Journal of Data Analysis and Information Processing, 2023, 11, 144-174
https://www.scirp.org/journal/jdaip
ISSN Online: 2327-7203
ISSN Print: 2327-7211

A Sufficient Statistical Test for Dynamic Stability

Muhammad Ashfaq Ahmed, Nasreen Nawaz

Federal Board of Revenue, Islamabad, Pakistan
Email: muhammad.ashfaq@fbr.gov.pk, nawaznas@msu.edu

How to cite this paper: Ahmed, M.A. and Nawaz, N. (2023) A Sufficient Statistical Test for Dynamic Stability. *Journal of Data Analysis and Information Processing*, **11**, 144-174.
https://doi.org/10.4236/jdaip.2023.112009

Received: March 3, 2023
Accepted: May 2, 2023
Published: May 5, 2023

Abstract

In the existing Statistics and Econometrics literature, there does not exist a statistical test which may test for all kinds of roots of the characteristic polynomial leading to an unstable dynamic response, *i.e.*, positive and negative real unit roots, complex unit roots and the roots lying inside the unit circle. This paper develops a test which is sufficient to prove dynamic stability (in the context of roots of the characteristic polynomial) of a univariate as well as a multivariate time series without having a structural break. It covers all roots (positive and negative real unit roots, complex unit roots and the roots inside the unit circle whether single or multiple) which may lead to an unstable dynamic response. Furthermore, it also indicates the number of roots causing instability in the time series. The test is much simpler in its application as compared to the existing tests as the series is strictly stationary under the null (C01, C12).

Keywords

Dynamic Stability, Real and Complex Roots, Unit Circle

1. Introduction

A univariate time series that can be written in the form

$$y_t = \mu + \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \cdots = \mu + \psi(L)\varepsilon_t,$$

with $\sum_{j=0}^{\infty} |\psi_j| < \infty$, roots of $\psi(\mathbb{Z}) = 0$ outside the unit circle, and $\{\varepsilon_t\}$ a white noise process with mean zero and variance σ^2 is a covariance stationary time series with 1) $E(y_t) = \mu$ and 2) $cov(y_t, y_{t-j}) = \gamma_j$, for all t where j is an integer. A process is said to be strictly stationary if the joint distribution of $(Y_t, Y_{t+j_1}, Y_{t+j_2}, \cdots, Y_{t+j_n})$ depends only on the intervals separating the dates, *i.e.*, $j_1, j_2, \cdots, j_n$ and not on the date itself, *i.e.*, t. The most important feature of a stable dynamic response which distinguishes it from an unstable response is its convergence property, *i.e.*, the future forecasts $\hat{y}_{t+s|t} \equiv \hat{E}(y_{t+s} | y_t, y_{t-1}, \cdots)$ converge to the unconditional mean, *i.e.*,

$$\lim_{s\to\infty} \hat{y}_{t+s|t} = \mu.$$

Thus if someone is trying to forecast a series farther into the future, it becomes very important to know whether the series is dynamically stable or not. The most common approach in time series literature is testing for a real positive unit root. A number of tests have been proposed, including [1]-[27] and [28].

The use of CUSUM for the stability (or stationarity) of regression equations was introduced by [29]. The CUSUM test which is based on the residuals from the recursive estimates is normally used to test the parameter change in time series model. A number of tests for testing the structural change are available in the existing literature, such as [30] [31] [32], etc. There has been some work related to complex unit roots as well, e.g., [33], etc.

However, if the null hypothesis of a unit root is rejected, the time series can still be unstable due to presence of other kinds of roots leading to instability as unit root is not the sole root as a cause of concern regarding stability. If the null hypothesis of unit root is rejected, forecast/prediction of time series can still be seriously flawed unless there is a test available to guarantee that the series is dynamically stable.

In the existing literature no test has been proposed which may test for all kinds of roots leading to an unstable response, *i.e.*, real as well as complex unit roots along with the roots (of $\psi(\mathbb{Z}) = 0$) which lie inside the unit circle.

This paper develops a test which is sufficient to prove the dynamic stability (in the context of roots of the characteristic polynomial) of a univariate as well as a multivariate time series without having a structural break. It covers all roots (positive and negative real unit roots, complex unit roots and roots inside the unit circle whether single or multiple) which may lead to an unstable dynamic response. Furthermore, it also indicates the number of roots causing instability in the time series. The test is much simpler in its application as compared to the existing tests as the series is strictly stationary under the null.

The remainder of this paper is organized as follows: Section 2 provides the background of the test explaining the Routh's stability criterion, formation of the Routh array, the theorems of the Routh test and the bilinear transformation. Section 3 discusses the hypothesis testing. Section 4 explains the methodology regarding the constrained minimization with inequality constraints. Section 5 consists of the theorems regarding the distributions under the null for various kinds of stability tests. Section 6 provides the power and size performance of the test. Section 7 consists of a Monte Carlo study. Section 8 provides an empirical application of the test. Section 9 comprises of the conclusion and finally the proofs of the theorems and the derivation of the null hypothesis for a VAR is provided in the appendix.

2. Background

The main idea behind the test is to expolit the Routh-Hurwitz stability criterion which is a mathematical test that provides a necessary and sufficient condition for the stability of a linear time invariant (LTI) control system. The Routh test was proposed by an English mathematician Edward John Routh in 1876 which is an efficient recursive algorithm to determine whether all the roots of the characteristic polynomial of a linear system have negative real parts.

German mathematician Adolf Hurwitz arranged the coefficients of the polynomial into a square matrix in 1895, and showed that the polynomial is stable if and only if the sequence of determinants of its principal submatrices is all positive. These two procedures are exactly equivalent. Routh stability criterion provides a more efficient way to compute the Hurwitz determinants than computing them directly.

For discrete systems, the corresponding stability test can be handled through the bilinear transformation, the Jury test or the Bistritz test which are all equivalent, however the bilinear transformation is much simpler in its use.

[34] compares graphically, using the Arnold Tongues, some sufficient criteria for the stability of periodic differential equations.

For using these stability criteria, the parameters need to be estimated for which the procedure has been described in the later part.

2.1. Routh's Stability Criterion

The Routh test is a purely algebraic method for determining how many roots of the characteristic equation have positive real parts; from this it can also be determined whether the system is stable, for if there are no roots with positive real parts, the system is stable in continuous time framework. The algorithm for applying Routh's stability criterion requires the order of the polynomial (the characteristic equation) to be finite and is as follows:

Write the characteristic equation in the form

$$a_0 w^n + a_1 w^{n-1} + a_2 w^{n-2} + \cdots + a_n = 0, \tag{1}$$

where a_0 is positive (if a_0 is originally negative, both sides are multiplied by -1). In this form, it is necessary that all the coefficients

$$a_0, a_1, a_2, \cdots, a_{n-1}, a_n$$

be positive if all the roots are to lie in the left half plane. If any coefficient is negative, the system is definitely unstable, and the Routh test is not needed to answer the question of stability. However, in this case, the Routh test will tell us the number of roots in the right half plane. If all the coefficients are positive, the system may be stable or unstable. It is then necessary to apply the following procedure to determine stability.

2.1.1. Routh Array

Arrange the coefficients of Equation (1) into the first two rows of the Routh array (Table 1) as follows.

The array has been filled in for $n = 7$ in order to simplify the discussion. For any other value of n, the array is prepared in the same manner. In general, there

Table 1. Routh array.

Row					
1	a_0	a_2	a_4	a_6	…
2	a_1	a_3	a_5	a_7	…
3	b_1	b_2	b_3	…	
4	c_1	c_2	c_3	…	
5	d_1	d_2	…		
6	e_1	e_2	…		
7	f_1	…			
$n+1$	g_1	…			

are $(n+1)$ rows. For n even, the first row has one more element than the second row. The elements in the remaining rows are found from the formulas

$$\begin{array}{llll} b_1 = \dfrac{a_1a_2 - a_0a_3}{a_1} & b_2 = \dfrac{a_1a_4 - a_0a_5}{a_1} & b_3 = \dfrac{a_1a_6 - a_0a_7}{a_1} & \cdots \\ c_1 = \dfrac{b_1a_3 - a_1b_2}{b_1} & c_2 = \dfrac{b_1a_5 - a_1b_3}{b_1} & c_3 = \dfrac{b_1a_7 - a_1b_4}{b_1} & \cdots \\ d_1 = \dfrac{c_1b_2 - b_1c_2}{c_1} & d_2 = \dfrac{c_1b_3 - b_1c_3}{c_1} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{array}$$

The elements for the other rows are found from formulas that correspond to those just given. The elements in any row are always derived from the elements of the two preceding rows. During the computation of the Routh array, any row can be divided by a positive constant without changing the results of the test. (The application of this rule often simplifies the arithmetic.)

Having obtained the Routh array, the following theorems are applied to determine stability.

2.1.2. Theorems of the Routh Test

Theorem 1 *The necessary and sufficient condition for all the roots of the characteristic equation to have negative real parts (stable system) is that all elements of the first column of the Routh array be positive and non-zero.*

Theorem 2 *If some of the elements in the first column are negative, the number of roots with a positive real part (in the right half plane) is equal to the number of sign changes in the first column.*

Theorem 3 *If one pair of roots is on the imaginary axis, equidistant from the origin, and all other roots are in the left half plane, all the elements of the nth row will vanish and none of the elements of the preceding row will vanish. The location of the pair of imaginary roots can be found by solving the equation*

$$Cw^2 + D = 0, \tag{2}$$

where the coefficients C and D are the elements of the array in the (n − 1)th row *as read from left to right, respectively.*

The algebraic method for determining stability is limited in its usefulness in that all we can learn from it is whether a system is stable. It does not give us any idea of the degree of stability or the roots of the characteristic equation.

Example: Given the characteristic equation

$$w^4 + 3w^3 + 5w^2 + 4w + 2 = 0,$$

let's determine the stability by the Routh criterion as follows:

Since all the coefficients are positive, the system may be stable. To test this, form the following Routh array (Table 2).

Since there is no change in sign in the first column, there are no roots having positive real parts, and the system is stable.

2.2. Bilinear Transformation

The Routh test which is often used to examine the roots of the characteristic equation of a continuous system may also be used to examine the roots of the characteristic equation of a discrete data system. The Routh test detects the presence of roots in the right half of w-plane. Since the criterion of stability of a discrete data system requires that all roots fall within the unit circle of the z-plane (or outside the unit circle of the $\mathbb{Z} = L = z^{-1}$ plane), one must first apply a transformation that will map the inside of the unit circle of the z-plane into the left half of the w-plane. One can then apply the Routh test to discover roots in the right half of the w-plane, and if none are found, we know that the roots of the characteristic equation fall within the unit circle and that the discrete data system is stable.

A transformation that will map the inside of the unit circle of the z-plane into the left half of the w-plane is

$$z = \frac{w+1}{w-1}. \tag{3}$$

This transformation is called the bilinear-transformation. The regions involving the transformation are shown in Figure 1.

3. Hypothesis Testing

Let us consider the following AR(1) process:

$$(1 - \phi_1 L) y_t = c + \epsilon_t, \tag{4}$$

where $\epsilon_t \sim N(0,1)[iid]$.

The characteristic polynomial of the above expression is as follows: $1 - \phi_1 L = 0$ which can be written in the z-plane as: $1 - \phi_1 z^{-1} = 0$, which implies that

$$z - \phi_1 = 0.$$

Applying the bilinear transformation on the above expression gives:

$$\frac{w+1}{w-1} - \phi_1 = 0.$$

Table 2. Example of routh array.

Row			
1	1	5	2
2	3	4	
3	11/3	6/3	
4	26/11	0	
5	2		

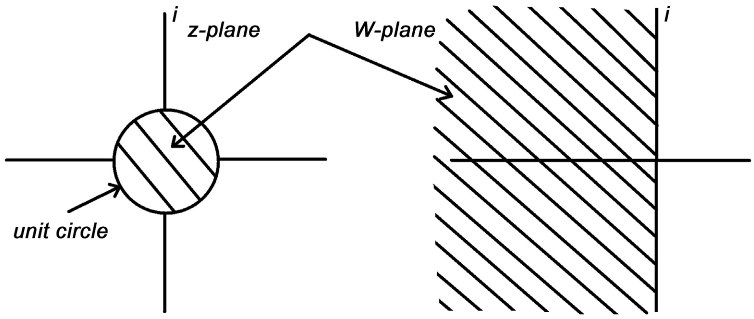

Figure 1. Bilinear transformation.

After rearranging, we get:

$$(1-\phi_1)w+(1+\phi_1)=0. \tag{5}$$

The Routh Array for the above expression is as follows:

Row	
1	$1-\phi_1$
2	$1+\phi_1$

The null hypothesis of stability of above AR(1) process is as follows:

$$H_0:\begin{bmatrix}1-\phi_1\\1+\phi_1\end{bmatrix}>0,\text{ or }\; H_0:-1<\phi_1<1.$$

As another example, suppose there is an AR(2) process:

$$(1-\phi_1 L-\phi_2 L^2)y_t=c+\epsilon_t, \tag{6}$$

where $\epsilon_t \sim N(0,1)[iid]$.

The characteristic polynomial of the above expression is as follows: $1-\phi_1 L-\phi_2 L^2=0$ which can be written in the *z*-plane as: $1-\phi_1 z^{-1}-\phi_2 z^{-2}=0$, which implies that

$$z^2-\phi_1 z-\phi_2=0.$$

Applying the bilinear transformation on the above expression gives:

$$\left(\frac{w+1}{w-1}\right)^2-\phi_1\left(\frac{w+1}{w-1}\right)-\phi_2=0.$$

After rearranging, we get:

$$\left(1-\phi_1-\phi_2\right)w^2+\left(2+2\phi_2\right)w+\left(1+\phi_1-\phi_2\right)=0. \quad (7)$$

The Routh Array for the above expression is as follows:

Row		
1	$1-\phi_1-\phi_2$	$1+\phi_1-\phi_2$
2	$2+2\phi_2$	
3	$1+\phi_1-\phi_2$	

The null hypothesis of stability of above AR(2) process is as follows:

$$H_0:\begin{bmatrix}1-\phi_1-\phi_2\\2+2\phi_2\\1+\phi_1-\phi_2\end{bmatrix}>0\text{, or }\ H_0:\begin{bmatrix}-\left(1-\phi_2\right)<\phi_1<\left(1-\phi_2\right)\\-1<\phi_2<1\end{bmatrix}.$$

Now to express the hypothesis testing in a general form, we can write a univariate time series as follows:

$$y_t=\boldsymbol{x}_t'\boldsymbol{\beta}+\epsilon_t, \quad (8)$$

where $\epsilon_t\sim N\left(0,1\right)\left[iid\right]$; y_t is a scalar time series variable, $\boldsymbol{x}_t$ is a $\left(k\times1\right)$ vector of regressors (each regressor represents some lagged value of y_t), $\boldsymbol{\beta}$ is a $\left(k\times1\right)$ vector of parameters of interest, and ϵ_t represents the unexplained part. After estimating the above model, we are interested in knowing whether the roots of the characteristic equation associated with the above model satisfy the stability criterion or not. Our null hypothesis is as follows:

$$H_0:\boldsymbol{a}<R\boldsymbol{\beta}<b,$$

against the alternative

$$H_1:\boldsymbol{b}\le\boldsymbol{R\beta}\text{, or }\boldsymbol{R\beta}\le\boldsymbol{a},$$

where the matrix of constraints $\boldsymbol{R}$ is a $\left(p\times k\right)$ matrix of rank *p*, where $p\le k$. $\boldsymbol{a}$ and $\boldsymbol{b}$ are known $\left(p\times1\right)$ vectors.

4. Minimization Problem with Inequality Constraints

The main challenge in the implementation of the above test is that in order to get the constrained estimates of $\boldsymbol{\beta}$'s, we need to do the following constrained minimization:

$$\min_{\beta}\sum_{t=1}^{T}\left(y_t-\boldsymbol{x}_t'\boldsymbol{\beta}\right)^2$$

subject to $\boldsymbol{a}<\boldsymbol{R\beta}<\boldsymbol{b}$.

There are various techniques in the current Mathematics literature that allow for constrained optimization with inequality constraints such as Linear programming (for linear objective function) [35], Quadratic programming or characterizing the problem in terms of the Karush-Kuhn-Tucker conditions (for non-linear objective function). However, these techniques require the inequality constraints to have an equality sign as well. On account of relying on these techniques, the current Statistics and Econometrics literature only allows a null hypothesis with an inequality sign if the equality sign is also there. Furthermore, this type of inequality constraint is seldom found in hypothesis testing where the parameters are bounded by a lower as well as an upper limit. To avoid a strict inequality sign in the null hypothesis, usually the problem is reframed in such a way that the null hypothesis involves a less than or equal to, or greater than or equal to sign. However, in this situation, if we try to make the alternative hypothesis as the null, the series can be unstable for a variety of reasons making the distribution under the null nearly impossible to calculate.

In general, a constraint with an upper and lower bound on parameters is more convenient and practical to test as compared to an equality constraint since the exact value of a parameter is hardly known. If $\boldsymbol{a} = \mathbf{0}$, and $\boldsymbol{b} = \infty$, then the above constraint is equivalent to the positiveness constraint of parameters. As the difference between $\boldsymbol{a}$ and $\boldsymbol{b}$ gets small, the inequality constraint approaches the equality constraint. In this regard, the above constraint is a general version of other type of constraints, and a methodology to handle this could have implications for hypothesis testing in general in econometrics, and solving the optimization problems with inequality constraints in Mathematics.

4.1. Methodology

4.1.1. β a Scalar

The inequality constraint is $a < \beta < b$. Let

$$k = \ln\left[\frac{\beta - a}{b - \beta}\right].$$

Taking the partial differential of the above expression, we get:

$$\delta k = \frac{b-a}{[\beta - a][b-\beta]}\delta\beta, \tag{9}$$

$\delta\beta$ is calculated as follows:

$$\delta\beta = \left(\sum_{t=1}^{T}\boldsymbol{x}_t\boldsymbol{x}_t'\right)^{-1}\left(\sum_{t=1}^{T}\left(\boldsymbol{x}_t y_t - \boldsymbol{x}_t'\beta_0\right)\right). \tag{10}$$

(See the detail of this step in the next section.)

Plugging this in Equation (9), we get δk which we can use in the following expression:

$$k = k_0 + \delta k, \tag{11}$$

k_0 is the initial value of *k*, *i.e.*,

$$k_0 = \ln\left[\frac{\beta_0 - a}{b - \beta_0}\right].$$

From Equation (11), we have

$$\ln\left[\frac{\beta - a}{b - \beta}\right] - \ln\left[\frac{\beta_0 - a}{b - \beta_0}\right] = \delta k,$$

$$\Rightarrow \ln\left[\frac{(\beta-a)(b-\beta_0)}{(b-\beta)(\beta_0-a)}\right]=\delta k,$$

$$\Rightarrow \left[\frac{(\beta-a)(b-\beta_0)}{(b-\beta)(\beta_0-a)}\right]=\exp(\delta k),$$

$$\Rightarrow (\beta-a)(b-\beta_0)=(b-\beta)(\beta_0-a)\exp(\delta k),$$

$$\Rightarrow \beta\left[(b-\beta_0)+(\beta_0-a)\exp(\delta k)\right]=a(b-\beta_0)+b(\beta_0-a)\exp(\delta k),$$

$$\Rightarrow \beta^{updated}=\frac{a(b-\beta_0)+b(\beta_0-a)\exp(\delta k)}{(b-\beta_0)+(\beta_0-a)\exp(\delta k)}.$$

Now treat the updated β as β_0 in Equation (10) and repeat the whole procedure until the estimate converges.

4.1.2. β a Vector

Let $\boldsymbol{A}\equiv\underset{(p\times p)}{\begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_p \end{bmatrix}}$; $\boldsymbol{B}\equiv\underset{(p\times p)}{\begin{bmatrix} b_1 & 0 & \cdots & 0 \\ 0 & b_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_p \end{bmatrix}}$;

$$\boldsymbol{D}\equiv\underset{(p\times p)}{\begin{bmatrix} (\boldsymbol{R}ß)_1 & 0 & \cdots & 0 \\ 0 & (\boldsymbol{R}ß)_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & (\boldsymbol{R}ß)_p \end{bmatrix}};\ \boldsymbol{e}_p\equiv\underset{(p\times 1)}{\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix}};\ \boldsymbol{a}=\boldsymbol{A}\boldsymbol{e}_p=\underset{(p\times 1)}{\begin{bmatrix}a_1\\a_2\\\vdots\\a_p\end{bmatrix}};\ \boldsymbol{b}=\boldsymbol{B}\boldsymbol{e}_p=\underset{(p\times 1)}{\begin{bmatrix}b_1\\b_2\\\vdots\\b_p\end{bmatrix}};$$

and $\boldsymbol{R}\boldsymbol{ß}=\boldsymbol{D}\boldsymbol{e}_p=\underset{(p\times 1)}{\begin{bmatrix}(\boldsymbol{R}ß)_1\\(\boldsymbol{R}ß)_2\\\vdots\\(\boldsymbol{R}ß)_p\end{bmatrix}}$.

The inequality constraint $\boldsymbol{a}<\boldsymbol{R}\boldsymbol{ß}<\boldsymbol{b}$ implies that $a_1<(\boldsymbol{R}\boldsymbol{ß})_1<b_1$, $a_2<(\boldsymbol{R}\boldsymbol{ß})_2<b_2$, $\cdots$, $a_p<(\boldsymbol{R}\boldsymbol{ß})_p<b_p$.

Let $k_1=\ln\left[\frac{(\boldsymbol{R}ß)_1-a_1}{b_1-(\boldsymbol{R}ß)_1}\right], k_2=\ln\left[\frac{(\boldsymbol{R}ß)_2-a_2}{b_2-(\boldsymbol{R}ß)_2}\right],\cdots,k_p=\ln\left[\frac{(\boldsymbol{R}ß)_p-a_p}{b_p-(\boldsymbol{R}ß)_p}\right]$,

$$\boldsymbol{K}\equiv\underset{(p\times p)}{\begin{bmatrix} k_1 & 0 & \cdots & 0 \\ 0 & k_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & k_p \end{bmatrix}}=\underset{(p\times p)}{\begin{bmatrix} \ln\left[\frac{(\boldsymbol{R}ß)_1-a_1}{b_1-(\boldsymbol{R}ß)_1}\right] & 0 & \cdots & 0 \\ 0 & \ln\left[\frac{(\boldsymbol{R}ß)_2-a_2}{b_2-(\boldsymbol{R}ß)_2}\right] & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \ln\left[\frac{(\boldsymbol{R}ß)_p-a_p}{b_p-(\boldsymbol{R}ß)_p}\right] \end{bmatrix}}$$

or

$$\boldsymbol{K} = \ln d\left[\left[\boldsymbol{B}-\boldsymbol{D}\right]^{-1}\left[\boldsymbol{D}-\boldsymbol{A}\right]\right], \tag{12}$$

where $\ln d[.] \equiv$ log of diagonal of the matrix

(Note: This is just a defining expression. It does not say that we can take the log of only the diagonal of a matrix.)

$$\boldsymbol{k} = \boldsymbol{K}\boldsymbol{e}_p = \underset{(p\times 1)}{\begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_p \end{bmatrix}} = \underset{(p\times 1)}{\begin{bmatrix} \ln\left[\dfrac{(\boldsymbol{R}ß)_1 - a_1}{b_1 - (\boldsymbol{R}ß)_1}\right] \\ \ln\left[\dfrac{(\boldsymbol{R}ß)_2 - a_2}{b_2 - (\boldsymbol{R}ß)_2}\right] \\ \vdots \\ \ln\left[\dfrac{(\boldsymbol{R}ß)_p - a_p}{b_p - (\boldsymbol{R}ß)_p}\right] \end{bmatrix}} = \ln d\left[\left[\boldsymbol{B}-\boldsymbol{D}\right]^{-1}\left[\boldsymbol{D}-\boldsymbol{A}\right]\right]\boldsymbol{e}_p,$$

$$\left[\boldsymbol{B}-\boldsymbol{D}\right]^{-1} = \begin{bmatrix} \left[b_1 - (\boldsymbol{R}ß)_1\right]^{-1} & 0 & \cdots & 0 \\ 0 & \left[b_2 - (\boldsymbol{R}ß)_2\right]^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \left[b_p - (\boldsymbol{R}ß)_p\right]^{-1} \end{bmatrix}; \text{ and}$$

$$\left[\boldsymbol{D}-\boldsymbol{A}\right] = \begin{bmatrix} \left[(\boldsymbol{R}ß)_1 - a_1\right] & 0 & \cdots & 0 \\ 0 & \left[(\boldsymbol{R}ß)_2 - a_2\right] & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \left[(\boldsymbol{R}ß)_p - a_p\right] \end{bmatrix}.$$

Taking the partial differential of the expression for k_1, we get:

$$\delta k_1 = \frac{b_1 - a_1}{\left[(\boldsymbol{R}ß)_1 - a_1\right]\left[b_1 - (\boldsymbol{R}ß)_1\right]}\delta(\boldsymbol{R}ß)_1.$$

Similarly

$$\delta k_2 = \frac{b_2 - a_2}{\left[(\boldsymbol{R}ß)_2 - a_2\right]\left[b_2 - (\boldsymbol{R}ß)_2\right]}\delta(\boldsymbol{R}ß)_2,$$

$$\vdots$$

$$\delta k_p = \frac{b_p - a_p}{\left[(\boldsymbol{R}ß)_p - a_p\right]\left[b_p - (\boldsymbol{R}ß)_p\right]}\delta(\boldsymbol{R}ß)_p.$$

δk_1 and $\delta(\boldsymbol{R}\boldsymbol{ß})_1$ have the same sign as the multiplier of $\delta(\boldsymbol{R}\boldsymbol{ß})_1$ is positive. In matrix notation, we can write:

$$\underset{(p\times 1)}{\delta\boldsymbol{k}} = \underset{(p\times p)}{\left[\boldsymbol{D}-\boldsymbol{A}\right]^{-1}}\underset{(p\times p)}{\left[\boldsymbol{B}-\boldsymbol{D}\right]^{-1}}\underset{(p\times p)}{\left[\boldsymbol{B}-\boldsymbol{A}\right]}\underset{(p\times 1)}{\delta(\boldsymbol{R}ß)}, \text{ or}$$

$$\underset{(p\times 1)}{\delta \boldsymbol{k}} = \begin{bmatrix} \delta k_1 \\ \delta k_2 \\ \vdots \\ \delta k_p \end{bmatrix} = \begin{bmatrix} \dfrac{b_1 - a_1}{\left[(\boldsymbol{R}\beta)_1 - a_1\right]\left[b_1 - (\boldsymbol{R}\beta)_1\right]}\delta(\boldsymbol{R}\beta)_1 \\ \dfrac{b_2 - a_2}{\left[(\boldsymbol{R}\beta)_2 - a_2\right]\left[b_2 - (\boldsymbol{R}\beta)_2\right]}\delta(\boldsymbol{R}\beta)_2 \\ \vdots \\ \dfrac{b_p - a_p}{\left[(\boldsymbol{R}\beta)_p - a_p\right]\left[b_p - (\boldsymbol{R}\beta)_p\right]}\delta(\boldsymbol{R}\beta)_p \end{bmatrix}, \tag{13}$$

$$[\boldsymbol{D} - \boldsymbol{A}]^{-1} = \begin{bmatrix} \left[(\boldsymbol{R}\beta)_1 - a_1\right]^{-1} & 0 & \cdots & 0 \\ 0 & \left[(\boldsymbol{R}\beta)_2 - a_2\right]^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \left[(\boldsymbol{R}\beta)_p - a_p\right]^{-1} \end{bmatrix}; \text{ and}$$

$$\delta(\boldsymbol{R\beta}) = \begin{bmatrix} \delta(\boldsymbol{R}\beta)_1 \\ \delta(\boldsymbol{R}\beta)_2 \\ \vdots \\ \delta(\boldsymbol{R}\beta)_p \end{bmatrix}.$$

$$\delta(\boldsymbol{R\beta}) = \boldsymbol{R} \cdot \delta\boldsymbol{\beta}$$

Similarly

$$\delta \boldsymbol{K} = \underset{(p\times p)}{\begin{bmatrix} \delta k_1 & 0 & \cdots & 0 \\ 0 & \delta k_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \delta k_p \end{bmatrix}}, \text{ or}$$

$$\delta \boldsymbol{K} = \underset{(p\times p)}{\begin{bmatrix} \dfrac{(b_1 - a_1)\cdot\delta(\boldsymbol{R}\beta)_1}{\left[(\boldsymbol{R}\beta)_1 - a_1\right]\left[b_1 - (\boldsymbol{R}\beta)_1\right]} & 0 & \cdots & 0 \\ 0 & \dfrac{(b_2 - a_2)\cdot\delta(\boldsymbol{R}\beta)_2}{\left[(\boldsymbol{R}\beta)_2 - a_2\right]\left[b_p - (\boldsymbol{R}\beta)_2\right]} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \dfrac{(b_p - a_p)\cdot\delta(\boldsymbol{R}\beta)_p}{\left[(\boldsymbol{R}\beta)_p - a_p\right]\left[b_p - (\boldsymbol{R}\beta)_p\right]} \end{bmatrix}} \tag{14}$$

$\delta\boldsymbol{\beta}$ is calculated as follows:

Let $f(\boldsymbol{x}_t, y_t; \boldsymbol{\beta}) = \sum_{t=1}^{T}(y_t - \boldsymbol{x}_t'\beta)^2$, which we want to minimize with respect to $\boldsymbol{\beta}$, *i.e.*, we want to find

$$\frac{\partial f}{\partial \beta} = 0,$$

$$\underset{(k\times 1)}{\frac{\partial f}{\partial \beta}} = -2\sum_{t=1}^{T} \underset{(k\times 1)}{\boldsymbol{x}_t} \underset{(1\times 1)}{(y_t - \boldsymbol{x}_t'\beta)}; \text{ and } \underset{(k\times k)}{\frac{\partial^2 f}{\partial\beta\partial\beta'}} = 2\sum_{t=1}^{T} \boldsymbol{x}_t\boldsymbol{x}_t'.$$

4.1.3. Newton-Raphson Method

The equation of the tangent line to the curve

$$Z = \frac{\partial f}{\partial ß},$$

at point $ß_0$ is

$$Z(ß) = \left.\frac{\partial^2 f}{\partial ß \partial ß'}\right|_{ß=ß_0} (ß - ß_0) + \left.\frac{\partial f}{\partial ß}\right|_{ß=ß_0}. \tag{15}$$

Setting $Z(ß) = 0$, and $ß = ß_1$ gives:

$$0 = \left.\frac{\partial^2 f}{\partial ß \partial ß'}\right|_{ß=ß_0} (ß_1 - ß_0) + \left.\frac{\partial f}{\partial ß}\right|_{ß=ß_0},$$

$$\Rightarrow ß_1 = ß_0 - \left(\left.\frac{\partial^2 f}{\partial ß \partial ß'}\right|_{ß=ß_0} \right)^{-1} \left(\left.\frac{\partial f}{\partial ß}\right|_{ß=ß_0} \right),$$

$$\Rightarrow ß_1 = ß_0 + \left(\sum_{t=1}^{T} \boldsymbol{x}_t \boldsymbol{x}_t'\right)^{-1} \left(\sum_{t=1}^{T} \boldsymbol{x}_t \left(y_t - \boldsymbol{x}_t' ß_0\right)\right).$$

The above equation can also be written as

$$\delta ß = \left(\sum_{t=1}^{T} \boldsymbol{x}_t \boldsymbol{x}_t'\right)^{-1} \left(\sum_{t=1}^{T} \boldsymbol{x}_t \left(y_t - \boldsymbol{x}_t' ß_0\right)\right). \tag{16}$$

Putting this in Equation (14), we get $\delta \boldsymbol{K}$ which we can use in the following expression:

$$\boldsymbol{K} = \boldsymbol{K}_0 + \delta \boldsymbol{K}, \tag{17}$$

where $\boldsymbol{K}_0$ is the initial value of $\boldsymbol{K}$, *i.e.*,

$$\boldsymbol{K}_0 = \begin{bmatrix} \ln\left[\frac{(\boldsymbol{R}ß_0)_1 - a_1}{b_1 - (\boldsymbol{R}ß_0)_1}\right] & 0 & \cdots & 0 \\ 0 & \ln\left[\frac{(\boldsymbol{R}ß_0)_2 - a_2}{b_2 - (\boldsymbol{R}ß_0)_2}\right] & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \ln\left[\frac{(\boldsymbol{R}ß_0)_p - a_p}{b_p - (\boldsymbol{R}ß_0)_p}\right] \end{bmatrix}_{(p \times p)}.$$

Putting Equation (12) into Equation (17), we get:

$$\ln d\left[[\boldsymbol{B} - \boldsymbol{D}]^{-1}[\boldsymbol{D} - \boldsymbol{A}]\right] - \ln d\left[[\boldsymbol{B} - \boldsymbol{D}_0]^{-1}[\boldsymbol{D}_0 - \boldsymbol{A}]\right] = \delta \boldsymbol{K},$$

$$\Rightarrow \ln d\left[[\boldsymbol{D}_0 - \boldsymbol{A}]^{-1}[\boldsymbol{B} - \boldsymbol{D}_0][\boldsymbol{B} - \boldsymbol{D}]^{-1}[\boldsymbol{D} - \boldsymbol{A}]\right] = \delta \boldsymbol{K},$$

$$\Rightarrow [\boldsymbol{D}_0 - \boldsymbol{A}]^{-1}[\boldsymbol{B} - \boldsymbol{D}_0][\boldsymbol{B} - \boldsymbol{D}]^{-1}[\boldsymbol{D} - \boldsymbol{A}] = \exp d(\delta \boldsymbol{K}),$$

where $\exp d(.) \equiv$ exponential of the diagonal of the matrix

(Note: This is just a defining expression. It does not say that we can take the exponential of only the diagonal of a matrix.)

$$\Rightarrow [\boldsymbol{D} - \boldsymbol{A}] = [\boldsymbol{B} - \boldsymbol{D}][\boldsymbol{B} - \boldsymbol{D}_0]^{-1}[\boldsymbol{D}_0 - \boldsymbol{A}] \exp d(\delta \boldsymbol{K}),$$

$$\Rightarrow \boldsymbol{D} = \boldsymbol{A} + [\boldsymbol{B} - \boldsymbol{D}][\boldsymbol{B} - \boldsymbol{D}_0]^{-1}[\boldsymbol{D}_0 - \boldsymbol{A}]\exp d(\delta\boldsymbol{K}),$$

$$\Rightarrow \boldsymbol{D}\left\{\boldsymbol{I} + [\boldsymbol{B} - \boldsymbol{D}_0]^{-1}[\boldsymbol{D}_0 - \boldsymbol{A}]\exp d(\delta\boldsymbol{K})\right\} = \boldsymbol{A} + \boldsymbol{B}[\boldsymbol{B} - \boldsymbol{D}_0]^{-1}[\boldsymbol{D}_0 - \boldsymbol{A}]\exp d(\delta\boldsymbol{K}),$$

$$\Rightarrow \boldsymbol{D}\{[\boldsymbol{B} - \boldsymbol{D}_0] + [\boldsymbol{D}_0 - \boldsymbol{A}]\exp d(\delta\boldsymbol{K})\} = [\boldsymbol{B} - \boldsymbol{D}_0]\boldsymbol{A} + \boldsymbol{B}[\boldsymbol{D}_0 - \boldsymbol{A}]\exp d(\delta\boldsymbol{K}),$$

$$\left(\text{as } \boldsymbol{B}[\boldsymbol{B} - \boldsymbol{D}_0]^{-1} = [\boldsymbol{B} - \boldsymbol{D}_0]^{-1}\boldsymbol{B}\right),$$

$$\Rightarrow \boldsymbol{D} = \{[\boldsymbol{B} - \boldsymbol{D}_0] + [\boldsymbol{D}_0 - \boldsymbol{A}]\exp d(\delta\boldsymbol{K})\}^{-1}\{[\boldsymbol{B} - \boldsymbol{D}_0]\boldsymbol{A} + \boldsymbol{B}[\boldsymbol{D}_0 - \boldsymbol{A}]\exp d(\delta\boldsymbol{K})\},$$

$$\Rightarrow \boldsymbol{D}\boldsymbol{e}_p = \{[\boldsymbol{B} - \boldsymbol{D}_0] + [\boldsymbol{D}_0 - \boldsymbol{A}]\exp d(\delta\boldsymbol{K})\}^{-1}\{[\boldsymbol{B} - \boldsymbol{D}_0]\boldsymbol{A} + [\boldsymbol{D}_0 - \boldsymbol{A}]\exp d(\delta\boldsymbol{K})\cdot\boldsymbol{B}\}\boldsymbol{e}_p,$$

$$\left(\text{as } \boldsymbol{B}[\boldsymbol{D}_0 - \boldsymbol{A}]\exp d(\delta\boldsymbol{K}) = [\boldsymbol{D}_0 - \boldsymbol{A}]\exp d(\delta\boldsymbol{K})\cdot\boldsymbol{B}\right).$$

This implies that

$$\boldsymbol{R}ß^{updated} = \{[\boldsymbol{B} - \boldsymbol{D}_0] + [\boldsymbol{D}_0 - \boldsymbol{A}]\exp d(\delta\boldsymbol{K})\}^{-1}\{[\boldsymbol{B} - \boldsymbol{D}_0]\boldsymbol{a} + [\boldsymbol{D}_0 - \boldsymbol{A}]\exp d(\delta\boldsymbol{K})\cdot\boldsymbol{b}\}. \quad (18)$$

In expanded form, the above expression can be written as follows:

$$\boldsymbol{R}ß^{updated} = \begin{bmatrix} (\boldsymbol{R}ß^{updated})_1 \\ (\boldsymbol{R}ß^{updated})_2 \\ \vdots \\ (\boldsymbol{R}ß^{updated})_p \end{bmatrix}_{(p\times 1)} = \begin{bmatrix} \dfrac{a_1[b_1 - (\boldsymbol{R}ß_0)_1] + b_1[(\boldsymbol{R}ß_0)_1 - a_1]\exp(\delta k_1)}{[b_1 - (\boldsymbol{R}ß_0)_1] + [(\boldsymbol{R}ß_0)_1 - a_1]\exp(\delta k_1)} \\ \dfrac{a_2[b_2 - (\boldsymbol{R}ß_0)_2] + b_2[(\boldsymbol{R}ß_0)_2 - a_2]\exp(\delta k_2)}{[b_2 - (\boldsymbol{R}ß_0)_2] + [(\boldsymbol{R}ß_0)_2 - a_2]\exp(\delta k_2)} \\ \vdots \\ \dfrac{a_p[b_p - (\boldsymbol{R}ß_0)_p] + b_p[(\boldsymbol{R}ß_0)_p - a_p]\exp(\delta k_p)}{[b_p - (\boldsymbol{R}ß_0)_p] + [(\boldsymbol{R}ß_0)_p - a_p]\exp(\delta k_p)} \end{bmatrix}_{(p\times 1)}. \quad (19)$$

We need to keep the restricted estimate as far from the boundaries as if the unrestricted estimate was equal to a boundary value, *i.e.*, either $\boldsymbol{a}$ or $\boldsymbol{b}$, we were able to reject the null. Therefore, we can allow the restricted estimate to take a value at the most at that point, e.g., if **ß** is a scalar, and the restricted estimate is converging toward a value close to *b*, we need a t-statistic (at 95% confidence level and 100 degrees of free1dom) as follows:

$$t = \frac{b - \tilde{\beta}}{se(\hat{\beta})} \geqq 1.984.$$

This implies that

$$\tilde{\beta} \le b - 1.984 * se(\hat{\beta}).$$

This leads to the intended performance of the test, *i.e.*, if the unrestricted estimate is less than *b*, we shall be able "not to reject" the null, whereas if the unrestricted estimate is greater than or equal to *b*, we will be able to reject the null.

4.1.4. Summary

Now let us summarize the step-wise methodology as follows:

1) Assume some initial values of ß, *i.e.*, $ß_0$ in the interval $a < Rß < b$.

2) Calculate $\delta ß$ from Equation (16) using the assumed initial values $ß_0$.

3) Plug this value of $\delta ß$ along with the initial values $ß_0$ in Equation (14) to get δK.

4) Plug the value of δK in Equation (18) to get updated ß.

5) Now treat the updated ß as $ß_0$ and repeat steps 2 to 5.

6) Stop the algorithm (if necessary) at the point discussed above.

5. Stability Test

Consider an ARMA model of the form

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q},$$

with ϵ_t as white noise:

$$E(\epsilon_t) = 0,$$

$$E(\epsilon_t \epsilon_\tau) = \begin{cases} \sigma^2 & \text{for } t = \tau, \\ 0 & \text{otherwise.} \end{cases}$$

Let $\boldsymbol{\beta} \equiv (c, \phi_1, \cdots, \phi_p, \theta_1, \cdots, \theta_q, \sigma^2)'$ denote the $k \times 1$ vector of population parameters.

Theorem 4 *Suppose that* $y_1, y_2, \cdots, y_T$ *have the joint probability density*

$$f_{Y_T, Y_{T-1}, \cdots, Y_1}(y_T, y_{T-1}, \cdots, y_1; \boldsymbol{\beta}),$$

$\boldsymbol{\beta} \in \Theta$, *and* $f_{Y_T, Y_{T-1}, \cdots, Y_1}(y_T, y_{T-1}, \cdots, y_1; \boldsymbol{\beta})$ *satisfies the following assumptions*.

Assumption 1: *Identifiability*; $\boldsymbol{\beta}_1 \neq \boldsymbol{\beta}_2$ *implies that*

$$F_{Y_T, Y_{T-1}, \cdots, Y_1}(y_T, y_{T-1}, \cdots, y_1; \boldsymbol{\beta}_1) \neq F_{Y_T, Y_{T-1}, \cdots, Y_1}(y_T, y_{T-1}, \cdots, y_1; \boldsymbol{\beta}_2).$$

Assumption 2: *For each* $\boldsymbol{\beta} \in \boldsymbol{\Theta}$, $F_{Y_T, Y_{T-1}, \cdots, Y_1}(y_T, y_{T-1}, \cdots, y_1; \boldsymbol{\beta})$ *has the same support not depending on* $\boldsymbol{\beta}$.

Assumption 3: *For each* $\boldsymbol{\beta} \in \boldsymbol{\Theta}$, *the first three partial derivatives of*

$$\left(\log f_{Y_T, Y_{T-1}, \cdots, Y_1}(y_T, y_{T-1}, \cdots, y_1; \boldsymbol{\beta})\right),$$

with respect to β *exist for* $y_T, y_{T-1}, \cdots, y_1$ *in the support of*

$$F_{Y_T, Y_{T-1}, \cdots, Y_1}(y_T, y_{T-1}, \cdots, y_1; \boldsymbol{\beta}).$$

Assumption 4: *For each* $\boldsymbol{\beta} \in \boldsymbol{\Theta}$, *there exists a function* $g(y_T, y_{T-1}, \cdots, y_1)$ (*possibly depending on* $\boldsymbol{\beta}$), *such that in a neighborhood of the given* $\boldsymbol{\beta}$ *and for all* $l, m, n \in \{1, \cdots, k\}$,

$$\left| \frac{\partial^3}{\partial \beta_l \partial \beta_m \partial \beta_n} \log f_{Y_T, Y_{T-1}, \cdots, Y_1}(y_T, y_{T-1}, \cdots, y_1; \boldsymbol{\beta}) \right| \leq g(y_T, y_{T-1}, \cdots, y_1),$$

where $\int g(y_T, y_{T-1}, \cdots, y_1) \mathrm{d}F_{Y_T, Y_{T-1}, \cdots, Y_1}(y_T, y_{T-1}, \cdots, y_1; \boldsymbol{\beta}) < \infty$.

Assumption 5: *For each* $\boldsymbol{\beta} \in \boldsymbol{\Theta}$, $E\left[\partial \log f(y_T, y_{T-1}, \cdots, y_1; \boldsymbol{\beta}) / \partial \boldsymbol{\beta}\right] = 0$,

$$\begin{aligned} \boldsymbol{I}(\boldsymbol{\beta}) &= E\left[\frac{\partial}{\partial \boldsymbol{\beta}} \log f(y_T, y_{T-1}, \cdots, y_1; \boldsymbol{\beta}) \frac{\partial}{\partial \boldsymbol{\beta}'} \log f(y_T, y_{T-1}, \cdots, y_1; \boldsymbol{\beta})\right] \\ &= E\left[-\frac{\partial^2}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'} \log f(y_T, y_{T-1}, \cdots, y_1; \boldsymbol{\beta})\right], \end{aligned}$$

and $\boldsymbol{I}(\boldsymbol{\beta})$ *is nonsingular.*

$\hat{\boldsymbol{\beta}}$ *satisfies* $S(\hat{\boldsymbol{\beta}}) = \partial L(\beta)/\partial \beta\big|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}} = 0$ *where*

$$L(\boldsymbol{\beta}) = \log f_{Y_T, Y_{T-1}, \cdots, Y_1}(y_T, y_{T-1}, \cdots, y_1; \boldsymbol{\beta}),$$

and $\hat{\boldsymbol{\beta}} \xrightarrow{p} \beta$ *as* $T \to \infty$.

Theorem 5 *Suppose that* $y_1, y_2, \cdots, y_T$ *have the joint probability density*

$$f_{Y_T, Y_{T-1}, \cdots, Y_1}(y_T, y_{T-1}, \cdots, y_1; \boldsymbol{\beta}),$$

$\boldsymbol{\beta} \in \Theta$, *and all the assumptions of theorem* 4 *hold, then under the null*

$$H_0 : \boldsymbol{\beta} = \tilde{\boldsymbol{\beta}},$$

where $\tilde{\boldsymbol{\beta}}$ *is the restricted estimator,*

$$LR = -2\log\left(L(\tilde{\boldsymbol{\beta}})\big/L(\hat{\boldsymbol{\beta}})\right) = 2\left(\log L(\hat{\boldsymbol{\beta}}) - \log L(\tilde{\boldsymbol{\beta}})\right) \xrightarrow{d} \chi_p^2 \left(\text{under } H_0\right),$$

where $L(\tilde{\boldsymbol{\beta}})$ *and* $L(\hat{\boldsymbol{\beta}})$ *denote the values of the log likelihood function at the restricted (* $\tilde{\boldsymbol{\beta}}$ *) and the unrestricted (* $\hat{\boldsymbol{\beta}}$ *) estimates respectively.*

Theorem 6 *If all the assumptions of theorem* 4 *hold, then under the null hypothesis that the restrictions are true*

$$LM = T^{-1}\boldsymbol{S}(\tilde{\boldsymbol{\beta}})' \boldsymbol{I}(\tilde{\boldsymbol{\beta}})^{-1} \boldsymbol{S}(\tilde{\boldsymbol{\beta}}) \xrightarrow{d} \chi_p^2 \left(\text{under } H_0\right).$$

Theorem 7 *If all the assumptions of theorem* 4 *hold and the null hypothesis is true, then*

$$Wald = T\left(\hat{\boldsymbol{\beta}} - \tilde{\boldsymbol{\beta}}\right)' \left[\boldsymbol{I}(\hat{\boldsymbol{\beta}})^{-1}\right]^{-1} \left(\hat{\boldsymbol{\beta}} - \tilde{\boldsymbol{\beta}}\right) \xrightarrow{d} \chi_p^2 \left(\text{under } H_0\right).$$

For Equation (8), *and the null hypothesis of the form* $\boldsymbol{R\beta} = \boldsymbol{R}\tilde{\boldsymbol{\beta}}$, *the above Wald statistic can be written as follows:*

$$W = \sqrt{T}\left(\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{R}\tilde{\boldsymbol{\beta}}\right)' \left[\hat{\sigma}^2 \boldsymbol{R}\left(T^{-1}\sum_{t=1}^{T} \boldsymbol{x}_t \boldsymbol{x}_t'\right)^{-1} \boldsymbol{R}'\right]^{-1} \sqrt{T}\left(\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{R}\tilde{\boldsymbol{\beta}}\right)$$
$$\xrightarrow{d} \chi_p^2 \left(\text{under } H_0\right),$$

where $\hat{\boldsymbol{\beta}}$ *and* $\tilde{\boldsymbol{\beta}}$ *are the unrestricted and the restricted estimators respectively.*

Theorem 8 *If an ARMA* (*p*, *q*) *process is misspecified and there is some serial correlation in the noise term, we need to use the robust inference given in* [36].

Theorem 9 *Let us consider a VAR of the form*

$$\boldsymbol{Y}_t = \boldsymbol{c} + \Phi_1 \boldsymbol{Y}_{t-1} + \Phi_2 \boldsymbol{Y}_{t-2} + \cdots + \Phi_p \boldsymbol{Y}_{t-p} + \boldsymbol{\epsilon}_t,$$

where Y_t *denote an* $(n \times 1)$ *vector containing the values that n variables take at date t.*

$$\boldsymbol{\epsilon}_t \sim i.i.d.(\boldsymbol{0}, \boldsymbol{\Omega}).$$

Let $\Pi' = \begin{bmatrix} \boldsymbol{c} & \Phi_1 & \Phi_2 & \cdots & \Phi_p \end{bmatrix}$ *which is* $n \times (np+1)$ *matrix.* $\boldsymbol{\beta} \equiv Vec\Pi$ *denote the* $\left((n^2 p + n) \times 1\right)$ *vector of population parameters. Suppose that we have observed each of n variables for* $(T+p)$ *time periods and* $y_1, y_2, \cdots, y_T$ *have the* *conditional joint probability density*

$$f_{\boldsymbol{Y}_T,\boldsymbol{Y}_{T-1},\cdots,\boldsymbol{Y}_1|\boldsymbol{Y}_0,\boldsymbol{Y}_{-1},\cdots,\boldsymbol{Y}_{-p+1}}\left(\boldsymbol{y}_T,\boldsymbol{y}_{T-1},\cdots,\boldsymbol{y}_1 \mid \boldsymbol{y}_0,\boldsymbol{y}_{-1},\cdots,\boldsymbol{y}_{-p+1};\boldsymbol{\beta}\right),$$

$\boldsymbol{\beta}\in\Theta$, *and the above conditional joint density satisfies the following assumptions*:

Assumption 6: *Identifiability*; $\boldsymbol{\beta}_1 \neq \boldsymbol{\beta}_2$ *implies that*

$$F_{\boldsymbol{Y}_T,\boldsymbol{Y}_{T-1},\cdots,\boldsymbol{Y}_1|\boldsymbol{Y}_0,\boldsymbol{Y}_{-1},\cdots,\boldsymbol{Y}_{-p+1}}\left(\boldsymbol{y}_T,\boldsymbol{y}_{T-1},\cdots,\boldsymbol{y}_1 \mid \boldsymbol{y}_0,\boldsymbol{y}_{-1},\cdots,\boldsymbol{y}_{-p+1};\boldsymbol{\beta}_1\right)$$
$$\neq F_{\boldsymbol{Y}_T,\boldsymbol{Y}_{T-1},\cdots,\boldsymbol{Y}_1|\boldsymbol{Y}_0,\boldsymbol{Y}_{-1},\cdots,\boldsymbol{Y}_{-p+1}}\left(\boldsymbol{y}_T,\boldsymbol{y}_{T-1},\cdots,\boldsymbol{y}_1 \mid \boldsymbol{y}_0,\boldsymbol{y}_{-1},\cdots,\boldsymbol{y}_{-p+1};\boldsymbol{\beta}_2\right).$$

Assumption 7: *For each* $\boldsymbol{\beta}\in\Theta$,

$$F_{\boldsymbol{Y}_T,\boldsymbol{Y}_{T-1},\cdots,\boldsymbol{Y}_1|\boldsymbol{Y}_0,\boldsymbol{Y}_{-1},\cdots,\boldsymbol{Y}_{-p+1}}\left(\boldsymbol{y}_T,\boldsymbol{y}_{T-1},\cdots,\boldsymbol{y}_1 \mid \boldsymbol{y}_0,\boldsymbol{y}_{-1},\cdots,\boldsymbol{y}_{-p+1};\boldsymbol{\beta}\right)$$

has the same support not depending on $\boldsymbol{\beta}$.

Assumption 8: *For each* $\boldsymbol{\beta}\in\Theta$, *the first three partial derivatives of*

$$\log f_{\boldsymbol{Y}_T,\boldsymbol{Y}_{T-1},\cdots,\boldsymbol{Y}_1|\boldsymbol{Y}_0,\boldsymbol{Y}_{-1},\cdots,\boldsymbol{Y}_{-p+1}}\left(\boldsymbol{y}_T,\boldsymbol{y}_{T-1},\cdots,\boldsymbol{y}_1 \mid \boldsymbol{y}_0,\boldsymbol{y}_{-1},\cdots,\boldsymbol{y}_{-p+1};\boldsymbol{\beta}\right),$$

with respect to $\boldsymbol{\beta}$ *exist for* $y_T, y_{T-1},\cdots,y_1$ *in the support of*

$$F_{\boldsymbol{Y}_T,\boldsymbol{Y}_{T-1},\cdots,\boldsymbol{Y}_1|\boldsymbol{Y}_0,\boldsymbol{Y}_{-1},\cdots,\boldsymbol{Y}_{-p+1}}\left(\boldsymbol{y}_T,\boldsymbol{y}_{T-1},\cdots,\boldsymbol{y}_1 \mid \boldsymbol{y}_0,\boldsymbol{y}_{-1},\cdots,\boldsymbol{y}_{-p+1};\boldsymbol{\beta}\right).$$

Assumption 9: *For each* $\boldsymbol{\beta}\in\Theta$, *there exists a function* $g\left(\boldsymbol{y}_T,\boldsymbol{y}_{T-1},\cdots,\boldsymbol{y}_1\right)$ (*possibly depending on* $\boldsymbol{\beta}$), *such that in a neighborhood of the given* $\boldsymbol{\beta}$ *and for all* $l,m,n\in\left\{1,\cdots,\left(n^2p+n\right)\right\}$,

$$\left|\frac{\partial^3}{\partial\beta_l\partial\beta_m\partial\beta_n}\log f\left(\boldsymbol{y}_T,\boldsymbol{y}_{T-1},\cdots,\boldsymbol{y}_1 \mid \boldsymbol{y}_0,\boldsymbol{y}_{-1},\cdots,\boldsymbol{y}_{-p+1};\boldsymbol{\beta}\right)\right| \leq g\left(\boldsymbol{y}_T,\boldsymbol{y}_{T-1},\cdots,\boldsymbol{y}_1\right),$$

where

$$\int g\left(\boldsymbol{y}_T,\boldsymbol{y}_{T-1},\cdots,\boldsymbol{y}_1\right)\mathrm{d}F\left(\boldsymbol{y}_T,\boldsymbol{y}_{T-1},\cdots,\boldsymbol{y}_1 \mid \boldsymbol{y}_0,\boldsymbol{y}_{-1},\cdots,\boldsymbol{y}_{-p+1};\boldsymbol{\beta}\right) < \infty.$$

Assumption 10: *For each* $\boldsymbol{\beta}\in\Theta$,

$$E\left[\partial\log f\left(\boldsymbol{y}_T,\boldsymbol{y}_{T-1},\cdots,\boldsymbol{y}_1 \mid \boldsymbol{y}_0,\boldsymbol{y}_{-1},\cdots,\boldsymbol{y}_{-p+1};\boldsymbol{\beta}\right)/\partial\boldsymbol{\beta}\right] = 0,$$

$$\begin{aligned}\boldsymbol{I}\left(\boldsymbol{\beta}\right) &= E\left[\frac{\partial}{\partial\boldsymbol{\beta}}\log f\left(\boldsymbol{y}_T,\boldsymbol{y}_{T-1},\cdots,\boldsymbol{y}_1 \mid \boldsymbol{y}_0,\boldsymbol{y}_{-1},\cdots,\boldsymbol{y}_{-p+1};\boldsymbol{\beta}\right)\right.\\ &\quad\left.\cdot\frac{\partial}{\partial\boldsymbol{\beta}'} f\left(\boldsymbol{y}_T,\boldsymbol{y}_{T-1},\cdots,\boldsymbol{y}_1 \mid \boldsymbol{y}_0,\boldsymbol{y}_{-1},\cdots,\boldsymbol{y}_{-p+1};\boldsymbol{\beta}\right)\right]\\ &= E\left[-\frac{\partial^2}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'}\log f\left(\boldsymbol{y}_T,\boldsymbol{y}_{T-1},\cdots,\boldsymbol{y}_1 \mid \boldsymbol{y}_0,\boldsymbol{y}_{-1},\cdots,\boldsymbol{y}_{-p+1};\boldsymbol{\beta}\right)\right],\end{aligned}$$

and $\boldsymbol{I}\left(\boldsymbol{\beta}\right)$ *is nonsingular*.

$\hat{\boldsymbol{\beta}}$ *satisfies* $S\left(\hat{\boldsymbol{\beta}}\right) = \partial L\left(\boldsymbol{\beta}\right)/\partial\boldsymbol{\beta}\big|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}} = 0$ *where*

$$L\left(\boldsymbol{\beta}\right) = \log f_{\boldsymbol{Y}_T,\boldsymbol{Y}_{T-1},\cdots,\boldsymbol{Y}_1|\boldsymbol{Y}_0,\boldsymbol{Y}_{-1},\cdots,\boldsymbol{Y}_{-p+1}}\left(\boldsymbol{y}_T,\boldsymbol{y}_{T-1},\cdots,\boldsymbol{y}_1 \mid \boldsymbol{y}_0,\boldsymbol{y}_{-1},\cdots,\boldsymbol{y}_{-p+1};\boldsymbol{\beta}\right),$$

and $\hat{\boldsymbol{\beta}} \xrightarrow{p} \boldsymbol{\beta}$ *as* $T\to\infty$.

Theorem 10 *Suppose that* $\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_T$ *have the conditional joint probabil-*

ity density

$$f_{\boldsymbol{Y}_T,\boldsymbol{Y}_{T-1},\cdots,\boldsymbol{Y}_1|\boldsymbol{Y}_0,\boldsymbol{Y}_{-1},\cdots,\boldsymbol{Y}_{-p+1}}\left(\boldsymbol{y}_T,\boldsymbol{y}_{T-1},\cdots,\boldsymbol{y}_1 \mid \boldsymbol{y}_0,\boldsymbol{y}_{-1},\cdots,\boldsymbol{y}_{-p+1}\right)$$

$\beta \in \Theta$, *and all the assumptions of theorem* 9 *hold, then under the null*

$$H_0 : \boldsymbol{\beta} = \tilde{\boldsymbol{\beta}},$$

where $\tilde{\boldsymbol{\beta}}$ *is the restricted estimator*

$$LR = -2\log\left(L\left(\tilde{\boldsymbol{\beta}}\right)/L\left(\hat{\boldsymbol{\beta}}\right)\right) = 2\left(\log L\left(\hat{\boldsymbol{\beta}}\right) - \log L\left(\tilde{\boldsymbol{\beta}}\right)\right) \xrightarrow{d} \chi_p^2\left(\text{under } H_0\right),$$

where $L\left(\tilde{\boldsymbol{\beta}}\right)$ *and* $L\left(\hat{\boldsymbol{\beta}}\right)$ *denote the values of the log likelihood function at the restricted* ($\tilde{\boldsymbol{\beta}}$) *and the unrestricted* ($\hat{\boldsymbol{\beta}}$) *estimates respectively.*

Theorem 11 *If all the assumptions of theorem* 9 *hold, then under the null hypothesis that the restrictions are true*

$$LM = T^{-1}\boldsymbol{S}\left(\tilde{\boldsymbol{\beta}}\right)' \boldsymbol{I}\left(\tilde{\boldsymbol{\beta}}\right)^{-1} \boldsymbol{S}\left(\tilde{\boldsymbol{\beta}}\right) \xrightarrow{d} \chi_p^2\left(\text{under } H_0\right).$$

Theorem 12 *If all the conditions of theorem* 9 *hold and the null hypothesis is true, then*

$$Wald = T\left(\hat{\boldsymbol{\beta}} - \tilde{\boldsymbol{\beta}}\right)' \left[\boldsymbol{I}\left(\hat{\boldsymbol{\beta}}\right)^{-1}\right]^{-1} \left(\hat{\boldsymbol{\beta}} - \tilde{\boldsymbol{\beta}}\right) \xrightarrow{d} \chi_p^2\left(\text{under } H_0\right).$$

If each equation of the VAR mentioned in theorem 6 *is expressed in the form of Equation* (8), *and the null hypothesis is of the form* $\boldsymbol{R}\boldsymbol{\beta} = \boldsymbol{R}\tilde{\boldsymbol{\beta}}$, *the above Wald statistic can be written as follows:*

$$W = \sqrt{T}\left(\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{R}\tilde{\boldsymbol{\beta}}\right)' \left[\boldsymbol{R}\left[\hat{\Omega}_T \otimes \left(T^{-1}\sum_{t=1}^{T}\boldsymbol{x}_t\boldsymbol{x}_t'\right)^{-1}\right]\boldsymbol{R}'\right]^{-1} \sqrt{T}\left(\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{R}\tilde{\boldsymbol{\beta}}\right)$$
$$\xrightarrow{d} \chi_p^2\left(\text{under } H_0\right),$$

where $\hat{\boldsymbol{\beta}}$ *and* $\tilde{\boldsymbol{\beta}}$ *are the unrestricted and the restricted estimators respectively.*

6. Power and Size of the Test

Since the distribution under the alternative is unknown, therefore the power of *ST* has been estimated through Monte Carlo simulations as follows:

$$\widehat{Pow} = \frac{1}{N}\sum_{i=1}^{N} I\left(ST_i > c_\alpha\right),$$

where ST_i is the test statistic for the *i*th Monte Carlo sample, c_α is a given critical value, and *I* is the indicator function having value 1 if $ST_i > c_\alpha$ and 0 otherwise. The following data generating process has been used:

$$y_t = \mu + \phi_1 y_{t-1} + \epsilon_t,$$
$$\mu = 0,$$
$$\epsilon_t \sim N(0,1).$$

The values of parameters have been estimated through least squares with the help of methodology described earlier. A power plot for a range of alternatives starting from a unit root, *i.e.*, $\phi_1 = 1$, to $\phi_1 = 1.08$ for $N = 1000$ and $T = 100$ is shown in **Figure 2**. The Sufficient test has the minimum power for a unit root. The power increases as we move farther from a unit root. The null rejection probabilities are reported in **Table 3**. For the true values satisfying the null hypothesis, the test has the correct size.

7. Monte Carlo Simulations

Some specific examples are listed below in order to highlight the performance of the test in a variety of situations, *i.e.*, real, complex, single and multiple roots. The examples also illustrate how the test identifies the number of roots causing instability in the dynamic response. A comparison with the Augmented Dickey Fuller and ADF-GLS test is also provided for the unit root cases.

7.1. AR(1)

Suppose that the model is correctly specified, *i.e.*, we are estimating an AR(1) process. The null hypothesis for stability of an AR(1) process is as follows:

$$H_0 : \begin{bmatrix} 1-\phi_1 \\ 1+\phi_1 \end{bmatrix} > 0 \text{, or}$$

$$H_0 : -1 < \phi_1 < 1 .$$

Under the null hypothesis, the dynamic response is stable and the t-statistic follows an asymptotic normal distribution. The Sufficient test statistic is defined as follows:

$$ST = \frac{\sqrt{T}\left(\hat{\beta} - \tilde{\beta}\right)}{\sqrt{\hat{\sigma}^2 \left(T^{-1} \sum_{t=1}^{T} \boldsymbol{x}_t \boldsymbol{x}_t'\right)^{-1}}} .$$

The following data generating process has been used for generating data:

$$y_t = \mu + \phi_1 y_{t-1} + \epsilon_t ,$$

$$\epsilon_t \sim N(0,1) .$$

7.1.1. Case (a)

$$\mu = 0, \phi_1 = 1.2$$

For 1000 replications ($T = 100$), the number of rejections has been recorded, and the result is as follows:

	Reject	Not reject
ST	1000	

The null hypothesis is rejected for all runs in the simulation, and the estimated response is dynamically unstable.

7.1.2. Case (b)

$$\mu = 0, \phi_1 = 1$$

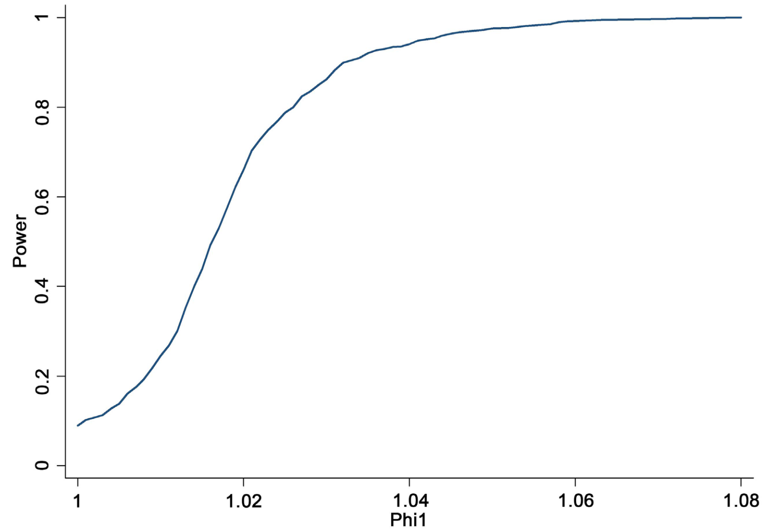

Figure 2. Power plot for the sufficient test for an AR(1) process.

Table 3. Empirical null rejection probabilities 5% level.

T = 100, rep = 1000, Two-Tailed Test of $H_0: -1 < \phi_1 < 1$		
DGP: $y_t = \mu + \phi_1 y_{t-1} + \epsilon_t, \epsilon_t \sim N(0,1)$		
S. No.	ϕ_1	Null Rejection Probability
1	0.60	0.045
2	0.70	0.043
3	0.80	0.057
4	0.90	0.052
5	0.91	0.050
6	0.92	0.053
7	0.93	0.061
8	0.94	0.058
9	0.95	0.060
10	0.96	0.059
11	0.97	0.066
12	0.98	0.063
13	0.99	0.059

For 1000 replications ($T = 100$), the number of rejections by the ST and the number of "No rejections" by the ADF test have been recorded, and the result is as follows:

	Reject	Not reject
ST	1000	

7.1.3. Case (c)

$$\mu = 0, \phi_1 = 0.99$$

For 1000 replications ($T = 100$), the number of "No rejections" by the ST and the number of rejections by the ADF test (which is optimal in the absence of an intercept and time trend in terms of power) have been recorded, and the result is as follows:

	Reject	Not reject
ST		962
ADF	47	

Although we cannot directly compare the power of ST with that of the ADF test as the null hypotheses are different, however, cases (c) and (d) somehow give a reflection of the power of ST as compared to that of the Augmented Dickey Fuller test. The ST has the minimum power in case (c) (a unit root case), *i.e.*, 0.089. As the roots move farther from a value of one, the power of ST increases. In contrast, in case (d), the ADF test has only a power of 0.047.

7.2. AR (2)

Suppose the model is correctly specified, *i.e.*, we are estimating an AR(2) process. The null hypothesis for stability of an AR(2) process is as follows:

$$H_0 : \begin{bmatrix} 1-\phi_1-\phi_2 \\ 2+2\phi_2 \\ 1+\phi_1-\phi_2 \end{bmatrix} > 0 \text{, or}$$

$$H_0 : \begin{bmatrix} -(1-\phi_2) < \phi_1 < (1-\phi_2) \\ -1 < \phi_2 < 1 \end{bmatrix}.$$

Under the null hypothesis, the dynamic response is stable and the Wald statistic follows a Chi squared distribution with two degress of freedom. The Sufficient test statistic is defined as follows:

$$ST = \sqrt{T}\left(\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{R}\tilde{\boldsymbol{\beta}}\right)' \left[\hat{\sigma}^2 \boldsymbol{R}\left(T^{-1}\sum_{t=1}^{T} \boldsymbol{x}_t \boldsymbol{x}_t'\right)^{-1} \boldsymbol{R}'\right]^{-1} \sqrt{T}\left(\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{R}\tilde{\boldsymbol{\beta}}\right)$$
$$\xrightarrow{d} \chi_2^2 \left(\text{under } H_0\right).$$

The following data generating process has been used for generating data:

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \epsilon_t,$$
$$\epsilon_t \sim N(0,1).$$

7.2.1. Case (a)

$$\mu = 0, \phi_1 = 0.85, \phi_2 = 0.3.$$

For 1000 replications ($T = 100$), the number of rejections has been recorded, and the result is as follows:

	Reject	Not reject
ST	1000	

The null hypothesis is rejected for all runs in the simulation, and the estimated response is dynamically unstable. In order to see how many roots are causing instability, a t-test is performed for all the three elements of the null hypothesis vector as follows:

$$H_0 : \begin{bmatrix} 1-\phi_1-\phi_2 \\ 2+2\phi_2 \\ 1+\phi_1-\phi_2 \end{bmatrix} > 0,$$

$$H_{01} : 1-\phi_1-\phi_2 > 0,$$

$$H_{02} : 2+2\phi_2 > 0,$$

$$H_{03} : 1+\phi_1-\phi_2 > 0.$$

For 1000 replications ($T = 100$), the number of rejections has been recorded, and the result is as follows:

	Reject
$t(H_{01})$	981
$t(H_{02})$	0
$t(H_{03})$	25

The t-statistic for mainly the first element of null hypothesis vector is rejected (which is equivalent to one sign change in the first column of the Routh array), which implies that only one root is causing instability in the AR(2) process. As the two roots are 1.118271 and −0.268271, therefore the test correctly determines the number of roots causing instability in the response.

7.2.2. Case (b)

$$\mu = 0, \phi_1 = 1.4, \phi_2 = -0.4.$$

For 1000 replications ($T = 100$), the number of rejections by the ST and the number of "No rejections" by the Augmented Dickey Fuller test have been recorded, and the result is as follows:

	Reject	Not reject
ST	310	
ADF		952

7.2.3. Case (c)

$$\mu = 0, \phi_1 = 1.2, \phi_2 = -0.21.$$

For 1000 replications ($T = 100$), the number of "No rejections" by the ST and the number of rejections by the Augmented Dickey Fuller test have been recorded, and the result is as follows:

	Reject	Not reject
ST		929
ADF	48	

Again the cases (c) and (d) provide an indirect comparison of the ST and the ADF test. The ST does not perform as badly for a unit root as the ADF test does for a root close to one.

7.2.4. Case (d)

$$\mu = 1, \phi_1 = 0.8, \phi_2 = -1.2.$$

This is a case of complex roots causing instability. For 1000 replications ($T = 100$), the number of rejections has been recorded, and the result is as follows:

	Reject	Not reject
ST	1000	

The null hypothesis is rejected for all runs in the simulation, and the estimated response is dynamically unstable. In order to see how many roots are causing instability, a t-test is performed for all the three elements of the null hypothesis vector as follows:

$$H_0 : \begin{bmatrix} 1-\phi_1-\phi_2 \\ 2+2\phi_2 \\ 1+\phi_1-\phi_2 \end{bmatrix} > 0,$$

$$H_{01} : 1-\phi_1-\phi_2 > 0,$$

$$H_{02} : 2+2\phi_2 > 0,$$

$$H_{03} : 1+\phi_1-\phi_2 > 0.$$

For 1000 replications ($T = 100$), the number of rejections has been recorded, and the result is as follows:

	Reject
$t(H_{01})$	0
$t(H_{02})$	1000
$t(H_{03})$	0

The t-statistic for the second element of null hypothesis vector is rejected (which is equivalent to two sign changes in the first column of the Routh array), which implies that two roots are causing instability in the AR(2) process. As the two roots are $0.4+1.019804i$ and $0.4-1.019804i$, therefore the test correctly determines the number of roots causing instability in the response.

7.2.5. Case (e)

$$\mu = 1, \phi_1 = 0.8, \phi_2 = -1.02.$$

This is a case of complex unit roots causing instability. For 1000 replications ($T = 100$), the number of rejections has been recorded, and the result is as follows:

	Reject	Not reject
ST	647	

The null hypothesis is rejected for 64.7 percent of the times, and the estimated response is dynamically unstable. In order to see how many roots are casuing instability, a t-test is performed for all the three elements of the null hypothesis vector as follows:

$$H_0 : \begin{bmatrix} 1-\phi_1-\phi_2 \\ 2+2\phi_2 \\ 1+\phi_1-\phi_2 \end{bmatrix} > 0,$$

$$H_{01} : 1-\phi_1-\phi_2 > 0,$$

$$H_{02} : 2+2\phi_2 > 0,$$

$$H_{03} : 1+\phi_1-\phi_2 > 0.$$

For 1000 replications ($T = 100$), the number of rejections has been recorded, and the result is as follows:

	Reject
$t(H_{01})$	0
$t(H_{02})$	1000
$t(H_{03})$	0

The t-statistic for the second element of null hypothesis vector is rejected. For a pair of unit roots in an AR(2) process, the third theorem of the Routh test is applicable, which implies that there are two roots which are causing instability in the AR(2) process. As the two roots are $0.4+0.927362i$ and $0.4-0.927362i$, therefore the test correctly determines the number of roots causing instability in the response.

7.2.6. Case (f)

$$\mu = 1, \phi_1 = 2, \phi_2 = -1.$$

This is a case of multiple unit roots. For 1000 replications ($T = 100$), the number of rejections by the ST and the number of "No rejections" by ADF test have been recorded, and the result is as follows:

	Reject	Not reject
ST	1000	
ADF		982

It is evident from the above table that ST has a higher power in case of multiple unit roots. In order to see how many roots are casuing instability, a t-test is performed for all the three elements of the null hypothesis vector as follows:

$$H_0:\begin{bmatrix}1-\phi_1-\phi_2\\2+2\phi_2\\1+\phi_1-\phi_2\end{bmatrix}>0,$$

$$H_{01}:1-\phi_1-\phi_2>0,$$

$$H_{02}:2+2\phi_2>0,$$

$$H_{03}:1+\phi_1-\phi_2>0.$$

For 1000 replications ($T=100$), the number of rejections has been recorded, and the result is as follows:

	Reject
$t(H_{01})$	1000
$t(H_{02})$	448
$t(H_{03})$	0

The t-statistic for the first two elements of the null hypothesis vector is rejected, which implies that there are two roots which are causing instability in the AR(2) process. As both the roots are equal to 1, therefore the test correctly determines the number of roots causing instability in the response.

7.3. VAR

Suppose we want to test the stability of the following VAR:

$$y_{1t}=\mu_1+\phi_{11}y_{1,t-1}+\phi_{12}y_{2,t-1}+\phi_{13}y_{3,t-1}+\epsilon_{1t},$$

$$y_{2t}=\mu_2+\phi_{21}y_{1,t-1}+\phi_{22}y_{2,t-1}+\phi_{23}y_{3,t-1}+\epsilon_{2t},$$

$$y_{3t}=\mu_3+\phi_{31}y_{1,t-1}+\phi_{32}y_{2,t-1}+\phi_{33}y_{3,t-1}+\epsilon_{3t},$$

$$\varepsilon_{1t},\varepsilon_{2t}\text{ and }\varepsilon_{3t}\sim N(0,1).$$

The null hypothesis for stability of the above VAR process is as follows (see appendix for detailed derivation of the null hypothesis):

$$H_0:\begin{bmatrix}1+A_0+A_1-A_2\\3-3A_0-A_1-A_2\\1-A_1-A_0^2-A_0A_2\\1-A_0+A_1+A_2\end{bmatrix}>0\text{; r}$$

$$H_0:\begin{bmatrix}a_{11}<\phi_{11}<b_{11}\\a_{12}<\phi_{12}<b_{12}\\a_{13}<\phi_{13}<b_{13}\\a_{21}<\phi_{21}<b_{21}\\a_{22}<\phi_{22}<b_{22}\\a_{23}<\phi_{23}<b_{23}\\a_{31}<\phi_{31}<b_{31}\\a_{32}<\phi_{32}<b_{32}\\a_{33}<\phi_{33}<b_{33}\end{bmatrix},$$

or $H_0:\boldsymbol{a}<vec\left(\Phi^{\mathrm{T}}\right)<\boldsymbol{b}$; where

$$\boldsymbol{a} = \begin{bmatrix} a_{11} \\ a_{12} \\ a_{13} \\ a_{21} \\ a_{22} \\ a_{23} \\ a_{31} \\ a_{32} \\ a_{33} \end{bmatrix}; \text{ and } \boldsymbol{b} = \begin{bmatrix} b_{11} \\ b_{12} \\ b_{13} \\ b_{21} \\ b_{22} \\ b_{23} \\ b_{31} \\ b_{32} \\ b_{33} \end{bmatrix}.$$

Under the null hypothesis, the dynamic response is stable and the Wald statistic follows a Chi squared distribution with nine degress of freedom. The Sufficient test statistic is defined as follows:

$$ST = \sqrt{T}\left(\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{R}\tilde{\boldsymbol{\beta}}\right)' \left[\boldsymbol{R}\left[\hat{\Omega}_T \otimes \left(T^{-1}\sum_{t=1}^{T} \boldsymbol{x}_t \boldsymbol{x}_t'\right)^{-1}\right]\boldsymbol{R}'\right]^{-1} \sqrt{T}\left(\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{R}\tilde{\boldsymbol{\beta}}\right)$$

$$\xrightarrow{d} \chi_9^2 \left(\text{under } H_0\right),$$

$$\hat{\Omega}_T = T^{-1}\sum_{t=1}^{T} \hat{\epsilon}_t \hat{\epsilon}_t',$$

$$\hat{\epsilon}_t' = \left[\hat{\epsilon}_{1t} \quad \hat{\epsilon}_{2t} \quad \hat{\epsilon}_{3t}\right].$$

7.3.1. Case (a)

$$\mu = 1, \phi_{11} = 0.5, \phi_{12} = 0.4, \phi_{13} = 0.3, \phi_{21} = 0.3,$$
$$\phi_{22} = 0.4, \phi_{23} = 0.2, \phi_{31} = 0.1, \phi_{32} = 0.1, \phi_{33} = 0.99.$$

For 1000 replications ($T = 100$), the number of rejections by ST has been recorded, and the result is as follows:

	Reject	Not reject
ST	1000	

The null hypothesis is rejected for all 1000 runs in the simulation, and the estimated response is dynamically unstable. In order to see how many roots are casuing instability, a Wald test is performed for all the four elements of the null hypothesis vector as follows:

$$H_0 : \begin{bmatrix} 1 + A_0 + A_1 - A_2 \\ 3 - 3A_0 - A_1 - A_2 \\ 1 - A_1 - A_0^2 - A_0 A_2 \\ 1 - A_0 + A_1 + A_2 \end{bmatrix} > 0,$$

$$H_{01} : 1 + A_0 + A_1 - A_2 > 0,$$

$$H_{02} : 3 - 3A_0 - A_1 - A_2 > 0,$$

$$H_{03} : 1 - A_1 - A_0^2 - A_0 A_2 > 0,$$

$$H_{04} : 1 - A_0 + A_1 + A_2 > 0.$$

For 1000 replications ($T = 100$), the number of rejections has been recorded, and the result is as follows:

	Reject
$Wald(H_{01})$	1000
$Wald(H_{02})$	0
$Wald(H_{03})$	0
$Wald(H_{04})$	0

The Wald statistic for only the first element of null hypothesis vector is rejected (which is equivalent to one sign change in the first column of the Routh array), which implies that only one root is causing instability in the VAR. As the three roots are 1.137951, 0.652049 and 0.1, therefore the test correctly determines the number of roots causing instability in the response.

7.3.2. Case (b)

$$\mu = 1, \phi_{11} = 0.1, \phi_{12} = 0.1, \phi_{13} = 0.1, \phi_{21} = 0.1,$$
$$\phi_{22} = 0.1, \phi_{23} = 0.1, \phi_{31} = 0.1, \phi_{32} = 0.1, \phi_{33} = 0.98.$$

This is a case of a real unit root. For 1000 replications ($T = 100$), the number of rejections by ST has been recorded, and the result is as follows:

	Reject	Not reject
ST	702	

The null hypothesis is rejected for 702 runs in the simulation, and the estimated response is dynamically unstable. In order to see how many roots are casuing instability, a Wald test is performed for all the four elements of the null hypothesis vector as follows:

$$H_0 : \begin{bmatrix} 1 + A_0 + A_1 - A_2 \\ 3 - 3A_0 - A_1 - A_2 \\ 1 - A_1 - A_0^2 - A_0 A_2 \\ 1 - A_0 + A_1 + A_2 \end{bmatrix} > 0,$$

$$H_{01} : 1 + A_0 + A_1 - A_2 > 0,$$
$$H_{02} : 3 - 3A_0 - A_1 - A_2 > 0,$$
$$H_{03} : 1 - A_1 - A_0^2 - A_0 A_2 > 0,$$
$$H_{04} : 1 - A_0 + A_1 + A_2 > 0.$$

For 1000 replications ($T = 100$), the number of rejections has been recorded, and the result is as follows:

	Reject
$Wald(H_{01})$	995
$Wald(H_{02})$	0
$Wald(H_{03})$	0
$Wald(H_{04})$	0

The Wald statistic for only the first element of null hypothesis vector is rejected (which is equivalent to one sign change in the first column of the Routh array), which implies that only one root is causing instability in the VAR. As the three roots are 10.175 and 0, therefore the test correctly determines the number of roots causing instability in the response.

7.3.3. Case (c)

$$\mu = 1, \phi_{11} = 0.7, \phi_{12} = -0.5, \phi_{13} = -0.6, \phi_{21} = 0.4,$$
$$\phi_{22} = 0.1, \phi_{23} = -0.2, \phi_{31} = 0.7, \phi_{32} = 0.8, \phi_{33} = 0.35.$$

This is a case of complex unit roots. For 1000 replications ($T = 100$), the number of rejections by ST has been recorded, and the result is as follows:

	Reject	Not reject
ST	622	

The null hypothesis is rejected for 622 runs in the simulation, and the estimated response is dynamically unstable. In order to see how many roots are casuing instability, a Wald test is performed for all the four elements of the null hypothesis vector as follows:

$$H_0 : \begin{bmatrix} 1 + A_0 + A_1 - A_2 \\ 3 - 3A_0 - A_1 - A_2 \\ 1 - A_1 - A_0^2 - A_0 A_2 \\ 1 - A_0 + A_1 + A_2 \end{bmatrix} > 0,$$

$$H_{01} : 1 + A_0 + A_1 - A_2 > 0,$$

$$H_{02} : 3 - 3A_0 - A_1 - A_2 > 0,$$

$$H_{03} : 1 - A_1 - A_0^2 - A_0 A_2 > 0,$$

$$H_{04} : 1 - A_0 + A_1 + A_2 > 0.$$

For 1000 replications ($T = 100$), the number of rejections has been recorded, and the result is as follows:

	Reject
$Wald(H_{01})$	0
$Wald(H_{02})$	0
$Wald(H_{03})$	643
$Wald(H_{04})$	0

The Wald statistic for only the third element of null hypothesis vector is rejected and the third theorem of the Routh test is applicable, which implies that there are two roots which are causing instability in the VAR. As the three roots are $0.126425, 0.511788 + 0.859458i$ and $0.511788 - 0.859458i$, therefore the test correctly determines the number of roots causing instability in the response.

8. Empirical Application

This empirical application is quite similar to the one in Marco Del Negro, and Frank Schorfheide (2004). The three variables used in the trivariate VAR (with a lag length of one quarter) are the US interest rates (effective federal funds rate for the first month of the quarter) and the quarterly data of percentage changes in real output growth and inflation for the period 1955:I to 2013:II. The data for real output growth come from the Bureau of Economic Analysis; the data for inflation come from the Bureau of Labor Statistics and the EFFR data source (not seasonally adjusted) is Board of Governors of the Federal Reserve System.

	%∇P_1	EFFR_1	%∇GDP_1
%∇P (UNRES)	−0.074	−0.001	−0.004
%∇P (RES)	−0.074	−0.001	−0.004
	(0.065)	(0.028)	(0.027)
EFFR (UNRES)	−0.029*	0.95	0.026
EFFR (RES)	−0.029*	0.95	0.026
	(0.049)	(0.021)	(0.02)
%∇GDP (UNRES)	−0.008	−0.106	−0.003
%∇GDP (RES)	−0.008	−0.106	−0.003
	(0.157)	(0.069)	(0.064)
Wald	5.90528e−20		
*significant at 5% level			

The null hypothesis is not rejected and the response of the VAR is dynamically stable.

9. Conclusions

In this paper, a *sufficient test* for dynamic stability (in the context of the roots of the characteristic polynomial) of a univariate as well as a multivariate time series has been proposed, which may test for all kinds of roots (positive and negative real unit roots, complex unit roots and roots inside the unit circle whether single or multiple) causing instability in the dynamic response. The test is much simpler in its application as the response is dynamically stable under the null. The test also indicates the number of roots causing instability in the dynamic response. In order to formulate the null hypothesis, Routh Hurwitz stability criterion (a mathematical test) is exploited which provides a necessary and sufficient condition for the stability of a dynamic response. To use the Routh stability test in the discrete data framework, bilinear transformation has been used which maps the inside of the unit circle of the z-plane into the left half of the w-plane. In order to find the restricted estimators which satisfy the Routh Hurwitz stability criterion (given the data), an algorithm for minimization of the regression objective function subject to the inequality constraints has been devised.

For the sufficient test, *a*, *t*, *LR*, *LM* and a *Wald* statistic are used. The *t* statistic follows an asymptotic normal distribution and *LR*, *LM* and *Wald* follow an asymptotic chi squared distribution under the null with degrees of freedom equal to the number of restrictions, when the model is correctly specified. In case of serial correlation, robust test in Nawaz (2020) has been proposed

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Dickey, D.A. and Fuller, W.A. (1979) Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association*, **74**, 427-431. https://doi.org/10.1080/01621459.1979.10482531

[2] Bhargava, A. (1986) On the Theory of Testing for Unit Roots in Observed Time Series. *Review of Economic Studies*, **53**, 369-384. https://doi.org/10.2307/2297634

[3] Park, J.Y. and Choi, B. (1988) A New Approach to Testing for a Unit Root. Working Paper 88. Center for Analytic Economics, Cornell University, New York, 23.

[4] Guilkey, D.K. and Schmidt, P. (1989) Extended Tabulations for Dickey-Fuller Tests. *Economics Letters*, **31**, 355-357. https://doi.org/10.1016/0165-1765(89)90028-1

[5] Kim, K. and Schmidt, P. (1990) Some Evidence on the Accuracy of Phillips-Perron Tests Using Alternative Estimates of Nuisance Parameters. *Economics Letters*, **34**, 345-350. https://doi.org/10.1016/0165-1765(90)90143-O

[6] Kwiatkowski, D. and Schmidt, P. (1990) Dickey-Fuller Tests with Trend. *Communications in Statistics-Theory and Methods*, **19**, 3645-3656. https://doi.org/10.1080/03610929008830402

[7] Schmidt, P. (1990) Dickey-Fuller Tests with Drift. In: Fomby, T.B. and Rhodes, J., Eds., *Advances in Econometrics*: *A Research Annual*. *Volume* 8: *Cointegration, Spurious Regression, and Unit Roots*, JAI Press, Greenwich, 161-200.

[8] Guilkey, D.K. and Schmidt, P. (1991) An Interpolation Test for a Unit Root in the Presence of Drift. *Biometrika*, **78**, 809-815. https://doi.org/10.1093/biomet/78.4.809

[9] Schmidt, P. and Lee, J. (1991) A Modification of the Schmidt-Phillips Unit Root Test. *Economics Letters*, **36**, 285-289. https://doi.org/10.1016/0165-1765(91)90034-I

[10] Stock, J.H. (1991) Confidence Intervals for the Largest Autoregressive Root in U.S. Macroeconomic Time Series. *Journal of Monetary Economics*, **28**, 435-459. https://doi.org/10.1016/0304-3932(91)90034-L

[11] Kwiatkowski, D., Phillips, P., Schmidt, P. and Shin, Y. (1992) Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root? *Journal of Econometrics*, **54**, 154-178. https://doi.org/10.1016/0304-4076(92)90104-Y

[12] Perron, P. and Vogelsang, T.J. (1992) Testing for a Unit Root in a Time Series with a Changing Mean: Corrections and Extensions. *Journal of Business and Economic Statistics*, **10**, 467-470. https://doi.org/10.1080/07350015.1992.10509923

[13] Schmidt, P. and Phillips, P.C. (1992) LM Tests for a Unit Root in the Presence of Deterministic Trends. *Oxford Bulletin of Economics and Statistics*, **54**, 257-287.

https://doi.org/10.1111/j.1468-0084.1992.tb00002.x

[14] Shin, Y. and Schmidt, P. (1992) The KPSS Stationarity Test as a Unit Root Test. *Economics Letters*, **38**, 387-392. https://doi.org/10.1016/0165-1765(92)90023-R

[15] Hwang, J. and Schmidt, P. (1993) On the Power of Point Optimal Tests of the Trend Stationarity Hypothesis. *Economics Letters*, **43**, 143-147. https://doi.org/10.1016/0165-1765(93)90027-A

[16] Kim, K. and Schmidt, P. (1993) Unit Root Tests with Conditional Heteroskedasticity. *Journal of Econometrics*, **59**, 287-300. https://doi.org/10.1016/0304-4076(93)90027-3

[17] Schmidt, P. (1993) Some Results on Testing for Stationarity Using Data Detrended in Differences. *Economics Letters*, **41**, 1-6. https://doi.org/10.1016/0165-1765(93)90101-H

[18] Lee, J. and Schmidt, P. (1994) Unit Root Tests Based on Instrumental Variables Estimation. *International Economic Review*, **35**, 449-462. https://doi.org/10.2307/2527063

[19] Hwang, J. and Schmidt, P. (1996) Alternative Methods of Detrending and the Power of Unit Root Tests. *Journal of Econometrics*, **71**, 227-248. https://doi.org/10.1016/0304-4076(94)01702-6

[20] Lee, D. and Schmidt, P. (1996) On the Power of the KPSS Test of Stationarity against Fractionally Integrated Alternatives. *Journal of Econometrics*, **73**, 285-302. https://doi.org/10.1016/0304-4076(95)01741-0

[21] Vogelsang, T.J. and Perron, P. (1998) Additional Tests for a Unit Root Allowing for a Break in the Trend Function at an Unknown Time. *International Economic Review*, **39**, 1073-1100. https://doi.org/10.2307/2527353

[22] Xiao, Z. and Phillips, P.C. (1998) An ADF Coefficient Test for a Unit Root in ARMA Models of Unknown Order with Empirical Applications to the US Economy. *The Econometrics Journal*, **1**, 27-43. https://doi.org/10.1111/1368-423X.12016

[23] Vogelsang, T.J. (1999) Two Simple Procedures for Testing for a Unit Root When There Are Additive Outliers. *Journal of Time Series Analysis*, **20**, 237-252. https://doi.org/10.1111/1467-9892.00135

[24] De Jong, R.M., Amsler, C. and Schmidt, P. (2007) A Robust Version of the KPSS Test Based on Indicators. *Journal of Econometrics*, **137**, 311-333. https://doi.org/10.1016/j.jeconom.2006.01.001

[25] Amsler, C., Schmidt, P. and Vogelsang, T.J. (2009) The KPSS Test Using Fixed-*b* Critical Values: Size and Power in Highly Autocorrelated Time Series. *Journal of Time Series Econometrics*, **1**. https://doi.org/10.2202/1941-1928.1027

[26] Su, J.-J., Amsler, C. and Schmidt, P. (2012) A Note on the Size of the KPSS Unit Root Test. *Economics Letters*, **117**, 697-699. https://doi.org/10.1016/j.econlet.2012.08.019

[27] Vogelsang, T.J. and Wagner, M. (2013) A Fixed-*b* Perspective on the Phillips-Perron Unit Roottests. *Econometric Theory*, **29**, 609-628. https://doi.org/10.1017/S0266466612000485

[28] Zeren, F. and Kızılkaya, F. (2021) A New Combination of Fourier Unit Root Tests: A PPP Application for Fragile Economies. *Applied Economics Letters*, **28**, 1707-1711. https://doi.org/10.1080/13504851.2020.1851647

[29] Brown, R.L., Durbin, J. and Evans, J.M. (1975) Techniques for Testing the Constancy of Regression Relationships over Time. *Journal of the Royal Statistical Society B*, **37**, 149-163. https://doi.org/10.1111/j.2517-6161.1975.tb01532.x

[30] Chow, G.C. (1960) Tests of Equality between Sets of Coefficients in Two Linear Regressions. *Econometrica*: *Journal of the Econometric Society*, **28**, 591-605. https://doi.org/10.2307/1910133

[31] Hansen, B.E. (1992) Testing for Parameter Instability in Linear Models. *Journal of Policy Modeling*, **14**, 517-533. https://doi.org/10.1016/0161-8938(92)90019-9

[32] Andrews, D.W.K. and Ploberger, W. (1994) Optimal Tests When a Nuisance Parameter Is Present Only under the Alternative. *Econometrica*, **62**, 1383-1414. https://doi.org/10.2307/2951753

[33] Bierens, H.J. (2001) Complex Unit Roots and Business Cycles: Are They Real? *Econometric Theory*, **17**, 962-983. https://doi.org/10.1017/S0266466601175055

[34] Franco, C.A. and Collado, J. (2017) Comparison on Sufficient Conditions for the Stability of Hill Equation: An Arnold's Tongues Approach. *Applied Mathematics*, **8**, 1481-1514. https://doi.org/10.4236/am.2017.810109

[35] Schrijver, A. (1998) Theory of Linear and Integer Programming. John Wiley & Sons, Hoboken.

[36] Nawaz, N. (2020) Robust Inference by Sub-Sampling. *Journal of Quantitative Economics*, **18**, 657-681. https://doi.org/10.1007/s40953-020-00207-x