European Online Journal of Natural and Social Sciences 2018;
Vol.7, No 2 pp. 494-500
ISSN 1805-3602

www.european-science.com

Comparison of Residual based Cointegration Tests: Evidence from Monte Carlo

Asad Ul Islam Khan[1]*, Yasir Riaz[1], Ali Raza[2], Mudassar Rashid[3]
[1]International Islamic University, Islamabad, Pakistan; [2]University of Haripur, Khyber Pakhtunkhwa, Pakistan; [3]COMSATS University Islamabad, Pakistan
*Email: asadkhan1983@gmail.com
Tel.: +923245277926

Received for publication: 14 March 2018.
Accepted for publication: 04 May 2018.

Abstract

In this article ten cointegration tests based on residuals of cointegrating equation are compared on basis of stringency criterion: a robust technique for comparison of tests using Monte Carlo simulations. Two tests i.e. Phillips and Ouliaris' $\hat{P}_u$ and Choi Durbin-Hausman statistic are the leading performers and are recommended for any sample size. The remaining eight tests are recommended for only large sample sizes of 200 or greater. The use of all these ten tests is not recommended when presence of both intercept and linear time trend is assumed in cointegrating equation unless the sample size is very large i.e. greater than 200.

Keywords: Comparison; Cointegration Tests; Monte Carlo; Stringency

Introduction

Since the path breaking paper by (Engle and Granger 1987), a variety of cointegration tests have been proposed to assess the long run relationship between economic variables and specially macroeconomic variables. Most of these tests are residual based techniques that are the enhancements of unit root tests like: Engle and Granger's Augmented Dickey-Fuller (EGADF) test (Engle and Granger 1987), Phillips and Ouliaris' $\hat{Z}_\alpha$ (POZA) test (Phillips and Ouliaris, 1990) and many more. However, some proposed tests were based on system estimation like: Johansen Maximum Eigen Value (JME) and Johansen Trace (JT) tests (Johansen and Juselius 1990). Several tests were developed on the mechanism of error correction model like: Autoregressive Distributed Lag Bounds (ADLB) test (Pesaran, Shin et al., 2001) and Boswijk Wald (BW) test (Boswijk 1989). All this variety of tests created a dilemma of contradictory results as one test tells us that two variables are cointegrated while the other says they are not. Therefore, to provide guidelines to practitioners and applied researchers, several comparisons like (Bewley and Yang, 1998), (Gabriel, 2003), (Haug, 1996), (Kremers, Ericsson et al., 1992), (Mariel, 1996), (Pesavento, 2004) and many more have been carried out to assess the size and power properties of cointegration tests. (Banerjee, Dolado et al., 1986) compared two tests i.e. Cointegrating Regression Durbin-Watson (CRDW) and t-test on error correction term of an error correction model (TECM) and they found that the later has better powers with slightly high size. (Kremers, Ericsson et al., 1992) evaluated the size and power properties of EGADF, CRDW and TECM and concluded that EGADF has lesser powers than rest of two. (Boswijk and Franses, 1992) examined the performance of three tests (BW, EGADF and JME) and BW was found to be better test both in terms of size and power. In a comprehensive comparative study by (Haug, 1996) performance of nine tests was evaluated and it was concluded that Phillips and Ouliaris $\hat{P}_z$ (POPZ) and POZA have higher powers with high size distortion. Similarly, (Mariel 1996) compared nine tests and concluded that tests with null of cointegration have lesser size distortions with lesser powers as compared to tests with null of no cointegration having higher powers with high size distortion. (Mariel 1996) recommended the use of both type of tests having null of no cointegration and null of cointegration for strong empirical evidence. (Pesavento, 2004) evaluated the performance of four tests on basis of analytic methods and Monte Carlo. (Pesavento, 2004) concluded that JME and TECM are better performers. In same manner, many more comparative studies were carried out, but these comparisons created no definite answer that which test is better than the rest. This is because all comparisons did not consider a universal set of alternative space and they used asymptotic critical values also. Hence, these comparisons failed to provide clear-cut guidelines to practitioner. As (Zaman, Zaman et al., 2017) pointed out that stringency is the robust technique for comparison of tests because it considers the whole alternate space, therefore, this article uses stringency criterion for comparison of ten residual based cointegration tests to fill the gap in literature. A detailed discussion of stringency criterion can be found in (Zaman 1996) and (Zaman, Zaman et al., 2017).

Tests to be Compared

The primary test with null hypothesis of no cointegration is Engle and Granger's Augmented Dickey-Fuller (EGADF) proposed by (Engle and Granger 1987). It is simply an Augmented Dickey Fuller (ADF) test of unit root on residuals μ_t of cointegrating Equation (1)

$$y_t = \delta\psi_t + \sum_{i=1}^{m}\beta_i x_{it} + \mu_t \qquad t=1,2,..........,T \qquad (1)$$

Two tests i.e. Phillips and Ouliaris' $\hat{Z}_\alpha$ (POZA) and Phillips and Ouliaris' $\hat{Z}_t$ (POZT) are using the long run variance of residuals μ_t of cointegrating Equation (1). Similarly, Phillips and Ouliaris $\hat{P}_u$ (POPU) is the variance ratio test of long run variance and contemporaneous variance of residuals of Eq. (1). While, Phillips and Ouliaris $\hat{P}_z$ (POPZ) is a trace statistic based on long run variance of residuals of Equation (1). The detailed discussion on all four Philips Ouliaris tests can be viewed in (Phillips and Ouliaris 1990).

(Hansen 1990) used (Cochrane and Orcutt 1949) procedure to obtain a bias adjusted estimate ρ^+ of autoregressive coefficient of residuals μ_t of cointegrating equation (1) and then ρ^+ is used to obtain bias adjusted residuals $\hat{\mu}_t^+$. (Hansen 1990) showed that simple t- test on autoregressive coefficient of $\hat{\mu}_t^+$, ADF and POZA statistics can be used on these bias adjusted residuals $\hat{\mu}_t^+$. These three tests are Hansen's Cochrane-Orcutt ρ^+ (HCO), Hansen's variation of the ADF (HADF) test and Hansen's variation of the $\hat{Z}_\alpha$ (HZA) test.

(Choi, 1994) proposed Durbin-Hausman statistic (CDHS) based on two estimators estimated from the residuals of Equation (1). According to (Sargan and Bhargava, 1983) a modified Durbin Watson statistic (Cointegrating Regression Durbin-Watson(CRDW)) from the residuals of Equation (1) may be used for an initial assessment of null of no cointegration.

Methodology

Data Generation

The data generating process is taken from (Jansson 2005).Consider two-time series, y_t and x_t of length *T*, that are generated by the process

$$y_t = D_t\delta' + x_t + \upsilon_t \qquad (2)$$
$$x_t = x_{t-1} + \mu_t^x \qquad (3),$$
$$\upsilon_t = \theta\upsilon_{t-1} + \mu_t^y \qquad (4)$$

$\mu_t = \left(\mu_t^y, \mu_t^x\right)$ and $\mu_t \sim N\left(0, \Sigma\right)$ where Σ represents an identity matrix of order in conformity with μ_t. D_t denotes the deterministic part comprising of intercept and trend i.e.

$$D_t = \begin{bmatrix} 1 & 1 & . & . & . & 1 \\ 1 & 2 & . & . & . & T \end{bmatrix}'$$

and δ is coefficient vector stating the nature of deterministic part. In this article three cases of deterministic part are considered i.e.

i. without Intercept and Linear Time Trend (I^0T^0): For this case $\delta = \begin{bmatrix} 0 & 0 \end{bmatrix}$,

ii. with Intercept and without Linear Time Trend (I^1T^0): For this case $\delta = \begin{bmatrix} 1 & 0 \end{bmatrix}$ and

iii. with Intercept and Linear Time Trend (I^1T^1): For this case $\delta = \begin{bmatrix} 1 & 1 \end{bmatrix}$.

For $\theta = 1$ in Eq. (4) y_t and x_t are generated under null hypothesis of no cointegration and for $0 \le \theta < 1$ in Eq. (4), y_t and x_t are generated under alternative hypothesis of cointegration. The values of θ are taken from the set $\{0, 0.1, 0.2, 0.3, ..., 0.9, 1\}$.

Stringency criterion

If Π_j^τ is used to denote the power of a test τ at a specific alternative hypothesis *j* for $\tau = 1, 2, -----, k$ and $j = 1, 2, -----, l$ where *"k"* represents total number of tests and *"l"* represents total number of point alternative hypothesis. Similarly, if Λ_j is used to represent the power of respective point optimal test at specific alternative hypothesis *j*, then shortcoming of a test τ i.e. Γ_j^τ at a point alternative hypothesis *j* is

$$\Gamma_j^\tau = \Lambda_j - \Pi_j^\tau$$

Stringency of a test τ i.e. Θ^τ is

$$\Theta^\tau = \max_{j=1,2,-----,l} \left(\Gamma_j^\tau\right)$$

The test with the minimum stringency from all tests is the most stringent test. See (Zaman, Zaman et al. 2017) for detailed discussion.

Cointegration tests are categorized into three; better performers, average/mediocre performers and worst performers on basis of their stringencies. Tests having stringency around or less than 30 are considered as better performers, tests having stringency greater than 30 and less or around than 50 are considered as average/mediocre performers and tests having stringency greater than 50 are considered as worst performers.

Results and Discussion

In this article four sample sizes are considered i.e. $T = 30, 60, 120, 240$ and a Monte Carlo sample size of 10000 is taken. To control for size of test, simulated critical values are used as when asymptotic critical values were used the size of tests was not around nominal size of 5%.

It is evident from Table 1 that for case I^0T^0 of deterministic part, stringency of all ten tests tends to decrease with increase in sample size with varying rates of decrease.

Table 1: Stringencies of Tests with Null of No Cointegration for I^0T^0

Tests	T=30	T= 60	T= 120	T= 240
POPU	30.26**	26.56**	22.27**	0.33**
CRDW	37.98*	27.71**	22.88**	1.16**
HZA	38.87*	25.22**	15.34**	3.58**
HCO	39.18*	25.54**	14.58**	3.51**
HADF	70.3	38.04*	19.33**	7.77**
POZA	41.84*	42.34*	41*	3.4**
POZT	42.77*	39.53*	39.1*	3.02**
CDHS	45.46*	43.28*	41.22*	3.34**
POPZ	46.33*	45.82*	46.05*	7.03**
EGADF	59.96	49.71*	45.18*	3.45**

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

Table **1** clearly depicts that from ten residual based tests only a single test i.e. POPU is better performer at all four sample sizes. Three residual based tests i.e. CRDW, HZA and HCO are average performers at smallest sample size of 30 however, these three are better performers at sample sizes of 60, 120 and 240. Five residual based tests i.e. POZA, POZT, CDHS, POPZ and EGADF are average performers up-to sample size of 120 however, these five are better performers at sample size of 240. The single remaining residual based test i.e. HADF is a worst performer at sample size of 30, an average performer at sample size of 60 and a better performer at sample size of 120 and 240.

The stringencies of tests with null of no cointegration for second case of deterministic part i.e. I^1T^0 are displayed in

Table 2: Stringencies of Tests with Null of No Cointegration for I^1T^0

Tests	T=30	T= 60	T= 120	T= 240
POPU	41.18*	36.42*	33.81*	2.83**
CDHS	47.51*	51.19*	49.85*	8.62**
HCO	55.87	37.51*	22.91**	4.81**
HZA	56.4	40.53*	26.38**	5.1**
HADF	80.99	50.31*	28.98**	15.76**
POZA	52.03	52.63	53.56	11.13**
CRDW	56.44	57.18	54.43	12.68**
POZT	58.11	61.49	58.32	16.95**
EGADF	67.68	58.2	59.62	14.51**
POPZ	68.34	58.56	51.25	13.57**

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

Table 2. From ten residual based tests only two tests i.e. POPU and CDHS are average performers at the smallest sample size of 30 and these tests continue to be average performers up-to sample size of 120. However, these two tests are better performers at sample size of 240. Three residual based tests i.e. HCO, HZA and HADF are worst performers at sample size of 30, average performers at sample size of 60 and better performers at sample sizes of 120 and 240. Rest of five residual based tests i.e. POZA, CRDW, POZT, EGADF and POPZ are worst performers up-to sample size of 120 however, these five are better performers at sample size of 240.

Table 3: Stringencies of Tests with Null of No Cointegration for I^1T^1

Tests	T=30	T= 60	T= 120	T= 240
CDHS	63.8	64.9	63.99	24.48**
POPU	64.51	56.22	56.73	15.79**
POZA	65.86	67.15	65.8	26.25**
CRDW	67.92	68.3	64.69	25.62**
HADF	92.7	76.09	69.08	27.4**
POZT	68.81	70.14	67.27	34.49*
EGADF	74.9	69.72	65.57	31.64*
HCO	92.39	94.45	94.07	94.25
HZA	94.21	94.98	94.9	95.14
POPZ	94.86	95.08	95.3	95.02

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

In Table 3 stringencies of tests with null of no cointegration for the third case of deterministic part i.e. I^1T^1 are displayed. It is clearly evident from *Table 3* that none of ten tests is a better or average performer up-to sample size of 120 rather; all of ten tests are worst performers up-to sample size of 120. However, at sample of 240 from ten residual based tests five i.e. CDHS, POPU, POZA, CRDW and HADF are better performers. From the remaining five residual based tests, two (POZT and EGADF) are average performers and three (HCO, HZA and POPZ) are worst performers.

Conclusions and Recommendations

In absence of nuisance parameters i.e. intercept and trend (I^0T^0) in cointegrating equation only POPU performs better at all four sample sizes. However, with increase in sample size some other tests like CRDW, HCO and HZA become also better performers. The remaining six tests are either worst or average performers at small sample sizes however; these six are also better performers at large sample size of 240. So, the use of POPU test for any sample size is recommended. However, CRDW, HCO and HZA are recommended for moderate small sample sizes of 60 and above. The remaining six tests i.e. HADF, POZA, POZT, CDHS, POPZ and EGADF are recommended for sample size of 240 or above only.

An increase in stringencies of all ten tests is observed when presence of one nuisance parameter i.e. intercept (I^1T^0) in cointegrating equation is assumed as compared to first case of I^0T^0. This increase in stringencies is due to decrease of powers of all tests. Due to this increase in stringencies, now POPU and CDHS are average performers up to sample size of 120 and these two are better performers at sample size of 240. Despite this increase in stringencies again POPU is the leading performer along with CDHS. Three tests i.e. HCO, HZA and HADF are worst performers at smallest sample size of 30, average performers at sample size of 60 and better performers at sample size of 120 and 240. The remaining five tests are only better performers at largest sample size of 240. So, in presence of intercept in cointegrating equation POPU and CDHS are recommended for small sizes as well as for large sample sizes. Three tests i.e. HCO, HZA and HADF are recommended for sample sizes greater than 60. The remaining five tests i.e. POZA, POPZ, POZT, CRDW and EGADF are recommended for only very large sample sizes of 240 or greater.

Again, an increase in stringencies of all ten tests is observed when presence of two nuisance parameters i.e. intercept and trend (I^1T^1) in cointegration is observed as compared to earlier cases of I^1T^0 and I^0T^0. This increase is the result of an overall decrease in powers of all ten tests. Now all of ten tests are worst performers up-to sample size of 120. Five tests i.e. CDHS, POPU, POZA, CRDW and HADF are better performers at the largest sample size of 240. So only these five tests are recommended for large sample sizes.

In general, two tests i.e. POPU and CDHS are recommended for all three cases of deterministic part. However, for the last case of deterministic part i.e. I^1T^1 these two tests are recommended for only very large sample sizes of 240 or greater.

References

Banerjee, A., J. J. Dolado, D. F. Hendry and G. W. Smith (1986). Exploring Equilibrium Relationships in Econometrics through Static Models: Some Monte Carlo Evidence*. Oxford Bulletin of Economics and Statistics 48(3): 253-277.

Bewley, R. and M. Yang (1998). On the size and power of system tests for cointegration. Review of economics and statistics 80(4): 675-679.

Boswijk, H. P. (1989). Estimation and testing for cointegration with trended variables: A comparison of a static and a dynamic regression procedure.

Boswijk, P. and P. H. Franses (1992). Dynamic Specification and Cointegration*. Oxford Bulletin of Economics and Statistics 54(3): 369-381.

Choi, I. (1994). Durbin-Hausman tests for cointegration. Journal of Economic Dynamics and Control 18(2): 467-480.

Cochrane, D. and G. H. Orcutt (1949). Application of least squares regression to relationships containing auto-correlated error terms. Journal of the American Statistical Association 44(245): 32-61.

Engle, R. F. and C. W. Granger (1987). Co-integration and error correction: representation, estimation, and testing. Econometrica: journal of the Econometric Society: 251-276.

Gabriel, V. J. (2003). Tests for the null hypothesis of cointegration: a Monte Carlo comparison. Econometric reviews 22(4): 411-435.

Hansen, B. (1990). A powerful, simple test for cointegration using Cochrane-Orcutt, Rochester Center for Economic Research, University of Rochester.

Haug, A. A. (1996). Tests for cointegration A Monte Carlo comparison. Journal of Econometrics 71(1): 89-115.

Jansson, M. (2005). Point optimal tests of the null hypothesis of cointegration. Journal of Econometrics 124(1): 187-201.

Johansen, S. and K. Juselius (1990). Maximum likelihood estimation and inference on cointegration—with applications to the demand for money. Oxford Bulletin of Economics and statistics 52(2): 169-210.

Kremers, J. J., N. R. Ericsson and J. J. Dolado (1992). The power of cointegration tests. Oxford bulletin of economics and statistics 54(3): 325-348.

Mariel, P. (1996). A comparison of cointegration tests. Applications of Mathematics 41(6): 411-431.

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.

Pesavento, E. (2004). Analytical evaluation of the power of tests for the absence of cointegration. Journal of Econometrics 122(2): 349-384.

Phillips, P. C. and S. Ouliaris (1990). Asymptotic properties of residual based tests for cointegration. Econometrica: Journal of the Econometric Society: 165-193.

Sargan, J. D. and A. Bhargava (1983). Maximum likelihood estimation of regression models with first order moving average errors when the root lies on the unit circle. Econometrica: Journal of the Econometric Society: 799-820.

Zaman, A. (1996). Statistical foundations for econometric techniques, Academic Press.

Zaman, A., A. Zaman and A.-u. Rehman (2017). The Concept of Stringency for Test Comparison: The Case of a Cauchy Location Parameter. International Econometric Review (IER) 9(1): 1-18.