DISCUSSION PAPERS IN ECONOMICS

Working Paper No. 23-01

A Multivariate Autoregressive Distributed Lag unit Root Test

Robert McNown
University of Colorado Boulder

July 27, 2023

Revised July 27, 2024

Department of Economics

University of Colorado Boulder
Boulder, Colorado 80309

A MULTIVARIATE AUTOREGRESSIVE DISTRIBUTED LAG UNIT ROOT TEST

CHUNG YAN SAM[a*], ROBERT MCNOWN[b], SOO KHOON GOH[a], KIM-LENG GOH[c]
[a] Centre for Policy Research & International Studies, Universiti Sains Malaysia, 11800 Penang, Malaysia
[b] University of Colorado, Boulder, CO, 256 UCB, 80309, United States
[c] Faculty of Business and Economics, University of Malaya, 50603 Kuala Lumpur, Malaysia

Abstract

We propose a unit root test using the multivariate ARDL framework. This new test yields higher power properties compared to the existing multivariate unit root tests based on Covariate Augmented Dickey-Fuller (CADF) as well as a few commonly used univariate unit root tests. The main advantage of the new test over the CADF test is its consideration of possible cointegration relationships between the variable of interest and the explanatory variables in the process of testing for unit roots while the latter does not allow cointegration. Therefore, the proposed ARDL unit root test avoids model misspecification by imposing a valid common factor restriction to improve the power of the test. Several sets of experiments for size and power are conducted to check the reliability and robustness of the test. The experiments also discover that univariate unit root tests face serious size distortions in testing a cointegrated process. An empirical example using the proposed multivariate ARDL unit root test is demonstrated in this paper.

JEL Code: C3

Keywords: Multivariate Unit Root Test; ARDL; Cointegration; Bootstrap; Model Mispecification; Size and Power Analysis

1. Introduction

In applied economic research, it is recommended to pre-test the series for unit roots to decide on the next step in the analysis. However, it is commonly acknowledged that existing unit root tests are often trapped into inconclusive results due to size distortion and low power problems. Inconclusive results may lead researchers to make arbitrary decisions based on prior evidence or theoretical arguments. This strategy may lead to an incorrect specification in the subsequent analysis, which would adversely affect the estimator properties and therefore implications for economic policies.

Currently, conventional unit root testing applications generally employ univariate methods. A weakness of univariate frameworks is that they ignore important variables that explain the variable of interest. Consequently, there is a loss of information from related time series, producing low levels of fit in the test equation, higher variances of parameter estimates, and a loss of power. The notorious low power property in unit root tests could be due to the simplicity of their framework.

Hansen (1995) is the first to address the problems of the univariate unit root tests through a multivariate procedure, advocating the addition of related variables to the univariate tests to improve their power. The Covariate Augmented Dickey-Fuller (CADF) test proposed by Hansen (1995) includes related covariates (explanatory variables) in the Augmented Dickey-Fuller (ADF) equation to improve the fit of the regression. Results have shown substantial gains in power over the univariate ADF test. Although the multivariate unit root test methodology is not new, it is uncommon. A search of the literature shows that only a few approaches based on the multivariate framework, for instance, Phillips and Durlauf (1986), Flores et al. (1996), Juhl and Xiao (2003), Elliott and Jansson (2003), and Kurozumi et al. (2012). These works represent innovations that are based on the CADF framework.

Building on the CADF framework, the autoregressive distributed lag (ARDL) model is a more comprehensive multivariate framework that considers a wider range of dynamics for inferring unit root properties with greater power in the tests. This is due to the inclusion of the lagged explanatory variables in levels in the ARDL framework, that are not found in CADF equation. The omission of lagged level explanatory variables makes the CADF test exclude the possibility of cointegration between the testing series and its explanatory variables. Kremer (1992) states that the error correction model (ECM, or equivalently the subject of interest here, the ARDL model) could yield a more powerful test than the Dickey-Fuller test when cointegration exists because the latter imposes an invalid common factor restriction. Pesaran and Shin (1999) show that the ARDL-based OLS estimator for the long-run parameters (determining stationarity) is super-consistent with a faster convergent rate if there is cointegration. Thus, the ARDL test could yield better inferences than the CADF test when cointegration exists. In addition, by including appropriate lagged changes of variables, the residual serial correlation problem can be resolved without affecting the asymptotic properties of the OLS estimator (Pesaran and Shin, 1999).

This paper intends to add to the literature by introducing a new unit root test using the multivariate ARDL framework. This ARDL unit root test adopts the framework proposed for cointegration testing by Pesaran et al. (2001) (PSS henceforth). The existing multivariate unit root tests are primarily based on the CADF frameworks that are typically misspecified when cointegration exists. In testing for unit roots, the ARDL framework provides a more comprehensive approach that considers possible cointegration relationships to avoid model misspecification from imposing an invalid common factor. Our proposed new test relies on the t-test for the significance of lagged level of the dependent variable and the F-test for the joint significance of the lagged levels of independent variables. The flexibility of the ARDL framework helps to determine the best fit of the regression describing the dynamics of the dependent variable. More importantly, the test allows the related explanatory variables to be endogenous where there is contemporaneous correlation and cross-autocorrelation between the error and lagged changes of the variables. When the integration order of the explanatory variables is $I(1)$, the limiting distributions of the statistics under the null are complicated. It is a mixture of non-standard and standard distributions. The limiting distributions of the tests are uncertain and depend on nuisance parameters, which are due to the contemporaneous correlations between the cross-equation errors and the correlations between the ARDL equation error and the explanatory variables. Therefore, to determine the underlying distribution effectively during the hypothesis testing, the bootstrap technique is used. The underlying distribution can be characterized by bootstrapping because the bootstrap distribution is generated using the information of each empirical data set. By determining the underlying distribution, bootstrap is known to correct size distortions and improve the power of the tests. Numerous sets of simulation experiments are conducted to examine the performance of this new test with finite samples, and a comparison with the conventional unit root tests is provided. From the experiments, we found that the univariate unit root tests fail to capture the unit root that arises from the nonstationary covariates when the process is cointegrated. Hence, these tests face serious size distortions in rejecting the null hypothesis of unit root.

2. The Methodology of Multivariate Unit Root Tests

Today, many unit root tests have been developed to determine the integration order of a time series y_t. Generally, these tests are univariate and they share common weaknesses, namely, size distortion and low power properties. These unit root tests have low power against $I(0)$ alternatives that are close to being $I(1)$, or have size distortions when the DGP is not an autoregressive (AR) representation. For instance, Schwert (1989) finds that if Δy_t is an ARMA process with a large and negative moving average (MA) component, then the ADF and Phillips-Perron (PP) tests are severely size distorted. The well-known power weakness of these tests has led to the development of more powerful alternatives, and the multivariate tests are one strategy towards this end.

In time series analysis, a series' DGP may be a more complex generating mechanism, including multivariate processes. In macroeconomics a multivariate framework is more informative than a univariate model, and this framework could provide greater explanatory power of the variable in question. In 1995, Hansen (1995) published 'Rethinking the Univariate Approach to Unit Root Testing,' emphasizing this point. By including correlated stationary covariates (explanatory variables) in the unit root testing equation, he finds this increases the power of the test substantially. To be specific, Hansen begins with the univariate Augmented Dickey-Fuller regression, $\Delta y_t = \beta_1 y_{t-1} + \sum_{i=1}^{p-1} \delta_i \Delta y_{t-i} + u_t$. By adding related stationary covariates, Δx_t, the test equation becomes $\Delta y_t = \beta_1 y_{t-1} + \sum_{i=1}^{p-1} \delta_i \Delta y_{t-i} + \sum_{j=0}^{q-1} \phi_j \Delta x_{t-j} + u_t$. Then, the t-statistic on the estimated coefficient β_1 provides Hansen's Covariates Augmented Dickey-Fuller (CADF) test for a unit root in y_t. Elliott and Jansson (2003) acknowledge the importance of including related covariates into the regression, but they found that CADF is not the point optimal test in general. They then extend Hansen's approach by introducing a feasible point optimal (CPT) test based on VAR models. The CPT approach gains more power than the CADF approach but at a cost of a small size distortion (see Lupi, 2010).

Hansen's work shows that a multivariate framework for unit root can improve the precision of estimators, producing test statistics that are more powerful than univariate alternatives. Building on this principle, the ARDL framework is proposed for multivariate unit root testing. The advantage of the ARDL model is that it is capable of accommodating a wider range of DGPs, in particular the existence of cointegration between y_t and its covariates. Note that Hansen's CADF framework is similar to the ARDL framework of PSS: $\Delta y_t = \beta_1 y_{t-1} + \beta_2 x_{t-1} + \sum_{i=1}^{p-1} \delta_i \Delta y_{t-i} + \sum_{j=0}^{q-1} \phi_j \Delta x_{t-j} + u_t$, but it imposes a restriction that the cointegrating vector is (1, 0). In other words, it rules out the possibility of cointegration between y_t and x_t or the absence of long-run effect between y_t and the covariates x_t. Kremer (1992) discusses how the ECM could yield more powerful tests than the Dickey-Fuller test when cointegration exists because the latter imposes an invalid common factor restriction. This issue applies to the CADF framework. The CADF test imposes an invalid common factor restriction if y_t has a cointegrating relationship with x_t, adversely affecting its testing power.

Pesaran and Shin (1999) and PSS developed a cointegration test based on the ARDL framework. In presenting the cointegration test, they consider the joint significance test of the lagged levels of dependent and independent variables as well as the *t*-test on the lagged level of the dependent variable. PSS prove that these test statistics are consistent and follow nonstandard distributions regardless of whether the underlying regressors are purely *I*(0), purely *I*(1) or mutually cointegrated. To improve the PSS ARDL cointegration test, McNown et al. (2018) propose a complementary test on the lagged levels of independent variables using a bootstrap method. The introduction of this complimentary test allows the use of ARDL framework to test the stationarity of a process. Through a similar procedure by using the *t*-test to test the significance of the lagged level of the dependent variable and the complementary *F*-test for the joint significance of lagged levels of independent variables, we can infer the representation of y_t as stationarity or integrated. A stationary process is concluded if the lagged level dependent variable is significant but the lagged level independent variables are jointly insignificant. Otherwise, y_t is nonstationary. This is shown in the following section.

The limiting distributions of the *t*- and *F*-tests proposed by McNown et al. (2018) under the null are uncertain. They depend on the dimension (number of explanatory variables) and the nuisance parameters of contemporaneous correlation of the errors and cross-autocorrelation of the error and explanatory variables. It is difficult to determine the distribution without prior knowledge of these nuisance parameters. However, the bounds testing procedure, as in the ARDL cointegration test, is not conclusive because the test statistic(s) could fall between the bounds for the critical values. Thus, to resolve the issue of uncertain distributions under the null of the tests, McNown, et al. (2018) propose the bootstrap method. It is common to use bootstrap for unit root testing to improve its performance, as in Chang and Park (2003), Zou and Politis (2019), Hansen and Racine (2018), Cavaliere and Taylor (2008), Davidson (2007), Park (2003), and others. Most importantly, Chang et al. (2017) demonstrate that the bootstrap method works effectively to improve the CADF approach, while McNown et al. (2018) show that bootstrap is useful in the ARDL test.

It is rare to see unit root testing with a multivariate framework. Exceptions are Phillips and Durlauf (1996), Flores et al. (1996), Juhl and Xiao (2003), Elliott and Jansson (2003), and Kurozumi et al. (2012). These approaches underscore the importance of including covariates in developing new unit root tests. The ARDL framework uses a single equation that allows the lag length of each of its explanatory variables to be different. It also covers a wider range of DGPs, including cointegrating relationships.

3. The Model Framework, Statistical Distributions, and Testing Procedures

3.1 The ARDL Model and Assumptions

Let $\mathbf{z}_t$ denote a (1 + *k*)-vector random process that can be partitioned into $\mathbf{z}_t = \left(y_t, \mathbf{x}_t\right)$. The DGP of $\mathbf{z}_t$ is the VAR model of order *p*:

$$\mathbf{\Phi}\left(L\right)\left(\mathbf{z}_t - \boldsymbol{\mu} - \boldsymbol{\gamma} t\right) = \boldsymbol{\varepsilon}_t, \quad t = 1, 2, ...$$

where *L* is the lag operator, $\boldsymbol{\mu}$ and $\boldsymbol{\gamma}$ are unknown (1 + *k*)-vectors of intercept and trend coefficients, $\mathbf{\Phi}\left(L\right) = \mathbf{I}_{1+k} - \sum_{i=1}^{p} \mathbf{\Phi}_i \mathbf{z}^i$ is a (1+*k*, 1+*k*) matrix lag polynomial, $\mathbf{I}_{1+k}$ is an identity matrix of order (1+*k*), and $\mathbf{\Phi}_i$ are (1+*k*, 1+*k*) matrices of unknown coefficients (see PSS). Instead of the VAR model, the DGP of y_t and $\mathbf{x}_t$ can also be represented in the form of multivariate conditional ECM specification as follows:

$$\Delta y_t = c_1 + c_2 t + \beta_1 y_{t-1} + \boldsymbol{\beta}_2 \mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta y_{t-i} + \sum_{j=1}^{q-1} \mathbf{\Phi}_j^{'} \Delta \mathbf{x}_{t-i} + \boldsymbol{\omega}^{'} \Delta \mathbf{x}_t + u_t, \tag{1}$$

$$\Delta \mathbf{x}_t = c_{x,1} + c_{x,2} t + \mathbf{\Pi}_x \mathbf{x}_{t-1} + \sum_{k=1}^{r-1} \phi_{x,i} \Delta y_{t-i} + \sum_{l=1}^{s-1} \mathbf{\Phi}_{x,j}^{'} \Delta \mathbf{x}_{t-i} + \mathbf{u}_{x,t}, \tag{2}$$

where $\mathbf{x}_t$ **is a** $k \times 1$ **vector of *I*(1) forcing variables** and $\Delta \mathbf{x}_t$ is a set of first-differenced stationary variables. The single-equation (1) is the focus and the subject of the study. Equations (1) and (2) together form a system of equations that is identical to the system in the PSS ARDL bounds test for cointegration. However, instead of the cointegration test, it is now used for unit root testing. For simplicity, we refer to the conditional model (1) as ARDL. Note that this single-equation ARDL framework (1) is similar to a generalized Dickey-Fuller regression and Hansen's CADF regression. The LHS of the regression is the first difference of the testing series y_t and the RHS of the regression consists of deterministic term c_1 and t, the long-run multipliers $\boldsymbol{\beta} = \left(\beta_1, \boldsymbol{\beta}_2\right)'$, the short-run multipliers $\mathbf{\Phi}_j$, and $\boldsymbol{\omega} = \mathbf{\Omega}_{xx}^{-1} \boldsymbol{\omega}_{xy}$ where the variance-covariance matrix for the system of equations (1) and (2) is $\mathbf{\Omega} = \begin{pmatrix} \omega_{yy} & \boldsymbol{\omega}_{yx} \\ \boldsymbol{\omega}_{xy} & \mathbf{\Omega}_{xx} \end{pmatrix}$. Similar assumptions in the PSS ARDL framework apply to equation (1) to ensure the stability of the model:

Assumption 1: The roots of $\left| \mathbf{I}_{1+k} - \sum_{i=1}^{p} \mathbf{\Phi}_i \mathbf{z}^i \right| = 0$ are either outside the unit circle $\left|z\right| = 1$ or satisfy $z = 1$.

Assumption 2: The vector error process $\mathbf{u}_t = (u_t, \mathbf{u}_{x,t})$ is $IN(\mathbf{0}, \mathbf{\Omega})$, $\mathbf{\Omega}$ is positive definite.

Assumption 3: The $\Delta \mathbf{x}_t$ is a vector of stationary processes.

Assumption 4: No feedback effect from the level of y_t onto $\mathbf{x}_t$.

Assumption 1 permits the possibility of $I(2)$ if roots for y_t and $\mathbf{x}_t$ are both unity but excludes the possibility of seasonal and explosive roots. Assumption 3 allows a degree of contemporaneous correlation between u_t and $\mathbf{u}_{x,t}$ with $u_t = \boldsymbol{\omega}_{yx} \mathbf{\Omega}_{xx}^{-1} \mathbf{u}_{x,t} + \xi_t$, $\xi_t \sim IN(0, \omega_{uu})$ and ξ_t is independent of $\mathbf{u}_{x,t}$. Note that equation (2) complies with Assumption 4, which restricts the level effect from y_t impacting $\mathbf{x}_t$ but it allows an effect from Δy_t. In addition, equation (2) also allows the regressors, $\mathbf{x}_t$, to be cointegrated among themselves. These situations allow a level of endogeneity in $\mathbf{x}_t$. Under all these assumptions, the OLS estimators for the parameters in equation (1) are consistent. [1] The estimated coefficients $\hat{\boldsymbol{\beta}} = (\hat{\beta}_1, \hat{\boldsymbol{\beta}}_2)'$ are super-consistent while the coefficients $\hat{\mathbf{\Phi}}_j$ are $\sqrt{T}$-consistent (see discussion in Pesaran and Shin, 1999, and PSS). The ARDL model (1) has the same properties as the PSS framework. We infer the unit root property of y_t through analysis of the long-run multipliers or the coefficients $\boldsymbol{\beta}$ in equation (1).

3.2 Possible Cases

Consider the ADF and CADF frameworks:

ADF: $$\Delta y_t = c_1 + c_2 t + \beta_1 y_{t-1} + \sum_{i=1}^{p-1} \beta_i \Delta y_{t-i} + u_t, \tag{3}$$

CADF: $$\Delta y_t = c_1 + c_2 t + \beta_1 y_{t-1} + \sum_{i=1}^{p-1} \beta_i \Delta y_{t-i} + \sum_{j=0}^{q-1} \beta_j \Delta x_{t-j} + u_t. \tag{4}$$

To test for the presence of a unit root in y_t, the ADF and CADF depend on the significance of the term β_1. If β_1 is insignificantly different from zero , it indicates y_t is an $I(1)$ process; If β_1 is significantly different from zero, it indicates y_t is an $I(0)$ process. However, as mentioned in Section 2, both frameworks (3) and (4) are restrictive and do

[1] In PSS the discussion in the first line of page 308 indicates that if the lagged dependent variable y_{t-1} does not enter the sub-VAR model for $\mathbf{x}_t$, while the inclusion of lagged changes of dependent variable, Δy_t, is not ruled out, then $\mathbf{x}_t$ serves as forcing variables. Hence, the ARDL model can be estimated consistently by OLS.

not allow the possibility of cointegrating relationship between y_t and x_t. The cointegrating vector is restricted to be (1, 0) for the long-run components. The ARDL regression (1) relaxes this restriction where the cointegrating vector is not constrained. Thus, to determine the presence of a unit root in y_t, we need to test the significance of both β_1 and $\boldsymbol{\beta}_2$ in equation (1): $\Delta y_t = c_1 + c_2 t + \beta_1 y_{t-1} + \boldsymbol{\beta}_2 \mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta y_{t-i} + \sum_{j=1}^{q-1} \boldsymbol{\Phi}'_j \Delta \mathbf{x}_{t-i} + \boldsymbol{\omega}' \Delta \mathbf{x}_t + u_t$. The significance of β_1 and $\boldsymbol{\beta}_2$ tells the integration order of the y_t. There are four possible cases in this testing framework.

Case I: Nonstationary Process / No Cointegration $[\beta_1 = \boldsymbol{\beta}_2 = 0]$.

For a nonstationary process with no cointegration, the ARDL regression in (1) is reduced to $\Delta y_t = c_1 + c_2 t + \sum_{i=1}^{p-1} \phi_i \Delta y_{t-i} + \sum_{j=1}^{q-1} \boldsymbol{\Phi}'_j \Delta \mathbf{x}_{t-i} + \boldsymbol{\omega}' \Delta \mathbf{x}_t + u_t$. (5)

It conveys the nonstationarity more clearly if we rewrite the equation as

$$y_t = c_1 + c_2 t + y_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta y_{t-i} + \sum_{j=1}^{q-1} \boldsymbol{\Phi}'_j \Delta \mathbf{x}_{t-j} + \boldsymbol{\omega}' \Delta \mathbf{x}_t + u_t. \quad (6)$$

Because the coefficient on y_{t-1} is unity, y_t **is an *I*(1) process**. Equation (5) is also equivalent to Hansen's CADF model under the null of nonstationarity. Note that Case I allows y_t to have a short-run relationship with $\mathbf{x}_t$ through the terms $\Delta \mathbf{x}_{t-j}$. Of course, it can also be described as the ADF equation if all the coefficients of $\Delta \mathbf{x}_t$ are zero, i.e., $\boldsymbol{\omega}' = \boldsymbol{\Phi}'_j = \mathbf{0}$.

Case II: Stationary Process / Degenerate Lagged Independent Variables $[\beta_1 < 0, \boldsymbol{\beta}_2 = 0]$.

The ARDL regression in (1) is reduced to

$$\Delta y_t = c_1 + c_2 t + \beta_1 y_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta y_{t-i} + \sum_{j=1}^{q-1} \boldsymbol{\Phi}'_j \Delta \mathbf{x}_{t-i} + \boldsymbol{\omega}' \Delta \mathbf{x}_t + u_t, \quad (7)$$

where $\beta_1 < 0$. Rewriting (7) by adding y_{t-1} to both sides of the equation and we obtain

$$y_t = c_1 + c_2 t + \lambda y_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta y_{t-i} + \sum_{j=1}^{q-1} \boldsymbol{\Phi}'_j \Delta \mathbf{x}_{t-i} + \boldsymbol{\omega}' \Delta \mathbf{x}_t + u_t, \quad (8)$$

where $\lambda - 1$. Regression (7) is similar to the stationary CADF regression under the alternative hypothesis or a stationary ADF regression if none of the $\Delta \mathbf{x}_t$ terms are present.

Thus, in Case II, y_t **is an *I*(0) process**, and its stationarity is not affected by the presence of the stationary variables $\Delta\mathbf{x}_t$.

Case III: Second Order Integration Process / Degenerate Lagged Dependent Variable $\left[\beta_1 = 0, \boldsymbol{\beta}_2 \neq 0\right]$**.**

Under Case III, the ARDL regression (1) is:

$$\Delta y_t = c_1 + c_2 t + \boldsymbol{\beta}_2' \mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta y_{t-i} + \sum_{j=1}^{q-1} \boldsymbol{\Phi}_j' \Delta \mathbf{x}_{t-i} + \boldsymbol{\omega}' \Delta \mathbf{x}_t + u_t, \qquad (9)$$

or equivalent to

$$y_t = c_1 + c_2 t + y_{t-1} + \boldsymbol{\beta}_2' \mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta y_{t-i} + \sum_{j=1}^{q-1} \boldsymbol{\Phi}_j' \Delta \mathbf{x}_{t-i} + \boldsymbol{\omega}' \Delta \mathbf{x}_t + u_t. \qquad (10)$$

This case is special as y_t involves two sources of non-stationarity, that is from the marginal process itself (y_t) and from $\mathbf{x}_t$. Since there are two unit root properties in the equation, y_t **is an *I*(2) process**.

Case IV: Nonstationary Process / Cointegration $\left[\beta_1 < 0, \boldsymbol{\beta}_2 \neq 0\right]$**.**

This is the exact case of equation (1) with the representation of

$$\Delta y_t = c_1 + c_2 t + \beta_1 y_{t-1} + \boldsymbol{\beta}_2' \mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta y_{t-i} + \sum_{j=1}^{q-1} \boldsymbol{\Phi}_j' \Delta \mathbf{x}_{t-i} + \boldsymbol{\omega}' \Delta \mathbf{x}_t + u_t, \qquad (1)$$

where $\beta_1 < 0$ and $\beta_2 \neq 0$. Although the marginal process of y_t itself is stationary (with coefficient $\beta_1 < 0$), y_t is also a function of the $\mathbf{x}_t$. Since $\mathbf{x}_t \sim I(1)$, the nonstationarity of y_t follows from that of $\mathbf{x}_t$ and therefore y_t is an $I(1)$ process. The non-stationarity of y_t is due to its movements that follow the path of the non-stationary variables $\mathbf{x}_t$ through cointegration.

3.3 The Null Hypotheses and Limiting Distributions for the ARDL Unit Root Tests

Based on the four cases presented in section 3.2, individual tests on both β_1 and $\boldsymbol{\beta}_2$ are necessary to infer the stationarity properties of a time series using the ARDL framework. A *t*-test is proposed to test the significance of the coefficient of the lagged dependent variable β_1 and an *F*- (Wald) test for testing the joint significant of the coefficients of the *k* independent variables $\boldsymbol{\beta}_2$. We define the null and alternative hypotheses of the tests as follows.

The ARDL test equation is:

$$\Delta y_t = c_1 + c_2 t + \beta_1 y_{t-1} + \boldsymbol{\beta}_2 \mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta y_{t-i} + \sum_{j=1}^{q-1} \boldsymbol{\Phi}_j^{'} \Delta \mathbf{x}_{t-i} + \boldsymbol{\omega}^{'} \Delta \mathbf{x}_t + u_t \quad (1)$$

The hypothesis for *t*-test on β_1: $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 < 0$ (11)

The hypothesis for *F*-test on $\boldsymbol{\beta}_2$: $H_0 : \boldsymbol{\beta}_2 = \mathbf{0}'$ vs $H_1 : \boldsymbol{\beta}_2 \neq \mathbf{0}'$ (12)

These *t*- and *F*-tests are conducted separately to identify which of the four cases is the correct representation for inferring the stationarity of y_t. However, the distributions for the *t*-test and *F*-test statistics under the null are not simple and straightforward. As discussed by Banerjee et al. (1998), the null distribution of the *t*-statistic for the lagged level dependent variable within the ARDL framework, in general, is complicated and dependent on the nuisance parameter $\boldsymbol{\omega}_{xy}$, the matrix of the covariance between the errors of the equations for y_t and $\mathbf{x}_t$.[2] Unless $\mathbf{x}_t$ is strictly exogenous in model (1), the limiting distribution of the *t*-statistic under the null of $H_0 : \beta_1 = 0$ is not asymptotically similar. To overcome the lack of similarity, Banerjee et al. (1998) follow the strategy proposed by Phillips and Loretan (1991) and Saikkonen (1991) by adding the future values of $\Delta \mathbf{x}_t$ into the equation to correct for endogeneity.

Nevertheless, this distribution for the *t*-statistic under the null still imposes the untested subsidiary hypothesis of $H_0 : \boldsymbol{\beta}_2 = \mathbf{0}'$ (or equivalent to $H_0^{\pi_{yx.x}} : \boldsymbol{\pi}_{yx.x} = \mathbf{0}'$ in PSS), that is, the limiting distribution of the Banerjee test statistic is obtained under the joint hypothesis of $\beta_1 = 0$ and $\boldsymbol{\beta}_2 = \mathbf{0}'$ ($H_0^{\pi_{yx.x}} : \boldsymbol{\pi}_{yx.x} = \mathbf{0}'$ and $\boldsymbol{\pi}_{yx.x} = \mathbf{0}'$ in PSS). The limiting distribution of the *t*-statistic under $H_0 : \beta_1 = 0$ is not asymptotically similar if $\boldsymbol{\beta}_2 \neq \mathbf{0}'$.[3] The case becomes more complicated if Δy_t Granger-causes $\Delta \mathbf{x}_t$ as in equation (2) because the limiting distribution is dependent on nuisance parameters based on the ratio of short-run multipliers.[4]

Note that the PSS distribution of the *t*-statistic under the null of $H_0 : \beta_1 = 0$ follows the procedures of Banerjee et al. (1998) that impose the untested subsidiary hypothesis of $H_0 : \boldsymbol{\beta}_2 = \mathbf{0}'$, as can be seen from Theorem 3.2 in their paper. However, the restriction $\boldsymbol{\beta}_2 = \mathbf{0}'$ will not be imposed on the proposed *t*-test of $H_0 : \beta_1 = 0$ in the ARDL unit root test, as $\boldsymbol{\beta}_2$ is also under investigation in this new testing procedure. This means the distribution for the ARDL unit root *t*-test is not subject to the restriction $\boldsymbol{\beta}_2 = \mathbf{0}'$ under the null, andthe distribution under the null is uncertain and dependent on nuisance

---

[2] See Banerjee et al. (1998), Section 5.

[3] See PSS p. 295, the paragraph after equation (11) and the discussion in p. 303, the paragraph before Table CII or the last paragraph in Section 4.

[4] See PSS p. 303, the paragraph before Table CII.

parameters. The argument is similar for the distribution of the *F*-statistic for testing $\boldsymbol{\beta}_2 = \mathbf{0}'$ in equation (1). The distribution of the *F*-statistic depends on nuisance parameters as in the distribution of the *t*-statistic. The distribution of the *F*-statistic under the null is therefore uncertain as well.

To overcome the uncertainty of the limiting distributions of the *t*- and *F*-tests under the null, we use the bootstrap method. The bootstrap can determine the exact empirical distribution under the null without information about the nuisance parameters. As stated by Davidson (2007), the basic idea of bootstrap testing is that when a test statistic of interest has an unknown distribution under the null hypothesis, that distribution can be characterized by using the information in the data set that is being analyzed. Furthermore, Palm et al. (2010) and McNown et al. (2018) previously established the usefulness of the bootstrap in the ARDL framework.

4. Methodology of the Monte Carlo Experiments

4.1 Data Generating Process (DGP)

In the experiments, the DGP is based on the ARDL model for generating series with multivariate relationships. For simplicity but without loss of generality, a bivariate case of y_t and x_t is considered in the study. Experiments with bivariate setups are common, for example, Hansen (1995), Palm et al. (2010), and McNown et al. (2018).

The experiments cover a range of DGPs for y_t including models with no intercept, with intercept, and with intercept and trend. These models are generated with the properties of no cointegration between y_t and x_t, stationary process for y_t and x_t, and cointegration between y_t and x_t. However, the experiments exclude processes with quadratic trend. Without loss of generality, the DGPs for y_t are limited to ARDL(1,1) with one lagged difference of both y_t and x_t. Two sets of simulation experiments are conducted to study the performance of the ARDL unit root test. The first set is based on simulations of simple relations between y_t and x_t, while the second set of experiments generates data with greater complexity.

In the first set of experiments, y_t is generated according to (13) and (14):

$$\Delta y_t = c_1 + c_2 t + \beta_1 y_{t-1} + \beta_2 x_{t-1} + 0.5\Delta y_{t-1} + 0.5\Delta x_{t-1} + \varepsilon_{y,t}, \tag{13}$$

$$y_t = y_{t-1} + \Delta y_t. \tag{14}$$

Throughout the experiments, the short-run multipliers are set at 0.5, the intercept $c_1 = 0$ for the DGPs without intercept, and $c_1 = 0.5$ for the DGPs with intercept. The coefficient on the deterministic trend is $c_2 = 0.5$ for the DGPs with a trend. x_t, is generated as a simple random walk

$$x_t = x_{t-1} + \varepsilon_{x,t}. \tag{15}$$

The error terms $\varepsilon_{y,t}$ and $\varepsilon_{x,t}$ both are drawn from *i.i.d.* normal distributions and are uncorrelated with each other.

For the second set of experiments endogeneity is introduced. The DGP for y_t employs the following models

$$\Delta y_t = c_1 + c_2 t + \beta_1 y_{t-1} + \beta_2 x_{t-1} - 0.3\Delta y_{t-1} - 0.3\Delta x_{t-1} + \varepsilon_{y,t}, \tag{16}$$

$$y_t = y_{t-1} + \Delta y_t. \tag{17}$$

Throughout the second set of experiments, the short-run multipliers are set at -0.3, and the intercept and the coefficient on the deterministic trend are set at $c_1 = 0.5$ and $c_2 = 0.5$, respectively, when they are present. To introduce the endogeneity property in this second experiment, x_t is generated as a nonstationary process with lagged differenced terms on both y_t and x_t:

$$\Delta x_t = -0.3\Delta y_{t-1} - 0.3\Delta x_{t-1} + \varepsilon_{x,t}, \tag{18}$$

$$x_t = x_{t-1} + \Delta x_t, \tag{19}$$

with the covariance matrix for the two equation errors given by

$$\begin{pmatrix} \varepsilon_{y,t} \\ \varepsilon_{x,t} \end{pmatrix} = \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{pmatrix} \begin{pmatrix} u_t \\ v_t \end{pmatrix}, \tag{20}$$

Here u_t and v_t are the structural innovations in Δy_t and Δx_t and they are drawn from *i.i.d.* normal distributions. Through Choleski decomposition, (20) can be orthogonalized into

$$\varepsilon_{y,t} = \rho\varepsilon_{x,t} + u_t \text{ and} \tag{21}$$

$$\varepsilon_{x,t} = v_t, \tag{23}$$

where $\rho = \omega_{12}\omega_{22}^{-1}$. The coefficient ρ is the contemporaneous correlation between the equation errors for y_t and x_t. In the experiment, we set $\rho = 0.3$.

4.2 Residual Bootstrap

The bootstrap method is well-known and extensively used for hypothesis testing in econometric analysis as the bootstrap critical values are often more accurate than the asymptotic critical values (Singh, 1981; Beran, 1988). McNown et al. (2018) demonstrated that the bootstrap method is also useful to determine the empirical distribution of the test statistic given that its time series properties are unknown. Thus, bootstrap is useful in the ARDL unit root test to simulate the empirical distribution when we do not have prior knowledge about the true distribution.

To bootstrap, first, we obtain an optimal ARDL regression that best describes Δy_t in (13) or (16):

$$\Delta y_t = \hat{c}_1 + \hat{c}_2 t + \hat{\beta}_1 y_{t-1} + \hat{\beta}_2 x_{t-1} + \sum_{i=1}^{p-1} \hat{\phi}_i \Delta y_{t-i} + \sum_{j=1}^{q-1} \hat{\phi}_j \Delta x_{t-i} + u_{y,t}, \tag{24}$$

with lag length according to the DGP and estimated by OLS while excluding the β_1 (or β_2) terms to impose the restrictions of the null hypothesis for the four cases described above. The estimated restricted regression residuals are saved and resampled with replacement and recentered to produce the bootstrap residuals. Then these bootstrap residuals are added to the restricted estimated equation to produce the bootstrap observations on the dependent variable. After the bootstrap dependent variable, y_t^*, is generated, equation (1) is re-estimated using y_t^* and x_t to compute the test statistics. These resampling and re-estimation steps are repeated many times (bootstrap replications) to construct a bootstrap distribution. Bootstrap critical values are obtained from the percentiles of the bootstrap distribution.

4.3 Size and Power Analysis

To analyze the statistical size or power of each test, we calculate the rejection frequency of the test statistic based on the critical value from the corresponding bootstrap distribution described above. In analyzing the statistical sizes of the bootstrap test, the Monte Carlo simulation first generates y_t with the DGP of (13) or (16), while imposing the null hypothesis of the test under consideration. Then, regression (24) is estimated using these null restricted series to obtain the test statistic. The simulation repeats the process *N* times to obtain *N* test statistics. Next, we calculate the proportion of these *N* test statistics that reject a true null hypothesis. This proportion is the size of the test. To calculate the power of the test, similar steps are followed except the DGP of y_t is generated by imposing the alternative hypothesis instead of the null to compute the proportion of times that a false null hypothesis is rejected.

5. Simulation Results and Discussion

This section presents the results for the size and power analysis of the proposed tests. The number of simulation replications used is $N = 1000$ and bootstrap replications is $B = 1000$. Experiments with a sample size of $n = 100$ are examined. Nominal testing level is set at $\alpha = 0.05$ or 5% level. Therefore, the *t*-test critical value is obtained from the 5% percentile of bootstrap distribution and 95% percentile for the *F*-test critical value. Two sets of experiments are conducted. The first set of experiments covers cases of DGPs without contemporaneous correlation and with exogenous regressors, while the second set iinvolves cases with contemporaneous correlation and endogenous regressors. Besides the ARDL unit root test analysis, we also include the conventional univariate unit root tests such as ADF and DF-GLS for comparison. Following the general practice, a regression with intercept is used for time series without atime trend, and a regression with intercept and trend is used for processes with a time trend. This principle fits the general specification, a plausible description of the data under both the null and the alternative hypotheses (see Hamilton, 1994, p. 501). The ADF and DF-GLS test equation are set at one lagged difference, the closest specification to the simulation DGPs. The simulation experiments study cases of nonstationary process, cointegrated process, and stationary processes, but exclude the case of a degenerate lagged dependent variable, which implies that y_t is an $I(2)$ process.

5.1 Size and Power Analysis for DGPs without Contemporaneous Correlation

The rejection rate for the null hypothesis of each test is recorded and summarized in tables 1 and 2. The null hypotheses for the bootstrap *t*-test for lagged level of the dependent variable and bootstrap *F*-test for lagged levels of the independent variables are $H_0 : \beta_1 = 0$ and $H_0 : \boldsymbol{\beta}_2 = \mathbf{0}'$, respectively. The null hypotheses for the ARDL, ADF and DF-GLS tests imply nonstationary processes. The rejection frequency is recorded if the test implies a stationary process. For the ARDL test, stationarity is concluded if the *t*-test is rejected and *F*-test is not rejected, and the values reported in the "ARDL" column indicate the power of this test. Otherwise, nonstationarity is implied, and the reported values in this column show the empirical size. The bold entries indicate that the size of the test is reported for that case. Non-bold entries record the power of each test for that case.

Table 1. Size and power analysis at 5% nominal level for y_t with intercept, *T*=100, *N*=1000, *B*=1000.

DGP		Integration	β_1	β_2	ARDL	ADF	DF-GLS
1	$\beta_1 = 0, \beta_2 = 0$	$I(1)$	**0.0450**	**0.0510**	**0.0300**	**0.0440**	**0.0580**
2	$\beta_1 = -0.05, \beta_2 = 0$	$I(0)$	0.4330	**0.0550**	0.3840	0.3390	0.6170
3	$\beta_1 = -0.1, \beta_2 = 0$	$I(0)$	0.8610	**0.0490**	0.8120	0.8010	0.9030
4	$\beta_1 = -0.3, \beta_2 = 0$	$I(0)$	1.0000	**0.0430**	0.9570	1.0000	0.9980
5	$\beta_1 = -0.6, \beta_2 = 0$	$I(0)$	1.0000	**0.0430**	0.9570	1.0000	1.0000
6	$\beta_1 = -0.05, \beta_2 = 0.3$	$I(1)$	0.9110	0.9690	**0.0010**	**0.0360**	**0.0350**
7	$\beta_1 = -0.1, \beta_2 = 0.3$	$I(1)$	0.8780	0.9950	**0.0000**	**0.0800**	**0.1300**
8	$\beta_1 = -0.3, \beta_2 = 0.3$	$I(1)$	1.0000	1.0000	**0.0000**	**0.5770**	**0.6210**
9	$\beta_1 = -0.6, \beta_2 = 0.3$	$I(1)$	1.0000	1.0000	**0.0000**	**0.9210**	**0.9040**
10	$\beta_1 = -0.05, \beta_2 = 0.6$	$I(1)$	0.8520	1.0000	**0.0000**	**0.0500**	**0.0580**
11	$\beta_1 = -0.1, \beta_2 = 0.6$	$I(1)$	0.9990	1.0000	**0.0000**	**0.0750**	**0.1150**
12	$\beta_1 = -0.3, \beta_2 = 0.6$	$I(1)$	1.0000	1.0000	**0.0000**	**0.3050**	**0.3900**
13	$\beta_1 = -0.6, \beta_2 = 0.6$	$I(1)$	1.0000	1.0000	**0.0000**	**0.6410**	**0.6690**

Note: To bootstrap *t*-test, the restriction of null $\beta_1 = 0$ is imposed. For the bootstrap *F*-test, the restriction of $\beta_2 = 0$ is imposed. The entries are the rejection frequencies and entries in bold indicate the size of the test.

Table 1 summarizes the size and power analysis for the model with an intercept. The column headed as "Integration" indicates the true order of integration of y_t for each DGP. From the results, we can see both the bootstrap *t*- and *F*-tests on the individual coefficients have proper sizes with rejection rates close to the 5% nominal level when the null is true. The tests have high powers when the null is false. Let us focus on the *t*-test (column β_1). In Case 1 when the null hypothesis is true, the *t*-test has a reasonable size of 4.5% rejection rate. In Case 2 when the null hypothesis is false, the rejection rate is high with 43.3% despite the β_1 value of -0.05 is close to the null value of zero. The rejection rate increases rapidly with the absolute value of β_1, reaching 86.1% at value of -0.1 of β_1 and converging to 100% as the coefficient increases in absolute value. This convergence is extremely fast when cointegration exists, that is, from Case 6 to Case 13. Case 6 with a coefficient of -0.05 has 91.1% chance of rejecting the null, which is high despite the coefficient value being very close to the null of zero. This is explainable because the OLS estimates are super-consistent with a faster converging rate within the ARDL framework when there is cointegration. Now consider the *F*-test (column β_2). The *F*-test maintains a rejection rate around 5% for both the nonstationary process in Case 1 and the stationary DF processes in Case 2 to Case 5. For cointegration cases 6 to 13, similar to the *t*-test, the *F*-test also has a very high rejection rate when the null hypothesis is false. Based on these results, we can confirm that the bootstrap performs well for the *t*- and *F*-tests under the ARDL framework. These results on the individual coefficients of the ARDL equation provide preliminary information on the performance of the bootstrap tests, which are now combined to examine the tests for unit roots.

Next, we combine the *t*- and *F*-tests to determine the overall rejection frequencies of the ARDL test, which is compared to those of the conventional unit root tests, the ADF and DF-GLS. Focus first on the DGPs without cointegration in the first five cases. Overall, the results show that both the ADF and DF-GLS tests perform well for these cases, despite the univariate framework's exclusion of information from the lagged independent variable. In Case 1 of a nonstationary process, the ADF test has a reasonable size of 4.4% given the null of nonstationarity is true, and its power increment is rapid in Cases 2 to 5 as β_1 increases in absolute value. In these same cases, the DF-GLS test has higher power and faster rate of convergence to 100% than the ADF test.

For the ARDL test, the size of the test is underestimated with a 3% rejection rate in Case 1 of a nonstationary process. The underestimated size indicates that the test may not be exact with a lower Type I error. This is problematic because usually there is a trade-off between Type I and Type II errors. A test that has a smaller size faces a larger Type II error when the critical value is not adjusted to make the nominal size and empirical size the same. Nevertheless, this trade-off does not arise with the ARDL test, which still shows power that is comparable to the other two tests.

Next, consider the DGPs with cointegration, Cases 6 to 13. Both the ADF and DF-GLS tests suffer severe size distortions. The tests have smaller size when the β_1 value is close to zero. However, the sizes of the tests are severely biased upwards as the coefficient β_1 increases in absolute value. A similar problem exists for Cases 10 – 13 with larger values of β_2, although the size distortion is not as severe. The reason for the size distortions with these two univariate tests becomes clear from examination of the ARDL DGP, which includes the lagged level of the dependent variable. The coefficient on this term, β_1, is the focus of the ADF and DF-GLS tests. When the data on y_t is generated from the ARDL equation with a non-zero value of β_1, naturally its computed *t*-statistic is likely to be significant with an increasing frequency as β_1 becomes larger in absolute value. It is remarkable, however, that cointegration, which is a multivariate property, has such a strong adverse effect on univariate tests for a unit root.

The ARDL test does not face the same problem as the ADF and DF-GLS tests, but exhibits extremely low sizes for cointegrated process. It seems contradictory that the rejection rates for β_1 and β_2 individually are high, but the rejection rate for the null of I(1) with the ARDL test is low. Note that rejecting both null hypotheses of $\beta_1 = 0$ and $\beta_2 = 0$ implies y_t is *I*(1). Therefore, the null hypothesis of nonstationary process in ARDL test is rejected with a very small frequency when the rejection rates for $\beta_1 = 0$ and $\beta_2 = 0$ are high. Similar results are obtained for models with and without intercept and deterministic trend; therefore we do not report those results here to save space.[5]

In summary, the combined bootstrap *t*- and *F*-tests have reasonably good size and high power properties in any environment. The experiments also reveal that the misspecification of ignoring the short-run effects from the covariates has minor impact on the performance of the ADF and DF-GLS tests. However, the univariate framework faces seriously biased inferences when the nonstationary process is cointegrated and the unit root property is driven by the covariates in the long run relation. The univariate framework is unable to detect the unit root driven by the covariates and therefore misleads one to conclude the process is stationary. Finally, the experiments show that the ARDL test has low Type I and Type II errors in any environment. Although the ARDL test may not be exact, the low Type I error does not sacrifice the power of the test. It shows that the ARDL test outperforms the conventional univariate unit root tests, especially when the processes are cointegrated.

5.2 Size and Power Analysis for DGPs with Contemporaneous Correlation

The first set of experiments examined the proposed *t*- and *F*-tests, ADF, DF-GLS and ARDL tests for the DGPs without contemporaneous correlations between the equation errors. The second set of experiments with DGPs with equation errors that are contemporaneously correlated with $\rho = 0.3$ is examined in this section. Table 2 summarizes the results for the model with an intercept. Like the case without correlation,

[5] Results can be obtained by request.

the individual bootstrap *t*- and *F*-tests have proper sizes when the null is true and their powers are generally high when the null is false. The *t*-test has lower power than for the DGPs without correlation in Cases 2 and 3. Nevertheless, the power increased dramatically when the coefficient β_1 approaches -0.3. The tests are notably high in power for the cointegrated processes because OLS is super-consistent under cointegration.

Table 2. Size and power analysis at 5% nominal level for y_t with intercept, *T*=100, *N*=1000, *B*=1000, $\rho = 0.3$.

DGP		Integration	β_1	β_2	ARDL	ADF	DF-GLS
1	$\beta_1 = 0, \beta_2 = 0$	*I*(1)	**0.0500**	**0.0670**	**0.0180**	**0.0270**	**0.0270**
2	$\beta_1 = -0.05, \beta_2 = 0$	*I*(0)	0.1060	**0.0640**	0.0690	0.0700	0.1610
3	$\beta_1 = -0.1, \beta_2 = 0$	*I*(0)	0.1850	**0.0590**	0.1460	0.1360	0.3690
4	$\beta_1 = -0.3, \beta_2 = 0$	*I*(0)	0.7750	**0.0600**	0.7150	0.8450	0.8880
5	$\beta_1 = -0.6, \beta_2 = 0$	*I*(0)	0.9960	**0.0690**	0.9270	1.0000	0.9560
6	$\beta_1 = -0.05, \beta_2 = 0.3$	*I*(1)	0.6690	0.9550	**0.0130**	**0.0370**	**0.0020**
7	$\beta_1 = -0.1, \beta_2 = 0.3$	*I*(1)	0.8280	0.9460	**0.0160**	**0.0270**	**0.0080**
8	$\beta_1 = -0.3, \beta_2 = 0.3$	*I*(1)	0.9540	0.9530	**0.0250**	**0.0920**	**0.1630**
9	$\beta_1 = -0.6, \beta_2 = 0.3$	*I*(1)	0.9990	0.9780	**0.0220**	**0.4220**	**0.4780**
10	$\beta_1 = -0.05, \beta_2 = 0.6$	*I*(1)	0.8710	1.0000	**0.0000**	**0.0410**	**0.0050**
11	$\beta_1 = -0.1, \beta_2 = 0.6$	*I*(1)	0.9800	1.0000	**0.0000**	**0.0310**	**0.0090**
12	$\beta_1 = -0.3, \beta_2 = 0.6$	*I*(1)	0.9970	1.0000	**0.0000**	**0.0340**	**0.0550**
13	$\beta_1 = -0.6, \beta_2 = 0.6$	*I*(1)	1.0000	1.0000	**0.0000**	**0.1320**	**0.2000**

Note: To bootstrap *t*-test, the restriction of null $\beta_1 = 0$ is imposed. For the bootstrap *F*-test, the restriction of $\beta_2 = 0$ is imposed. The entries are the rejection frequencies and entries in bold indicate the size of the test.

The estimated size of the ADF is slightly smaller in Case 1 and its power is low when the coefficient is closer to the null (Case 2 and Case 3). Its power increases quickly when the coefficient β_1 is more negative (Case 4 and Case 5). The DF-GLS test has slightly higher power than the ADF test in Case 2 and Case 3 but generally both tests' performances are comparable. Similar to the DGPs without correlation, the ADF and DF-GLS tests perform poorly when processes are cointegrated (Case 6 to Case 13). The ADF test is slightly undersized when β_1 has smaller negative values, but its size is seriously distorted as β_1 takes on larger negative values. The DF-GLS test has high power and low size properties even in some of the cointegration cases. However, this advantage does not persist. The size of the test is seriously distorted when the negative values of β_1 become larger in magnitude (see Cases 8, 9 and 13). Therefore, although the DF-GLS test performs better than the ADF test, it is still performing poorly in some cointegration cases.

For the ARDL test, its performance is similar to the DGPs without correlation. The test has low Type I and Type II errors in any situation. The ARDL test is excellent to detect unit root in cointegrated processes, and the result is almost certainly correct with extremely low Type I error. Again, results for the models with and without intercept and deterministic trend are not reported as they have results similar to the model with intercept.[6]

To conclude the findings of the experiments, the simulations show that the bootstrap ARDL test works well in estimating the empirical distribution of the test statistics. The bootstrap *t*- and *F*-tests have proper sizes and high power properties. For the ADF test, it performs well for the stationary DF and nonstationary processes if there is no cointegration. The unspecified short-run effect from the covariates or the lagged level covariates in the univariate framework does not bring significant impact to the estimated power. However, the test performs poorly when there is a cointegrating relationship. The univariate test is unable to identify the unit root property from the covariates. This is similar for the DF-GLS test. The test works well in testing the stationary and nonstationary processes without cointegration, although it may have misspecified the short-run effect. The test also performs relatively well in some of the cointegration cases but is seriously biased when the speed of adjustment gets larger in magnitude. The test will mislead one to believe the process is stationary but in fact it is a nonstationary process where its nonstationary property comes from the covariates. The size distortions for the two univariate tests are understood from the ARDL DGP, which includes the lagged level of the dependent variable. Under cointegration the coefficient on this term, β_1, is non-zero, and the ADF and DF-GLS test will find its computed t-statistic to be significant with high frequency. This leads to incorrect rejection of the unit root hypothesis and large Type I errors with a frequency that increases as β_1 becomes larger in absolute value. It is remarkable that cointegration, which is a multivariate property, has strong adverse effects on univariate tests for a unit root.

The ARDL test, however, is not vulnerable to this problem and performs well in all environments. The test has lower Type I and Type II errors than the ADF and DF-GLS tests in testing the stationary and nonstationary processes in most cases and extremely low Type I error when cointegration exists. The test is reliable in detecting the unit root property that is driven by the long run relation with covariates, whereas tests in univariate framework, such as ADF and DF-GLS tests, fail to do so.

6. An Empirical Application

The inflation rate is a key variable in many economic models (Basher and Westerlund, 2008). However, the degree of persistence and order of integration of inflation rate is widely debated. Unit root test results vary across countries and time periods and depend on the particular statistical methods adopted. The integration order of inflation has implications for Fisher's real interest parity hypothesis, the expectations augmented Phillips curve model, the accelerationist hypothesis, the traditional capital asset pricing model, monetary policies for controlling money supply growth, among

[6] Results can be obtained by request.

others. Knowledge on inflation rate stationarity is essential in the implementation of monetary policy as well. As discussed by Cecchetti and Debelle (2006), a nonstationary inflation rate implies that shocks to inflation have a permanent effect, resulting in higher costs for central banks to control inflation through monetary policy. Moreover, the selection of the correct econometric method depends on its stationarity as well. If the inflation rate is stationary, standard regression methods can be used. Otherwise, the cointegration framework may be necessary.

Many methods are used to investigate the stationarity of the inflation rate, including the conventional univariate unit root tests (e.g. Rose, 1988; Lai, 1997; Koustas and Lamarche, 2009; Rapach and Weber, 2004) and sophisticated approaches (e.g. Baillie, 1989; Ozcan, 2013; Culver and Papell, 1997; Gregoriou and Kontonikas, 2006; Cook 2009). The conclusions on the inflation rate's order of integration are mixed. In this section the multivariate ARDL unit root test is applied to examine the stationarity of Malaysia's inflation rate.

To apply the ARDL unit root test, it is necessary to identify a potential covariate that can help to explain the dynamics of inflation within the ARDL framework. As discussed by Lee and Tsong (2011), international financial markets have become increasingly integrated, through globalization and the liberalization of international trade and investment flows. These changes improve the degree of integration of markets across countries, causing real interest rates, and by extension nominal interest rates and inflation rates, to be related across countries. Lee and Tsong (2011) investigate the stationarity of the nominal interest rates and inflation rate using the bootstrap CADF test, using data from other countries as covariates.

In this empirical application, the Malaysian inflation rate, measured by the GDP deflator (MAS-I) is the subject of the study, and the covariate is the US inflation rate. The sample period is from 1961 to 2019. The data are from the World Bank database. Figure 1 depicts these two time series.

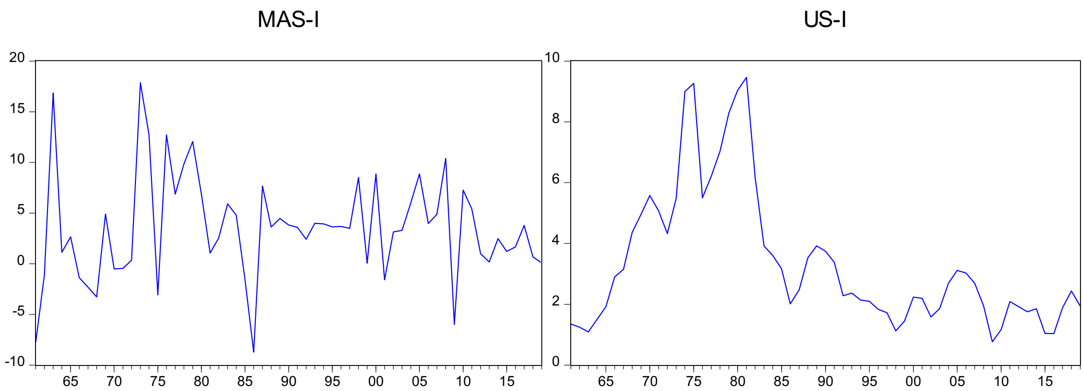

Figure 1. Time series plots for Malaysia and the US inflation rates.

A visual inspection of Figure 1 shows that the Malaysia inflation rate (MAS-I) may be stationary with huge volatility whereas the US inflation rate (US-I) is strongly persistent throughout the period. Neither series shows an apparent trend. Hence, a unit root test regression is implemented with only an intercept as deterministic component. Five conventional univariate unit root tests were applied to MAS-I with results summarized in Table 3. The unit root tests suggest different results for MAS-I. The ADF test indicates it is *I*(1) process, while PP (Phillips-Perron) and KPSS tests suggest it is *I*(0). Even more problematic is that the DF-GLS and NP (Ng-Perron) tests suggest that MAS-I has a higher integration order, which does not seem reasonable. For US-I, the result is consistent across all univariate tests, concluding that it is *I*(1). The contradictory results on MAS-I could be due to its complex interaction with global inflation, so that a univariate framework is inadequate to represent the data, while the multivariate ARDL unit root test may be more appropriate for testing this series.

Table 3. Univariate unit root tests.

Unit Root Tests	MAS-I		US-I	
	Level	1st Difference	Level	1st Difference
ADF	-2.5064	-12.0795***	-1.6249	-5.8058***
DF-GLS	-0.3703	-0.0411	-1.3927	-5.8284***
PP	-7.5917***	-	-1.8726	-5.7435***
KPSS	0.0982	-	0.4231*	0.1641
NP	0.3496	0.0693	0.3952	26.8098***

Note: The lag length selection for ADF, DF-GLS, and NP tests are according to modified AIC. For the PP and KPSS tests, the spectral estimation method is based on Bartlett kernel and bandwidth with Newey-West selection. For the NP test, the spectral estimation method is based on AR GLS-detrended method. The null hypotheses for ADF, DF-GLS, PP, and NP tests are H_0: series contains a unit root, while the null hypothesis for KPSS is H_0: series is stationary. *, **, *** indicate significance at 10%, 5% and 1% levels, respectively.

A bootstrap program is developed to run the ARDL unit root test with MAS-I as the dependent variable and US-I as its covariate. The maximum lag length is set at 8, and the optimal lag length is selected according to modified AIC. The number of bootstrap replications is set at B = 5000 to generate the bootstrap critical values. The ARDL equation includes a constant but no trend term. Table 4 summarizes the test results. According to the modified AIC, the optimal ARDL model is ARDL(0,2). The estimated *t*-statistic for the lagged level dependent variable and *F*-statistic for the lagged level independent variable are -7.6179 and 3.5897, respectively. Compared to the bootstrap critical values, both the *t*- and *F*-statistics are significant at 1% level and 10% level, respectively. It indicates that the MAS-I is cointegrated with US-I, and the finding of a cointegration relationship shows that the movement of MAS-I responds to the movement of US-I. Since the US-I is *I*(1), therefore MAS-I is *I*(1) also. The ARDL result is in line with the ADF test.

Table 4. Multivariate ARDL unit root test for RINT.

ARDL unit root test in level			
Testing variable	MAS-I	Optimal model	ARDL(0, 2)
Covariate	US-I	Modified AIC	6.0279
Regression type	Intercept	Bootstrap Rep.	5000

t-statistic	-7.6179
F-statistic	3.5897

Significance levels	0.100	0.050	0.025	0.010
t-critical values	-2.8721	-3.2202	-3.5369	-3.9095
F-critical values	2.8703	4.2725	5.4865	7.2723

In sum, the conventional unit root tests give mixed results on the order of integration for the Malaysia inflation rate. The DF-GLS and NP tests, which are generally believed to have higher power properties, suggest that the Malaysia inflation rate has an order of integration higher than one. This is unacceptable as the series obviously does not display any visual evidence of second order integration. The t- and F-unit root tests in the ARDL framework suggest that the Malaysia inflation rate is $I(1)$ and it is cointegrated with the US inflation rate. The ARDL framework discovers the cointegration relationship between the Malaysia and the US inflation, and this complex dynamic relationship could not be detected using the univariate framework. This explains why the conventional unit root tests lead to conflicting outcomes. The unit root in the Malaysian inflation rate is related to the integrated property of the US inflation rate though cointegration, which undermines the validity of the univariate tests.

7. Conclusion

The stationarity of a time series is important in applied economic research because it informs the methodology that is appropriate for subsequent analysis. Standard regression methods are used if the underlying variables are stationary, and a cointegration method may be more appropriate with nonstationary variables. Standard unit root tests can often lead to contradictory results, and they can be either size distorted or low in power. The size and power problems could be due to the univariate framework of the tests as demonstrated in this paper. Hansen (1995) is the first to propose a multivariate framework unit root test, the Covariate Augmented Dickey-Fuller (CADF) test, which includes relevant variables to capture model dynamics. However, the CADF unit root test still faces model misspecification problem as it excludes the possibility of cointegration with the covariates. If there is a cointegrating relationship, the CADF framework imposes an invalid common factor, thus affecting its power. Therefore, to avoid imposing an invalid common factor, a multivariate unit root test within autoregressive distributed lag (ARDL) framework is proposed and analyzed in this study.

The multivariate ARDL unit root test uses two tests to infer the order of integration of a time series, namely, the t-test on the lagged level of the dependent variable and the F-test on the lagged levels of the independent variables. The stationarity of a process can be determined according to the significance of the proposed tests. However, the null distributions for the two tests are complicated and nuisance parameter dependent. To overcome the uncertainty of the test distributions, the bootstrap method is used. The ARDL unit root test using the bootstrap has correct size and excellent power in a variety of environments, regardless of whether the time series of interest is a nonstationary,

stationary, or cointegrated process. The ARDL unit root test outperforms the conventional unit root tests, especially when cointegration exists. The experiments show that when there is a cointegrating relationship, the ARDL test has a lower size property which indicates a lower Type I error. Usually, this is an issue for a test because with any tests, there is a trade-off between Type I and Type II errors. However, the undersized ARDL test still shows higher powers compared to the ADF or DF-GLS tests. The experiments also show the sizes of the ADF and DF-GLS tests are seriously distorted in testing a cointegrated process because they fail to capture the unit root that arises from the nonstationary covariates. The existence of cointegration is manifested in a significant coefficient on the lagged level of the dependent variable, which leads the standard univariate unit root tests to reject the true null hypothesis of a unit root.

In the re-examination of the inflation rate stationarity, we show the weaknesses of the univariate unit root tests with the evidence of data from Malaysia. Five conventional unit root tests yield mixed results for the Malaysian inflation rate. Some of the tests even suggest that the series has an integration order higher than $I(1)$, which is not plausible based on the time series plot. Through the ARDL test, the evidence suggests the Malaysian inflation rate is $I(1)$, and its nonstationary property comes from its cointegration with the US inflation rate.

References

Baillie, R.T. (1989). Commodity Prices and Aggregate Inflation: Would Commodity Price Rule be Worthwhile? *Carnegie Rochester Conference Series on Public Policy*, **31**, 185-240.

Banerjee, A., Dolado, J., and Mestre, R. (1998). Error-Correction Mechanism Tests for Cointegration in Single-Equation Framework. *Journal of Time Series Analysis*, **19**, 267-283.

Basher, S.A. and Westerlund, J. (2008). Is There Really a Unit Root in the Inflation Rate? More Evidence from Panel Data Models. *Applied Economics Letters*, **15**(3), 161-164.

Beran, R. (1988). Prepivoting Test Statistics: A Bootstrap View of Asymptotic Refinements. *Journal of the American Statistical Association*, **48**, 253-277.

Cavaliere, G. and Taylor, A.M.R. (2008). Bootstrap Unit Root Tests for Time Series with Nonstationary Volatility. *Econometric Theory*, **24**(1), 43-71.

Cecchetti, S.G. and Debelle, G. (2006). Has the Inflation Process Changed? *Economic Policy*, **21**(46), 311-352.

Chang, Y. and Park, J.Y. (2003). A Sieve Bootstrap for the Test of a Unit Root. *Journal of Time Series Analysis*, **24**(4), 379-400.

Chang, Y., Sickles, R.C. and Song, W. (2017). Bootstrapping Unit Root Tests with Covariates. *Econometric Reviews*, **36**, 136-155.

Cook, S. (2009). A Re-Examination of the Stationarity of Inflation. *Journal of Applied Econometrics*, **24**, 1047-1053.

Culver, S.E. and Papell, D.H. (1997). Is There a Unit Root in the Inflation Rate? Evidence from Sequential Break and Panel Data Models. *Journal of Applied Econometrics*, **12**, 435-444.

Davidson, R. (2007). Bootstrapping Econometric Models. Departmental Working Papers 2007-13, McGill University, Department of Economics.

Elliott, G. and Jansson, M. (2003). Testing for Unit Roots with Stationary Covariates. *Journal of Econometrics*, **115**, 75-89.

Flores, R.G., Jorion, P., Preumont, P.Y., and Szafarz, A. (1996). Multivariate Unit Root Tests. CEME-ULB, Discussion Paper 9601.

Gregoriou, A. and Kontonikas, A. (2006). Inflation Targeting and the Stationarity of Inflation: New Results from an ESTAR Unit Root Test. *Bulletin of Economic Research*, **58**(4), 3307-3378.

Hamilton, J.D. (1994). Time Series Analysis. Princeton University Press, Princeton.

Hansen, B.E. (1995). Rethinking the Univariate Approach to Unit Root Testing: Using Covariates to Increase Power. *Econometric Theory*, **11**, 1148-1171.

Hansen, B.E., and Racine, J.S. (2018). Bootstrap Model Averaging Unit Root Inference. Department of Economics Working Paper No. 2018-09.

Juhl, T. and Xiao, Z. (2003). Power Functions and Envelopes for Unit Root Tests'. *Econometric Theory*, **19**, 240-253.

Kokoszka, P. and Gabrield, Y. (2016). KPSS Test for Functional Time Series. *A Journal of Theoretical and Applied Statistics*, **50**, 957-973.

Koustas, Z. and Lamarche, J. (2009). 'Evidence of Nonlinear Mean Reversion in the Real Interest Rate'. *Applied Economics*, **42**, 237-248.

Kremers, J.J.M., Ericsson, N.R. and Dolado, J.J. (1992). The Power of Cointegration Test. *Oxford Bulletin of Economics and Statistics*. **54**, 325-348.

Kurozumi, E., Yamazaki, D. and Hadri, K. (2012). Covariate Unit Root Test for Cross-Sectionally Dependent Panel Data. Global COE Hi-Stat Discussion Paper Series gd12-256, Institute of Economic Research, Hitotsubashi University.

Lai, K.S. (1997). 'Long-Term Persistence in the Real Interest Rate: Evidence of a Fractional Unit Root'. *International Journal of Finance and Economics*, **2**, 225-235.

Lee, C.F. and Tsong, C.C. (2011). 'Do Real Interest Rates Really Contain a Unit Root? More Evidence from a Bootstrap Covariate Unit Root Test'. *Pacific Economic Review*, **16**(5), 616-637.

Lupi, C. (2010). Unit Root CADF Testing with R. Journal of Statistical Software, 32(2), 1-19.

McNown, R., Sam, C.Y., Goh, S.K. (2018). Bootstrapping the Autoregressive Distributed Lag Test for Cointegration. *Applied Economics*, **50**, 1509-1521.

Ozcan, B. (2013). Are Inflation Rates Stationary in 11 Mediterranean Countries? Evidence from Univariate and Panel Unit Root Tests. *Eurasian Journal of Business and Economics*, **6**(12), 79-96.

Palm, Franz C., Smeekes, S. and Urbain, J-P. (2010). A Sieve Bootstrap Test for Cointegration in a Conditional Error Correction Model. *Econometric Theory*. **26**, 647-681.

Park, J.Y. (2003). Bootstrap Unit Root Tests. *Econometrica*, **71**(6), 1845-1895.

Pesaran, M.H., and Shin, Y. (1999). An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis. Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium, (Ed.) S. Storm, Cambridge University Press, Cambridge, 1-33.

Pesaran, M.H., Shin Y. and Smith, R.J. (2001). Bounds Testing Approaches to the Analysis of Level Relationships. *Journal of Applied Econometrics*, **16**, 289-326.

Phillips, P.C.B. and Durlauf, S.N. (1986). Multiple Time Series Regression with Integrated Process. Review of Economic Studies, 53, 473-495.

Phillips, P.C.B., and Loretan, M. (1991). Estimating Long-Run Economic Equilibria. *Review of Economic Studies*, **58**, 407-436.

Rapach, D.E. and Weber, C.E. (2004). Are Real Interest Rates Really Nonstationary? New Evidence from Tests with Good Size and Power. *Journal of Macroeconomics*, **26**, 409-430.

Rose, A.K. (1988). 'Is the Real Interest Rate Stable?'. *Journal of Finance*, **43**, 1095-1112.

Saikkonen, P. (1991). Asymptotically Efficient Estimation of Cointegration Regressions. *Econometric Theory*, **7**, 1-21.

Schwert, G.W. (1989). Testing for Unit Roots: A Monte Carlo Investigation. *Journal of Business and Economic Statistics*, **7**, 147-159.

Singh, K. (1981). On the Asymptotic Accuracy of Efron's Bootstrap. *Annals of Statistics*. **9**, 1187-1195.

Zou, N. and Politis, D.N. (2019). Linear Process Bootstrap Unit Root Test. *Statistics and Probability Letters*, **145**, 74-80.

FUNDING :

This work was supported by the Ministry of Higher Education Malaysia, FRGS/1/2020/SSO/USM/02/25.