

For an integrated order two I(2) series $\{y_i\}$, it is conditionally sufficient that $|\beta_1| > 1$, $|\beta_2| \ge 1$ and $|\beta_3| < 1$. We can as well use (6) to ascertain whether a series $\{y_i\}$ is still non-stationary after first difference. Obtaining the absolute value of β_1 to be greater than or equal to unity, we say that $\{y_t\}$ is integrated order two I(2).

Generally, for clarity purpose, (7) should be used for order of integration test. The parametric boundary conditions for I(1) are as follows; $|\beta_1| \ge 1$, $|\beta_2| < 1$, $|\beta_3| < 1$

and
$$\frac{|\beta_1|}{|\beta_2|} > 1$$
 and for I(2) it is expected that

 $|\beta_1| > 1$, $|\beta_2| \ge 1$, $\frac{|\beta_1|}{|\beta_2|} > 1$ and $\frac{|\beta_2|}{|\beta_2|} \ge 1$. However, detecting the order of integration for non-stationary series is synonymous with detecting the number of unit roots present in the same series. If a non-stationary series is I(1), it means there is only one unit root in that series. Where a non-stationary series is I(2), it means the series has two unit roots.

E. Proposition 1

Consider an AAR(3) model as given in (7), if $\{y_t\}$ is I(1) and $\{\beta_2, \beta_3\} \in (-1, 1) \ni : \forall \beta_3 < \beta_2 \text{ then, it suffice that}$ $|\beta_1| \ge 1$.

Proof:

AAR(3) as presented in (7) can be re written as:

$$(1 - \beta_1 L - \beta_2 L^2 - \beta_3 L^3) y_t = \beta_0 + \varphi t r + e_t$$
 (7.1)

Let $\Phi = \beta_0 + \varphi tr$ be the deterministic term so that our focus can be on the polynomial of order 3. Then (7.1) becomes:

$$(1 - \beta_1 L - \beta_2 L^2 - \beta_3 L^3) y_t = \Phi + e_t \tag{7.2}$$

Equating the polynomial to zero, we have:

$$1 - \beta_1 L - \beta_2 L^2 - \beta_3 L^3 = 0 (7.3)$$

Equation (7.3) becomes:

$$\beta_1 L + \beta_2 L^2 + \beta_3 L^3 = 1 \tag{7.4}$$

For $\beta_2 = 0$ and $\beta_3 = 0$ in (7.4),

$$\beta_1 L = 1
\beta_1 = \frac{1}{L}$$
(7.5)

From (7.5), it implies that $|\beta_1| \ge 1$ iff $L \in [-1, 1]$. So, assuming that $-1 \le L \le 1$ then,

$$|\beta_1|$$
 = 1 if L = ±1
= 0 if L = 0
> 1 otherwise (7.6)

Note that L is the lag polynomial at order 1 and cannot be zero. Hence, $|\beta_1| \ge 1$ and without loss of generality, $|\beta_1| \in [1, \infty).$

Corollary1: For an I(1) series, given that $|\beta_1| \ge 1$ and $\{\beta_2, \beta_3\} \in (-1, 1) \ni : \forall \beta_3 < \beta_2, \text{ then } \left| \frac{\beta_1}{\beta_2} \right| > 1.$

Proof:

For any $\varepsilon \in (0, 1)$, given that $|\beta_1| \ge 1$ then,

$$\frac{\left|\beta_{1}-\varepsilon\right|}{\left|\beta_{2}-\varepsilon\right|} < \frac{\left|\beta_{1}-\varepsilon\right|^{2}}{\left|\beta_{2}-\varepsilon\right|^{2}} < \frac{\left|(\beta_{1}+\varepsilon)\otimes(\beta_{1}-\varepsilon)\right|}{\left|(\beta_{2}+\varepsilon)\otimes(\beta_{2}-\varepsilon)\right|} < \frac{\left|\beta_{1}+\varepsilon\right|^{2}}{\left|\beta_{2}+\varepsilon\right|^{2}} \le \frac{\left|\beta_{1}\right|}{\left|\beta_{2}\right|}$$

$$(7.7)$$

Given that β_2 , $\beta_3 \in (-1, 1)$ $\ni : \forall \beta_3 < \beta_2$, it suffice to say that,

$$1 \le e^{\left|\frac{\beta_2}{\beta_2}\right|} < \frac{\left|\beta_1\right|}{\left|\beta_2\right|} \tag{7.8}$$

If (7.8) is true, then $\frac{|\beta_1|}{|\beta_1|} > 1$.

F. Proposition 2

Consider an AAR(3) model as given in (7), if $\{y_i\}$ is I(2) and $|\beta_3| \in (0, 1)$ then it is sufficient that, $1 \le |\beta_2| < |\beta_1|$. **Proof:**

Consider (7.4), it can be shown that:

$$\beta_{2} = \frac{1 - \beta_{1}L - \beta_{3}L^{3}}{L^{2}} \Rightarrow |\beta_{2}| = \left| \frac{1 - \beta_{1}L - \beta_{3}L^{3}}{L^{2}} \right|$$
(7.7)

Since β_2 is the coefficient of L^2 (which invariably is the notation for lag 2). It is not confusing to let $L=\pm 2$ and given that $|\beta_1| > 1$. Then, at any admissible value of $\hat{\beta}_3 \in (0, 1)$ it can be shown that:

$$\left| \frac{1 - \beta_1 L - \beta_3 L^3}{L^2} \right| > 1$$

Hence, $|\beta_2| \ge 1$.

Corollary 2: For an I(2) series, If $1 \le |\beta_2| < |\beta_1|$ and $|\beta_3| \in (0, 1)$ it is necessary that $\frac{|\beta_2|}{|\beta_1|} > 1$.

Proof: Given that $1 \le |\beta_2| < |\beta_1|$ and $\tilde{\beta}_3 \in (0, 1)$, for any $\varepsilon \in (0, \frac{1}{2})$ then,

$$\frac{\left|\beta_{1}\right|}{\left|\beta_{2}\right|} \leq \frac{\left|\beta_{1}\right| - \epsilon}{\left|\beta_{2}\right| - \epsilon} < \frac{\left[\left|\beta_{1}\right| - \epsilon\right]^{2}}{\left[\left|\beta_{2}\right| - \epsilon\right]^{2}} < \frac{\left|\beta_{2}\right| - \epsilon}{\left\|\beta_{3}\right| - \epsilon\right|}$$

$$(7.8)$$

Since, $\frac{|\beta_1| - \epsilon}{\|\beta_2\| - \epsilon\|} > 1$, it can be verified therefore that,

$$\frac{\left|\beta_{2}\right|}{\left|\beta_{3}\right|} > e^{\varepsilon}.$$

Corollary 3: For an I(2) series, if $1 \le |\beta_2| < |\beta_1|$ and $|\beta_3| \in (0, 1)$ then, $|\beta_1|^2 - |\beta_2|^2 \ge |\beta_2|^2 - |\beta_3|^2 > 0$.

Proof: Given that $1 \le |\beta_2| < |\beta_1|$ and $\tilde{\beta}_3 \in (0, 1)$, for any $\varepsilon \in (0, \frac{1}{2})$ then it is sufficient to show that the $\frac{\left|\beta_1^2 - \beta_2^2\right|}{\left|\beta_2^2 - \beta_2^2\right|} > 1$. That is,

$$\frac{\left|\beta_1^2 - \beta_2^2 + \varepsilon^2\right|}{\left|\beta_2^2 - \beta_3^2 + \varepsilon^2\right|} \ge \frac{\left|\beta_1^2 - \beta_2^2\right| + \varepsilon^2}{\left|\beta_2^2 - \beta_3^2\right| + \varepsilon^2}.$$

$$\geq \frac{\left|\beta_{1}\right|^{2} - \left|\beta_{2}\right|^{2} + \varepsilon^{2}}{\left|\beta_{2}\right|^{2} - \left|\beta_{3}\right|^{2} + \varepsilon^{2}}$$

$$\geq e^{\frac{\varepsilon}{4}} \tag{7.9}$$

However, since
$$\frac{|\beta_{1}|^{2} - |\beta_{2}|^{2}}{|\beta_{2}|^{2} - |\beta_{3}|^{2}} \ge \frac{|\beta_{1}|^{2} - |\beta_{2}|^{2} + \varepsilon^{2}}{|\beta_{2}|^{2} - |\beta_{3}|^{2} + \varepsilon^{2}}, \quad \text{it}$$
 suffices to say that
$$\frac{|\beta_{1}|^{2} - |\beta_{2}|^{2}}{|\beta_{2}|^{2} - |\beta_{3}|^{2}} \ge e^{\frac{\varepsilon}{4}}.$$

Note that corollary 3 can be proved using triangular inequalities as follows.

Since $|\beta_1|^2 - |\beta_2|^2 \ge |\beta_2|^2 - |\beta_3|^2 \Rightarrow |\beta_1^2 - \beta_2^2| \ge |\beta_2^2 - \beta_3^2|$, it adequate to show that $\left\lceil \beta_2^2 - \beta_3^2 \right\rceil > 0$. So based on triangular inequalities $\left|\beta_2^2 - \beta_3^2\right| \le \left|\beta_2^2 - \beta^2\right| + \left|\beta^2 - \beta_3^2\right|$.

Taking the probability and expectation of both sides of the inequality and for any $\varepsilon > 0$.

$$p\left\{E\left[\left|\beta_{2}^{2}-\beta_{3}^{2}\right|>\varepsilon\right]\right\} \leq p\left\{E\left[\left|\beta_{2}^{2}-\beta^{2}\right|>\frac{\varepsilon}{2}\right]\right\} + p\left\{E\left[\left|\beta^{2}-\beta_{3}^{2}\right|>\frac{\varepsilon}{2}\right]\right\} \to 1$$
And thus,
$$\lim_{\varepsilon \to 0} p\left\{E\left[\left|\beta_{2}^{2}-\beta_{3}^{2}\right|>\varepsilon\right]\right\} = 1$$

Alternative way of proving proposition 2 is by differencing the non-stationary series $\{y_t\}$ under study once so that (7.1) can be rewritten as:

$$(1 - \beta_1 L - \beta_2 L^2 - \beta_3 L^3) \nabla y_t = \beta_0 + \varphi t r + e_t$$
 (7.10)

Then simply repeat proposition 1 and Corollary1, respectively.

G. Order of Integration Test Hypotheses

In order to test the null hypothesis that the integration order of the variable y is equal to zero against the one sided alternative that it is greater than or equal to unity. Using (6), we state the following hypotheses:

$$H_{01}: \beta_1 < 1$$
 (it is integrated order zero) testing for I(1) $H_{11}: \beta_1 \ge 1$ (it is integrated order one) testing for I(1) $H_{02}: \beta_2 < 1$ (it is integrated order zero) $H_{12}: \beta_2 \ge 1$ (it is integrated order two)

The test statistic for the above hypotheses is given as:

$$t_j = \frac{\hat{\beta}_j}{s.e(\beta_j)} \tag{8}$$

which has a Student's t-distribution with (n - p - 1)degrees of freedom. The test is carried out by comparing the observed value with the appropriate critical value $t_{(n-p-1,\alpha)}$, which is obtained from the t-table. Where α is the significance level. H_0 is to be rejected at the significance α level if:

$$\left|t_{j}\right| \geq t_{\text{(n-p-1},\alpha)} \tag{9}$$

where $p(|t_j|)$ is the **p-value** of the test, which is the probability that a random variable having a Student tdistribution, with (n - p - 1) is greater than $|t_i|$ (the absolute value of the observed value of the t-test). A criterion similar to that in (8) is to compare the p-value of the test with α and reject H₀ if this probability value is less than the size of the test, say 0.05, we reject the null hypothesis.

H. Method of Estimation Using Generalized Least Squares The method of estimation adopted is the generalized least squares method and it is presented briefly as follows:

(7) can be rewritten in matrix form as:

$$y = Y^*\beta + e$$
 (10)
where the Y^* is the vector of lagged dependent variables.
The $E(e) = 0$ and $var-cov(e) = \sigma^2 V$. And where σ^2 is

known, V represents the assumed structure of variances and covariances among the random errors e_t . Under the given condition of the variance-covariance of the error terms e, β 's can be estimates using:

$$\beta = \left(Y^{\bullet'}V^{-1}Y^{\bullet}\right)^{-1}Y^{\bullet'}V^{-1}y \tag{11}$$

where $\operatorname{var-cov}(\beta) = \sigma^2 \left(Y'^{\bullet} V^{-1} Y^{\bullet} \right)^{-1}$. If it is assumed that the variance σ^2 of each error term is the same over time and the error terms are mutually uncorrelated, then the V matrix reduces to identity matrix. And if there is heteroscedasticity as well as autocorrelation, then the V matrix will have entries on the main diagonal as well as on the off diagonal.

I. Unit Root Test

Unit root test is a conventional test for ascertaining the order of integration of any underlying time series data or unstructured data. In other words, unit root test analysis specifies whether any given data structure is stationary or nonstationary.

There are lots of unit root tests developed over time, for the literatures and technical details, see [1], [3], [4], [6], [10]-[13]. These unit root test sometimes are divergent and contradict each other on the order of integration of some data series, most especially logarithm transformed data

Explicit example where results of unit root test on a time series data as regards these unit root tests will be presented in section III.

III. DATA ANALYSIS AND RESULTS

The two data sets used in this study include simulated sample of 2000 observations and an empirical data on service sector output obtain from [2] statistical bulletin consisting of 39 observations spanning from 1981 to 2019. The empirical data on services sector output used is known to be integrated order two I(2).

A. Integration of Order One I(1) Analysis Using Simulated

The simulated data with drift is analysed for order of integration and the result is presented below:

$$y_{1t} = 0.3369 + 0.0006 + 1.3022 y_{1t-1} + 0.0135 y_{1t-2} - 0.3154 y_{1t-3} + e_t$$

prob. (0.0000) (0.1495) (0.0000) (0.7078) (0.0000) (12)

The result in (12) indicates that y_{1t} is I(1) since $|\beta_1| = 1.3022 > 1$, $|\beta_2| = 0.0135 < 1$, $|\beta_3| = |-0.3154| < 1$ and $\frac{|\beta_1|}{|\beta_2|} = \frac{|1.3022|}{|0.0135|} = 96.4592 > 1$ and the p-value (0.0000) of β_1 is significant under 5% level as expected. This result reveals that there is only one unit root present in the simulated series y_{1t} . Therefore, first differencing can be used to achieve stationarity of y_{1t} . It was stated earlier that β_0 (the constant term) is optional. Not including the constant term does not affect the integration order of the variable under study. The result without the constant term is presented in (13) below:

$$\begin{aligned} y_{1t} &= 0.0019 + 1.4049 y_{1t-1} + 0.0140 y_{1t-2} - 0.4208 y_{1t-3} + e_t \\ \text{prob.} \quad & (0.0000) \quad & (0.0000) \quad & (0.7082) \quad & (0.0000) \end{aligned}$$

This result also confirms the integration order of y_t . $|\beta_1| = 1.4049 > 1$, $|\beta_2| = 0.0140 < 1$, Giving that $|\beta_3| = |-0.4208| < 1$ and $\frac{|\beta_1|}{|\beta_2|} = \frac{|1.4049|}{|0.0140|} = 100.35 > 1$. And since the *p-value* (0.0000) of β_1 is significant under 5% level, the hypothesis that (H_{01} : $\beta_1 < 1$) the variable is I(0) is rejected in favour of the alternative. Hence, the series y_{1t} is I(1).

B. Integration of order two I(2) Analysis using Empirical Data.

The analysis of detecting I(2) or the presence of two unit roots in a set of data is demonstrated using service sector output data from CBN and the results is presented in (14)

$$y_{2t} = -684.9050 + 69.4891 + 1.8847 y_{2t-1}$$

prob. (0.1276) (0.0239) (0.0000)
 $-1.2293 y_{2t-2} + 0.3586 y_{2t-3} + e_t$ (14)
(0.0029) (0.1292)

The regression result in (14) reveals that the trend parameter value (69.4891) is significant under 5% level. The coefficients $|\beta_1| = 1.8847 > 1$, $|\beta_2| = |-1.2293| > 1$, $\left|\beta_{3}\right| = 0.3586 < 1,$ $\frac{\left|\beta_{1}\right|}{\left|\beta_{2}\right|} = \frac{\left|1.8847\right|}{\left|-1.2293\right|} = 1.5331 > 1$ $\frac{|\beta_2|}{|\beta_1|} = \frac{|-1.2293|}{|0.3586|} = 3.4281 > 1$. The values of β_1 and β_2 are significant under 1% and 5% levels, respectively. This result entails that y_{2t} is I(2). Conventionally, y_{2t} can be stationary should the second differencing be applied. So, there are two unit roots present in service sector output data. Omitting the constant term in (14), the result becomes as giving in (15) below:

$$y_{2t} = 27.6504 + 1.9532 y_{2t-1} - 1.2736 y_{2t-2} + 0.3436 y_{2t-3} + e_t$$

prob. (0.0296) (0.0000) (0.0025) (0.1536) (15)

The choice to omit the β_0 as shown in (15) does not change the order of integration or frequency of unit root present in y_{2t} . The trend parameter is also significant under 5% level. Nevertheless, It is clearer that $|\beta_1| = 1.9532 > 1$, $|\beta_2| = |-1.2736| > 1$, $|\beta_3| = 0.3436 < 1$, $\frac{|\beta_1|}{|\beta_2|} = \frac{|1.9532|}{|-1.2736|} = 1.5336 > 1$ and $\frac{|\beta_2|}{|\beta_3|} = \frac{|-1.2736|}{|0.3436|} = 3.7066 > 1$. The values of β_1 and β_2 are significant under 1% and 5% level,

respectively. This result indicates that y_{2t} in the above (14) and (15) is I(2).

C. Unit Root Test Analysis

Here, few selected unit root tests results such as Augmented Dickey Fuller (ADF) test, Phillips-Perron (PP) Test; Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) Test; Elliot, Rothenberg, and Stock Point Optimal (ERS) Test and Ng and Perron (NP) Tests are being applied to compare the findings made using the new proposed AA(3) method of detecting integration order (frequency of unit root) in a data series.

The results of the ADF, PP, KPSS, ERS and NP unit root test in Table I - Table III indicate that the simulated series y_1 is integrated order one I(1) at the level series and I(0) or stationary at first difference and it is significant under 5% level. The empirical data on services sector output y_2 is integrated order two I(2) and stationary at the second difference. These results implied that simulated series y_1 has one unit root and y_2 has two unit roots.

TABLE I: ANALYSIS OF UNIT ROOT TEST USING ADF AND P	P
---	---

Test	Series	DT	Lags /BW	Test Value	1% level 5%level 10%level	Prob.	Rmk
	y_1	C,T	22	-3.1697	-3.9627 -3.4121 -3.1280	0.0909	I(1)
	∇y_1	C,T	20	-5.0720	-3.9627 -3.4121 -3.1279	0.0001	I(0) at 1 st diff.
ADF	y_2	C,T	1	-0.0104	-4.2268 -3.5366 -3.2003	0.9946	I(1)
	∇y_2	C,T	1	-2.9978	-4.2350 -3.5403 -3.2024	0.1468	I(2) at 1 st diff
	$\nabla^2 y_2$	C	4	-5.1682	-3.6268 -2.9458 -2.6115	0.0001	I(0) at 2 nd diff.
	y_1	C,T	22	-3.1697	-3.9627 -3.4121 -3.1279	0.0909	I(1)
	∇y_1	C,T	20	-5.0720	-3.9627 -3.4121 -3.1279	0.0001	I(0) at 1 st diff.
PP	y_2	C,T	3	1.1837	-4.2191 -3.5331 -3.1983	0.9999	I(1)
	∇y_2	C,T	4	-2.5684	-4.2268 -3.5366 -3.2003	0.2960	I(2) at 1st diff
	$\nabla^2 y_2$	C	15	-5.2650	-3.6268 -2.9458 -2.6115	0.0001	I(0) at 2 nd diff.

TABI	LE II: ANA	LYSIS O	F UNIT	ROOT TEST U	JSING KPSS	AND ERS
	y_1	C,T	34	0.6199	0.2160 0.1460	I(1) I(0)
	21	C,1	54	0.01))	0.1190	at 1st
					0.2160	diff
	∇y_1	C,T	30	0.1150	0.1460	I(1)
	J 1				0.1190	under 5%
					0.2160	5% I(1)
KPSS	y_2	C,T	5	0.1965	0.1460	under
	J Z	-,-			0.1190	5%
					0.2160	I(2)
	∇y_2	C,T	4	0.1464	0.1460	at 1st
					0.1190	diff.
					0.7390	I(0)
	$\nabla^2 y_2$	C	0	0.2033	0.4630	at 2 nd
	_				0.3470	diff
					3.9600	
	\boldsymbol{y}_1	C,T	7	19.5312	5.6200	I(1)
					6.8900	
	∇				3.9600	I(0)
	∇y_1	C,T	7	1.1760	5.6200	At 1st
					6.8900	Diff
EDC	.,	ОТ		150.1254	4.2200	T/1)
ERS	y_2	C,T	1		5.7200	I(1)
					6.7700	1(2)
	∇y_2	C,T	1	6.2547	4.2200 5.7200	I(2) at 1 st
	v y ₂	C,1	1	0.2347	6.7700	diff
					1.8700	I(0)
	$\nabla^2 y_2$	C	0	2.2252	2.9700	at 2 nd
	. 72	Ü	•		3.9100	diff

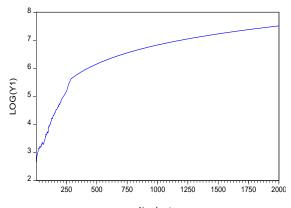
Series	D T	1	$\alpha\%$	Mza	MZt	MSB	MPT	Rmk		
				-4.6959	-1.5323	0.3263	19.405			
	C, T	7	1%	-23.800	-3.4200	0.1430	4.0300	T(1)		
y_1	T	/	5%	-17.300	-2.9100	0.1680	5.4800	I(1)		
			10%	-14.200	-2.6200	0.1850	6.6700			
				-63.705	-5.6402	0.0885	1.4473			
∇y_1	C, T	7	1%	-23.800	-3.4200	0.1430	4.0300	T(0)		
$\mathbf{v} y_1$	T	T	T	/	5%	-17.300	-2.9100	0.1680	5.4800	I(0)
			10%	-14.200	-2.6200	0.1850	6.6700			
			1%	1.4534	1.4554	1.0014	238.06			
.,	C, T	6	6		-23.800	-3.4200	0.1430	4.0300	T(1)	
y_2	T			0	0	5% 10%	-17.300	-2.9100	0.1680	5.4800
			1070	-14.200	-2.6200	0.1850	6.6700			
			1%	-9.0238	-2.1180	0.2347	10.121			
∇y_2	C, T	<u>;</u> 0	5%	-23.800	-3.4200	0.1430	4.0300	T(1)		
v y 2	T		5% 10%	-17.300	-2.9100	0.1680	5.4800	I(1)		
					1070	-14.200	-2.6200	0.1850	6.6700	
$\nabla^2 y_2$	C	C 0	1%	-17.919	-2.7531	0.1536	2.2057	I(0)		
				-13.800	-2.5800	0.1740	1.7800	. ,		
v y ₂			5%	-8.1000	-1.9800	0.2330	3.1700	at		
			10%	-5.7000	-1.6200	0.2750	4.4500	5%		

(service sector output) The symbol ∇ represents first difference and ∇^2 indicates second difference.

D. Behaviour of AAR(3) Model in Log Transformed Series

In this section, pictorial representation of logarithm of and the simulated data y_1 empirical data (service sector output) y_2 are presented in Fig. 1 below.

The time series plots in Fig. 1 and 2 exhibit presences of drift in the log transformation of y_1 and y_2 and as such both series are non-stationary. The plots in Figure 3 and 4 exhibit a stationary outlook as the first log differencing transformation has eliminated the presence of drift in both variable y_1 and y_2 .



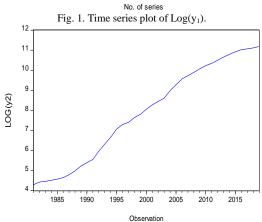


Fig. 2. Time series plot of $Log(y_2)$.

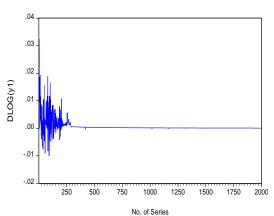


Fig. 3. Time series plot of first log difference of y₁.

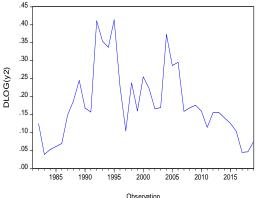


Fig. 4. Time series plot of first log difference of y2.

E. Order of Integration Test Using AAR(3) Model in the Log Transformed Series

The results of the order of integration test for y_1 and y_2 via AAR(3) model are in Eq.(16) - (19) below:

$$\begin{split} \log(y_{\rm lr}) &= 0.0197 + 1.87E - 06 + 1.0886 \log(y_{\rm lr-1}) \\ \text{prob.} & (0.0000) \ \ \, (0.0000) \ \ \, (0.0000) \\ & + 0.0815 \log(y_{\rm lr-2}) - 0.1732 \log(y_{\rm lr-3}) + e_{\rm r} \\ & (0.0131) \ \ \, (0.0000) \end{split}$$

The result in (12) indicates that the trend is significant and y_{1t} is I(1) since $|\beta_1| = 1.0886 > 1$, $|\beta_2| = 0.0815 < 1$, $\left|\beta_3\right| = \left|-0.1732\right| < 1$ and $\frac{\left|\beta_1\right|}{\left|\beta_2\right|} = \frac{\left|1.0886\right|}{\left|0.0815\right|} = 13.3571 > 1$ and the p-value (0.0000) of β_1 is significant under 5% level. This result reveals that there is only one unit root present in the simulated series $log(y_1)$ Therefore, first differencing can be used to achieve stationarity of $log(y_{1t})$. It was stated earlier that β_0 (the constant term) is optional. Not including the constant term does not affect the integration order of the variable under study. When there is no constant term, the result becomes:

$$\begin{split} \log(y_{1t}) &= -2.58\text{E}\text{-}06 + 1.1988\log(y_{1t-1}) \\ \text{prob.} & (0.0000) \quad (0.0000) \\ & + 0.0828\log(y_{1t-2}) - 0.2810\log(y_{1t-3}) + e_t \\ & (0.0159) \quad (0.0000) \end{split}$$

This result also confirms the integration order of $log(y_{1})$. $|\beta_1| = 1.1988 > 1, |\beta_2| = 0.0828 < 1,$ Giving that $|\beta_3| = |-0.2810| < 1$ and $\frac{|\beta_1|}{|\beta_2|} = \frac{|1.1988|}{|0.0828|} = 14.4783 > 1$ the series $log(y_{1})$ is I(1). And since the *p-value* (0.0000) is significant under 5% level, the hypothesis that (H_{01} : β_1 < 1) the $log(y_{lt})$ is I(0) is rejected in favour of the alternative. Hence, $\log(y_1)$ is I(1).

The result of order of integration test for $\log \log y_2$ is presented below:

$$\log(y_{2t}) = 0.3096 + 0.0126 + 1.5158 \log(y_{2t-1})$$
prob. (0.0985) (0.3136) (0.0000)
$$-0.4340 \log(y_{2t-2}) - 0.1481 \log(y_{2t-3}) + e_t$$
(0.1712) (0.4209)

The regression result in Equation (18) reveals that $|\beta_1| = 1.5158 > 1$, $|\beta_2| = |-0.4340| < 1$, $\left|\beta_3\right| = 0.1481 < 1 \quad \frac{\left|\beta_1\right|}{\left|\beta_2\right|} = \frac{\left|1.5158\right|}{\left|-0.1481\right|} = 10.2350 > 1$. The values of β_1 is significant under 5% level. This result entails that $\log(y_{2t})$ is I(1). Conventionally, $\log(y_{2t})$ can be stationary should the first logarithm differencing be applied. So, there is one unit root present in service sector output data. Omitting the constant term in (18), the result becomes as giving in (19) below:

$$\begin{split} \log(y_{2t}) &= -0.0075 + 1.6126 \log(y_{2t-1}) \\ \text{prob.} & (0.0663) \quad (0.0000) \\ & -0.5234 \log(y_{2t-2}) - 0.0609 \log(y_{2t-3}) + e_t \\ & (0.1057) & (0.7361) \end{split}$$

The choice to omit the β_0 as shown in (19) does not change the order of integration present in $log(y_{2t})$ or frequency of unit root present in the data. It is clearer that $|\beta_1| = 1.6126 > 1$, $|\beta_2| = |-0.5234| < 1$, $|\beta_3| = -0.0609 < 1$

and $\frac{|\beta_1|}{|\beta_2|} = \frac{|1.6126|}{|-0.5234|} = 3.0810 > 1$. The value of β_1 is significant under 5% level. This result indicates that $log(y_{2t})$ in (18) and (19) is I(1).

It is good to note that variation in y_{2t} has been abridged by it logarithm transformation.

F. Unit Root Tests for the Log Transformed Variables

The section presents several unit root tests for the log transformed variables and the results are shown in Table IV and Table V below.

Test	Series	DT	1 200		1% level		
	~	DT	Lags	Test	5%level	Prob.	Rmk
		2.	/BW	Value	10%level	1100.	
					-3.9627		
	$log(y_1)$	C, T	22	-5.8803	-3.4121	0.000	I(0)
	-801/	T	22	3.0003	-3.1280	0.000	1(0)
					-4.2191		
	$\log(y_2)$	C, T	0	-0.0867	-3.5331	0.993	I(1)
	108(72)	T	U	-0.0007	-3.1983	0.773	1(1)
ADF					-4.2268		
	$\nabla \log(y_2)$	C, T	0	-2.7117	-3.5366	0.238	I(2)
	106(32)	T	U	-2./11/	-3.2003	0.236	1(2)
					-3.6268		
	$\nabla^2 \log(y_2)$	C	0	7 5555		0.000	I(0)
	$\mathbf{v} = \log(y_2)$	C	0	-7.5555	-2.9458	0.000	at 1%
					-2.6115		
	1()	C.		10 1505	-3.9627	0.000	7(0)
	$\log(y_1)$	C, T	22	-12.4535	-3.4121	0.000	I(0)
					-3.1279		
	1(C,		4 0	-4.2191	0.000	
	$\log(y_2)$	T,	4	-1.2576	-3.5331	0.883	I(1)
PP		•			-3.1983		
11		C,			-4.2268		
	$\nabla \log(y_2)$	T,	2	-2.5756	-3.5366	0.293	I(2)
		1			-3.2003		
					-3.6268		I(0)
	$\nabla^2 \log(y_2)$	C	3	-7.9043	-2.9458	0.000	at 1%
					-2.6115		at 170
		C			0.2160		
	$\log(y_1)$	C,	34	0.9989	0.1460		I(1)
		T			0.1190		
					0.2160		
	$\nabla \log(y_1)$	C,	26	0.9714	0.1460		I(2)
T/DGG		T			0.1190		. ,
KPSS					0.7390		T(0)
	$\nabla^2 \log(y_1)$	C	131	0.1758	0.4630		I(0)
	- w 1				0.3470		at 1%
		~			0.2160		I(0)
	$\log(y_2)$	C,	4	0.1383	0.1460		under
	÷ :- 2/	T	-		0.1190		5%
		~			3.9600		
	$\log(y_1)$	C, T	22	2201.725	5.6200		I(1)
	0.01/	T			6.8900		-(1)
		_			3.9600		
	$\nabla \log(y_1)$	C,	7	2.4857	5.6200		I(0)
	90017	T	,	2.4031	6.8900		1(0)
					4.2200		
ERS	$\log(y_2)$	C,	1	54.3193	5.7200		I(1)
LIG	~ B(/2/	T	1	JT.J17J	6.7700		1(1)
					4.2200		
	$\nabla \log(y_2)$	C,T	0	8.8847	5.7200		1(2)
	V 105(y ₂)	C, I	U	0.004/			I(2)
					6.7700		
	$\nabla^2 \log(\cdots)$	~		2.0400	1.8700		T/0\
	$\nabla^2 \log(y_2)$	C	0	2.0408	2.9700 3.9100		I(0) at 5%

Note that 'D,T' represents deterministic term and 'Rmt' represents remarks.

The unit root results of the logarithm transformed variables y₁ and y₂ as presented in Table II. The ADF and PP unit root tests indicate that $log(y_1)$ is I(0) and $log(y_2)$ is I(2) and both are significant under 5% level. KPSS and NP report that $log(y_1)$ is I(2) significant at 5% and 10%

respectively. For KPSS, log(y₂) is I(0) significant under 5% level while NP reports that $log(y_2)$ is I(2) significant under 5% level. For ERS, Log(y₁) is I(1) significant under 5% level and $log(y_2)$ is I(2) significant under 5% level.

G. Discussion of Results

The results of the introduced order of integration test obtained in (12) and (13) revealed that the simulated data y_1 variable is I(1) or has one unit root. And the test results in (14) and (15) indicated that y_2 variable is I(2) or has two unit roots. These findings agree with the unit root test results of [3], [11], [6], [4], [1] and [10] as shown in Table I.

In the case of the logarithm transformed variables $log(y_1)$ and log(y2) as presented in Table IV and Table V. The introduced order of integration test using AAR(3) model in (16) and (17) revealed that $Log(y_1)$ is I(1). This finding agrees with that of ERS but disagrees with that of ADF and PP unit root tests that reported that $log(y_1)$ is I(0) and that of KPSS and NP whose results indicated that $Log(y_1)$ is I(2). For $log(y_2)$, AAR(3) results in (18) and (19) revealed that $Log(y_2)$ is I(1). This finding is contrary to that of ADF, PP, NP, ERS that reported I(2) and KPSS whose result indicted that $Log(y_2)$ is I(0).

The pictorial representation of $log(y_1)$ reveals more or less like S-shaped curve trend pattern and unlikely to be I(0) or stationary as claimed by [3], [11] and neither can it be I(2) as claimed by [6]. And the plot of first difference of natural logarithm of y₁ as shown in Fig. 1 reveals that $\nabla \log(y_1)$ is stationary or I(0), this confirms the results as presented in (16) and (17) that logarithm transformation of y_1 is I(1). This conspicuous reason implies that the introduced AAR(3) method of order of integration test performed better than other unit roots as exemplified in this paper.

IV. CONCLUSION

The paper presents a simple procedure for order of integration test that can be used as an alternative to unit root test. This procedure uses auxiliary autoregressive AAR(3) model framework to detect the order of integration for any given non-stationary variable. The generalized least square (GLS) was used to estimate the model parameters and hypotheses testing were based on t-statistic.

The order of integration test was applied to a simulated non-stationary series $\{y_{1t}\}$ with a sample size of 2000 and the result indicates that y_{1t} is integrated order one I(1). The test procedure was also applied to a known non-stationary variable y_{2t} with two unit roots and the test detected the presence of two unit roots, in other words, y_{2t} is integrated

The uniqueness of the test procedure lies in its simplicity within the univariate regression framework. The test procedure is apt for testing unit root and can compete favourably with any conventional unit root test.

REFERENCES

- Bhargava, A. On the theory of testing for unit roots in observed time series. Review of Economic Studies. 1986; 53, 369–384. Available: https://doi.org/10.2307/2297634.
- Central Bank of Nigeria. Statistical bulletin. 2020. Available:Central Bank of Nigeria | Annual Statistical Bulletin (cbn.gov.ng).
- Dickey, D.A. and W.A. Fuller. "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," Journal of the American Statistical Association, 1979; 74, 427-431. Available: https://doi.org/10.2307/2286348.
- Elliott, G., Rothenberg, T.J. and Stock, J.H. Efficient test for an autoregressive unit root. Econometrica, 1996; 64(4): 813-836. Available: http://dx.doi.org/10.2307/2171846.
- Harris, R and Sollis, R., Applied Time Series Modelling and Forecasting, Wiley and Sons Ltd, 2003.
- Kwiatkowski, D., P.C.B. Phillips, P. Schmidt, Y. Shin. Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root, Journal of Econometrics, 1992; 54, pp. 159-178, Available: http: DOI: 10.12691/ijefm-4-2-1.
- MacKinnon, James G. "Critical Values for Cointegration Tests," Chapter 13 in R. F. Engle and C. W. J. Granger (eds.), Long-run Economic Relationships: Readings in Cointegration, Oxford: Oxford University Press, 1991.
- MacKinnon, James G. "Numerical Distribution Functions for Unit Root and Cointegration Tests," Journal of Applied Econometrics, 1996; 11, 601-618. Available: DOI: 10.12691/ajams-4-5-1.
- Nelson, C.R., Plosser, C.I. Trends and random walks in macroeconomic time series. Journal of Monetary Economics, 1982; 10. 139-162. Available: http://dx.doi.org/10.1016/0304-3932(82)90012-5.
- [10] Ng, S. and Perron, P. Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power. Econometrica, 2001; 69, $1519\text{-}1554.\ http://dx.doi.org/10.1111/1468-0262.00256.$
- [11] Phillips, P.C.B., P. Perron. Testing for a Unit Root in Time Series Regression, Biometrika, 1988; 75, pp. 335- 346. Avilable: DOI: 10.12691/jfe-4-2-3.
- [12] Said, S.E. and Dickey, D.A. (1984) Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order. Biometrika, 599-607. http://dx.doi.org/10.1093/biomet/71.3.599.
- [13] Schwert, G.W. Tests for unit roots: A Monte Carlo investigation. Journal of Business and Economic Statistics, 1989; 7, 147-159. Available: https://doi.org/10.2307/1391432.



Amaefula Chibuzo Gabriel is from Umuoloo-Umuokanne in Ohaji-Egbema LGA of Imo state, Nigeria. He had B. Sc degree in statistics (1998) from Usmanu Danfodiyo University Sokoto, M. Sc degree in statistics(2002) from University of Ibadan and Ph. D in statistics(2014) from Usmanu Danfodiyo University Sokoto, all in Nigeria.

He was a lecturer at the Federal Polytechnic Kaura-Namoda, Zamfara state, Nigeria (2008-2015), the head of Department of Mathematics and

Statistics (2013-2015) and the Director of School of Information and Communication Technology (SICT) from 2013-2015. Presently, he is a senior lecturer in the Department of Mathematics and Statistics, Federal University Otuoke, Bayelsa state, Nigeria. He has published many articles in the field of time series and econometrics both in national and international Journals. Google Amaefula C.G for some information concerning previous publications.

Dr Amaefula C.G is a member of Nigerian Statistical Association (NSA) and International Research and Development Institute (IRDI).