COMMUNICATIONS IN STATISTICS - SIMULATION AND COMPUTATION®
https://doi.org/10.1080/03610918.2019.1691229

Taylor & Francis
Taylor & Francis Group

Most stringent test of null of cointegration: a Monte Carlo comparison

Asad ul Islam Khan[a], Waqar Muhammad Khan[b], and Mehmood Hussan[c]

[a]Department of Economics, Ibn Haldun University, Istanbul, Turkey; [b]Department of Economics, Women University of Azad Jammu and Kashmir, Bagh, Pakistan; [c]School of Economics, International Institute of Islamic Economics, International Islamic University Islamabad, Pakistan

ABSTRACT

To test for the existence of long run relationship, a variety of null of cointegration tests have been developed in literature. This study is aimed at comparing these tests on basis of size and power using stringency criterion: a robust technique for comparison of tests as it provides with a single number representing the maximum difference between a test's power and maximum possible power in the entire parameter space. It is found that in general, asymptotic critical values tends to produce size distortion and size of test is controlled when simulated critical values are used. The simple LM test based on KPSS statistic is the most stringent test at all sample sizes for all three specifications of deterministic component, as it has the maximum difference approaching to zero and lesser than 20% for the entire parameter space.

ARTICLE HISTORY
Received 18 June 2019
Accepted 5 November 2019

KEYWORDS
Comparison; Cointegration tests; Power; Size; Stringency

JEL CLASSIFICATION
C01; C15

1. Introduction

Engle and Granger (1987) were first to introduce the concept of cointegration. Cointegration means the existence of long run relationship between two or more than two integrated of order one variables. For two variables say *X* and *Y* having order of integration as one, if their linear combination has order of integration as zero, then these two variables are said to be cointegrated i.e., they have a long run relationship. After, its development numerous tests of cointegration have been developed for the last three decades such as tests developed by Phillips and Ouliaris (1990), tests developed by Johansen and Juselius (2009), test developed by Pesaran et al. (2001) and many more. At first these cointegration tests were developed for the null of no cointegration. However, later on Leybourne and McCabe (2009) were the pioneers of developing cointegration tests for null of cointegration. Following them, null of cointegration tests were developed in great diversity such as test developed by Shin (1994), tests developed by Fernández-Macho and Mariel (1994) and many others. All these tests were developed on basis of different characteristics of population assuming different data generation of a cointegrated system.

CONTACT Asad ul Islam Khan asadkhan1983@gmail.com, asad.khan@ihu.edu.tr Department of Economics, Ibn Haldun University, Istanbul, Turkey.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/lssp.

These developed cointegration tests produce contradicting results to each other for a same empirical problem, due to different underlying assumptions of each test. Hence, there was a need to evaluate the performance of these developed tests. Therefore, to fill this gap and to evaluate the performance of cointegration tests, comparisons have been carried out in literature in great diversity. Most of these comparisons were based on Monte Carlo simulations (MCS) such as Banerjee et al. (2009), Kremers et al. (1992) Boswijk and Franses (1992), Haug (1996), Mariel (1996), Pekmezci and Dilek (2016) and many more. However, few comparison were based on real data such as Pesavento (2004) and Gonzalo and Lee (1998). All these MCS-based comparisons assessed the performance of tests using two criteria: one is the size of the test and the other is power of the test. The size of test is defined as "The probability of rejecting a null hypothesis when actually it is true" and the power of a test is defined as "Probability of rejecting a null hypothesis when actually it is false." A test is said to be better than other if it has a controlled size (i.e., empirical size must be around nominal size, when test is conducted) and better power than rest. For a detailed survey of these comparisons, Khan and Zaman (2017) may be consulted. Majority of these MCS-based comparison considered one or two alternative hypotheses even though the alternative space has infinite values. However, among all these comparisons, the studies of Haug (1996) and Mariel (1996) were the most comprehensive and broad ones, comparing nine different tests over 10 different alternative hypotheses, covering almost all of the alternative space. These two studies assumed different data generating processes (DGPs). However, according to Haug (1996) and Mariel (1996) there is no conclusive answer that which test or tests perform better than rest. In a recent study by Pekmezci and Dilek (2016), only two cointegration tests were compared using only three points in alternative space.

The performance of six tests having null of cointegration was evaluated by Gabriel (2003) on basis of size and power using MCS. Gabriel (2003) assumed a variety of DGPs. The study of Gabriel (2003) is the only one, carrying comparison of only null of cointegration tests, to the best of our knowledge. Rest of all studies in literature are either comparing only the tests having null hypothesis of no cointegration or comparing both types of tests, i.e., tests with null of no cointegration and null of cointegration.

All these MCS-based comparative studies used the asymptotic critical values, producing size distortions. However, even though the size of tests was not controlled, yet these tests were compared on basis of their powers. It is the basic principle in statistics and econometrics that two or more tests cannot be compared on basis of power if their sizes are not controlled. Because, if a test has empirical size way larger than the nominal size then it is more than likely to have higher power. Moreover, there is a need of framework through which tests are compared for the whole alternative space. Therefore, this article is aimed to fill two gapes in literature: first the size of tests is controlled around a fixed nominal size. Secondly, the whole alternative space is considered for power comparison. For this purpose, the stringency criterion: a robust technique for comparison of tests is used. The detailed discussion on stringency criterion can be found in Zaman et al. (2017). The stringency criterion has been used for comparison of tests in a limited number of studies including Khan and Khan (2018) and Islam (2017).

Eight tests with null of cointegration are compared based on MCS in this study. These tests are compared on basis of two criteria, i.e., size of test and stringency. This study will enable practitioners, applied researchers and statisticians to select an appropriate better performing test on basis of size and power. This study will also shed light on worse performing tests, so those may be avoided. It will also enable the practitioners to decide, whether to use asymptotic critical values or simulated ones.

The structure of the rest of study is as follows. Section 2 discusses the whole methodological framework adapted to carry out the study. Section 3 have the results explained. Section 4 draws conclusions and provide recommendations to the readers. At the end "References" are listed.

2. Methodology

This section gives details of eight tests to be compared in "tests to be compared," the system of equations used to generate the artificial data in "artificial data generation," the steps to be followed to get α the size of test, using asymptotic critical values as well as using simulated critical values in "MCS design for empirical size of test α," the steps to be followed for obtaining simulated critical values in "MCS design for simulated critical values" by fixing α, the steps to be followed to estimate power of test in "MCS design for power," the detail of point optimal test in "point optimal test" and the details of stringency in "stringency criterion." Moreover, the details of an empirical example using the quarterly data of GDP and household consumption expenditure for USA from 1957-Q1 to 2013-Q3 are also laid out. All the codes for carrying out this study are compiled and run in MATLAB.

2.1. Tests to be compared

Eight tests with null hypothesis of cointegration are compared in this article whose details are:

2.1.1. LM Test of cointegration based on KPSS statistic (LMKPSS)

It is supposed that there are two classes of variables, say y_t and x_{it}, $i = 1, 2, \ldots\ldots\ldots\ldots, m$ which are separately $I(1)$, i.e., Integrated of order one. To carry out the test, first an ordinary least square (OLS) regression is fitted

$$y_t = \delta\psi_t + \sum_{i=1}^{m} \beta_i x_{it} + \mu_t \quad t = 1, 2, \ldots, T \tag{1}$$

Where, y_t is the dependent and $x_{it}, i = 1, 2, \ldots\ldots\ldots\ldots, m$ are m independent variables. The term ψ_t denotes the deterministic part containing intercept and linear time trend. Leybourne and McCabe (2009) suggested that following Kwiatkowski et al. (1992), LM type test statistic can be used for cointegration testing under null hypothesis of cointegration. This LM type test statistic is

$$\mathrm{LM} = T^{-2} \frac{\sum_{t=1}^{T} S_t^2}{s^2(l)}$$

where $S_t = \sum_{t=1}^{i} \hat{\mu}_t$ and

$$s^2(l) = \frac{\hat{\mu}_t{}'\hat{\mu}_t}{T}$$

And $\hat{\mu}_t$ are OLS residuals estimated from Eq. (1). In this article, critical values of this test have been taken from McCabe et al. (1997).

2.1.2. Leybourne and McCabe's LBI test of cointegration (LMLBI)

This test was also developed by Leybourne and McCabe (2009). According to them the same LM type test statistic can be used as a test of cointegration with null of cointegration with a slight nonparametric modification. The test statistic is

$$\mathrm{LBI} = T^{-2}\frac{\sum_{t=1}^{T} S_t^2}{s^2(l)}$$

where $S_t = \sum_{t=1}^{i} \hat{\mu}_t$ and

$$s^2(l) = T^{-1}\sum_{t=1}^{T}\hat{\mu}_t^2 + 2T^{-1}\sum_{s=1}^{l}\sum_{t=s+1}^{T}\hat{\mu}_t\hat{\mu}_{t-s}$$

And $\hat{\mu}_t$ are OLS residuals estimated from Eq. (1). The selection of lag truncation parameter l is a crucial issue in real world applications. Different values of l lead to different results. The power and size of test depends on value of l. In this $l = 4$ has been taken as proposed by Mariel (1996). According to him at this value, size of test is controlled and gives reasonable powers. In this the critical values of this test have been used from McCabe et al. (1997).

2.1.3. Shin's C test of cointegration (SC)

Shin (1994) proposed to use OLS residuals $\hat{\mu}_t$ from regression Eq. (2) instead of Eq. (1)

$$y_t = \delta\psi_t + \gamma x_t + \sum_{i=-k}^{k}\pi_i\Delta x_{t-i} + \mu_t \qquad t = 1, 2, \ldots, T \tag{2}$$

Where, y_t is the dependent, x_t is vector of independent variables and k is the maximum number of lags or leads. The term ψ_t denotes the deterministic part containing intercept and linear time trend. The test statistic is

$$C = T^{-2}\frac{\sum_{t=1}^{T} S_t^2}{s^2(l)}$$

Where $S_t = \sum_{t=1}^{i} \hat{\mu}_t$ and

$$s^2(l) = T^{-1}\sum_{t=1}^{T}\hat{\mu}_t^2 + 2T^{-1}\sum_{s=1}^{l}\left(1 - s(l+1)^{-1}\right)\sum_{t=s+1}^{T}\hat{\mu}_t\hat{\mu}_{t-s}$$

And $\hat{\mu}_t$ are OLS residuals estimated from Eq. (2). The selection of lag truncation parameter l is again a crucial issue as stated in Sec. 2.1.3. In this study, $l = 4$ has been used as proposed by Mariel (1996) and $k = 5$ as proposed by Shin (1994). Critical values are given in Shin (1994).

2.1.4. McCabe–Leybourne–Shin test of cointegration (MLS)

McCabe et al. (1997) proposed a different estimation methodology using maximum likelihood estimation instead of OLS. They suggested that first OLS residuals $\hat{\mu}_t$ may be estimated from Eq. (2) and then these residuals $\hat{\mu}_t$ may be used to estimate residuals $\hat{\eta}_t$ by applying maximum likelihood estimation to equation

$$\hat{\mu}_t = \sum_{i=1}^{p} \lambda_i \Delta \hat{\mu}_{t-i} + \eta_t$$

Where p is selected with minimum Akaike information criterion (AIC). The test statistic is

$$Ls = \frac{\hat{\eta}_t' \Omega \hat{\eta}_t}{T^2 S^2(l)}$$

Where $\Omega = \Lambda\Lambda'$ with Λ is the lower triangular matrix of ones and

$$S^2(l) = \frac{\hat{\eta}_t' \hat{\eta}_t}{T}$$

In this study, the maximum value of p is taken as 4 and then AIC is used to choose appropriate value of p. The critical values of MLs are given in McCabe et al. (1997).

2.1.5. Hausman H_1 test of cointegration (HH1)

Fernández-Macho and Mariel (1994) proposed a test statistic that compares two estimators. These two estimators are consistent under the null hypothesis of cointegration; however, one is inconsistent under the alternative hypothesis of no cointegration. The OLS estimation of Eq. (2) will give us an estimate of γ say $\hat{\gamma}_L$. Define

$$\upsilon_t = y_t - \sum_{j=-k}^{k} \hat{\pi}_j \Delta x_{t-j} \tag{3}$$

Where $\hat{\pi}_j$ are OLS estimates of Eq. (2). Then estimate following equation using OLS to obtain estimator $\hat{\gamma}_D$

$$\Delta \upsilon_t = \gamma_D \Delta x_t + \varepsilon_t \tag{4}$$

These two estimators are used to define Hausman like tests statistic under null hypothesis of cointegration

$$H_1 = (\hat{\gamma}_L - \hat{\gamma}_D)'(\hat{V}_D + \hat{V}_L)^{-1}(\hat{\gamma}_L - \hat{\gamma}_D)$$

where $\hat{V}_D$ and $\hat{V}_L$ are estimates of the covariance matrices of $\hat{\gamma}_D$ and $\hat{\gamma}_L$, respectively. Fernández-Macho and Mariel (1994) also gave critical values for the sample sizes of $T = 10, 20, \ldots, 500$ and for 1–4 regressors. In our study we have settled for $k = 5$.

2.1.6. Hausman H_2 test of cointegration (HH2)

Fernández-Macho and Mariel (1994) proposed a second test statistic H_2 based on same estimators $\hat{\gamma}_L$ and $\hat{\gamma}_D$ under null hypothesis of cointegration. These two estimators are estimated using Eqs. (2)–(4). The test statistic is given as

$$H_2 = (\hat{\gamma}_L - \hat{\gamma}_D)' \ \hat{V}_D{}^{-1}(\hat{\gamma}_L - \hat{\gamma}_D)$$

where $\hat{V}_D$ is estimate of the covariance matrix of $\hat{\gamma}_D$. Again, Fernández-Macho and Mariel (1994) also gave critical values for the same sample sizes of $T = 10, 20, \ldots, 500$ and for 1–4 regressors. In our study, again we settled for $k = 5$.

2.1.7. Hansen's L_c test of cointegration (HLC)

Hansen (1992) developed the L_c test which is based on the fully modified estimation method of Phillips and Hansen (1990). This procedure follows with estimation of Eq. (1) by OLS and finding out of estimated residuals $\hat{\mu}_t$. Then by taking first difference

$$\Delta x_{it} = \upsilon_{it} \text{ for } \quad i = 1, 2, \ldots, m,$$

with the υ_{it} representing mean-zero random errors. Then ξ_t is defined by

$$\xi_t = (\mu_t \ \ \upsilon'_t)$$

VAR (1) model is estimated using equation

$$\xi_t = \Theta \xi_{t-1} + \Upsilon_t$$

The estimated residuals matrix $\hat{\Upsilon}_t$ is used to calculate

$$\hat{\Lambda}_\Upsilon = \sum_{s=0}^{T} w(s/\hat{M}) \frac{1}{T} \sum_{t=s+1}^{T} \hat{\Upsilon}_{t-s} \hat{\Upsilon}'_t$$

$$\hat{\Omega}_\Upsilon = \sum_{s=-T}^{T} w(s/\hat{M}) \frac{1}{T} \sum_{t=s+1}^{T} \hat{\Upsilon}_{t-s} \hat{\Upsilon}'_t$$

Where $w(s/\hat{M})$ denotes an appropriate weighting Kernel. In our study, we used quadratic spectral kernel, i.e.,

$$w(s/\widehat{M}) = \frac{25}{12\pi^2 (s/\widehat{M})^2} \left\{ \frac{\sin\left(6\pi(s/\widehat{M})\right)/5}{\left(6\pi(s/\widehat{M})\right)/5} - \cos\left(\left(6\pi(s/\widehat{M})\right)/5\right) \right\}$$

and automatic bandwidth estimator M as

$$\hat{M} = 1.3221 \left\{ \left(\hat{\alpha}(2)\right) T \right\}^{1/5}$$

where

$$\hat{\alpha}(2) = \sum_{a=1}^{p} \frac{4\hat{\rho}_a^2 \hat{\sigma}_a^4}{(1-\hat{\rho}_a)^8} \Big/ \sum_{a=1}^{p} \frac{\hat{\sigma}_a^4}{(1-\hat{\rho}_a)^4} \tag{5}$$

In Eq. (5), $\hat{\rho}_a$ and $\hat{\sigma}_a^2$ are the estimated AR coefficient and estimated variance of residuals of endogenous variable "*a*." The above estimates are recolored to get the required covariance estimates

$$\hat{\Omega} = (I - \hat{\Theta})^{-1} \hat{\Omega}_\Upsilon (I - \hat{\Theta}')^{-1}$$

$$\hat{\Lambda} = (I - \hat{\Theta})^{-1}\hat{\Lambda}_{\Upsilon}(I - \hat{\Theta}')^{-1}(I - \hat{\Theta})^{-1}\hat{\Theta}\hat{\Sigma}$$

Where $\hat{\Sigma} = \frac{1}{T}\xi_t\xi'_t$
Then matrices $\hat{\Omega}$ and $\hat{\Lambda}$ are partitioned in conformity with ξ_t :

$$\Omega = \begin{bmatrix} \Omega_{\mu\mu} & \Omega_{\mu\upsilon} \\ \Omega_{\upsilon\mu} & \Omega_{\upsilon\upsilon} \end{bmatrix} \quad \text{and } \Lambda = \begin{bmatrix} \Lambda_{\mu\mu} & \Lambda_{\mu\upsilon} \\ \Lambda_{\upsilon\mu} & \Lambda_{\upsilon\upsilon} \end{bmatrix}$$

Further define

$$\Omega_{\mu.\upsilon} = \Omega_{\mu\mu} - \Omega_{\mu\upsilon}\Omega_{\upsilon\upsilon}^{-1}\Omega_{\upsilon\mu} \tag{6}$$

and $\Lambda_{\upsilon\mu}^{+} = \Lambda_{\upsilon\mu} - \Lambda_{\upsilon\upsilon}\Omega_{\upsilon\upsilon}^{-1}\Omega_{\upsilon\upsilon}$
The fully modified OLS (FMOLS) estimator

$$\hat{\beta}^{+} = \left[\sum_{t=1}^{T}\left(y_t^{+}x'_t - (0 \quad \hat{\Lambda}_{\upsilon\mu}^{+\prime})\right)\right]\left[\sum_{t=1}^{T}x_t x'_t\right]^{-1}$$

Where $y_t^{+} = y_t - \hat{\Omega}_{\mu\upsilon}\hat{\Omega}_{\upsilon\upsilon}^{-1}\Delta x_t$
and the FMOLS residuals are

$$\hat{\mu}_t^{+} = y_t^{+} - \hat{\beta}^{+}x_t \tag{7}$$

The L_c test statistic is a Lagrange multiplier (LM) test:

$$L_c = trace\left[\left(\sum_{t=1}^{T}x_t x'_t\right)^{-1}\sum_{t=1}^{T}\hat{S}_t\hat{\Omega}_{\mu\bullet\upsilon}^{-1}\hat{S}'_t\right]$$

where $\hat{S}_t = \sum_{i=1}^{t}\left(x_i\hat{\mu}_i^{+} - \begin{bmatrix} 0 \\ \hat{\Lambda}_{\upsilon\mu}^{+} \end{bmatrix}\right)$

Critical values are given in Boswijk and Franses (1992).

2.1.8. Xiao fluctuation test of cointegration (XFT)

Xiao (1999) derived a residual based test for the null hypothesis of cointegration based on the fluctuation of the residuals $\hat{\mu}_t$ from the cointegrating regression Eq. (1). Xiao (1999) used FMOLS to construct a test statistic given as

$$R_T = \max_{i=1,\ldots,T}\frac{i}{\sqrt{\Omega_{\mu\cdot\upsilon}T}}\left|\frac{1}{i}\sum_{t=1}^{i}\hat{\mu}_t^{+} - \frac{1}{T}\sum_{t=1}^{T}\hat{\mu}_t^{+}\right|$$

where $\hat{\mu}_t^{+}$ are the residuals as given in Eq. (7) and $\Omega_{\mu.\upsilon}$ is given as in Eq. (6). Critical values of R_T can be found in Xiao (1999).

Table 1 lists the cointegration tests to be compared with their abbreviations to be used in the rest of paper.

Table 1. Cointegration tests and their abbreviations.

S. no	Name of test	Abbreviation used
1	LM test of cointegration based on KPSS statistic	LMKPSS
2	Leybourne and McCabe's LBI test of cointegration	LMLBI
3	Shin's C test of cointegration	SC
4	McCabe–Leybourne–Shin test of cointegration	MLS
5	Hausman H_1 test of cointegration	HH1
6	Hausman H_2 test of cointegration	HH2
7	Hansen's L_c test of cointegration	HLC
8	Xiao fluctuation test of cointegration	XFT

2.2. Artificial data generation (ADG)

The model in artificial data generation is derived from Jansson (2005) with some modifications for deterministic part. Let us consider time series y_t and x_t of length T, following the DGP given by

$$\begin{aligned} y_t &= D_t\delta' + x_t + \upsilon_t \\ x_t &= x_{t-1} + \mu_t^x \\ \upsilon_t &= \upsilon_{t-1} - \varphi\mu_{t-1}^y + \mu_t^y \end{aligned} \tag{17}$$

$\mu_t = (\mu_t^y, \mu_t^x)$ and $\mu_t : N(0, \Sigma)$ where Σ is an identity matrix of same order as μ_t. Under null hypothesis of cointegration and alternative hypothesis of no cointegration

$$\begin{array}{ll} H_0 : \varphi = 1 & \text{(Cointegration)} \\ H_A : 0 \leq \varphi < 1 & \text{(No cointegration)} \end{array}$$

In our study, we have taken a set of values of φ under alternative hypothesis and this set is

$$\varphi = (0, 0.1, 0.2, 0.3, 0.4, --------0.9)$$

D_t denotes the deterministic part comprising of intercept and linear time trend i.e.,

$$D_t = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ \bullet & \bullet \\ \bullet & \bullet \\ \bullet & \bullet \\ 1 & T \end{bmatrix}$$

and δ is its coefficient vector. In our study we are considering three cases of deterministic part,

i. Without intercept and linear time trend (*denoted as* D^0T^0): For this case $\delta = [0 \quad 0]$
ii. With intercept and without linear time trend (denoted as D^1T^0): For this case $\delta = [1 \quad 0]$
iii. With intercept and linear time trend (*denoted as* D^1T^1): For this case $\delta = [1 \quad 1]$

There are 10 point alternative hypotheses, one null hypothesis, three combinations of deterministic part and four sample sizes, i.e., $T = 30,\ T = 60,\ T = 120$ *and* $T = 240$

considered in this study. It means, 132 different artificial data generation processes have been used to assess the performance of each test.

2.3. MCS design for empirical size of test α

There will be two α's calculated for each test: one using the asymptotic critical values and the other using the simulated. However, their MCS design is same as:

i. Data are generated using ADG under null hypothesis.
ii. Tests statistic is calculated for this generated data.
iii. Decision (rejection of null or not) is made on basis of asymptotic critical value or simulated one, depending upon the α calculated.
iv. Above three steps are repeated for a fixed Monte Carlo sample size say M and number of rejections are counted.
v. The proportion of rejections out of M in terms of percentage is the α.

2.4. MCS design for simulated critical value

In this study, significance level for testing the null hypothesis of cointegration is taken as the commonly and routinely used one, i.e., 5%. To obtain simulated critical value or values of a test, following design is followed:

i. Data are generated using ADG under null hypothesis.
ii. Tests statistic is calculated for this generated data.
iii. Above two steps are repeated for a fixed Monte Carlo sample size M and test statistics are saved in an array say A.
iv. According to nature of test, simulated critical value is found. If the test is two tailed then 2.5th and 97.5th percentiles of A are lower and upper simulated critical values, if the test is right tailed, then 95th percentile of A is simulated critical value and if the test is left tailed then 5th percentile of A is simulated critical value.

2.5. MCS design for power

To obtain power of a test, following steps have been used:

i. Data are generated using ADG under a point alternative hypothesis.
ii. Tests statistic is calculated for this generated data.
iii. Decision (rejection of null or not) is made on basis of simulated critical value.
iv. Above three steps are repeated for a fixed M and number of rejections are counted.
v. As power of a test is defined as

$$Power = P(Rejection\ H_0\ /\ H_0\ is\ False)$$

So, proportion of rejections out of M in terms of percentage is the power of test at that specific alternative.

2.6. Point optimal test

Jansson (2005) proposed a point optimal test for null of cointegration. He considered the model comprising of two series y_t and x_t as

$$\begin{aligned} y_t &= x_t + \upsilon_t \\ x_t &= x_{t-1} + \mu_t^x \\ \upsilon_t &= \upsilon_{t-1} - \theta\mu_{t-1}^y + \mu_t^y \end{aligned}$$

$\mu_t = (\mu_t^y, \mu_t^x)$ and $\mu_t \sim N(0, \Sigma)$ where Σ is a positive definite matrix. The null hypothesis of cointegration and alternative hypothesis of no cointegration are

$$\begin{aligned} H_0 &: \theta = 1 && \text{(Cointegration)} \\ H_A &: 0 \le \theta < 1 && \text{(No cointegration)} \end{aligned}$$

Jansson (2005) assumed $H_A : \theta = \theta^*$ where $0 \le \theta^* < 1$ and $R = (x_t, D_t)$ where D_t denotes the deterministic component (intercept and trend). He partitioned Σ in conformity with μ_t as $\Sigma = \begin{bmatrix} \sigma_{yy} & \sigma'_{xy} \\ \sigma_{xy} & \Sigma_{xx} \end{bmatrix}$ and defined $\Psi_\theta = \Psi_\theta^{1/2}\Psi_\theta^{1/2'}$ where $\Psi_\theta^{1/2}$ is a lower triangular matrix of order $T \times T$ given as

$$\Psi_\theta^{1/2} = \begin{bmatrix} 1 & 0 & 0 & \bullet & \bullet & 0 \\ 1-\theta & 1 & 0 & \bullet & \bullet & 0 \\ 1-\theta & 1-\theta & 1 & \bullet & \bullet & 0 \\ \bullet & \bullet & \bullet & \bullet & \bullet & 0 \\ \bullet & \bullet & \bullet & \bullet & \bullet & 0 \\ 1-\theta & 1-\theta & 1-\theta & \bullet & \bullet & 1 \end{bmatrix}$$

$$\sigma_{yy\bullet x} = \sigma_{yy} - \sigma_{xy}{}'\Sigma_{xx}{}^{-1}\sigma_{xy}$$

$$L(\theta) = \log|R'\Psi_\theta^{-1}R| + \sigma_{yy\bullet x}{}^{-1}Y_\theta{}'\left(\Psi_\theta^{-1} - \Psi_\theta^{-1}R(R'\Psi_\theta^{-1}R)^{-1}R'\Psi_\theta^{-1}\right)Y_\theta$$

Where $Y_\theta = Y - \theta\ \Psi_\theta^{-1/2}x_t\Sigma_{xx}{}^{-1}\sigma_{xy}$.
According to Jansson (2005) that the test statistic PO based on log-likelihood ratio is the point optimal test based on Neyman Pearson Lemma (Neyman and Pearson 1992).

$$PO_{\theta^*} = L(1) - L(\theta^*)$$

2.7. Stringency criterion

The shortcoming of a test a (SC_m^a) at a point alternative hypothesis m is the difference of power of point optimal test (PO_m) and the power of the test a (P_m^a) at that alternative m.

$$\mathrm{SC}_m^a = \mathrm{PO}_m - P_m^a \quad \forall\ m = 1,\ 2,\ 3,\ \ldots\ldots\ldots\ldots,\ k \text{ and } a = 1,\ 2,\ 3,\ \ldots\ldots\ldots\ldots,\ l$$

The maximum shortcoming (MSC) or stringency of a test a (Π^a) is

$$\Pi^a = \max_m\ (\mathrm{SC}_m^a)$$

The most stringent test will be the test having the minimum of these maximum shortcomings, i.e.,

$$\text{Most stringent test} = \min_{a} \ (\Pi^a)$$

For detailed discussion on stringency criterion, see (Zaman et al. 2017).

The cointegration tests are classified into three, depending upon their maximum shortcomings. If a test has maximum shortcoming of 30% or less, then it is categorized as better performer. If a test has maximum shortcoming of 50% or less but greater than 30% then it is classified as average or mediocre performer and if a test has maximum shortcoming of greater than 50%, then it is categorized as worst performer.

2.8. An empirical example

In order to investigate the cointegration between household consumption expenditure (denoted by C hereafter) and gross domestic product, i.e., GDP (denoted by Y hereafter) for the USA, quarterly data in billions of current US $ are retrieved from World Bank IFS Data set, from 1957Q1 to 2013Q3 ($T = 227$). The cointegration between these two series was also investigated by Shin (1994). The natural logarithms of these two series (LnC and LnY) are obtained for further analysis. First two tests, investigating whether the two series have unit root or stationary are used and these two tests are KPSS stationary test and ADF unit root test. Then the three relatively better performing tests from eight tests with null of cointegration and one test with null of no cointegration, i.e., Philips Ouliaris' Z_α test Phillips and Ouliaris (1990) are used to assess the existence of cointegration between LnC and LnY.

3. Discussion of results

The eight cointegration tests are compared using four sample sizes of $T =$ 30, 60, 120 and 240 at three specifications of deterministic part (D^0T^0, D^1T^0 and D^1T^1) and for eleven points of null and alternative hypotheses (One null and ten point alternative hypotheses). To obtain size, power and simulated critical values of tests, 50,000 simulations have been carried out.

It is evident from Table 2 that size of tests is not controlled around nominal size of 5% when asymptotic critical values are used. There are only 14 (marked as BOLD in Table 2) out of 96 cases when the size is controlled around nominal size of 5%. LMKPSS faces over rejection problem for D^0T^0, has controlled size for D^1T^0 and faces problem of under rejection for D^1T^1. The two Hausman tests HH1 and HH2 have their sizes way smaller than the nominal size. In same manner, from the remaining five tests, two tests XFT and HLC have uncontrolled size at all sample sizes for three specifications of deterministic part. SC has under rejection problem for D^0T^0, has controlled size for D^1T^0 and has controlled size for D^1T^1 also, except at $T = 30$. In general, LMLBI tends to control the size with increase in sample size, hence it attains controlled size for D^1T^0 at $T = 240$. But this the only case where LMLBI has controlled size. Similarly, MLS also tends to decrease its size with increase in sample size, but it has controlled size for only two cases.

As the size of tests is not controlled around nominal size of 5% when asymptotic critical values are used, so simulated critical values are obtained and then using these simulated critical values again, sizes of tests are estimated and displayed in Table 3. It is

Table 2. Size in % for asymptotic critical values.

	D^0T^0				D^1T^0				D^1T^1			
	Sample size T				Sample size T				Sample size T			
Tests	30	60	120	240	30	60	120	240	30	60	120	240
LMKPSS	22.71	22.56	21.94	22.32	**5.89**	**5.34**	**5.75**	**5.33**	0.02	0.07	0.04	0.02
HH1	0.76	0.84	0.65	0.49	0.23	0.46	0.51	0.61	0.33	1.14	1.48	1.71
HH2	1.11	0.74	0.63	0.55	0.54	0.61	0.63	0.64	1.93	1.60	1.66	1.77
SC	0.47	1.85	2.38	2.51	**6.38**	**5.36**	**4.98**	**5.28**	20.34	**5.72**	**5.20**	**5.01**
LMLBI	34.07	28.4	24.64	23.84	21.34	12.99	7.97	5.94	10.68	**3.75**	0.55	0.05
XFT	0.31	2.23	1.84	1.82	18.72	17.41	18.32	18.71	41.32	43.74	44.34	40.91
HLC	53.24	43.52	29.84	15.52	55.8	40.74	33.42	17.72	58.98	42.08	34.14	21.88
MLS	65.81	60.91	58.62	54.23	58.91	41.21	33.51	29.82	36.41	11.71	**6.51**	**6.32**

Note: Marked BOLD shows that size is controlled around nominal size of 5%.

Table 3. Size in % for simulated critical values.

	D^0T^0				D^1T^0				D^1T^1			
	Sample size T				Sample size T				Sample size T			
Tests	30	60	120	240	30	60	120	240	30	60	120	240
LMKPSS	4.76	4.99	5.11	4.78	5.14	5.38	4.82	4.79	4.54	4.67	4.68	4.87
HH1	5.23	5.02	4.91	4.43	5.06	4.81	5.02	5.01	4.08	5.07	5.11	5.05
HH2	5.09	5.58	4.84	4.85	5.18	5.13	4.42	4.91	5.61	5.12	4.72	4.84
SC	5.09	5.29	5.19	4.73	4.85	5.48	5.21	5.01	5.42	4.98	4.85	4.81
LMLBI	4.89	5.05	4.98	5.11	4.92	5.13	4.97	5.13	4.16	5.17	4.52	4.73
XFT	4.67	4.41	5.25	4.72	3.91	5.17	5.42	4.82	5.36	5.42	5.24	4.59
HLC	4.42	4.78	4.92	5.04	4.84	4.92	4.96	4.91	4.78	5.02	5.34	5.44
MLS	3.73	4.31	4.54	6.46	4.78	4.67	6.32	3.81	3.82	5.35	5.74	5.72

clear from Table 3 that now all eight tests have controlled size around nominal size of 5%. These simulated critical values are used further in power comparison based on stringency criterion.

For the power comparison, first the simplest model, i.e., D^0T^0 is considered and the maximum shortcomings for this case at all sample sizes are displayed in Table 4. It is clearly portrayed from Table 4 that from two residual based tests, one test (LMKPSS) is consistently a better performer at all sample sizes as it has stringencies of 30.12%, 26.88%, 22.78% and 14.37% at $T = 30$, 60, 120 and 240, respectively. The other residual based test i.e., LMLBI is a worst performer up to sample size of 60 (having stringency of 87.82% and 68.13% at $T = 30$ and 60) and an average performer at sample sizes of 120 and 240 (having stringencies of 51.17% and 35.77%, respectively). From three tests based on DOLS estimation a single test, i.e., SC is a worst performer up-to sample size of 60 (stringencies of 76.81% and 56.96% at $T = 30$ and 60, respectively), an average performer at sample size of 120 (stringency of 36.97%) and a better performer (stringency of 26.59%) at sample size of 240. The other two tests based on DOLS estimation, i.e., HH1 and HH2are worst performers (stringencies are greater than 65%) at all sample sizes. Similarly, from two tests based on FMOLS estimation one test i.e., HLC is a worst performer up-to sample size of 60 (stringencies of 90.01% and 71.24% at $T = 30$ and 60) and an average performer at sample sizes of 120 and 240 (stringencies of 49.62% and 33.96%). The other test based on FMOLS estimation i.e., XFT is worst performer at all sample sizes (stringencies greater than 84%). The single test based on

Table 4. Maximum shortcomings in % of cointegration tests for D^0T^0.

Tests	$T = 30$	$T = 60$	$T = 120$	$T = 240$
LMKPSS	30.12**	26.88**	22.78**	14.37**
SC	76.81	56.96	36.97*	26.59**
LMLBI	87.82	68.13	51.17*	35.77*
HLC	90.01	71.24	49.62*	33.96*
HH2	66.86	70.05	73.20	81.42
HH1	70.85	70.63	74.38	81.98
XFT	84.80	87.80	87.10	88.10
MLS	87.50	84.30	88.70	86.40

Note: ** and * shows that a test is better and average performer, respectively.

Table 5. Maximum shortcomings in % of cointegration tests for D^1T^0.

Tests	$T = 30$	$T = 60$	$T = 120$	$T = 240$
LMKPSS	23.53**	16.81**	15.34**	10.90**
SC	80.11	55.26	32.74*	20.67**
LMLBI	93.63	78.66	49.51*	29.03**
HH2	73.13	71.06	78.67	83.91
XFT	76.60	82.04	83.01	81.40
HH1	78.59	72.97	78.94	83.63
MLS	80.60	64.70	57.30	54.30
HLC	93.52	88.02	75.88	62.68

Note: ** and * shows that a test is better and average performer, respectively.

Table 6. Maximum shortcomings in % of cointegration tests for D^1T^1.

Tests	$T = 30$	$T = 60$	$T = 120$	$T = 240$
LMKPSS	18.90**	13.88**	11.53**	12.15**
SC	88.62	57.64	34.24*	23.06**
MLS	94.60	75.40	49.20*	41.60*
LMLBI	91.66	94.99	62.92	31.60*
XFT	80.50	80.50	86.70	84.10
HH2	86.12	82.60	84.96	88.29
HH1	88.73	83.67	86.20	88.02
HLC	93.14	95.70	92.60	83.98

Note: ** and * shows that a test is better and average performer, respectively.

ML estimation, i.e., MLS is also a worst performer at all sample sizes (stringencies greater than 84%).

Moving to the second case of deterministic part, i.e., D^1T^0, it is evident from Table 5 that only one test, i.e., LMKPSS which is residual based is a better performer at all sample sizes (stringencies of 23.52%, 16.81%, 15.34% and 10.90% at $T = 30$, 60, 120 and 240, respectively). The other residual based test i.e., LMLBI is worst performer up-to sample size of 60 (stringencies of 93.63% and 78.66%), an average performer at sample size of 120 (stringency of 49.51%) and a better performer at sample size of 240 (stringency of 29.03%). From three tests based on DOLS estimation two (HH1 and HH2) are worst performers at all sample size (stringencies greater than 70%) and one, i.e., SC is worst performer up-to sample size of 60 (stringencies of 80.11% and 52.26%), an average performer at sample size of 120 (stringnecy of 32.74%) and a better performer at sample size of 240 (stringency of 20.67%). Rest of three tests out which two are based on FMOLS estimation (HLC and XFT) and one is based on ML estimation i.e., MLS are all worst performers at all four sample sizes (stringencies are greater than 64%).

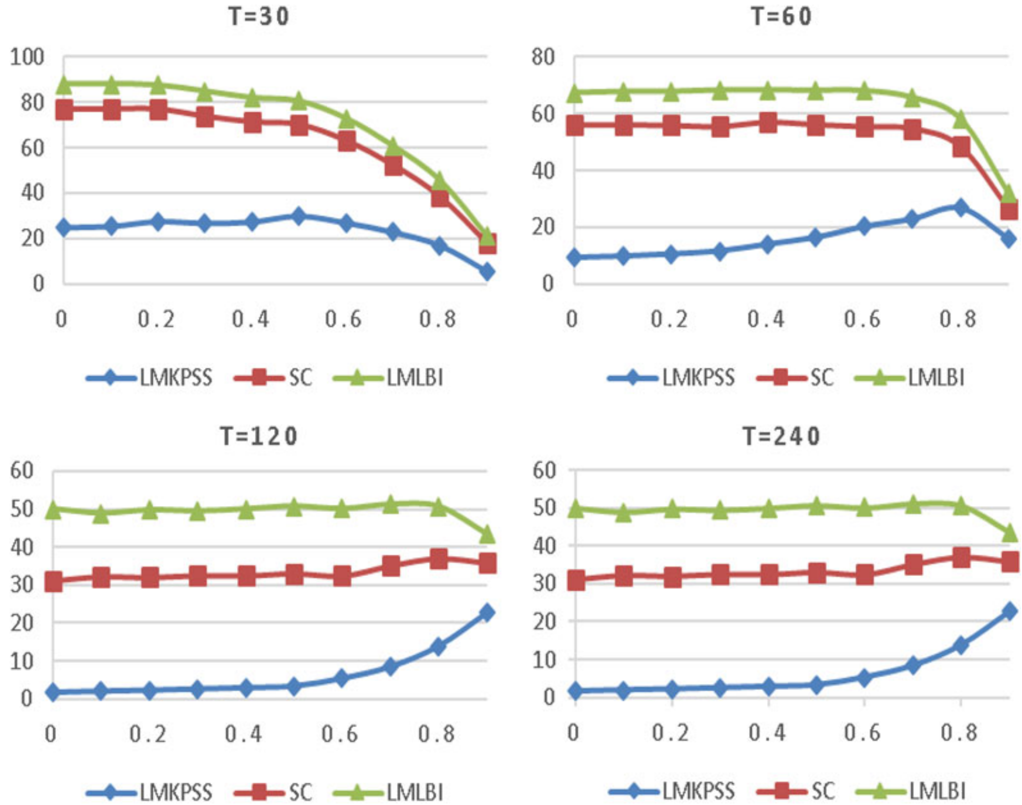

Figure 1. Shortcomings for D^0T^0. Note: Along *x*-axis are the parameter θ and along *y*-axis are the shortcomings in %.

For the last case of deterministic part, i.e., D^1T^1, maximum shortcomings of the tests are displayed in Table 6. It is clear that again a sole residual based test, i.e., LMKPSS is a better performer at all sample sizes (stringencies of 18.90%, 13.88%, 11.53% and 12.15% at $T = 30$, 60, 120 and 240, respectively). The other residual based test, i.e., LMLBI is worst performer up-to sample size of 120 (stringencies greater than 62%). However, it is an average performer at sample size of 240 (stringency of 31.60%). From three tests based on DOLS estimation two tests, i.e., HH1 and HH2 are worst performers at all sample sizes (stringencies are greater than 82%). However, the third one, i.e., SC is worst performer up-to sample size of 60 (stringencies are 88.62% and 57.64%), an average performer at sample size of 120 (stringency of 34.24%) and a better performer at sample size of 240 (stringency of 23.06%). The single test based on ML estimation, i.e., MLS is worst performer up-to sample size of 60 (stringencies of 94.60% and 75.40%) and an average performer at sample sizes of 120 and 240 (stringencies of 49.20% and 41.60%). The rest of two tests based on FMOLS estimation, i.e., HLC and XFT are worst performers at all sample sizes (stringencies are greater than 80%).

Three tests i.e., LMKPSS, SC and LMLBI are selected for further discussion as they are performing relatively way better than the rest of five. The shortcomings of these three tests are plotted against the parameter θ, representing the alternative hypotheses. For the first case of deterministic part i.e., D^0T^0, the shortcomings are displayed in

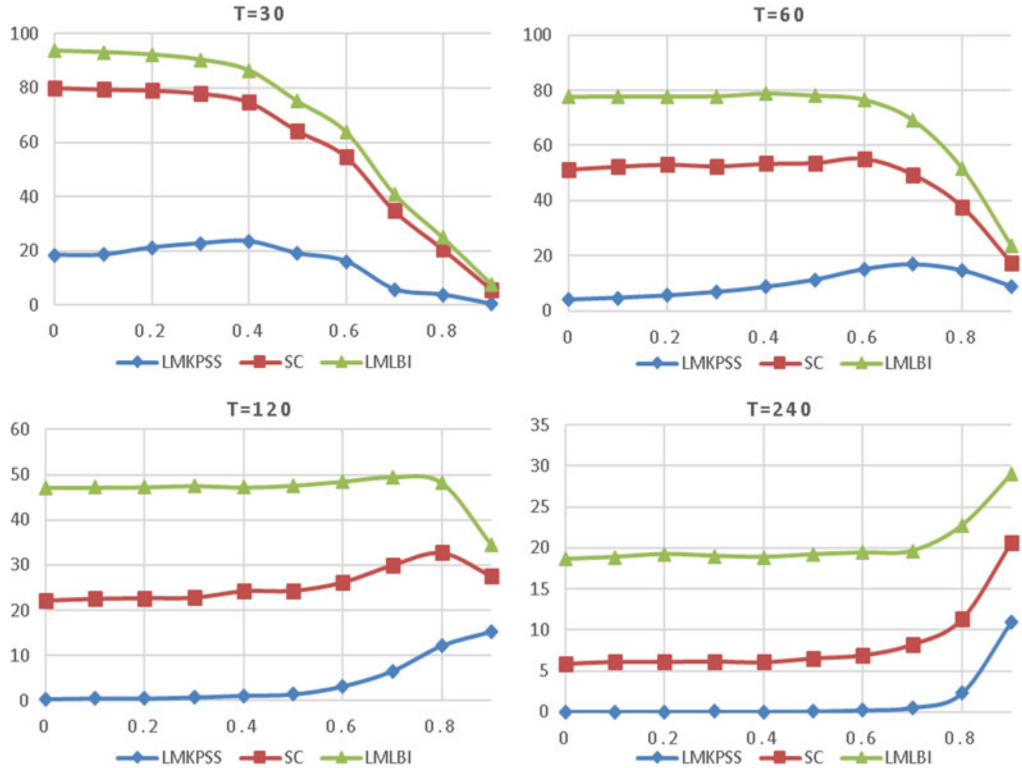

Figure 2. Shortcomings for D^1T^0. Note: Along *x*-axis are the parameter θ and along *y*-axis are the shortcomings in %.

Figure 1. It is evident from the Figure 1 that at T = 30, all three tests have increasing shortcomings with decrease in parameter θ. However, LMKPSS test's shortcomings are less than 30% for the entire paramter space. At T = 60, two tests i.e., LMLBI and SC have increasing shortcomings for the entire parameter space. However LMKPSS has increasing shortcomings in [0.8, 0.9] but has decreasing shortcomings in [0,0.8). For $T = 120$ and 240, LMLBI and SC have shortcomings between 30% and 50% for the whole parameter space. However, LMKPSS has sharply decreasing shortcomings with decreasing parameter in whole parameter space.

For the second and third cases of deterministic part, i.e., D^1T^0 and D^1T^1, it is evident from Figures 2 and 3, that at $T = 30$, again the three tests have increasing shortcoming with decreasing parameter value in parameter space. However, LMKPSS test's shortcomings are never the less lesser than 20% in the entire parameter space. The shortcomings of LMKPSS decrease sharply with increase in parameter value and it approaches to zero for [0,0.4) for $T = 120$ and [0, 0.6) for $T = 240$.

Before investigating the existence or absence of cointegration between LnC and LnY, these two series are assessed for stationarity and unit root by two tests, i.e., KPSS test of stationary null hypothesis and augmented dickey fuller (ADF) test of unit root null hypothesis. For KPSS two specifications are taken; one is the user specified value of bandwidth, i.e., $l = 10$, (following (Shin 1994)) and the other is Automatic bandwidth l using Newey–West method. However, when using the automatic bandwidth, it turns out its value is $l = 11$. For both types, the Bartlett Kernel has been used to estimate the

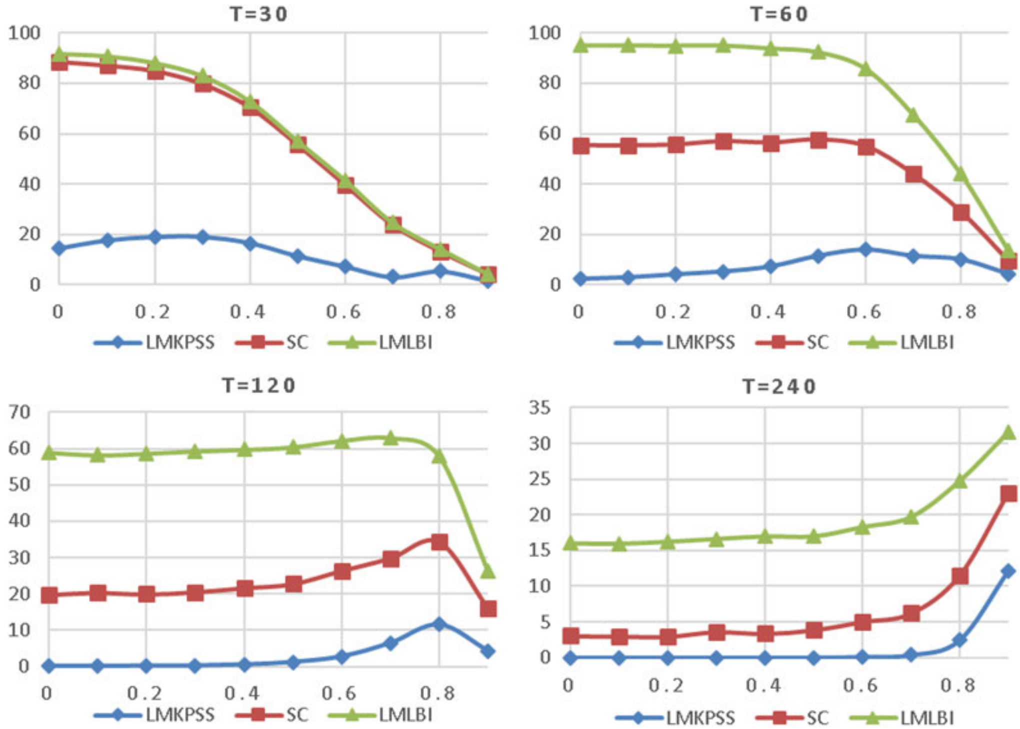

Figure 3. Shortcomings for D^1T^1. Note: Along *x*-axis are the parameter θ and along *y*-axis are the shortcomings in %.

long run variance. For ADF test five augmentations are taken (following (Shin 1994)). For each test, two cases of deterministic part, i.e., with constant and with constant and trend (trend stationary) are taken. The results of the stationarity and unit root tests are displayed in Table 7. There is strong evidence that both of the series, i.e., LnC and LnY have unit root except one case. This exceptional one case is according to ADF, LnY is stationary at 10% level of significance when only there is intercept in deterministic part. However, even LnY is unit root at 5% level of significance.

The cointegration between LnC and LnY has been assessed using three relatively better performing tests (LMKPSS, SC and LMLBI) with null of cointegration and one test with null of no cointegration, i.e., Philips Ouliaris' Z_α test (Phillips and Ouliaris 1990), denoted by PO Z_α, hereafter. Two specifications of deterministic component, i.e., demeaned (D^1T^0) and detrended (D^1T^1) have been analyzed (following (Shin 1994)). The results of the four tests are displayed in Table 8. It is evident from Table 8 that for detrended case all three tests with null of cointegration reach to the same conclusion of existence of cointegration between LnC and LnY. The PO Z_α also rejects the null hypothesis of no cointegration at 10% level of significance and concludes in favor of existence of cointegration between LnC and LnY. However, for the Demeaned case, SC and LMLBI concludes in favor of existence of cointegration, while, LMKPSS concludes in favor of existence of cointegration at 1% and 5% level of significance, but it also concluded against the existence of cointegration at 10% level of significance. The PO Z_α has the same conclusion as for the detrended case.

Table 7. Stationarity and unit root tests.

Tests	Stationarity test				Unit root test	
	KPSS ($l = 10$)		KPSS (automatic l)		ADF	
Series/deterministic part	Constant	Constant and trend	Constant	Constant and trend	Constant	Constant and trend
LnY	2.1485***	0.4797***	1.9774***	0.4427***	−2.7908*	1.3984
LnC	2.1515***	0.4654***	1.9802***	0.4295***	−2.3152	1.1047

Note: ***, ** and * represent the rejection of null hypothesis at 1%, 5% and 10% level of significance, respectively.

Table 8. Tests for cointegration and no cointegration.

	D^1T^0		D^1T^1	
Tests/deterministic component	Test statistic	p Value	Test statistic	p Value
LMKPSS	1.1573*	0.0831	0.698	0.1011
SC	0.1323	0.3113	0.0857	0.1488
LMLBI	0.2085	0.1872	0.0992	0.2268
PO Z_α	−18.9345*	0.0635	−23.9634*	0.079

Note: * represents the rejection of null hypothesis at 10% level of significance.

4. Conclusions and recommendations

The use of asymptotic critical values generally produces size distortions as most of tests have either too much high or too much low sizes as compared to the nominal size. However, the use of simulated critical values controls the sizes of tests around nominal size. Therefore, it is highly recommended that asymptotic critical values may not be used and only simulated critical values may be used in empirical studies.

As for as the power comparison of tests is concerned, if both nuisance parameters, i.e., intercept and linear time trend are absent from cointegrating equation (D^0T^0) then from eight tests, a residual based test, i.e., LMKPSS is the sole better performer at all sample sizes as it has stringencies lesser than 30% and with decrease in parameter value in parameter space [0,0.9], it narrows down the difference between its power and maximum possible power and at larger sample sizes of 120 and 240, this difference approaches zero. Another test based on DOLS estimation, i.e., SC is only better performer at large sample sizes of 240 or higher. Four tests (HH1, HH2, XFT and MLS) are worst performers at all sample sizes. Similarly, if only one of the nuisance parameters i.e., intercept is present in cointegrating equation (D^1T^0) then from eight tests, again the same residual-based test, i.e., LMKPSS is the sole better performer at all sample sizes. The LMKPSS test has shortcomings around or majority of times lesser than 20% at all sample sizes and specially the difference between its powers and possible maximum power decreases with decrease of the parameter in the entire parameter space. Another residual based test, i.e., LMLBI along with a test based on DOLS estimation, i.e., SC are only better performers at large sample sizes of 240 or higher. The remaining five tests are worst performers at all sample sizes. In same manner, if both nuisance parameters are present in cointegrating equation then again, the same residual-based test, i.e., LMKPSS is the sole better performer at all sample sizes. However, another test based on DOLS estimation, i.e., SC is only a better performer at large sample sizes of 240 or higher. LMKPSS test has the shortcomings lesser than 20% for sample sizes of 30 and 60, however, at larger sample sizes of 120 and 240, the difference between its powers and maximum possible power even approaches zero with decreasing parameter value in parameter space. The remaining tests are worst performers. Hence, LMKPSS is the most stringent test at all sample sizes for all three specifications of deterministic component as the difference between its powers and maximum possible power is the minimum and even it approaches zero for the parameter in [0, 0.6). Therefore, the use of LMKPSS is highly recommended for any sample size and any specification of deterministic component. In addition to LMKPSS, LMLBI and SC may be used for very large sample sizes of 240 or greater. However, the use of remaining five tests may be avoided in empirical studies.

The empirical example assessing the cointegration between household consumption expenditure and GDP of USA, using three cointegration tests with null of cointegration (LMKPSS, SC and LMLBI), performing relatively better in our MCS comparison and one test with null of no cointegration, i.e., Philips Ouliaris' Z_α test, concludes that the two series are cointegrated for most of cases and specifications.

ORCID

Waqar Muhammad Khan http://orcid.org/0000-0003-0938-8075

References

Banerjee, A., J. J. Dolado, D. F. Hendry, and G. W. Smith. 2009. Exploring equilibrium relationships in econometrics through static models: some Monte Carlo evidence*. *Oxford Bulletin of Economics and Statistics* 48 (3):253–77. doi:10.1111/j.1468-0084.1986.mp48003005.x.

Boswijk, P., and P. H. Franses. 1992. Dynamic specification and cointegration*. *Oxford Bulletin of Economics and Statistics* 54 (3):369–81. doi:10.1111/j.1468-0084.1992.tb00007.x.

Engle, R. F., and C. W. J. Granger. 1987. Co-integration and error correction: Representation, estimation, and testing. *Econometrica* 55:251–76. doi:10.2307/1913236.

Fernández-Macho, F. J., and P. Mariel. 1994. *Testing the Null of Cointegration, Hausman-like Tests for Regressions with a Unit Root.* Universidad Del País Vasco, Facultad de Ciencias Económicas.

Gabriel, V. J. 2003. Tests for the null hypothesis of cointegration: a Monte Carlo comparison. *Econometric Reviews* 22 (4):411–35. doi:10.1081/ETC-120025897.

Gonzalo, J., and T.-H. Lee. 1998. Pitfalls in testing for long run relationships. *Journal of Econometrics* 86 (1):129–54. doi:10.1016/S0304-4076(97)00111-5.

Hansen, B. E. 1992. Tests for parameter instability in regressions with I (1) processes. *Journal of Business & Economic Statistics* 10:321–35. doi:10.2307/1391545.

Haug, A. A. 1996. Tests for cointegration a Monte Carlo comparison. *Journal of Econometrics* 71 (1-2):89–115. doi:10.1016/0304-4076(94)01696-8.

Islam, T. U. 2017. Stringency-based ranking of normality tests. *Communications in Statistics - Simulation and Computation* 46 (1):655–68. doi:10.1080/03610918.2014.977916.

Jansson, M. 2005. Point optimal tests of the null hypothesis of cointegration. *Journal of Econometrics* 124 (1):187–201. doi:10.1016/j.jeconom.2004.02.011.

Johansen, S., and K. Juselius. 2009. Maximum likelihood estimation and inference on cointegration—with applications to the demand for money. *Oxford Bulletin of Economics and Statistics* 52 (2):169–210. doi:10.1111/j.1468-0084.1990.mp52002003.x.

Khan, A. ul I., and A. Zaman. 2017. *Theoretical and Empirical Comparisons of Cointegration Tests*. Islamabad, Pakistan: International Islamic University.

Khan, W. M., and A. ul I. Khan. 2018. Most stringent test of independence for time series. *Communications in Statistics – Simulation and Computation*, 1–19. doi:10.1080/03610918.2018.1527350.

Kremers, J. J. M., N. R. Ericsson, and J. J. Dolado. 1992. The power of cointegration tests. *Oxford Bulletin of Economics and Statistics* 54 (3):325–48. doi:10.1111/j.1468-0084.1992.tb00005.x.

Kwiatkowski, D., P. C. B. Phillips, P. Schmidt, and Y. Shin. 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 (1–3):159–78. doi:10.1016/0304-4076(92)90104-Y.

Leybourne, S. J., and B. P. M. McCabe. 2009. Practitioners corner: a simple test for cointegration. *Oxford Bulletin of Economics and Statistics* 56 (1):97–103. doi:10.1111/j.1468-0084.1994.mp56001008.x.

Mariel, P. 1996. A comparison of cointegration tests. *Applied Mathematics* 41:411–31.

McCabe, B. P. M., S. J. Leybourne, and Y. Shin. 1997. A parametric approach to testing the null of cointegration. *Journal of Time Series Analysis* 18 (4):395–413. doi:10.1111/1467-9892.00058.

Neyman, J., and E. S. Pearson. 1992. *On the Problem of the Most Efficient Tests of Statistical Hypotheses*. New York: Springer.

Pekmezci, A., and M. Dilek. 2016. The comparison of performances of widely used cointegration tests. Communications *in Statistics – Simulation and Computation* 45 (6):2070–80. doi:10.1080/03610918.2014.889157.

Pesaran, M. H., Y. Shin, and R. J. Smith. 2001. Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics* 16 (3):289–326. doi:10.1002/jae.616.

Pesavento, E. 2004. Analytical evaluation of the power of tests for the absence of cointegration. *Journal of Econometrics* 122 (2):349–84. doi:10.1016/j.jeconom.2003.10.025.

Phillips, P. C. B., and B. E. Hansen. 1990. Statistical inference in instrumental variables regression with I (1) processes. *The Review of Economic Studies* 57 (1):99–125. doi:10.2307/2297545.

Phillips, P. C. B., and S. Ouliaris. 1990. Asymptotic properties of residual based tests for cointegration. *Econometrica* 58 (1):165–93. doi:10.2307/2938339.

Shin, Y. 1994. A residual-based test of the null of cointegration against the alternative of no cointegration. *Econometric Theory* 10 (1):91–115. doi:10.1017/S0266466600008240.

Xiao, Z. 1999. A residual based test for the null hypothesis of cointegration. *Economics Letters* 64 (2):133–41. doi:10.1016/S0165-1765(99)00079-8.

Zaman, A., A. Zaman, and A. Rehman. 2017. The concept of stringency for test comparison: the case of a Cauchy location parameter. International *Econometric Review* 9:1–18. doi:10.33818/ier.319909.