MONTE CARLO COMPARISON OF PANEL DATA COINTEGRATION TESTS AND THEIR ECONOMIC APPLICATION

PhD Econometrics

Researcher:	**Mehmood Hussan**
Registration No.	**18-SE/PhD (ET)/S12**
Supervisor:	**Dr. Asad Zaman**
Co-supervisor:	**Dr. Abdul Jabbar**

School of Economics

INTERNATIONAL INSTITUTE OF ISLAMIC ECONOMICS

INTERNATIONAL ISLAMIC UNIVERSITY ISLAMABAD

Panel cointegration tests

Economics

MONTE CARLO COMPARISON OF PANEL DATA COINTEGRATION TESTS AND THEIR ECONOMIC APPLICATION

Mehmood Hussan

Registration No. 18-SE/PhD (ET)/S12

Submitted in partial fulfillment of the requirements for the

Doctor of Philosophy in Econometrics

at International Institute of Islamic Economics (IIIE)

International Islamic University Islamabad (IIUI)

Supervisor: Dr. Asad Zaman

Co-supervisor: Dr. Abdul Jabbar

July 2019

بسم الله الرحمن الرحيم

DECLARATION

I hereby declare that this thesis, neither as a whole nor as a part thereof, has been copied out from any source. It is further declared that I have carried out this research by myself and have completed this thesis based on my personal efforts under the guidance and help of my supervisor. If any part of this thesis is proven to be copied out or earlier submitted, I shall stand by the consequences. No portion of work presented in this thesis has been submitted in support of any application for any other degree or qualification in International Islamic University or any other university or institute of learning.

Mehmood Hussan

DEDICATION

To my Parents (Abba g, Amma g)

&

Wife (Javeria Kausar)

APPROVAL SHEET

Monte Carlo Comparisons of Panel Cointegration Tests and their Economic Applications

by

Mehmood Hussan

Reg. No: 18-SE/PhD(Et)/S12

Accepted by the International Institute of Islamic Economics (IIIE), International Islamic University Islamabad (IIUI), as partial fulfillment of the requirements for the award of degree of DOCTOR OF PHILOSOPHY IN ECONOMETRICS.

Supervisor: Dr. Asad Zaman

Co-Supervisor: Dr. Abdul Jabbar
Assistant Professor, IIIE, IIUI

Internal Examiner: Dr. Arshad Ali Bhatti
Assistant Professor, IIIE, IIUI

External Examiner: Dr. Atiq ur Rehman
Assistant Professor, PIDE Islamabad

External Examiner: Dr. Mumtaz Ahmed
Assistant Professor, Department of Economics, COMSATS Islamabad

Dr. Hamid Hassan
Chairman School of Economics

Dr. Atiq-uz-Zafar
Director General

International Institute of Islamic Economics (IIIE)
International Islamic University Islamabad (IIUI)

Date of Viva-Voce Examination: **July 22, 2019**

ACKNOWLEDGEMENTS

All praises to Allah, Almighty; Peace and Blessing of Allah on his last Prophet Muhammad (PBUH) who taught the humanity to human beings.

All achievements by me are nothing but a blessing of Allah. I am most obliged to my parents, sisters (Sakina, Naseem), brothers (Lala Ramzan, Zafar Iqbal, Muhammad Saleem) and my kids (Muhammad, Ahmad Hassan and Saad Hassan). Particularly, I am very grateful to my kind and gentle mother (AMMA). She always encourages me to complete all the assigned task gently and I could not forgive her special prayers throughout my success. Very special thanks to my beloved and always caring wife; Javeria Kausar. She always joint her hands with me for completing this journey smoothly.

I extended my kinds words to Dr Asad Zaman who always helps me morally, spiritually, technically and academically throughout my PhD journey. I could not also forget the support of Dr Abdul Jabbar during Doctoral research.

At this occasion, I am thankful to all my teachers who taught me at any stage of my studies. I would acknowledge the teaching of Dr. Faiz-ur-Raheem for their kind style of teaching. I must acknowledge cooperation and support of my departmental faculty including Dr. Atiq-uz-Zafar, Dr. Hamid Hassan, Dr. Hafiz Muhammad Yasin, Dr. Mumtaz Ahmad, Dr. Attiq-ur-Rahman, Dr. Ashad Ali Bhatti and Dr Abdul Rasheed.

Prominently, I want to extend my gratitude thankfulness to Dr Asad-UL-Islam Khan for his intellectual, kind and naughty support of friendship during whole period of stay in Islambad, Pakistan. I would like to acknowledge my best friends and class fellows Yasir Riaz Shahid Akbar, Dr Waqar Khan, Ali Raza, Dr Imadullah, Abdul Waheed, Abdul Ghafar khan Bhahadar Abdul Rehman, Gulfam Haider, and Dr Imran Hussain Kosa whose company I always enjoyed.

I also acknowledge the cooperation of the staff of IIIE including Mr. Syed Niaz Ali Shah, Mr. Zareef and Mr. Tauqeer Ahmed. I would like to thanks my working colleagues and friends; Aamir Shabir Javed Ali Bhidwal, Asif Nawaz, Dr Abdul Waheed, Fida Hussain, Nazar Abbas Asif Kamal, Malik Ameer, Ameer Mukhtar, Abdul Majid Shah, Zubair Khan and Shazia Jabeen. I am not able to find words to say thanks to my elder brother Lala Ramzan, Zafar iqbal and Muhammad Saleem. and my cousin Muhammad Haroon continuous cooperation, support, encouragement and lot more. Last but not the least I am thankful to my parents and wife for their prayers, sacrifices, physical contribution, patience and immortal love that encouraged me at every flash of study period.

Mehmood Hussan

Table of Contents

List of Figures ix

List of Tables xii

List of Acronyms xiv

Abstract xvi

Chapter 01: Introduction 1

1.1. Background of Study 1

1.2. Motivation 4

1.3. Significance of the Study 5

1.4. Objectives of Study 6

1.5. Scheme of Dissertation 8

Chapter 02: Literature Review 9

2.1. Regressive Review of Findings 9

2.2. Tests having Null Hypothesis of No Cointegration 18

2.2.1. Residual-Based Tests 18

2.2.1.1. Pedroni Tests 18

2.2.1.2. Kao Tests (1999) 24

2.2.1.3. Average Phillips Zt Statistics for Panel Cointegration 29

2.2.2. Error Correction Based Tests 31

2.2.2.1. Westerlund Cointegration Tests (2007) 31

2.2.3. Average Weighted Symmetric Test (AWS) 35

2.2.4. Fisher Type Tests 37

2.2.4.1. Fisher Augmented Dickey-Fuller Test (FADF) 37

2.2.4.2. Fisher Weighted Symmetric Test (FWS) 39

2.2.5. Maximum Likelihood Based Test 40

2.3. Panel cointegration Tests for Null Hypothesis of Cointegration 42

List of Figures

Figure no	Title	Page no
4.1	EMPIRICAL SIZE OF TESTS HAVING NULL OF NO COINTEGRATION, N=02	58
4.2	EMPIRICAL SIZE OF TESTS HAVING NULL OF NO COINTEGRATION, N=08	60
4.3	EMPIRICAL SIZE OF TESTS HAVING NULL OF NO COINTEGRATION, N=16	63
4.4	EMPIRICAL SIZE OF TESTS HAVING NULL OF NO COINTEGRATION, N=32	65
4.5	EMPIRICAL SIZE OF TESTS HAVING NULL OF COINTEGRATION, N=02	67
4.6	EMPIRICAL SIZE OF TESTS HAVING NULL OF COINTEGRATION, N=08	69
4.7	EMPIRICAL SIZE OF TESTS HAVING NULL OF COINTEGRATION, N=16	70
4.8	EMPIRICAL SIZE OF TESTS HAVING NULL OF COINTEGRATION, N=32	71
5.1	SIMULATED SIZE OF TESTS HAVING NULL OF NO COINTEGRATION, N=02	75
5.2	SIMULATED SIZE OF TESTS HAVING NULL OF NO COINTEGRATION, N=08	76

5.3	SIMULATED SIZE OF TESTS HAVING NULL OF NO COINTEGRATION, N=16	77
5.4	SIMULATED SIZE OF TESTS HAVING NULL OF NO COINTEGRATION, N=32	78
5.5	SIMULATED SIZE OF TESTS HAVING NULL OF COINTEGRATION, N=02	79
5.6	SIMULATED SIZE OF TESTS HAVING NULL OF COINTEGRATION, N=08	80
5.7	SIMULATED SIZE OF TESTS HAVING NULL OF COINTEGRATION, N=16	80
5.8	SIMULATED SIZE OF TESTS HAVING NULL OF COINTEGRATION, N=32	81
6.1	BEST PERFORMING TESTS HAVING NULL OF NO COINTEGRATION	86
6.2	MEDIOCRE PERFORMING TESTS HAVING NULL OF NO COINTEGRATION	89
6.3	WORST PERFORMING TESTS HAVING NULL OF NO COINTEGRATION	92
6.4	MAXIMUM SHORTCOMINGS OF TESTS HAVING NULL OF COINTEGRATION, N=02	99
6.5	MAXIMUM SHORTCOMINGS OF TESTS HAVING NULL OF COINTEGRATION, N=08	101
6.6	MAXIMUM SHORTCOMINGS OF TESTS	103

	HAVING NULL OF COINTEGRATION, N=16	
6.7	MAXIMUM SHORTCOMINGS OF TESTS HAVING NULL OF COINTEGRATION, N=32	105

List of Tables

Table number	Title	Page number
6.1	Best, Mediocre and Worst Performing Tests According to their MS	85
6.2	Best Performing Tests having Null of No Cointegration	88
6.3	Mediocre Performing Tests having Null of No Cointegration	91
6.4	Worst Performing Tests having Null of No Cointegration	95
6.5	Power of Tests having Null of Cointegration, N=02	100
6.6	Power of Tests having Null of Cointegration, N=08	102
6.7	Power of Tests having Null of Cointegration, N=16	104
6.8	Power of Tests having Null of Cointegration, N=32	106
7.1	Bootstrap Empirical Power for N=08	115
7.2	Bootstrap Empirical Power for N=16	116
	Appendix Tables	
A-1	Empirical Size of Tests having Null Hypothesis of No Cointegration using Asymptotic Critical value N=02	136
A-2	Empirical Size of Tests having Null Hypothesis of No Cointegration using Asymptotic Critical values N=08	137
A-3	Empirical Size of Tests having Null Hypothesis of No Cointegration using Asymptotic Critical values N=16	138
A-4	Empirical Size of Tests having Null Hypothesis of No Cointegration using Asymptotic Critical values N=32	139

A-5	Empirical Size of Tests having Null Hypothesis of Cointegration using Asymptotic Critical values N= 02, 08, 16, 32	140
A-6	Maximum Shortcomings of Tests having Null Hypothesis of No Cointegration N=02	141
A-7	Maximum Shortcomings of Tests having Null Hypothesis of No Cointegration N=08	142
A-8	Maximum Shortcomings of Tests having Null Hypothesis of No Cointegration N=16	143
A-9	Maximum Shortcomings of Tests having Null Hypothesis of No Cointegration N=32	144
A-10	Maximum Shortcomings of Tests having Null Hypothesis of Cointegration N=02	145
A-11	Maximum Shortcomings of Tests having Null Hypothesis of Cointegration N=08	145
A-12	Maximum Shortcomings of Tests having Null Hypothesis of Cointegration N=16	145
A-13	Maximum Shortcomings of Tests having Null Hypothesis of Cointegration N=32	146

List of Acronyms

PDFT : Kao Dickey Fuller t-statistic, that is, DF_t Cointegration test

PDFrho : Kao Dickey Fuller rho-statistic, that is, DF_ρ Cointegration test

PDFTstar : Kao Dickey Fuller t*-statistic, that is, DF_{t*} Cointegration test

PDFrhostar : Kao Dickey Fuller rho*-statistic, that is, $DF_{\rho*}$ Cointegration test

PADF: Kao Augmented Dickey Fuller statistic Cointegration test

PdGtnp : Pedroni Group-t non/semi parametric, that is, $Z_{tN,T}$ Cointegration test

PdGtp : Pedroni Group-t parametric, that is, $Z^*_{tN,T}$ Cointegration test

PdGrho: Pedroni Group_rho $Z_{\rho N,T^{-1}}$ Cointegration test

PdPtp: Pedroni Panel _t parametric, that is, $Z^*_{tN,T}$ Cointegration test

PdPtnp : Pedroni Panel _t non/semi parametric, that is, $Z_{tN,T}$ Cointegration test

PdP_V : Pedroni Panel-V, that is, $Z_{\hat{v}N,T}$ Cointegration test

PdPrho : Pedroni Panel _rho, that is, $Z_{\rho N,T^{-1}}$ Cointegration test

PhZt : Average Phillips Zt Statistics for panel cointegration

PAWS : Average Weighted Symmetric based Cointegration test

Pfadf : Fisher Augmented Dickey Fuller Cointegration test

Pfaws : Fisher Weighted Symmetric based Cointegration test

LR : Likelihood Ratio cointegration tests

W_Gt : Wester Lund Group-t, that is, G_T cointegration test

W_Ga : Wester Lund Group-α, that is, G_α cointegration test

W_Pt : Wester Lund Panel-t, that is, P_T Cointegration test

W_Pa : Wester Lund Panel-α , that is, P_α Cointegration test

OLS : Ordinary Least Square

LM_OLS : Lagrange Multiplier Ordinary Least Square Cointegration test

LM_DOLS : Lagrange Multiplier dynamic Ordinary Least Square Cointegration test

LM_FMOLS : Lagrange Multiplier Fully Modified Ordinary Least Square Cointegration test

PO : Point Optimal Cointegration test

LLR : Log Likelihood ratio test

DGP : Data Generating Process

MSC : Maximum Shortcomings

MCSS : Monte-Carlo Sample Size

ECM : Error Correction Model

DF : Dickey Fuller

ADF : Augmented Dickey Fuller

MA : Moving Average

Asymptotic Critical Values: ACV

Simulated Critical Values: SCV

Abstract

Mainly, the current study demonstrates the comparison of panel cointegration tests. This study comprised of comparative assessments of 24 tests and their comparison using stringency criterion. Among the compared tests, 21 tests are based on the null hypothesis of no cointegration, while 3 tests are based on the null hypothesis of cointegration. As far as novelty of the study is concerned, none of the existing studies have used this stringency criterion for the power comparison of panel cointegration tests. within the framework of null of no cointegration, this study evaluates the performance of 21 tests included; Residual based tests (parametric/nonparametric), Maximum Likelihood based tests, Fisher type tests, Average Weighted Symmetric test and Error Correction Based tests.

The current study also extended toward comparison of power 3 panel cointegration tests having null hypothesis of cointegration. This study also evaluates the empirical size of the under consideration panel cointegration tests using asymptotic critical values and found that most of the tests are oversized. Therefore, size of all these tests have been controlled by using simulated critical values. The results of current study depicts that two residual based tests (PdP_V, PdPrho) and average weighted symmetric based test (PAWS) performed in paramount way throughout all time and cross sectional dimensions. current study also reveals that maximum likelihood based test LR and second generation test W_Gt perform worst in the current scenario. The LM_OLS test having null hypothesis of cointegration performed better than LM_DOLS and LM_FMOLS. Three better performing tests have reasonably high bootstrap powers based on Fisher hypothesis.

Chapter 01: Introduction

1.1. Background of Study

The use of cointegration techniques in economic empirical literature has been widely recognized during the last three decades. In 1987 Economist (Engle and Granger) introduced the core concept of cointegration firstly. The phenomena of cointegration elaborates the extended relationship between integrated variables. The concept of cointegration was introduced first time by Engle and Granger (1987). The cointegration describes the long run relationship between the integrated variables that is, if the variables are integrated of order *d* then their linear combination must have the order smaller than *d*. For example, variables are said to be cointegrated if components of p-dimensional process Z_t are integrated of order one that is, $Z_t \sim I(1)$ and the linear combination of the components of Z_t are integrated of order zero (stationary) that is, $\gamma' Z_t \sim I(0)$ for $\gamma \neq 0$.

Cointegration was also described by Lütkepohl (2005) in another way. "if at least one of the components of a *p*-dimensional process Z_t is integrated of order *d* that is, $Z_t \sim I(d)$, and their linear combination $\gamma' Z_t$ is integrated of order less than *d*, then the components of the process Z_t are cointegrated" (for detail reference is provided; Lütkepohl (2005). From the provided reference it can be concluded that all components of Z_t may not necessarily have the same order of integration. It is known that the regression of an uncorrelated integrated process. Factually, variance of regression cannot be estimated consistently. However, spurious regression has non-stationary residuals. Moreover, residuals of regression are present in stationary form, it may results cointegration of integrated process. Theoretically for explaining economic models in best ways, researchers relies more on cointegration technique as compared to conventional methodologies.

Within comparative assessment of panel data models and time series analysis cointegration showed complex behavior in former data models. Complexity in panel data models occurs due to multiple factors such as sectional dependence, non-homogeneity, unbalanced data and many others. Mainly, two prominent categories of cointegration tests exist primarily based on residuals and maximum likelihood technique. Former category tests consist of single equation model which detect extended relationship while later category consists system of equations, detecting not only the presence of extended relationship, but also determines the numeral relationships of cointegration among the variables of the system.

Panel cointegration tests are the extension of panel unit root tests. Pedroni (1999), Kao (1999) and Pedroni (2004) have introduced the first panel cointegration residuals based tests having null hypothesis of no cointegration of Dickey and Fuller (1979) (DF) and Said and Dickey (1984) (ADF) statistics which are the extension of panel unit root of Levin and Lin (1993). Furthermore, Pedroni (1999) also introduced the panel cointegration tests having null hypothesis of no cointegration. These tests are the extensions of the variance ratio tests proposed by Phillips and Ouliaris (1990) and Phillips and Perron (1988).

Firstly, deliberate discovery of panel cointegration tests (mainly residual based) was done by McCoskey and Kao (1998) using the principle; null of cointegration. Basically mentioned test (McCoskey and Kao) was an extension of test proposed by Harris and Inder (1994) and Shin (1994). According to Harris and Inder (1994) and Shin (1994), the tests having null hypothesis of cointegration can be very appealing in application, where the cointegration is predicted prior by the economic theory.

Keeping in mind heterogeneous panel, Larsson, Lyhagen et al. (2001) introduced cointegration tests based on maximum likelihood technique. Another test proposed by Groen and Kleibergen (2003) for cointegration also have maximum likelihood approach.

Breitung (2005) introduced a new procedure of estimation of panel cointegration tests. Breitung (2005) suggested Wald, Lagrange multiplier and LR tests, based on the procedure of Saikkonen (1999). Westerlund (2007) also developed four panel cointegration tests that may explain the generalized posture of tests given by Banerjee, Dolado et al. (1998). All these four tests are based on the Error Correction Model (ECM). Out of these four, two tests are group mean tests, which assume that error correction terms are varied in cross sections. While, the other two panel tests assume that error correction terms are homogenous in cross sections.

The empirical size and power of any test represents the performance of that test. The performance of any test is being considered the best, if empirical size of the test is close to the nominal size with high power. In the literature, different Monte Carlo studies have been conducted to compare the size and power of the panel cointegration tests. Usually, whenever someone proposes a new test for panel cointegration, also compares the size and power of this new test with the existing ones in literature (Monte Carlo simulations). In most of the Monte Carlo comparisons of size and power, two or three points of alternative hypothesis have been considered. Similarly, in these comparisons most of the researchers have used asymptotic critical values., The problem with the asymptotic critical values is that the size of tests is not controlled and does not converge to the nominal size.

In this study two classes of panel cointegration tests have been compared that are , the tests having null hypothesis; no cointegration and cointegration by using stringency criterion discussed by Asad Zaman (1996). For this scenario two point optimal tests with opposite null hypothesis have been used

1.2. Motivation

Cointegration technique is most popular, significant and important technique in the economic literature for testing the long run relationship among the economic variables as well as financial variables. The popularity of cointegration technique has been growing in the empirical literature because it is used to test the presence of long run relationship among the integrated variables. In recent years, there has been an increase in the importance and use of panel data sets, where economic variables have been observed over extended periods of time across a large number of cross-sectional units, such as countries, industries and households. This development has in turn given rise to a great amount of interest in econometric techniques for dealing with potentially non stationary panel data variables. There has been a rapid growth in the use of cointegration methods with large panel data sets, to empirically evaluate important economic theories.

Multiple studies have been conducted in literature to evaluate the performance of tests using simulation to prime factors; power and size. Comparative studies from existing literature depicts that tests showed fluctuating behavior regarding their performance. To tackle this, an integrated framework may be developed for their comparison on whole alternative space

In all these studies, different Monte Carlo simulation designs are used. In the economic literature, there has not been such study where both types of panel cointegration tests have been compared using stringency criterion. The above gape in literature inspired us to carry out a comprehensive evaluation of cointegration (24 tests). The mentioned tests have been assessed with robust technique; stringency criterion.

In this study, two different classes of tests have been compared with opposite null hypothesis of (no) cointegration. simulated critical values have been used for both classes of tests. This study has compared the Panel cointegration tests which are residual based, maximum likelihood based, average weighted symmetric based, Fisher type tests and those tests which are based on the vector error correction model. So, this study has filled this gap in the literature and provides the guidelines to practitioners about the best panel cointegration tests.

1.3. Significance of the Study

As researchers and practitioners are aware of the importance of the concept of cointegration in the empirical literature of panel data, therefore they have frequently used this concept to check the long run relationship among the integrated economic as well as financial variables. But the core problem faced by the practitioners and researchers in this regard is that there is no clear cut guideline in the existing literature that which test can be applied according to the length of time dimension and cross sectional dimension. Usually in the literature practitioners and researchers have applied the cointegration tests without knowing the flaws of the cointegration tests, that is, size distortion and low power.

Practitioners also have made the decision about the rejection of null hypothesis by using the asymptotic critical values, which usually leads towards misleading results.

There exist a few studies in which the size and power comparisons have been done. These include but are not limited to, McCoskey and Kao (1998), Kao (1999), Gutierrez (2003), Pedroni (2004) and Örsal (2007). But these studies have limitations, only few tests have been considered in each study. In most of the studies size of the tests has not been controlled before comparing their power leading to meaningless results. Another thing which is common in these exiting comparisons in the literature is that no one has used whole alternative space, tests have been compared on few points of alternative hypothesis.

This study has tried to overcome all these problems and has compared most of the tests in the existing literature by using Monte Carlo simulation experiment. In spite of Asymptotic critical values (ACV), simulated critical values (SCV) have been used in this Monte Carlo experiment. Critically, ACV have not been used due to unstable size, therefore SCV might be used.

In present study when the empirical size of tests has been checked. Study also evaluate SCV have been used to control/ stabilize the size of tests and then their power comparison is carried out using stringency criterion. This study has also investigated the bootstrap empirical powers of best cointegration tests based on Fisher hypothesis by using the OECD countries quarterly data.

1.4. Objectives of Study

- The first most important aim of this current study is to evaluate the pragmatic size of panel cointegration test keeping in view ACV To complete this two different classes will be compared based on opposite null hypothesis of (no) cointegration.

- As researchers/practitioners know that the power of two tests can only be compared if their size is stable and close to the assumed nominal size. So, in this regard, to evaluate whether the under consideration tests are stable sized, oversized or under sized, empirical size of tests has been estimated. Within this objective, the impact of time and cross sectional dimensions are also assessed on empirical size of test.

  The Second prime and integral objective of the study is to stabilize or control size of panel cointegration tests by means of SCV. As discussed earlier, this study should compare the tests of stable size using stringency criteria. Within this objective, the impact of time and cross sectional dimensions are also assessed on stabilizing size of test.

- The third main and core aim is to investigate power of both classes of panel cointegration tests and assess the pattern of power of each test at all alternative space. Within the sphere of third objective, the impact of time and cross sectional dimensions are also assessed on power of test.
- The fourth main and vital objective of this study is to investigate the most stringent test of all under consideration tests in each type of tests.
- The fifth and last objective is to investigate the bootstrap empirical powers of best panel cointegration tests based on Fisher Hypothesis. For this, OECD countries quarterly data have been used.

1.5. Scheme of Dissertation

The structure of current dissertation consists of workpackeges; Chapter 1 consists of motivation, significance of the study and objectives of the study. Chapter 2 dissects the contents into two parts; regressive review on panel cointegration tests and their comparison with existing methodologies. Chapter 3 provides the details of methodology regarding Monte Carlo comparison using stringency criteria. Chapter 4 gives an overview of empirical size of the cointegration tests. In Chapter 5, size of the tests has been controlled because most of the tests have not stable size. Chapter 6 is the discussion on the power of the tests by using stringency criteria. Chapter 7 discusses the bootstrap empirical powers of the best tests using real data. The final and last part (chapter 8) constitutes the core finding and future directives..

Chapter 02: Literature Review

2.1. Regressive Review of Findings

Mainly, two types of the panel cointegration tests exist namely residual based cointegration and maximum likelihood tests. Researchers have developed residual based cointegration test for two classes of tests; tests having the opposite null hypothesis of (no) cointegration. Core theme of current tests is to evaluate the unit root presence in residuals of cointegration regression. The existence of (no) cointegration in the integrated variables can be checked via presence of unit root in residuals and stationary residuals respectively. Principally, if only one cointegration relationship exist among the variables then residual based tests are used if vice versa (more than one) relationship exist then mentioned tests cannot be used.

The second most important category of tests to hand panel data is maximum likelihood based tests which is extended form of Johansen (1988). Advantages of mentioned tests are; former type determine more than one cointegration relationship among the variables if more than one relationship exists. The phenomena of autonomous behavior (free from choice of variable) is prominent outcomes of maximum likelihood based tests.

Multiple comparative studies exist for panel cointegration tests (Kao, 1999; McCoskey and Kao,1998; Wu and Yin, 1999; Larsson et al, 2001; Pedroni, 2004). Whenever, researchers have developed the new panel cointegration tests, mostly, these developed tests have been compared with some existing panel cointegration tests. Kao (1999) has developed five tests of panel cointegration having null of no cointegration, in which four are Dickey-Fuller (DF) type and one is augmented Dickey-Fuller type (ADF) that is, $DF_t, DF_\rho, DF_t^*, DF_\rho^*, ADF$. Two Dickey-Fuller (DF) tests; DF_t and DF_ρ have assumed strong exogeneity where two tests DF_t^* and DF_ρ^* have relaxed this assumption. Kao (1999) has also developed the asymptotic distribution of these tests which converges to standard normal distribution when N and T converge to infinity.

McCoskey and Kao (1998) proposed the residual based Lagrange Multiplier (LM) tests for the null hypothesis of cointegration. McCoskey and Kao, 1998 also performed the Monte Carlo comparison for size and power properties of these tests. McCoskey and Kao concluded that the empirical size of LM-FMOLS and LM-DOLS are closed to the nominal size of 5% (even in small samples). According to McCoskey and Kao (1998) power of these two tests is quite good when time dimension is greater than 50. When cross sectional and time dimensional length is fixed at 50, the presence of endogeneity and moving average term, these two tests performed differently. In general LM-DOLS test performed better, nevertheless, LM-FMOLS test is more powerful in some cases.

Kao (1999) has developed residual based five panel cointegration tests

Authors also conducted simulation experiments to compare size and power of these five tests using asymptotic critical values. In these five tests, comparative study concluded that empirical size of tests; DF_t^* and DF_ρ^*, is close to the nominal size (of 5%) when both time "T" and cross sectional "N" dimensions are large. Whereas, the empirical size of ADF test is greater than 9% for all values of N and T, similarly, the size of DF_t and DF_ρ is greater than 7% for all values of N and T. Kao (1999) concluded that when time dimension "T" is small (around 25) then size of all tests distorts even though the N is large enough. Kao (1999) also described the unadjusted powers of these five tests for power comparison using asymptotic critical values and considered only one point for alternative hypothesis.

According to Kao (1999), all these five tests have small powers when T and N are small, however, DF_ρ^* has small power even though the N is large. Once, the time dimension is increased DF_ρ^* dominates the DF_t^*. Regressors are endogenous in the data generating process (DGP) while the DF_t and DF_ρ perform well even though tests are misspecified. Hence, experiments suggest that the distribution of DF_t^*, DF_ρ^* and ADF can be far different from standard normal N(0,1) when the DGP contains the moving average (MA) components. It is also evident that the DF_t and DF_ρ tests are substantially robust (even though the model is misspecified).

McCoskey et al., 1999 have compared power and size characteristics of five tests; $(ADF^*, PO_t^*, PO_\alpha^*, APG^*$ and $LM^*)$,where first four tests have the null hypothesis of no cointegration while last one has the null hypothesis of cointegration. In this comparison firstly, authors compared 04 tests having null of no cointegration subsequently, select two best tests regarding size and power. Laterally, these 02 best tests were compared with the test having null hypothesis of cointegration. The empirical size of these four tests having null hypothesis of no cointegration, ADF^* performed best as compared to other, In the power comparison of these four test, the two tests, ADF^* and APG^* performed best. Lastly, these two tests have been compared with LM^*. Also, these three tests ADF^*, APG^* and LM^* have been compared with each other by using two different DGPs where LM^* performed best as compared to other two.

Wu and Yin (1999) have compared empirical size and power of ADF and maximum Eigen value based tests in pooling information on means and p-values respectively. Wu and Yin (1999) concluded that the average ADF performs better with respect to power and their maximum eigen value based p-value perform better with regards to size.

Firstly, for panel framework, maximum likelihood based test proposed by Larsson et al. (2001), subsequently, also compared size and power of standard trace test with the standardized LR-bar test. Comparison revealed that standard trace test has better size than standardized LR-bar test.

Pedroni (2004) studied the sample properties (small) of the five tests which were initially developed by the same research group (Pedroni 1999). Author conducted that Monte Carlo experiments to analyze the empirical size and power of the test $\left(Z_{vNT}, Z_{\rho NT^{-1}}, Z_{\hat{t}NT}, Z_{\rho NT^{-1}} \text{ and } Z_{\hat{t}NT}\right)$. In this experiment Pedroni employed two DGPs one for empirical size and the other for power comparison of the tests. First the empirical size of these tests is observed as time dimension "T" varies and cross sectional "N" size is fixed. When cross sectional dimension is 20 and time dimension varies from 40 to onward then two tests based on t-statistics converge to nominal size (5%) from above whereas, three tests $Z_{vNT}, Z_{\rho NT^{-1}}$ and $Z_{\rho NT^{-1}}$ converge to nominal size (5%) from below.

In other words, when time dimension is small; two tests are oversized and three tests are undersized but when time dimension approaches to 150 then the size of all five tests lie in the range (4% to 7.5%). Proceeding this further, author also considered the reverse process that was fixed the time dimension (T=250) and varied the cross sectional dimension "N". Pedroni also observed that in the start the empirical size of all five tests lie in the range (4.5% to 8.5%) as cross-sectional dimension (N=150) then range becomes little narrow (3% to 6%).

Next Pedroni analyzed the empirical size of only two tests; panel-rho and group-rho $\left(Z_{\rho NT^{-1}} \text{ and } \tilde{Z}_{\rho NT^{-1}}\right)$ respectively and varied both cross-sectional dimension (N) and time dimension (T) from $N = T^{\frac{1}{2}}$ to $N = T^{\frac{5}{6}}$. Pedroni observed that the empirical size of panel-rho $\left(Z_{\rho NT^{-1}}\right)$ converges quickly to the nominal size (5%) when $N = T^{\frac{3}{4}}$ then converges slowly when $N = T^{\frac{1}{2}}$. Similarly, the empirical size of group-rho test $\left(\tilde{Z}_{\rho NT^{-1}}\right)$ converges very quickly when rate of expansion is $N = T^{\frac{5}{6}}$ and converges slowly when the rate of expansion is $N = T^{\frac{1}{2}}$.

Pedroni (2004) also examined the empirical power of the tests $Z_{vNT}, Z_{\rho NT^{-1}}, Z_{\hat{t}NT}, \tilde{Z}_{\rho NT^{-1}}$ and $\tilde{Z}_{\hat{t}NT}$ using different combination of time dimension and cross sectional size. Pedroni observed that when cross sectional size is fixed (N=20) and time dimension varied from 20 with increment of 10 and AR(1) coefficient for the regression residuals is $\phi = 0.9$ then most of the tests approach to 100% power at T=50. When the value of $\phi = 0.95$, then all the tests attain 100% power for larger value of "T" and N=20 is fixed. Panel-v test has achieved the 100% power quickest at T=90. When author considered the extreme case $\phi = 0.99$ very close to the unit root, resultantly, all these five tests gained 100% power for large value of T keeping the fixed cross section dimension (N=20. Similarly, Pedroni observed the power of these tests by reversing the process, fixing T=250 and then N varied, resultantly, the behavior of these tests remained approximately same. Very importantly, Pedroni (2004) also applied these tests to check the purchasing power parity hypothesis.

Gutierrez (2003) investigated the powers of different tests which were proposed by Kao (1999), Pedroni (1999) and Larsson, Lyhagen et al. (2001) by using Monte Carlo simulation experiment. The DGP used for this study was based similar to one proposed by Engle and Granger (1987) and the same DGP was also reported by Gonzalo (1994), Haug (1996) and Kao (1999). Gutierrez (2003) used different combination of time dimension "T" and cross sectional "N" length where the value of T and N are $T = \{10, 25, 50, 100\}$ $N = \{10, 25, 50, 100\}$. Gutierrez displayed those tests in his article which has the best power inside the group of Kao's tests, that is, DF_ρ and DF_ρ^*. Furthermore, Pedroni tests that is Panel-rho and group-rho and also inserted LR-bar test which is proposed by Larsson et al.,2001.

Gutierrez used the simulated critical value at 5% level of significance. When time dimension is 10 and cross section dimension varied then Kao's tests showed higher power as compared to other tests but power of other tests was still fairly low even the number of cross section is reached to 100. Gutierrez showed that Kao (1999) tests performed better than Pedroni (1999) tests when time dimension is small but when time dimension is stimulated the Pedroni's tests performed better than Kao's tests. For small sample size, LR-bar test has very low power because for small-T, the test size is distorted. Generally, the power of the tests increases when the time and cross section dimension is stimulated.
Hoang (2006) has proposed three new tests that are, average weighted symmetric test (AWS), fisher weighted symmetric test (FWS) and fisher augmented dickey fuller test (FADF) and also compared the size and power of these three tests with Pedroni $group-t$, $group-\rho$ and Kao ADF test by using Monte Carlo's simulation. In above mentioned six tests comparison, author considered four different groups of DGPs by taking and

relaxing the assumption of endogeneity and autocorrelation of the residuals. Overall, three new tests such as AWS, FWS and FADF performed well having good size, converged to the nominal size 5% for all four sets of experiments. The AWS and FWS have performed consistently most powerful tests in all data generating process.

Hoang (2006) concluded that two tests dominate the other four tests in all aspects. The AWS test has little bit more power than FWS but AWS test only works with balance panel data whereas the FWS can be applied to both kind of panel data balance or unbalance.

Örsal (2007) has compared the size and size-adjusted power of five panel cointegration tests by using Monte Carlo's simulation study, in which four are residual based which are Pedroni (1999) tests that is, $\text{panel}-\rho$, $\text{panel}-\text{t}$, $\text{group}-\rho$ and $\text{group}-\text{t}$ and one is maximum likelihood based LR-bar statistic of Larsson et al., 2001.

Orsal used the DGP of Toda (1995) which has been used in the literature. Author reported in this experiment that pragmatic size of $\text{Group\_}\rho$ and $\text{Panel\_}\rho$ are always zero (T=10,25, $N \geq 1$). Whereas, severe size distortion has occurred for the other test when time dimension is small and N is large and tests statistics oversized when T is small and N is increasing. Hence author concluded that the empirical size of tests is not appropriate when time dimension is much smaller than cross sectional dimension. Orsal also concluded that the empirical size of $panel-t$ and LR-bar statistics converges to the nominal size when time dimension is 200 and cross sectional size is greater than five whereas the $group-\rho$ and $panel-\rho$ converge to the nominal size when time dimension is 100 and 50 respectively and cross sectional size is greater than five.

Orsal also observed the size-adjusted power properties of these five tests. When the time dimension is 10, the cointegrating parameter $\psi = 0.5$, the LR-bar statistic and $group-t$ statistic have least power for cointegration rank of 1 and 2 whereas the power of $panel-\rho$ reached 89% and 68% for cointegrating rank 1 and 2 respectively. If the cointegrating parameter is $\psi = 0.95$ near the unity then all the tests have the power less than 8% at T=10. The LR-bar statistics have gained the highest power among these entire five tests when time dimension is 100 and cointegration rank is three.

Örsal (2009) has also applied these cointegration tests to check Fisher hypothesis. Örsal (2009) has also proposed a maximum likelihood-based new panel cointegration tests that is, Panel SL test. Örsal compared this Panel SL test with Larsson, Lyhagen et al. (2001) test that is, LLL test by using Monte Carlo's experiment. Örsal (2009) considered three different DGPs to study the size and power properties of these two tests. The simulation results have indicated that the size of these two tests distorted for lower value of T, whereas size of mentioned two tests converged to nominal size (5%) when time dimension is increased. When the phenomena of uncorrelated component of DGP exist then panel SL test has more advantageous characteristic regarding size as compared to LLL test evident from Orsal findings.Generally, the power of both the tests approached to 100% if cross sectional dimension is increased and time dimension is small.

There are also many empirical application of panel cointegration technique exist in the literature, for example testing of purchasing power parity hypothesis, Fisher hypothesis, energy consumption and economic growth and so forth.

In this study discussed, the empirical application of Fisher hypothesis, it is not only tested via cointegration technique in time series case but also tested via cointegration technique in panel case. Crowder (2003), Westerlund (2008), Örsal (2007), Toyoshima and Hamori (2011), and Badillo et al., 2011 have used panel data analysis to test the Fisher hypothesis via panel cointegration. Crowder (2003) have studied on panel data having nine industrialized countries to test the Fisher hypothesis. Westerlund (2008) have used the data from 1980–2004 for 20 OECD economies and examined the relationship between the nominal interest rate and the inflation rate. The results did not reject the existence of the Fisher effect.

Örsal (2007) have also tested the Fisher hypothesis by using two data sets of OECD countries having different time dimension and cross sectional size. and concluded that Fisher hypothesis exist. Similarly Badillo, Reverte et al. (2011) have also investigated the Fisher hypothesis for the 15 European Union countries via the cointegration test data from 1983–2009. Researchers found evidence in support of a weak Fisher effect. Toyoshima and Hamori (2011) have also performed panel data analysis by using monthly data of United States, United Kingdom, and Japan (January 1990–December 2010), and concluded strong evidence in favour of the Fisher hypothesis.

Presently, this study defines the Fisher hypothesis, it states that the real interest rate is the difference between the nominal interest rate and expected inflation.

$$r_{it} = n_{it} - \pi_{it}^{e} \qquad (2.1)$$

where r_{it}, n_{it} and π_{it}^{e} represent the real interest rate, nominal interest rate and expected inflation respectively. It is elaborated that no personal and intuitional entity lends at nominal interest rate than the expected inflation in reduced order. In other word real interest rate could also be defined; additive product of expected inflation and borrowing capacity. Fisher (1930) stated that real interest rates show little trend in the long run or real interest rates are constant.

2.2. Tests having Null Hypothesis of No Cointegration

In this study, twenty one panel cointegration tests are included having null hypothesis of no cointegration, which are given below with details. These twenty one tests are divided into five sub-categories

1. Residual-based tests (parametric/ non parametric)
2. Error correction based tests
3. Maximum likelihood based tests
4. Fisher type tests
5. Average weighted symmetric test

2.2.1. Residual-Based Tests

Residual based panel cointegration tests have been introduced by Pedroni and Kao. Furthermore, Pedroni (1999) has also introduced both parametric and non parametric type tests the detail of tests are give below;

2.2.1.1. Pedroni Tests

There are two well-known types of tests in panel cointegration frame work, one type is residual based tests and other is likelihood based tests. Residual base panel cointegration tests have been introduced first by Pedroni (1996), further Pedroni (1999) and Pedroni (2004) has extended this work by using multiple regressors in the panel cointegration tests. Pedroni (1999) has developed 07 cointegration tests having null of no cointegration, where, 04 statistics are within dimension based whereas three statistics are between dimensions based. The four within dimension based statistics are calculated by summing up the denominator and numerator separately over N cross section, these four within dimension based statistics are $Panel - v$, $Panel\_\rho$, semi parametric $Panel\_t$ and parametric $Panel\_t$.

The three between dimension statistics are calculated by dividing the numerator and denominator first and then summing up over the N cross section , these three between dimension statistics are $Group\_\rho$, $Group\_t$ parametric and $Group\_t$ non- parametric. It is to be noted that $Panel - v$ is the extended form calculated by ratio of variance introduced by famous authors; Phillips and Ouliaris (1990). Whereas authors such as Phillips and Perron (1988) and Phillips and Ouliaris (1990). proposed non parametric rho-statistics, which is further extend to $Panel\_\rho$. In a same way, t-statistic (non parametric) proposed by Phillips and Perron (1988), further extended to $Panel\_t$. $Panel\_t$ (parametric) is the extended form of ADF t-statistics. To calculate the seven Pedroni panel cointegration tests we first compute residuals from the regression (2.2) ;

$$y_{it} = \alpha_{0i} + \alpha_{1i} t + x_{it}\beta_i + e_{it} \quad \text{..(2.2)}$$

where $i = 1,2,3,.........., N;\ t = 1,2,3,..........., T$

T denotes the time length and N denotes the number of cross section, y_{it} and x_{it} are assumed integrated of order one I (1) that is

$$x_{it} = x_{it-1} + v_{it} \quad \text{..(2.3)}$$

where x_{it} has K-dimension vector of independent variables. The α_{0i}, and α_{1i} denote the individual specific intercept and trend parameter respectively which vary across the cross sections, $\beta_i = (\beta_{1i}, \beta_{2i}, ..., \beta_{Ki})'$ denotes the cointegration vector which also vary across the cross section. Furthermore, error process $w_{it} = (e_{it}, v'_{it})'$ is assumed independently identically distributed (IID) with respect to cross section. The ordinary least square (OLS) method is used to estimate the cointegration regression (2.2) separately for each individual. Moreover, differenced equation (2.4) is used to calculate tests; $\text{Panel} - v$ and parametric Panel_t

$$\Delta y_{it} = c_{1i}\Delta x_{1it} + c_{2i}\Delta x_{2it} + \text{..........} + c_{Ki}\Delta x_{Kit} + \zeta_{it} \text{ ....(2.4)}$$

After estimating the above differenced equation (2.4), calculate the long run variance of residual ζ_{it} which is denoted by L^2_{11i}

$$L^2_{11i} = \frac{1}{T}\sum_{t=1}^{T}\zeta^2_{it} + \frac{2}{T}\sum_{s=1}^{M_i}(1 - \frac{s}{M_i + 1})\sum_{t=s+1}^{T}\zeta_{it}\zeta_{i,t-s} \text{......................(2.5)}$$

For non parametric statistics, the residual of cointegration regression (2.2) is used in the following regression equation and then compute contemporaneous and long run variance

$$\hat{e}_{it} = \rho_i \hat{e}_{i,t-1} + u_{it} \text{ ....(2.6)}$$

$$\hat{s}^2_i = \frac{1}{T}\sum_{t=1}^{T} u^2_{it} \text{ ....(2.7)}$$

$$\sigma^2_i = \hat{s}^2_i + 2\lambda_i \text{ ....(2.8)}$$

where $\lambda_i = \frac{1}{T}\sum_{s=1}^{M_i}\left(1-\frac{s}{M_i+1}\right)\sum_{t=s+1}^{T} u_{it}\, u_{i,t-s}$ ....(2.9)

$\hat{s}_i^2$ and σ_i^2 denote the contemporaneous and long run variance of u_{it} respectively.

$$\sigma_{NT}^{*2} = \frac{1}{N}\sum_{i=1}^{N} L_{11i}^{-2}\sigma_i^2 \quad \text{...(2.10)}$$

For parametric test statistics that is, $Group\_t$ which is also known as $Group-ADF$ and $Panel\_t$ which is also known as $Panel-ADF$, residual of cointegration regression (2.2) is used in the following regression equation

$$e_{it} = \rho_i \hat{e}_{i,t-1} + \varphi_{i1}\Delta\hat{e}_{i,t-1} + \varphi_{i2}\Delta\hat{e}_{i,t-2} + \ldots\ldots\ldots + \varphi_{ip}\Delta\hat{e}_{i,t-p} + u_{it}^{*} \quad \text{.......(2.11)}$$

$$\tilde{s}_i^{*2} = \frac{1}{T}\sum_{t=1}^{T} u_{it}^{*2} \quad \text{..(2.12)}$$

$$\tilde{s}_{NT}^{*2} = \frac{1}{N}\sum_{i=1}^{N} \tilde{s}_i^{*2} \quad \text{..(2.13)}$$

$\tilde{s}_i^{*2}$ denotes the contemporaneous variance of u_{it}^{*}.

Pedroni's seven test statistics are described as follows:

1. **Panel − v** Statistic:

$$T^2N^{3/2}Z_{vN,T} = T^2N^{3/2}\left(\sum_{i=1}^{N}\sum_{t=1}^{T} L_{11i}^{-2}\hat{e}_{i,t-1}^{2}\right)^{-1} \quad \text{...................(2.14)}$$

2. Panel _ ρ Statistic:

$$TN^{1/2} Z_{\rho N,T^{-1}} = TN^{1/2} \left(\sum_{i=1}^{N} \sum_{t=1}^{T} \hat{e}_{i,t-1}^{2} \right)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} \left(\hat{e}_{i,t-1} \Delta \hat{e}_{it} - \lambda_i \right) \text{..............(2.15)}$$

3. Panel _ t Statistic non parametric

$$Z_{tN,T} = \left(\sigma_{NT}^{2} \sum_{i=1}^{N} \sum_{t=1}^{T} \hat{e}_{i,t-1}^{2} \right)^{-1/2} \sum_{i=1}^{N} \sum_{t=1}^{T} \left(\hat{e}_{i,t-1} \Delta \hat{e}_{it} - \lambda_i \right) \text{...................(2.16)}$$

4. Panel _ t Statistic parametric

$$Z_{tN,T}^{*} = \left(\tilde{S}_{NT}^{*2} \sum_{i=1}^{N} \sum_{t=1}^{T} L_{11i}^{-2} \hat{e}_{i,t-1}^{2} \right)^{-1/2} \sum_{i=1}^{N} \sum_{t=1}^{T} L_{11i}^{-2} \hat{e}_{i,t-1} \Delta \hat{e}_{it} \text{.................(2.17)}$$

5. Group _ ρ Statistic:

$$TN^{-1/2} Z_{\rho N,T^{-1}} = TN^{-1/2} \sum_{i=1}^{N} \left[\left(\sum_{t=1}^{T} \hat{e}_{i,t-1}^{2} \right)^{-1} \sum_{t=1}^{T} \hat{e}_{i,t-1} \Delta \hat{e}_{it} - \lambda_i \right] \text{.........(2.18)}$$

6. Group _ t Statistic non parametric

$$N^{-1/2} Z_{tN,T} = N^{-1/2} \sum_{i=1}^{N} \left[\left(\sigma_i^{2} \sum_{t=1}^{T} \hat{e}_{i,t-1}^{2} \right)^{-1/2} \sum_{t=1}^{T} \hat{e}_{i,t-1} \Delta \hat{e}_{it} - \lambda_i \right] \text{...........(2.19)}$$

7. Group _ t Statistic parametric:

$$N^{-1/2} Z_{tN,T}^{*} = N^{-1/2} \sum_{i=1}^{N} \left[\left(\hat{s}_i^{*2} \sum_{t=1}^{T} \hat{e}_{i,t-1}^{2} \right)^{-1/2} \sum_{t=1}^{T} \hat{e}_{i,t-1} \Delta \hat{e}_{it} \right] \text{.........(2.20)}$$

If we describe the summary of above mentioned Pedroni's seven statistics, we will perform the following steps

1. Estimate the regression equation (2.2) and collect the residuals $\hat{e}_{it}$.

2. For every cross section, estimate the difference regression (2.4) separately and collect the residuals ζ_{it} .

3. Calculate the long run variance L_{11i} of ζ_{it} by using any kernel estimator.

4. Using the residuals $\hat{e}_{it}$ of regression (2.2), estimate the regression (2.6) and ((2.11) for the parametric and non parametric statistics.

The asymptotic distributions of above seven statistics are standard normally distributed

$$\frac{\chi_{NT} - \mu\sqrt{N}}{\sqrt{v}} \Rightarrow N(0,1)$$

Where χ_{NT} denotes the suitably standardized form of the above seven test statistics with respect to the time dimension T and cross section dimension N whereas μ and v denote the moments of Brownian motion functional. Respective values of variance and mean; μ and v respectively are reported in Pedroni (1999).

Null of no cointegration for all above mentioned 07 statistics,

$$H_0 : \rho_i = 1, \forall \ i = 1,2,\ldots\ldots., N$$

The alternate hypothesis is different for both between dimension and within dimension. The alternate hypothesis for between dimension test statistics is

$$H_1 : \rho_i < 1, \quad \forall \ i = 1, 2, \ldots\ldots, N$$

The alternate hypothesis for within dimension statistics is,

$$H_1 : \rho_i = \rho < 1, \quad \forall \ i = 1, 2, \ldots\ldots, N$$

Under the alternate hypothesis only the test statistics $Panel-v$ diverges to positive infinity while other six statistics diverge to negative infinity. So. right tail of the standard normal distribution to reject the null hypothesis of $Panel-v$ is used whereas left tail is used to reject the null hypothesis of other six statistics.

2.2.1.2. Kao Tests (1999)

Parametric residual based (panel cointegration) tests have been proposed by Kao (1999) having null of no cointegration. These tests are extended form of DF and ADF unit root tests to panel cointegration. Kao's tests have based on the spurious least square dummy variable panel regression equation with one single regressor

$$y_{it} = \alpha_i + x_{it}\beta + e_{it} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2.21)$$

Where $i = 1, 2, \ldots\ldots\ldots, N \quad t = 1, 2, \ldots\ldots, T$

y_{it} and x_{it} are integrated of order one that is I (1) series such that,

$$x_{it} = x_{i,t-1} + \nu_{it} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2.22)$$

$$y_{it} = y_{i,t-1} + \varsigma_{it} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2.23)$$

where $\varsigma_{it} \sim N(0, \sigma_\varsigma^2)$ i.i.d. and $v_{it} \sim N(0, \sigma_v^2)$ i.i.d. and also assumed that the error process $w_{it} = (\varsigma_{it}, v_{it})'$ is independent across the individuals i, and it has satisfied the invariance principle. α_i Denotes the intercept of *ith* cross section and which varies for each cross section, in other words heterogeneous across i and β denotes the slope parameter of the *ith* cross section which remains unchanged/fixed for each cross section, in other words heterogeneous across i. Fixed effect Model of regression (2.21) can be estimated using Least square dummy variable or within dimension approach. Kao (1999) has proposed DF type tests using the auto regressive least square dummy variable residuals of regression (2.21),

$$\hat{e}_{it} = \rho \hat{e}_{i(t-1)} + u_{it} \quad \text{...(2.24)}$$

Where ρ is homogeneous.

Null hypothesis of no cointegration is defined as $H_0 : \rho = 1$ which implies that the residuals $\hat{e}_{it}$ are unit root, and Alternate hypothesis is defined as $H_1 : 0 \leq \rho < 1$.

Estimator of ρ is given as,

$$\hat{\rho} = \frac{\sum_{i=1}^{N} \sum_{t=2}^{T} \hat{e}_{it} \hat{e}_{i(t-1)}}{\sum_{i=1}^{N} \sum_{t=2}^{T} \hat{e}_{i(t-1)}^{2}} \quad \text{..(2.25)}$$

The t-statistic to test $\rho = 1$, t_ρ is as follow

$$t_\rho = \frac{\left(\hat{\rho} - 1\right)\sqrt{\sum_{i=1}^{N}\sum_{t=2}^{T}\hat{e}_{i(t-1)}^2}}{S_e} \quad \text{.......................................(2.26)}$$

Where $$s_e^2 = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=2}^{T}\left(\hat{e}_{it} - \hat{\rho}\hat{e}_{i(t-1)}\right)^2 \quad \text{..............................(2.27)}$$

The test statistics of DF type are defined as follows:

1. $$DF_\rho = \frac{\sqrt{N}T\left(\hat{\rho} - 1\right) + 3\sqrt{N}}{\sqrt{10.2}} \quad \text{....................................(2.28)}$$

2. $$DF_t = \sqrt{1.25}t_\rho + \sqrt{1.875N} \quad \text{.....................................(2.29)}$$

In these two DF type tests that is, DF_ρ and DF_t are assumed the strict exogeneity of the regressors with respect to residuals, while the other two tests DF_ρ^* and DF_t^* are allowed for endogeneity of the regressors, the test statistics of DF_ρ^* and DF_t^* are given as,

3. $$DF_\rho^* = \frac{\sqrt{N}T\left(\hat{\rho} - 1\right) + \frac{3\sqrt{N}\hat{\sigma}_u^2}{\hat{\sigma}_{ou}^2}}{\sqrt{3 + \frac{36\hat{\sigma}_u^4}{5\hat{\sigma}_{ou}^2}}} \quad \text{...(2.30)}$$

4. $$DF_t^* = \frac{t_\rho + \frac{\sqrt{6N}\hat{\sigma}_u}{2\hat{\sigma}_{ou}}}{\sqrt{\frac{\hat{\sigma}_{ou}^2}{2\hat{\sigma}_u^2} + \frac{3\hat{\sigma}_u^2}{10\hat{\sigma}_{ou}^2}}} \quad \text{...(2.31)}$$

Where $\hat{\sigma}_u^2$ and $\hat{\sigma}_{ou}^2$ are defined as follow,

$$\hat{\sigma}_u^2 = \hat{\Sigma}_\varsigma + \hat{\Sigma}_{\varsigma v}\Sigma_v^{-1} \text{ ....(2.32)}$$

$$\hat{\sigma}_{ou}^2 = \hat{\Omega}_\varsigma - \hat{\Omega}_{\varsigma v}\hat{\Omega}_v^{-1} \text{....................................(2.33)}$$

$\hat{\sigma}_u^2$ and $\hat{\sigma}_{ou}^2$ are consistent estimator of contemporaneous variance and long run conditional variance respectively of the $w_{it} = (\varsigma_{it}, v_{it})'$. So for this obtained the estimate of $w_{it} = (\varsigma_{it}, v_{it})'$ that is, w_{it} and use for estimation of Σ and Ω.

$$\Sigma = \begin{pmatrix} \sigma_\varsigma^2 & \sigma_{\varsigma v} \\ \sigma_{v\varsigma} & \sigma_v^2 \end{pmatrix} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} w_{it} w_{it}' \text{ ....(2.34)}$$

$$\Omega = \begin{pmatrix} \sigma_{o\varsigma}^2 & \sigma_{o\varsigma v} \\ \sigma_{ov\varsigma} & \sigma_{ov}^2 \end{pmatrix} = \frac{1}{N}\sum_{i=1}^{N}\left\{\frac{1}{T}\sum_{t=1}^{T} w_{it} w_{it}' + \frac{1}{T}\sum_{\tau=1}^{l}\bar{\omega}_{\tau l}\sum_{t=\tau+1}^{T}\left(w_{it} w_{i,t-\tau}' + w_{i,t-\tau} w_{it}'\right)\right\} \text{.....(2.35)}$$

Where $\bar{\omega}_{\tau l}$ denotes the weight function or a kernel, so use appropriate bandwidth and a kernel estimator for the estimation of long run variance and covariances. The asymptotic distribution of DF type tests that is, DF_ρ, DF_t, DF_ρ^*, DF_t^* converge to standard normal distribution $N(0,1)$ as $T \to \infty, N \to \infty$

5. ADF type panel cointegration test has been based on the following auto regressive regression,

$$\hat{e}_{it} = \rho\,\hat{e}_{i(t-1)} + \sum_{j=1}^{p}\varphi_j \Delta \hat{e}_{i(t-j)} + u_{it}^{*} \text{.................................(2.36)}$$

$H_0 : \rho = 1$ (no cointegration) and Alternate hypothesis is defined as, $H_1 : 0 \le \rho < 1$.

$$t_{ADF} = \frac{(\hat{\rho}-1)\left(\sum_{i=1}^{N} e_i' Q_{xp} e_i\right)^{1/2}}{s_{u^*}} \quad \ldots\ldots(2.37)$$

Where ρ is the estimated value of ρ by using OLS and also noted that

$$(\rho-1) = \left(\sum_{i=1}^{N}(e_i' Q_{xp} e_i)\right)^{-1}\left(\sum_{i=1}^{N}(e_i' Q_{xp} \vartheta_i)\right) \quad \ldots\ldots(2.38)$$

$$\Rightarrow \qquad t_{ADF} = \frac{\left(\sum_{i=1}^{N} e_i' Q_{xp} \vartheta_i\right)}{s_{u^*}\left(\sum_{i=1}^{N} e_i' Q_{xp} e_i\right)^{1/2}} \quad \ldots\ldots(2.39)$$

Where e_i is the vector of observation of $\hat{e}_{i,t-1}$, ϑ_i denotes the vector of observations of $\hat{u}^*_{it}$ which is the estimate of u^*_{it}.

$$\text{Where } Q_{xp} = I + X_p\left(X_p'X_p\right)^{-1} X_p' \quad \ldots\ldots(2.40)$$

$$s^2_{u^*} = \frac{1}{T}\sum_{t=1}^{T} \hat{u}_{it}^{*2} \quad \ldots\ldots(2.41)$$

where X_p denotes the matrix of observations on the having p regressors that is, $\left(\Delta\hat{e}_{t-1},\ldots\Delta\hat{e}_{t-p}\right)$

$$ADF = \frac{t_{ADF} + \frac{\sqrt{6N}\,\hat{\sigma}_u}{2\hat{\sigma}_{ou}}}{\sqrt{\frac{\hat{\sigma}_{ou}^2}{2\hat{\sigma}_u^2} + \frac{3\hat{\sigma}_u^2}{10\hat{\sigma}_{ou}^2}}} \quad \ldots\ldots(2.42)$$

Under the null hypothesis of no cointegration ADF statistics is converged to standard normal distribution $N(0,1)$ as $T \to \infty, N \to \infty$.

2.2.1.3. Average Phillips Zt Statistics for Panel Cointegration

As Phillips and Ouliaris (1990) have suggested that the autocorrelation and contemporaneous correlation can be removed through differencing and Augmented Dickey Fuller t-statistics method or through non parametric correction. So McCoskey and Kao (1998) used this idea and considered the test which is based on the average of the Phillips Z_t statistics, which is calculated as given below:

The model is defined as

$$y_{it} = \alpha_i + x_{it}\beta_i + e_{it} \quad \text{...(2.43)}$$

Where $i = 1,2.......,N \quad t = 1,2,.............T$

$$\hat{e}_{it} = \rho_i \hat{e}_{i(t-1)} + u_{it} \quad \text{...(2.44)}$$

$H_0 : \rho_i = 1$ (no cointegration) and Alternate hypothesis is defined as, $H_1 : 0 \le \rho_i < 1$

$$\text{Define} \qquad s_{iu}^2 = \frac{1}{T}\sum_{t=1}^{T} \hat{u}_{it}^2 \quad \text{...(2.45)}$$

$$s_{iT^*}^2 = T^{-1}\sum_{1}^{T} \hat{u}_{it}^2 + 2T^{-1}\sum_{s=1}^{T} w\left(s/\hat{M}\right)\sum_{t=s+1}^{T} u_{it}u_{i(t-s)} \quad \text{................(2.46)}$$

Where $w\left(s/\hat{M}\right)$ denotes the kernel or weighting function, in this case we have used quadratic spectral kernel Andrews (1991) which is given as:

$$w\left(s/\hat{M}\right)=\frac{25}{12\pi^2\left(s/\hat{M}\right)^2}\left\{\frac{\sin\left(6\pi\left(s/\hat{M}\right)\right)/5}{\left(6\pi\left(s/\hat{M}\right)\right)/5}-\cos\left(\left(6\pi\left(s/\hat{M}\right)\right)/5\right)\right\}\text{...........}(2.47)$$

and the associated automatic plug-in bandwidth estimator (Andrews 1991) is given as

$$\hat{M}=1.3221\left\{\frac{4T\hat{r}_1^2}{(1-\hat{r}_1)^4}\right\}^{1/5}\text{..................................}(2.48)$$

Where, $\hat{r}_1$ is the first-order autoregressive coefficient estimate of $\hat{u}_{it}$

$$\hat{Z}_{it}=\frac{(\hat{\alpha}-1)}{s_{iT^*}\left(\sum_{t=2}^{T}\hat{e}_{i(t-1)}^2\right)^{-1/2}}-\frac{\frac{1}{2}\left(s_{iT^*}^2-s_{iu}^2\right)}{s_{iT^*}\left(\frac{1}{T^2}\sum_{t=2}^{T}\hat{e}_{i(t-1)}^2\right)^{1/2}}\text{.......................}(2.49)$$

$\hat{Z}_{it}$ can be simply written in the following form

$$\hat{Z}_{it}=\left\{(\alpha-1)\left(\sum_{2}^{T}\hat{e}_{i(t-1)}^2\right)^{1/2}-\sum_{s=1}^{T}w\left(s/\hat{M}\right)\sum_{t=s+1}^{T}u_{it}u_{i(t-s)}\left(\sum_{t=2}^{T}\hat{e}_{i(t-1)}^2\right)^{-1/2}\right\}\Bigg/\left\{T^{-1}\sum_{1}^{T}u_{it}^2+2T^{-1}\sum_{s=1}^{T}w\left(s/\hat{M}\right)\sum_{t=s+1}^{T}u_{it}u_{i(t-s)}\right\}\text{....}(2.50)$$

The average of cross section $\hat{Z}_{it}$ statistic can be defined as $\overline{Z_t}$.

$$\overline{Z_t}=\frac{1}{T}\sum_{i=1}^{N}\hat{Z}_{it}\text{..}(2.51)$$

Asymptotic distribution of $\overline{Z}_t$ is normal distribution.

2.2.2. Error Correction Based Tests

Error correction based panel cointegration tests are developed by Westerlund (2007) the detail of tests are given below.

2.2.2.1. Westerlund Cointegration Tests (2007)

Westerlund (2007) has developed four panel cointegration tests which are based on structural rather than residual dynamic and Westerlund (2007) has not imposed the common factor restriction, these tests are the generalized version of tests proposed by the Banerjee, Dolado et al. (1998). All these four tests have null hypothesis of no cointegration and are the error correction based tests. These four tests accommodate serially correlated error term and also non-strictly exogenous regressors. These four tests are classified into two categories; two tests have alternative hypothesis that at least one unit/ one individual is cointegrated, these are called group mean tests and two tests have the alternative hypothesis that panel is cointegrated as a whole, these are called panel tests. These four tests are normally distributed. These tests have based on the following error correction model.

$$\Delta y_{it} = \delta_i' d_t + \alpha_i (y_{i,t-1} - \beta_i' x_{i,t-1}) + \sum_{j=1}^{p_i} \alpha_{ij} \Delta y_{i,t-j} + \sum_{j=-q_i}^{p_i} \gamma_{ij} \Delta x_{i,t-j} + e_{it} \quad .....(2.52)$$

where $t = 1, 2, ..., T$ and $i = 1, 2, ..., N$, d_t denotes the deterministic part and there are three possible cases for d_t, If $d_t = 0$ so the deterministic part has been eliminated in the regression (2.52), if $d_t = 1$ then only the intercept will be included in the regression (2.52), if $d_t = (1, t)'$ then both intercept and trend will be included in regression. The vector x_{it} having dimension K is integrated of order one. Moreover it is assumed that the error process is independent across both i and t. Further added that dependency across the cross section will be handled through bootstrapping. Regression (2.52) can be written as

$$\Delta y_{it} = \delta_i' d_t + \alpha_i y_{i,t-1} + \lambda_i' x_{i,t-1} + \sum_{j=1}^{p_i} \alpha_{ij} \Delta y_{i,t-j} + \sum_{j=-q_i}^{p_i} \gamma_{ij} \Delta x_{i,t-j} + e_{it} \quad \text{..........(2.53)}$$

where $\lambda_i' = -\alpha_i \beta_i'$. The parameter α_i has determined the rate of speed at which system converges to the equilibrium relationship $y_{i,t-1} - \beta_i' x_{i,t-1}$ after the sudden shock. If $\alpha_i < 0$ then $x_{i,t}$ *and* $y_{i,t}$ are cointegrated, so error correction term will exist, if $\alpha_i = 0$ then $x_{i,t}$ *and* $y_{i,t}$ are not cointegrated, thus error correction term will not exist. Thus null hypothesis is stated as $H_0 : \alpha_i = 0 \ \forall$ i. the alternative hypothesis is assumed in two different ways depend upon the homogeneity of α_i. Two tests which are called group mean tests, have not assumed α_i's to be equal, so the alternative hypothesis is $H_1 : \alpha_i < 0$ for at least one i, the other pair of the tests which are called panel tests, are assumed that all α_i are equal and hence the alternative hypothesis is $H_1 : \alpha_i = \alpha < 0$ for all *i*.

Estimation Procedure for Westerlund (2007) Group Mean Tests

To estimate the group mean tests we follow the three steps, the first step is to estimate the regression equation (2.53) using OLS for every cross section separately, which is given

$$\Delta y_{it} = \hat{\delta}_i d_t + \hat{\alpha} y_{i(t-1)} + \hat{\lambda}_i x_{i(t-1)} + \sum_{j=1}^{p_i} \hat{\alpha}_{ij} \Delta y_{i(t-j)} + \sum_{j=-q_i}^{p_i} \hat{\gamma}_{ij} \Delta x_{i(t-j)} + \hat{e}_{it} \quad .....(2.54)$$

where p_i and q_i denote the lag and lead, which can be varied across the individuals and lag and lead are determined by any information criterion. After obtaining $\hat{e}_{it}$ and $\hat{\gamma}_{ij}$ the second step is computed

$$\hat{\varepsilon}_{it} = \sum_{j=-q_i}^{p_i} \hat{\gamma}_{ij} \Delta x_{i(t-j)} + \hat{e}_{it} \quad ....(2.55)$$

which we use to estimate the $\hat{\alpha}_i(1) = \frac{\hat{w}_{\varepsilon i}}{\hat{w}_{yi}}$ where $\hat{w}_{\varepsilon i}$ and $\hat{w}_{yi}$ are the long run variance usual Newey and West (1994) based on $\hat{\varepsilon}_{it}$ and Δy_{it} respectively.

Bandwidth parameter selection problem has been occurred during the estimation procedure for α_i, so for this when intercept or both intercept and trend are included in regression model (2.53) , the estimation of $\hat{w}_{yi}$ using a kernel estimator instead of Δy_{it}, residuals from first stage regression of Δy_{it} on d_t is being used.

Third step is to estimate the group means statistics in the following way:

$$G_T = \frac{1}{N} \sum_{i=1}^{N} \frac{\hat{\alpha}_i}{SE(\hat{\alpha}_i)} \quad ....(2.56)$$

where $SE(\hat{\alpha}_i)$ standard error of $\hat{\alpha}_i$.

$$G_\alpha = \frac{1}{N} \sum_{i=1}^{N} \frac{T \hat{\alpha}_i}{\hat{\alpha}_i(1)} \quad ....(2.57)$$

Estimation Procedure for Westerlund (2007) Panel Tests:

Three steps procedure is followed to estimate the Panel tests,

First step is to regress the following two regressions separately by using OLS, and collect the residuals of both regressions that is, $\hat{e}_{it}$ and $\hat{\psi}_{it}$ which are given as below,

$$\hat{e}_{it} = \Delta y_{it} - \hat{\delta}_i' d_t - \hat{\lambda}_i' x_{i(t-1)} - \sum_{j=1}^{p_i} \hat{\alpha}_{ij} \Delta y_{i(t-j)} - \sum_{j=-q_i}^{p_i} \hat{\gamma}_{ij} \Delta x_{i,t-j} \text{............}(2.58)$$

$$\hat{\psi}_{it} = y_{i(t-1)} - \hat{\delta}_i' d_t - \hat{\lambda}_i' x_{i(t-1)} - \sum_{j=1}^{p_i} \hat{\alpha}_{ij} \Delta y_{i(t-j)} - \sum_{j=-q_i}^{p_i} \hat{\gamma}_{ij} \Delta x_{i(t-j)} \text{.........}(2.59)$$

Second step is to estimate the common error correction parameter α and its standard error by using $\hat{e}_{it}$ and $\hat{\psi}_{it}$;

$$\hat{\alpha} = \left(\sum_{i=1}^{N} \sum_{t=2}^{T} \hat{\psi}_{it}^2 \right)^{-1} \sum_{i=1}^{N} \sum_{t=2}^{T} \frac{1}{\hat{\alpha}_i(1)} \hat{\psi}_{it} \hat{e}_{it} \text{..............................}(2.60)$$

$$SE\left(\hat{\alpha}\right) = \left(\left(\hat{S}_N^2 \right)^{-1} \sum_{i=1}^{N} \sum_{t=2}^{T} \hat{\psi}_{it}^2 \right)^{-1/2} \text{...........................}(2.61)$$

Where $SE\left(\hat{\alpha}\right)$ denotes the standard error of $\hat{\alpha}$.

where $$\hat{S}_N^2 = \frac{1}{N} \sum_{i=1}^{N} \hat{S}_i^2 \text{..}(2.62)$$

$$\hat{S}_i = \frac{\hat{\sigma}_i}{\hat{\alpha}_i(1)} \text{...................................}(2.63)$$

$\hat{\sigma}_i$ denotes the regression standard error in equation (2.53).

Third step is to estimate the panel statistics as follows

$$P_T = \frac{\hat{\alpha}}{SE\left(\hat{\alpha}\right)} \qquad \text{.................................(2.64)}$$

$$P_{\alpha} = T\hat{\alpha} \qquad \text{.......................................(2.65)}$$

The asymptotic distribution of these four tests is normal distribution as the N and T approach to infinity sequentially. In other words these four statistics are normally distributed after standardizing with appropriate moments and also these asymptotic distributions and moments have depended upon the deterministic part and number of regressors included in the regression model.

The test statistics P_T, P_α, G_T and G_α diverge to negative infinity under the alternative hypothesis, so the test decision has been made on the left tail of the standard normal distribution. Furthermore, the Westerlund (2007) has suggested using bootstrap approach of Chang (2004) for panel cointegration testing, which compensate for cross-sectional dependence.

2.2.3. Average Weighted Symmetric Test (AWS)

Average weighted symmetric (AWS) test proposed by Hoang (2006), AWS test is based on the idea of average test statistics for each cross section which is used in Im, Pesaran et al. (2003) for panel unit root. The procedure of AWS test is similar to the McCoskey and Kao (1998) test. In the AWS test weighted symmetric estimation is used for every cross section. Run the following regression for the estimation of WS panel statistics,

$$y_{it} = \alpha_i + x_{it}\beta + e_{it} \qquad \text{...(2.66)}$$

Where $i = 1, 2, \ldots\ldots\ldots, N \quad t = 1, 2, \ldots\ldots, T$

After estimating the regression (2.66) the residuals $\hat{e}_{it}$ is obtained, then estimate the following equation,

$$\hat{e}_{it} = \rho_i \hat{e}_{i(t-1)} + u_{it} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2.67)$$

where $i \in N,\ t \in T$.

Null hypothesis of no cointegration is given as, $H_o : \rho_i = 1$

Under the alternative hypothesis $H_1 : \rho_i < 1$.

$$\hat{\rho}_i = \frac{\sum_{t=2}^{T} \hat{e}_{it} \hat{e}_{i(t-1)}}{\sum_{t=2}^{T-1} \hat{e}_{it}^2 + \frac{1}{T}\sum_{t=1}^{T} \hat{e}_{it}^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2.68)$$

where $\hat{e}_{it}$ denotes the residual of *ith* cross section.

Weighted symmetric test statistic of *ith* cross section is given as is:

$$t_{iws} = \left(\hat{\sigma}_{u_{it}}^2 \left(\sum_{t=2}^{T-1} \hat{e}_{it}^2 + \frac{1}{T}\sum_{t=1}^{T} \hat{e}_{it}^2 \right)^{-1} \right)^{\frac{-1}{2}} \left(\hat{\rho}_i - 1 \right) \ldots\ldots\ldots\ldots\ldots\ldots(2.69)$$

$$\text{where } \hat{\sigma}_{u_{it}}^2 = \frac{1}{T}\sum_{t=2}^{T} \left(\hat{e}_{it} - \hat{e}_{it} \right)^2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2.70)$$

Finally weighted symmetric test statistic for panel is computed after taking the average of all cross section weighted symmetric test statistic which is given as follows:

$$\bar{t}_{ws} = \frac{1}{N}\sum_{i=1}^{N} t_{iws} \quad \text{..................................(2.7 1)}$$

Whereas the WS panel statistics converge to N (0, 1) by using the central limit theorem such that

$$\sqrt{N}\left(\frac{\bar{t}_{ws} - \mu_{ws}}{\sigma_{ws}}\right) \xrightarrow{L} N(0,1), \text{ Where } E[t_{iws}] = \mu_{ws}, \ var(t_{iws}) = \sigma^2_{ws}$$

The value of μ_{ws} and σ^2_{ws} is found through simulation and given in the table 1 Hoang (2006).

If the residuals e_{it} of (2.67) are correlated then use the augmented equation (2.72) instead of using (3.67)

$$\hat{e}_{it} = \rho_i \hat{e}_{i(t-1)} + \sum_{j=1}^{p} \varphi_j \Delta \hat{e}_{i(t-j)} + u^*_{it} \text{..................................(2.72)}$$

2.2.4. Fisher Type Tests

The Fisher type panel cointegration tests have also developed byHoang (2006) and the detail of the tests are discussed below

2.2.4.1. Fisher Augmented Dickey-Fuller Test (FADF)

Hoang (2006) has developed the Fisher type tests and based on the Fisher approach to combine the p-value of all test statistics of each cross sections which has introduced in Maddala and Wu (1999). The Procedure of Fisher Augmented Dickey-Fuller (FADF) test is similar to the panel unit root test of Maddala and Wu (1999). The FADF tests have also been based on the results of ADF t-statistics for cointegration from Phillips and Ouliaris (1990), which converge to the standard Wiener process.

Run the following regression (2.73) for each cross section and obtained the residuals

$$y_{it} = \alpha_i + x_{it}\beta + e_{it} \quad \text{...(2.73)}$$

Where $i = 1, 2, \ldots\ldots\ldots, N \quad t = 1, 2, \ldots\ldots\ldots, T$

After obtaining the residual from (15), estimate the following regression.

$$\hat{e}_{it} = \rho_i \hat{e}_{i(t-1)} + \sum_{j=1}^{p} \varphi_j \Delta \hat{e}_{i(t-j)} + u_{it}^{*} \text{..................................(2.74)}$$

Null hypothesis of no cointegration is given as, $H_o : \rho_i = 1$

Under the alternative hypothesis $H_1 : \rho_i < 1$.

Then find t-statistics of regression (2.74) for each cross section,

$$t_{iADF} = \frac{\left(\hat{\rho}_i - 1\right)}{\hat{\sigma}_i} \quad \text{..(2.75)}$$

The FADF test is required to drive the distribution of the DF t- statistics, so for this simulated t- statistics values are generated for each cross section, and then calculate the p-values π_{iadf} for each t-statistics. Finally FADF test statistics is calculated as;

$$P_{FADF} = -2\sum_{i=1}^{N} \log_e \pi_{iadf} \sim \chi^2_{2N} \text{..................................(2.76)}$$

We will use χ^2 table for the critical value of the test statistics of FADF.

2.2.4.2. Fisher Weighted Symmetric Test (FWS)

The procedure of FWS test is approximately similar to the procedure of the FADF tests, only the disparity in procedure, use the WS estimation procedure for each cross section rather than the ADF estimation procedure. The WS estimation procedure is already defined above in panel AWS test. The t_{iws} is calculated for each cross section which is given as follows:

$$t_{iws} = \left(\hat{\sigma}^2_{u_{it}} \left(\sum_{t=2}^{T-1} \hat{e}^2_{it} + \frac{1}{T} \sum_{t=1}^{T} \hat{e}^2_{it} \right)^{-1} \right)^{\frac{-1}{2}} \left(\hat{\rho}_i - 1 \right) \text{.........................} (2.77)$$

$$\text{where } \hat{\sigma}^2_{u_{it}} = \frac{1}{T} \sum_{t=2}^{T} \left(\hat{e}_{it} - \hat{e}_{it} \right)^2 \text{.................................} (2.78)$$

Where $\hat{e}_{it}$ are the residuals of the regression (2.66).

Null hypothesis of no cointegration is given as, $H_o : \rho_i = 1$

Under the alternative hypothesis $H_1 : \rho_i < 1$.

To derive the distribution of FWS test, generate the simulated t_{iws} values and then calculate the p-values π_{iws} for each cross section. Then final Panel FWS statistics is given as,

$$P_{FWS} = -2 \sum_{i=1}^{N} \log_e \pi_{iws} \sim \chi^2_{2N} \text{...} (2.79)$$

We use the χ^2 table for the critical values of the table.

2.2.5. Maximum Likelihood Based Test

Maximum-likelihood based test for panel cointegration has been proposed by the Larsson, Lyhagen et al. (2001). In this test standardized LR-bar statistics has been proposed, which are based on the average of individual rank trace statistics of Johansen (1995). Heterogeneous VAR model for K-dimensional process has been considered by the Larsson, Lyhagen et al. (2001), which is given as follows:

$$Y_{it} = \sum_{j=1}^{q_i} \pi_{ij} Y_{i,t-j} + e_{it} \ldots\ldots\ldots\ldots\ldots\ldots\ldots 2.80)$$

Where $i = 1,2,3,\ldots\ldots, N$, $t = 1,2,\ldots\ldots, T$, the residuals e_{it} are i.i.d, $e_{it} \sim N(0, \Omega_i)$.

Johansen (1995). has also introduced error correction representation of k-dimensional VAR model, so the regression (2.80) has been given in heterogeneous error correction model representation,

$$\Delta Y_{it} = \Pi_i Y_{i,t-1} + \sum_{j=1}^{q_i-1} \Gamma_{ij} \Delta Y_{i,t-j} + e_{it} \ldots\ldots\ldots\ldots\ldots\ldots (2.81)$$

$i = 1,2,\ldots\ldots\ldots, N \quad t = 1,2,\ldots\ldots\ldots\ldots, T$

Where the order of Π_i is $p \times p$, p denotes the number of regressors in each cross section $\Pi_i = -\left(I_k - A_{i1} - \ldots\ldots\ldots - A_{i,p_i}\right)$, If Π_i is of reduced rank then Π_i can be decomposed to $\Pi_i = \alpha_i \beta_i'$ where α_i (the adjustment parameter) and β_i (the long-run coefficients) be the matrices of order $p \times r_i$ with full column rank, Note that the dimension of time T should large enough , so that the regression (81) can be estimated for each cross section separately.

Larsson, Lyhagen et al. (2001) has considered the null hypothesis that there are at most r cointegrating relations among the p variables in all cross section N. So the null hypothesis can be expressed as,

$$H_0 : \text{rank}(\Pi_i) < r_i \leq r, \quad H_1 : \text{rank}(\Pi_i) = p \quad \forall\ i$$

The Likelihood ratio test which is called trace statistic, denoted by LR_{iT} for the *ith* cross section which is given as below:

$$LR_{iT}\{H_0 \mid H_1\} = -2\ln Q_{iT}\{H_0 \mid H_1\} = -T\sum_{i=r+1}^{p}\ln(1-\lambda_i)\text{...........}(2.82)$$

Where λ_i is the Eigen value of the *ith* cross section. Now define the trace statistics for panel cointegration by taking the average of N cross section,

$$\overline{LR}_{NT}\{H_0 \mid H_1\} = \frac{1}{N}\sum_{i=1}^{N} LR_{iT}(H_0 \mid H_1)\text{......................}(2.83)$$

The standardized LR-bar statistics is given as follows:

$$\gamma_{\overline{LR}_{NT}}\{H_0 \mid H_1\} = \frac{\sqrt{N}\left(\overline{LR}_{NT}\{H_0 \mid H_1\} - E(Z_K)\right)}{\sqrt{var(Z_K)}}\text{..................}(2.84)$$

Where $E(Z_k)$ and $var(Z_k)$ denote the mean and variance respectively of the asymptotic trace statistics. Asymptotic distribution of panel trace statistics will converge to standard normal N (0, 1) as T and N converge to infinity.

2.3. Panel cointegration Tests for Null Hypothesis of Cointegration

In panel framework, residual based cointegration tests was proposed by McCoskey and Kao (1998) for null of cointegration. Harris and Inder (1994) and also Shin (1994) have proposed the univariate LM test, which is extended in panel cointegration frame work by McCoskey and Kao (1998). Suppose y_{it} and x_{it} are integrated of order one.

$$y_{it} = \alpha_i + x_{it}\beta_i + e_{it} \quad \text{.............................}(2.85)$$

$$x_{it} = x_{i,t-1} + v_{it} \quad \text{.....................................}(2.86)$$

$$e_{it} = r_{it} + u_{it} \quad \text{...}(2.87)$$

$$r_{it} = r_{i,t-1} + \varphi u_{it} \quad \text{......................................}(2.88)$$

where $u_{it} \sim N(0, \sigma_u^2)$

Where $i = 1, 2, ..., N \quad t = 1, 2, ..., T$

By using backward substitution in the above system we can get the following equation

$$y_{it} = \alpha_i + x_{it}\beta_i + \varphi\sum_{j=1}^{t} u_{ij} + u_{it} \quad \text{...........................}(2.89)$$

$H_0 : \varphi = 0$ (Null Hypothesis of Cointegration).

$H_1 : |\varphi| \neq 0$ (Null Hypothesis of no Cointegration).

The test statistics is defined as follows,

$$LM = \frac{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=1}^{T} S_{it}^2}{S^2} \quad \text{((2.90)}$$

Where $$S_{it} = \sum_{j=1}^{t} \hat{e}_{ij} \quad \text{(2.91)}$$

Where S_{it} denotes the partial sum of the residuals

Where $$S^2 = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t}^{T} \hat{e}_{it}^2 \quad \text{(2.92)}$$

I f $\varphi = 0$ t hen residuals are stationary and x_{it} and y_{it} are cointegrated.

Furthermore authors assumed that the error process u_{it} and ξ_{it} is dependent weakly and heterogeneously. OLS estimator is asymptotically biased when the residuals are serially correlated and regressors are endogenous, for unbiased estimation McCoskey and Kao (1998). have proposed to use optimal estimator, in which one is fully-modified OLS (FMOLS) estimator of Phillips and Hansen (1990) and other one is dynamic OLS (DOLS) estimator of Saikkonen (1991).

McCoskey and Kao (1998) have proposed the panel LM test based on the FMOLS estimator, for this long run covariance matrix of w_{it} is defined as:

$$\Omega = \begin{pmatrix} w_1^2 & w_{12} \\ w_{21} & \Omega_{22} \end{pmatrix} \quad \text{(2.93)}$$

$$u_{it}^{+} = u_{it} - \hat{w}_{12}\,\hat{\Omega}_{22}\,\xi_{it} \quad \text{(2.94)}$$

$$y_{it}^{+} = y_{it} - \hat{w}_{12}\,\hat{\Omega}_{22}^{-1}\,\xi_{it} \quad \text{(2.95)}$$

Where $\hat{\Sigma}_{22}$ and $\hat{w}_{12}$ are the consistent estimators of Σ_{22} and w_{12} respectively.

The panel LM test statistics using FMOLS is defined as follows:

$$LM^{+} = \frac{\frac{1}{N}\sum_{i=1}^{N}\left[\frac{1}{T^2}\sum_{t=1}^{T}\left(\sum_{j=1}^{t}\hat{e}_{ij}^{+}\right)^2\right]}{\hat{w}_{1.2}^{2}} \quad \text{.................................(2.96)}$$

Where $\hat{e}_{it}^{+} = y_{it}^{+} - \alpha_i - x_{it}^{'}\beta_{iFM}^{+}$ are the fully modified OLS residuals and β_{iFM}^{+} denotes FMOLS beta, where $w_{1.2}^{2}$ denotes the long run variance for all T and N, which is defined as follows

$$\hat{w}_{1.2}^{2} = \hat{w}_{1}^{2} - \hat{w}_{12}^{2}\hat{\Omega}_{22}^{-1}w_{21} \quad \text{......................(2.97)}$$

$$\beta_{iFM}^{+} = (x_{it}^{'}x_{it})^{-1}(x_{it}^{'}y_{it}^{+} - T\delta_i^{+}) \quad \text{.............................(2.98)}$$

$$\text{Where } \delta_i^{+} = \Pi_{21} - \Pi_{22}\Omega_{22}w_{21} \text{.......................(2.99)}$$

$$\Pi = \Sigma + \Gamma \text{......................................(2.100)}$$

$$\Sigma = \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}E(w_{it}w_{it}^{'}) \text{.......................................(2.101)}$$

$$\Gamma = \lim_{T\to\infty}\frac{1}{T}\sum_{k=1}^{T-1}\sum_{t=k+1}^{T}E(w_{it}w_{i,t-k}^{'}) \quad \text{......................(2.102)}$$

Σ and Γ denote the contemporaneous variance and long run auto covariances respectively.

Average of individuals of locally best unbiased invariant (LBUI) test is known as LM^{+} test, which is extension of LBUI (locally best unbiased invariant) test of Harris and Inder

(1994) and also Shin (1994). Similarly LM^{+}_DOLS can also be computed using DOLS regression, for this LM^{+} test statistics based on the following dynamics regression,

$$y_{it} = \alpha_i + x_{it}\beta_i + \sum_{j=-p}^{p} c_{ij}\Delta x_{i,t+j} + e_{it}^{*} \qquad \text{............(2.103)}$$

Thus in (2.96) instead of using $\hat{e}_{it}^{+}$, the residual e_{it}^{*} of (2.103) is used to calculate LM^{+}.

The standardized Panel LM statistic for DOLS and FMOLS is calculated as given below:

$$Z^{+} = \frac{\sqrt{N}\left(LM^{+} - \mu\right)}{\sqrt{\sigma^2}} \qquad \text{..................................(2.104)}$$

$\sqrt{N}\left(LM^{+} - \mu\right) \Rightarrow N(0, \sigma^2)$

Chapter 03:Methodology

For the assessment of power and size of under consideration twenty four panel cointegration tests, in which twenty one tests have no cointegration in their null hypothesis and three tests have cointegration in their null hypothesis, this study uses artificial data generating experiments using time dimension $T = 10, 25, 50, 100$ and cross sectional length $N = 02, 08, 16, 32$. As this study analyzes both types of tests in which some have no cointegration as null hypothesis whereas some have cointegration as null hypothesis, therefore, two data generating processes (DGP) for Monte Carlo experiments are introduced., From these two DGPs, one is for no cointegration as null hypothesis and the other is for null cointegration as null hypothesis. This study compares tests using stringency criterion because this method considers the whole alternate space. In this chapter, firstly details of the both types of DGP are laid down. Then point optimal tests and stringency criterion are listed. Lastly, the step by step algorithms for simulated critical values, empirical size and power curve are discussed.

3.1 Artificial Data Generation (DGP)

This study is considering the simplest DGP for both types of tests. If a test performs worst in simplest case, it is obvious that it will perform worst in all scenarios. However, if the test is robust in this simplest case then it may not be robust in more complex DGPs. It can be explored in further studies.

To compare the cointegration tests in heterogeneous panels, this study has considered two different artificial data generations (DGP).

3.1.1. DGP-A: No Cointegration as Null Hypothesis

Let x_{it} and y_{it} be two series where $i \in N$ and $t \in T$. These two series are generated using following statistical/ econometric model

$$y_{it} = \alpha_i + \beta_i x_{it} + e_{it} \quad \text{.......................................(3.1)}$$

$$e_{it} = \rho e_{it-1} + u_{it} \quad \text{.......................................(3.2)}$$

$$x_{it} = x_{it-1} + v_{ixt} \quad \text{.......................................(3.3)}$$

Where $u_{it} \sim N(0,1)$, $v_{ixt} \sim N(0,\sigma_i^2)$.

Where "i" represents the individual and "t" denotes the time dimension. α_i, β_i and σ_i can take any value such that $\beta_i \neq 0$ and $\sigma_i > 0$. But in this case, assume that $\alpha_i \sim U(0,10), \beta_i \sim U(0,2)$ and $\sigma_i^2 \sim U(.5,1.5)$.

Under no cointegration as null hypothesis $\rho = 1$, whereas under cointegration as alternate hypothesis $0 \leq \rho < 1$. In our analysis we will take the values of ρ under the alternative hypothesis as $\rho = \{0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1, 0\}$

3.1.2.D GP-B: Null of Cointegration

Let x_{it} and y_{it} be two series where $i \in N$ and $t \in T$. These two series are generated using following statistical/ econometric model,

$$y_{it} = \alpha_i + \beta_i x_{it} + e_{it} \quad (3.4)$$

$$e_{it} = e_{i(t-1)} - \phi u_{iy(t-1)} + u_{iyt} \quad (3.5)$$

$$x_{it} = x_{it-1} + v_{ixt} \quad (3.6)$$

Where $\eta_{it} = (u_{iyt}, v_{ixt})$ and $\eta_{it} \sim N(0, \Omega_i)$ where $\Omega_i = \begin{pmatrix} 1 & 0 \\ 0 & \sigma_i^2 \end{pmatrix}$.

Where "i" represents the individual and "t" denotes the time dimension. α_i, β_i and σ_i can take any value such that $\beta_i \neq 0$ and $\sigma_i > 0$. But in this case, assume that $\alpha_i \sim U(0,10), \beta_i \sim U(0,2)$ and $\sigma_i^2 \sim U(.5,1.5)$.

Under cointegration as the Null hypothesis $H_0 : \phi = 1$ where under no cointegration as the alternate hypothesis $H_A : 0 \leq \phi < 1$. Under the alternative hypothesis we will take the values of ϕ as $\phi = \{0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1, 0\}$

The case A's null hypothesis is same as case B's alternate hypothesis., Similarly, case B's null hypothesis is same as Alternative hypothesis in case A. Now before describing the point optimal tests for the above stated DGP-A and DGP-B, first Neyman-Pearson Lemma is stated.

3.2. Neyman-Pearson Lemma

Suppose $X_1, X_2, X_3, \ldots\ldots, X_n$ denote the random sample. The likelihood functions of these random samples under the null and alternative hypothesis are $f^X(x, \xi_0)$ and $f^X(x, \xi_1)$ respectively. Then there exist a constant η_α such that $P(PO(X) \leq \eta_\alpha / H_0) = \alpha$ where $\alpha > 0$

and $PO(X)$ denotes Likelihood ratio, defined as $PO(X)=\frac{f^X(x,\xi_0)}{f^X(x,\xi_1)}$. According to Neyman-Pearson lemma, for a given size α this $PO(X)$ is the test of maximum power at point null and alternative ξ_0 and ξ_1 respectively. In literature, also known as Point Optimal test.

3.3. Point Optimal (PO) Test for No Cointegration as Null Hypothesis

To derive the PO test for no cointegration as null hypothesis, in panel frame work, this study has modified the PO test of Jansson (2008). This study is considering the simple case by taking only two series x_t and y_t having time length T.

$$y_{it} = \alpha_i + \beta_i x_{it} + e_{it} \quad \text{...} (3.7)$$

$$e_{it} = \rho e_{it-1} + u_{it} \quad \text{...} (3.8)$$

$$x_{it} = x_{it-1} + v_{it} \quad \text{...} (3.9)$$

where $u_{it} \sim N(0,1)$, $v_{it} \sim N(0,\sigma_i^2)$.

Where "i" represents the individual and "t" denotes the time dimension.

$H_0 : \rho = 1,\ H_A : \rho < 1$

Based on Neyman Pearson lemma, PO test is:

$$LLR = \frac{1}{2}\left\{\sum_{i=1}^{N}\left(\sum_{t=2}^{T}\left(e_{it} - e_{i(t-1)}\right)^2\right) - \sum_{i=1}^{N}\left(\sum_{t=2}^{T}\left(e_{it} - \rho e_{it-1}\right)^2\right)\right\} \text{........} (3.10)$$

Where N denotes the cross sectional dimension. Hence in this case, LLR is the panel PO test for no cointegration as null hypothesis.

3.4. Point Optimal (PO) Test for Cointegration as the Null Hypothesis

To describe, in panel framework, the PO test for cointegration as the null hypothesis, this study has used and modified the concept of Jansson (2005). For this, Point optimal test is described for a single cross section and then extended this point optimal test for panel data. The following procedure is adopted for two series x_t and y_t having time length "T".

$$y_{it} = \alpha_i + \beta_i x_{it} + e_{it} \text{ ....(3.11)}$$

$$e_{it} = e_{i(t-1)} - \phi u_{iy(t-1)} + u_{iyt} \text{ ....(3.12)}$$

$$x_{it} = x_{it-1} + v_{ixt} \text{ ....(3.13)}$$

Where $\eta_{it} = (u_{iyt}, v_{ixt})$ and $\eta_{it} \sim N(0, \Omega_i)$ where $\Omega_i = \begin{pmatrix} 1 & 0 \\ 0 & \sigma_i^2 \end{pmatrix}$.

Where "i" represents the individual and "t" denotes the time dimension. α_i, β_i and σ_i can take any value such that $\beta_i \neq 0$ and $\sigma_i > 0$.

Under cointegration as the Null hypothesis $H_0 : \phi = 1$ where under no cointegration as the alternate hypothesis $H_A : 0 \leq \phi < 1$.

Assume that, $H_A : \phi = \phi^*$ where $0 \leq \phi^* < 1$ and $R = (x_{it}, D_{it})$ where D_{it} denotes the deterministic component. Partition Ω_i in conformity with η_{it} as

$$\Omega_i = \begin{bmatrix} \sigma_{yy} & \sigma'_{xy} \\ \sigma_{xy} & \Omega_{xx} \end{bmatrix} \text{ ....(3.13)}$$

and define $\Sigma_\phi = \Sigma_\phi^{1/2}\Sigma_\phi^{1/2\prime}$ where $\Sigma_\phi^{1/2}$ is a lower triangular matrix of order $T\times T$ given as

$$\Sigma_\phi^{1/2} = \begin{bmatrix} 1 & 0 & 0 & \bullet & \bullet & 0 \\ 1-\phi & 1 & 0 & \bullet & \bullet & 0 \\ 1-\phi & 1-\phi & 1 & \bullet & \bullet & 0 \\ \bullet & \bullet & \bullet & \bullet & \bullet & 0 \\ \bullet & \bullet & \bullet & \bullet & \bullet & 0 \\ 1-\phi & 1-\phi & 1-\phi & \bullet & \bullet & 1 \end{bmatrix} \text{..(3.14)}$$

$$\sigma_{yy\bullet x} = \sigma_{yy} - \sigma_{xy}'\Omega_{xx}^{-1}\sigma_{xy} \text{ ....(3.15)}$$

Then Likelihood function given by Jansson (2005) is

$$LH(\phi) = \left|R'\Sigma_\phi^{-1}R\right| * \exp(\sigma_{yy\bullet x}^{-1}Y_\phi'\left(\Sigma_\phi^{-1} - \Sigma_\phi^{-1}R\left(R'\Sigma_\phi^{-1}R\right)^{-1}R'\Sigma_\phi^{-1}\right)Y_\phi) \text{ ....(3.16)}$$

Where $Y_\phi = Y_{it} - \phi\Sigma_\phi^{-1/2}x_{it}\Sigma_{xx}^{-1}\sigma_{xy}$ ....(3.17) .

As there are " $i = 1,2,3.........,N$ " independent cross section so for the " ith " cross section, Likelihood function is

$$LH(\phi_i) = \left|R_i'\Sigma_{\phi_i}^{-1}R_i\right| * \exp(\sigma_{yy\bullet x(i)}^{-1}Y_{\phi_i}'\left(\Sigma_{\phi_i}^{-1} - \Sigma_{\phi_i}^{-1}R_i\left(R_i'\Sigma_{\phi_i}^{-1}R_i\right)^{-1}R_i'\Sigma_{\phi_i}^{-1}\right)Y_{\phi i}) \text{.........(2.18)}$$

As it is assumed that these entire "N" cross sections are independent, so their joint Likelihood function is given as

$$JLH(\phi) = LH(\phi_1) \times LH(\phi_2) \times \text{.........} \times LH(\phi_N) \text{ ....(3.19)}$$

$$JLH(\phi) = \prod_{i=1}^{N}\left(\left|R'\Sigma_\phi^{-1}R\right|\right) \times \exp\left(\sum_{i=1}^{N}\left(\sigma_{yy\bullet x(i)}^{-1}Y_{\phi_i}'\left(\Sigma_{\phi_i}^{-1} - \Sigma_{\phi_i}^{-1}R_i\left(R_i'\Sigma_{\phi_i}^{-1}R_i\right)^{-1}R_i'\Sigma_{\phi_i}^{-1}\right)Y_{\phi i}\right)\right) \text{.....(3.20)}$$

Taking log both sides we get,

$$\ell(\phi) = \sum_{i=1}^{N} \log\left(\left|R'\Sigma_{\phi}^{-1}R\right|\right) + \sum_{i=1}^{N}\left(\sigma_{yy\bullet x(i)}{}^{-1}Y_{\phi_i}'\left(\Sigma_{\phi_i}^{-1} - \Sigma_{\phi_i}^{-1}R_i\left(R_i'\Sigma_{\phi_i}^{-1}R_i\right)^{-1}R_i'\Sigma_{\phi_i}^{-1}\right)Y_{\phi i}\right) \quad \text{..........(3.21)}$$

Panel data PO test having cointegration as the null hypothesis is given by

$$PO = \ell(1) - \ell(\phi^*) \quad \text{......................................(3.22)}$$

where $0 \leq \phi^* < 1$

3.5. Stringency Criterion

- **Shortcoming of the Test**

  Suppose Υ_τ^i indicates the test τ's power estimated at a specific point alternative hypothesis *"i"* and PO^i denotes the PO test's power at the same specific point alternative hypothesis *"i"*, then τ's shortcoming $S(\tau)$ is the gap between the PO test's power and test τ's power at *"i"*, i.e. $S(\tau)=PO^i - \Upsilon_\tau^i$ where $S(\tau) \geq 0$.

- **Stringent Test**

Suppose there are N tests, having K number of point alternative hypotheses. Then, shortcomings of each test are calculated at K points of alternative hypotheses. Therefore, the maximum shortcoming or stringency of each test at all given K points is,

$$\Pi_\tau = \max_{i=1,2,\ldots,K} S_i(\tau) \text{ where } \tau \in N ,$$

The test which has the minimum of these stringencies is declared to be the most stringent test, defined as, $MT = \min(\max_{i=1,2,\ldots,K} S_i(\tau) / \forall \ \tau \in N)$

3.6. Mote Carlo Simulation (MCS) Design

This section explains and elaborates the step by step algorithms used for calculating the critical values based on artificial data generations, empirical size based on asymptotic as well as simulated critical values and empirical power based on simulated critical values, for both types of tests having cointegration as the null hypothesis and no cointegration as the null hypothesis.

3.6.1. Monte Carlo Simulation Design for Simulated Critical Value

1. This study adopts the following procedure to calculate the critical values based on artificial data generations. Generate the data using the DGP A under the null hypothesis $H_0 : \rho = 1$.
2. Apply the tests having the same null hypothesis on the Data one by one and obtain the test statistic
3. Above two steps are repeated according to fix number of Monte Carlo Sample Size (MCSS) M.
4. Find the 5^{th} or 95^{th} percentile according to the nature of test whether it is left tail or right tail. However, for two tailed test, calculate 2.5th and 97.5th percentile as lower and upper critical value .

For tests having cointegration as null hypothesis, only the first step changes i.e. instead one has to generate the data following DGP-B.

3.6.2. Monte Carlo Simulation Design for Obtaining Empirical Size

1. . The empirical size i.e. probability of committing Type-II Error $(Rejecting\ H_0 \ | H_0\ isTrue)$ is calculated using the algorithm detailed as:Generate the data according to above stated DGP A under the null hypothesis $H_0 : \rho = 1$
2. Apply the under consideration cointegration tests on the data one by one and obtain the test statistic of each cointegration test.
3. According to above test statistic of each cointegration test, decision about the rejection or no rejection of null hypothesis is taken based on asymptotic critical value.
4. The steps mentioned in 1, 2 and 3 are carried out over and over again for a fixed number of times M (MCSS) and the number of rejections of null hypothesis are totalled.
5. Calculate the empirical size of test as $A.S. = (\frac{K}{M}) * 100$ where A.S. is the empirical size and K is the number of rejections. For the empirical size of the test having null of cointegration the above simulation design is followed except the first step because in this case data is generated using DGP B.

For stabilizing/controlling the empirical size of tests having the "null hypothesis of no cointegration/ null hypothesis of cointegration" same Monte Carlo Simulation design is adopted as above except the third step. In the third step simulated critical values are used instead of asymptotic ones.

3.6.3. Monte Carlo Simulation Design for Power Curves

Power is the proportion of rejections of the null hypothesis when data are artificially generated under a specific point alternative hypothesis. Thus, power of a test is defined as

$Power = P(\text{Reject } H_0/H_A \text{ is true})$. Power curve of a test is obtained by plotting the powers of test against the points of alternative hypotheses.

The following simulation procedure is adopted to obtain the power curves having null hypothesis of no cointegration.

1. Generate the data according to DGP-A under the point alternative hypothesis $0 \leq \rho < 1$.
2. Apply the tests having no cointegration as the null hypothesis on the data one by one and calculate the test statistic of each test.
3. Decision has been drawn regarding the null hypothesis by using simulated critical value of each test.
4. The steps 1, 2 and 3 are repeated for a fixed number of Monte Carlo Sample Size (MCSS) "M" and number of rejections are totaled.
5. Calculate the power of test as $\text{P.S.} = (\frac{K}{M}) * 100$ where P.S. is the power and K is the total number of rejections. To obtain power for each point alternative hypothesis, the steps 1 to 5 are repeated for each point in alternative space i.e. $\rho = \{0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1, 0\}$.
6. Plot the power of test against each point in alternate space to obtain the power curve of test.

For cointegration tests with cointegration as null hypothesis, the same above algorithm is used to estimate their power curves with one difference i.e. the artificial data generation is done using DGP-B.

Power envelop is obtained by plotting the power of PO test (DGP-A, DGP-B) against different alternatives. For the power envelop above same simulation design is adopted which is illustrated in the case of power curves of the test.

3.7. Evaluation Based on Bootstrap Empirical Power

After analyzing the characteristics of cointegration tests based on their size and power, when data are artificially generated, there is a need for the comparison of better performing tests based on some real economic theory. To serve this purpose, this study uses better performing cointegration tests on the real economic data based on Fisher hypothesis and evaluate the better performing tests using bootstrap empirical powers. Few studies have already investigated about the existence of Fisher hypothesis. These studies include but are not limited to Crowder (2003), Westerlund (2008), Örsal (2007), Toyoshima and Hamori (2011), and Badillo, Reverte et al. (2011). All these have found that Fisher hypothesis holds. In this study, we pick best three tests of cointegration and then their bootstrap empirical powers are assessed. This study has used two quarterly data sets of OECD countries of different time and cross sectional dimension.

Chapter 04: Empirical Size of Panel Cointegration Tests

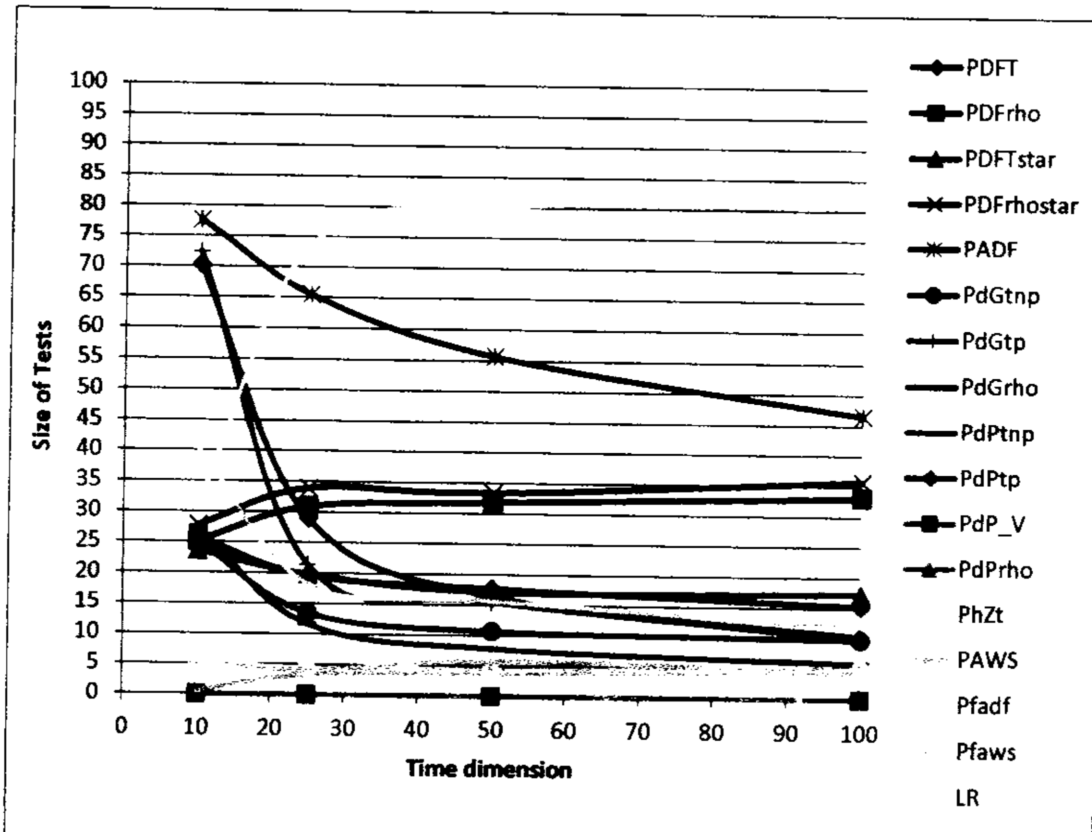
In this chapter of dissertation, size of under consideration panel cointegration tests are analyzed using asymptotic critical values. Some cointegration tests have null hypothesis of no cointegration whereas others have null of cointegration. In order to calculate the size of tests, data are generated under null hypothesis and the number of rejections are counted. Nominal size is considered 5% as this is taken as frequent. Empirical size of two different classes of panel cointegration tests are displayed and discussed from Figure 4.1 to Figure 4.8 which are also displayed in tables section in appendix from Table A-1 to Table A-8.

4.1. Empirical Size of Tests having Null Hypothesis of No Cointegration

In this study, twenty one tests are included which have null of no cointegration. These under consideration tests are from different categories that are, residual based (parametric, non-parametric) , error correction based, maximum likelihood based, average weighted symmetric based and fisher type tests. In this study, four different time (T= 10, 25, 50, 100) and four cross sectional dimensions (N = 2, 8, 16, 32).

In Figure 4.1; number of cross sections are fixed N=2, along x-axis time dimension is displayed and along y-axis size of test is displayed. From Figure 4.1; indicates dissimilar behavior of different tests. When time dimension is 10, three tests, namely PdPtp, PdGtp and PADF have size of 70%, 72% and 77% respectively. When time dimension is increased from 10 to 25 then size of these three tests is decreased. However, PdPtp and PdGtp have sharp rate of decreasing whereas, PADF does not depict sharp rate of decreasing. When time dimension is 25, the size of PdPtp, PdGtp and PADF tests is 21%, 29% and 65% respectively. Similarly, when time dimension is increased, the size of these three tests converges to nominal size (5%) with different rate of convergence. The size of LR test has totally different pattern as compared to the rest of tests, when time dimension is 10 the size of LR test is 31% but when the time dimension is increased, the size of LR

Figure 4.1: Empirical Size of Tests having Null of No Cointegration, N=02



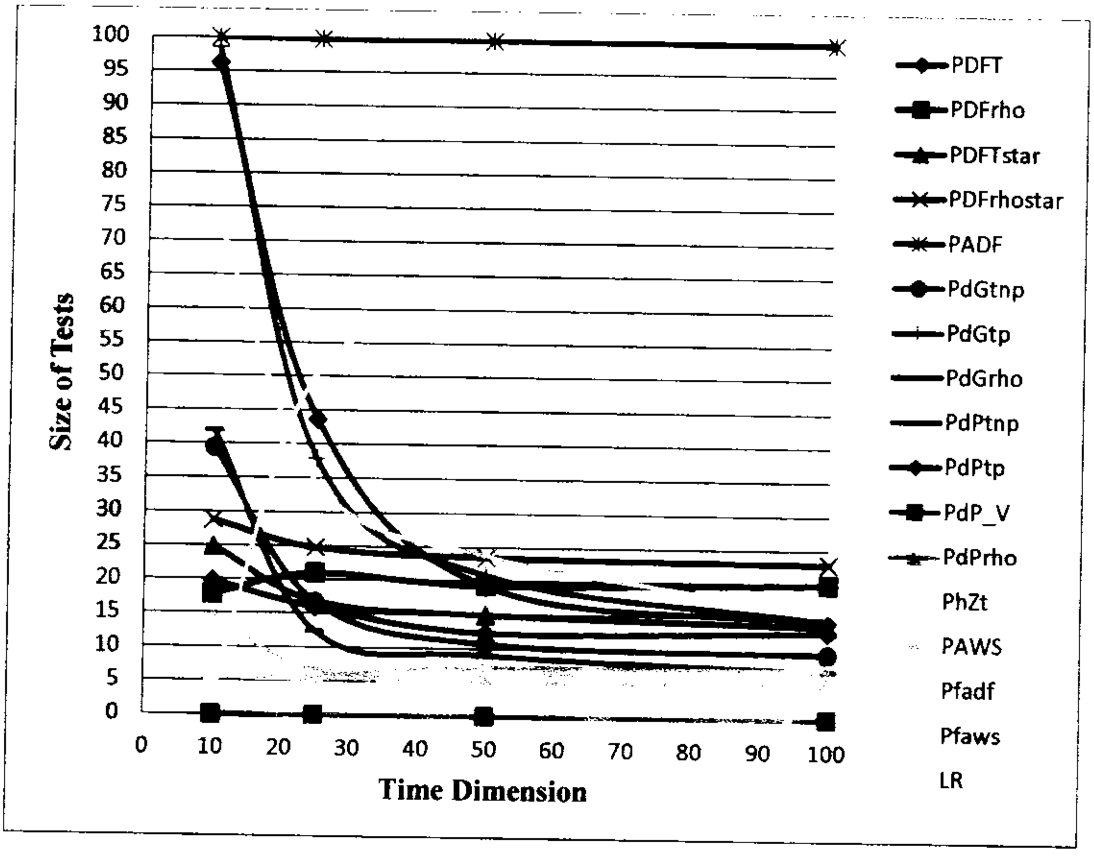
test is increased and hence size of LR test shows a divergent behavior. The size of PhZt is 40% when time dimension is 10 and shows a decreasing pattern when time dimension is increased. The size of six tests ranges from 20% to 30% when time dimension is 10, out of which four are Kao's tests such as, PDFT, PDFrho, PDFTstar, PDFrhostar and two are Pedroni's tests like, PdGtnp, PdPtnp. When time dimension is 25 then PDFrho and

PDFrhostar have increased as compared to time dimension 10 and size of tests 30% and 34% respectively. A slight change occurs in the size of these two tests when time dimension is 50 and 100. The size of PDFT and PDFTstar have diminished from 20% to 15% when time dimension increases from 25 to 100. The size of PdGtnp and PdPtnp converges to nominal size when time dimension is increased from 25 to 100. Only two tests out of these compared tests have nominal size 5%, which are fisher's tests such as, Pfadf, Pfaws for all of four time dimensions.

The size of PAWS starts from 0.6% and converges to nominal size 5% as time dimension increases from 10 to 100. Three tests which are PdPrho, PdP_V and PdGrho have not a single rejection throughout the time dimension. Comparative analysis in term of size means that these three tests have zero size at all time dimensions considered in the study. Briefly, size of three tests out of compared ones tests remains zero throughout all the four time dimensions which are PdGrho, PdPrho and PdP_V. Four tests that are, Pfadf, Paws, PAWS and PdPtnp converge to nominal size when time dimension gradually increased and size of four tests that are, PhZt, PdPtp, PdGtp and PdGtnp converge to 10% when time dimension is 100. While, size of remaining six tests is more than 15% when time dimension is even 100.

Figure 4.2, size of tests are displayed having cross sections fixed (N=8),where time dimension T is varied along x-axis and size of test is along y-axis. From the starting point of time dimension (T=10), three tests that are, PADF, PdGtp and PdPtp have size 100% approximately. Out of these three, PADF shows consistent size 100% at all four time dimensions; however, remaining two tests have convergent behavior towards nominal size 5% with increase in time dimension. When time dimension is 10, then size of PhZt test is 73%, whereby, time dimension is increased size of PhZt is decreased quickly, as it is observed when time dimension is 25 the size of PhZt becomes 34%. Similarly, when time dimension is further increased the size of PhZt test is decreased and it demonstrates a convergent pattern towards nominal size 5%. Maximum likelihood based LR test, there is a positive relationship between size and time dimension, that is, when time dimension is increased, size of LR test is also amplified. When time dimension is 10, size of LR test is 3.6% which is considered as stable around nominal size (5%) but when time dimension is increased from 10 to 25, then size of LR test becomes 50%.

Figure 4.2: Empirical Size of Tests having Null of No Cointegration, N=08



Similarly, when the time dimension further increased the size of LR test also increased, so the pattern of size of LR test is divergent with increase in time dimension. Three tests, that are, PdPrho, PdGrho and PdP_V out of these compared tests, behave consistently thoroughly for time dimension as showed in Figure 4.1. These three tests are under sized and even have not a single rejection, hence have zero size. Two tests are consistent as time dimension varied from 10 to100, throughout the time dimension the size of these tests is around nominal size, and these are Pfadf and Pfaws tests.

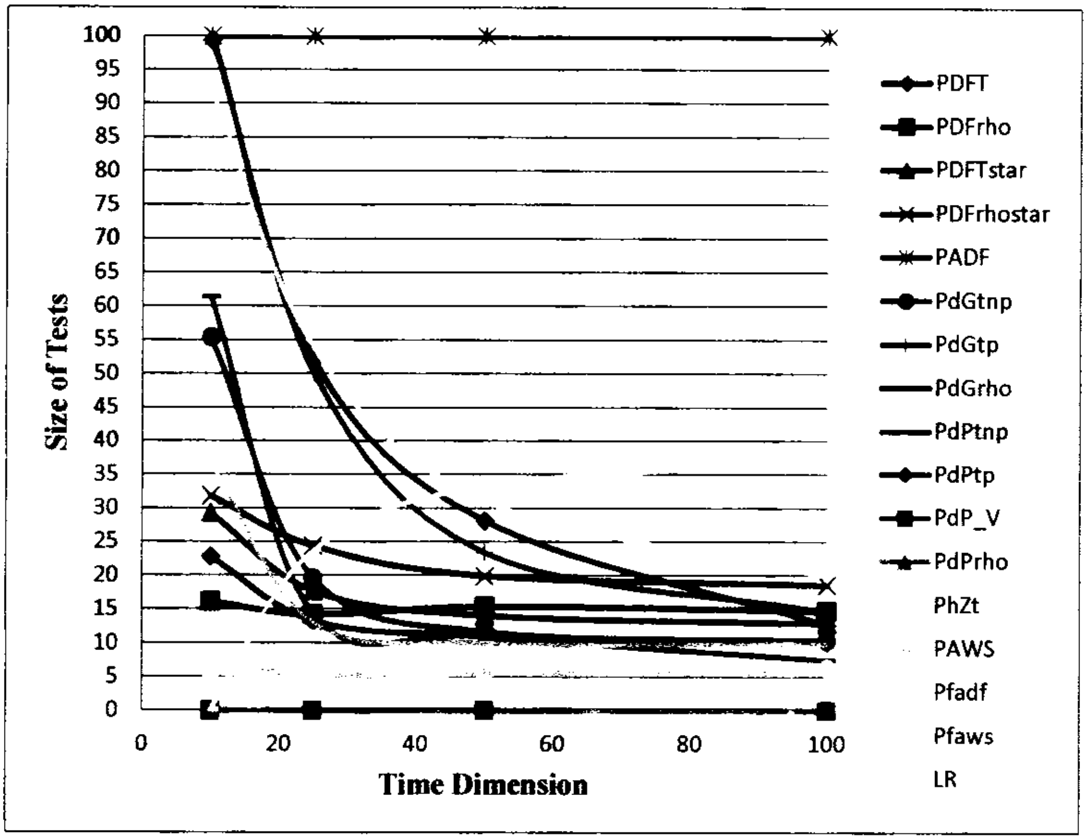
Average weighted symmetric based test PAWS test is also convergent towards nominal size (5%) as time dimension increased from 10 to 100. When time dimension is 10 its size is 14% but when time dimension is 25 then its size becomes 6% which is very close to nominal size, and hence showed convergent behavior. PDFrho test behaves better as compared to Figure 4.1 while numbers of cross sections are fixed (N=8), it shows non convergent behavior. In this case, when time dimension is 10, size becomes 17% but when time dimension is increased, its size is also increased and become 21%. When time dimension is 100, its size becomes 19%. However, its size does not converge to nominal size (5%). The remaining five tests exhibit convergent behavior when time dimension is increased. Figure 4.2; it is observed that when time dimension is 10 the size of PdGtnp and PdPtnp is round about 40% and the size of PDFT, PDFTstar and PDFrhostar is 19%, 24% and 28% respectively.

When time dimension is increased from 10 to 25 then the size of PdGtnp and PdPtnp becomes 16% and 12% respectively and the size of PDFT, PDFTstar and PDFrhostar becomes 15%, 16% and 24% respectively, which is relatively lesser decrease as compared to PdGtnp and PdPtnp. Hence these five tests show convergent behavior. In general, it can be summarized that size of four tests that are, PdPrho, PdGrho, PdP_V and PADF lie on extreme points 0% and 100% throughout time dimension, in which three tests PdPrho, PdGrho, PdP_V are at one extreme (0%) and PADF lies on other extreme (100%). Size of LR tests starts from nominal size and increases as time dimension varies. Size of two tests that are, Pfadf and Pfaws remains around nominal size throughout the time dimensions. All other tests except PDFrho pursue the convergent pattern but rate of convergence of each test is variable.

In Figure 4.3, the number of cross section is fixed N=16, along x-axis time dimension is labeled and along y-axis size of tests is labeled. In this case few tests behave like earlier as in Figure 4.2. When time dimension is 10, the size of PdPrho, PdGrho, and PdP_V is zero and size of these does not change even though when the time dimension varies. So the pattern of size of these tests is totally same with previous case where number of cross section was 8. It means that cross sectional variation does not have an impact on these tests. The behavior of PADF test is also same as in the previous case,, that is, the size of PADF is 100% throughout the time dimension in this case also. Size of two tests, that is, PdGtp and PdPtp is 100% approximately and PhZt has 91% of size when time dimension is 10. If compare current stage of these three tests with the previous case when time dimension was 10 and cross sectional size was 8, two tests, PdGtp and PdPtp have approximately same size as they had in previous case but the third one, PhZt test faces an increase in size.

In Figure 4.3, as increase the time dimension from 10 to 25 then size of these three tests become 50% and in same manner, when time dimension is further increased the size of these three tests converge to nominal size 5%, however, each test has different rate of convergence. The size pattern of LR test is similar as shown in Figure 4.2. In Figure 4.3 when time dimension is 10, size of LR test is zero but when time dimension increases size of LR test increases. When time dimension is 100 size of LR test becomes 36%, size of LR test has increasing pattern with respect to time dimension and hence does not converge to nominal size 5%. In Figure 4.3, there are only two tests, that is, Pfadf and Pfaws, exist which have approximately nominal size throughout the time dimension.

Figure 4.3: Empirical Size of Tests having Null of No Cointegration, N=16



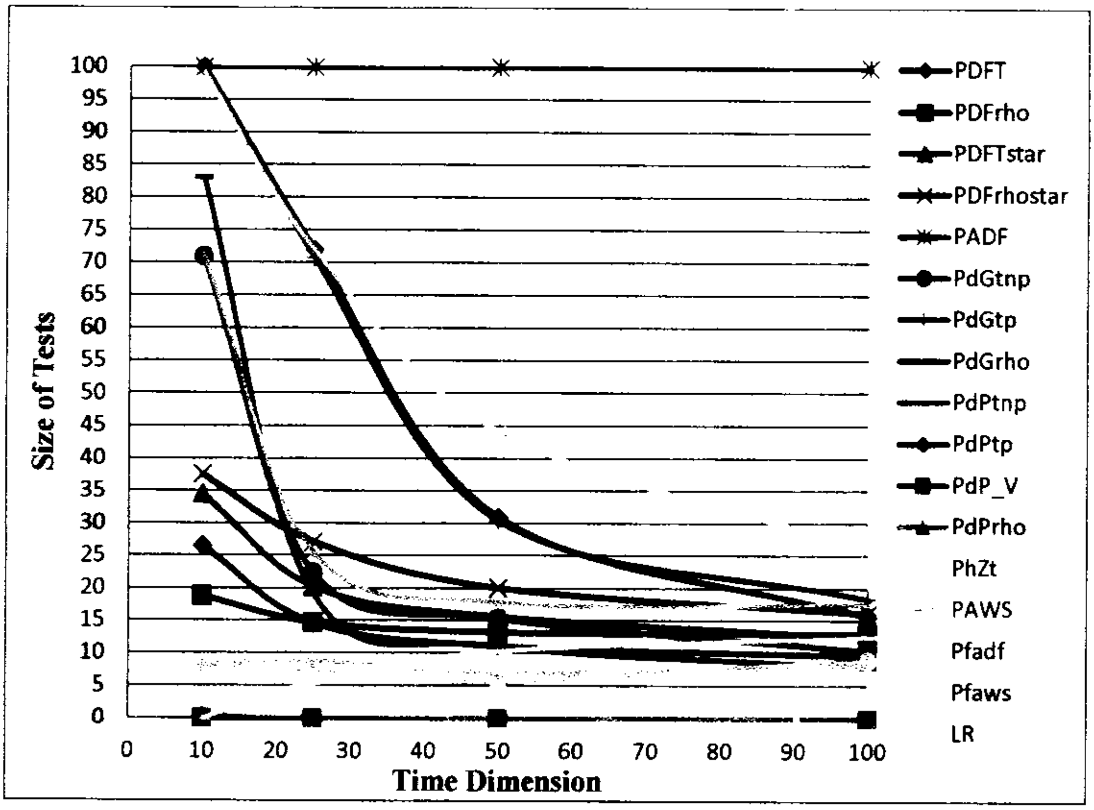
When size behavior of these two tests are compared with previous Figure 4.2 then a similar results are obtained. It means that cross sectional dimension do not have much impact on the size of these two tests. The pattern of size of PDFrho test is changed as compared to Figure 4.2. In Figure 4.3, size of PDFrho has decreasing pattern. When time dimension is 10 its size is 16% but when time dimension is increased, its size becomes decreased and shows convergence towards nominal size 5%. It can be easily observed from Figure 4.3 that the remaining six tests have convergent behavior with different rate of convergence. When time dimension is 10, then size of PDFT, PDFTstar, PDFrhostar and PAWS are 22%, 29%, 31% and 36% respectively and size of Pedroni non parametric tests, that is, PdGtnp and PdPtnp have sizes of 55% and 61% respectively. However, when time dimension increases, the size of these six tests decreases with different magnitude.

While summarizing Figure 4.3 may be stated as; two tests have size around nominal size throughout the time dimension which is fisher's tests, that is, Pfadf and Pfaws. Four tests remain on extreme throughout the time dimension, in which three tests, that is, PdGrho, PdPrho and PdP_V remain on lower extreme and the fourth test, that is, PADF remains on upper extreme. Maximum likelihood based LR test shows clear cut increasing pattern throughout the time dimension. Size of three tests, that is, PdPtp, PdGtp and PhZt start from approximately upper extreme when time dimension starts from 10 but converge towards nominal size of 5% with different convergence rate as time dimension increases. Size of five tests start between 15% to 35% and then converge towards nominal size and the size of remaining two tests start between 55% to 60% and then converge towards nominal size with different rate of convergence.

In Figure 4.4, size of test is displayed along y-axis and time dimension along x-axis. In Figure 4.4, 32 numbers of cross sections have been taken, which is fixed throughout the time dimension. Cumulatively, four tests have extreme size of zero and 100%, the size of PdGrho, PdPrho and PdP_V are zero when time dimension is 10 and do not change when time dimension increases and hence these three have zero size throughout the time dimension like as in previous case in Figure 4.3. While the size of fourth test, which is

PADF, lies on the upper extreme, that is, 100% throughout the time dimension. The size of three tests which are PdGtp, PdPtp and PhZt becomes 100% when time dimension is 10, however, size of these three tests decreases when time dimension is increased just like when time dimension is 25 then size of these three tests becomes 70% approximately. Hence, the size of these three tests converges towards nominal size 5% when time dimension is increased. Two tests out of these three tests have equal rate of convergence approximately. The size of two tests, that is, PdGtnp, PAWS are around 70% and size of PdPtnp is 83% when time dimension is 10, but the size of these three tests decreases sharply and becomes approximately 20% when time dimension is 25.

Figure 4.4: Empirical Size of Tests having Null of No Cointegration, N=32



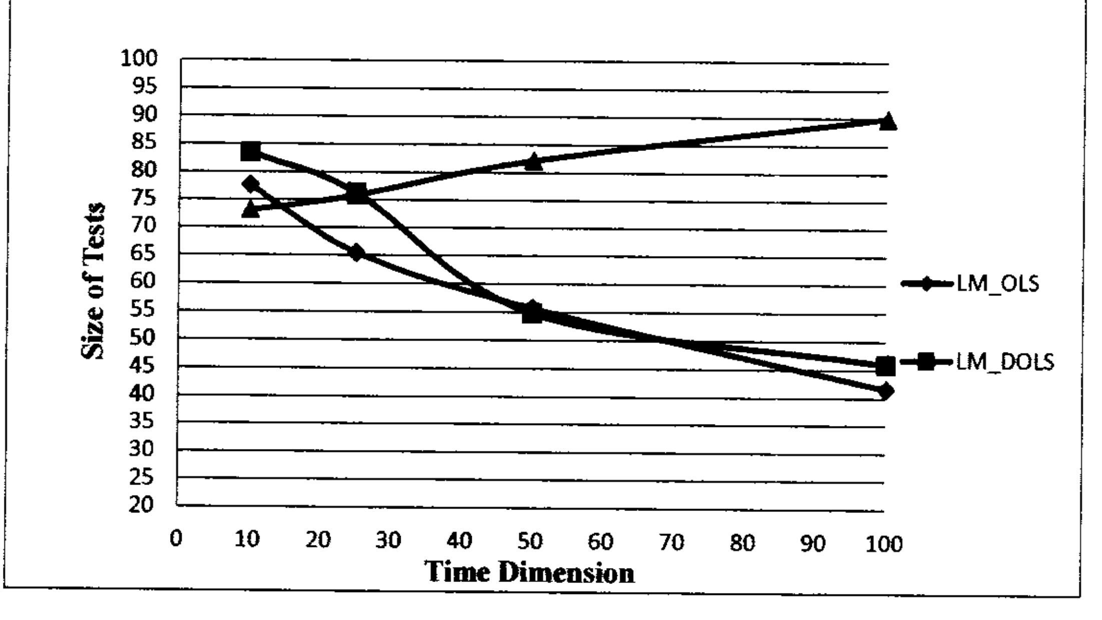
Similarly, when time dimension has increased the size of these three tests have decreased, and hence converge to nominal size with different rate of convergence. The size of Pfadf and Pfaws is around 7% and 8% respectively when time dimension is 10, the size of Pfadf is exactly 5% when time dimension is 50, however, with varying time dimension, the size of these two tests fluctuate from 7% to 9%. The size pattern of LR tests is slightly different from the previous pattern, when time dimension is 10, size of LR test is zero but when time dimension becomes 25, its size converges to nominal size of 5% and when time dimension is increased to 50, the size of LR test increases and becomes 10%. In Figure 4.4, increasing pattern of size of LR test is slow as compared to Figure 4.3. In Figure 4.4, the size of remaining four tests (PDFT, PDFTstar, PDFrho, and PDFrhostar) behave moderately, when time dimension is 10 then the size of these four tests lies between 20% to 30%, but when time dimension is increased the size of these four tests slightly decrease. Moreover, the size of these four tests lies in the range of 10% to 20% for three time dimensions, that are, 25, 50 and 100.

To summarize, Figure 4.4, in which three tests, that are, PdGrho, PdPrho and PdP_V have zero size and PADF have 100% size throughout the time dimension. Three tests PdGtp, PdPtp, and PhZt have size 100% when time dimension is 10, with increase in time dimension to 100 these tests have size in range 25% to 15%. Size of LR test lies in the range of zero to 10%, when time dimension increases from 10 to 100. Size of all other tests lies in the range of 10% to 20% when time dimension is 100 but when time dimension is different than 100 then size lies in multiple range.

4.2. Empirical Size of Tests having Null Hypothesis of Cointegration

In this section, the empirical size of tests having null of cointegration is discussed. In Figure 4.5, along x-axis time dimension along y-axis size of test are displayed respectively and number of cross section is fixed N=2. The size of three tests, OLS based (LM-OLS), Dynamic OLS based (LM-DOLS), and Fully Modified OLS based (LM-FMOLS) tests are displayed. When time dimension is 10, the size of OLS based test (LM-OLS) and Dynamic OLS based (LM-DOLS) tests is around 80% while size of Fully Modified OLS based test (LM-FMOLS) test is around 73%. Figure 4.5 shows that the size of LM-OLS and LM-DOLS tests have decreasing pattern whereas the size of LM-FMOLS has increasing pattern.

Figure 4.5: Empirical Size of Tests having Null of Cointegration, N=02

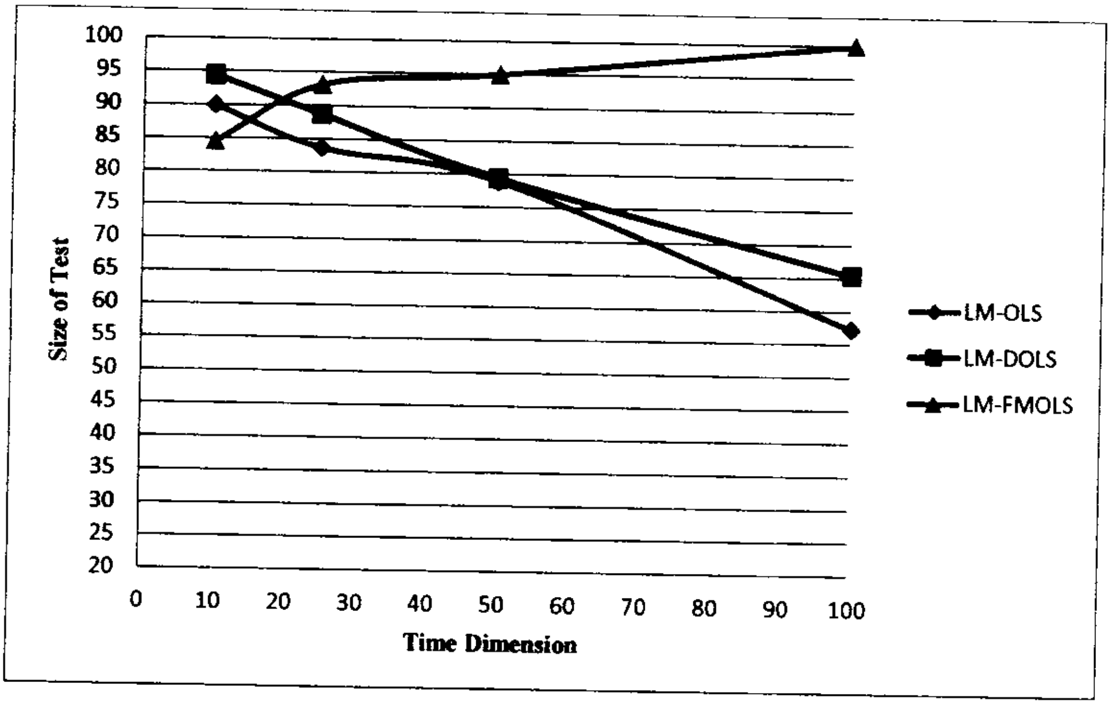


When the time dimension is 25, the size of LM-OLS and LM_DOLS tests are 65% and 76% respectively, whereas the size of LM-FMOLS test is 75%. As time dimension increases to100 the size of LM-FMOLS test becomes 90%. Similarly, the size of LM-OLS and LM-DOLS becomes 41% and 46% respectively.

Figure 4.5 summarized that these three tests do not converge to nominal size when time dimension becomes large. Although, two tests have decreasing pattern while third test has increasing pattern which clearly diverges when time dimension is increased. The size of LM-OLS and LM-DOLS tests have decreasing pattern and lies in range (80% to 40%) for time dimension range (10 to 100). Similarly, LM-FMOLS test lies in the range (73% to 90%) for time dimension range (10 to 100).

In Figure 4.6, number of cross sections is 8 which are fixed and time dimension is varied from 10, 25, 50 and 100. In Figure 4 6, time dimension and size are labeled along x-axis y-axis, respectively. When time dimension is 10, size of LM-OLS and LM-DOLS tests is 90% and 94% respectively whereas size of LM-FMOLS test is 84%. When time dimension increases size of LM-OLS and LM-DOLS tests slightly decreases. When time dimension is 25, size of LM-OLS and LM-DOLS tests is 85 % around and size of LM-FMOLS test is around 94%. The size pattern of these three tests are not much more different from the previous Figure 4.5, where size of LM-OLS and LM-DOLS tests has slightly decreasing pattern and size of LM-FMOLS has increasing pattern and attains the maximum value 100%. Empirical size of these three tests are much more away from the nominal size. These three tests have over rejection problem when asymptotic value are used.

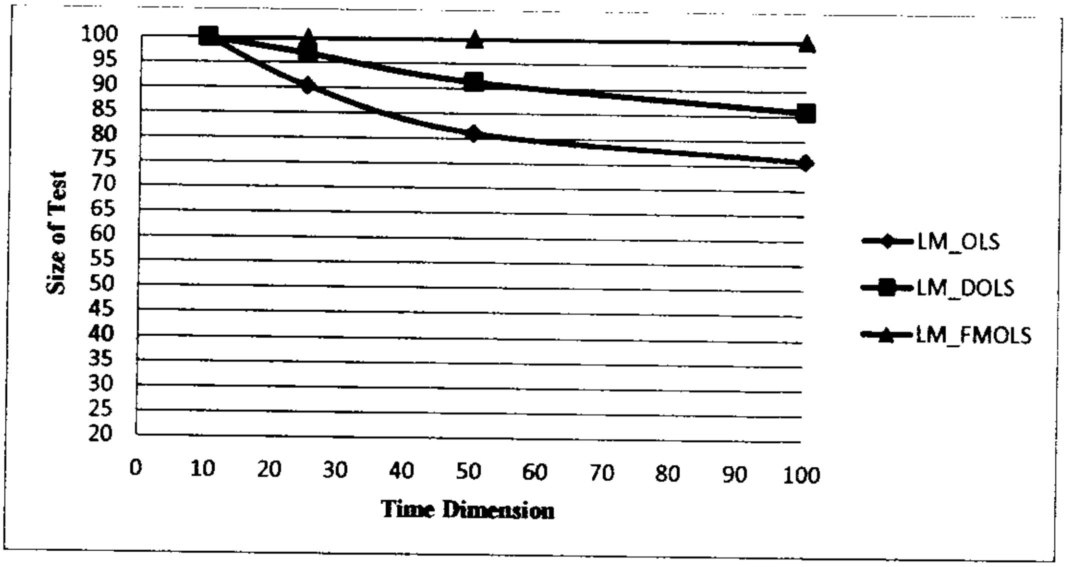
Figure 4.6: Empirical Size of Tests having Null of Cointegration, N=08



The size of LM-OLS and LM-DOLS tests lie in the range (94% to 60%) for all time dimensions, (10 to 100), while the size of third tests LM-FMOLS lies in the range (84% to 100%) for all time dimensions. Hence, it is concluded from Figure 4.6 that the size of these three tests have ambiguous behavior, although the size of two tests have slightly decreasing pattern but not much useful while the size of the third test has divergent pattern for the given time dimension.

Figure 4.7 presents the results when number of cross sections is 16. Size of three tests (LM-OLS, LM-DOLS, and LM-FMOLS) have been displayed in which two tests have little bit decreasing pattern but the size of third test lies on the upper extreme throughout the time dimension. When time dimension is 10, size of LM-OLS and LM-DOLS tests are at upper extreme (100%.) Similarly, the size of LM-FMOLS test is also on upper extreme. When time dimension increases to 25 then size of LM-OLS and LM-DOLS tests slightly decreases but still lies around upper extreme.

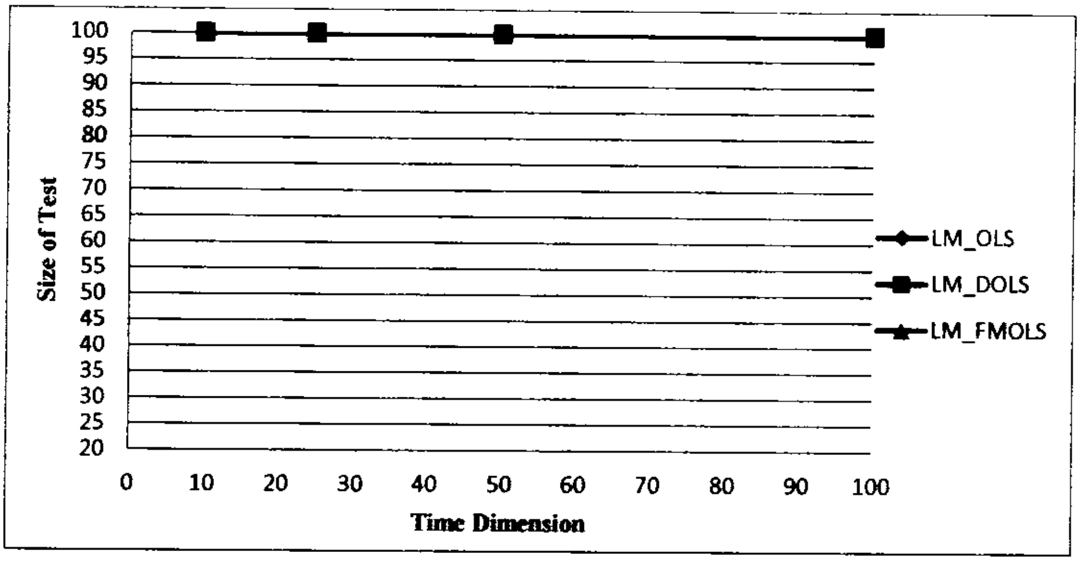
Figure 4.7: Empirical Size of Tests having Null of Cointegration, N=16



Whereas, size of LM-FMOLS test has remained fixed on the upper extreme. Overall, from the analysis of Figure 4.7, it is summarized that size of LM-OLS and LM-DOLS tests lie in the range (100% to 80%) throughout the time dimensions. Whereas, size of third test, LM-FMOLS, exists on the upper extreme throughout the time dimension. Hence, all three tests have over rejection in other words over size throughout the time dimension.

In Figure 4.8, size of three tests (LM-OLS, LM-DOLS and LM-FMOLS) has been analyzed by taking 32 numbers of cross sections. It is observe that these three tests have exactly same size pattern throughout the time dimension. The size of these three tests is at extreme upper bound throughout the time dimension.

Figure 4.8: Empirical Size of Tests having Null of Cointegration, N=32



When time dimension increases from 10 to 25, then size of LM-OLS, LM-DOLS and LM-FMOLS tests remain at 100% and does not change even when time dimension is changed. Hence, it is concluded from Figure 4.8 that these three tests have over rejection in all cases, in other words these tests are over sized throughout the dimension.

4.3. Conclusion

The above two sections of this chapter describe the results of both classes of tests, in which one section describes results of empirical size of tests having null of no cointegration while other section describes the results of empirical size of tests having null of cointegration. Firstly, Figure 4.1 to Figure 4.4 explains the results of empirical size of tests having null hypothesis of no cointegration. The aim of the comparison of empirical size of these tests is to analyze the asymptotic properties of these under consideration tests with respect to time and cross-sectional dimensions variations. In each figure number of cross sections is fixed and time dimension is varied.

In the first section, the empirical size of six types of tests, such as, parametric residual based, non parametric residual based, maximum likelihood based, weighted symmetric based, and fisher type tests have been displayed from Figure 4.1 to Figure 4.4 when asymptotic critical values are used. From Figure 4.1 to Figure 4.4, it is evident that residual based parametric tests show mix behavior, that is, either increasing, decreasing or constant when time and cross sectional dimension are varied. From eleven parametric residual based tests, only three tests, that are, PdGrho, PdPrho, and PdP_V have constant size of 0% at all time and cross sectional dimensions. Whereas, another parametric residual based test, that is, PADF has also constant behavior of size 100% for N=8, 16, 32. However, when N=02 then the residual based test PADF has a decreasing pattern of size with increase in time dimensions.

The remaining seven parametric residual based tests,(PdGtp, PdPtp, PDFT, PDFTstar, PDFrho, PDFrhostar, and PhZt) have decreasing pattern of empirical size when time dimensions are increased. From these seven parametric residual based tests, three tests, that are, PdGtp, PdPtp, and PhZt have sharp rate of decrease as time dimension increases as compared to other four tests, that are, PDFT, PDFTstar, PDFrho, and PDFrhostar. However, none of these seven tests have achieved the nominal size of 5% throughout the given time and cross sectional dimensions. Nevertheless, these seven parametric residual based tests have increasing pattern of size as cross sectional dimension increases. Nonetheless, both the non-parametric residual based tests have a decreasing pattern of size as the time dimension increases. Very interestingly, only one of them (PdPtnp) test achieves nominal size of 5% at time dimension 100 or more.

Maximum likelihood based test (LR) shows a divergent behavior as it has an increasing pattern of size with increase in time and cross sectional dimension. Two fisher type test, (Pfadf and Pfaws) have their size around nominal size of 5% at all time and cross sectional dimensions. While, the single weighted symmetric type of test (PAWS) shows a unique behavior as it has decreasing pattern of size with increase in time dimension for N=08, 16, 32. However, it has increasing pattern with increase in time dimension for N=02. Although it achieves nominal size of 5% at time dimension of 100 for all given cross section.

In the second section, the empirical size of tests is described having the null of cointegration . In this section, there are only three type of tests, that are, OLS based (LM-OLS), Dynamic OLS based (LM-DOLS), and Fully Modified OLS based (LM-FMOLS). The size of these three type of tests have been analyzed for different time and cross sectional dimension, which are portrayed in Figure 4.5 to Figure 4.8. Fully Modified OLS based test, (LM-FMOLS) has an increasing pattern of size as the time dimension increases for N = 02, 08. However, for N = 16, 32 this test has maximum size of 100% at all time dimensions. The rest of two type of tests, that are, LM-OLS and LM-DOLS have decreasing pattern of size with increasing time dimension for N = 02, 08, 16. However, these two tests show constant behavior of size 100% at all time dimension for N=32.

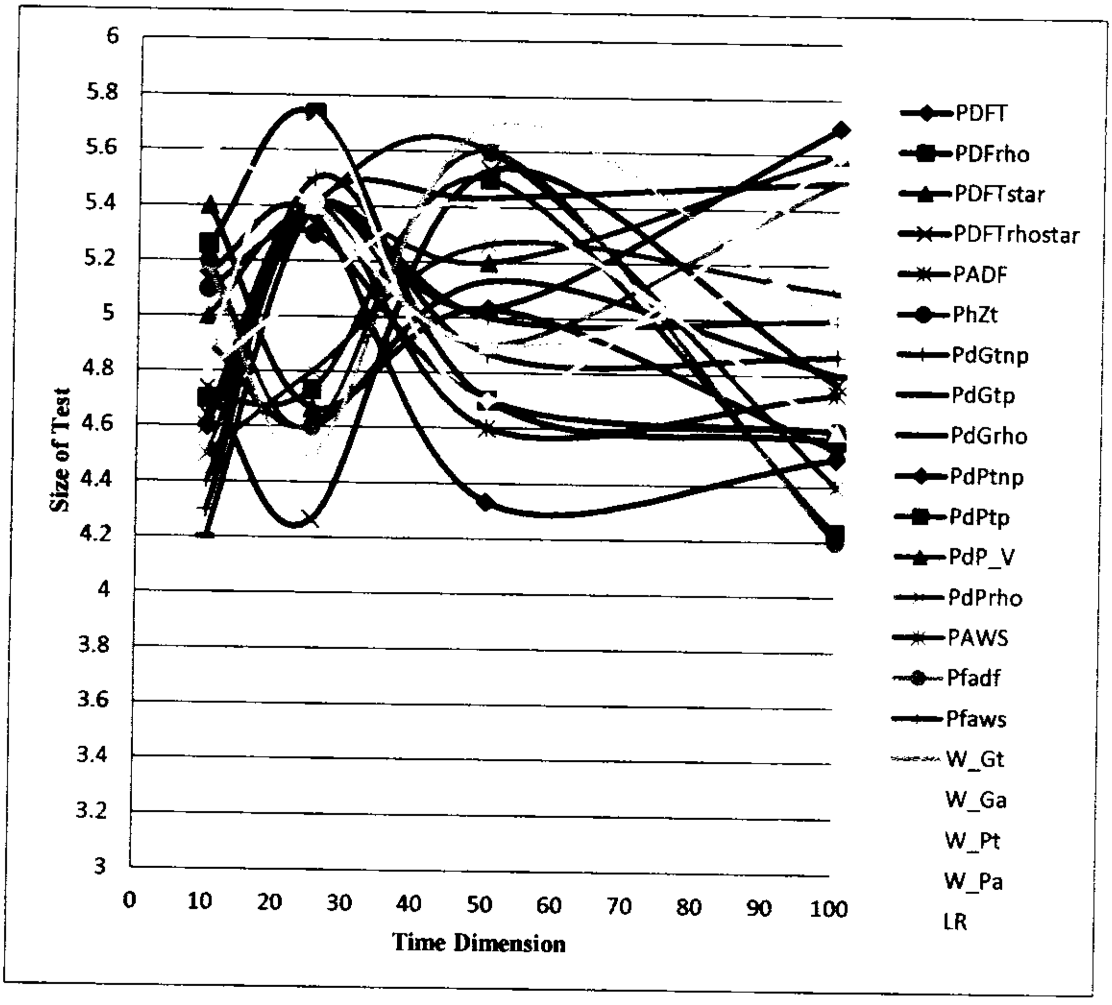
Chapter 05: Controlling the Size of Tests

In this chapter, size of tests are discussed by using simulated critical values. As mentioned earlier, this study has considered two classes of tests, one class of tests have the null of no cointegration while other class of tests have null of cointegration. In the previous chapter, it is observed that both classes of tests do not have empirical size around the nominal size when asymptotic critical values are used for the given time and cross sectional dimension. In order to stabilize the size, simulated critical values are used for both classes of tests under consideration.

5.1. Stabilizing the Size of Tests having Null of No Cointegration

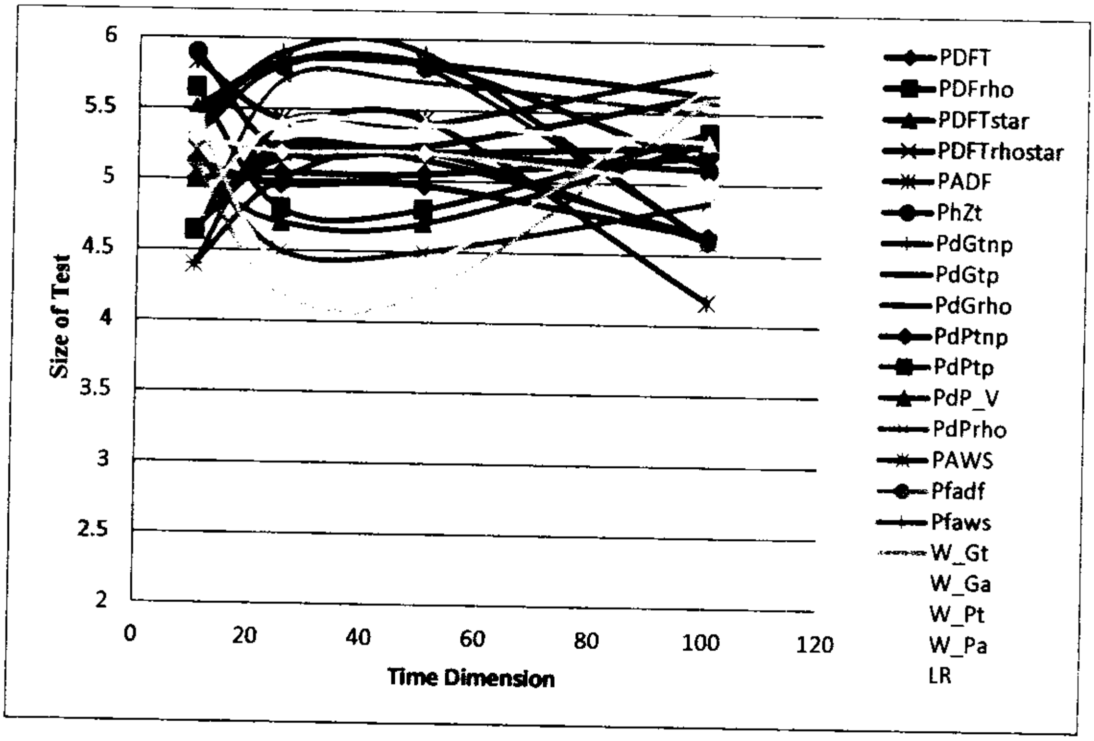
In this section, size of twenty-one tests are analyzed having null of no cointegration by using simulated critical values. In Figure 5.1 along x-axis time dimension has been displayed which is 10, 25, and 50 and 100 while, along y-axis the size of test has been displayed and the number of cross sections is 2 which is fixed in this case. Now the size of each test in Figure 5.1 is around the nominal size, size of all tests lie in the range between 4% to 6% for all given time dimension, hence now size of all tests are stabilized around 5%. Hence, size of tests has been stabilized when simulated critical values are used while using the asymptotic critical values size of tests was not stable.

Figure 5.2, Simulated Size of Tests having Null of No Cointegration, N=08



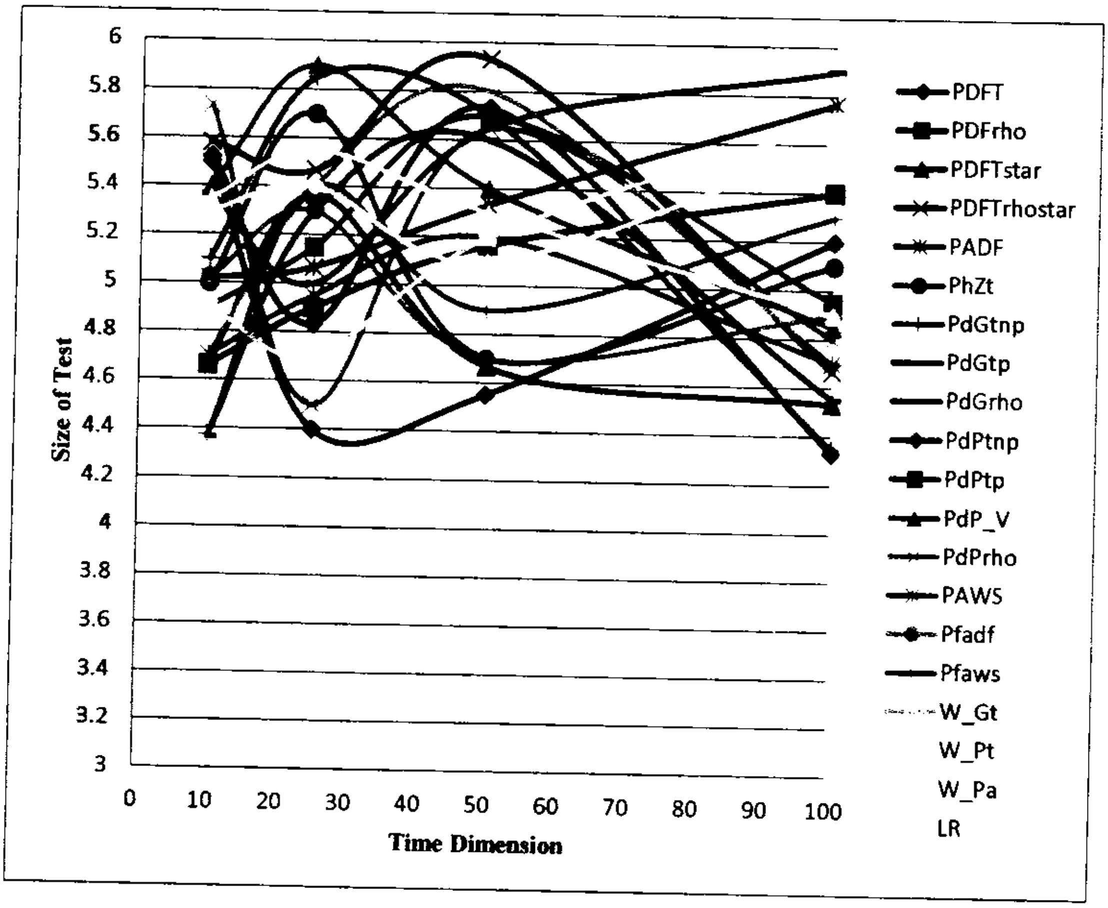
In the Figure 5.3, size of 21 tests have been displayed which have null of no cointegration. In this figure, time dimension has been labeled along x-axis and size of test has been labeled along y-axis for the number of cross section 16. It is observed that size of all tests lie between the range of 4% to 6% because simulated critical values are used instead of asymptotic critical values. Now, the size of each test is stable which is approximately around the nominal size. In the previous chapter, when asymptotic critical values were used, most of the tests have not stabilized the size of tests. Hence, in this study, simulated critical values are used for stabilizing the size of tests.

Figure 5.3, Simulated Size of Tests having Null of No Cointegration, N=16



The following Figure 5.4, size of 21 tests have been showed in which time dimension is labeled along x-axis and the size of test has been displayed along y-axis for N=32. Here, simulated critical values are used because when asymptotic values was used then size of tests was not stable as have seen in the previous chapter for the same time dimension and number of cross section. Here, the size of each test remains between 4% to 6% throughout the time dimension and cross section. Hence, size of each test is stable and remains around nominal size of 5% when simulated critical values are used.

Figure 5.4, Simulated Size of Tests having Null of No Cointegration, N=32

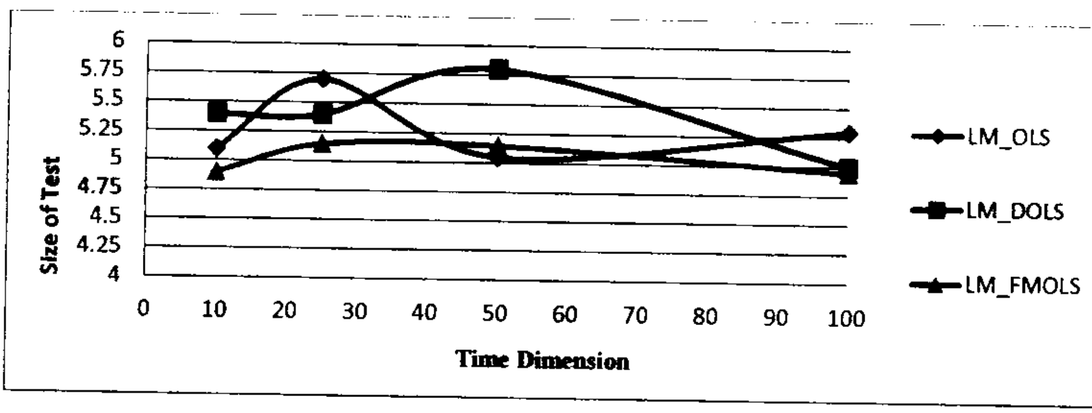


5.2. Stabilizing the Size of Tests having Null of Cointegration

In this section, the size of tests that have null of cointegration is analyzed. Here, simulated critical values are used instead of using asymptotic critical values as have seen the distortive size behavior of tests in the previous chapter. In this section, three tests having the null of cointegration are considered which are LM-OLS, LM-DOLS, and LM-FMOLS. All of these three tests are over sized for all time dimensions and number of cross sections when asymptotic critical values were used. Here, simulated critical values for time dimension 10, 25, 50 and 100 and cross sectional size 2, 8, 16 and 32. The cross section is fixed in each figure and varying the time dimension to analyze the size of three tests.

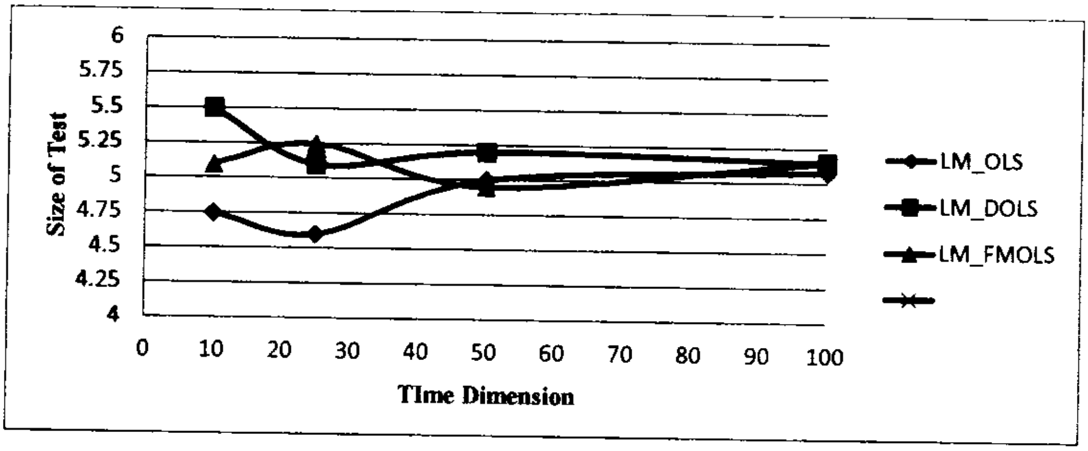
In Figure 5.5, size of three tests LM-OLS, LM-DOLS, and LM-FMOLS are displayed. The time dimension which is 10, 25, 50, and 100 have been displayed along x-axis and the size of tests have been displayed along y-axis for the number of cross section two. The size of each test remains in the bound 4% to 6%, which shows that the size of these tests are stable now and approximately around the nominal size.

Figure 5.5, Simulated Size of Tests having Null of Cointegration, N=02



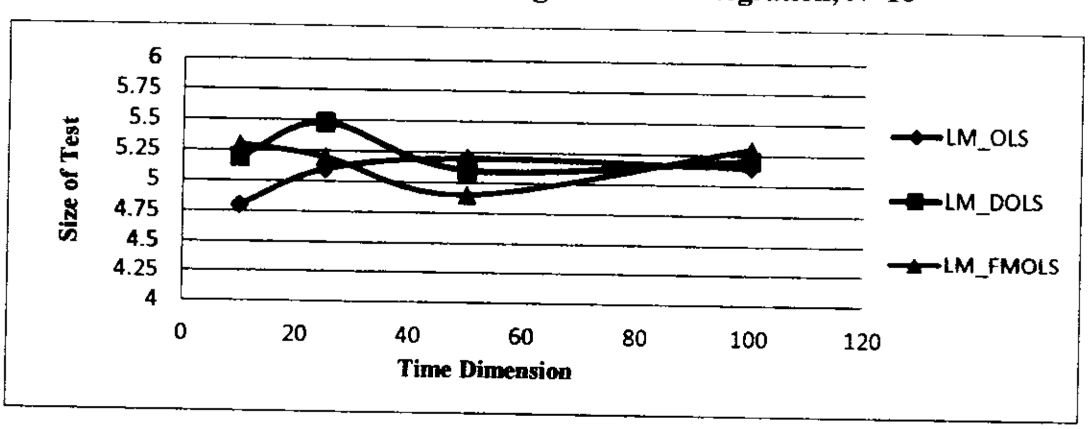
In Figure 5.6, size pattern of tests under consideration have been displayed. The number of cross sections in this case is eight which is fixed and time dimension varies. Here, time is labeled along x-axis whereas the size of test is labeled along y-axis. It is observed Figure 5.6 when simulated critical values are used size of each test remains between 4% to 6% throughout the time dimension, whereas in case of asymptotic critical value all three tests were oversized. Hence, sizes of all three under consideration tests are stable around 5% under simulated critical values.

Figure 5.6, Simulated Size of Tests having Null of Cointegration, N=08



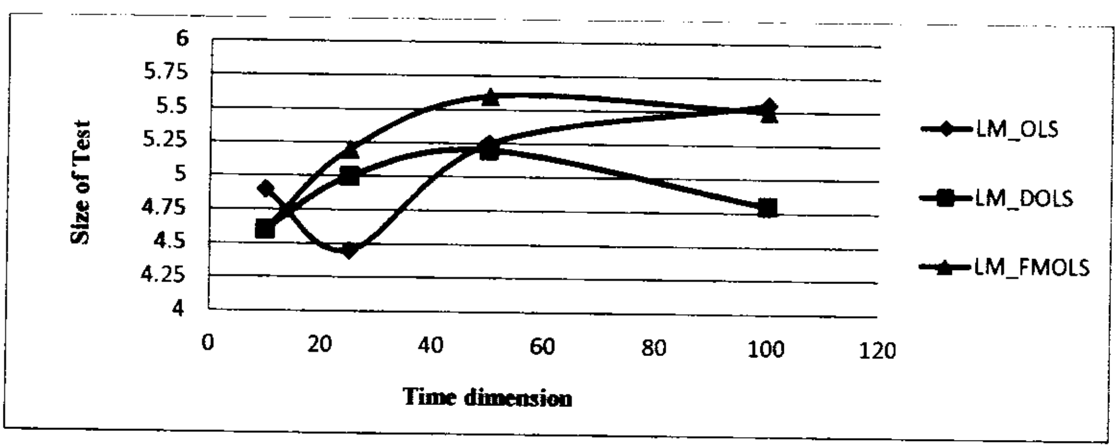
In Figure 5.7, the time dimension is displayed along x-axis and the size of test is displayed along y-axis whereas the number of cross sections is 16. In Figure 5.7, simulated critical values are used instead of using asymptotic critical values. As, it has seen in the previous chapter that each test was oversized throughout the time dimension when asymptotic critical values were used. But, in this case when simulated critical values are used the size of each test remains between 4% to 6% throughout the time dimension and size of all three tests lies around the nominal size 5%. Hence now size of all three tests are stabled which is clearly evident from Figure 5.7.

Figure 5.7, Simulated Size of Tests having Null of Cointegration, N=16



The following Figure 5.8 shows the simulated size of three tests having the null of cointegration. In this figure, time dimension is displayed along x-axis and size of test is displayed along y-axis, whereas the number of cross section is 32. It is observed from the following figure that size of all three test lies between 4% to 6% throughout the time dimension which shows that size of tests is stable. Whereas, size of these tests was oversized for all time dimensions when asymptotic critical values are used. Hence, now the size of each test lies around the nominal size throughout the time dimension.

Figure 5.8, Simulated Size of Tests having Null of Cointegration, N=32



5.3. Conclusion

From the overall conclusion of above two sections, in which one section has briefly discussed the size of those tests which have null of no cointegration whereas other section has discussed the size of those tests which have null of cointegration by taking simulated critical values. In both section of this chapter the time dimension is 10, 25, 50, and 100 which is varied in each figure whereas the cross sections are varied from figure to figure which are 2, 8, 16 and 32. From Figure 5.1 to Figure 5.8, it is evident the size of both type of tests having null of no cointegration and null of cointegration have stable size when simulated critical values are used. Hence, from Figure 5.1 to Figure 5.8, it can be observed that the size of all under consideration tests lie in the range 4% to 6% for all given time and cross sectional dimension and converge to the nominal size of 5%. So, simulated critical values are used for power comparison.

Chapter06: Power Comparison

As the two characteristics of the test are very important for assessing the test in which one is the size of the test and other is the power of test. It is commonly accepted concept that the tests having equivalent size level should be comparable on basis of power. In this chapter, the power of different panel cointegration tests having null of no cointegration and null of cointegration are discussed. Furthermore in this power comparison, this study has considered both first generation and second generation tests.

As from the previous chapter, it is observed that size of both classes of tests have been stable by using simulated critical values. Now, to compare the power of tests using stringency criterion which is discussed by (Zaman (1996)) and also (Zaman, Zaman et al. (2017)). According to Zaman, Zaman et al. (2017), if the best tests have stringency above 50% then there is need to search an alternative methods for testing whereas if the best tests have the stringency between 5% to 10% then there is no need to search furthermore.

In the first section of this chapter, a Monte Carlo power comparison of 21 tests is briefly discussed which have null of no cointegration with figures and tables. In second section of the chapter the Monte Carlo power comparison of three test having the null of cointegration is discussed and in the end of the chapter the overall summary and conclusion of the both sections are described. The time dimension T=10, 25, 50, 100 and cross section dimension N=02, 08, 16, 32 are taken to make a power comparison. A Monte Carlo sample size of 50,000 is used for all tests except PAWS, Pfadf, and Pfaws, for these three tests a Monte Carlo sample size of 10,000 is taken. In all tables, tests are categorized into six categories from A to F according to their maximum shortcomings values, the test having maximum shortcomings less than 10% is assigned in category A and test having the maximum shortcomings between 10% to 20% ($10\% \leq MS < 20\%$, where MS denote the maximum shortcomings of the test) is assigned in category B, test having maximum shortcomings between 20% to 30% ($20\% \leq MS < 30\%$) is assigned in category C, test having maximum shortcomings between 30% to 40% ($30\% \leq MS < 40\%$) is assigned in category D, test having maximum shortcomings between 40% to 50% ($40\% \leq MS < 50\%$) is assigned in category E, and test having the shortcoming greater than 50% is assigned in category F. Maximum shortcomings of two classes of tests null of no cointegration and null of cointegration are displayed section appendix from Tables A-6 to Tables A-13.

6.1. Power of Tests having Null of No Cointegration

To describe the power analysis of panel cointegration tests having null hypothesis of no cointegration, all tests are classified into three categories according to their maximum shortcomings; best performing, mediocre performing, and worst performing tests. After finding the power of cointegration tests at all time and cross sectional dimensions, it is observed that majority of the tests perform poorly at low time and cross sectional dimensions. Moreover, these tests have approximately same power of 100% at high time and cross sectional dimensions. So, in order to classify all tests in three categories, the appropriate time and cross sectional dimension of 50 and 8 respectively is taken as these dimensions are neither very low nor very high.

According to this time dimension and cross sectional dimension, that is, T=50, N=08, the tests having maximum shortcomings less than 20% declared as best tests, tests having maximum shortcomings between 20% to 50% declared as mediocre tests, and tests having maximum shortcomings greater than 50% declared as worst tests. According to this benchmark, two tests have performed best in which one is residual -based test, that is, PdP_V and other best test is average weighted symmetric based test, that is, PAWS.

Table 6.1: Best, Mediocre and Worst Performing Tests According to their MS

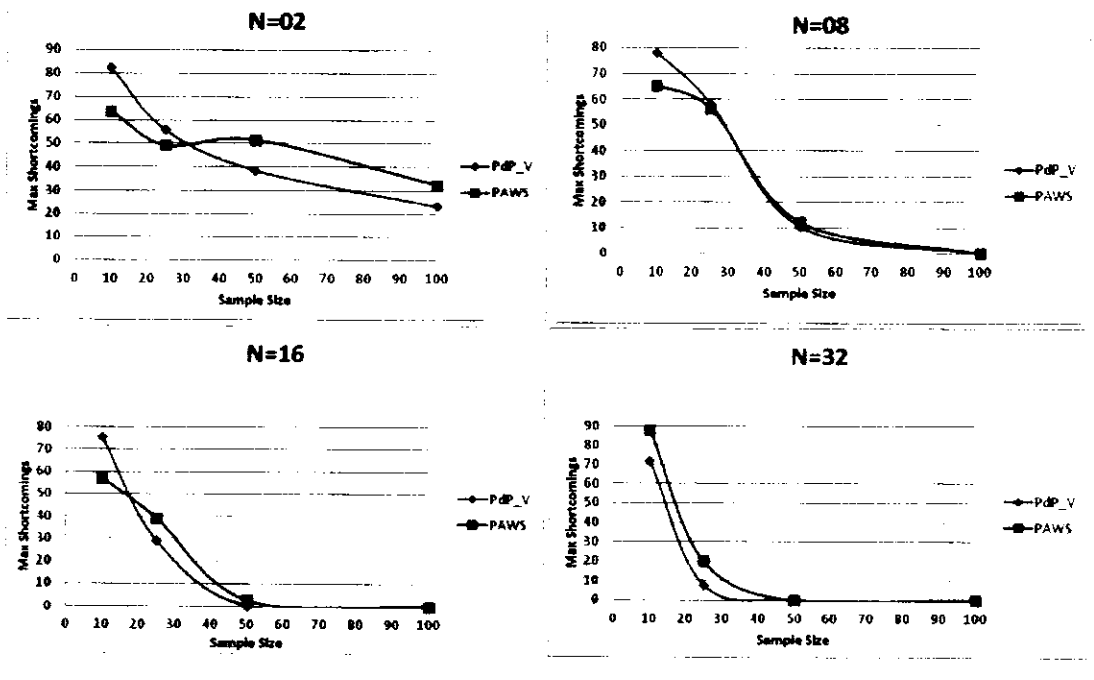
Category	Tests	Maximum Shortcomings
Best Performing Tests	PdP_V	10.5**
	PAWS	12.4**
Mediocre Performing Tests	PdPrho	38.6*
	Pfaws	44.8*
	PdGrho	47.66*
Worst Performing Tests	PdPtnp	56.66
	PdPtp	58
	PhZt	62.1
	PDFrho	62.56
	W_Pa	63.6
	PdGtp	66.3
	PdGtnp	66.66
	PDFrhostar	69.06
	W_Ga	71.5
	PADF	72.46
	PDFT	76.16
	PDFTstar	79.43
	Pfadf	81
	W_Pt	83.7
	W_Gt	91.2
	LR	93.8

There are three tests which perform mediocre in which two tests are residual based, that is, PdPrho, PdGrho and one Fisher type test. According to this benchmark, sixteen tests perform worst. All best, mediocre, and worst tests are displayed in Table 6.1.

Next, the performance of these best, mediocre and worst tests are analyzed by varying the time and cross sectional dimensions and check whether these two best tests also perform best in other time and cross section dimensions. Also, the performance of mediocre and worst tests are checked whether these tests are improved or worsen when time and cross sectional dimensions increase or decrease.

Figure 6.1 portrays the maximum shortcomings of best two tests. At T=10, both parametric residual-based test, that is, PdP_V and average weighted symmetric based test, that is, PAWS perform worst for all given number of cross section. At the same time dimension their maximum shortcomings are greater than 50%.

Figure 6.1: Best Performing Tests Having Null of No Cointegration



When time dimension is 25 then maximum shortcomings of these two tests lie between 50% to 60% for cross sectional size 02 and 08. As cross sectional dimension is 16 then both these parametric residual-based test, that is, PdP_V and average weighted symmetric based test, that is, PAWS perform mediocre corresponding to their maximum shortcomings are 30% and 40% respectively. At this time dimension, T=25, both tests perform best at cross sectional dimension 32, where parametric residual based test PdP_V has maximum shortcomings less than 10% while average weight symmetric test has maximum shortcomings less than 20% .

When time dimension is 50 then these two tests perform best for all given cross section except for cross sectional size 02, where these two tests perform mediocre. These two tests, that is, parametric residuals based PdP_V and average weighted symmetric test PAWS have zero maximum shortcomings and most stringent tests when cross sectional dimension is 16 or above sixteen. When time dimension is 100 then both these tests perform best and have zero shortcomings for all given cross section except for the cross sectional dimension 02. At this cross sectional dimension these two tests perform mediocre and have maximum shortcomings 20% to 30%.

Overall, Figure 6.1 summarized that when time dimension is 10 then cross sectional dimension does not improve the performance of these two tests. When time dimension is 25 then the performance of these two tests approximately equal for cross sectional dimension 02 and 08 but improve when cross sectional dimension is 16 and above. At time dimension 50, these two tests perform mediocre at cross sectional size 2 and perform best at cross sectional size 8 and also equally perform best at cross sectional dimension 16 and above. When time dimension is 50 these two tests, that is, parametric residual based PdP_V and average weighted symmetric based PAWS perform equally best at cross sectional dimension 8 and above.

These two best tests are explained in table form by classifying into six categories from 'A' to 'F' and **Table 6.2** shows that when cross sectional dimension is 02 then two tests, that is, residual based PdP_V and average weighted symmetric based PAWS fail to lie in the top two categories 'A' and 'B'.

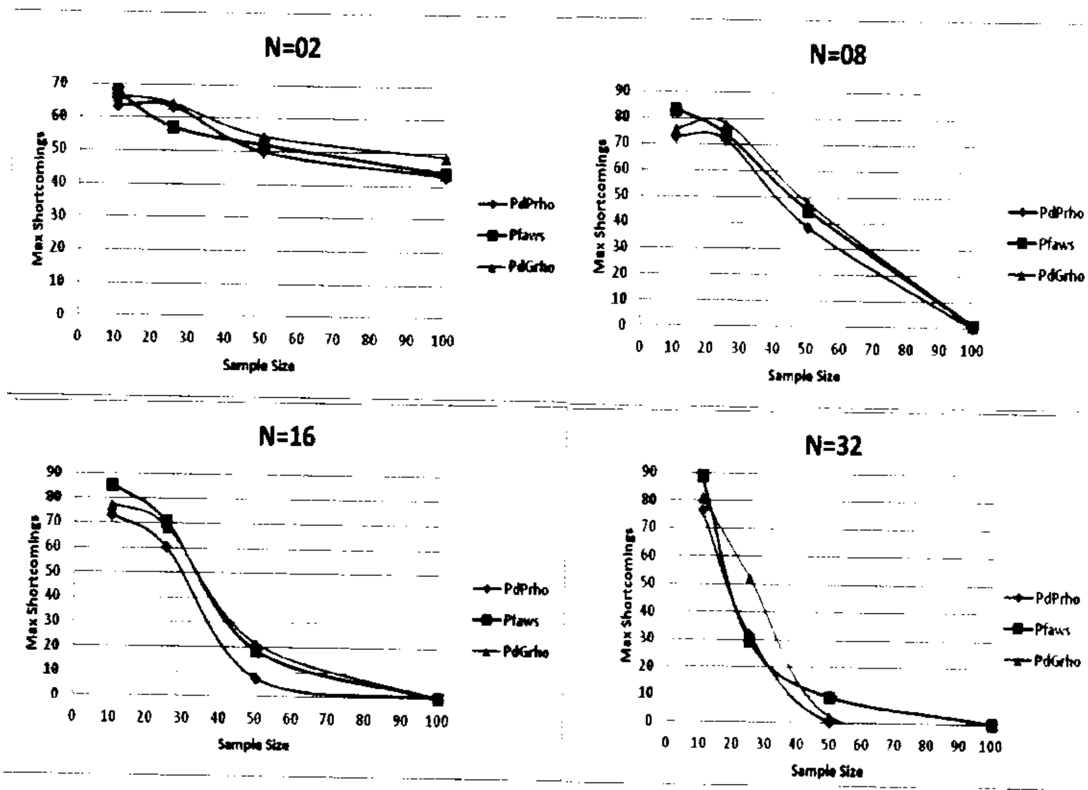
Table 6.2 : Best Performing Tests having Null of No Cointegration

N=02			N=08		
SS \ Tests	PdP_V	PAWS	SS \ Tests	PdP_V	PAWS
10	F	F	10	F	F
25	F	E	25	F	F
50	D	E	50	B	B
100	C	D	100	A	A
N=16			**N=32**		
SS \ Tests	PdP_V	PAWS	SS \ Tests	PdP_V	PAWS
10	F	F	10	F	F
25	C	D	25	A	B
50	A	A	50	A	A
100	A	A	100	A	A

When cross sectional dimension is 08, these tests lies in category 'B' and 'A' for time dimension 50 and 100 respectively. In case of cross sectional dimension 16 residual based test PdP_V lie in the category 'C' whereas average weighted symmetric based test lie in the category 'D' at time dimension 25 but both tests switch in the category 'A' at time dimension 50 and onward. In the last case, when cross sectional dimension is 32 then residual based test PdP_V and average weighted symmetric based test PAWS lie in the category 'A' and 'B' respectively at time dimension 25. But, both tests lie in the category 'A' at time dimension 50 and onward.

Figure 6.2 portrays the maximum shortcoming of three tests in which two tests are residual based, that is, PdPrho , PdGrho and one test is Fisher type test, these three tests are performed mediocre according to benchmark at N=08 and T=50. **Figure 6.2** depicts the performance of these three tests at different time and cross sectional dimension. When cross sectional dimension is 02, then all these three tests perform worst for all given time dimension except 100, where these three perform mediocre.

Figure 6.2: Mediocre Performing Tests having Null of No Cointegration



Although, at time dimension 50 these are performed worst but their maximum shortcomings are close to 50%. When cross section dimension is 08, all these three tests perform worst for time dimension less than 50. However, at the same time dimension, residual based test PdPrho perform little bit better as compared to other two tests but as whole these three tests perform mediocre at this time dimension. But, as the time dimension increased from 50 maximum shortcoming of all these three converge to zero. At the time dimension 100, all these three perform best and their performance is equal to the above best two test which were discussed in **Figure 6.1**.

When cross sectional dimension is 16, all these three tests perform worst when time dimension is 25 or less but the residual based test PdPrho perform batter as compared to other two throughout the given time dimensions. When time dimension is increased the maximum shortcomings of these three tests decrease, and at time dimension 50 all three tests perform best. However, residual based test PdPrho perform better than the other two tests which clearly evident from the figure. When time dimension is 100 then all these three tests equally perform best and have the zero percent maximum shortcomings. Hence, at this stage the performance of these three tests equal to the above two tests, that is, residual based test PdP_V and average weighted symmetric test PAWS which were discussed in the **Figure 6.1**.

At last cases, when cross sectional dimension is 32 at time dimension 10 all three perform worst but when time dimension is 25 then these three tests perform mediocre. Although, the residual based test PdPrho and Fisher type test Pfaws perform equally and better then the third one test PdGrho which is residual based test. When time dimension is 50 then all three tests perform best, the residual based tests PdPrho and PdGrho perform equally and have the zero percent maximum shortcomings whereas the Fisher type test Pfaws has maximum shortcomings around 10%. At the same stage, the performance of residual based tests PdPrho and PdGrho are equal with the performance of two best tests, that is, PdP_V and PAWS. At time dimension 100, all three tests perform best having zero percent maximum shortcomings and the performance of these three tests is equal to best two tests which were discussed in **Figure 6.1**.

Three mediocre performance tests are also discussed in the **Table 6.3** by varying time and cross sectional dimension. When cross sectional dimension is 2 then residual based test PdPrho performs better a little bit from the other two and lie in category 'E' at time dimension 50. Whereas other two test perform worst and lie in category 'F'. All three tests lie in the category 'E' at time dimension 100. Similarly when cross sectional dimension is 08 then all these three tests lies in the category 'A' at time dimension 100. At time dimension 50 residual based test PdPrho lies in the category 'E'.

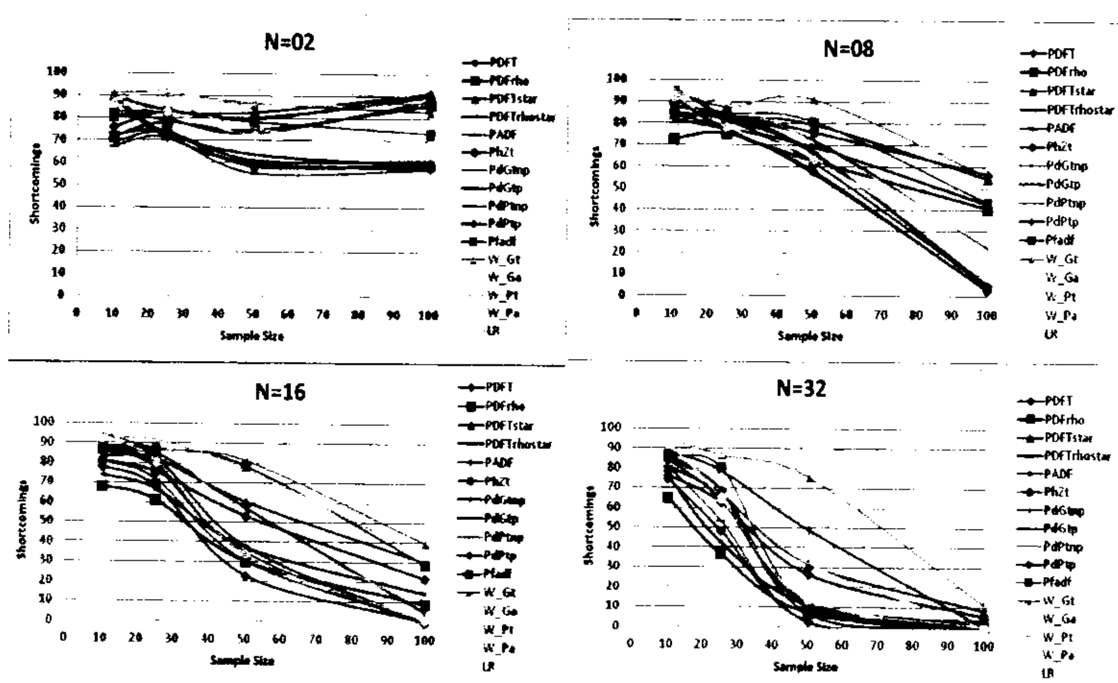
Table 6.3: Mediocre Performing Tests having Null of No Cointegration

N=02				N=08			
Tests \ SS	**PdPrho**	**Pfaws**	**PdGrho**	Tests \ SS	**PdPrho**	**Pfaws**	**PdGrho**
10	F	F	F	10	F	F	F
25	F	F	F	25	F	F	F
50	E	F	F	50	D	E	E
100	E	E	E	100	A	A	A
N=16				**N=32**			
Tests \ SS	**PdPrho**	**Pfaws**	**PdGrho**	Tests \ SS	**PdPrho**	**Pfaws**	**PdGrho**
10	F	F	F	10	F	F	F
25	F	F	F	25	D	C	F
50	A	B	C	50	A	A	A
100	A	A	A	100	A	A	A

When cross sectional dimension is 16, all these three tests lie in the category 'F' when time dimension is 25 or less than 25. When time dimension is 50 then residual based tests PdPrho lies in category 'A' whereas Fisher type test Pfaws lies in category 'B' and residual based test PdGrho lies in the category 'C'. At time dimension 100, all these three lie in the category 'A'. In the last case of the Table 6.3, when cross sectional dimension is 32 then all these three tests lie in the highest category 'A' at time dimension 50 and onward.

Figure 6.3 depicts the shortcomings of sixteen tests of different categories which perform worst according to benchmark at different time and cross sectional dimension. When the cross sectional dimension is 02 then all the test perform worst throughout the given time dimension. When cross section dimension is 08, it is observed from the figure that all tests is worst at the time dimension 50 or less than 50 having maximum shortcomings above 50%.

Figure 6.3: Worst Performing Tests having Null of No Cointegration



When the time dimension increase and become 100 then six tests perform best and have maximum shortcomings less than 20% in which one test is error correction base, that is, W_Pa and two tests are non parametric residual based, that is, PdPtnp, PdGtnp and three tests are parametric residual based, that is, PhZt, PdPtp, PdGtp, At this cross sectional and time dimensions, five tests perform mediocre in which one test is error correction based, that is, W_Ga , one test is Fisher type, that is, Pfadf and three tests are residual based, that is, PDFrho, PDFrhostar, PADF. Whereas five tests still perform worst at this time and cross sectional dimension. In these five worst performing tests two tests are error correction based, that is, W_Pt, W_Gt and one test is maximum likelihood hood based, that is, LR and two tests are residual based tests, that is, PDFT, PDFTstar.

In the above **Figure 6.3** when cross sectional dimension is 16, all these sixteen tests perform worst when time dimension is 25 or less than 25. When time dimension is 50 then eight tests perform mediocre and eight tests perform worst, in which eight mediocre performance tests one is error correction based test, that is, W_Pa whereas two non-parametric residual based PdPtnp, PdGtnp and five residual based which are PDFrho, PDFrhostar, PhZt, PdPtp, PdGtp. The eight worst performance tests consist one maximum likelihood based test, that is, LR, one Fisher type test, that is, Pfadf, three error correction based tests, that is, W_Pt, W_Gt , W_Ga and three residual based tests PDFT, PDFTstar, PADF.

When time dimension is increased and become 100 then nine tests perform best, five tests perform mediocre and one test perform worst which is maximum likelihood based test, that is, LR test. The nine best tests consist two error correction based tests, that is, W_Pa, W_Ga , two non-parametric residual based tests, that is, PdPtnp, PdGtnp and six residual based, that is, PDFrho, PDFrhostar, PhZt, PdPtp, PdGtp whereas five mediocre performance tests consist two error correction based tests, that is, W_Pt, W_Gt and two residual based tests, that is, PDFT, PDFTstar and one Fisher type test, that is, Pfadf. When cross sectional dimension is 32, in this case when time dimension is 10 then all these sixteen tests perform worst but when time dimension is 25 then only three tests perform mediocre while thirteen tests perform still worst. The three mediocre performance tests are residual based, that is, PDFrho, PDFrhostar, PdPtp.

When time dimension is 50 then nine tests perform best, four tests perform mediocre and three tests perform worst. Nine best performing tests consist one is error correction based test, that is, W_Pa, one Fisher type test, that is, Pfadf whereas two non-parametric residual based PdPtnp, PdGtnp and five residual based which are PDFrho, PDFrhostar, PhZt, PdPtp, PdGtp. In the four mediocre performing tests three consist residual based, that is, PDFT, PDFTstar, PADF and one error correction based, that is, W_Ga whereas three worst performing tests are error correction based, that is, W_Pt, W_Gt and maximum likelihood based test LR. When time dimension is 100 then fifteen tests perform best and only one test still perform worst which is maximum likelihood based test.

Table 6.4 have also discussed the worst performance sixteen tests at different time and cross sectional dimension. When cross sectional dimension is 02 then all sixteen tests lie in lowest category 'F' throughout the given time dimension. When cross sectional dimension is 08 then all these sixteen tests lie the category 'F' at time dimension 50 or less than 50, but at time dimension 100 five tests much improve and switch in the

Table 6.4: Worst Performing Tests having Null of No Cointegration

N=02					N=08				
Tests \ SS	10	25	50	100	Tests \ SS	10	25	50	100
PDFT	F	F	F	F	PDFT	F	F	F	F
PDFrho	F	F	F	F	PDFrho	F	F	F	E
PDFTstar	F	F	F	F	PDFTstar	F	F	F	F
PDFrhostar	F	F	F	F	PDFrhostar	F	F	F	E
PADF	F	F	F	F	PADF	F	F	F	C
PhZt	F	F	F	F	PhZt	F	F	F	A
PdGtnp	F	F	F	F	PdGtnp	F	F	F	A
PdGtp	F	F	F	F	PdGtp	F	F	F	A
PdPtnp	F	F	F	F	PdPtnp	F	F	F	A
PdPtp	F	F	F	F	PdPtp	F	F	F	A
Pfadf	F	F	F	F	Pfadf	F	F	F	E
W_Gt	F	F	F	F	W_Gt	F	F	F	F
W_Ga	F	F	F	F	W_Ga	F	F	F	C
W_Pt	F	F	F	F	W_Pt	F	F	F	F
W_Pa	F	F	F	F	W_Pa	F	F	F	B
LR	F	F	F	F	LR	F	F	F	F

N=16					N=32				
Tests \ SS	10	25	50	100	Tests \ SS	10	25	50	100
PDFT	F	F	F	C	PDFT	F	F	C	A
PDFrho	F	F	D	A	PDFrho	F	F	A	A
PDFTstar	F	F	F	C	PDFTstar	F	F	D	A
PDFrhostar	F	F	D	B	PDFrhostar	F	F	B	A
PADF	F	F	F	A	PADF	F	F	E	A
PhZt	F	F	D	A	PhZt	F	F	A	A
PdGtnp	F	F	D	A	PdGtnp	F	F	B	A
PdGtp	F	F	D	A	PdGtp	F	F	A	A
PdPtnp	F	F	D	A	PdPtnp	F	F	A	A
PdPtp	F	F	C	A	PdPtp	F	F	A	A
Pfadf	F	F	F	C	Pfadf	F	F	B	A
W_Gt	F	F	F	D	W_Gt	F	F	F	B
W_Ga	F	F	F	A	W_Ga	F	F	D	A
W_Pt	F	F	F	D	W_Pt	F	F	F	A
W_Pa	F	F	D	A	W_Pa	F	F	B	A
LR	F	F	F	F	F	F	F	F	F

category 'A' and one error correction based test W_Pa lies in the category 'B'. In these five tests, two tests are non parametric residual based and three tests are residual based.

When cross sectional dimension is 16 then all of these sixteen tests lie in the category 'F' at time dimension 25 and less than 25, few tests a little bit improve and switched the higher category at time dimension 50. When time dimension is 100 then nine tests lie in the category 'A' in which two tests are error correction based, that is, W_Pa, W_Ga, two tests are non parametric residual based, that is, PdPtnp, PdGtnp and five tests are residual based, that is, PDFrho, PDFrhostar, PADF, PdPtp, PdGtp, maximum likelihood based test, that is, LR still lie in the lowest category 'F'. In the last case when cross sectional dimension is 32, all these sixteen tests lie in the worst category 'F' at time dimension 25 but when time dimension is 100, all these tests lie in the category 'A' except the maximum likelihood based test, that is, LR which is still perform worst and lie in the category F.

6.1.1. Summary

Overall, the summary of the above sections consists of three figures and three tables. In the above section, the power of 21 of tests is analyzed of different categories. These categories include: residual based parametric/nonparametric, error correction based, maximum likelihood based, Fisher type tests which are based on p-values, average weighted symmetric based tests having null of no cointegration using stringency criteria. A benchmark is marked on the performance of these 21 tests at cross sectional dimension 08 and time dimension 50. At this benchmark, two tests perform best, that is, residual based test PdP_V and average weighted symmetric based test PAWS. From the **Figure 6.1** and **Table 6.2,** the performance of these two tests is observed on different time and cross sectional dimension. When, comparing Figure 6.1 with other two Figures 6.2 and Figure 6.3 or comparing Table 6.2 with other two Table 6.3 and Table 6.4, and the performance of these two best tests residual based test PdP_V and average weighted symmetric based test PAWS at different time and cross sectional dimension. These two tests are performed better as compared to all other tests at all time and cross sectional dimension. Also these two tests, that is, residual based tests PdP_V and average weighted symmetric based test PAWS are most stringent test. Furthermore, the stringencies of the most tests approach to zero at time dimension 100 and cross sectional dimension 32. Their performance become equal to each other and also the stringencies of more than half tests approach to zero at cross sectional dimension 16 and time dimension 100 in all above three figures. From the **Figure 6.2,** three mediocre performance tests are discussed, in these three tests the residual based test PdPrho perform better the other two tests. In other words, from **Table 6.3** the residual based test PdPrho lies in the higher category as compared to other two tests but equally perform when time dimension is 50 and onward at cross sectional dimension 32.

Next, comparing Figure 6.2 with Figure 6.3 or Table 6.3 with Table 6.4, overall these three mediocre performance tests perform better than the other sixteen worst performing tests. Although, these half of the worst performing tests equally well and their stringencies approach to zero like mediocre and best performing tests when time dimension is 100 and cross sectional dimension is 32. Form our above power of the tests analysis, first two best performing tests are residual based tests PdP_V and average weighted symmetric test PAWS, third best performing test is also residual based test, that is, PdPrho, the maximum likelihood based test, that is, LR perform worst throughout the time and cross sectional dimension.

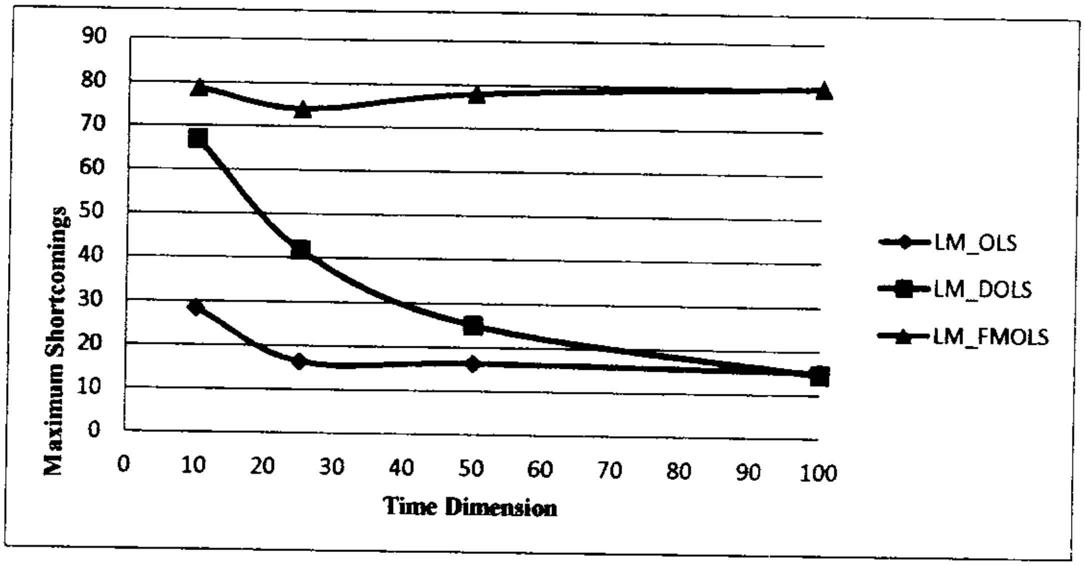
6.2. Power of Tests having Null of Cointegration

In this section, the analysis of power of tests having null of cointegration is explained using stringency criterion. Here, tests are explained using figures and tables. In the figures the maximum shortcomings of three tests OLS based test, that is, LM_OLS, Dynamic OLS based test, that is, LM_DOLS and Fully Modified OLS based test, that is, LM_FMOLS have been drawn. In the table, all the under consideration tests are classified into six categories from A to F as have discussed above. As, already mentioned that the time length is 10, 25, 50 and 100, and number of cross section are 2, 8, 16 and 32 to compare tests.

In Figure 6.4, the maximum shortcomings of three tests have been displayed. The time dimension is labeled along x-axis and along y-axis maximum shortcomings have been labeled. In Figure 6.4, two numbers of cross sections have been taken in this case. When time dimension is 10, one test LM_OLS performs better than the other two tests. It has maximum shortcomings around 30% where as LM_DOLS and LM_FMOLS tests have the maximum shortcomings above 65% but when the time length is increased and becomes 25, then maximum shortcomings of LM_OLS test is around 15% whereas the maximum shortcomings of LM_DOLS is around 40%. The third test LM_FMOLS still performs badly and has the shortcoming above 70%.

When the time dimension is 50 then LM_OLS does not change the behavior, it has still shortcoming around 15% but LM_DOLS improves the position as compared to previous time length position and has shortcoming around 25%. But, LM_FMOLS test still performs badly and has shortcoming above 70%.

Figure 6.4: Maximum Shortcomings of Tests having Null of Cointegration, N=02



Finally, when time dimension is 100 then LM_OLS test has still same maximum shortcomings as it was in the previous time length. LM_DOLS again improve its position and has shortcomings around 15%. The LM_FMOLS still performs badly and has maximum shortcomings above 70%. Figure 6.4 concludes that over all LM_OLS performs best as compared to other two and LM_FMOLS performs worst throughout the time dimension.

Now we summarize the Figure 6.4 in table form which is expressed in Table 6.5. Again, all three tests LM_OLS, LM_DOLS, and LM_FMOLS are classified into the six categories from 'A' to 'F'. When time dimension is 10, the test LM_OLS lies in the category 'C' whereas the other two tests LM_DOLS and LM_FMOLS lie in the lowest category 'F'. When time dimension is 25 then LM_OLS has switched from the category

'C' to 'B' where the test LM_DOLS has switched from the category 'F' to 'E', but the test LM_FMOLS still performs worst and lie in the category 'F'. When time dimension becomes 50, the test LM_OLS maintains same position as in category 'B', the test LM_DOLS has switched the category from 'E' to 'C' whereas the test LM_FMOLS still remained in the lowest category 'F'.

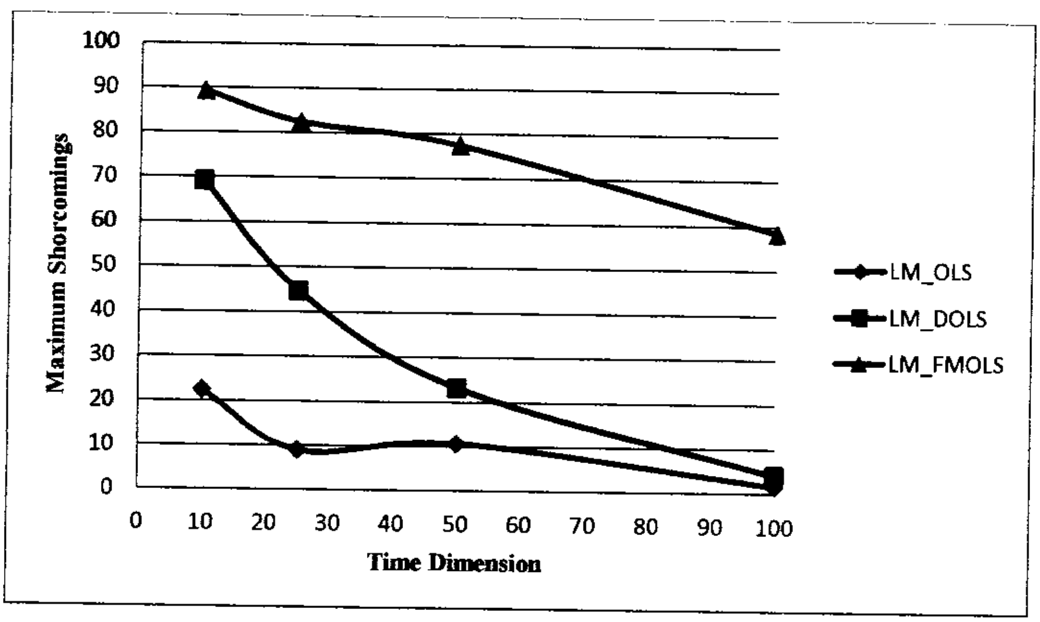
Table 6.5: Power of Tests having Null of Cointegration (N=02)

Tests / SS	LM_OLS	LM_DOLS	LM_FMOLS
10	C	F	F
25	B	E	F
50	B	C	F
100	B	B	F

When the time dimension is 100 then the tests LM_OLS does not improve and still remains in the category 'B' and LM_DOLS has improved and switched from the category 'C' to 'B'. Whereas the test LM_FMOLS has still performed worst and lied in the category 'F'. Overall, summery of Table 6.5, shows that LM_OLS performs best as compared to other two tests where as the test LM_FMOLS performs worst for all time dimensions and given cross section.

In the following Figure 6.5, the maximum shortcomings has been displayed. In this case, a numbers of cross sections 8 is discussed where the time dimension has been labeled along x-axis and maximum shortcomings along y-axis. From the figure when time dimension is 10, LM_OLS test has maximum shortcomings around 22%, LM_DOLS and LMFMOLS have maximum shortcomings around 70% and 90% respectively. When time dimension is 25, LM-OLs performs well and has the maximum shortcomings around 10% whereas the test LM_DOLS has shortcoming between 40% and 50%, LM_FMOLS still performs badly and has shortcoming around 80%. Similarly, when time dimension is 50, the maximum shortcomings of LM_OLS remains around 10% whereas the test LM_DOLS has improved the position and has maximum the shortcoming around 20%. But the test LM_FMOLS still performs worst and has maximum shortcoming between 75% and 80%. Finally, when time dimension is 100, then the maximum shortcoming of LM_OLS converges to zero.

Figure 6.5: Maximum Shortcomings of Tests having Null of Cointegration, N=08



Similarly the maximum shortcoming curve of LM_DOLS has also approached to zero but the maximum shortcoming of test FM_OLS has remained at 60%. Overall, the performance of all three tests, the LM_OLS performs best as compared to other tests in all time dimensions and the LM_FMOLS performs worst though out the time dimension. These findings are also explained in table form, where the tests have been classified into six different categories from 'A' to 'F'. In Table 6.6, it is observed that when time dimension is 10 then LOM_OLS test lies in the category 'C' and the other two LM_DOLS and LM_FMOLS lie in the category 'F'. When time dimension is 25 then LM_OLS lies in the category 'A' whereas test LM_DOLS lies in the category 'E', but LM_FMOLS still performs badly and lies in the category 'F'. When time length further increases and becomes 50 then LM_OLS lies in category 'A' and test LM_DOLS improves and switched from category 'E' to 'C'. But, test LM_FMOLS still lies in the lowest category 'F'. Finally, when time dimension becomes 100, then both tests LM_OLS and LM_DOLS lie in the category 'A', whereas test LM_FMOLS still performs worst and lies in the lowest category 'F'.

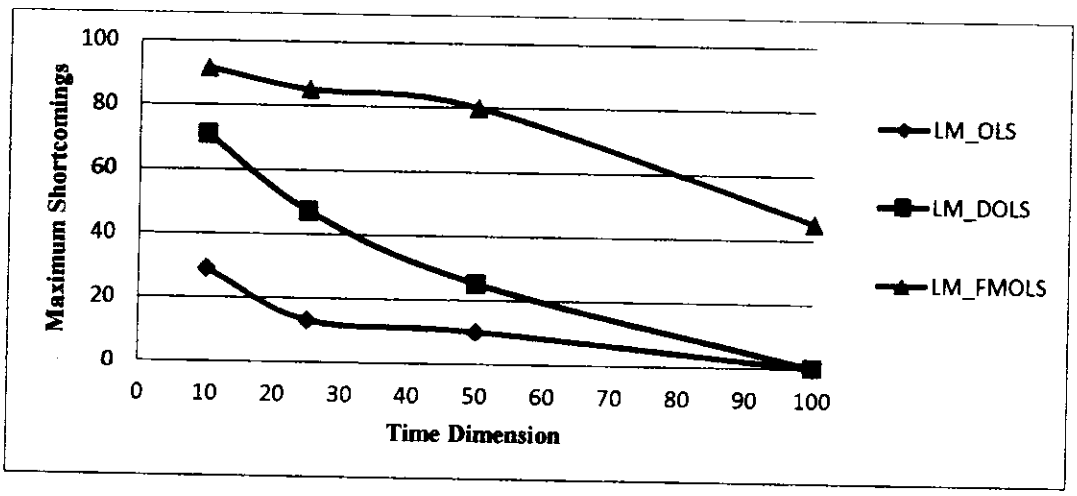
Table 6.6. Power of tests having null of cointegration (N=08)

Tests / SS	LM_OLS	LM_DOLS	LM_FMOLS
10	C	F	F
25	A	E	F
50	A	C	F
100	A	A	F

In Table 6.6, overall performance of LM_OLS test is best and most stringent test as compared to all other two tests in all time length, whereas test LM_FMOLS performs worst throughout the time dimension.

In Figure 6.6, maximum shortcomings of three tests have been displayed. In this case 16 number of cross sections have been taken, time dimension has been labeled along x-axis whereas maximum shortcomings have been labeled along y-axis. It is seen from Figure 6.6 that when time dimension is 10, two tests LM_DOLS and L_FMOLS perform worst and have maximum shortcomings above 70% whereas the LM_OLS has maximum shortcomings around 30%. As the time length increases and becomes 25 then the test LM_DOLS improves and it has maximum shortcoming around 47% whereas the maximum shortcomings of LM_OLS is also more decreased and around 13%, the test LM_FMOLS still performs worst. When the time dimension is 50 then the test LM_OLS has maximum shortcomings around 10% where the test LM_DOLS has maximum shortcomings around 25 %, the test LM_FMOLS still performs badly and has maximum shortcomings around 80%. Similarly, when time dimension is 100 then both the tests, that is, LM_OLS and LM_DOLS have the zero maximum shortcomings and the test LM_FMOLS has the maximum shortcomings around 45%.

Figure 6.6: Maximum Shortcomings of Tests having Null of Cointegration, N=16



Overall, Figure 6.6 indicates that LM_OLS test performs best as compared to all other two tests and is most stringent test in this figure, whereas the test LM_FMOLS performs worst throughout the time dimension.

In Table 6.7, the summary discussion of Figure 6.6 is explained. In Table 6.7, three tests are classified into six different categories from 'A' to 'F'. As we can see from the following Table 6.7, when time dimension is 10 then the test LM_OLS lies in the category 'C' whereas, the other two tests LM_DOLS and LM_FMOLS lie in the lowest category 'F'. But when the time length is increased and becomes 25 then test LM_OLS lies in the category 'B' whereas the test LM_DOLS switches from the category 'F' to 'E', test LM_FMOLS still lies the lowest category 'F'. When time length is 50 then LM_OLS remains in the same category 'B' whereas test LM_DOLS switches from the category 'E' to 'C' and the test LM_FMOLS still performs worst and lies in the category 'F'.

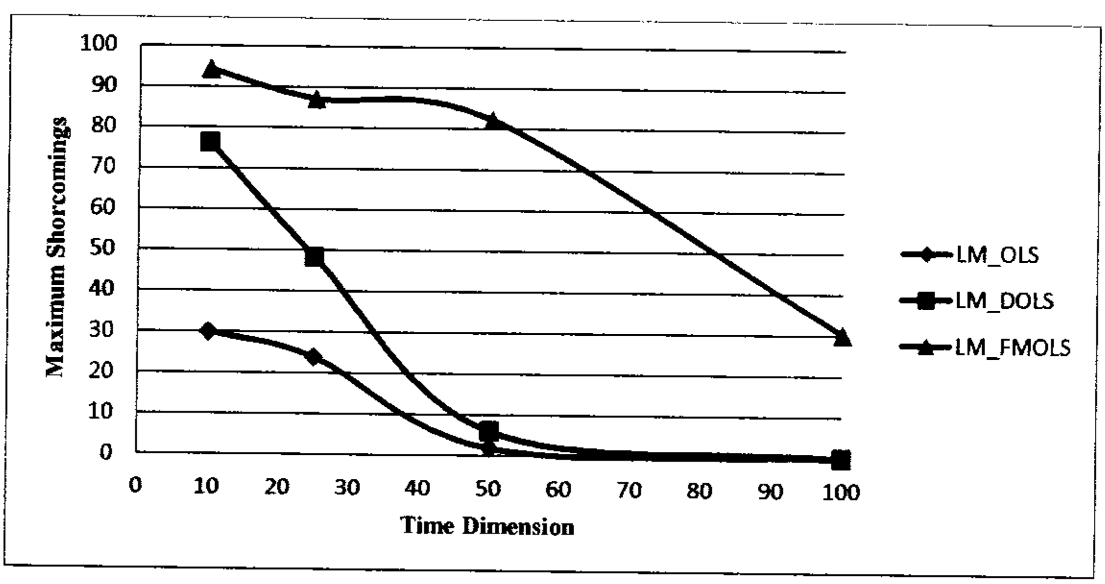
Table 6.7 Power of Tests having Null of Cointegration (N=16)

Tests SS	LM_OLS	LM_DOLS	LM_FMOLS
10	C	F	F
25	B	E	F
50	B	C	F
100	A	A	E

Finally, when time dimension is 100 then both tests LM_OLS and LM_DOLS lie the highest category 'A', whereas, the test LM_FMOLS lies in category 'E'. Hence, in this case again LM_OLS test is best and most stringent test.

Figure 6.7 explains the results for N=32, where time dimension has been taken along x-axis and maximum shortcomings of tests have been taken along y-axis. It is evident from Figure 6.7 that when time dimension is 10 then the test LM_OLS has maximum shortcomings around 30%. Whereas, the other two tests LM_DOLS and LMFMOLS have performed badly and have maximum shortcomings around 75% and 95% respectively. When time length increases and becomes 25 then maximum shortcomings of LM_OLS is around 23%. At the same time level, LM_DOLS has maximum shortcomings around 50% and test LM_FMOLS has performed badly and has maximum shortcomings about 87%. Similarly, when time dimension is 50 then LM_OLS performs well and has maximum shortcomings less than 5% whereas the other test LM_DOLS also performs well and has maximum shortcomings around 5%. LM_FMOLS test still performs worst and has maximum shortcomings around 80%.

Figure 6.7: Maximum Shortcomings of Tests having Null of Cointegration, N=32



At last when time dimension is 100 then both the tests LM_OLS and LM_DOLS have a zero maximum shortcomings, test LM_FMOLS has a maximum shortcomings around 30%. In Figure 6.7, test LM_OLS performs best as compared to other tests throughout the time dimension and LM_FMOLS performs worst throughout time dimension except at 100 sample size. Hence, LM_OLS test is the most stringent test.

Table 6.8 also explains the results shown in Figure 6.7. Three tests have been classified into six different categories from 'A' to 'F'. It is seen from Table 6.8, when time dimension is 10 then test LM_OLS lies in the category 'C' whereas the other two tests LM_OLS and LM_FMOLS lie in the lowest category 'F'. When time dimension is 25 then LM_OLS still lies in the category 'C' but test LM_DOLS switches the category 'F' to 'E', test LM_FMOLS still performs worst and lies in the category 'F'. As the time length increases and becomes 50 then both tests LM_OLS and LM_DOLS lie in the highest category 'A', but the test LM_FMOLS still lies in the lowest category 'F'

Table 6.8. Power of Tests having Null of Cointegration (N=32)

Tests / SS	LM_OLS	LM_DOLS	LM_FMOLS
10	C	F	F
25	C	E	F
50	A	A	F
100	A	A	D

Finally, when time length is 100 then again both tests LM_OLS and LM_DOLS lie in the highest category 'A' and the test LM_FMOLS lies in the category 'D'. Hence again the test LM_OLS performs better as compared to other tests and LM_FMOLS performs worst throughout the time dimension. Hence, the test LM_OLS is the best and most stringent test.

6.2.1. Summary

Second section explains analysis of tests having null of cointegration according to power using stringency criterion. This section consisted of four figures and four tables. When time dimension is 10, then two tests, that is, LM_DOLS and LM_FMOLS perform worst for all cross sections and have maximum shortcomings between 70% and 90%. But, test LM_OLS performs better as compared to other two tests for all cross sectional dimension and lies between in range of 20% to 30%. At this stage LM_DOLS and LM_FMOLS tests lie in the category 'F' for all number of cross sections and LM_OLS lies in the category 'C' for all number of cross sections.

When time dimension is 25 then maximum shortcomings of LM_OLS and LM_DOLS tests move downward but the cross sectional dimension does not affect the range of maximum shortcomings of these two tests, and LM_FMOLS test performs worst for all cross sections. At this stage two tests LM_OLS and LM_DOLS lie in the category 'B' and 'E' respectively for all cross sections, whereas test LM_FMOLS lies in the category 'F' for all number of cross sections.

When the time length is 50, then the maximum shortcomings of test LM_OLS is around 10% for cross sections 2, 8, and 16 but its maximum shortcomings is less than 5% when cross sections are 32. In other words, it lies in the category 'B' and 'A' for cross section 2, 8, 16, and 32 respectively. Whereas the maximum shortcomings of LM-DOLS test lies between the range 20% to 30% for the cross sections 2, 8, and 16 but its maximum shortcomings is less than 5% for cross section 32. In other words, it lies in the category 'C' and 'A' for the cross section 2, 8, 16, and 32 respectively. The test LM_FMOLS still lies in the range of 70% to 80% for all cross sections, it lies in the lowest category 'F'. In this case the cross section length has not much effect on the maximum shortcomings of LM_OLS and LM_DOLS tests except when its length is 32.

Finally, when time length is 100, then it is observed that cross sectional variation does not affect the maximum shortcomings of all these three tests, the maximum shortcomings of tests LM_OLS and LM_DOLS are zero for all cross sections. In other words, these tests lie in the highest category 'A', where the maximum shortcomings of LM_FMOLS test lies in category 'F' except the case when cross section length is 32. Hence, overall analysis shows that test LM_OLS performs best and most stringent test for time dimension and cross section whereas the test LM_FMOLS performs worst for all time lengths and for all cross sectional lengths.

6.3. Conclusion

Now, a general conclusion of both sections is explained in which first section deals with the power comparison using stringency criterion of tests having null of no cointegration whereas the other section deals with the power comparison using stringency criterion of tests having null of cointegration. In the first section, it is observed that most of the tests have maximum shortcomings downward as the time length and cross section length increases.

For the sake of simplicity, tests are classified into six categories from 'A' to 'F'. A cutoff point is marked at time dimension 50 and cross sectional dimension 08. At this cutoff point, it is observed that two tests perform best and have stringencies between 0% to 20% and three tests perform mediocre and have a stringencies between 20% to 50%. Whereas, sixteen worst performing tests have a stringencies above 50%.

From **Figure 6.1,** the performance of these two best tests is observed, that is, Parametric residual based test PdP_V and average weighted symmetric based test PAWS at different time and cross sectional dimension. Both of these two tests perform worst when time dimension is 10 in all four cases of cross sectional dimensions. In **Figure 6.1,** when cross sectional dimension is 2 these two tests perform mediocre at time dimension 50 and onward. When cross sectional dimension is 08 these two tests perform mediocre at time dimension 50 and perform best at time dimension 100. However, when cross sectional dimension is 16 these two tests are performed mediocre at time dimension 25 but performed best at time dimension 50 and onward. When cross sectional dimension is 32 then residual based test, that is, PdP_V and average weighted symmetric based test PAWS perform best at time dimension 25 and onward.

Figure 6.2 portrays the stringencies of three mediocre tests in which two tests are Parametric residual based, that is, PdPrho, PdGrho and one test is Fisher type test, that is, Pfaws. if the performance of these three tests are compared, then it is concluded that Parametric residual based test PdPrho perform better the other two tests. When cross sectional dimension is 02 all these three tests perform worst at time dimension 25 and less than 25 and at time dimension 50 only parametric residual based test PdPrho perform mediocre whereas the other two tests perform worst, at time dimension 100 all three perform mediocre. When cross sectional dimension is 08 these three tests perform mediocre at time dimension 50 and all these three tests perform best at time dimension 100. When cross sectional dimension is 16 the parametric residual based test PdPrho and Fisher type test perform best at time dimension 50 and onward but residual based test PdGrho perform mediocre at time dimension 50 and it performs best at time dimension 100. When cross sectional dimension is 32 then parametric residual based test PdPrho and Fisher type test Pfaws perform mediocre at time dimension 25 whereas the third residual based test perform worst at this time dimension. When time dimension is 50 or more, all these three tests perform best.

From **Figure 6.3,** stringencies of sixteen worst performance tests of different types, that is, parametric/non parametric residual based tests, error correction based tests, maximum likelihood based tests, Fisher type test are displayed and their performance are observed at different time and cross sectional dimension. All of these sixteen different types of tests perform worst throughout the given cross section and time dimension when cross sectional dimension is 2. At time dimension 100 and cross section dimension 08, six tests perform best in which three tests are parametric residual based tests PhZt, PdPtp, PdGtp, one test is error correction based test W_Pa and two tests are non-parametric residual based tests PdPtnp, PdGtnp. Seven parametric/non parametric residual based tests, that is, PDFrho, PDFrhostar, PhZt, PdPtp, PdGtp, PdPtnp, PdGtnp and one error correction based test, that is, W_Pa perform mediocre while one maximum likelihood based test, three error correction based tests, one Fisher type test and three parametric residual based tests perform worst at cross section dimension 16 and time dimension 50.

When time dimension is increased from 50 then most of the tests improve their performance in which eight parametric /nonparametric residual based tests and two error correction based tests perform best at time dimension 100. Five tests perform mediocre and maximum likelihood based test, that is, LR still perform worst. When cross sectional dimension is 32 and time dimension is 50 then seven tests perform best in which one test is Fisher type test, one tests is error correction based test and five tests are parametric /non parametric residual based test, while seven tests perform mediocre and two tests perform worst in which one is maximum likelihood based and one is error correction based test, at time dimension 100 all these tests perform best except maximum likelihood based test, that is, LR.

In the second section, three tests having null of cointegration have been analyzed. In this section Figure 6.4 to Figure 6.7 and Tables 6.5 to Table 6.8 have been explained. When time length is 10 then the maximum shortcomings of test LM_OLS have been lied in the range 20% to 30% for all cross section, length of cross section does not have impact on the performance of this test. Where the tests LM_DOLS and LM_FMOLS perform worst for all cross section at this time length. At this time length in Table 6.5 to Table 6.8, LM_OLS test has been lied in the 'C' category and tests LM_DOLS and LM_FMOLS have been lied in the category 'F'. When time length is 25 then LM_OLS has been lied in the range 10% to 20% for all cross section and LM_DOLS test has been lied in the range 40% to 50% for all cross sections. Third test LM_FMOLS perform worst for all cross sections. At this time length, LM_OLS test has been lied in category 'B' and test LM_DOLS has been lied in category 'E' and the test LM_FMOLS has been lied in the category 'F'. In this section, two tests LM_OLS and LM_DOLS improve their performance when time length increases but the cross section length does not have impact on these two tests. Where the test LM_FMOLS test performs worst for all situation in the analysis. In the second section of analysis, LM_OLS test performs best throughout and the LM_DOLS test performs better except when time length is ten.

Chapter 07: Bootstrap Empirical Power

As in the previous chapters, the size and power properties of different panel cointegration tests are discussed using simulated critical values. In this chapter, the empirical power of best performing tests, that are, PdP_V, PAWS, PdPrho are demonstrated using the example of Fisher hypothesis. If the series of nominal interest rate and inflation rate are non-stationary then the application of panel cointegration technique is a suitable choice to test for the presence of long run relationship between the nominal interest rate and inflation rate. In the existing empirical literature, the evidence on the existence of Fisher hypothesis is found. There are several studies in the literature of time series analysis and panel data analysis finding that the Fisher hypothesis holds. The Fisher hypothesis describes that the real interest rate is the difference between the nominal interest rate and expected inflation rate that is

$$r_{it} = n_{it} - \pi_{it}^{e} \quad \text{..(7.1)}$$

where r_{it}, n_{it} and π_{it}^{e} denote the real interest rate, nominal interest rate and expected inflation, respectively. The Fisher hypothesis is tested by using the following equation that is,

$$n_{it} = \alpha_i + \beta_i \pi_{it} + \varepsilon_{it} \quad \text{...(7.2)}$$

where n_{it} and π_{it} represent the nominal interest rate and observed inflation rate for *ith* cross section at time t respectively. Similarly, α_i represents the country specific constant which also denotes the mean of the ex-ante real interest rate. To evaluate the performance of three best performing tests on the basis of real data, it is assumed that the cointegration relation does exist between the nominal interest rate and the inflation rate, that is, Fisher hypothesis holds. For the evaluation of performance of the said three tests, the concept of empirical power is used which is derived from real data using bootstrap method. Therefore, bootstrap critical values are used instead of simulated critical values.

Here, two different data sets are considered, that is, two cross sectional dimensions of OECD countries for three different time dimensions. For this purpose, quarterly data of nominal interest rate and the inflation rate are used. The data are taken from Organization of Economics Co-operation and Development (OECD). The first data set is called "Data-A", which consists of eight OECD countries (N=8), "Australia, Belgium, Canada, Germany, Spain, France, the United Kingdom, Italy" having three time dimension, that is, from 2004Q2-2016Q4 (T=50), 2010Q4-2016Q4 (T=25), 2014Q2-2016Q1 (T=10). Whereas, the second data set is called "Data-B", which consists of sixteen OECD countries (N=16) "Australia, Austria, Belgium, Canada, Germany, Spain, France, Finland, Ireland, Iceland, Korea, the United Kingdom, Italy, Norway, Sweden" having three time dimension from 2004Q2-2016Q4 (T=50), 2010Q4-2016Q4 (T=25), 2014Q2-2016Q1 (T=10).

7.1. Empirical Powers (N=08)

Table 7.1 depicts the bootstrap empirical powers of three best performing tests, that is, PdP_V, PdPrho and PAWS for eight countries and three time dimensions, that is, 10, 25 and 50. It is evident that at the smallest time dimension of 10, two tests, that is, PAWS and PdPrho perform far better than the third test, that is, PdP_V. When time dimension is increased from 10 to 25 then bootstrap empirical powers of all three tests increase with significant magnitude, as empirical power of PAWS increases from 56.93% to 99.8%, empirical power of PdP_V increases from 45.03% to 77.8% and the empirical power of PdPrho increases from 56.63% to 99.3%. This behavior shows that when time dimension is increased then empirical powers of all three tests are also increased which is analogous with the conclusion earlier drawn in Chapter 6. Again, when time dimension is increased from 25 to 50 then empirical power of all three tests are 100% which is again in accordance with the conclusions earlier drawn in Chapter 6. From these three better performing tests, two tests, that is, PAWS and PdPrho are performing superior than the third one, that is, PdP_V in terms of bootstrap empirical power.

Table 7.1: Bootstrap Empirical power for N=08

N=08			
T	**PAWS**	**PdP_V**	**PdPrho**
10	56.93	45.03	56.63
25	**99.8**	**77.8**	**99.3**
50	99.96	99.96	100

7.2. Empirical Powers (N=16)

Table 7.2 portrays the bootstrap empirical powers of three best performing tests in which two tests are residual based, that is, PdP_V, PdPrho and one test is average weighted symmetric based, that is, PAWS. In this case, "Data-B" is taken which consists of sixteen countries (cross sections) and three time dimensions, that is, 10, 25 and 50. When time dimension is 10 then bootstrap empirical power of average weighted symmetric based test, that is, PAWS is 77.7%. Whereas bootstrap empirical power of residual based test, that is, PdP_V is around 60% and empirical power of other residual based test, that is,75.66% at this time dimension. When time dimension increases from 10 to 25 then bootstrap empirical power of all these three best performing tests increase and converge to 100%. Also, bootstrap empirical power of these three tests, that is, two are residual based and one test is average weighted symmetric based converge to 100% when time dimension increase from 25 to 50. Which is also again the similar conclusion which is drawn in Chapter 6, that is, time dimension increase power of test is also increased.

Table 7.2: Bootstrap Empirical power for N=16

N=16			
T	PAWS_V	PdP_V	PdPrho
10	77.7	59.13	75.66
25	99.96	99.86	99.96
50	100	100	100

From careful inspection of both Tables 7.1 and Table 7.2, it is evident that with increase in cross sectional dimension empirical powers of test increase. It is observed that the average weighted symmetric based test, that is, PAWS has bootstrap empirical power 56% at time dimension 10 and cross sectional dimension is 8. However, when cross sectional dimension is 16 then average weighted symmetric based test, that is, PAWS has bootstrap empirical power 77% at time dimension 10, same effect of cross sectional dimension enlarge happened with the bootstrap empirical power of other two residual based tests, that is, PdP_V and PdPrho in the Table 7.1 and Table 7.2.

7.3. Conclusion

In this chapter, bootstrap empirical power of three best tests have been compared in which two tests are parametric residual based , that is, PdP_V and PdPrho and third test is average weighted symmetric based test, that is, PAWS. The performance of these three tests is evaluated on Fisher hypothesis and bootstrap method is used to obtain the empirical power of tests at different cross sectional and time dimensions. It is concluded that with increase in time dimension, the empirical power of tests increase as mentioned in Table 7.1 that the two parametric residual based tests, that is, PdP_V and PdPrho have bootstrap empirical power 45% and 56 %, respectively, at time dimension 10. But, when time dimension increases from 10 to 25 then these two parametric residual tests have bootstrap empirical power 77% and 99%, respectively.

The bootstrap empirical power of these two residual based tests converge to 100% at time dimension 50, a similar situation is observed for these two residual based tests in table 7.2 when time dimension is large. The average weighted symmetric based test, that is, PAWS has bootstrap empirical power 56% at time dimension 10. When time dimension increases from 10 to 25 then average weighted symmetric based test, that is, PAWS has bootstrap empirical power 99% and converge to 100% at time dimension 50, a similar situation is observed for PWS in Table 7.2 when time dimension is large. This is similar to the conclusions stated earlier in Chapter 6.

Moreover, when cross sectional dimension is increased then again empirical powers of tests also increase as has been observed in Table 7.1 and Table 7.2. In Table 7.1, when time dimension is 10 then average weighted symmetric test, that is, PAWS has 56% empirical power whereas in Table 7.2. When cross sectional dimension is 16 then PAWS has 77% bootstrap empirical power at time dimension 10. The same pattern is followed by the other two residual based tests, that is, PdP_V and PdPrho in Table 7.1 and Table 7.2. Here, when cross sectional dimension increase their bootstrap empirical powers also increase. This is again in accordance with the conclusions of Chapter 6. However, among three best tests, PdPrho and PAWS tests perform far better than the third residual based test, that is, PdP_V at small time dimensions. But, these three tests have similar performance at larger time dimensions.

Chapter: 08 Conclusions and Recommendations

In this chapter, the overall conclusions are drawn about the simulated comparison of panel cointegration tests of both types, that is, the tests having null hypothesis of no cointegration and the tests having null hypothesis of cointegration. Based on these conclusions, some useful recommendations are given for applied researchers and practitioners. At the end, some suggestions are stated for further research direction.

8.1. Conclusions

In Chapter 4, the empirical size of panel cointegration tests are assessed by using the asymptotic critical values. For this purpose, four time dimensions and four cross sectional dimensions are used. The results are displayed from Figure 4.1 to Figure 4.8, Figure 4.1 to Figure 4.4 depicts the empirical size of panel cointegration tests having null hypothesis of no cointegration. Whereas, Figure 4.5 to Figure 4.8 have portrayed the empirical size of three panel cointegration tests having null hypothesis of cointegration. In all figures, the cross sectional dimension is taken fixed and time dimension is varied.

The empirical size of tests having null of no cointegration, when asymptotic critical values are used are portrayed in the Appendix from Table 9-9 to 9-12. The empirical size of two Fisher type tests, that is, Pfadf and Pfaws remains near the assumed nominal size of 5% throughout all the time and cross sectional dimensions. The empirical size of these two tests lies in the range 4.4% to 9% for all time dimensions and cross sectional dimensions. Three parametric residual based tests, that is, PdGrho, PdPrho, and PdP_V have remained under sized throughout the time dimensions and cross sectional dimensions. These three parametric residual based tests have empirical size zero for all time dimensions and cross sectional dimensions. The empirical size of parametric residual based PADF test, has remained over the nominal size throughout the time dimensions and cross sectional dimensions. Size of parametric residual based test PADF, is approximately 100% for all time and cross sectional dimensions except when cross sectional size is very low. Empirical size of other four parametric residual based tests also have remained over nominal size for all time and cross sectional dimensions. Empirical size of these four parametric residual based tests PDF, PDFrho, PDFTstar, and PDFrhostar have lied in the range 10% to 40% throughout the time and cross sectional lengths.

The empirical size of maximum likelihood based LR test behave differently as compared to all other panel cointegration tests, as time dimension increases its empirical size is also increased. One more thing about empirical size of maximum likelihood based LR test is noted, that its empirical size has decreased as cross sectional size increases.

The empirical size of two nonparametric residual based tests, that is, PdGtnp and PdPtnp converged to nominal size of 5% as time dimension increases. The empirical size of non-parametric residual based tests, PdGtnp and PdPtnp, is far from the nominal size of 5% when time dimension is close or less than the cross sectional dimension. But, when time dimension is increased and is much greater than the cross sectional dimension then empirical size of these two non-parametric residual based tests converges to the nominal size of 5%. The decreasing pattern of empirical size of parametric residual based tests, that is, PdGtp and PdPtp are also same as the pattern of non-parametric residual based tests. Although, the empirical size of PdGtp and PdPtp tests have also converges to nominal size 5% as time dimension is increased.

In the second section of the Chapter 04, the empirical size of three tests is analyzed which have the null hypothesis of cointegration. For this, again four time dimension and four cross sectional dimensions are used and the results are displayed in Figure 4.5 to Figure 4.8. In this analysis, it is concluded that all three tests, that is, fully modified based, dynamic OLS based and OLS based tests are over sized for each case. All three tests have over rejection for all time dimensions and cross sectional dimensions and do not converge to the nominal size of 5%.

In Chapter 05, the size of both types of panel cointegration tests are evaluated by using simulated critical values. In Chapter 04, it is seen that different panel cointegration tests have different size and most of the tests have size far from the nominal size 5%. As, it is known that two or more tests are comparable if their size are equal, usually equal to nominal size of 5%. The results of size based on simulated critical values are displayed from Figure 5.1 to Figure 5.8.

Figure 5.1 to Figure 5.4 depict the results of the twenty one panel cointegration tests having null hypothesis of no cointegration. Whereas, Figure 5.5 to Figure 5.8 depict the results of three panel cointegration tests having null hypothesis of cointegration. For controlling the size of tests, four time and four cross sectional dimensions are taken. Form Figure 5.1 to Figure 5.4 time dimension is fixed and cross sectional size is varied. In the case of panel cointegration tests having null hypothesis of no cointegration the size of all the twenty tests have lied in the range of 4% to 6% for all time and cross sectional dimensions when simulated critical values are used instead of using asymptotic critical values.

A similar procedure is adopted in the case of panel cointegration tests having null hypothesis of cointegration. Figure 5.5 to Figure 5.8 have shown the size of three panel cointegration tests having null hypothesis of cointegration by using simulated critical values. These four figures have depicted that size of all these three tests lie in the range 4% to 6% throughout the time dimensions and cross sectional dimensions. In the Chapter 06, the Monte Carlo power comparisons of the panel cointegration tests are performed using stringency criterion. In this comparison, simulated critical values are used in order to get nominal size of a test equal to 5%. As, under the asymptotic critical values the size of different panel cointegration tests are different and most of the tests are over sized and few tests are under sized. So, to overcome this problem we have used simulated critical values.

Figure 6.1 to Figure 6.7 depict the results of maximum shortcomings of both types of panel cointegration tests having null hypothesis of no cointegration and the null hypothesis of cointegration. The power comparisons of tests are also explained in alternative way of table form by classifying the tests into different categories 'A' to 'F' according to their maximum shortcomings. First section deals with the power comparison using stringency criterion of tests having null of no cointegration. Whereas, the other section deals with the power comparison of test having null of cointegration using stringency criterion.

In the first section, it is observed that most of the tests' maximum shortcomings are downward trending as the time length and cross section length increases. For the sake of simplicity, tests are classified into six categories from 'A' to 'F'. A cutoff point is marked at time dimension 50 and cross sectional dimension 08. It has been observed that two tests perform best and have stringencies between 0% to 20% and three tests perform mediocre and have stringencies between 20% to 50%. Whereas, sixteen worst performing tests have stringencies above 50%.

Figure 6.1shows the performance of these two best tests, that is, parametric residual based test PdP_V and average weighted symmetric based test PAWS at different time and cross sectional dimensions. In Figure 6.1, four cases are discussed by varying the cross sectional dimension, that is, N = 2, 8, 16, 32. Figure 6.1 depicts that both of these two tests perform worst when time dimension is 10 in all four cases of cross sectional dimension. When cross sectional dimension is 2 then residual based test, that is, PdP_V performs better and average weighted symmetric based test, that is, PAWS performs mediocre at time dimension 25. These two tests PdP_V and PAWS perform mediocre at time dimension 50 and onward. When, cross sectional dimension increases from 02 to 08 then these two tests perform worst at time dimension 25 and less than 25. But these tests improve their performance when time increase and, hence at time dimension 50 these two tests perform best. When cross sectional dimension is more increased and become 16 then these two residual based tests and average weighted symmetric based test perform mediocre at time dimension 25.,However,they perform best at time dimension 50 and onwards. When cross sectional dimension is 32 then parametric residual based test, that is, PdP_V and average weighted symmetric based test PAWS perform best at time dimension 25 and onwards. Hence, it is concluded that when cross sectional dimension increases the power of these two tests also increases.

Figure 6.2 portrays the stringencies of three mediocre performing tests in which two tests are parametric residual based, that is, PdPrho, PdGrho and one test is Fisher type test, that is, Pfaws. The performance of these three tests are compared and it is concluded that residual based test PdPrho perform better than the other two tests. In Figure 6.2, four cases are discussed in which cross sectional dimension vary in each case. Moreover, all these three tests perform worst at time dimension 10 in all four cases. When cross sectional dimension is 02, all these three tests perform worst at time dimension 25 and less than 25 and at time dimension 50 only residual based test PdPrho performs mediocre. Whereas, at time dimension 100 all these three tests perform mediocre. When cross sectional dimension increase from 02 to 08, then these three tests perform worst at time dimension 25 and less than 25. However, these three tests perform mediocre at time dimension 50 and these three tests perform best at time dimension 100. When cross sectional dimension increase from 08 to 16 then again, all these three tests perform worst at time dimension 25 and less than 25 but the parametric residual based test PdPrho and Fisher type test perform best at time dimension 50 and onwards. The third test, that is, parametric residual based test PdGrho performs mediocre at time dimension 50 and best at time dimension 100. When cross sectional dimension is more increased and it becomes 32, then parametric residual based test PdPrho and Fisher type test Pfaws perform mediocre at time dimension 25. Whereas, the third parametric residual based test perform worst at time dimension 25 or less than 25. All three tests perform best at time dimension 50 and onwards. In Figure 6.2, it is noted that not only the increase in time dimension effect the power of tests but also the power of tests is impacted by the cross sectional dimension.

Figure 6.3 displays the stringencies of sixteen worst performing tests according to our benchmark, that is, T=50 and N=08. From these sixteen worst performing tests, two tests are nonparametric residual based, that is, PdPtnp, PdGtnp, one Maximum Likelihood based test, that is, LR, four error correction based tests, that is, W_Pa, W_Pt, W_Ga, W_Gt, one Fisher type test, that is, Pfadf, and eight parametric residual based test, that is, PDFrho, PDFrhostar, PDFT, PDFTstar, PADF, PhZt, PdPtp, PdGtp. In Figure 6.3, four cases are discussed in which cross sectional dimension varies and time dimension remains fix in all four cases, that is, T=10, 25, 50, 100.

All these sixteen different types of tests perform worst when time dimension is less than or equal to 25 throughout the cross sectional dimensions. Also, all these sixteen tests perform worst throughout the time dimensions when cross sectional dimension is 02 . Three parametric residual based tests, that is, PhZt, PdPtp, PdGtp, two non parametric residual based tests, that is, PdPtnp, PdGtnp and one error correct based test, that is, W_Pa perform best at time dimension 100 for cross sectional dimension 08. When cross sectional dimension is 16, then eight tests perform mediocre in which seven tests are parametric or nonparametric residual based and one error correction based test. While, eight tests perform worst at time dimension 50. When time is increased from 50, then most of the tests improve their performance. Ten tests perform best at time dimension 100, five tests perform mediocre and one test, that is, Maximum Likelihood based test LR still performs worst. When cross sectional dimension is 32 and time dimension is 50 then nine tests perform best. From these nine, seven tests are parametric or nonparametric residual based, one error correction based, and one is Fisher type test. While, four tests perform mediocre in which three are parametric residual based tests and one is error correction based test. Whereas, three tests perform worst in which two are error correction based tests and one is maximum likelihood based test. At time dimension 100 all these tests perform best except maximum likelihood based test, that is, LR.

Further, the conclusions of power comparison of three panel cointegration tests having null hypothesis of cointegration are drawn using stringency criterion. Figure 6.4 to Figure 6.7 depict maximum shortcomings of these three types of tests, that is, Fully Modified Based (LM_FMOLS), Dynamic OLS based (LM-DOLS) and OLS based (LM-OLS) tests and also these results have been explained in Table 6.5 to Table 6.8. For the power comparison analysis of these three tests, the same four time dimensions and four cross sectional dimensions are used.

Figure 6.4 clearly shows that Fully Modified Based (LM_FMOLS) test has worst performance as its maximum shortcomings lie between 70% to 80% throughout the time and cross sectional dimensions. Whereas, the OLS based (LM-OLS) test has performed better as compared to Dynamic OLS based (LM-DOLS). Although, the curves of maximum shortcoming of both these tests have a decreasing pattern. When time dimension is 10 then LM-DOLS and LM-OLS tests have maximum shortcomings around 70% and 30% respectively. When time dimension is increased and become 25 then maximum shortcomings become around 40% and 16% respectively. But when time dimension is more increased, then the LM-DOLS has improved more but maximum shortcoming of LM-OLS remains same around 15%. When time dimension is 100 then the maximum shortcomings of LM-DOLS and LM-OLS tests are equal around 15%. From Figure 6.5, it is observed that LM-FMOLS test still performs worst in this case, and maximum shortcoming of this tests remains between 60% to 90% throughout the time dimensions. Whereas, the maximum shortcomings of other two tests LM-DOLS and LM-

OLS have decreasing pattern. When time dimension is 10, then maximum shortcomings are 20% and 70% respectively. But when time dimension is 50 then maximum shortcomings are 20 % and 10% respectively. Moreover, maximum shortcomings of these two tests LM-DOLS and LM-OLS have been approximately equal and less than 5% when time dimension is 100.

From Figure 6.6, the maximum shortcomings behavior of these three tests have remained same as has been behaved in Figure 6.5. In Figure 6.5, it is seen that LM-FMOLS test has still performed worst and its maximum shortcomings have lied in the range 45% to 90% throughout the time dimensions. The other two tests, that is,, LM-DOLS and LM-OLS have performed not well when time dimension is 10. But when time dimension is increased and it becomes 25 then both tests have improved and their maximum shortcomings become 47% and 13% respectively. When time dimension is 100 then maximum shortcomings of LM-DOLS and LM-OLS tests have been zero. From Figure 6.7, it is concluded that LM-FMOLS test has performed worst and its maximum shortcomings range is 30% to 94%. LM-DOLS has also performed worst when time dimension is less than 50. LM-OLS test has performed not well when time dimension is 10 and 25 but has performed better relatively and is the most stringent test as compared to other two tests that is, LM-FMOLS and LM-DOLS.

In Chapter 07, bootstrap empirical powers of three best tests are compared using real data in which two tests are parametric residual based tests, that is, PdP_V and PdPrho and third test is average weighted symmetric based test, that is, PAWS.

The performance of these three tests based on Fisher hypothesis are evaluated using the bootstrap method to obtain the empirical power of tests at different cross sectional and time dimensions. It is concluded that with increase in time dimension, the empirical power of tests increases as mentioned in Table 7.1. Here, two parametric residual based tests, that is, . PdP_V and PdPrho have bootstrap empirical power 45% and 56% respectively at time dimension 10. But when time dimension increases from 10 to 25 then these two residual based tests have bootstrap empirical power 77% and 99% respectively. While, bootstrap empirical power of these two residual based tests converge to 100% at time dimension 50. Similar results are obtained for two parametric residual based tests in Table 7.2. In Table 7.2, the empirical power of these two residual based tests, that is,, PdP_V and PdPrho are 59% and 77% respectively at time dimension 10. But when time dimension is 25 then these two residual based tests have 100% empirical power.

Third average weighted symmetric based test, i.e PAWS has bootstrap empirical power 56% at time dimension 10 in Table 7.1. But when time dimension increases from 10 to 25, then average weighted symmetric based test, i. PAWS has bootstrap empirical power 99% and converge to 100% at time dimension 50. Similar results are obtained for the average weighted symmetric test, i.e PWS in Table 7.2. This is very similar to conclusions stated earlier in Chapter 06. Moreover, when cross sectional dimension is increased then again empirical power of tests also increases as depicted from Table 7.1 and Table 7.2.

However, from the three tests, two tests in which one is residual based test, that is,, PdPrho and other is average weighted symmetric based test, that is,, PAWS perform far better than the third residual based test, that is,, PdP_V at small time dimensions. But, these three tests have similar performance at larger time dimensions. Overall, these three tests perform best and have high bootstrap empirical powers using real data.

Finally, above whole conclusion is summarized about the empirical size, power of tests, and bootstrap empirical power of first best three tests. The only two Fisher type tests having stable empirical size throughout the time and cross-sectional dimensions, i.e Pfaws, Pfadf, when asymptotic critical values are used. Whereas, residual based tests, that is,, PdP_V and PdPrho have zero empirical size throughout the time and cross-sectional dimensions when asymptotic critical values are used. Empirical size of maximum likelihood based test LR is increasing when time dimension increases throughout cross sectional dimensions. All other tests are oversized throughout the time and cross sectional dimensions when asymptotic critical values are used. The OLS based test, i.e LM-OLS, dynamic OLS based test, i.e LM-DOLS and Fully modified OLS based test, i.e LM-FMOLS have remained over sized throughout the time and cross sectional dimensions when asymptotic critical values are used. The size of all tests having null hypothesis of no cointegration and tests having null hypothesis of cointegration have remained around the nominal size 5% when simulated critical values are used instead of asymptotic critical values.

Based on power comparison, residual based test, i.e PdP_V and average weighted symmetric test, i. PAWS have outperformed all other tests except at time dimension 10. Another better performer is residual based test PdPrho.

Maximum likelihood based test, that is, LR and error correction based test, i. W_Gt performed worst throughout the time and cross section dimensions. Note that all tests have performed worst when time dimension is 10 for all cross sections. Similarly, when cross section dimension is 02, all tests performed worst throughout the time dimensions. Most of the tests outperformed when time dimension is 100 and cross-sectional dimension is greater than or equal to 8. Similarly, all the tests performed worst when time dimension is 25 and cross section dimension is less than or equal to 8. But, when cross section dimension increases and become 16 then residual based test, that is, PdP_V and average weighted symmetric test, i.e PAWS performed mediocre. Whereas, all other performed worst, at this stage of time dimension, i.e 25.

When cross sectional size more increases and becomes 32 then residual based test, that is,, PdP_V and average weighted symmetric test, that is,, PAWS performed best. Four residual based tests, that is,, PdPrho, PdPtp, PDFrho, PDFrhostar and one Fisher type test, that is,, Pfaws performed mediocre while all other performed worst at this stage of time and cross-sectional dimension. When time dimension is 50 and cross sectional dimension is 8 then residual based test, that is,, PdP_V and average weighted symmetric test, that is,, PAWS performed best. Three other tests performed mediocre in which two tests are residual based, that is,, PdGrho, PdPrho and one test is Fisher type test, that is,, Pfaws. Whereas, all tests perform worst. When cross sectional dimension increases at this stage of time dimension and becomes 16 then residual based tests, that is,, PdP_V, PdPrho and average weighted symmetric based test, that is,, PAWS outperformed. Six parametric residual based tests, that is,, PdPtp, PdGrho, PdGtp, PhZt, PDFrho, PDFrhostar, two nonparametric residual based tests, that is,, PdPtnp, PdGtnp, one Fisher type test, that is,, Pfaws and one error correction based test, that is,, W_Pa performed mediocre. Whereas, the remaining eight tests performed worst. Similarly, when cross-sectional dimension increases and becomes 32, at this time dimension 50 then maximum likelihood based test, that is,, LR and two error correction based tests W_ Pt, W_Gt performed worst. Seven different types of tests in which parametric residual based tests, that is,, PDFtstar, PADF, PDFrhostar, error correction based tests, that is,, W_Ga, W_Pa, nonparametric residual based test, that is,, PdGtnp, and Fisher type test Pfadf performed mediocre. Whereas, the remaining eleven different types of tests in which residual based tests, that is,, PdP_V, PdPrho, PdPtp, PdGrho, PdGtp, PhZt, PDFrho, PDFrhostar average weighted symmetric based test, that is,, PAWS, Fisher type test, that is,, Pfaws and nonparametric residual based test, that is,, PdPtnp performed best equally.

When time dimension is 100 and cross sectional dimension is 8 then eleven different types of tests, residual based tests, that is,, PdP_V, PdPrho, PdGrho, PhZt, PdPtp, PdGtp, nonparametric residual based tests, that is,, PdGtnp, PdPtnp, error correction based test, that is,, W_Pa, average weighted symmetric based test, that is,, PAWS and Fisher type test, that is,, Pfaws performed best. Five different types of tests performed mediocre in which three tests are residual based, that is,, PDFrho, PDFrhostar, PADF, one Fisher type test, that is,, Pfadf, and one error correction based test, that is,, W_Gt. Whereas, five tests performed worst in which one test is maximum likelihood based, that is,, LR, two tests are error correction based, that is,, W_Ga, W_Pt and two tests are residual based, that is,, PDFT, PDFtstar. When cross sectional dimension increases and becomes 16 and time dimension is 100 then only maximum likelihood based test, that is,, LR performed worst. Six different types of tests performed mediocre in which three tests are residual based, that is,, PDFT, PDFtstar, PDFrhostar, one Fisher type test, that is,, Pfadf, and two tests are error correction based tests W_Pt, W_Gt. Whereas, all other tests performed best at this stage of time and cross-sectional dimension. When cross sectional dimension is 32 and time dimension is 100 then only maximum likelihood based test, that is,, LR performed worst and error correction based test, that is,, W_Gt performed mediocre while all other tests performed equally best.

Overall, now summarizing the other type of tests having null of cointegration. The OLS based test, that is,, LM-OLS performed mediocre throughout the time dimension when cross sectional dimension is 2. LM-OLS test performed best when time dimension is greater than or equal to 50 throughout the cross sectional dimensions. Fully Modified OLS based test, that is,, LM-FMOLS performed worst throughout the time and cross-sectional dimensions except when time dimension is 100 and cross-sectional dimension is greater than or equal to 16.,At this stage it performed mediocre. Dynamic OLS based test, that is,, LM-DOLS performed worst throughout the cross sections when time dimension is 10. LM-DOLS outperformed when time dimension is 100 and cross sectional dimension lies in the range 8 to 16. But it outperformed when time dimension is 50 and cross sectional dimension is greater than or equal to 32.

Finally, bootstrap empirical power of best three tests are summarized which is concluded from our power analysis. This study concludes from power analysis using simulated data that the three tests in which two tests are residual based, that is,, PdP_V, PdPrho and one average weighted symmetric based test, that is,, PAWS tests performed best as compared to other tests. It is observed that these three tests also perform best on real data, that is,, based on Fisher hypothesis and these three tests have also high bootstrap empirical powers for the real data.

8.2. Recommendations

As it has been concluded from analysis that when asymptotic critical values are used for assessing the size of test of panel cointegration tests then most of the tests have unstable size except the two Fisher type tests, that is,, Pfaws and Pfadf throughout the time and cross sectional dimensions. Whereas, the size of tests has been stable for all tests when simulated critical values are used. Keeping in view our analysis, it is strongly recommended to use simulated critical values instead of using asymptotic critical values. For the other type of panel cointegration tests which have null hypothesis of cointegration, all three tests in which Fully Modified OLS based test, that is,, LM-FMOLS, Dynamic OLS based test, that is,, LM-DOLS and OLS based test, that is,, LM-OLS have been over sized when asymptotic critical values are used. Whereas, all three tests have stable size when simulated critical values are used. So, it is also recommended to use simulated critical values for panel tests having null hypothesis of cointegration.

From Monte Carlo simulation power comparison, it has concluded that residual based test, that is,, PdP_V and average weighted symmetric test, that is,, PAWS are the most stringent tests and performed best as compared to all other tests. So, it is recommended that these two tests should be used for the detection of long run relationship. Beside these two tests, three tests, that is,, two are residual based PdPrho, PdGrho and one is Fisher type test, that is,, Pfaws have also performed better as compared to remaining tests. In this regard, it is also recommended to use these three tests for the detection of cointegration. From the analysis of this study, it is also concluded that when time and cross section dimensions are very small and less than 25 then all of the tests performed worst except parametric residual based test, that is,, PdP_V and average weighted symmetric test, that is,, PAWS. These tests have performed a little bit better as compared to all other tests at small time and cross sectional dimensions. Hence, for the detection of cointegration, it is suggested that time and cross sectional dimension should be appropriate. Based on analysis, for the detection of cointegration it is recommended to use $T \times N \geq 800$ at least dimension for parametric/nonparametric residual based tests and average weighted symmetric based tests. Whereas for error correction based tests and maximum likelihood based tests a high time and cross sectional dimensions should be needed for detection of long run relationship. Although, it has not a theoretical background. It is also strongly recommended that the error correction based Westerlund tests should be used when time and cross section dimensions are large.

Moreover, it is recommended to the practitioners that when time dimension is 100 and cross sectional is greater than or equal to 8, then seven parametric residual based tests, that is,, PdP_V, PdPrho, PdPtp, PdGrho, PdGtp, PhZt, PADF, two nonparametric residual based tests, that is,, PdPtnp, PdGtnp, one error correction based test W_Pa, one average weighted symmetric based test, that is,, PAWS and one Fisher type test, that is,, Pfaws are used. Because, all equally performed best at this time and cross sectional dimension. From tests with null of cointegration, OLS based test, that is,, LM-OLS is best test as compared to other tests and the Fully Modified OLS based test, that is,, LM-FMOLS performed worst throughout the time and cross sectional dimensions. Although, all three tests performed worst when time and cross sectional dimension is small and less than 10 but when time dimension is 50 then LM-OLS and LM-DOLS have performed equally. So, it is recommended to the researcher to use LM-OLS and LM-DOLS panel cointegration tests for the detection of cointegration. From the evaluation of performance of three better performing cointegration tests (average weighted symmetric based test, that is,, PAWS, and residual based tests, that is,, PdP_V, PdPrho) on basis of Fisher hypothesis, it is recommended that these three tests may be used for real data investigations.

8.3. Directions for Future Research

In this study, size and power of panel cointegration tests are compared by using simple DGP of two regressors and heterogeneous panel considering the cross sectional independence. The size and power of existing panel cointegration tests can be analyzed by using more than one DGPs with respect to cross sectional dependence, heterogeneity, number of regressors, more than one cointegrating vectors etc. In this study, it is observed that still there is huge gap between the power of point optimal test and existing cointegration tests, when time and cross sectional dimension are small. There is a need to develop a new cointegration test which has an appropriate power when time and cross sectional dimension are small and has the ability to detect the genuine long-run relationship.

Appendix

Table A-1: Empirical Size of Tests having Null Hypothesis of No Cointegration using Asymptotic Critical values N=02

Tests \ SS	10	25	50	100
PDFT	25.8	19.65	17.75	15.35
PDFrho	25.2	30.9	31.8	33.1
PDFTstar	23.7	19.85	17.15	17.3
PDFrhostar	27.8	33.9	33.4	35.75
PADF	77.7	65.5	55.7	46.6
PdGtnp	25.2	13.5	10.7	9.8
PdGtp	72.4	21.4	15.5	10.4
PdGrho	0	0	0	0
PdPtnp	27.4	11.6	7.8	5.9
PdPtp	70.3	29	16.5	9.8
PdP_V	0	0	0	0
PdPrho	0	0	0	0
PhZt	40.3	18.3	15.7	11.9
PAWS	0.6	3.3	3.8	4.5
Pfadf	5.3	4.4	4.4	5.1
Pfaws	5	4.1	5.8	4.7
LR	31.1	73.7	80.2	80.7

Table A-02: Empirical Size of Tests having Null Hypothesis of No Cointegration using Asymptotic Critical values N=08

Tests \ SS	10	25	50	100
PDFT	19.8	15.95	12.35	12.8
PDFrho	17.8	21	19.3	19.9
PDFTstar	24.9	16.85	15	13.8
PDFrhostar	28.8	24.8	23.5	22.9
PADF	100	99.85	99.8	99.65
PdGtnp	39.4	16.7	10.7	9.6
PdGtp	98.5	37.8	20.8	14.5
PdGrho	0	0	0	0
PdPtnp	41.9	12.4	9	7.1
PdPtp	96.2	43.6	19.2	14.4
PdP_V	0	0	0	0
PdPrho	0.1	0	0	0
PhZt	73.2	34.4	23.7	15.8
PAWS	14.9	6.7	7.3	7.3
Pfadf	5.7	5.4	6.6	4.8
Pfaws	5.6	5.7	5	5
LR	3.6	50.7	57.9	58.5

Table A-03: Empirical Size of Tests having Null Hypothesis of No Cointegration using Asymptotic Critical values N=16

Tests \ SS	10	25	50	100
PDFT	22.8	13.3	11.05	10.1
PDFrho	16.1	14.2	15.4	14.75
PDFTstar	29.3	17.7	14.1	12.95
PDFrhostar	31.8	24.35	19.9	18.6
PADF	100	100	100	100
PdGtnp	55.4	19.6	11.7	10.4
PdGtp	100	50.1	23.4	14.6
PdGrho	0	0	0	0
PdPtnp	61.4	14.3	11	7.5
PdPtp	99.4	52	28.1	12.7
PdP_V	0	0	0	0
PdPrho	0.1	0	0	0
PhZt	91.8	53.7	31.1	19.9
PAWS	36.7	13	10	9.9
Pfadf	5.6	4.9	5.5	6.3
Pfaws	7.3	5.4	5.5	5.5
LR	0.1	23.7	33.3	36.9

Table A-04: Empirical Size of Tests having Null Hypothesis of No Cointegration using Asymptotic Critical values N=32

Tests \ SS	10	25	50	100
PDFT	26.5	14.6	11.15	9.65
PDFrho	18.8	14.65	13.2	13
PDFTstar	34.6	20.15	15.6	12.45
PDFrhostar	37.6	27.1	20.05	16.1
PADF	100	100	100	100
PdGtnp	71	22.4	15.3	10.4
PdGtp	100	70.7	30.4	18.4
PdGrho	0	0	0	0
PdPtnp	83	19.1	11.1	7.9
PdPtp	100	72	31	15.7
PdP_V	0	0	0	0
PdPrho	0.3	0	0	0
PhZt	99.3	71.9	44.1	26
PAWS	71.3	25.5	17.8	17.5
Pfadf	7.7	7.3	5.7	9
Pfaws	8.2	8.8	7	8
LR	0	5.4	10.3	12.5

Table A-05 Empirical Size of Tests having Null Hypothesis of Cointegration using Asymptotic Critical values N= 02, 08, 16, 32

N=2					N=08				
Tests \ SS	10	25	50	100	Tests \ SS	10	25	50	100
LM_OLS	77.7	65.5	55.7	41.6	LM-OLS	90	83.6	78.9	57.3
LM_DOLS	83.5	76.3	54.8	46	LM-DOLS	94.4	88.7	79.4	65.4
LM_FMOLS	73.2	75.9	82.2	90	LM-FMOLS	84.5	93.2	95	100
N=16					**N=32**				
Tests \ SS	10	25	50	100	Tests \ SS	10	25	50	100
LM_OLS	100	90.3	81.1	76	LM_OLS	100	100	100	100
LM_DOLS	100	96.8	91.3	85.9	LM_DOLS	100	100	100	100
LM_FMOLS	100	100	100	100	LM_FMOLS	100	100	100	100

Table A-06 Maximum Shortcomings of Tests having Null Hypothesis of No Cointegration N=02

Tests \ SS	10	25	50	100
PDFT	75.96	82.73333	80.22667	89.76667
PDFrho	71.4	78.16667	74.58667	86.26667
PDFTstar	80.94	82.66667	82.82667	91.5
PDFrhostar	89.3	83.76667	79.70667	91.03333
PADF	80.34	79.2	73.22667	87.86667
PhZt	72.3	74.3	58.26667	59.6
PdGtnp	69.2	73.8	61.06667	61.06667
PdGtp	89.25	73.06667	63.66667	58.86667
PdGrho	66.6	64.3	54.66667	48.86667*
PdPtnp	67.2	71.4	55.36667	57.16667
PdPtp	88.25	71.56667	59.66667	57.56667
PdP_V	82.55	55.73333	38.56667*	23.4*
PdPrho	63.3	63.36667	49.76667*	42.96667*
PAWS	63.8	49.03333*	45*	32.6*
Pfadf	82.65	82.83333	78.46667	73.4
Pfaws	67.5	57.23333	51.96667	43.9
W_Gt	91.35	91.46667	86.86667	83.2
W_Ga	89.05	87.26667	75.46667	69.2
W_Pt	88.15	91.5	86.16667	87.2
W_Pa	86.2	82.83333	72.56667	68.9
LR	87.35	92.81667	92.71667	89.8

Note: "**" represents the best performing tests when test has maximum shortcomings less than 20%, "*" represents mediocre performing test when test has maximum shortcomings between 20% to 50%, maximum shortcoming without "*" represents worst performing tests when maximum shortcoming is greater than 50%.

Table A-07 Maximum Shortcomings of Tests having Null Hypothesis of No Cointegration N=08

Tests \ SS	10	25	50	100
PDFT	83.15	83.5	76.16667	56.36667
PDFrho	72.79	75.23333	62.56667	40.63333*
PDFTstar	84.37	82.96667	79.43333	54.63333
PDFrhostar	81.55	77.33333	69.06667	42.7*
PADF	88.47	82.4	72.46667	22.7**
PhZt	85.31667	81.6	62.1	4.833333**
PdGtnp	87.55	84.13333	66.66667	5.6**
PdGtp	96.5	83.03333	66.3	5.133333**
PdGrho	75.75	77.83333	47.56667*	1.733333**
PdPtnp	83.75	77.53333	56.66667	3.333333**
PdPtp	91.8	76.5	58	2.9**
PdP_V	78.15	58.2	10.5**	0**
PdPrho	72.9	71.53333	38.6*	0.333333**
PAWS	65.3	56.2	12.4**	0**
Pfadf	88.05	84.8	81	43.4*
Pfaws	83.25	73.6	44.8*	1.2**
W_Gt	95.4	87.7	91.2	55.9
W_Ga	93.15	89.2	71.5	23.6*
W_Pt	94.5	90	83.7	60.5
W_Pa	92.65	78.4	63.6	13.9**
LR	91.8	93.75	93.8	75.95

Note: "**" represents the best performing tests when test has maximum shortcomings less than 20%, "*" represents mediocre performing test when test has maximum shortcomings between 20% to 50%, maximum shortcoming without "*" represents worst performing tests when maximum shortcoming is greater than 50%.

Table A-08 Maximum Shortcomings of Tests having Null Hypothesis of No Cointegration N=16

Tests \ SS	10	25	50	100
PDFT	80.93	75.5	53.2	21.93333*
PDFrho	68.31	61.23333	30.23333*	8.933333**
PDFTstar	83.21	82	60.46667	29.9*
PDFrhostar	74.17	66.66667	37.56667*	14.43333**
PADF	88.67	84.73333	57.23333	4.7**
PhZt	86.65	77.5	33.9*	0**
PdGtnp	85.35	82.4	37.8*	0**
PdGtp	87.9	78.95	34.7*	0**
PdGrho	77.5	68.65	21.65*	0**
PdPtnp	81.05	71.7	30.5*	0**
PdPtp	78.45	68.05	23*	0**
PdP_V	75.55	29.2*	0.25**	0**
PdPrho	73.3	60.4	7.55**	0**
PAWS	57.15	39.2*	2.8**	0**
Pfadf	87	87	78.6	29*
Pfaws	85.55	70.4	18.8**	0**
W_Gt	94.6	87	81.2	39.9*
W_Ga	91.65	91.45	53.45	7.05**
W_Pt	92.4	91.2	76.7	33.2*
W_Pa	92.85	80.6	34.2*	1.2**
LR	94.4	94.9	92.15	75.85

Note: "**" represents the best performing tests when test has maximum shortcomings less than 20%, "*" represents mediocre performing test when test has maximum shortcomings between 20% to 50%, maximum shortcoming without "*" represents worst performing tests when maximum shortcoming is greater than 50%.

Table A-09 Maximum Shortcomings of Tests having Null Hypothesis of No Cointegration N=32

Tests \ SS	10	25	50	100
PDFT	76.36	62.86667	26.66667*	5.333333**
PDFrho	65.24	36.83333*	7.4**	0.766667**
PDFTstar	81.08	66.73333	31.83333*	8.266667**
PDFrhostar	75.38	42.66667*	11.13333**	1.766667**
PADF	88.26	78.73333	49.06667*	0.333333**
PhZt	84.8	68.5	6.8**	0**
PdGtnp	84.6	63.85	11.35**	0**
PdGtp	86.6	69.05	10**	0**
PdGrho	81.26667	52.3	2.3**	0**
PdPtnp	81.06667	53.65	2.9**	0**
PdPtp	74.7	48.6*	2.85**	0**
PdP_V	72	7.8**	0**	0**
PdPrho	76.4	31.95*	0.4**	0**
PAWS	88.1	20.3*	0**	0**
Pfadf	86.375	80.3	11.6**	3.2**
Pfaws	88.75	29.8*	9.4**	0**
W_Gt	94	86.7	75.6	11.3**
W_Ga	92.2	68.3	32.2*	0.2**
W_Pt	93.6	86.4	56.6	9**
W_Pa	92.6	64.8	13.8**	0**
LR	94.3	94.15	90.55	81.55

Note: "**" represents the best performing tests when test has maximum shortcomings less than 20%, "*" represents mediocre performing test when test has maximum shortcomings between 20% to 50%, maximum shortcoming without "*" represents worst performing tests when maximum shortcoming is greater than 50%.

Table A-10 Maximum Shortcomings of Tests having Null Hypothesis of Cointegration N=02

SS / Tests	10	25	50	100
LM_OLS	28.45*	16.45**	16.35**	15.2**
LM_DOLS	67.1	41.7*	25.1*	14.6**
LM_FMOLS	78.85	74.1	78	80

Table A-11 Maximum Shortcomings of Tests having Null Hypothesis of Cointegration N=08

SS / Tests	10	25	50	100
LM_OLS	22.5*	9**	10.7**	1.7**
LM_DOLS	69.2	44.7*	23.2*	4.1**
LM_FMOLS	89.4	82.5	77.55	58.55

Table A-12 Maximum Shortcomings of Tests having Null Hypothesis of Cointegration N=16

SS / Tests	10	25	50	100
LM_OLS	29.2*	13.25**	10.15**	0**
LM_DOLS	71.1	47.2*	25.2*	0.2**
LM_FMOLS	91.7	85.2	80.1	45.2

Note: "**" represents the best performing tests when test has maximum shortcomings less than 20%, "*" represents mediocre performing test when test has maximum shortcomings between 20% to 50%, maximum shortcoming without "*" represents worst performing tests when maximum shortcoming is greater than 50%.

Table A-13 Maximum Shortcomings of Tests having Null Hypothesis of Cointegration N=32

SS / Tests	10	25	50	100
LM_OLS	29.95	23.9*	2**	0**
LM_DOLS	76.3	48.4*	6.1**	0**
LM_FMOLS	94.4	87.2	82.3	30.2*

Note: "**" represents the best performing tests when test has maximum shortcomings less than 20%, "*" represents mediocre performing test when test has maximum shortcomings between 20% to 50%, maximum shortcoming without "*" represents worst performing tests when maximum shortcoming is greater than 50%.

References

1. Andrews, D. W. (1991). "Heteroskedasticity and autocorrelation consistent covariance matrix estimation." Econometrica: Journal of the Econometric Society: 817-858.
2. Badillo, R., C. Reverte and E. Rubio (2011). "The Fisher Effect in the EU Revisited: New Evidence using Panel Cointegration Estimation with Global Stochastic Trends." Applied Economics Letters **18**(13): 1247-1251.
3. Banerjee, A., J. Dolado and R. Mestre (1998). "Error-Correction Mechanism Tests for Cointegration in a Single-Equation Framework." Journal of time series analysis **19**(3): 267-283.
4. Breitung, J. (2005). "A Parametric Approach to the Estimation of Cointegration Vectors in Panel Data." Econometric Reviews **24**(2): 151-173.
5. Chang, Y. (2004). "Bootstrap Unit Root Tests in Panels with Cross-Sectional Dependency." Journal of econometrics **120**(2): 263-293.
6. Crowder, W. J. (2003). "Panel Estimates of the Fisher effect." manuscript, University of Texas at Arlington.
7. Dickey, D. A. and W. A. Fuller (1979). "Distribution of the Estimators for Autoregressive Time Series with a Unit Root." Journal of the American statistical association **74**(366a): 427-431.
8. 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.
9. Fisher, I. (1930). "The theory of interest." New York **43**.

10. Gonzalo, J. (1994). "Five Alternative Methods of Estimating Long-Run Equilibrium Relationships." Journal of econometrics **60**(1-2): 203-233.

11. Groen, J. J. J. and F. Kleibergen (2003). "Likelihood-Based Cointegration Analysis in Panels of Vector Error-Correction Models." Journal of Business & Economic Statistics **21**(2): 295-318.

12. Gutierrez, L. (2003). "On the Power of Panel Cointegration Tests: a Monte Carlo comparison." Economics Letters **80**(1): 105-111.

13. Harris, D. and B. Inder (1994). "oA test of the null hypothesis of cointegrationp." Nonstationary Time Series Analysis and Cointegration. Oxford University Press, Oxford.

14. Haug, A. A. (1996). "Tests for Cointegration a Monte Carlo Comparison." Journal of econometrics **71**(1-2): 89-115.

15. Hoang, N. T. (2006). "New Tests for Cointegration in Heterogeneous Panels." Center for Economic Analysis.

16. Im, K. S., M. H. Pesaran and Y. Shin (2003). "Testing for Unit Roots in Heterogeneous Panels." Journal of econometrics **115**(1): 53-74.

17. Jansson, M. (2005). "Point Optimal Tests of the Null Hypothesis of Cointegration." Journal of Econometrics **124**(1): 187-201.

18. Jansson, M. (2008). "Semiparametric Power Envelopes for Tests of the Unit Root Hypothesis." Econometrica **76**(5): 1103-1142.

19. Johansen, S. (1988). "Statistical Analysis of Cointegration Vectors." Journal of economic dynamics and control **12**(2-3): 231-254.

20. Johansen, S. (1995). Likelihood-Based Inference in Cointegrated Vector Autoregressive Models, Oxford University Press on Demand.

21. Kao, C. (1999). "Spurious Regression and Residual-Based Tests for Cointegration in Panel Data." Journal of econometrics **90**(1): 1-44.

22. Larsson, R., J. Lyhagen and M. Löthgren (2001). "Likelihood-Based Cointegration Tests in Heterogeneous Panels." The Econometrics Journal **4**(1): 109-142.

23. Levin, A. and C.-F. Lin (1993). "Unit Root Tests in Panel Data: New Results." University of California at San Diego, Economics Working Paper Series.

24. Lütkepohl, H. (2005). New Introduction to Multiple Time Series Analysis, Springer Science & Business Media.

25. Maddala, G. S. and S. Wu (1999). "A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test." Oxford Bulletin of Economics and statistics **61**(S1): 631-652.

26. McCoskey, S. and C. Kao (1998). "A Residual-Based Test of the Null of Cointegration in Panel Data." Econometric reviews **17**(1): 57-84.

27. Newey, W. K. and K. D. West (1994). "Automatic Lag Selection in Covariance Matrix Estimation." The Review of Economic Studies **61**(4): 631-653.

28. Örsal, D. D. K. (2007). Comparison of Panel Cointegration Tests, SFB 649 discussion paper.

29. Örsal, D. D. K. (2009). Essays on panel cointegration testing, Humboldt-Universität zu Berlin, Wirtschaftswissenschaftliche Fakultät.

30. Pedroni, P. (1996). "Fully modified OLS for heterogeneous cointegrated panels and the case of purchasing power parity." Manuscript, Department of Economics, Indiana University.

31. Pedroni, P. (1999). "Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors." Oxford Bulletin of Economics and statistics **61**(S1): 653-670.

32. Pedroni, P. (2004). "Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time series Tests with an Application to the PPP Hypothesis." Econometric theory **20**(3): 597-625.

33. Phillips, P. C. and B. E. Hansen (1990). "Statistical inference in Instrumental Variables Regression with I (1) Processes." The Review of Economic Studies **57**(1): 99-125.

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

35. Phillips, P. C. and P. Perron (1988). "Testing for a Unit Root in Time Series Regression." Biometrika **75**(2): 335-346.

36. Said, S. E. and D. A. Dickey (1984). "Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order." Biometrika **71**(3): 599-607.

37. Saikkonen, P. (1991). "Asymptotically Efficient Estimation of Cointegration Regressions." Econometric theory **7**(1): 1-21.

38. Saikkonen, P. (1999). "Testing Normalization and Overidentification of Cointegrating Vectors in Vector Autoregressive Processes." Econometric Reviews **18**(3): 235-257.

39. Shin, Y. (1994). "A rResidual-Based Test of the Null of Cointegration Against the Alternative of No Cointegration." Econometric theory **10**(1): 91-115.

40. Toda, H. Y. (1995). "Finite Sample Performance of Likelihood Ratio Tests for Cointegrating Ranks in Vector Autoregressions." Econometric theory **11**(5): 1015-1032.

41. Toyoshima, Y. and S. Hamori (2011). "Panel Cointegration Analysis of the Fisher effect: Evidence from the US, the UK, and Japan." Economics Bulletin **31**(3): 2674-2682.

42. Westerlund, J. (2007). "Testing for Error Correction in Panel Data." Oxford Bulletin of Economics and statistics **69**(6): 709-748.

43. Westerlund, J. (2008). "Panel Cointegration Tests of the Fisher effect." Journal of Applied Econometrics **23**(2): 193-233.

44. Wu, S. and Y. Yin (1999). Tests for Cointegration in Heterogeneous Panel: A Monte Carlo Study.", Working paper, Department of Economics, State University of New York at ….

45. Zaman, A. (1996). Statistical Foundations for Econometric Techniques, Academic Press.

46. 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.