DOKUZ EYLÜL UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

# IMPROVING COINTEGRATION TESTS UNDER STRUCTURAL BREAKS IN MULTIVARIATE GARCH MODELS

by
Berhan ÇOBAN

February, 2018
İZMİR

# IMPROVING COINTEGRATION TESTS UNDER STRUCTURAL BREAKS IN MULTIVARIATE GARCH MODELS

**A Thesis Submitted to the**
**Graduate School of Natural And Applied Sciences of Dokuz Eylül University**
**In Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Statistics**

**by**
**Berhan ÇOBAN**

**February, 2018**
**İZMİR**

## Ph.D. THESIS EXAMINATION RESULT FORM

We have read the thesis entitled "**IMPROVING COINTEGRATION TESTS UNDER STRUCTURAL BREAKS IN MULTIVARIATE GARCH MODELS**" completed by **BERHAN ÇOBAN** under supervision of **PROF. DR. ESİN FİRUZAN** and we certify that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy.

...................................................................

Prof. Dr. Esin FİRUZAN

Supervisor

...................................................................

Assoc. Prof. Dr. Hamdi EMEÇ

Thesis Committee Member

...................................................................

Assoc. Prof. Dr. A. Fırat ÖZDEMİR

Thesis Committee Member

...................................................................

Prof. Dr. Hüseyin TATLIDİL

Examining Committee Member

...................................................................

Doç. Dr. Çağdaş H. ALADAĞ

Examining Committee Member

Prof. Dr. Kadriye ERTEKİN
Director
Graduate School of Natural and Applied Sciences

## ACKNOWLEDGEMENTS

I owe appreciations to many people for their help and support which made this work possible and existing.

Above all, I would like to express my deep and sincere gratitude to my dissertation chair Prof. Dr. Esin FİRUZAN, who has been supporting my scientific career as my supervisor since 2008. If it had not been for her true mentorship and academic guidance this dissertation would not have been written. It has been a great chance for me to work with her.

I am very thankful to members of my committee who generously contributed me; namely to Doç. Dr. Hamdi EMEÇ and to Doç. Dr. A. Fırat ÖZDEMİR, for sparing their precious time to serve on my committee and giving their valuable suggestions throughout this thesis.

Special thanks to all my friends and colleagues especially I am very grateful to Dr. Şerife ÖZKAR for contribution in designing this report. I would also like to thank Dr. Hanife TAYLAN SELAMLAR and Bahar YALÇIN for sharing all their experience and time with me generously.

I will forever be thankful to my family, for their endless love and encouragements in whole of my life.

Last, but the most, I would like to express my great appreciation to my invaluable wife Sevil GÜLBUDAK ÇOBAN for her infinite patience, support, and understanding. Her love has always given me unlimited strength.

Berhan ÇOBAN

**IMPROVING COINTEGRATION TESTS UNDER STRUCTURAL BREAKS IN MULTIVARIATE GARCH MODELS**


## ABSTRACT

In time series, data may experience a sudden change / break in its dynamics. That is, there may be breaks in the mean, slope, trend function or some other characteristics that occur at random times. Structural changes may occur in time series due to policy changes, financial crises and natural disasters. Another destructive attribute is heteroscedasticity of error term. Econometric and financial time series are well known with a fat-tailed distribution that has the large skewness and/or excess kurtosis. The existence of the structural breaks and heteroscedastic error term may cause various problems such as biased, inconsistent estimations and poor predictions. Many studies demonstrated that heteroscedaticity and structural breaks in a cointegration relationship significantly influence the performance of cointegration tests.

This dissertation proposes a more powerful test among residual based test, the new test's name is the "RALS(2)-LM GARCH", which is suitable for GARCH effects under the level or/and slope structural breaks. The test formed using the residual based Lagrange Multiplier method. New test relieves the practitioners from conducting extra analyses by considering the structural breaks and heteroscedasticity together. The proposed tests lies in the invariance feature that the distribution does not depend on the different values of GARCH parameter in the presence of multiple level and trend-breaks. The suggested test is adaptable to multiple breaks and heteroscedatic error terms.

In simulation study, the performances of Engle-Granger (1987), Gregory-Hansen (1996), Westerlund-Edgerton LM (2007) and RALS(2)-LM GARCH are examined under structural break and GARCH effect. When the results of the simulation are examined in general, it can be argued that when there is a structural break in the series, the RALS(2)-LM GARCH and WE-LM test yielded better results. In addition, simulation result showed that GH test exhibits a liberal behavior and diverges from the nominal size while the EG test has more conservative results in different break scenarios. Furthermore, the simulation shows that the newly suggested RALS-based cointegration tests utilizing higher moment conditions exhibit substantial power gains even if the errors are GARCH distributed. Final evidence from the simulation study is that the position of the break and the type of cointegration model in the experiment design influenced the performances of the tests.

**Keywords:** Cointegration tests, structural break, heteroscedaticity, GARCH effect

**ÇOK DEĞİŞKENLİ GARCH MODELLERİNDE YAPISAL KIRILMALAR OLMASI DURUMUNDA EŞBÜTÜNLEŞME TESTLERİNİN GELİŞTİRİLMESİ**

## ÖZ

Zaman serisi verilerinde ve serinin dinamiklerinde ani değişimler / kırılmalar olabilir. Bu kırılmalar herhangi bir zamanda ortalamada, eğimde, trend fonksiyonunda ve diğer bileşenlerde ortaya çıkabilir. Yapısal kırılmalar, politika değişikliği, finansal krizler ve doğal afetler gibi nedenlerden dolayı meydana gelmektedir. Zaman serilerindeki diğer bozucu özellik ise farklı varyanslı hata terimleridir. Ekonometrik ve finansal zaman serileri genellikle ağır kuyruklu, aşırı çarpık ve basık olarak tanımlanırlar. Farklı varyanslı hata terimleri ve yapısal kırılmaların olması yanlı ve tutarsız parametre kestirimlerine ve zayıf tahminleme gibi önemli problemlere yol açarlar. Birçok çalışmada farklı varyanslığın ve yapısal kırılmaların eşbütünleşme testlerinin performansını etkilediği gösterilmiştir.

Bu tez artık temelli eşbütünleşme literatürüne, GARCH etkisi ve sabit terimde/ eğimde yapısal kırılma olduğu durumda daha güçlü bir test olan “RALS(2)-LM GARCH” testini önermektedir. Test artık tabanlı Lagranj Çarpanları yöntemini kullanmaktadır. Önerilen test farklı varyanslık ve yapısal kırılma olduğu durumda alternatif ve daha etkin bir test sunmaktadır. Yeni test kullanıcıların ek analiz yapmalarına gerek kalmadan yapısal kırılma ve GARCH etkilerini aynı anda değerlendirmektedir. Önerilen test farklı GARCH parametreleri ve yapısal kırılmalar için değişmeyen özelliklere sahiptir. Yeni geliştirilen test kesmede ve trendde birden fazla kırılmaya ve farklı varyanslı hata terimlerine uyarlanabilir.

Simulasyon çalışmasında, yapısal kırılma ve GARCH etkileri altında Engle-Granger (1987), Gregory-Hansen (1996), Westerlund-Edgerton LM (2007) ve RALS(2)-LM GARCH testlerinin performansı karşılaştırılmıştır. Genel olarak simulasyon sonuçları incelendiğinde Westerlund-Edgerton LM (2007) ve RALS(2)-LM GARCH testleri daha iyi sonuçlar vermektedir. Buna ek olarak farklı kırılma senaryolarında Engle-Granger (1987) testi tutucu/temkinli sonuçlara sahip iken Gregory-Hansen (1996) testi liberal sonuçlar sergilemektedir. Bu sonuçlara ek olarak hata terimi GARCH dağıldığında, yeni önerilen test hata terimlerinin yüksek dereceden momentlerini kullandığından önemli derecede etkinlik kazancı göstermektedir. Son olarak simülasyon çalışması kırılmanın konumunun ve eşbütünleşme model çeşidinin testlerin performansına etki ettiğini göstermiştir.

**Anahtar kelimeler:** Eşbütünleşme testleri, yapısal kırılma, farklı varyanslık, GARCH etkisi

**CONTENTS**

## LIST OF FIGURES

**Page**

## LIST OF TABLES

# CHAPTER ONE
# INTRODUCTION

## 1.1 Introduction

Time series analysis is useful technique for identifying the nature of the phenomenon representing by the sequences of observation. The aim of the time series analysis is to extrapolate the identified pattern and to predict future events. One of the multivariate time series technique is the cointegration analysis. Cointegration analysis is a method developed to reveal whether there is a long term linear correlation between time series.

Cointegration tests have been among the most important and influential tools in empirical economics and finance since the information about tests is based on over three decades ago. In time series, some risky components such as breaks and non-normal error terms may cause some problems. Structural breaks may occur in time series due to policy changes, technological improvement, financial crises and natural disasters. If there are structural breaks in the time series used in cointegration analysis, the analysis is conducted without considering the break; so the results of the test might be unreliable.

Another damaging attribute is heteroscedasticity of error term. Econometric and financial time series are well known with a fat-tailed distribution that has the large skewness or excess kurtosis. In these circumstances Ordinary Least Squares estimator is still unbiasedness and consistency, but not efficient. This time-varying volatility is also known as "conditional heteroscedasticity". Heteroscedasticity can be expressed as volatility of error term. In other words, it is the downward or upward speed changes of the variance. This changeability in error term influences the model parameter estimates. These destructive behaviors lead to periodic volatility of error term and non-normality in series.

The existence of the structural breaks and heteroscedastic error term may cause various problems such as biased and inconsistent estimation results and poor predictions. In these circumstances, estimators of residual based tests are inefficient and have reducing power.

Structural breaks and heteroscedasticity affect the performance of not only unit root tests, but also the cointegration tests. Therefore, various tests have been developed in order to determine the cointegration relationship between the time series under these destructive properties.

Suggested residual based cointegration test in this thesis originated from Lagrange Multiplier approach. LM based test has more general and flexible assumptions about error term because of its asymptotic properties. From a structural break perspective, several studies have shown that the LM tests has great advantages with compared to other tests.

To obtain more powerful estimator or/and test Residual Augmented Least Squares (RALS) can be used. Residual Augmented Least Squares (RALS) approach which utilizes information on the skewness and excess kurtosis is more efficient than classical OLS estimation.

The question addressed in this thesis is whether structural breaks and GARCH innovation term in cointegration regression can affect the behavior of cointegration tests. Thereafter generate a new test which is take into account these negative components in cointegration regression.

This dissertation proposes a more powerful test among residual based test, the new test's name is the "RALS(2)-LM GARCH", which is suitable for GARCH effects under the level or/and slope structural breaks. The newly suggested RALS-based cointegration test utilizing higher moment conditions exhibits substantial power gains when the errors are GARCH distributed. Test is an extension and combination of the WE-LM(2007) cointegration test and RALS unit root test proposed by Meng, Lee & Payne (2016). Asymptotic distribution of RALS(2)-LM GARCH cointegration test is same as that of the Generalized Method of Moments (GMM) estimator as well as the test of Hansen (1995).

Theoretical framework and operational base of suggested test follows as;

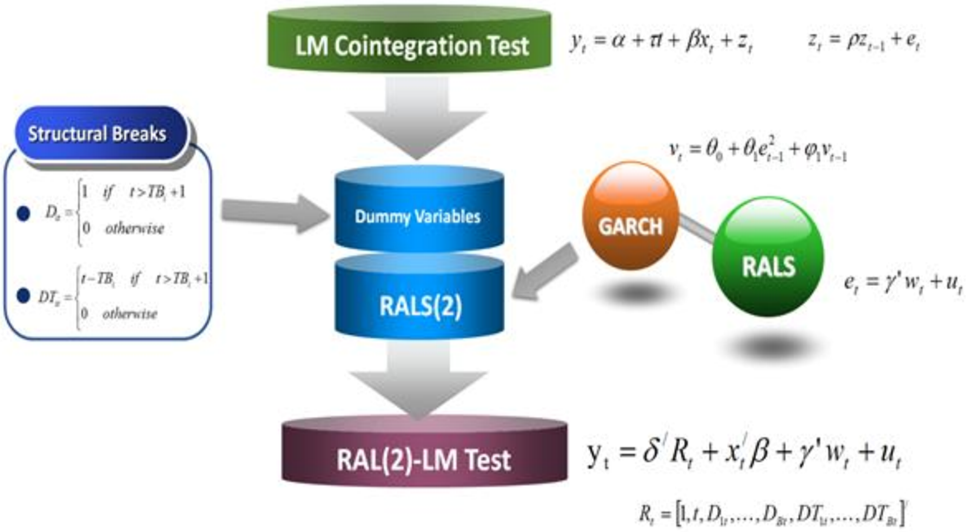


Figure 1.1 Operational and theoretical base of suggested test

## 1.2 Contributions and Constraints of Dissertation

The test formed using the residual based LM method has many approaches, which are original in many ways.

- The test offers an alternative and a more efficient solution to the tests with decreasing performances due to structural breaks and heteroscedasticity.
- The test relieves the users from conducting extra analyses by considering the structural breaks and heteroscedasticity together.
- The test provides an increase in power for different GARCH models, in addition to its proven asymptotically for the GARCH(p,q) parameters.
- The effects of the structural breaks in the heteroscedastic error terms cannot be determined clearly. In addition, changes may occur in the variance of the series before or after the breaks. In such cases, the test is more applicable and exhaustive compared to other tests.

- It is seen, in many studies in the literature, that the Johansen test, a system based test, is more powerful in cases of heteroscedasticity. This study proposes a more powerful test among residual based test, in cases where the heteroscedastic error terms a present.

When the limitations of the test are examined, the following are observed:

- The test is a residual based test. Therefore, the model combinations should be implemented for the cointegrated relation of more than two series.
- The test is valid for the error terms distributed with the GARCH model as well as linear model. For different GARCH model, the test may require adding some assumptions.
- The parameter estimations of the test are conducted with the OLS and the restricted MLE methods. This estimation is limited with the characteristics of these estimators.

The rest of thesis is organized as follows. In Chapter 2, provides a brief description about some cointegration tests such as Engle-Granger (1987), Gregory-Hansen (1996) and Westerlund-Edgerton (2007) tests. Then this chapter gives information structural breaks and heteroscedasticity. Following part contains theoretical literature review on the cointegration test under the structural breaks and heteroscedaticity. In Chapter 3, the information about Lagrange-Based test for cointegration and asymptotic properties are given. This section covers the cointegration test in the presence of one and more structural breaks. Chapter 4 gives brief information of the RALS approach. This part establishes the construction of the new test named RALS(2)-LM GARCH Cointegration Test. Also this section presents application process of new test. In Chapter 5, there is a comparison of cointegration tests with RALS(2)-LM GARCH test under the presence of structural break and heteroscedastic innovations via Monte Carlo simulation. Lastly, in Chapter 6, conclusion is stated as a final chapter.

# CHAPTER TWO

# COINTEGRATION TESTS, STRUCTURAL BREAK, GARCH(P,Q) MODEL

## 2.1 Introduction

A time series is a set of observation $x_t$, each successive value represents consecutive measurement takes at equally spaced time intervals. Time series analysis is useful technique for identifying the nature of the phenomenon representing by the sequences of observation. The aim of the time series analysis is to extrapolate the identified pattern and to predict future events.

While time series analysis may depend on single variable analysis, modeling and analysis can also be performed on more than one series together. This analyis is called the multivariate time series in the literature. One of the multivariate time series analysis is the cointegration analysis. Cointegration analysis is a method developed to reveal whether there is a long term linear correlation between time series.

Cointegration tests have been among the most important and infuential tools in empirical economics since the information about tests is based on over three decades ago. Most of the time series, especially econometric and financial series, are usually nonstationary. This implies that modeling the econometric relation between these series often relies on cointegration techniques. Currently there are two fundamental approaches to investigate cointegration relationships, system based tests and residual based tests. System based cointegration tests include tests such as Johansen-Juselius (1990) using vector autoregressive model (VAR) in order to determine long-run relationship between series; whereas the residual based tests utilize error terms of a linear combination between two non-stationary time series.

In a time series, data may experience a sudden change in its dynamics. That is, there may be changes in the mean, slope, trend function or some other characteristics that occur at random times. Structural changes may occur in time series due to policy changes, financial crises and natural disasters. If there are structural breaks in the time series used in cointegration analysis, the analysis is conducted without considering the break; the results of the test might be unreliable. Another damaging attribute is heteroscedasticity of error term. Econometric and financial time series are well known with a fat-tailed distribution that has the large skewness or excess kurtosis. These destructive behaviors lead to aperiodic volatility of error term and non-normality in series. In these circumstances, estimators of residual based tests are inefficient.

This thesis focuses on residual based cointegration tests. Since, systems-based tests are more sensitive than residual based ones. When the structural breaks are present in cointegration model, the system based model can not evaluate easily. These tests suffer from one or more of the following problems: assumption of no breaks under the null, of known break dates, of the same break point for all time series, and accounting for only breaks in the intercept of the trend functions of time series.

The most common cointegration tests are based on the assumption that the individual variables have unit root processes. Unit root test is applied on the error terms obtained from the cointegration model, in order to determine whether the error terms are stationary or not. If the residuals are stationary (or has no unit root), i.e. $I(0)$, it means that the variables are cointegrated and have a long-run or equilibrium relationship.

Structural breaks and heteroscedasticity affect the performance of not only unit root tests, but also the cointegration tests performance. Therefore, various tests have been developed in order to determine the cointegration relationship between the time series under these destructive properties.

This chapter gives information about the Engle-Granger (1987), Gregory-Hansen (1996) and Westerlund-Edgerton (2007) tests, which are classical cointegration tests used widely. Structural break and heteroscedasticity, which affect the performance of the tests, are examined. Lastly, a theoretical literature review on the cointegration test, in presence of breaks, and volatility, are given.

## 2.2 General Overview on Cointegration Tests

### *2.2.1 Engle-Granger Cointegration Test*

Single-equation regression or residuals-based tests for cointegration, proposed by Engle & Granger (1987) have been quite popular in applied work. The basic approach in Engle-Granger method is that the error terms of a linear combination between two non-stationary time series have the property of stationarity. Single equation cointegration model is as follows,

$$y_t = \alpha + x_t'\beta + z_t \tag{2.1}$$

A general model that can be built between two series are presented as in equation (2.1). In this model, the $y_t$ demonstrates the dependent variable, $x_t$, the independent variable and $z_t$ the error term, respectively. In order that the variables in the model to be cointegrated, it is both assumed that the difference of both variables are obtained ($I(1)$ distributed) and the error term is non-differenced ($I(0)$ distributed). In other words, the error term is $z_t$ and $u_t \sim IN(0, \sigma^2)$ in equation (2.2). Test statistic is the *t-ratio* associated with the ordinary least squares (OLS) estimate of $\rho$ in the following equation:

$$\Delta z_t = \phi z_{t-1} + u_t \tag{2.2}$$

In order to determine the existence of the linear correlation between the series Engle & Granger (1987) proposed a procedure comprising of two steps. According to this procedure, first a linear equation (ordinary least squares, OLS) is built and the parameter estimations are obtained by using the least square method. As for the second step, the unit root test is applied on the error terms obtained from the model. In order to determine whether the error terms are stationary or not, the well-known Dickey-Fuller test is used.

The hypotheses for this test are;

$H_0 : \phi = 0$ means that $u_t$ has unit root. In other words, $x_t$ and $y_t$ are not cointegrated.
$H_1 : \phi \neq 0$ means that $u_t$ has not unit root. In other words, $x_t$ and $y_t$ are cointegrated.

$t - ratio = \dfrac{\phi}{se_(\phi)}$ is in the form of test statistics for these hypotheses. From now on, in this thesis this ratio is called as $\tau$. The critical values for this test statistics are compared to the values produced by Dickey-Fuller instead of the standard *t* table. A number of tests for cointegration have been proposed since then Engle-Granger cointegration test.

#### *2.2.2 Gregory-Hansen Cointegration Test (1996)*

The simulation results of Gregory & Hansen (1996) revealed that the power of the Engle-Granger (1987) test is significantly reduced when there is a break in the long run relationship. To solve this problem, Gregory-Hansen (1996) develop the Engle-Granger test to allow for breaks either in the intercept or the intercept and trend of the cointegration model at an unknown time.

Gregory-Hansen (1996) considered three models allowing structural change in the cointegrating relationship, being specified and denoted as follows:

*Model C: Level Shift*

$$y_t = \alpha + \alpha_1 D_{1t} + x_t' \beta + z_t, \ \ t = 1, ..., T \tag{2.3}$$

*Model C/T: Level Shift with Trend*

$$y_t = \alpha + \tau t + \alpha_1 D_{1t} + x_t' \beta + z_t, \ \ t = 1, ..., T \tag{2.4}$$

*Model C/S: Regime Shift (C/S)*

$$y_t = \alpha + \alpha_1 D_{1t} + x_t'\beta + x_t'\beta_1 DT_{1t} + z_t, \;\; t = 1, ..., T \tag{2.5}$$

where $\alpha_1, \beta_1$ represents the change after the break intercept and slope respectively, and $D_{1t}$, $DT_{1t}$ are dummy variables. The dummy variables $(D_{1t}, DT_{1t})$ which are included in the models to determine the structural breaks can be defined as below:

$$D_{1t} = \begin{cases} 1 & if \;\; t \geq TB + 1 \\ 0 & otherwise \end{cases} \quad , \quad DT_{1t} = \begin{cases} t - TB & if \;\; t \geq TB + 1 \\ 0 & otherwise \end{cases}$$

Here, $T$ is the number of observations while $TB$ is a coefficient which shows the break period between $(0.15T, 0.85T)$ and takes the value of 0 or 1. The GH(1996) test tries to reveal the position of the break and the cointegrated structure between the series using the residual-based test statistics. The residuals are obtained from the different cointegration models. Then GH(1996) test procedures utilize Dickey-Fuller and Phillips test statistics in order to find out the relationship between series.

Many cointegration tests have been developed in the literature in order to avoid inference problems in misleading decisions in cointegration under structural break. Gregory & Hansen (1996) and Campos et al. (1996) propose cointegration tests that allow for a structural break of unknown timing. Harris & Inder (1996), Lee & Strazicich (2001), and Carrion-I-Silvestre & Sanso (2006) developped tests that search for the null of cointegration with one structural break in the level and slope. Recent studies have been conducted considering the finite-sample properties of some cointegration tests that allow for a structural break in the cointegration relationship that is subject to either breaks in level or slope, when applied to independent series. (Hoglund & Ostermark, 2003; Noh and Kim, 2003; Leybourne & Newbold, 2003; Cook, 2004; Westerlund & Edgerton, 2007; Hillebrand & Medeiros, 2008; Tam 2012 etc.)

J.Masuda (2008) extended the GH(1996) test and considered multi-point breaks and obtained asymptotic distribution and critical values through the Monte Carlo simulation method. A. Hatemi-J (2008) deveoloped three residual-based test statistics for cointegration to the cases that take into account two possible regime shifts. These test statistics are modified versions of Gregory Hansen (1996) test. Suggested statistics showed that each test has a small size distortion and very good power properties. However, the GH(1996) method of cointegration analysis with structural breaks is very widespread but suffers from size distortion, when the series have conditional heteroscedasticity (Höglund & Östermark, 2003).

### *2.2.3 Westerlund Edgerton Cointegration Test (2007)*

Westerlund & Edgerton (2007) have offered a cointegration test that is derived from the Lagrange Multiplier (LM) techniques. The test for the null hypothesis of no cointegration considers the serially correlated errors, no break and the deterministic trends as well as the unknown breaks in both the intercept and slope.

General bivariate system of integrated series $(y_t, x_t)$ is:

$$y_t = \alpha + \tau t + \alpha_1 D_{1t} + x_t'\beta + DT_{1t}x_t'\beta_1 + z_t \tag{2.6}$$

$$z_t = \rho z_{t-1} + e_t \quad X_t = X_{t-1} + \omega_t$$

where $t = 1, ..., T$ is the time series index and $y_t$ and $x_t$ are time series of regression equation. $\alpha$ represents the intercept before the break and $\alpha_1$ denotes the change in the intercept at time of the break. $\beta_1$ represents the change in the trend (or slope) at the time of the break. If $TB$ $(1 < TB < T)$ shows the location of break, then dummy variables are defined as before.

The processes $e_t$ and $w_t$ satisfy the following set of conditions:

(a) The process $e_t$ is normal, independent and identically distributed (i.i.d.) with $E(e_t) = 0$ and $E(e_t^2) = \sigma^2 > 0$.

(b) $E(w_t) = 0$ and $E(w_t w_t') = \Omega$ is positive definite.

(c) $E(e_t w_j) = 0$ for all $t$ and $j$.

Condition a enables the application of the standard functional central limit theorem whereas Condition b rules out cointegration among the regressors greater than 1. Finally Condition c requires that the regressors are strictly exogenous.

In this equation $\alpha$ is the intercept, $\alpha_1$, $\beta$, $\beta_1$ and $\rho$ are coefficients, with $z_t$, $e_t$ and $\omega_t$ being the error processes. When $\alpha_1 = \beta_1 = 0$, there is no break in the cointegrating relationship between series. If $\beta_1 = 0$ there is only single intercept break in the cointegrating relationship and if $\alpha_1 = \beta_1 \neq 0$ then both slope and intercept break are allowed for in the relationship.

The Westerlund & Edgerton (2007) test uses a *t-statistics* for testing $\phi = 0$ against the alternative of $\phi < 0$ in following regression model:

$$\Delta \hat{S}_t = constant + \phi \hat{S}_{t-1} + error$$

with $\hat{S}_t$ being the regression error. Thus the regression error can be expressed as

$$\hat{S}_t = y_t - \hat{\alpha} + \hat{\tau} t + \hat{\alpha}_1 D_{1t} + x_t' \hat{\beta} + DT_{1t} x_t' \hat{\beta}_1 \;\; t = 2, 3, ..., T \; with \; \hat{S}_1 = 0.$$

Parameter estimates can be obtained by OLS estimation of the equation above. The asymptotic distribution of the $LM$ test statistics is the same regardless of with or without breaks in the cointegration relationship is the same (Westerlund & Edgerton, 2007). The test utilizes critical values which are computed by Schmidt & Phillips (1992).

The error $z_t$ in equation (2.6) is stationary when $x_t$ and $y_t$ are cointegrated but has a unit root if they are not cointegrated. The cointegration analysis depends on hypothesis testing for a null hypothesis of unit root in the residuals $(H_0 : \rho = 1)$ against the alternative hypothesis of no unit root in the residuals $(H_1 : |\rho| < 1)$. Two test statistics, the *t*-statistics are defined as $\tau = \dfrac{\phi}{se(\phi)}$ and $\phi_N = T\phi$.

These test statistics are the ones proposed in WE-LM(2007).

The model considers only one structural break. They find that, compared to other existing tests, their method performs quite well. They also show that since the asymptotic null distributions of the tests are independent of the nuisance parameters, they are associated with both the trend and the structural breaks. Detailed information about WE-LM(2007) test is presented in chapter 3.

Oh & Lee (2016) proved that the significant losses of power in the WE-LM(2007) cointegration tests when potential multiple breaks are ignored. They extend the WE-LM(2007) cointegration test by allowing for an unknown number of breaks. Then developed modified testing procedures which is restore power loss from the WE-LM(2007) cointegration tests reasonably in the presence of more than one break without affecting the properties of tests under the null.

### 2.3 Structural Breaks

Time series may experience a sudden change in its dynamics. That is, there may be changes in the mean, slope, trend function, variance, dependence structure, or some other characteristic that occur at random times due to policy changes, financial crises, technological improvements and natural disasters.

The existence of the structural breaks may cause various problems such as biases and inconsistent estimation results, biased parameter estimation, poor predictions and modelling of a linear model as a non-linear model. These results decrease the power of the test used.

From the unit root perspective Perron (1989) demonstrates that the ability to reject a unit root decreases when the alternative hypothesis of stationarity is true and an existing structural break is ignored. Perron (1989) used a modified Dickey-Fuller unit root test that inserts dummy variables to model. This method considers exogenous structural breaks. However, Zivot & Andrews (1992) modified the test to allow for one unknown breakpoint that is determined endogenously from the data. For multiple structural breaks, Lee - Strazicich (2003) proposed a two-break minimum LM test, which endogenously determines the location of two breaks in level and trend and tests the null of a unit root. Other unit root tests which take into account breaks are Christiano (1992), Banarjee, Lumsdaine & Stock (1992), Zivot & Andrews (1992), Perron & Vogelsang (1992), Lee and Strazicich (2003), and Perron (1997).

Researchers have paid increasing attention to the effects of structural breaks, not only on unit root testing, but also on cointegration testing. The structural breaks of cointegration relationship mean a significant change of the cointegration parameters or the change of the existence of cointegration relationships. In order to attain correct results in time series analysis, Quintos & Phillips (1993), Inder (1994), Shin (1994), Gregory & Hansen (1996), Hao (1996), , Hansen & Johansen (1999), Busetti (2002), Westerlund & Edgerton(2007) and Hatemi-J( 2008) tests which take structural breaks in time series into consideration, are used (Benerjee & Silvestre, 2006).

Consequentially, if there are structural breaks in the time series used in cointegration, the analysis is conducted without considering the break; the results of the test might be unreliable. These outcomes markedly affected the performance of the tests. Not taking into account the structural breaks from the cointegration analysis may yield incorrect results, since it causes the cointegration parameters to have different values before and after the breaks. The widely used classical residual based cointegration tests such as Engle-Granger may give false results, since they investigate long-term relations without considering structural breaks.

There are many possible forms of structural breaks such as level, slope, mean or variance. In the literature on unit root and cointegration analyses, structural breaks may be summarized as the following figure:

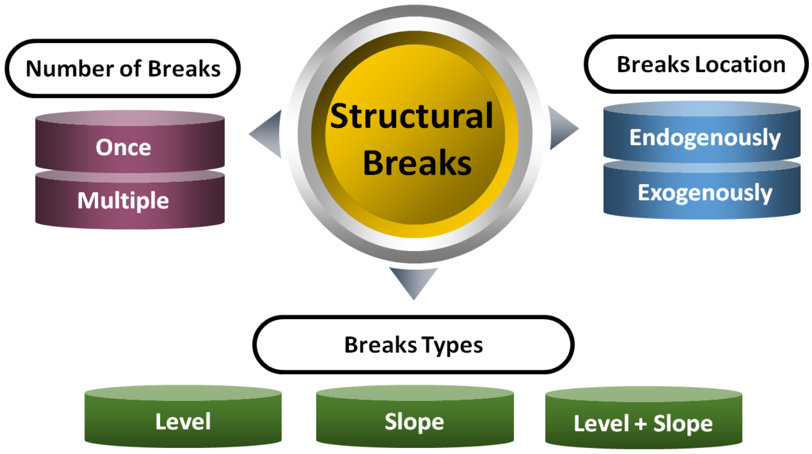


Figure 2.1 Behavior of structural breaks

The nature of a structural break may show different properties such as type of breaks, number of breaks and knowing (exogenously) or unknowing (endogenously) the break locations.

## 2.4 Heteroscedasticity

"Financial economists have long since known that volatility in returns, tends to cluster and that the marginal distributions of many asset returns are leptokurtic. Robert Engle's modelling of time-varying volatility by way of autoregressive conditional heteroskedasticity (ARCH) thus signified a genuine breakthrough. Since their advent, models built around this concept have become an indispensable tool for financial analysts, bankers and fund managers throughout the world" (Nobel Price Committe 2003).

Most econometric and financial time series and their returns usually display non-normality because of that skewness, leptokurtosis and heavy (fat) tail. This time-varying volatility is also known as "conditional heteroscedasticity". The Autoregressive Conditional Heteroskedasticity (ARCH) process is initially recommended by Engle (1982); and in this process, the unconditional variance is constant yet letting the conditional variance change over time "as a function of past errors".

Bollerslev (1986) improve this formulation to allow for a more general formulation that allows lags of the conditional variance to influence its current value, termed the Generalised Autoregressive Conditional Heteroskedasticity (GARCH) model.

The question addressed in this sub-section is whether GARCH error term in cointegration regression can affect the behavior of cointegration tests. Some study remarked that existing OLS-based cointegration tests have low power and that the use of the additional information contained in the heteroskedasticity of the cointegration regression errors should cause to more powerful tests. (Westerlund & Narayan, 2013).

Some different methods are used for specifying model volatility, e.g. the Box-Cox transformations, and ARCH (GARCH) models.

### *2.4.1 Autoregressive Conditional Heteroscedasticity Model (ARCH)*

Autoregressive Conditional Heteroscedasticity (ARCH) described by in 1982.

Using squares of the residuals

$$\hat{e}_t^2 = \theta_0 + \theta_1 \hat{e}_{t-1}^2 + \theta_2 \hat{e}_{t-2}^2 + ... + \theta_q \hat{e}_{t-q}^2 + \eta_t \tag{2.7}$$

The equation (2.7) is called autoregressive conditional heteroscedasticity model (ARCH).

The general ARCH model is;

$$e_t = \eta_t \sqrt{v_t}$$

$$v_t = \theta_0 + \theta_1 e_{t-1}^2$$

$\eta_t \sim WN(0, \sigma^2 = 1)$. $v_t$ and $e_{t-1}$ are independent from each other. In addition $\theta_0$, $\theta_1$ are constant, $\theta_0 > 0$, $0 < \theta_1 < 1$.

Thus, ARCH process have zero mean, with uncorrelated error term and unconditional constant variance (if exists) but conditional variance varies depending on time.

### *2.4.2 GARCH Model*

In 1986, Bollerslev considered the conditional variance model which represented by process moving average rather than autoregressive process. Thus, he added the lagged of conditional variance into the model,. In this case, ARCH model, which is extended by means of the lagged variable, is expressed by the GARCH model.

A generalized autoregressive conditional heteroscedastic (GARCH) model with order $q\ (\geq 1)$ and $p\ (\geq 0)$ is defined as

$$e_t = \eta_t \sqrt{v_t}$$

$$v_t = \theta_0 + \sum_{i=1}^{q} \theta_i e_{t-i}^2 + \sum_{j=1}^{p} \varphi_j v_{t-j}^2$$

where $\theta_0 \geq 0$ and $\varphi_j \geq 0$ are constants. $\{v_t\} \sim IID(0,1)$ and $\{v_t\}I$ are independent of $\{e_{t-j}\ j \geq 1\}$ for all $t$.

A stochastic process $\{e_t\}$ defined by the equations above is called GARCH$(p, q)$ process.

The necessary and sufficient condition for defining unique strictly stationary process

$$\{e_t\ \ t = 0, \pm 1, \pm 2\} \text{ with } E\{e_t\} < \infty \text{ is } \sum_{i=1}^{q} \theta_i e_{t-i}^2 + \sum_{j=1}^{p} \varphi_j v_{t-j}^2 < 1$$

$$E\{e_t\} = 0$$

$$Var\{e_t\} = \frac{\alpha_0}{1 - \sum_{i=1}^{q} \theta_i - \sum_{j=1}^{p} \varphi_j}$$

$$Cov(e_t, e_{t-j}) = 0 \text{ for any } j \neq 0$$

Beside these models, there are many models regarding the of modeling the variance of the error term. Commonly used ones are; ARCH-M (ARCH in mean), IGARCH (Integrated GARCH), EGARCH (Exponential GARCH), TGARCH (Treshold GARCH), FIGARCH (Fractionally integrated GARCH).

### *2.4.3 Test of Heteroscedasticity*

Various tests have been developed in order to determine the heteroscedasticity in time series model. These test are the; Goldfeld-Quandt Test (GQ), Breusch-Pagan-Godfrey Test (BPG) ARCH (LM) and GARCH, Mcleod and Li's Q. The most widely used, among these tests, is Engle's ARCH-LM.

General cointegration tests may lead to spurious cointegration relationships under heteroskedastic error term. Since some econometric and financial time series data are often modeled by heteroskedastic variances, practitioners should be careful when using cointegration tests allowing for ARCH or GARCH process. (D.Maki 2012)

Heterocedasticity can be expressed as volatility of error term. In other words, it is the downward or upward speed changes of the variance. This changeability in error term influences the model parameter estimates. Models should be established to take this effect into account. Otherwise an incorrect modeling would be obtained.

## 2.5 Literature of Cointegration Analysis in Presence of Structural Breaks and Heteroscedaticity

The empirical literature about cointegration tests has been growing over the last three decades. In this sub-section, theoretical literature of cointegration under structural breaks and GARCH effects are is presented shortly.

Leybourne & Newbold (2003) demonstrated that spurious cointegration might appear when breaks in level or slope of integrated series are ignored in the analysis.

The effects of structural breaks are investigated for three cointegration tests. These are the Engle-Granger test, Johansen trace test and Banerjee et al. (1986) test.

Tam (2012) explored the size properties of the Lagrange Multiplier (LM) cointegration tests in the presence of structural breaks. It is found that the misspecification of the types of breaks is liable to spurious rejection. The study showed that WE-LM(2007) test have better performance than alternative existing cointegrating tests.

Silvestre & Sanso (2006) proposed an LM-Type statistic to test the null hypothesis of cointegration allowing for the possibility of a structural break, in both the deterministic and the cointegration vector. They considered the cases of known and unknown break dates.

Maki (2012) investigated the influence of heteroskedastic variances on cointegration tests. Monte Carlo simulation results showed that cointegration tests allowing for structural breaks overreject the null hypothesis of no cointegration in the presence of GARCH errors and variance breaks. The Monte Carlo simulation results in the study showed that all cointegration tests, except for the variance ratio cointegration test, overreject the null hypothesis of no cointegration in the presence of heteroskedastic variances.

Krauss & Herrmann (2017) compaired the power and size properties of ten cointegration tests under the different GARCH process via Monte Carlo simulation. Authors suggested six different processes for cointegration residuals based on the current literature. As a result, they found PGFF and the PP tests exhibit the most favorable power properties.

Kurita (2013) investigated impacts of multivariate (GARCH) errors on hypothesis testing for cointegrating vectors. The experiments demonstrated that the regularity condition plays a crucial role in rendering the hypothesis testing operational. It is also shown that the Bartlett correction and wild bootstrapping are useful in improving the small-sample performance of the test statistic of interest.

Kosapattarapim, Lin & Mccrae (2013) investigated which cointegration test is the most powerful in detecting cointegration relationship when the cointegration errors follow GARCH model with non-normal error distribution. Simulation results indicated that the performance of power of the Johansen's tests in capturing cointegration between financial time series is still higher than alternative tests even when the cointegration errors are not normally distributed.

Westerlund & Narayan (2013) searched efficient market hypothesis for heteroscedastic spot and futures prices. Monte Carlo simulations showed that Weighted LS estimator results more efficient than OLS. They applied Weighted LS approach to test the slope restriction in four commodity spot and futures markets, namely, gold, silver, platinum and oil.

Lee & Strazicich (2003) proposed an endogenous two-break Lagrange multiplier unit root test which endogenously determines the location of two breaks in level and trend and tests the null of a unit root. As a result, rejection of the null unambiguously implies trend stationarity.

Maki (2012) introduced cointegration tests allowing for an unknown number of breaks. Suggested method is based on the test for structural breaks proposed by Bai and Perron (1998). Simulation results provided two main outcomes. First, the proposed tests perform as well as the tests of Gregory & Hansen (1996) and Hatemi-J(2008). Second, the proposed tests perform better than orther tests, when the cointegration relationship has more than three breaks or has persistent Markov switching (MS).

Höglund & Östemark (2003) examined the power and size distortions of a number of cointegration tests when the underlying system is subjected to regime shifts and conditional (ARCH-type) heteroskedasticity. The study demonstrates that the size of the cointegration tests is severely distorted by conditional heteroskedasticity. SW-method of Stock & Watson and Johansen test are severely biased by conditional heteroskedasticity. The author suggested that the Phillips (Za) and ADF-statistics could provide a suitable framework for measuring cointegration in bivariate systems.

Brooks & Rew (2002) considered the effect of GARCH errors on the tests proposed by Perron (1997) for a unit root in the presence of a structural break. They found that the Perron tests are reasonably robust to the presence of GARCH and do not suffer from severe over-or under-rejection of a correct null hypothesis

Silvapulle & Podivinsky (2000) investigated the effect of ARCH/GARCH type processes and non-normal disturbances on cointegration tests in finite samples. Results indicated that researchers should not be concerned by the possibility of small departures from non-normality when using Johansen's suggested techniques. However, ARCH and GARCH effects may be more problematic.

Cavaliere,Rahbek, & Taylor (2010) investigated the properties of the conventional Gaussian-based cointegrating rank tests of Johansen (1996) in the case where the vector of series under test is driven by possibly non-stationary, conditionally heteroskedastic innovations. They proposed wild bootstrap implementations of the cointegrating rank tests. The Monte Carlo evidence presented suggested that the proposed wild bootstrap cointegrating rank tests perform very well with infinite samples.

Lastly and importantly, in unit root framework, Narayan & Liu (2013) suggested a new unit root test which is simultaneously accounts for heteroskedasticity and two endogenous structural breaks. Test has relaxed classical assumption of iid distributed errors and propose a GARCH(1,1) unit root model that accommodates two endogenous breaks in the intercept in the presence of heteroskedastic errors. Parameter estimation of the model are obtained by joint maximum likelihood (ML) estimation. Since break dates are unknown and have to be substituted by their estimates, a sequential procedure is used for deriving the estimates of break dates.

# CHAPTER THREE
# LAGRANGE-BASED TEST FOR COINTEGRATION

## 3.1 Introduction

Cointegration is a statistical property of a long-run relationship between time series variables. If there exists a stationary linear combinations of nonstationary time series variables which move through time and roughly following a random walk (i.e. $I(1)$), the time series variables combined might be cointegrated. In order to find out whether they are cointegrated, firstly the least squares regression model is constructed, then the unit root analysis is performed to residuals of the model in residual-based cointegration test procedure. If the residuals are stationary (or has no unit root), i.e. $I(0)$, it means that the variables are cointegrated and have a long-run or equilibrium relationship.

Cointegration could be tested only on nonstationary time series variables. Applied cointegration tests to stationary data may lead to spurious inference. In this sense, the data must be tested for integration (nonstationary), because the cointegration assumes that the underlying time series variables are integrated. Augmented Dickey-Fuller-ADF (1979) test was common unit root test in literature. Perron (1989) indicated that ignoring an existing break leads to a bias and decreases to reject a false unit root null hypothesis. Zivot and Andrews (1992), Gregory-Hansen (1995) and Perron (1997) utilized ADF test which is determined the break point endogenously from the data.

Another common unit root test is Lagrange Multiplier Test (LM) based on the Lagrange Multiplier principle which was initially proposed by Schmidt and Phillips (1992). LM unit root test has more general and flexible assumptions about error term because of its asymptotic properties. From a structural break perspective, several studies have shown that the LM tests has great advantages with compared to DF-type tests (P.Schmidt and P.Phillips (1992), Phillips and Ouliaris (1990), J. Westerlund (2013). A valuable property of LM based tests is that their distributions under both the null and alternative hypothesis are independent of the intercept, trend coefficient and variance of error (Phillips & Schmitz, 1989). Amsler & Lee (1995) showed that the asymptotic distribution of the LM unit root test is invariant to a structural change in the intercept assuming no break(s) under the null hypothesis. These authors also proposed a modification to the LM test that allows for a break in the intercept at some known structural break. They showed that the limiting distribution of the test statistic under the null hypothesis of a unit root is the same as for the Schmidt and Phillips (1992) LM test where no break is considered.

### 3.2 The LM-Based Cointegration Test Statistics

This section covers the connection cointegration and unit root procedure under the structural break. The motivation comes from the article of Westerlund & Edgerton (2007). Authors suggested the LM-based tests statistics of existing one structural break. They make the assumptions that the data generating process (DGP) can be described by the following unobserved components representation

$$y_t = \alpha + \tau t + x_t' \beta + z_t \tag{3.1}$$

$$z_t = \rho z_{t-1} + e_t \tag{3.2}$$

$$x_t = x_{t-1} + w_t \tag{3.3}$$

where $t = 1, ..., T$ is the time series index and $x_t$ is a $K$-dimensional vector of regressors, $z_t$ is error term of cointegration relationship. $\alpha$ and $\tau t$ are deterministic components which represent the intercept and linear time trend in (3.1), respectively. For convenience in deriving the residual based tests and their asymptotic distributions, and the initial conditions $z_0$ and $x_0$ are assumed as fixed. The processes $e_t$ and $w_t$ satisfy the following set of conditions:

**Assumption 1:**

a. $E(e_t w_j) = 0$ for all $t$ and $j$;

b. The process $e_t$ is normal, independent and identically distributed (*i.i.d.*) with $E(e_t) = 0$ and $E(e_t^2) = \sigma^2 > 0$

c. $E(w_t) = 0$ and $E(w_t w_t') = \Omega$ is positive definite.

These conditions needed for improving the residual based cointegration tests.

This residual based test relies on the error terms ($z_t$) constructed from equation (3.1). The cointegration analysis depends on hypothesis testing for a null hypothesis of unit root in the residuals ($H_0 : \rho = 1$) against the alternative hypothesis of no unit root in the residuals ($H_1 : |\rho| < 1$).

$H_0$ : No cointegration; ($\rho = 1$ means that $z_t$ has unit root (integrated))
$H_1$ : Cointegration $x$ and $y$; ($|\rho| < 1$ means that $z_t$ has stationary process)

Rejecting null hypothesis means that the residuals have unit root, thereby the null of no cointegration is also rejected. In other words, error $z_t$ in equation (3.1) has to become stationary while $x_t$ and $y_t$ are cointegrated, otherwise but has a unit root if they are not cointegrated.

The simple form of LM statistics is written in equation (3.4).

$$LM = \left(\frac{\partial lnL}{\partial \rho}\right)^2 \left(\frac{\partial^2 lnL}{\partial \rho^2}\right)^{-1} \tag{3.4}$$

where $lnL = constant - \frac{T}{2} ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{t=1}^{T} e_t^2$ under the assumption 1.

LM score or principle can applied for testing the $\rho = 1$ restriction for $\rho$ in the equation (3.2). The score has zero mean when evaluated at the true parameter vector under the null hypothesis. The efficient LM score must be evaluated at $\hat{\rho} = 1$ under the null hypothesis. Direct calculation of the score vector of likelihood function is obtained by minimizing the sum of squared errors which then reduces to

$$SSE = e_1^2 + \sum_{t=2}^{T} e_t^2 \ \text{ or } \ SSE = S_1^2 + \sum_{t=2}^{T}(\Delta S_t)^2 \tag{3.5}$$

First of all, reduced SSE must be obtained into two parts by using equation (3.1) and (3.2) under $\rho = 1$.

$$e_1 = y_1 - \hat{\alpha} - \hat{\tau} - x_1'\hat{\beta} - z_0 \tag{3.6}$$

$e_t$ can be obtained as in equation (3.7)

$$e_t = y_t - \hat{\alpha} - \hat{\tau}t - x_t'\hat{\beta} - \hat{\rho}z_{t-1} \tag{3.7}$$

Score of the SSE with respect to $\rho$ is given by

$$\frac{\partial SSE}{\partial \rho} = -\frac{1}{2\sigma^2}\Big[\frac{\partial}{\partial \rho}(y_1 - \hat{\alpha} - \hat{\tau} - x_1'\hat{\beta} - \hat{\rho}z_0)^2 + \frac{\partial}{\partial \rho}\sum_{t=2}^{T}(\Delta y_t - \hat{\tau} - \Delta x_t'\hat{\beta} - \hat{\rho}\Delta z_{t-1})^2\Big]$$

$$\frac{\partial SSE}{\partial \rho} = -\frac{1}{\sigma^2}\Big[-z_0(y_1 - \hat{\alpha} - \hat{\tau} - x_1'\hat{\beta} - \hat{\rho}z_0) - \sum_{t=2}^{T}(-\Delta z_{t-1})(\Delta y_t - \hat{\tau} - \Delta x_t'\hat{\beta} - \hat{\rho}\Delta z_{t-1})\Big]$$

$$\frac{\partial SSE}{\partial \rho} = e_1 z_0 + \underbrace{\sum_{t=2}^{T} \Delta z_{t-1}\Delta \hat{S}_t}_{z_1e_2 + z_2e_3 + \ldots + z_{T-1}e_T = \sum_{t=2}^{T} e_t(z_{t-1})}$$

For detailed information see Appendix 1 Result 4.

$$\frac{\partial SSE}{\partial \rho} = e_1 z_0 + \sum_{t=2}^{T} e_t(\hat{S}_{t-1} + z_0) = e_1 z_0 + \sum_{t=2}^{T} e_t\hat{S}_{t-1} + \sum_{t=2}^{T} e_t z_0$$

$$\begin{aligned}
\frac{\partial SSE}{\partial \rho} &= \hat{S}_1 z_0 + \sum_{t=2}^{T}\Delta\hat{S}_t\hat{S}_{t-1} + z_0\sum_{t=2}^{T}\Delta\hat{S}_t \\
&= \hat{S}_1 z_0 + \sum_{t=2}^{T}\Delta\hat{S}_t\hat{S}_{t-1} + z_0\sum_{t=2}^{T}(\hat{S}_T - \hat{S}_1) \\
&= \hat{S}_1 z_0 + z_0\hat{S}_T - z_0\hat{S}_1 + \sum_{t=2}^{T}\Delta\hat{S}_t\hat{S}_{t-1}
\end{aligned}$$

$\hat{S}_1 z_0$ terms cancel out each other in the equation and $\hat{S}_T z_0$ equals to zero since $\hat{S}_T = 0$ (For detail see Appendix 1 Results 1-5). Hence,

$$\frac{\partial SSE}{\partial \rho} = \sum_{t=2}^{T} \Delta \hat{S}_t \hat{S}_{t-1}$$

By replacing this result into $lnL$ function,

$$\frac{\partial lnL}{\partial \rho} = \hat{\sigma}^{-2} \sum_{t=2}^{T} \Delta \hat{S}_t \hat{S}_{t-1}$$

Westerlund & Edgerton (2007) preferred to use $(.)_p$ notation for the errors from projecting observations onto the generic vector random variable of the projection variables. For example $(S_{t-1})_p = (S_{t-1})_{obs} - S'_{t-1}a$ where $a$ is a vector of projection parameters (i.e. mean value).

As a result,

$$\frac{\partial lnL}{\partial \rho} = \hat{\sigma}^{-2} \sum_{t=2}^{T} (\Delta \hat{S}_t)_p (\hat{S}_{t-1})_p \tag{3.8}$$

indicates that errors around projection parameter values (mean values).

The estimated $S_t$ can be rewritten by means of equation (3.7)

$$\hat{S}_t = y_t - \hat{\alpha} - \hat{\tau} t - x'_t \hat{\beta}$$

where the restricted maximum likelihood estimate (MLE) of $\hat{\alpha}$ equals to $y_1 - \hat{\alpha} - \hat{\tau} t - x'_t \hat{\beta}$ (for the derivation procedure see Appendix 1, result 1).

Ordinary Least Squares (OLS) regression in equation (3.9) which is used for obtaining restricted MLE of parameters $\tau$ and $\beta$.

$$\Delta y_t = \hat{\tau} + \Delta x'_t \hat{\beta} + error \tag{3.9}$$

The equation (3.10) can be constructed by the equation (3.7) which indicates that the likelihood score is proportional to the OLS estimate of $\phi$ parameter which is autoregressive parameter in the equation.

$$\Delta\hat{S}_t = intercept + \phi\hat{S}_{t-1} + error \tag{3.10}$$

where $\hat{S}_{t-1}$ is the residual from an Ordinary Least Squares regression of $y_{t-1}$ on an intercept and time trend. It obvious that a test of the hypothesis of $H_0$ $(\rho = 1)$ versus $H_1$ $(|\rho| < 1)$ is equivalent to a test of the hypothesis of $\phi = 0$ versus $\phi < 0$, respectively. The OLS estimate of can be tested by t-*test*.

Therefore, there are two statistics which might be called non-break LM-based cointegration tests statistics

$$t_N = \frac{\hat{\phi}}{\hat{\sigma}/\sqrt{\sum_{t=2}^{T}\hat{S}_{t-1}^2}} \text{ and } \phi_N = T\hat{\phi} \tag{3.11}$$

where $\hat{\phi}$ is the OLS estimate of $\phi$ in (3.10) and $\hat{\sigma}$ is the estimated standard error from the same regression.

### *3.2.1 Asymptotic Distribution LM-based test without Structural Break*

Test asymptotic distribution of Schmidt & Phillips (1992) LM-based test depends on the theory of functional weak convergence which is given in detail in Appendix 2. The limit distributions of $\phi_N$ and $t_N$ for without break can be obtained by standard Brownian motion. For this purpose, it is convenient to let $W(r)$ denote a standard Brownian motion on the unit interval $r \in [0, 1]$, to let $V(r) = W(r) - rW(1)$ denote a standard Brownian bridge and to let $U(r) = V(r) - \int_0^1 V(s)ds$ indicate a demeaned standard Brownian bridge.

**Theorem 1:** The limiting distribution under $H_0$ and Assumption 1, as $T \to \infty$ then

$$t_N \Rightarrow -\left(4\int_0^1 U(r)^2 dr\right)^{-1/2} \quad and \;\; \phi_N \Rightarrow -\left(2\int_0^1 U(r)^2 dr\right)^{-1} \tag{3.12}$$

The asymptotic distribution $\phi_N$ and $t_N$ are identical to those obtained by Schmidt & Phillips (1992) for their $\hat{\phi}_N$ and $\hat{t}_N$ unit root statistics, the critical values for the Westerlund-Edgerton LM (WE-LM) test are available in Table 3.1 in the paper.

The limiting distribution of $\phi_N$ and $t_N$ test statistics does not affected by the regressors. Hence, the same set of critical values are available regardless of the regressors, so it makes great advantages in computational step.

### *3.2.2 Obtaining LM-Based Test Statistics without Structural Break*

The limiting distribution of LM-based test statistics is derived under $H_0$.

Considering equation (3.1) and (3.2), $y_1$ and $y_t$ can be written easily as equation (3.13).

$$\begin{aligned}
y_1 &= \alpha + \tau + x_1'\beta + \rho z_0 + e_1 \\
y_{t-1} &= \alpha + \tau(t-1) + x_{t-1}'\beta + \rho z_{t-2} + e_{t-1} \\
y_t - \rho y_{t-1} &= \alpha + \tau t + x_t'\beta + z_t - \rho(\alpha + \tau(t-1) + x_{t-1}'\beta + \rho z_{t-1} + e_{t-1}) \\
y_t - \rho y_{t-1} &= \alpha(1-\rho) + \tau(t - \rho(t-1)) + \beta(x_t - \rho x_{t-1})' + (z_t - \rho z_{t-1})
\end{aligned}$$

$$y_t = \rho y_{t-1} + \alpha(1-\rho) + \tau(t-\rho(t-1)) + \beta(x_t - \rho x_{t-1})' + e_t \tag{3.13}$$

Pay attention please, there is no any $z_t$ term in equation (3.13). In this sense, log-likelihood function can be expressed as general form regardless of $z_t$ term, because this term is eliminated by differencing.

Therefore the score vector of the concentrated log-likelihood function is given by equation (3.14). Let define the true parameter vector $\gamma = (\tilde{\alpha}, \tau, \beta')$

$$\frac{\partial lnL}{\partial \gamma} = -\frac{1}{2\sigma^2}\sum_{t=1}^{T}\frac{\partial e_t^2}{\partial \gamma} \tag{3.14}$$

Parameters obtained by minimizing the sum of squared errors, which has the reduced form under the null in equation (3.5).

As shown in Appendix 1, Result 1, 2, and 3, the restricted MLE of $\gamma$ vector parameter is presented in equation (3.15).

$$-\frac{1}{2\sigma^2}\sum_{t=1}^{T}\frac{\partial e_t^2}{\partial \gamma} = \left[S_1, S_T, S_1 x_1' + \sum_{t=2}^{T}\Delta S_t \Delta x_t'\right] \tag{3.15}$$

To derive the LM-based test statistics, the second order derivation of SSE with respect to $\rho$ must be obtained

$$-\frac{1}{2\sigma^2}\sum_{t=1}^{T}\frac{\partial e_t^2}{\partial \rho^2} = -\frac{1}{2\sigma^2}\left[\frac{\partial}{\partial \rho}\left[z_0(y_1 - \alpha - \tau - x_1'\beta - \rho z_0)\right]\right.$$
$$\left. + \sum_{t=2}^{T}\frac{\partial}{\partial \rho}\Big(S_{t-1} \times [(y_t - \rho y_{t-1}) - \alpha(1-\rho) - \tau(t - \rho(t-1)) - \beta(x_t' - \rho x_{t-1}')]\Big)\right]$$

Since $e_t = \hat{S}_T = 0$ (Appendix 1, result 5), the appropriate normalizing order of these derivatives is determined by order of the second derivative with respect to $\rho$, which is given by

$$\frac{\partial SSE}{\partial \rho^2} = z_0^2 + \sum_{t=2}^{T} S_{t-1}^2 \tag{3.16}$$

By using equations (3.8) and (3.16), the simple form of LM-based test has the asymptotically representation of LM-based test as in equation (3.17).

$$LM = \left(\sum_{t=2}^{T}\Delta \hat{S}_t \hat{S}_{t-1}\right)^2 \left(\hat{\sigma}^2 \sum_{t=2}^{T}\hat{S}_{t-1}^2\right)^{-1} \tag{3.17}$$

For completing the derivation of the LM-based test statistics, it has to be shown that the Hessian matrix of the restricted log-likelihood function is indeed asymptotically block diagonal for all parameters $(\rho, \alpha, \beta$ and $\tau)$.

The Hessian Matrix is (Appendix 1, result 6)

$$H = \frac{1}{\sigma^2}\begin{pmatrix} \frac{\partial^2 e_t^2}{\partial \rho^2} & \frac{\partial^2 e_t^2}{\partial \alpha \partial \rho} & \frac{\partial^2 e_t^2}{\partial \tau \partial \rho} & \frac{\partial^2 e_t^2}{\partial \beta \partial \rho} \\ & \frac{\partial^2 e_t^2}{\partial \alpha^2} & & \vdots \\ \vdots & & \frac{\partial^2 e_t^2}{\partial \tau^2} & \\ & & & \frac{\partial^2 e_t^2}{\partial \beta^2} \end{pmatrix}$$

$$= \frac{1}{\sigma^2}\begin{pmatrix} z_0^2 + \sum_{t=2}^{T} \hat{S}_{t-1}^2 & z_0 & z_0 + \sum_{t=2}^{T} \hat{S}_{t-1} & x_T z_0 + \sum_{t=2}^{T} x'_{t-1}\Delta S_t + \sum_{t=2}^{T} \Delta x'_t S_{t-1} \\ & \frac{\partial^2 e_t^2}{\partial \alpha^2} & & \vdots \\ \vdots & & \frac{\partial^2 e_t^2}{\partial \tau^2} & \\ & & & \frac{\partial^2 e_t^2}{\partial \beta^2} \end{pmatrix}$$

To show that Hessian matrix is asymptotically block diagonal, it must be normalized by $T^{-2}$ (appropriate normalizing order). If the first term in the first row multiply by $T^{-2}$, $z_0^2 + \sum_{t=2}^{T} \hat{S}_{t-1}^2$ approaches a limiting distribution in light of convergence where $W(r)$ represents the standard Brownian motion on $[0, 1]$. Hence, this terms $\sigma^2 \int_0^1 W(r)^2 dr \rightarrow O_p(1)$ (approaches to a constant). And also $-\frac{z_0^2}{\sigma^2} \rightarrow o(1)$. All convergence properties are summarized in Table 3.2 (Convergence information which are around $S_t$). In the overall perspective, since both $T^{-1/2}\hat{S}_t$ and $T^{-1/2}x_t$ approach to $O_p(1)$, $T^{-3/2}\sum_{t=2}^{T} \hat{S}_{t-1}$, $T^{-1}\sum_{t=2}^{T} \Delta x_t \hat{S}_{t-1}$ and $T^{-1}\sum_{t=2}^{T} x_{t-1}\Delta \hat{S}_t$ also approach to $O_p(1)$. This shows that Hence, the Hessian matrix is an asymptotically block-diagonal matrix. According to this result, The LM-based test which is presented in equation (3.17) is the asymptotically representation of the LM-based test.

Having the asymptotically representation of the LM-based test, asymptotic distributions of $t_N$ and $\phi_N$ under the null and assumption 1 must be proved (equation (3.12)).

According to Schmidt (1992), $S_t$ could be written as

$$S_t = \sum_{j=2}^{t} \left( \rho^{j-1}(\rho - 1)z_0 + e_j + (\rho - 1) \sum_{i=2}^{j-1} \rho^i e_{j-i} - \bar{\omega} \right)$$

Under $H_0$ and $\rho = 1$ circumstances, $S_t$ could be written as $S_t = \sum_{j=2}^{t}(e_j - \bar{e})$. The advantage of this expression is that it is independent from $\alpha$ and $\tau$ parameters. And also Westerlund (2007) is used this information for proof of Theorem 1.

$$T^{-1/2}\hat{S}_t = T^{-1/2} \sum_{j=2}^{t}(e_j - \bar{e}) + o_p(1) \tag{3.18}$$

Let define the limiting distribution of $\hat{S}_t$ as follows (detail information see Appendix 1, Result 5)

$$\hat{S}_t = y_t - y_1 - \hat{\tau}(t-1) - (x_t - x_1)'\hat{\beta}$$

$$\hat{S}_t = \sum_{j=2}^{t} \Delta y_j - \hat{\tau}(t-1) - \sum_{j=2}^{t} \omega_j'\hat{\beta} \quad \text{where } \omega_j' = (x_t - x_1)' = \sum \Delta x_t'$$

$$S_t = e_t = \Delta y_t - \tau - \Delta x_t'\beta \quad \text{(3b in Table 3.1)}$$

$$\hat{S}_t - S_t = \sum_{j=2}^{t} \Delta y_j - \Delta y_j - (\hat{\tau} - \tau)(t-1) - \sum_{j=2}^{t} \omega_j'(\hat{\beta} - \beta)$$

$$\hat{S}_t - S_t = \sum_{j=2}^{t} e_j - (\hat{\tau} - \tau)(t-1) - \sum_{j=2}^{t} \omega_j'(\hat{\beta} - \beta) \text{ where } \sum_{j=2}^{t} e_j = \sum_{j=2}^{t} \Delta y_j - \bar{\Delta} y_j \tag{3.19}$$

Using almost surely (a.s) convergence theorem, $P\left(|\hat{S}_t - S_t| > \varepsilon\right) \xrightarrow{a.s} 0$,

$$(\hat{\tau} - \tau)(t-1) = \sum_{j=2}^{t} e_j - \sum_{j=2}^{t} \omega_j'(\hat{\beta} - \beta)$$

$$(\hat{\tau} - \tau) = \frac{\sum_{j=2}^{t} e_j}{t-1} - \frac{\sum_{j=2}^{t} \omega_j'(\hat{\beta} - \beta)}{t-1}$$

$$(\hat{\tau} - \tau) = \bar{e} - \bar{\omega}'(\hat{\beta} - \beta) \tag{3.20}$$

If equation (3.20) is replaced into equation (3.19),

$$\hat{S}_t = \sum_{j=2}^{t} e_j - \Big(\bar{e} - \omega_j'(\hat{\beta} - \beta)(t-1)\Big) - \sum_{j=2}^{t} \omega_j'(\hat{\beta} - \beta)$$

$$\hat{S}_t = \sum_{j=2}^{t} e_j - \bar{e}(t-1) + \omega_j'(\hat{\beta} - \beta)(t-1) - \sum_{j=2}^{t} \omega_j'(\hat{\beta} - \beta) \tag{3.21}$$

If the both side of equation multiply by $\sqrt{T}$,

$$\begin{aligned} T^{-1/2}\hat{S}_t &= T^{-1/2}\sum_{j=2}^{t} e_j - T^{-1/2}\bar{e}(t-1) - T^{-1/2}\sum_{j=2}^{t} \omega_j'(\hat{\beta} - \beta) + T^{-1/2}\omega'(\hat{\beta} - \beta)(t-1) \\ &= T^{-1/2}\sum_{j=2}^{t}(e_j - \bar{e}) - T^{-1/2}\sum_{j=2}^{t}(\omega_j - \bar{\omega})'(\hat{\beta} - \beta) \\ &= part\ I - part\ II \end{aligned} \tag{3.22}$$

For $part\ I$, we have $T^{-1/2}\sum_{j=2}^{t}(e_j - \bar{e}) \to \sigma V(r)$ as $T \to \infty$, which implies that $O_p(1)$. In order to explain this approximation, brief information about Brownian motion, continuous mapping theorem and functional central limit theorem is given in Appendix 2.

For $part\ II$, $T^{-1}\sum_{j=2}^{t}(\omega_j - \bar{\omega}) \to \Omega^{1/2}\overrightarrow{V}(r)$ as $T \to \infty$ where $\overrightarrow{V}(r)$ is a standard vector Brownian bridge process. $\Omega$ is the variance-covariance matrix, which is defined in Assumption 1.

If $T^{-1}\sum_{j=2}^{t}(\omega_j - \bar{\omega})'(\hat{\beta} - \beta)$ term is multiplied by $\sqrt{T}$, the convergence would be discussed for $T^{-1/2}$. For examining the convergence procedure for $\beta$, equation (3.22) could be thought as a linear regression model. The OLS estimate of $\beta$ is

$$\sum_{j=2}^{t} e_j = \sum_{j=2}^{t} \omega_j \beta \ \text{ or } \ \sum_{j=2}^{t} e_j = \sum_{j=2}^{t} (\Delta x_t)\beta \ \text{ where } \ \omega_j' = (x_t - x_1)' = \sum \Delta x_t'$$

$$\sqrt{T}(\hat{\beta}-\beta)=\sqrt{T}\left(\frac{\sum_{t=2}^{T}(\Delta x_t)_p(e_t)_p}{\sum_{t=2}^{T}(\Delta x_t)_p(\Delta x_t)'_p}\right) \text{ where } (.)_p \text{ is the projection.}$$

$\sqrt{T}\sum_{t=2}^{T}(\Delta x_t)_p(e_t)_p \rightarrow O_p(1)$ (Due to $e_t$ and $\omega_t$ are perpendicular each other according to Assumption 1a.)

$T^{-1}\sum_{t=2}^{T}(\Delta x_t)_p(\Delta x_t)'_p \xrightarrow{p} \Omega$ (Converges to variance-covariance in probability, due to Assumption 1c.)

Hence, $\sqrt{T}(\hat{\beta}-\beta) = \left(T^{-1}\sum_{t=2}^{T}(\Delta x_t)_p(\Delta x_t)'_p\right)^{-1}T^{-1/2}\sum_{t=2}^{T}(\Delta x_t)_p(e_t)_p$. This statement explained by asymptotic normality of OLS estimator Lindberg-Levy CLT.

$$\sqrt{T}(\hat{\beta}-\beta)\rightarrow O_p(1)$$

Therefore, $part\ II\ T^{-1/2}\sum_{j=2}^{t}(\omega_j-\omega)'\sqrt{T}(\hat{\beta}-\beta) = O_p(T^{-1/2})$. $O_p(1)\rightarrow o_p(1)$ converges to zero.

As a result, equation (3.22) turns into

$$T^{-1/2}\hat{S}_t = T^{-1/2}\sum_{j=2}^{t}(e_j-\bar{e}) - o_p(1)\rightarrow \sigma V(r)$$

To show the limiting distribution of $\phi_N$ and $t_N$ as $T\rightarrow\infty$, $\hat{S}_t$ can be rewritten as $\hat{S}_t = y_t - y_1 - \hat{\tau}(t-1) - (x_t-x_1)'\hat{\beta}$.

The result which is shown in Appendix, equation A6, $\Delta y - \Delta x\hat{\beta} = \hat{\tau}$ could be replaced in $\hat{S}_t$.

$$\hat{S}_t = y_t - y_1 - (\Delta y - \Delta x'\hat{\beta})(t-1) - (x_t-x_1)'\hat{\beta}$$
$$\hat{S}_t = \sum_{j=2}^{t}\Delta y_j - t\frac{\sum_{j=2}^{t}\Delta y_j}{t} + \frac{\sum_{j=2}^{t}\Delta y_j}{t} - t\frac{\sum_{j=2}^{t}\Delta x_j}{t}\hat{\beta} + \frac{\sum_{j=2}^{t}\Delta x_j}{t}\hat{\beta}$$
$$\hat{S}_t = \sum_{j=2}^{t}\Delta y_j - \sum_{j=2}^{t}\Delta y_j + \frac{\sum_{j=2}^{t}\Delta y_j}{t} - \sum_{j=2}^{t}\Delta x_j\hat{\beta} + \frac{\sum_{j=2}^{t}\Delta x_j}{t}\hat{\beta}$$

$$\hat{S}_t = \sum_{j=2}^{t} \Delta y_j - \sum_{j=2}^{t} \Delta y_j + \Delta y - \sum_{j=2}^{t} \Delta x_j \hat{\beta} + \Delta x \hat{\beta}$$
$$\hat{S}_t = \sum_{j=2}^{t} \Delta y_j - \left( \sum_{j=2}^{t} \Delta y_j - \Delta y \right) - \left( \sum_{j=2}^{t} \Delta x_j + \Delta x \right) \hat{\beta}$$

It is easy to see that $\left( \sum_{j=2}^{t} \Delta y_j - \Delta y \right) = 0$ and $\left( \sum_{j=2}^{t} \Delta x_j - \Delta x \right) = 0$ since $\Delta y$ and $\Delta x$ are the sample averages.

Finally, $\hat{S}_t$ has become with the projection form,

$$\hat{S}_t = \sum_{j=2}^{t} (\Delta y_j)_p - \sum_{j=2}^{t} (\Delta x_j)'_p \hat{\beta} \tag{3.23}$$

It is known that $\hat{S}_t = e_t$, this equality comes from 1a in Table 3.1.

$$e_t = \sum_{j=2}^{t} (\Delta y_j)_p - \sum_{j=2}^{t} (\Delta x_j)'_p \hat{\beta}$$

In terms of mean deviations, $\sum_{j=2}^{t} (\Delta y_j)_p - \sum_{j=2}^{t} (\Delta x_j)'_p \hat{\beta} + (e_t)_p$ and it is known that $\Delta \hat{S}_t = e_t = \Delta y_t - \hat{\tau} - \Delta x'_t \hat{\beta}$ (from 3a in Table 3.1) , equation can deduce to $\Delta \hat{S}_t = (e_t)_p$.

Now, consider $\phi_N$ from the following OLS regression (mentioned before in equation (3.10))

$$\Delta \hat{S}_t = intercept + \phi \hat{S}_{t-1} + error \tag{3.24}$$

The OLS estimate of $\phi_N$ with the projection form as follows

$$\phi_N = \left( \sum_{t=2}^{T} (\hat{S}_{t-1})_p^2 \right)^{-1} \sum_{t=2}^{T} (\hat{S}_{t-1})_p (\Delta \hat{S}_t)_p \tag{3.25}$$

If the numerator multiplies by $T^{-1}$ and the denominator multiplies by $T^{-2}$ in equation

(3.25), by this way, convergence properties could be available.

$$\phi_N = \frac{T^{-1}\sum_{t=2}^{T}(\hat{S}_{t-1})_p(\Delta\hat{S}_t)_p}{T^{-2}\left(\sum_{t=2}^{T}(\hat{S}_{t-1})_p^2\right)} \tag{3.26}$$

Let examine the convergence properties of $\phi_N$ in the equation (3.26) piece by piece

$$T^{-1}\sum_{t=2}^{T}(\hat{S}_{t-1})_p(\Delta\hat{S}_t)_p = -\frac{1}{2}\left(T^{-1}\sum_{t=2}^{T}(\Delta\hat{S}_t)_p^2\right)$$

The proof of the $\sum_{t=2}^{T}\hat{S}_{t-1}\Delta\hat{S}_t = -\frac{1}{2}\left(\sum_{t=2}^{T}\Delta\hat{S}_t^2\right)$ is shown in Appendix 1, result 7.

$$-\frac{1}{2}\left(T^{-1}\sum_{t=2}^{T}(\Delta\hat{S}_t)_p^2\right) = -\frac{1}{2}\left(T^{-1}\sum_{t=2}^{T}(e_t)_p^2\right) \xrightarrow{p} -\frac{1}{2}\sigma^2 \tag{3.27}$$

$$\left(T^{-2}\sum_{t=2}^{T}(\hat{S}_{t-1})_p^2\right) \to \sigma^2\int_0^1 U(r)^2dr \tag{3.28}$$

Equations (3.27) and (3.28) imply that the limit of $\phi_N$ as $T \to \infty$ is equal to

$$\phi_N = -\frac{1}{2}\frac{\sigma^2}{\sigma^2\int_0^1 U(r)^2dr} = -\frac{1}{2}\left(\int_0^1 U(r)^2dr\right)^{-1} \to O_p(T^{-1}) \tag{3.29}$$

For obtaining the limiting distribution of $t_N = \dfrac{\phi_N}{S_{\phi_N}}$, the convergence of $\hat{\sigma}^2$ must be examined as follows:

Like in the OLS procedure $\hat{\sigma}^2 = \dfrac{(Y-\beta X)^2}{n}$, the projection form of $\hat{\sigma}^2$ can be written as

$$\hat{\sigma}^2 = T^{-1}\sum_{t=2}^{T}\Big((\Delta\hat{S}_t)_p - \hat{\phi}(S_{t-1})_p\Big)^2$$

$$\hat{\sigma}^2 = T^{-1}\left[\sum_{t=2}^{T}(\Delta\hat{S}_t)_p^2 - 2\hat{\phi}\sum_{t=2}^{T}(\Delta\hat{S}_t)_p(\hat{S}_{t-1})_p + \hat{\phi}^2\sum_{t=2}^{T}(\hat{S}_{t-1})_p^2\right]$$

$$\hat{\sigma}^2 = T^{-1}\left[\sum_{t=2}^{T}(e_t)_p^2 - 2\hat{\phi}\frac{1}{2}\sum_{t=2}^{T}(\Delta\hat{S}_t)_p^2 + \hat{\phi}^2\sum_{t=2}^{T}(\hat{S}_{t-1})_p^2\right] \text{ by result 7 in Appendix 1}$$

$$\hat{\sigma}^2 = T^{-1}\sum_{t=2}^{T}(e_t)_p^2 + \hat{\phi}T^{-1}\sum_{t=2}^{T}(e_t)_p^2 + \hat{\phi}^2 T^{-1}\sum_{t=2}^{T}(\hat{S}_{t-1})_p^2 \tag{3.30}$$

There are three term in equation (3.30). The second term $T^{-1}\sum_{t=2}^{T}(e_t)_p^2 \xrightarrow{p} \sigma^2$ and the third term $T^{-1}\sum_{t=2}^{T}(\hat{S}_{t-1})_p^2 \rightarrow O_p(T)$ and both second and third term convergence to zero $(o_p(1))$. The first term also convergence to $\sigma^2$. By means of all results $\hat{\sigma}^2 \xrightarrow{p} \sigma^2$.

By using this result, limiting distribution of $t_N$ as $T \rightarrow \infty$ like $\phi_N$ can be obtained as following:

$$t_N = \frac{\sum_{t=2}^{T}(\hat{S}_{t-1})_p(\Delta\hat{S}_t)_p}{\sqrt{\sum_{t=2}^{T}(\hat{S}_{t-1})_p^2/\hat{\sigma}^2}}$$

As mentioned before, if the numerator multiplies by $T^{-1}$ and the denominator multiplies by $T^{-1}$ this way convergence properties could be available.

$$t_N = \left(\hat{\sigma}^{-2}T^{-2}\sum_{t=2}^{T}(\hat{S}_{t-1})_p^2\right)^{-1/2} T^{-1}\sum_{t=2}^{T}(\hat{S}_{t-1})_p(\Delta\hat{S}_t)_p \tag{3.31}$$

Let examine the convergence properties of $t_N$ in the equation (3.31) piece by piece

$$T^{-1}\sum_{t=2}^{T}(\hat{S}_{t-1})_p(\Delta\hat{S}_t)_p = -\frac{1}{2}T^{-1}\sum_{t=2}^{T}(\Delta\hat{S}_t)_p^2 = -\frac{1}{2}T^{-1}\sum_{t=2}^{T}(e_t)_p^2 \rightarrow -\frac{1}{2}\sigma^2$$

$$T^{-2}\sum_{t=2}^{T}(\hat{S}_{t-1})_p^2 \rightarrow \sigma^2\int_0^1 U(r)^2 dr$$

$$t_N = \hat{\sigma}^{-1}\left(\sigma^2\int_0^1 U(r)^2 dr\right)^{-1/2}\left(-\frac{1}{2}\sigma^2\right) = \hat{\sigma}^{-2}\left(\int_0^1 U(r)^2 dr\right)^{-1/2}\left(-\frac{1}{2}\sigma^2\right)$$

$$t_N \rightarrow -\frac{1}{2}\left(\int_0^1 U(r)^2 dr\right)^{-1/2} \tag{3.32}$$

Eventually, Theorem 1 has established step by step from the equation (3.18) to the equation (3.32).

## 3.3 LM-Based Cointegration Test in the presence of One and More Structural Breaks

The standard process to test the cointegration (the null of no cointegration) is estimated by OLS regression model, which is determined by adding dummy variables for each break. Type of structural break characterizes the dummy variables, i) level shift ii) level shift with trend iii) regime shift (both level shift and slope break). As mentioned in the previous section, LM based tests' distributions under both the null and alternative hypothesis are independent of the intercept $(\alpha)$, trend coefficient $(\tau)$ and variance of error term. Therefore, one structural break in level does not affect the asymptotic null distributions (Schmidt & Philips (1992), Westerlund & Edgerton (2007), Amsler & Lee (1995)).

### *3.3.1 Obtaining LM-Based Test Statistics with one or more structural breaks*

Let reconstruct the equation (3.1) under one structural break and more structural breaks,

$$\begin{aligned} y_t &= \alpha + \tau t + x_t'\beta + z_t \\ y_t &= \delta' R_t + x_t'\beta + z_t \\ z_t &= \rho z_{t-1} + e_t \ \ , \ \ X_t = X_{t-1} + w_t \end{aligned} \tag{3.33}$$

where $R_t$ covers exogeneous variables $R_t = [1, t, D_{1t}, ..., D_{Bt}, DT_{1t}, ..., DT_{Bt}]'$, where $D_{it}$ and $DT_{it}$ represent the dummy variables denoting the positions of the $i^{th}$ level (or intercept) and/ or trend breaks (or break in slope), respectively.

$$D_{it} = \begin{cases} 1 & if \ \ t \geq TB_i + 1 \ \ i = 1, ..., B \\ 0 & otherwise \end{cases}$$

$$DT_{it} = \begin{cases} t - TB_i & if \;\; t \geq TB_i + 1 \;\; i = 1, ..., B \\ 0 & otherwise \end{cases}$$

where $t = 1, 2, ..., T$ and $B$ is the maximum number of structural changes. $TB_i$ stands for the location of $i^{th}$ structural break $(1 < TB_i < T)$. "1" denotes the $\alpha$ and "$t$" stands for $\tau$ in $R_t = [1, t, D_{1t}, DT_{1t}]$. $\alpha$ and $\tau$ demonstrate deterministic components the intercept and trend before the $i^{th}$ break. $\delta$ denotes the changes in the intercept (level) and/or slope (trend) at the time of $i^{th}$ break.

The location of break $TB_i$ is an integer such that $TB_1 = \lambda_1 T$ for $i = 1$, $TB_i - TB_{i-1} = \lambda_i T$ for $i = 2, ..., B$ and $T - TB_B = \lambda_{B+1} T$ for $i = B$ where $\lambda \in [T, 1 - T]$ for $T \in (0, 1)$ denotes the location parameter which is the fraction of the ith sub-sample in each regime. Thus, the change point does not reach out too near to the beginning or end of the realization.

As in the no-break case, the LM-based (score) test statistics allowing for structural break are obtained from the following regression as in Schmidt & Philips (1992)

$$\Delta\hat{S}_t = constant + \phi\hat{S}_{t-1} + error$$

In this circumstance, in order to take account for the break, the variable $\hat{S}_t$ has some modification as shown in the following equation

$$\hat{S}_t = y_t - \hat{\delta}' R_t - x_t'\hat{\beta}$$

Cointegration test of null versus alternative is based on the OLS estimate of $\phi$ in equation (3.10) and its t-statistics. And it also has be to done some modification in equation (3.9) as shown in equation (3.34)

$$\Delta y_t = \hat{\delta}\Delta R_t + \hat{\phi}\hat{S}_{t-1} + \beta\Delta X_t' + error \tag{3.34}$$

$\hat{S}_t$ represents the LM-based detrended test of $y_t$ series, depending on the types of break. $\hat{\delta}$ is the coefficient vector obtained in the regression of equation (3.34) where the first difference of $y_t$ on $R_t$ and $x_t$.

$R_t = [1, t, D_t]'$ means that only level-shift has occurred
$R_t = [1, t, DT_t]'$ means that only trend-shift has occurred
$R_t = [1, t, D_t, DT_t]'$ shows that both level and trend break have occurred.

When $R_t = [1, t]$ is valid, then the DGP is the same as that in the no-break test of Schmidt & Phillips (1992).

According to Westerlund & Edgerton (2007), adaptation by adding dummy variable into the general model (equation 3.1, 3.2 and 3.3) could be done and this modification does not affect the asymptotic null distribution of the $t_N$ and $\phi_N$. Additionally, misplacement or exclusion or disregarding the break does not have any impact on the asymptotic null distribution of LM-based statistics computed under the assumption of no break since $T^{-1/2}D_t$ or $T^{-1/2}DT_t$ vanish as $T$ goes to infinity. Westerlund & Edgerton addressed the problem that incorrect placement or exclusion of the break make test biased towards accepting null hypothesis and it does affect the power of tests, especially reducing effect. The conclusion of authors shows that the location of the structural break might be effective on the cointegration test.

As pointed out in Westerlund & Edgerton (2007), the limiting distributions of the LM-based cointegration test statistics are identical to those of the unit root tests of Schmidt & Phillips (1992), and invariant to types of breaks and the number of regressors (Oh & Lee, 2016). However, Lee & Strazicich (2003) showed that the asymptotic distribution of the LM-based test statistic (let say $\tau_{LM}$) depends on the nuisance parameters in the presence of the trend-breaks. Im, Lee & Tieslau (2011) consider a simple transformation which can make the unit root test statistic free of the dependency on the break location.

In the light of this information, the following transformation can remove the dependency on the nuisance parameter:

$$\hat{S}_t^* = \begin{cases} \dfrac{T}{TB_1}\hat{S}_t & for \;\; t \leq TB_1 \\ \\ \dfrac{T}{TB_2 - TB_1}\hat{S}_t & for \;\; TB_1 < t \leq TB_2 \\ \vdots & \\ \dfrac{T}{T - TB_B}\hat{S}_t & for \;\; TB_B < t \leq T \end{cases} \tag{3.35}$$

$\hat{S}_t^*$ is the transformed series. If $\hat{S}_{t-1}^*$ replaces into the regression in equation (3.34) instead of $\hat{S}_t^*$, $\tau_{LM}^*$ can be t-statistic (null hypothesis $\phi = 0$) for a new testing regression

$$\Delta y_t = \hat{\delta}\Delta R_t + \hat{\phi}\hat{S}_{t-1}^* + \beta\Delta X_t' + error \tag{3.36}$$

With these modifications, the asymptotic distributions of the new test statistic $\tau_{LM}^*$ will be invariant to the nuisance parameter $\lambda_i$ (Lee,Strazicich, & Meng (2012)). $\tau_{LM}^*$ depends only on the number of trend breaks, since the distribution is given as the sum of $B+1$ with structural changes evenly distributed or $\lambda_i = \dfrac{i}{B+1}$. Therefore, there is no need to simulate critical values at all possible change point arrangements.

### *3.3.2 Asymptotic Distribution of LM-Based Cointegration test in presence of one or more structural breaks*

For the simplicity, let consider the one-time structural break in intercept and trend in the equation (3.35).

The first step testing regression (3.36) can be alternatively written as:

$$\Delta y_t = \tau + \beta\Delta X_t' + \delta_3 D_{1t} + \delta_4 D_{2t} + u_t \tag{3.37}$$

$D_{1t}$ denote $D_{1t} = 1 \;\; t \leq TB$ as before the break and 0 otherwise, $D_{2t} = 1 \;\; t \geq TB + 1$ as after the break.

Since $\tau$ and $\beta \Delta X_t'$ asymptotically converge to $o_p(1)$ by similar arguments used in the proof of Theorem 1, these variables are negligible without a loss of generality:

$$\Delta y_t = \delta_3 D_{1t} + \delta_4 D_{2t} + u_t$$

For $t \leq T_B$,

$$\hat{\delta}_3 = \frac{1}{TB-1}\sum_{t=2}^{TB} \Delta y_t = \frac{1}{TB-1}\sum_{t=2}^{TB}(\delta_3 D_{1t} + \delta_4 D_{2t} + u_t) = \delta_3 + \frac{1}{TB-1}\sum_{t=2}^{TB} u_t$$

and

$$(\hat{\delta}_3 - \delta_3) = \frac{1}{TB-1}\sum_{t=2}^{TB} u_t$$

$$\bar{u}_t = T^{-1}\sum_{t=1}^{T} a_t = \bar{a}_t \qquad \sqrt{T}\bar{u}_t \to N(0, \sigma^2)$$

Under the Functional Central Limit Theorem and Continuous Mapping Theorem, (see Appendix 2 for detailed information)

$$\lambda = \frac{TB}{T} \text{ in } (\hat{\delta}_3 - \delta_3) = \frac{1}{TB-1}\sum_{t=2}^{TB} u_t = \frac{1}{\lambda T}\sum_{t=1}^{\lambda T} u_t$$

Asymptotic distribution of $u_t$ up to $\lambda T$;

$$\sqrt{T}(\hat{\delta}_3 - \delta_3) = \frac{1}{\lambda}\frac{1}{T}\sqrt{T}\sum_{t=1}^{\lambda T} u_t = \frac{1}{\lambda}\frac{1}{\sqrt{T}}\sum_{t=1}^{\lambda T} u_t = \frac{1}{\lambda}\frac{\sqrt{\lambda T}}{\sqrt{T}}\frac{1}{\sqrt{\lambda T}}\sum_{t=1}^{\lambda T} u_t$$

$$\sqrt{T}(\hat{\delta}_3 - \delta_3) \to \sigma \frac{W(\lambda)}{\lambda} \tag{3.38}$$

Similarly, we can obtain

$$\begin{aligned}\delta_4 &= \frac{1}{T-TB}\sum_{t=TB+1}^{T}\Delta y_t \\ &= \frac{1}{T-TB}\sum_{t=TB+1}^{T}(\delta_3 D_{1t} + \delta_4 D_{2t} + u_t) \\ &= \delta_4 + \frac{1}{T-TB}\sum_{t=TB+1}^{TB} u_t\end{aligned}$$

So

$$\sqrt{T}(\hat{\delta}_4 - \delta_4) \to \sigma\frac{W(1-\lambda)}{(1-\lambda)} = \sigma\sqrt{1-\lambda}W(1) \tag{3.39}$$

Finally, the asymptotic null distribution is

$$T^{-2}\sum_{t=2}^{T}\hat{S}_t^2 \to \sigma^2\int_0^1 V(r)^2 dr = \sigma^2\left[\lambda\int_0^{\lambda} V(r/\lambda)^2 dr + (1-\lambda)\int_{\lambda}^{1} V((r-\lambda)/(1-\lambda))^2 dr\right]$$

$$= [part\ 1\ +\ part\ 2]$$

For $part\ 1$

$$\begin{aligned}\frac{r}{\lambda} =& r^* & dr^* =& \frac{1}{\lambda}dr \Rightarrow \lambda dr^* = dr \\ r =& 0 \quad \frac{r}{\lambda} = r^* \to 0 & r =& 1 \quad \frac{\lambda}{\lambda} = r^* \to 1\end{aligned}$$

For $part\ 2$

$$\begin{aligned}\frac{r-\lambda}{1-\lambda} =& r^+ & dr^+ =& \frac{1}{1-\lambda}dr \Rightarrow (1-\lambda)dr^+ = dr \\ r =& \lambda \quad \frac{\lambda-\lambda}{1-\lambda} = 0 & r =& 1 \quad \frac{1-\lambda}{1-\lambda} = 1\end{aligned}$$

Thus we can rewrite equation,

$$T^{-2}\sum_{t=2}^{T}\hat{S}_t^2 \to \sigma^2\left[\lambda^2\int_0^1 V_1(r^*)^2dr^* + (1-\lambda)^2\int_0^1 V_2(r^+)^2dr^+\right] \tag{3.40}$$

Gregory & Hansen (1996) and Westerlund & Edgerton (2007) proposed cointegration tests with one structural break. Generally the break locations and number of break are rarely known as a priori. Ben-David, Lumsdaine, and Papell (1999) indicate that many time series might contain more than one structural break. Hence assuming only one structural break leads to problems about modelling and estimating.

Oh & Lee (2016) proved that the significant losses of power in the WE-LM(2007) cointegration tests when potential multiple breaks are ignored. They extend the WE-LM(2007) cointegration test by allowing for an unknown number of breaks. Then developed modified testing procedures which is restore power loss from the WE-LM(2007) cointegration tests reasonably in the presence of more than one break without affecting the properties of tests under the null.

For the case of multiple structural breaks, equations (3.34), (3.35) and (3.36) are still convenient, because they constructed for easy adaptation to multiple breaks whether in level term an/or in trend term in the series.

In the case of multiple changes in the series, equation (3.37) has adapted to equation (3.41) as follows:

$$\Delta y_t = \tau + \beta\Delta X_t' + \sum_{i=1}^{B}\delta_{3i}D_{it} + \sum_{i=1}^{B}\delta_{4i}DT_{it} + u_t \tag{3.41}$$

$\tau$ and $\beta\Delta X_t'$ are negligible without a loss of generality, the following testing regression based on LM (score) principle:

$$\hat{S}_t = y_t - \sum_{i=1}^{B}\hat{\delta}_{3i}D_{it} - \sum_{i=1}^{B}\hat{\delta}_{4i}DT_{it} \tag{3.42}$$

$$\Delta \hat{S}_t = \Delta y_t - \sum_{i=1}^{B} \hat{\delta}_{3i} D_{it} - \sum_{i=1}^{B} \hat{\delta}_{4i} DT_{it} \tag{3.43}$$

Brownian bridge can be defined as following equation for multiple structural breaks.

Let define $V_i(r)$, the weak limit of the partial sum residual process $\hat{S}_t$ in equation (3.42) as follows:

$$V_i^*(r) = \begin{cases} \sqrt{\lambda_1^*} V_1(r/\lambda_1) & for \;\; r \leq \lambda_1 \\ \\ \sqrt{\lambda_2^*} V_2[(r-\lambda_1)/(\lambda_2-\lambda_1)] & for \;\; \lambda_1 < r \leq \lambda_2 \\ \vdots \\ \sqrt{\lambda_{R+1}^*} V_{R+1}[(r-\lambda_R)/(1-\lambda_R)] & for \;\; \lambda_R < r \leq 1 \end{cases} \tag{3.44}$$

Hence, using a common argument $r$, following results can be obtained:

$$T^{-2} \sum_{t=2}^{T} \tilde{S}_t^2 \rightarrow \sigma^2 \int_0^1 V(r)^2 dr$$

$$= \sigma^2 \Bigg[ \lambda_1^* \int_0^{\lambda_1^*} V(r/\lambda_1^*)^2 dr + \lambda_2^* \int_{\lambda_1^*}^{\lambda_2^*} V\Big((r-\lambda_1^*)/(\lambda_2^*-\lambda_1^*)\Big)^2 dr$$
$$+ ... + \lambda_{R+1}^* \int_{\lambda_R^*}^{1} V\Big((r-\lambda_R^*)/(1-\lambda_R^*)\Big)^2 dr \Bigg]$$

$$= \sigma^2 \sum_{i=1}^{R+1} \lambda_1^{*2} \int_0^1 V_i(r)^2 dr.$$

Under the case of multiple structural break, to get the asymptotic null distribution of the $\tau_{LM}^*$ equation (3.36) reconstructed as follows:

$$\Delta y_t = \hat{\delta} \Delta R_t + \hat{\phi} \hat{S}_{t-1}^* + error$$

$$\hat{\phi} = \Big(\hat{S}_1' H_{\Delta R} \Delta y\Big)^2 \Big(\hat{S}_1' H_{\Delta R} \hat{S}_1\Big)^{-1}$$

where $\hat{S}_1 = [\hat{S}_1, ..., \hat{S}_{T-1}]'$, $\Delta R = [\Delta R_2, ..., \Delta R_T]'$ and $H_{\Delta R} = I - \Delta R(\Delta R'\Delta R)^{-1}\Delta R'$ is the hat matrix in the regression analysis.

It can be shown that

$$T^{-2}\hat{S}_1' H_{\Delta R}\hat{S}_1 \rightarrow \sigma^2 \sum_{i=1}^{B+1} \lambda_t^{*2} \int_0^1 \underline{V}_i(r)^2 dr$$

Here, $\underline{V}_i(r)$ is the projection of the process $V_i(r)$ on the orthogonal complement of the space spanned by the trend break function $dz(\lambda^*, r)$ as defined over the interval $r \in [0, 1]$. That is,

$$\underline{V}_i(r) = V_i(r) - dz(\lambda, r)\tilde{\delta},$$

$$\tilde{\delta} = \arg\min_{\delta} \int_0^1 \Big(V_i(r) - dz(\lambda, r)\delta\Big)^2 dr.$$

$$T^{-1}\hat{S}_1' H_{\Delta R}\Delta y = T^{-1}\hat{S}_1' H_{\Delta R} e = T^{-1}\hat{S}_1' e \rightarrow -\frac{1}{2}\sigma^2$$

Accordingly, the limiting distribution of $\hat{\tau}_{LM}$ is obtained as follows:

$$\hat{\tau}_{LM} \rightarrow -\frac{1}{2}\omega \left[\sum_{i=1}^{B+1} \lambda_t^{*2} \int_0^1 \underline{V}_i(r)^2 dr\right]^{-1/2}$$

Now, when $\hat{S}_t^2$ is divided by the fraction of each sub-sample, it is easy to see that:

$$T^{-2}\sum_{t=2}^{T} \hat{S}_t^2 \rightarrow \sigma^2 \Bigg[(1/\lambda_1^*)^2 \lambda_1^* \int_0^{\lambda_1^*} V(r/\lambda_1^*)^2 dr$$
$$+(1/\lambda_2^*)^2 \lambda_2^* \int_{\lambda_1^*}^{\lambda_2^*} V\Big((r - \lambda_1^*)/(\lambda_2^* - \lambda_1^*)\Big)^2 dr$$
$$+ ... + (1/\lambda_B^*)^2 \lambda_{B+1}^* \int_{\lambda_B^*}^{1} V\Big((r - \lambda_B^*)/(1 - \lambda_B^*)\Big)^2 dr\Bigg]$$

$$= \sigma^2 \sum_{i=1}^{B+1} \int_0^1 V_i(r)^2 dr.$$

$\hat{S}_1^* = [\hat{S}_1^*, ..., \hat{S}_{T-1}^*]$ is used.

$$T^{-2}\hat{S}_1^{*\prime} H_{\Delta R}\hat{S}_1^* \rightarrow \sigma^2 \sum_{i=1}^{B+1} \lambda_t^{*2} \int_0^1 \underline{V}_i(r)^2 dr$$

Then, it can be shown that the asymptotic distributions of $\hat{\tau}_{LM}^*$ become invariant to the nuisance parameter $\lambda$ as follows:

$$\hat{\tau}_{LM}^* \rightarrow -\frac{1}{2}\left[\sum_{i=1}^{B+1} \int_0^1 \underline{V}_i(r)^2 dr\right]^{-1/2}$$

where $V_i(r) = W_i(r) - rW(1)$ and $W_i(r)$ is a Wiener process for $i = 1, ..., B$. The asymptotic distributions of the test statistic $\tau_{LM}^*$ will be invariant to the nuisance parameter $\lambda_i$.

The above result shows that, following the transformation, the unit root test statistic no longer depends on the nuisance parameter anymore. The asymptotic distribution of $\hat{\tau}_{LM}^*$ depends only on the number of trend breaks.

## 3.4 Abbreviations and Equations

Table 3.1 Equations which are used in the thesis for $S_t$

| | |
|---|---|
| 1 | (a) $S_t = e_t = y_t - \alpha - \tau t - x_t'\beta - \rho z_{t-1}$ (b) $\hat{S}_t = e_t = y_t - \hat{\alpha} - \hat{\tau} t - x_t'\hat{\beta}$ |
| 2 | $S_{t-1} = e_{t-1} = y_{t-1} - \alpha - \tau(t-1) - x_{t-1}'\beta - \rho z_{t-2}$ |
| 3 | (a) $\Delta S_t = \Delta y_t - \tau - \Delta x_t'\beta - \rho \Delta z_{t-1}$ (b) $\Delta \hat{S}_t = e_t = \Delta y_t - \hat{\tau} - \Delta x_t'\hat{\beta}$ |
| 4 | $S_1 = e_1$ and $S_1^2 = e_1^2$ |
| 5 | $\hat{S}_t = 0$ |
| 6 | $\Delta S_t = S_t - S_1$ and $\Delta S_t = S_t - S_{t-1}$ |

Table 3.2 Convergence information which are used in the thesis for $S_t$

| | |
|---|---|
| | $V(r) = W(r) - rW(1) \qquad W(1) \sim N(0,1)$ and $W(r) \sim N(0,r)$ |
| | $U(r) = V(r) - \int_0^1 V(s)ds$ |
| 1 | $T^{-1/2}\hat{S}_t \to \sigma V(r) \to O_p(1)$ |
| 2 | $T^{-2} \times \sum_{t=2}^{T} \hat{S}_{t-1}^2 \to \sigma^2 \int_0^1 W(r)^2 dr \to O_p(1)$ |
| 3 | $-\frac{z_0^2}{\sigma^2} \to o(1)$ |
| 4 | $T^{-3/2} \sum_{t=2}^{T} \hat{S}_{t-1} \to \sigma \int_0^1 W(r)dr \to O_p(1)$ |
| 5 | $T^{-1} \sum_{t=2}^{T} x_{t-1}\Delta\hat{S}_t \to O_p(1)$ |
| 6 | $\left(T^{-2} \sum_{t=2}^{T} (\hat{S}_{t-1})_p^2\right) \to \sigma^2 \int_0^1 U(r)^2 dr$ |
| 7 | $T^{-1} \sum_{t=2}^{T} (\hat{S}_{t-1})_p (\Delta\hat{S}_t)_p \to -\frac{1}{2}\sigma^2$ |
| 8 | $T^{-1} \sum_{t=2}^{T} \Delta x_t \hat{S}_{t-1} \to O_p(1)$ |
| 9 | $T^{-1/2} x_t = O_p(1)$ |
| 10 | Further, for $r \leq \lambda$, by defining $r^* = r/\lambda$, $r^* \in [0,1]$, we have:<br>$W(r) - rW(\lambda)/\lambda = W(r^*\lambda) - r^*\lambda W(\lambda)/\lambda = \sqrt{\lambda}[W(r^*) - r^*W(1)]$<br>where we define<br>$V_1(r^*) \equiv W(r/\lambda) - (r/\lambda)W(1) = W(r^*) - r^*W(1)$ |
| 11 | Further, for $r > \lambda$, by defining $r^* = (r-\lambda)/(1-\lambda)$, $r^* \in [0,1]$, we have:<br>$W(r) - (r-\lambda)W(1-\lambda)/(1-\lambda) = W(r^*(1-\lambda)) - r^*(1-\lambda)W(1-\lambda)/(1-\lambda)$<br>$= \sqrt{1-\lambda}[W(r^*) - r^*W(1)]$<br>where we define<br>$V_2(r^*) \equiv W((r-\lambda)/(1-\lambda)) - ((r-\lambda)/(1-\lambda))W(1) = W(r^*) - r^*W(1)$ |
| 12 | Combining 10 and 11<br>$V^*(r) = \sqrt{\lambda}V_1(r/\lambda)$ for $r \leq \lambda$ and $V^*(r) = \sqrt{1-\lambda}V_2((r-\lambda)(1-\lambda))$ for $r > \lambda$ |

$W(r)$ : Standard Brownian Motion on the unit interval $r \in [0,1]$

$V(r)$ : Standard Brownian Bridge

$\overrightarrow{V}(r)$ : Standard vector Brownian Bridge

$U(r)$ : Demeaned Standard Brownian Bridge

# CHAPTER FOUR
# MORE EFFICIENT ESTIMATOR: RALS APPROACH

## 4.1 Residual Augmented Least Square (RALS) Estimator

Generally, regression analysis is used as time series method by treating time as the independent variable and the time series variable as a dependent variable. As might be expected, the Ordinary Least Squares (OLS) estimation technique is a major tool for estimating the parameter of the linear regression model. The accuracy and validity of regression model depends some basic error assumptions: independence, linearity, normality, homoscedasticity. The assumption of the homoscedasticity (or same variance) is the most important issues to consider.

Several noticeable features of econometric/economic time series are well known with a fat-tailed distribution that has the large skewness or excess kurtosis. As mentioned in Chapter 2, these destructive behaviors lead to aperiodic volatility of error term and non-normality in series. In these circumstances OLS estimator is still unbiasedness and consistency, but inefficient. A violation of the efficiency would result in incorrect signs of the OLS estimates' variance, leading to confidence intervals that are too wide or too narrow.

Cointegration, unit root, causality analyses are performed by least square estimation can also be affected by the violation of homoscedasticity and normality of error term. If the error term has the destructive attitudes, researchers should either remove or fix this effect by some transformations. Another way to measure the effect of the heteroscedastic error term on the model is to reflect this effect on the model by non-linear framework with GARCH models.

In last few decades, Residual Augmented Least Squares (RALS) approach which utilizes information on the skewness and excess kurtosis is more efficient than classical OLS estimation. RALS method first introduced by Im (1996), then extended by Im & Schmidt (2008) to the functional form of estimator under non-normality of error term.

### 4.2 Literature Review of RALS

From the theoretical perspective of RALS methodology, Im & Schmidt (2008) is the pioneer of the Residual Augmented Least Squares (RALS) method. They have developed the functional form of estimator under non-normal error terms. Im, Lee & Tieslau (2014) proposed unit root tests based on RALS (2&3) and RALS (t5). Additionally, in order to increase the power of the traditional Dickey-Fuller (1979, DF) test, they adopted the RALS estimation procedure to the Dickey-Fuller test. Meng et al. (2014) improved the RALS methodology connected with LM test, hereby they gained more powerful and robust test results with non-normal errors and some arrangements of nonlinearity, however, the results are not valid for ignoring structural breaks when they exist in the series. Meng et al(2016) improved the RALS-LM test which adapted by Meng et al (2014) while allowing for trend shifts. Suggested test's results are more powerful in comparison to LM test that does not integrate the information on non-normal errors terms.

Li & Lee (2015) applied the RALS estimator to AR models to utilize the additional restrictions on the skewness and excess kurtosis moments shaped in non-normal errors. Authors indicated that the RALS forecast has the best performance even if skewedness. The gap between the RALS and OLS forecasts tends to grow as the forecast horizon rises. By contrast, the LAD estimator performs well when the data contain potential outliers. When the shape of distribution is close to normal curve, the RALS forecast outperforms the LAD forecast. Researchers also investigate how much the RALS forecasts affect by the presence of conditional heteroskedastic error terms (ARCH). Monte Carlo simulation results show that RALS estimator is still consistent, but estimator can lead to a small efficiency loss, compared to the LS estimator.

From cointegration viewpoint, Taylor & Peel (1998) has suggested RALS cointegration test, which is robust to excess skewness and kurtosis of the error term. This test extensively used to study bubbles in financial markets. Lee et al. (2015)

developed present cointegration tests Engle-Granger, Modified Engle Granger (Modified-EG), Autoregressive Distributed Lag test (ADL) and Error Correction Model (ECM) owing to RALS approach and authors suggest that resulting estimators can be more efficient and the tests may become more powerful under the non-normal error term.

In the recent papers which are from different application platform, such as finance, actuary, economy, energy and tourism. A brief summary of the studies is given in the following.

Gallagher % Taylor (1999) employed VAR (Vector Autoregressive) model of US real stock prices and nominal interest rates in order to identify the temporary and permanent components. They investigated the sensitivity of VAR decomposition in comparison of OLS, Least Absolute Deviation (LAD) and (RALS) estimation methods. The non-normal residuals in the VAR model may cause the underestimated size of the mean-reverting component. However, RALS and LAD techniques have results that are more accurate.

Taylor & Peel (1998) compared the performance of a RALS cointegration test in the presence of periodically collapsing bubbles. RALS method applied to analyze a long run relationship of US real stock price and dividend data. Authors rejected the bubbles hypothesis accordingly RALS method.

Payne & Lee (2017) investigated the convergence of per capita renewable energy consumption across the 50 U.S. states employing LM and RALS-LM unit root tests with endogeneouly determined trend-breaks to detect for stochastic convergence. Results from LM and RALS-LM unit root tests under the two breaks provide strong support for the presence of stochastic convergence in renewable energy consumption among the U.S. states.

Pierdzioch, Risse,& Rohloff (2015) analyzed cointegration structure between gold and silver prices over time, possibly reflecting the impact of bubble and financial crises. Performing of the RALS cointegration test yielded stronger evidence against no cointegration with reference to Engle-Granger cointegration test.

Meng,Strazicich,& Lee (2017) examined of the hysteresis hypothesis for 14 OECD countries by utilizing both linear and nonlinear unit root tests. They employed ADF, ADF-GLS, RALS-LM and Fourier LM tests for checking thypothesis. Authors revealed that breaks in the unemployment have permanent effects in 11 of the 14 countries. Also, this reflection of permanent effect tends toward to biasedness when structural breaks and/or non-normal errors are ignored. This means that specification error in the model and less accurate conclusions.

Kristic, Dumancic,& Arcabic (2017) is attractive working paper. The aim of paper is to determine if the Euro Area (EUA) membership had a significant impact on the unemployment rates for the EUA countries. The paper employed LM and RALS-LM unit root tests to analyze the persistence of two breaks, to test the stochastic convergence and to determine the location of structural break(s) in unemployment rates.

Heuson & Hutchinson (2011) studied that assessments of hedge fund parameters estimated using OLS are inaccurate when fund returns exhibit skewness. To remedy this problem, the RALS estimator is the best method to give a reliable results when the data are highly skewed. Study provided that choosing a robust estimator to the skewness is the most important part when evaluating the persistence of hedge fund performance.

Pierdzioch (2010) investigated the cointegration between stock price index and the ratio of dividends to stock prices. Study implied that, if series have no bubbles, stock prices and dividends should be cointegrated.

Alda,Ferruz,& Gallagher (2012) analyzed the performance of 134 Spanish pension funds using a linear and nonlinear performance models. Because funds return have nonnormality properties along with the presence of both negative skewness and positive kurtosis. RALS estimator is more appropriate to analyze of the funds' performance. They also examine the sub-periods analyses via different model and risk factors.

Ozcan & Erdogan (2015) analyzed whether Turkey's 14 major tourist source markets converges to each other within the framework of the two-step LM and three-step RALS-LM unit root tests under the structural breaks. The obtained results provide strong support for the convergence hypothesis, indicating that 10 major tourism markets of Turkey make positively increasing contribution in tourist arrivals to Turkey.

Mishra & Smyth (2017) researched convergence structure in energy consumption per capita at the sector level in Australia through the LM and RAS-LM unit root tests under the two endogenously determined breaks. Study revealed that there is stochastic conditional convergence in energy consumption per capita for six of the seven sectors.

### 4.3 Theory of RALS Estimator

This sub-section gives brief information of the RALS estimator proposed by Im & Schmidt (2008), which is robust with respect to skewness and kurtosis and any other deviation from normal assumption. The RALS estimator, which is closely related to the GMM estimator, is one of a widespread variety of alternative robust estimation methods which have more efficiency for non-normal data.

Im & Schmidt (2008) regarded the efficiency gains that are possible in least squares regression under the assumption about additional moments of the residual term. These efficiency gains can be realized in a GMM framework if the additional moments are known, but also if the moments are simply assumed to be correlated with error term but not with the regressors (Im & Schmidt (2008)).

Consider the linear regression model

$$y_t = \alpha + x_t'\beta + e_t \tag{4.1}$$

where $\alpha$ is the intercept $\beta$ is the vector of parameters. For the moment, assume that both the dependent and independent variable are stationary. Im and Schmidt (2008) are concerned with improving an estimator, which is robust to skew and kurtosis in the distribution of the error term. Error term of regression equation, $e_t$, is normally distributed and independent of the other variables. $E(x_t e_t) = 0$.

Skewness implies a non-zero standardized third central moment

$$E(e_t^3 - \sigma^3) = E[e_t(e_t^2 - \sigma^2)] \neq 0$$

Similarly, excess kurtosis implies that the standardized fourth central moment of the series exceeds three, so that

$$E(e_t^4 - 3\sigma^4) = E[e_t(e_t^3 - 3\sigma^2 e_t)] \neq 0$$

Implying that $(e_t^3 - 3\sigma^2 e_t)$ is correlated with $e_t$ but not correlated independent variables which implies that $(e_t^2 - \sigma^2)$ is correlated with $u_t$ but not with the independent variables. Im and Schmidt (2008) suggested a two-step estimator which can be simply computed from OLS estimation method in respect to equation (4.1).

$$\hat{w}_t = [(\hat{e}_t^3 - 3\hat{\sigma}^2\hat{e}_t)(\hat{e}_t^2 - \hat{\sigma}^2)]'$$

$$y_t = \alpha + x_t'\beta + \gamma'\hat{w}_t + u_t$$

where $\hat{e}_t$ denotes the residual and $\hat{\sigma}^2$ the standard residual variance estimate obtained from equation (4.1). When both the independent and the dependent variables are stationary, the (RALS) estimator of $\beta, \beta^*$ say, has an asymptotic distribution given by

$$\sqrt{T}(\beta^* - \beta) \rightarrow N[0, \sigma_A^2 Var(x_t)^{-1}]$$

where

$$\sigma_A^2 = \sigma^2 - \frac{\mu_3^2\left(\mu_6 - 6\mu_4\sigma^2 + 9\sigma^6 - \mu_3^2\right) - 2\mu_3\left(\mu_4 - 3\sigma^4\right)\left(\mu_5 - 4\mu_3\sigma^2\right)\left(\mu_4 - 3\sigma^4\right)^2\left(\mu_4 - \sigma^4\right)}{\left(\mu_4 - \sigma^4\right)\left(\mu_6 - 6\mu_4\sigma^2 + 9\sigma^6 - \mu_3^2\right) - (\mu_5 - 4\mu_3\sigma^2)^2}$$

where $\mu_i$ denotes the $i$th central moment of $e_t$ (Taylor & Peel, (1998)).

The variance of RALS estimator is smaller than the one, which calculated from OLS method, so it is efficient. This efficiency gain expressed as follows, if there exist some $w_t$ related with equation (4.1) such that; $E[w_t' x_t] = 0$ and $E[w_t' e_t] \neq 0$, so estimating $\alpha$ and $\beta$ parameters gain efficiency when $w_t$ is included in the model:

$$y_t = \alpha + x_t'\beta + \gamma'\hat{w}_t + u_t \text{ where } e_t = \gamma'\hat{w}_t + u_t$$

$$\begin{aligned}
Var(u_t) &= Var(e_t - \gamma'\hat{w}_t) \\
&= Var(e_t) + \gamma^2 Var(w_t) - 2Cov(e_t, \gamma' w_t) \\
&= Var(e_t) + \left(\frac{\sigma_{ew}}{\sigma_w^4}\right)^2 \sigma_w^2 - 2\left(\frac{\sigma_{ew}}{\sigma_w^2}\right)\sigma_{ew} \\
&= Var(e_t) + \frac{\sigma_{ew}^2 \sigma_w^2}{\sigma_w^4} - 2\left(\frac{\sigma_{ew}\sigma_w^2}{\sigma_w^4}\right)\sigma_{ew} \\
&= Var(e_t) + \frac{\sigma_w^2(\sigma_{ew}^2 - 2\sigma_{ew}^2)}{\sigma_w^4} = Var(e_t) - \frac{\sigma_{uw}^2}{\sigma_w^2} = Var(u_t)
\end{aligned}$$

It is easy to show that the new error $u_t$ in the regression has smaller variance than $e_t$. Testing procedure of RALS estimation method is implemented in a linear framework that relies on least squares estimation.

The RALS estimator can be interpreted as a GMM estimator based on the moment conditions $E(x_t e_t) = E[x_t(e_t^3 - \mu^3)] = E[x_t(e_t^2 - \sigma^2)] = 0$ where $\mu_i$ denotes the $i$th central moment of $e_t$. Therefore, asymptotic distribution of RALS estimator is the same as that of the GMM estimator as well as the test of Hansen (1995) who suggested including stationary covariates (Im,Lee, & Tieslau,2014).

### 4.4 RALS-LM Unit Root Test

Im et al. (2014) proposed unit root tests based on RALS (2&3) and RALS (t5). Additionally, in order to increase the power of the traditional Dickey-Fuller (1979, DF) test, they adopted the RALS estimation procedure to the Dickey-Fuller test. Meng et al. (2014) improved the RALS methodology connected with LM test.

As mentioned in Chapter 3, Schmidt & Phillips (1992) LM unit root model can be rewritten as:

$$y_t = \alpha + \tau t + x_t'\beta + z_t \qquad z_t = \rho z_{t-1} + e_t$$

The unit root (or cointegration) null hypothesis implies $\rho = 1$ against the alternative that $\rho < 1$. In general, following the LM (score) principle, the LM unit root test statistic can be obtained from the following regression:

$$\Delta y_t = \delta' \Delta Z_t + \phi S_{t-1} + e_t \tag{4.2}$$

where $S_t$ is detrended series and *t*-statistic testing the null hypothesis $\phi = 0$. Then, the LM test statistic is given by:

$\tau_{LM} = t-$statistic testing the null hypothesis $\phi = 0$.

Meng et al. (2014) improve the power of the LM test utilize the information on non-normal errors in order to improve upon the power of the unit root test, making use of the RALS estimation procedure, To begin with,

Let define $\xi_t = (\Delta S_{t-1}, \Delta S_{t-2}, ..., \Delta S_{t-p}),\ f_t = (S_{t-1}, \xi_t')'$ and $F_t = (\Delta S_{t-1}', f_t')'$.

Suppose the following moment conditions:

$$E[g(e_t) \otimes F_t] = 0 \quad t = 1, 2, 3, ..., T$$

$g(e_t)$ is a $j \times l$ vector that satisfies the following assumption.

**Assumption (Lee,Lee, Im 2015):** $g(.)$ is differentiable and satisfies the first-order Lipschitz condition $|g_j'(x) - g_j'(y)| < M|x - y|$ for some constant $M$ for all $j$, where $g_j(.)$ is the $j$th element of $g_j(.)$. Also, $E[g(e_t)] = 0$, the second moment of $g(e_t)$ exist, and $E[g'(e_t)] < \infty$.

where $g(e_t)$ is a function defined as $g(e_t) = \Big(e_t, [h(e_t) - K]'\Big)$ with $K = E(e_t)$ and $h(e_t)$ is a nonlinear function which contains second and third moment of error term. We can divide this moment conditions into two parts. The first part is the usual moment condition of least squares estimation

$$E[e_t \otimes F_t] = 0$$

The second part, based on nonlinear functions of $e_t$, contains an additional $(J-1) \times (p+2)$ moment conditions given by

$$E[(h(e_t) - K) \otimes F_t] = 0$$

Therefore we have variance covariance matrix of error and error functions can be defined,

$$C = \begin{bmatrix} \sigma_e^2 & C_{21}' \\ C_{21} & C_{22} \end{bmatrix}, \quad \text{and} \quad D = \begin{bmatrix} 1 \\ D_2 \end{bmatrix}$$

where $C_{21} = E[e_t h(e_t)]$, $C_{22} = E[h(e_t)h(e_t)']$, and $D_2 = E[h'(e_t)]$.

Let $\hat{e}_t$ denote the residuals from the usual LM regression (4.2). Following Im & Schmidt (2008), the RALS procedure augments the following term $\hat{w}_t$ to testing regression (4.2)

$$\hat{w}_t = h(\hat{e}_t) - \hat{K} - \hat{e}_t \hat{D}_2$$

where $\hat{e}_t$ is the OLS residual from regression (4.2), $\hat{K} = \dfrac{1}{T}\sum_{t=1}^{T} h(\hat{e}_t)$ and $\hat{D}_2 = \dfrac{1}{T}\sum_{t=1}^{T} h'(\hat{e}_t)$.

To capture the information of non-normal errors, they define as $h(\hat{e}_t) = [\hat{e}_t^2, \hat{e}_t^3]'$, which involves the second and third moments of $\hat{e}_t^2$. Then, letting $\hat{m}_j = T^{-1}\sum_{t=1}^{T} \hat{e}_t^j$ the augmented term can be given as

$$\hat{w}_t = \begin{bmatrix} \hat{e}_t^2 & \hat{e}_t^3 \end{bmatrix}' - \frac{1}{T}\sum_{t=1}^{T} \begin{bmatrix} \hat{e}_t^2 & \hat{e}_t^3 \end{bmatrix}' - \hat{e}_t \frac{1}{T}\sum_{t=1}^{T} \begin{bmatrix} 2\hat{e}_t^2 & 3\hat{e}_t^3 \end{bmatrix}$$

$$\hat{w}_t = \begin{bmatrix} \hat{e}_t^2 - \frac{1}{T}\sum_{t=1}^{T}\hat{e}_t^2 - 2\frac{1}{T}\sum_{t=1}^{T}\hat{e}_t \\ \hat{e}_t^3 - \frac{1}{T}\sum_{t=1}^{T}\hat{e}_t^3 - 3\frac{1}{T}\sum_{t=1}^{T}\hat{e}_t^2\hat{e}_t \end{bmatrix}$$

$$= \begin{bmatrix} \hat{e}_t^2 - T^{-1}\sum_{t=1}^{T}\hat{e}_t^2 \\ \hat{e}_t^3 - T^{-1}\sum_{t=1}^{T}\hat{e}_t^3 - 3T^{-1}\sum_{t=1}^{T}\hat{e}_t^2\hat{e}_t \end{bmatrix} = \begin{bmatrix} \hat{e}_t^2 - m_2 \\ \hat{e}_t^3 - m_3 - 3m_2\hat{e}_t \end{bmatrix}$$

Thus $\hat{w}_t$ can be obtained as RALS(2&3) test as follows,

$$\hat{w}_t = [\hat{e}_t^2 - \hat{m}_2, \hat{e}_t^3 - \hat{m}_3 - 3\hat{m}_2\hat{e}_t]' \tag{4.3}$$

The new test's name is the "RALS(2&3)" test. The first term in $\hat{w}_t$ is associated with the moment condition $E[(e_t^2 - \sigma_e^2)y_{t-1}] = 0$. The second term in $\hat{w}_t$ improves efficiency unless $\mu_4 = 3\sigma^4$, where $\mu_j = E(e_t^j)$. This condition improves the efficiency of the estimator of $\phi$ when the distribution of error terms are not symmetric.

Generally, RALS approach is adaptable for developing various test under some conditions. It is possible to use higher moments using $h(\hat{e}_t) = [\hat{e}_t^2, \hat{e}_t^3, \hat{e}_t^4, ..., \hat{e}_t^k]$ with $k > 3$ and properly defined $\hat{w}_t$ in model that correspond to the higher moments. This extension requires the assumption that the higher moments exist. Authors leave it as a future research and accept power gain is already significant for RALS (2&3) test when using the augmented terms.

Error function of $h(e_t)$ is an appropriate function to obtain efficiency gain and non-normal information of data, whereas some other nonlinear functions can be used such as $exp(.)$ and $log(.)$ depend on the characteristics of the series non-normality.

The RALS-LM test statistic is obtained from the regression

$$\Delta y_t = \delta'\Delta Z_t + \phi\hat{S}_{t-1} + \gamma'\hat{w}_t + u_t \tag{4.4}$$

where the corresponding $t$-statistic for $\phi = 0$ is $\tau^*_{RALS-LM}$. One may relate equation (4.2) to this regression with $e_t = \gamma' \hat{w}_t + u_t$ where $\hat{w}_t$ is uncorrelated with $u_t$, as proved in Li & Lee (2015). We take $E(u_t^2) = E(e_t^2) - [E(e_t u_t)]^2 / E(\hat{w}_t^2)$ since $\gamma = E(\hat{w}_t e_t)/E(\hat{w}_t^2)$. Thus, $E(e_t^2) < E(u_t^2)$ implying that the variance of the error term in (4.4) is smaller than that in (4.2). This result yields the evidence of the asymptotic efficiency gain (thus, increase in power of the test) with non-normal errors (Meng et. al.(2014)).

The asymptotic distribution of the new test can be given as

$$\tau_{RALS-LM} \rightarrow \rho \tau_{LM} + \sqrt{1-\rho^2} Z$$

where $\rho \tau_{LM}$ denotes the limiting distribution of the t-statistic for the usual LM estimator in regression (4.2), and $\rho$ is the correlation between two error terms such that $\rho^2 = E(u_t^2)/E(e_t^2)$.

In the presence of normal error terms, in case $\rho^2 = 1$, RALS is asymptotically identical to the OLS estimator and there is no efficiency gain. Also if correlation term equals one ($\rho = 1$) RALS-LM test is same as classical LM test ($\tau_{RALS-LM} = \tau_{LM}$). Asymptotic distribution of RALS-LM unit root test is same as of the GMM estimator as well as the test of Hansen(1995) who suggested including stationary covariates (Im,Lee,& Tieslau 2014).

### 4.5 RALS Cointegration Test

The pioneer study of RALS cointegration test was suggested Taylor & Peel (1998). This test extensively used to study bubbles in financial markets (C. Pierdzioch (2010), C. Blancard & H. Raymond (2004), Kizys & Pierdzioch(2012)). Taylor & Peel (1998) test is adopted to cointegration test when the error term of an Engle Granger cointegration regression shows signs of skewness and excess kurtosis.

Remarkable paper which is produced by Lee et. al.(2015) extended present cointegration tests which are single equation model framework (Engle Granger, Modified Engle Granger, Autoregressive distributed lag test, Error Correction Model) owing to RALS approach and authors revealed that the estimators can be more efficient and the tests may become more powerful under the non-normal error term. Additionally, authors provided critical value of the tests and compared the performances with reference the former tests. In comparison of the empirical size of test and RALS test, results showed that correct sizes are close to 5% regardless of the shape of error distribution, either normal or various cases of non-normal distributions.

When the authors examined the power property of the proposed tests, they realized that the power of the RALS based tests had increased drastically in all cases when the error term follows any type of non-normal distribution. This result was core findings of RALS approach.

In application, most of the time series data have structural breaks (because of policy changes, financial crises, natural disasters etc.) and heteroscedasticity. In the cointegration framework, especially, neglecting structural breaks and non-normal error term induces spurious rejection and the performances of conventional cointegration tests are negatively affected. Therefore, researchers may make false decisions. Although the RALS cointegration test in the literature are robust, performance under both structural breaks and GARCH disturbances has not studied yet. RALS cointegration test which is subject to double destructive effects (presence of structural breaks and GARCH (p,q) effect ) has improved in this thesis.

## 4.6 Improved version of RALS-LM Test named as RALS(2)-LM GARCH Cointegration Test

A new cointegration test that allows for structural breaks in both the intercept and the slope has improved by adopting RALS-LM procedure to gain improved power when the error term has a GARCH effect.

As mentioned in the previous section, LM based tests' distributions under both the null and alternative hypothesis are independent of the intercept $(\alpha)$, trend coefficient $(\tau)$. Therefore, one structural break in level does not affect the asymptotic null distributions (Schmidt & Philips (1992), Westerlund & Edgerton (2007), Amsler & Lee (1995)).

The previous studies of cointegration test with structural break show that most tests suffered from reducing power if the number of structural breaks is more than one. (Oh & Lee, 2016) addressed the importance of power losses in the WE-LM(2007) cointegration tests when potential multiple breaks are ignored. Oh & Lee (2016) expanded the WE-LM(2007) cointegration test by allowing for more than one unknown level and slope breaks. However, in slope breaks, Lee & Strazicich (2003) showed that the asymptotic distribution of the LM-based test statistic (let say $\tau_{LM}$) depends on the nuisance parameters in the presence of the trend-breaks. Im et al.(2011) consider a simple transformation which can make the unit root test statistic free of the dependency on the break location.

Most of the time series have some negative properties such as excess kurtosis, unexpected volatility periods and skewed or fat tailed distribution. These features lead to volatility of error term and non-normality in series. Cointegration and unit root test which are based on least squares can be adversely affected by the presence of heteroskedasticity and non-normality of error term. If the series contain this negative error term components, models should be established take into account these effects.

Residual Augmented Least Squares (RALS) approach first introduced the by Im (1996), then Im & Schmidt (2008) which is utilize information on the moment of the error term. The RALS procedure utilizes the information that exists when the errors in the testing equation exhibit any departures from normality, such as non-linearity, heteroscedaticity, asymmetry, or fat-tailed distributions. Our work is extended to the cointegration testing procedure using the results in Im & Schmidt (2008),Meng et al(2016) and Oh & H.Lee (2016) under the breaks and GARCH effects. Li & Lee (2015) examined the performance of the RALS forecast. They found that RALS estimator still consistent but small efficiency loss under the ARCH disturbances.

RALS approach is easy adaptable to developing various tests under some conditions about error distribution. The classical RALS(2&3) test is well-known, but some assumptions are required for GARCH innovations.

In this thesis, RALS(2)-LM cointegration test which is suitable for GARCH effects under the level or/and slope structural breaks is developed. Suggested test does not need to additional assumptions like classical RALS(2&3) test under the GARCH effects. This test is thought as the adaptation of "RALS(2)-LM GARCH" cointegration test since it is an extension and combination of the WE-LM(2007) and RALS unit root test proposed by Meng et al(2016).

## 4.7 RALS(2)-LM Test under the presence Structural Breaks and GARCH(p,q) Disturbances

Define data generating process of cointegration equations under the structural breaks and GARCH(1,1) innovation as follows;

$$y_t = \delta' R_t + x_t'\beta + z_t \tag{4.5}$$

$$\begin{gathered} z_t = \rho z_{t-1} + e_t \quad e_t = \eta_t \sqrt{v_t} \quad X_t = X_{t-1} + w_t \\ v_t = \theta_0 + \theta_1 e_{t-1}^2 + \varphi_1 v_{t-1} \quad \eta_t \sim N(0,1) \end{gathered} \tag{4.6}$$

where $R_t$ covers exogeneous variables $R_t = [1, t, D_{1t}, ..., D_{Bt}, DT_{1t}, ..., DT_{Bt}]'$, where $D_{it}$ and $DT_{it}$ represent the dummy variables denoting the positions of the $i^{th}$ level (or intercept) and/ or trend breaks (or break in slope), respectively.

$$D_{it} = \begin{cases} 1 \quad if \ \ t \geq TB_i + 1 \ \ i = 1, ..., B \\ 0 \quad otherwise \end{cases}$$

$$DT_{it} = \begin{cases} t - TB_i & if \ \ t \geq TB_i + 1 \ \ i = 1, ..., B \\ 0 & otherwise \end{cases}$$

where $t = 1, 2, ..., T$ and $B$ is the maximum number of structural changes. $TB_i$ stands for the location of $i^{th}$ structural break $(1 < TB_i < T)$. "1" denotes the $\alpha$ and "$t$" stands for $\tau$ in $R_t = [1, t, D_{1t}, DT_{1t}]$. $\alpha$ and $\tau$ demonstrate deterministic components the intercept and trend before the $i^{th}$ break. $\delta$ denotes the changes in the intercept (level) and/or slope (trend) at the time of $i^{th}$ break.

The location of break $TB_i$ is an integer such that $TB_1 = \lambda_1 T$ for $i = 1$, $TB_i - TB_{i-1} = \lambda_i T$ for $i = 2, ..., B$ and $T - TB_B = \lambda_{B+1} T$ for $i = B$ where $\lambda \in [T, 1 - T]$ for $T \in (0, 1)$ denotes the location parameter which is the fraction of the $i$th sub-sample in each regime. Error term include in $z_t$ equation has distributed as GARCH(1,1) effect. For GARCH parameters, $\eta_t \sim WN(0, \sigma^2 = 1)$. $\eta_t$ and $e_{t-1}$ independent each other. In addition $\theta_1$, $\varphi_1$ are constants, and $\theta_0 > 0$ it is assumed that $e_t$ and $w_t$ are uncorrelated for all $t$.

The unit root test statistics are then obtained from the following regression

$$\Delta y_t = \delta \Delta R_t + \phi \hat{S}_{t-1} + \beta \Delta X_t' + e_t$$

$\hat{S}_t$ represents the LM-based detrended test of $y_t$ series, depending on the types of break. The location of break $TB_i$ is an integer such that $TB_i = \lambda T$, where $\lambda \in [T, 1 - T]$ and $T \in (0, 1)$.

The error $z_t$ becomes stationary when $x_t$ and $y_t$ are cointegrated but have a unit root process if the two series are not cointegrated. Then, the null of no cointegration can be tested using the null of $\rho = 1$ versus the alternative of $|\rho| < 1$ from (4.6). Unit root parameter $\rho = 1$ can be also tested using the null of $\phi = 0$ against the alternative of $\phi < 0$ from the following regression based on the LM (score) principle:

$$\Delta \hat{S}_t = d_t + \phi \hat{S}_{t-1} + e_t \tag{4.7}$$

where $d_t$ is the constant term and $\hat{S}_t$ is a detrended series depending on the types of breaks.

Parameters estimates are performed restricted MLE and OLS as in WE-LM(2007). Now the variable $\hat{S}_t$ is defined as

$$\hat{S}_t = y_t + \hat{\delta}' R_t - x_t' \hat{\beta} \tag{4.8}$$

We consider modified cointegration tests in a single equation model framework.

The following transformation can remove the dependency on the nuisance parameter:

$$\hat{S}_t^* = \begin{cases} \dfrac{T}{TB_1}\hat{S}_t & for\ \ t \leq TB_1 \\ \\ \dfrac{T}{TB_2 - TB_1}\hat{S}_t & for\ \ TB_1 < t \leq TB_2 \\ \vdots & \\ \dfrac{T}{T - TB_K}\hat{S}_t & for\ \ TB_K < t \leq T \end{cases}$$

where $\hat{S}_t$ is the untransformed series and $\hat{S}_t^*$ is the transformed series. New testing regression $\Delta y_t = \delta \Delta R_t + \phi \hat{S}_{t-1}^* + \beta \Delta X_t' + e_t$ and denote as $\tau_{LM}^*$ as the $t$-statistic testing the null hypothesis $\phi = 0$. Then, the asymptotic distributions of the test statistic $\tau_{LM}^*$ will be invariant to the nuisance parameter $\lambda$. (Lee, Strazicich, & Meng (2012)). Thanks to this transformation asymptotic distribution of $\tau_{LM}^*$ depend only on the number of breaks, since the distribution is given as the sum of $B + 1$ independent stochastic term. Based on the discussion mentioned chapter three, it is reasonable to believe the results of transformed test outcome are more accurate.

Two-step procedure is adopted to the "Residual Augmented Least Squares" (RALS) method, which can enable of nonlinear moment conditions driven by GARCH innovations to use.

With ARCH(1) distributed. Let $e_t^2$

$$e_t^2 = (\eta_t \sqrt{v_t})^2 = \theta_0 + \theta_1 e_{t-1}^2$$

$$h(e_t) = [\theta_0 + \theta_1 e_{t-1}^2]$$

$$\begin{aligned} w_t &= [\theta_0 + \theta_1 e_{t-1}^2] - T^{-1} \sum_{t=1}^{T} [\theta_0 + \theta_1 e_{t-1}^2] - e_t T^{-1} \sum_{t=1}^{T} [\theta_0 + \theta_1 e_{t-1}^2] \\ &= [\theta_0 + \theta_1 e_{t-1}^2] - T^{-1} \sum_{t=1}^{T} [\theta_0 + \theta_1 e_{t-1}^2] - e_t T^{-1} \sum_{t=1}^{T} [2\theta_1 e_{t-1}] \\ &= e_t^2 - m_2 - e_t T^{-1} \sum_{t=1}^{T} [2\theta_1 e_{t-1}] \end{aligned}$$

$$e_t T^{-1} \sum_{t=1}^{T} [2\theta_1 e_{t-1}] = T^{-1}(e_0 + e_1 + e_2 + ... + e_{T-1}) = 0$$

Thus,

$$w_t = e_t^2 - m_2 \text{ where } m_j = T^{-1} \sum_{t=1}^{T} e_t^j.$$

RALS(2) test with $h(e_t) = [e_t^2]'$ is adopted to implement the RALS procedure to test for cointegration. The expand procedure for ARCH(p) process, model can be defined as;

$$\begin{aligned} v_t &= \theta_0 + \theta_1 e_{t-1}^2 + \theta_2 e_{t-2}^2 + ... + \theta_p e_{t-p}^2 \\ \eta_t &\sim N(0,1) \\ h(e_t) &= [\theta_0 + \theta_1 z_{t-1}^2 + \theta_2 z_{t-2}^2 + ... + \theta_p z_{t-p}^2] \end{aligned}$$

$$
\begin{aligned}
\hat{w}_t &= [\theta_0 + \theta_1 e_{t-1}^2] - T^{-1}\sum_{t=1}^{T}[\theta_0 + \theta_1 e_{t-1}^2] - e_t T^{-1}\sum_{t=1}^{T}[\theta_0 + \theta_1 e_{t-1}^2 + ... + 2\theta_p e_{t-p}^2] \\
&= [\theta_0 + \theta_1 e_{t-1}^2] - T^{-1}\sum_{t=1}^{T}[\theta_0 + \theta_1 e_{t-1}^2] - e_t T^{-1}\sum_{t=1}^{T}[2\theta_1 e_{t-1}] + ... + [2\theta_p e_{t-p}] \\
&= e_t^2 - m_2 - e_t T^{-1}\sum_{t=1}^{T}[2\theta_1 e_{t-1}] + [2\theta_2 e_{t-2}] + ... + [2\theta_p e_{t-p}] \\
&= e_t^2 - m_2 - e_t T^{-1}\left(\sum_{t=1}^{T}[2\theta_1 e_{t-1}] + \sum_{t=1}^{T}[2\theta_2 e_{t-2}] + ... + \sum_{t=1}^{T}[2\theta_p e_{t-p}]\right) \\
&= e_t^2 - m_2 - e_t T^{-1}(0 + 0 + ... + 0) \\
&= e_t^2 - m_2
\end{aligned}
$$

For GARCH(1,1) disturbances model can be obtained as,

$$
\begin{aligned}
v_t &= \theta_0 + \theta_1 e_{t-1}^2 + \varphi v_{t-1} \\
\eta_t &\sim N(0,1) \\
e_t^2 &= (\eta_t \sqrt{v_t})^2 = \theta_0 + \theta_1 e_{t-1}^2 + \varphi v_{t-2} \\
h(e_t) &= [\theta_0 + \theta_1 e_{t-1}^2 + \varphi v_{t-1}]
\end{aligned}
$$

$$
\begin{aligned}
\hat{w}_t &= [\theta_0 + \theta_1 e_{t-1}^2] - T^{-1}\sum_{t=1}^{T}[\theta_0 + \theta_1 e_{t-1}^2 + \varphi v_{t-1}] - e_t T^{-1}\sum_{t=1}^{T}[\theta_0 + \theta_1 e_{t-1}^2 + \varphi v_{t-1}] \\
&= [\theta_0 + \theta_1 e_{t-1}^2] - T^{-1}\sum_{t=1}^{T}[\theta_0 + \theta_1 e_{t-1}^2 \varphi v_{t-1}] - e_t T^{-1}\sum_{t=1}^{T}[2\theta_1 e_{t-1}] + ... + [2\theta_p e_{t-p}] \\
&= e_t^2 - m_2 - e_t T^{-1}\sum_{t=1}^{T}[2\theta_1 e_{t-1}] + ... + [2\theta_p e_{t-p}] \\
&= e_t^2 - m_2 - e_t T^{-1}\left(\sum_{t=1}^{T}[2\theta_1 e_{t-1}] + \sum_{t=1}^{T}[2\theta_2 e_{t-2}] + ... + \sum_{t=1}^{T}[2\theta_p e_{t-p}]\right) \\
&= e_t^2 - m_2 - e_t T^{-1}(0 + 0 + ... + 0) \\
&= e_t^2 - m_2
\end{aligned}
$$

Asymptotic distribution of RALS(2)-LM GARCH cointegration test is same as that of the GMM estimator as well as the test of Hansen (1995) who suggested including stationary covariates. The basic benefit of the RALS(2)-LM GARCH cointegration tests lies in the invariance feature that the distribution does not depend on the different degree of GARCH parameter in the presence of multiple level-breaks.

Generally cointegration equations as follows,

$$y_t = \delta' R_t + x_t' \beta + z_t$$

$$\hat{w}_t = [e_t^2 - m_2]'$$

In light of these information cointegration equation can be rewriting as,

$$y_t = \hat{\delta}' R_t + x_t' \hat{\beta} + \gamma' \hat{w}_t + u_t$$

The first term in $\hat{w}_t$ is associated with the moment condition $E[(e_t^2 - \sigma_e^2) y_{t-1}] = 0$ where $e_t$ denotes the residual and $\hat{\sigma}^2$ the standard residual variance estimate obtained from ordinary least squares applied to Equation (4.1). When both the independent and the dependent variables are stationary, the resulting RALS estimator of $\beta, \beta^*$ say, has an asymptotic distribution given by

$$\sqrt{T}(\beta^* - \beta) \rightarrow N[0, \sigma_A^2 Var(x_t)^{-1}]$$

where

$$\sigma_A^2 = \sigma^2 - \left( \frac{\mu_3^2}{\mu_4 - \sigma^4} \right)$$

where $\mu_i$ denotes the $i$th central moment of $e_t$. (Taylor & Peel, 1998)). In this thesis it is followed the framework of generalized methods of moments (GMM) and sought to find more powerful tests by utilizing moment conditions implied by heteroscedastic errors. It is easy to show that the new error $u_t$ in the regression has smaller variance than $z_t$. RALS estimator variance is smaller than the OLS estimators. For this reason RALS cointegration test is more powerful than the classical LM based tests. Derivative of this efficiency gain is mentioned in Section 4.3.

## 4.8 Application process of RALS(2)-LM GARCH(p,q) Test

In this section, how the suggested cointegration test can be applied will be explained in detailed. This test can be implemented in four steps

- In the first step, the test procedure estimates the maximum likelihood and least-squares methods in order to get LM score for and determining the number and location of level and trend breaks. In order to find break location, the test uses Bai and Perron (1998) methodology, which utilizes grid search procedures to determine the location endogenously. This method examines each possible combination of B break point over the time interval, and choses break points by minimizing the SSR from the test regression.
- In the second step, the method performs the transformation proposed by Im et al. (2011) to the data set, which can make the cointegration test statistic free of dependency on the break locations.
- In the third step, the information of heteroscedastic errors are captured through Residuals-Augmented Least Squares (RALS) regression, using the estimated residuals from the cointegration relation
- In the fourth and final step, test procedure constitutes the RALS based test statistics and variance, and calculates the efficiency gain. As a result, null of no cointegration can be tested using the null of $\phi = 0$ versus the alternative of $\phi < 0$, and thus the error $z_t$ becomes stationary when $x_t$ and $y_t$ are cointegrated but have a unit root process, if the two series are not cointegrated.

# CHAPTER FIVE
# MONTE CARLO EXPERIMENT FOR RALS(2)-LM GARCH TEST UNDER STRUCTURAL BREAK AND GARCH EFFECT

## 5.1 Introduction

Conducting cointegration tests are often a core part for the interpretation of the empirical results for economic theories. The core stems from the basic idea stating that a linear combination of economic time series could be stationary even though these are non-stationary. In this case, there should be some long-run equilibrium relation between individual variables when combined. Engle & Granger (1987) used the term "cointegration" to describe this long-run relationship and provided a new theoretical base for analyzing non-stationary time series.

Most of the economic variables such as inflation, interest rates, exchange rates, real GDP and so forth appear to be highly persistent, and are typically described as unit root or non-stationary processes. Failure to allow for an existing break may lead to a bias that reduces the ability of a unit root test to reject a false unit root null hypothesis. Many cointegration tests have been developed in the literature in order to avoid inference problems in misleading decisions in cointegration under structural break. Gregory & Hansen (1996) and Campos et al. (1996) propose cointegration tests that allow for a structural break of unknown timing. Harris & Inder (1996), Lee & Strazicich (2001), and Carrion-I-Silvestre & Sanso (2006) develop tests that search for the null of cointegration with one structural break in the level and slope. Recent research has been conducted considering the finite-sample properties of some cointegration test allowing for structural break in the cointegration relationship when applied to independent series subject to either breaks in level or slope. (Höglund & Östermark, (2003); Noh & Kim, (2003); Leybourne & Newbold, (2003); Cook, (2004); Westerlund & Edgerton, (2007) Hillebrand & Medeiros, (2008); Tam, (2012) etc.)

These developed cointegration tests rely upon the strict unit-root assumption which is not easy to argue using economics theory. In fact, the structural break or regime shift is not the only problem when analyzing most economic or financial time series; another is inherent heteroscedasticity. In one approach to the problem, Busetti & Taylor (2003), Seo (2007); Herwartz & Lütkepohl (2011) examined stationarity test and cointegration parameters under conditional heteroscedasticity, respectively.

This study focuses on the existing empirical literature by further re-examining the efficiency gain and empirical size of Engle-Granger (EG), Gregory-Hansen (GH), Westerlund-Edgerton Langrange Multiplier (WE-LM) and RALS(2)-LM GARCH tests. The EG(1987) test was chosen because it is the pioneer cointegration test. The GH test which takes into account breaks in cointegration relationship is significant in literature. The WE-LM test not only has desirable features that render them straightforward and easy to implement, but also considers structural break. Lastly, we perform suggested RALS(2)-LM GARCH cointegration test which take into account structural breaks and GARCH disturbances in the series or error terms.

This paper aims to examine the behavior of the aforementioned tests under the presence of structural break and heteroscedastic innovations, because these issues may lead to invalid inference drawn from a cointegration test that does not take them into account. It is assumed that structural break in cointegration relationship can be modeled by level shift and regime shift. The sensitivity to sample size and some conditions which may affect finite-sample properties of EG(1987), GH(1996), WE-LM and RALS(2)-LM GARCH tests have been measured.

The remaining sections of the chapter have been organized accordingly. Section 5.2 establishes the construction of the experimental design for revealing the impact of break location, break magnitude and dependency degree parameters. In addition, the results of the Monte Carlo experiment in terms of finite sample size are discussed in this section. We have tried to clarify whether they are conservative or liberal tests. Finally, some concluding remarks are offered in the conclusion.

## 5.2 Finite-Sample Evidence-Monte Carlo Experiment

### *5.2.1 The test setting*

The impact of structural break and GARCH(1,1) innovations on the performances of the cointegration tests is examined in this study. Analogously the research conducted by Leybourne & Newbold, Cook and Tam, also this thesis is focused on size properties of cointegration tests under structural break with heteroskedastic innovations.

Simulation design with structural break and heteroscedastic innovations (GARCH (1, 1));

$$Y_t = bDU_t(\tau_y) + u_{yt} \quad u_{yt} = u_{yt-1} + \varepsilon_{yt}$$
$$X_t = bDU_t(\tau_x) + u_{xt} \quad u_{xt} = u_{xt-1} + \varepsilon_{xt}$$

$$e_{xt} = e_{yt} = \eta_t \sqrt{v_t} \quad v_t = \theta_0 + \theta_1 e_{t-1}^2 + \varphi v_{t-1}$$

where $\theta_0 \geq 0$ and $\theta_1 \geq 0$ are constants, $\eta_t \sim IID(0,1)$ and $\eta_t$ is independent of $\{e_{t-k}\},\ k \geq 1$ for all $t$. The bivariate process of simulation design provides the covariance-stationary feature (Sáez & Kunst, 1995).

Errors of model $(e_{yt}, e_{xt})$ are mutually independent standard normally distributed white noise processes. $DU_t(.), DT_t(.)$ is the break function of the time series, with $\{\tau_y, \tau_x\}$ denoting the break points and $b$ being the break magnitude. Breaks in level of the time series are

$$DU_t(TB_y) = \begin{cases} 1 & t > TB_y \\ 0 & otherwise \end{cases} \quad , \quad DU_t(TB_x) = \begin{cases} 1 & t > TB_x \\ 0 & otherwise \end{cases}$$

while breaks in slope are imposed using the following expressions:

$$DT_t(TB_y) = \begin{cases} t - TB_y & t > TB_y \\ 0 & otherwise \end{cases} \quad , \quad DT_t(TB_x) = \begin{cases} t - TB_x & t > TB_x \\ 0 & otherwise \end{cases}$$

Data generation process (DGP) is conducted using R and GAUSS 15.0 programs. Cointegrated $y_t$ and $x_t$ series are generated from either $AR(1) \sim WN(0, \sigma_e^2)$ or $AR(1) \sim GARCH(1,1)$ innovations with values of $\varphi = \theta_1 = 0.4999$ parameter for 100 sample size with 1000 replications.

The series are generated for single structural break in intercept, break in slope and their version of heteroscedastic innovations. The Cointegration relationships of two different models have been studied (Models C and C/S). Both breaks in intercept and breaks in slope have been considered separately. In this study, it is assumed that the break date is known or exogenous.

The nominal size of the tests is compared depending on part of the series where the break has occurred. The sections of the series are treated as a break in the first quarter (0.25T), in the second quarter (0.50T) and in the third quarter (0.75T) of the series. Since the magnitude of the structural break in the series may also affect the nominal size of the tests, the performance of tests is also evaluated with 2 and 10 unit break magnitude in slope and level, respectively. These values of break sizes coherent with those that led to very strong cases of spurious rejections by Dickey-Fuller tests in Leybourne et al.(2003). Maximum GARCH ($\theta_1 = \varphi = 0.4999$) parameters value are selected in order to determine the heteroscedastic variance of the error term. The significance level is 5%.

In short, the DGP design is summarized using the following parametric combinations:

$$\begin{aligned}
\text{Intercept and Slope before the break:}\quad & \mu_1 = 0, \beta_1 = 0 \\
\text{Intercept-change and Slope-change after the break:}\quad & \mu_2 = 10, \beta_2 = 2 \\
\text{Number of observation:}\quad & T = 100 \\
\text{Break function:}\quad & DU_t(.), DT_t(.) \\
\text{Break magnitude:}\quad & b = 2, 10 \\
\text{Break point:}\quad & TB_y = TB_x = (25, 50, 75) \\
\text{GARCH parameters:}\quad & \theta_0 = 0.0001, \theta_1 = \varphi = 0.4999
\end{aligned}$$

The results of the simulation are presented in Figure 5.1 to Figure 5.6.

### *5.2.2 The Test Results*

The simulation design is created based on the comparison of the empirical sizes of the time series under the conditions of structural breaks and heteroscedaticity. The residual based test performance dependent on the results of the unit root analysis of the residuals obtained from the model equations of the cointegration tests. Therefore, the power of the test is increased as the error terms obtained from the cointegration equations which subject to destructive effects such as trend, seasonality, structural break, and heteroscedaticity. The break points, break types and the test related structure of the error terms obtained as a result of the simulation design are shown in Figure 5.1 along with a couple of series. When Figure 5.1 is examined, it is observed that the break points of the $x_t$ and $y_t$ series for the level shift are the same, while the error terms of the EG(1987) test are affected by the break point; and the GH(1996) test and the WE-LM, which take a single break into consideration and the RALS(2)-LM GARCH, which takes multiple breaks into consideration, are not affected from this break. In addition, the RALS(2)-LM GARCH test generates residuals terms with a smaller variance. In cases in which the break is observed in different points, e.g. in the third quarter in $y_t$ series and the first quarter in $x_t$ series, the EG(1987) test has detected two break point to the error term, while the second break negatively affects the GH(1996) and WE-LM tests since they take only single break into consideration. As the RALS(2)-LM GARCH test considers multiple breaks, there is not any significant change in its performance.

When there is a break at the slope parameter of the time series, the cointegrated structure of the series are effected more when compared to the level shift. When examined in general, it can be argued that the test performances are in parallel with the level shift break. The results of the RALS(2)-LM GARCH test have a tendency to be closer to the stationary structure compared to other tests.

*5.2.2.1 No Break and No GARCH Effect*

The empirical sizes and efficiency gain of the test, when there are no structural breaks or GARCH effects in the series, are presented in Table 5.1 and Figure 5.2.

When the test performances are examined, some size distortions are observed in the GH(1996) test, while the RALS(2)-LM GARCH, WE-LM and EG(1987) tests are the closest ones to the nominal size value. The empirical size values of the WE-LM test show parallelism with the results in Tam (2012). It is also observed that the GH(1996) test experience spurious rejection issues arise from the break in the second quarter of $x_t$ and $y_t$ series.

Some irregular patterns may be observed in the series, which are generated by the GH(1996) test simulation process. Although the series which are shown in Figure 5.2 may be perceived as problematic, for example

$Y(0.5T)|X(0.5T)(emp.size_{Y(0.5T)|X(0.5T)} = 0.067)$,
$Y(0.5T)|X(0.75T)(emp.size_{Y(0.5T)|X(0.75T)} = 0.039)$, and
$Y(0.75T)|X(0.75T)(emp.size_{Y(0.75T)|X(0.75T)} = 0.036)$ cases,

the Zivot-Andrews unit root test results indicate that the series have no structural break. These extreme values were found to be natural results for the GH(1996) test simulation process.

*5.2.2.2 Level Break*

When the performances of the test for the C model with a break at the intercepts of the series are examined, it can be observed that the GH(1966) test diverges from the nominal size as mentioned in Cook (2004); while the EG(1987) test has more conservative results in different break scenarios as argued by Leybourne and Newbold (2003). On the other hand, the closest results to the nominal size values are presented by the WE-LM and RALS(2)-LM GARCH tests. The spurious reject issue of the GH(1996) test is observed in the break at the first quarter of the $X$ series and the second of the $Y$ series, at the most ((empirical value) p-value=0.243). Similarly, the breaks at the last quarter of the WE-LM test indicate that the test is liberal. However, although there is not such a problem in the RALS(2)-LM GARCH test, in cases where the break points are at the same location, there occurs the issue of having spurious regression in the test results.

*5.2.2.3 Slope Break*

When the performances of the tests for the C/S model with breaks in the slopes of the series are examined, it is apparent that the EG(1987) and GH(1996) tests have revealed an under-sizing problem. It appears that the EG(1987) test converges towards the nominal size value when the break has occurred at the same point in both series; however, this cannot refute the hypothesis in other break scenarios. On the other hand, the GH(1996) test diverges from the nominal value although it yields better results than the EG(1987) test. When the results of the WE-LM test are examined, it is clear that, contrary to the results of Tam (2012), it has conservative values when the break has occurred in the first quarter of the $Y_t$ series, and it yields values converging to nominal values in other break conditions. When the results of the RALS(2)-LM GARCH tests are examined, it is seen that results close to the nominal size are obtained. As the slope breaks in the $x_t$ and $y_t$ series being at different points create a multiple break effect in the residual terms, tests other than the RALS(2)-LM GARCH test become deviated from the nominal size values. With this result, it is understood that the location of the break in the series, in addition to having one break, is significant in the cointegration analyses.

*5.2.2.4 Level Break and GARCH Effect*

When the performances of the tests for the C model with GARCH effect and a structural break at the intercepts of the series are examined, it is apparent that the GH(1996) test has a spurious rejection issue and yields very high empirical size results in every break scenario. While the WE-LM test yields liberal results where the break points of the series are identical, it is observed that the results converge to nominal value. As for the RALS(2)-LM GARCH test, it yielded results closer to the nominal size values. However, it exhibits a liberal behavior when the break points are at the same locations. Unlike the other tests, the empirical values of the EG(1987) test are lower than the nominal size. It can be argued that the EG(1987) test yielded more consistent results when compared to the other tests, regardless of the low empirical size values. Despite the consistency of the EG(1987) test, the test with the best results under the level break and GARCH effect is the RALS(2)-LM GARCH test. On the other hand, it can be seen that GH(1996) and WE-LM tests have a spurious rejection issue.

*5.2.2.5 Slope Break and GARCH Effect*

When the performances of the tests for the C/S model with the GARCH effect and structural breaks at the slopes of the series are examined, it can be argued that EG(1987) and WE-LM tests generally experience an under-sizing problem. In addition, the WE-LM test has a spurious rejection issue when the break points are equal while results of the GH(1996) test converge towards the nominal size value at the first and last quarter of the $X$ series, and compared to other situations, serious distortions have occurred. The RALS(2)-LM GARCH test yields better results in general as deviations from the empirical size are observed in cases where the breaks occur at different locations. There are substantial deviations from the empirical size of the cointegration test except RALS(2)-LM GARCH, which are examined when breaks have occurred at the slopes of the series with GARCH effect. When examined in general, the structural break comprises of the "under-sizing" problem of the cointegration tests with the GARCH effect.

#### *5.2.2.6 Comparison of Efficiency Gain (Variance Ratio)*

When the variance ratios of the RALS(2)-LM GARCH and the EG(1987) test, which is a classical cointegration test, it can be argued that the RALS(2)-LM GARCH test have smaller variance by means of the moments It is also observed that the structural break and the GARCH effect influence this efficiency gain.

When there is not any break of GARCH effect, the variances of the $x_t$ and $y_t$ series are very close to each other since they are generated from a white noise process. Therefore, the variance ratio is expected to be close to 1. As a result of the study, this value is found above 97%. When the variance ratios for the level shift are compared, it is seen that the variance gain of the RALS(2)-LM GARCH test is between 13% and 35%. The finding, that the power values of the EG(1987) test is low when the break occurs at different locations, is supported with the empirical size results. The low variance of the RALS(2)-LM GARCH test increases the test statistics naturally. When the GARCH effect is added to the level shift the variance ratio of RALS(2)-LM GARCH test to the EG(1987) test decreases even more, and the efficiency gain increases. This result indicates that the GARCH effect increases the variance of the error terms of the cointegration tests. However, such negative effects can be eliminated with tests like RALS(2)-LM GARCH.

When the variance ratios for the breaks at the slope are compared, it is seen that the variance gain of the RALS(2)-LM GARCH test varies between 3% and 25%. Similar to the level shift results, while the variance is high for the series with breaks at the same location, the variance difference increases for the breaks at different locations. When the GARCH effect is added, the variance ratio of the classical RALS(2)-LM GARCH test is lower than the EG(1987) test.

This results shows that the proposed test is a more effective one, when compared to the present residual based tests.

## 5.3 Conclusion

This study aimed at extending the studies of Leybourne & Newbold (2003), Cook (2004) and Tam (2012), also considering the GARCH effect. The performances of Engle-Granger, Gregory-Hansen, Westerlund-Edgerton LM and RALS(2)-LM GARCH (finite sample properties) are examined under structural break and GARCH effect. When the results of the simulation are examined in general, it can be argued that when there is a structural break in the series, the RALS(2)-LM GARCH and WE-LM (2007) test yielded better results, the EG(1987) test has conservative results, and the GH(1996) tests have a spurious rejection issue. When the GARCH effect, in addition to structural break, is introduced to the series, size distortion or under-sizing issues are observed in the slightly more effective WE-LM (2007)test. However, RALS(2)-LM GARCH cointegration test resulted more powerfully. The Study shows that the newly suggested RALS-based cointegration tests utilizing higher moment conditions exhibit substantial power gains when the errors are GARCH distributed. It is also observed that the position of the break, the GARCH parameter value and the type of cointegration model in the experiment design influenced the performances of the tests. The different empirical values yielded by the tests may result in spurious cointegration values. In order to prevent such a situation, practitioners are advised to choose the tests allowing for the structural breaks in the series under the GARCH effect.

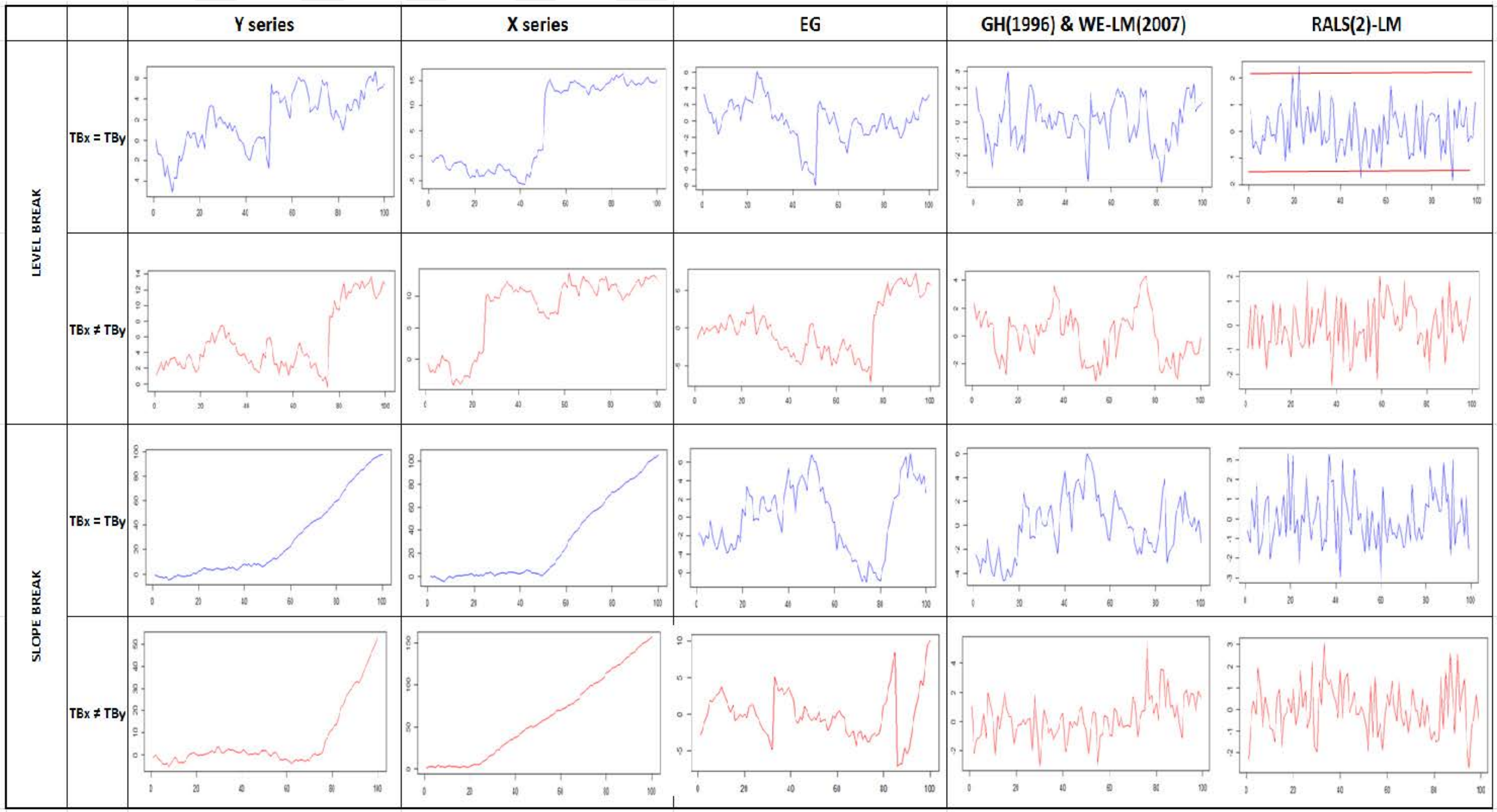


Figure 5.1 Example of simulation design for series and different residual terms

Table 5.1 The empirical size of no break and no GARCH effect

| | Engle-Granger | | | Gregory-Hansen | | | WE-LM | | | RALS(2)-LM GARCH | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Y\X | 0.25T-EG | 0.5T-EG | 0.75T-EG | 0.25T-GH | 0.5T-GH | 0.75T-GH | 0.25T-LM | 0.5T-LM | 0.75T-LM | 0.25T-RALS(2) | 0.5T-RALS(2) | 0.75T-RALS(2) |
| 0.25T | 0.048 | 0.060 | 0.043 | 0.055 | 0.052 | 0.047 | 0.050 | 0.051 | 0.050 | 0.054 | 0.06 | 0.043 |
| 0.5T | 0.047 | 0.058 | 0.042 | 0.044 | 0.067 | 0.039 | 0.051 | 0.050 | 0.051 | 0.058 | 0.05 | 0.045 |
| 0.75T | 0.048 | 0.060 | 0.056 | 0.048 | 0.046 | 0.036 | 0.049 | 0.048 | 0.050 | 0.052 | 0.05 | 0.045 |

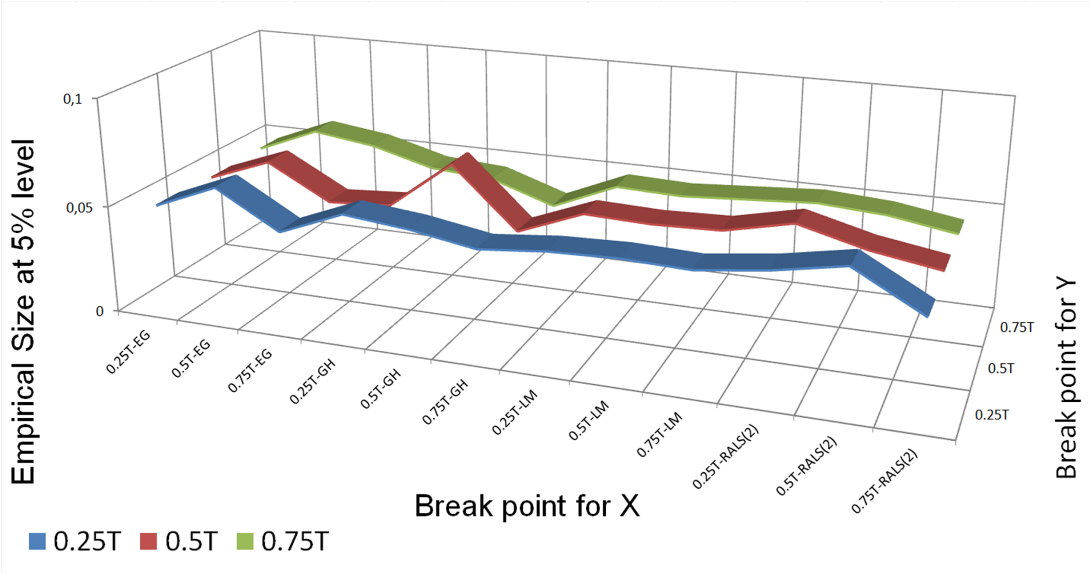


Figure 5.2 The empirical size of no break and no GARCH effect

Table 5.2 The empirical size of level break

| | Engle-Granger | | | Gregory-Hansen | | | WE-LM | | | RALS(2)-LM GARCH | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Y\X | 0.25T-EG | 0.5T-EG | 0.75T-EG | 0.25T-GH | 0.5T-GH | 0.75T-GH | 0.25T-LM | 0.5T-LM | 0.75T-LM | 0.25T-RALS(2) | 0.5T-RALS(2) | 0.75T-RALS(2) |
| 0.25T | 0.040 | 0.011 | 0.015 | 0.158 | 0.224 | 0.215 | 0.056 | 0.052 | 0.059 | 0.076 | 0.037 | 0.039 |
| 0.5T | 0.020 | 0.034 | 0.020 | 0.243 | 0.224 | 0.237 | 0.049 | 0.064 | 0.055 | 0.042 | 0.077 | 0.045 |
| 0.75T | 0.019 | 0.016 | 0.040 | 0.219 | 0.197 | 0.189 | 0.076 | 0.059 | 0.066 | 0.038 | 0.041 | 0.079 |

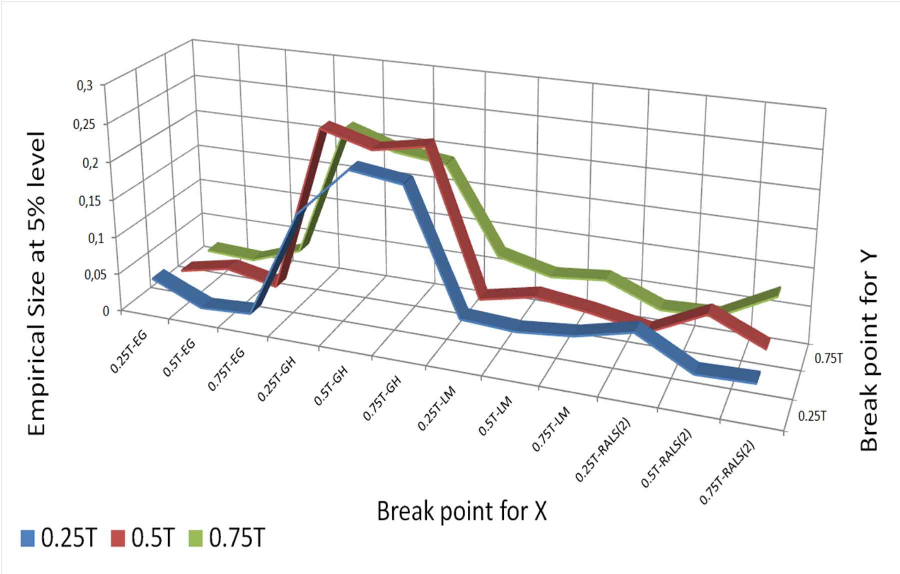


Figure 5.3 The empirical size of level break

Table 5.3 The empirical sizes of slope break

| | Engle-Granger | | | Gregory-Hansen | | | WE-LM | | | RALS(2)-LM GARCH | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Y\X | 0.25T-EG | 0.5T-EG | 0.75T-EG | 0.25T-GH | 0.5T-GH | 0.75T-GH | 0.25T-LM | 0.5T-LM | 0.75T-LM | 0.25T-RALS(2) | 0.5T-RALS(2) | 0.75T-RALS(2) |
| 0.25T | 0.064 | 0 | 0 | 0.018 | 0.028 | 0.002 | 0.005 | 0.008 | 0.002 | 0.082 | 0.021 | 0.015 |
| 0.5T | 0 | 0.071 | 0 | 0.007 | 0.025 | 0.002 | 0.028 | 0.006 | 0.029 | 0.034 | 0.076 | 0.038 |
| 0.75T | 0 | 0 | 0.032 | 0.012 | 0.011 | 0.028 | 0.058 | 0.045 | 0.030 | 0.041 | 0.039 | 0.073 |

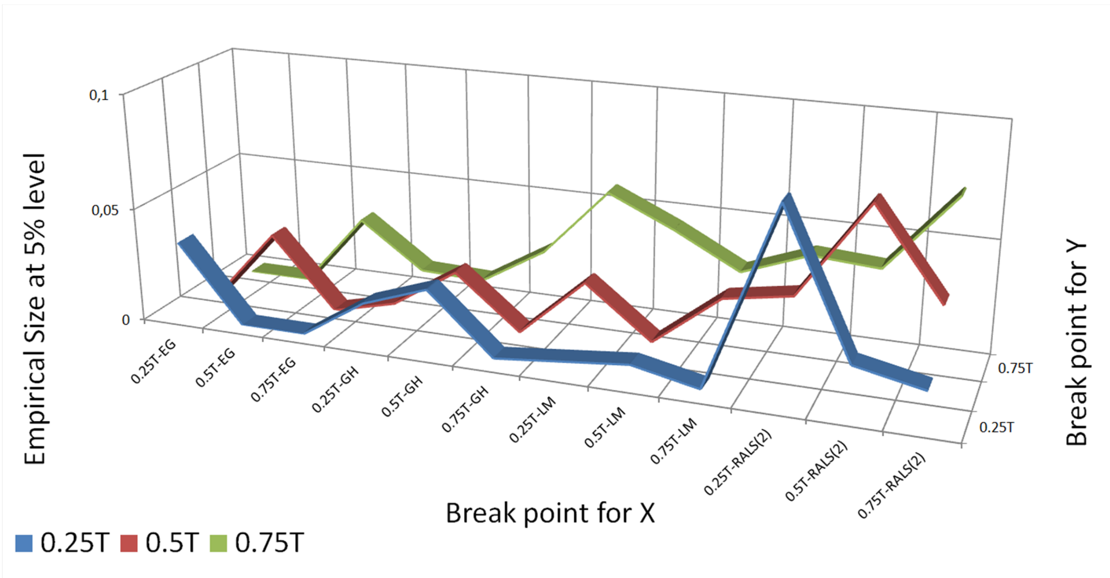


Figure 5.4 The empirical sizes of slope break

Table 5.4 Empirical size of level break and GARCH effect on C model

| | Engle-Granger | | | Gregory-Hansen | | | WE-LM | | | RALS(2)-LM GARCH | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Y\X | 0.25T-EG | 0.5T-EG | 0.75T-EG | 0.25T-GH | 0.5T-GH | 0.75T-GH | 0.25T-LM | 0.5T-LM | 0.75T-LM | 0.25T-RALS(2) | 0.5T-RALS(2) | 0.75T-RALS(2) |
| 0.25T | 0.036 | 0.016 | 0.029 | 0.346 | 0.502 | 0.544 | 0.103 | 0.058 | 0.062 | 0.098 | 0.036 | 0.033 |
| 0.5T | 0.013 | 0.040 | 0.019 | 0.531 | 0.307 | 0.526 | 0.065 | 0.099 | 0.064 | 0.047 | 0.106 | 0.041 |
| 0.75T | 0.025 | 0.014 | 0.038 | 0.515 | 0.548 | 0.375 | 0.067 | 0.078 | 0.101 | 0.037 | 0.043 | 0.085 |

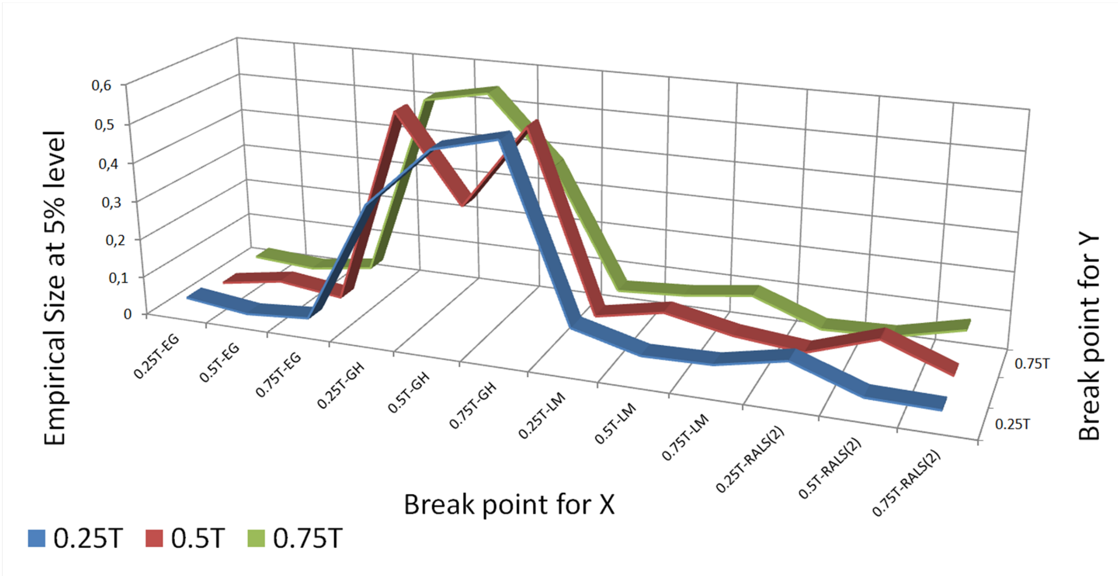


Figure 5.5 Empirical size of level break and GARCH effect on C model

Table 5.5 The empirical size of slope break and GARCH effect

| | Engle-Granger | | | Gregory-Hansen | | | WE-LM | | | RALS(2)-LM GARCH | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Y\X | 0.25T-EG | 0.5T-EG | 0.75T-EG | 0.25T-GH | 0.5T-GH | 0.75T-GH | 0.25T-LM | 0.5T-LM | 0.75T-LM | 0.25T-RALS(2) | 0.5T-RALS(2) | 0.75T-RALS(2) |
| 0.25T | 0.032 | 0 | 0 | 0.049 | 0.133 | 0.109 | 0.102 | 0.004 | 0 | 0.096 | 0.028 | 0.035 |
| 0.5T | 0 | 0.026 | 0 | 0.204 | 0.112 | 0.037 | 0.003 | 0.080 | 0 | 0.041 | 0.079 | 0.038 |
| 0.75T | 0.001 | 0 | 0.03 | 0.349 | 0.273 | 0.217 | 0.001 | 0.004 | 0.091 | 0.029 | 0.032 | 0.082 |

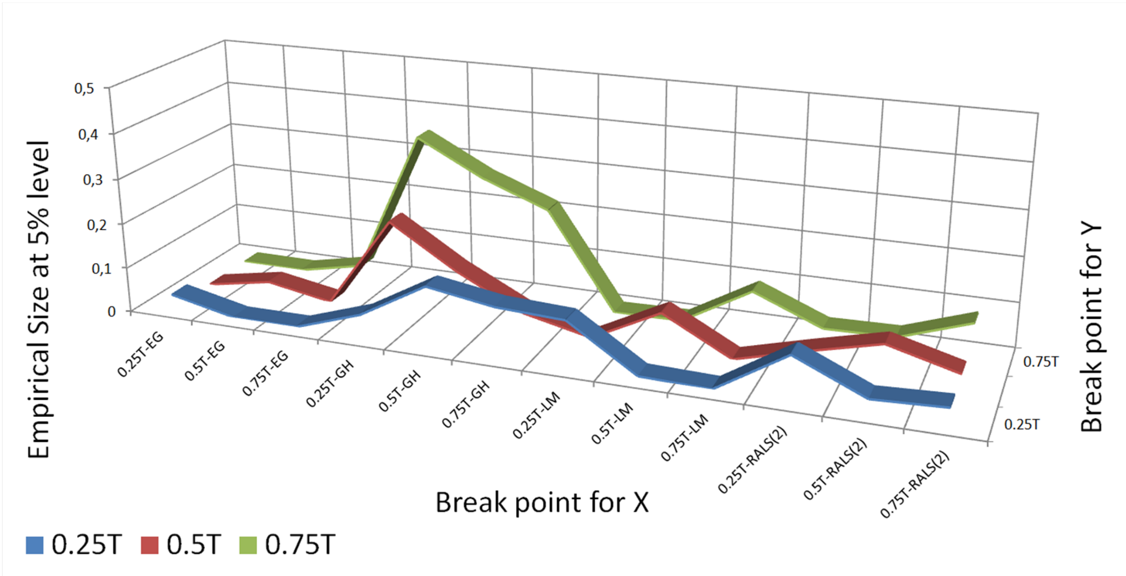


Figure 5.6 The empirical size of slope break and GARCH effect

Table 5.6 Efficiency Gain (Variance Ratio) of no break and no GARCH effect

| Y\X | 0.25T | 0.5T | 0.75T |
|---|---|---|---|
| 0.25T | 0.972 | 0.971 | 0.969 |
| 0.5T | 0.973 | 0.973 | 0.971 |
| 0.75T | 0.970 | 0.972 | 0.975 |

Table 5.7 Efficiency Gain (Variance Ratio) of level break

| Y\X | 0.25T | 0.5T | 0.75T |
|---|---|---|---|
| 0.25T | 0.848 | 0.727 | 0.656 |
| 0.5T | 0.684 | 0.867 | 0.685 |
| 0.75T | 0.653 | 0.693 | 0.843 |

Table 5.8 Efficiency Gain (Variance Ratio) of slope break

| Y\X | 0.25T | 0.5T | 0.75T |
|---|---|---|---|
| 0.25T | 0.973 | 0.785 | 0.756 |
| 0.5T | 0.764 | 0.971 | 0.788 |
| 0.75T | 0.743 | 0.804 | 0.973 |

Table 5.9 Efficiency Gain (Variance Ratio) of level break and GARCH effect

| Y\X | 0.25T | 0.5T | 0.75T |
|---|---|---|---|
| 0.25T | 0.839 | 0.711 | 0.659 |
| 0.5T | 0.687 | 0.857 | 0.685 |
| 0.75T | 0.659 | 0.685 | 0.844 |

Table 5.10 Efficiency Gain (Variance Ratio) of slope break and GARCH effect

| Y\X | 0.25T | 0.5T | 0.75T |
|---|---|---|---|
| 0.25T | 0.957 | 0.728 | 0.733 |
| 0.5T | 0.749 | 0.924 | 0.761 |
| 0.75T | 0.716 | 0.739 | 0.895 |

# CHAPTER SIX
# CONCLUSION

Cointegration introduced by Engle-Granger (1987) therefore plays a major role in time series analysis. This concept, the hypothesis that one stationary linear combination of individually non stationary variables exists, has been commonly used for empirical purposes in many areas. The cointegration framework has been evolved rapidly over the last three decades. A number of tests are available in the literature. Most of these tests are residual based and they are widely used due to their simplicity.

Many studies have demonstrated that structural breaks in a cointegration relationship significantly impact the performance of cointegration tests. Besides structural breaks, high frequency time series data is also characterized by GARCH model. The existence of the structural breaks and heteroscedastic error term may cause various problems such as biases and inconsistent estimation results, biased parameter estimation and poor predictions. In these circumstances, estimator of residual based tests is inefficient and decrease the power of the test used.

This dissertation proposes a more powerful test among residual based test, the new test's name is the "RALS(2)-LM GARCH", which is suitable for GARCH effects under the level or/and slope structural breaks. The test formed using the residual based LM method has many approaches, which are original in many ways. First, the test offers an alternative and a more efficient solution to the tests with decreasing performances due to structural breaks and heteroscedasticity. Second, the test relieves the users from conducting extra analyses by considering the structural breaks and heteroscedasticity together. Third, tests lies in the invariance feature that the distribution does not depend on the different degree of GARCH parameter in the presence of multiple level-breaks.

In simulation study, the performances of Engle-Granger (1987), Gregory-Hansen (1996), Westerlund-Edgerton LM (2007) and RALS(2)-LM GARCH are examined under structural break and GARCH effect. When the results of the simulation are examined in general, it can be argued that when there is a structural break in the series, the RALS(2)-LM GARCH and WE-LM test yielded better results. Furthermore, the simulation shows that the newly suggested RALS-based cointegration tests utilizing higher moment conditions exhibit substantial power gains when the errors are GARCH distributed. It is also observed that the position of the break, the GARCH parameter value and the type of cointegration model in the experiment design influenced the performances of the tests.

## REFERENCES


Amsler, C., & Lee, J. (1995). An LM test for a unit-root in the presence of a structural change. *Econometric Theory*, *11*, 359-368.

Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, *66*(1), 47-78. doi:10.2307/2998540

Banerjee, A., & Carrion-i-Silvestre, J. L. (2006). Cointegration in panel data with breaks and cross-section dependence. Working Paper; *European University Institute*.

Ben-David, D., Lumsdaine, R., & Papell, D. H. (2003). Unit root, postwar slowdowns and long-run growth: Evidence from two structural breaks. *Empirical Economics*, *28*(2), 303-319.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*, 307—327.

Brooks, C., & Rew, A. (2002). Testing for a unit root in a process exhibiting a structural break in the presence of GARCH errors. *Computational Economics*, *20*(3), 157-176.

Campos, J., Ericsson, N. R., & Hendry, D. F. (1996) Cointegration tests in the presence of structural breaks. *Journal of Econometrics*, *70*, 187-220.

Carrion-i-Silvestre, J. L., & Sanso. (2006). A Testing the null of cointegration with structural breaks. *Oxford Bulletin of Economics and Statistics*, *68*, 623-646.

Committee, Nobel Prize. (2003). Time-series econometrics: Cointegration and autoregressive conditional heteroskedasticity. *Nobel Prize in Economics documents*.

Cook, S. (2004). Spurious rejections by cointegration tests incorporating structural change in the cointegrating relationship. *Applied Economics Letters*, *11*, 879-884.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for the autoregressive time series with a unit root. *Journal of the American Statistical Association*, *74*, 427-431.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom Inflation. *Econometrica*, *50*, 987—1007.

Engle, R. F., & Granger, C. W. J. (1987). Cointegration and error correction: representation, estimation and testing. *Econometrica*, *55*, 251–276.

Gallagher, L. A., & Taylor, M. P. (2000). Measuring the temporary component of stock prices: Robust multivariate analysis. *Economics Letters*, *67*, 193-200.

Gregory, A. W., & Hansen, B. E. (1996). Residual-based tests for cointegration in models with regime shifts. *Journal of Econometrics*, *70*, 99-126.

Hansen, B. E. (1995). Rethinking the univariate approach to the unit root testing: using covariates to increase the power. *Econometric Theory*, *11*, 1148-1171.

Hatemi-J, A. (2008). Tests for cointegration with two unknown regime shifts with an application to financial market integration. *Empirical Economics*, 35, 497-505. doi:10.1007/s00181-007-0175-9

Höglund, R., & Östermark, R. (2003). Size and power of some cointegration tests under structural breaks and heteroskedastic noise. *Statistical Papers*, *44*, 1–22.

Im, K. S., Lee, J., & Tieslau, M. (2005). Panel LM unit-root tests with level shifts. *Oxford Bulletin of Economics and Statistics*, *67*, 393–419.

Im, K. S., & Schmidt, P. (2008). More efficient estimation under non-normality when higher moments do not depend on the regressors, using residual-augmented least squares. *Journal of Econometrics*, *144*, 219-233.

Im, K. S., Lee, J., & Tieslau, M. (2014). More powerful unit root tests with nonnormal errors. *The Festschrift in Honor of Peter Schmidt*, edited by R. Sickles and W. Horrace, 315–342, Berlin, Germany: Springer Publishing Co.

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

Kosapattarapim, C., Lin, Y. X., & McCrae, M. (2013). Size and power of cointegration tests with non-normal GARCH error distribution. Working Paper; *Centre for Statistical and Survey Methodology, School of Mathematics and Applied Statistics, University of Wollongong, Australia*.

Krauss, C., & Herrmann, K., (2017). On the power and size properties of cointegration tests in the light of high-frequency stylized facts. *Journal of Risk and Financial Management*, *10*(7), 1-24. doi:10.3390/jrfm10010007

Kurita, T. (2013). Exploring the impact of multivariate GARCH innovations on hypothesis testing for cointegrating vectors. *Communications in Statistics- Simulation and Computation*, *42*(8), 1785-1800. doi:10.1080/03610918.2012.677920

Lee, J., & Strazicich M. C. (2001). Break point estimation and spurious rejections with endogenous unit root tests. *Oxford Bulletin of Economics and Statistics*, *63*, 535-558.

Lee, J., & Strazicich, M. C. (2003). Minimum lagrange multiplier unit root test with two structural breaks. *Review of Economics and Statistics*, *85*(4), 1082-1089.

Lee, J., Meng, M., & Strazicich, M. C. (2012). Two-step LM unit root tests with trend-breaks. *Journal of Statistical and Econometric Methods*, *1*(2), 81-107.

Leybourne, S. J., & Newbold, P. (2003). Spurious rejections by cointegration tests induced by structural breaks. *Applied Economics*, *35*, 1117-1121.

Li, J., & Lee, J. (2015). Improved autoregressive forecasts in the presence of non-normal errors. *Journal of Statistical Computation and Simulation*, *85*(14), 2936-2952. doi:10.1080/00949655.2014.945930

Maki, D. (2012). Tests for cointegration allowing for an unknown number of breaks. *Economic Modelling*, *29*(5), 2011-2015.

Masuda, J. (2008). Residual-based tests for cointegration in models with multi-breaks. *Applied Economics Letters*, *15*(13), 1001-1006. doi:10.1080/13504850600972378

Meng, M., Im, K., Lee, J., & Tieslau, M. (2014). More powerful LM unit root tests with non-normal errors. *The Festschrift in Honor of Peter Schmidt*, edited by R. Sickles and W. Horrace, 343–357, Berlin, Germany: Springer Publishing Co.

Meng, M., Lee, J., & Payne, J. E. (2016). RALS-LM unit root test with trend breaks and non-normal errors: application to the Prebisch-Singer hypothesis, *Studies in Nonlinear Dynamics and Econometrics*, *21*(1). doi:10.1515/snde-2016-0050

Meng, M., Strazicich, M. C., & Lee, J. (2017). Hysteresis in unemployment? Evidence from linear and nonlinear unit root tests and tests with non-normal errors. *Empirical Economics*, 1-16.

Mercedes, A., Ferruz, L., & Gallagher, L. A. (2013). Performance of Spanish pension funds: robust evidence from alternative models. *Applied Financial Economics*, *23*(4), 297-314. doi:10.1080/09603107.2012.720011

Mishra, V., & Smyth, R. (2017). Conditional convergence in Australia's energy consumption at the sector level. *Energy Econ.*, *62*, 394-401.

Narayan, P. K., & Liu, R. (2013). New evidence on the weak-form efficient market hypothesis. Working Paper; *Centre for Financial Econometrics, Deakin University*.

Oh, D., & Lee, H. (2016). LM cointegration tests allowing for an unknown number of breaks: implications for the forward rate unbiasedness hypothesis. *Applied Economics*, *49*, 1194-1203.

Ozcan, B., & Erdogan, S. (2015). Are Turkey's tourism markets converging? Evidence from the two-step LM and three-step RALS-LM unit root tests. *Current Issues in Tourism*. doi:10.1080/13683500.2015.1040741

Param Sothy, S., & Podivinsky, J. M. (2000). The effect of non-normal disturbances and conditional heteroskedasticity on multiple cointegration tests. *Journal of Statistical Computation and Simulation*, *65*(1-4), 173-189. doi:10.1080/00949650008811997

Payne, J. E., Lee, J., & Vizek, M. (2017). Is there convergence in per capita renewable energy consumption across U.S. States? Evidence from LM and RALS-LM unit root tests with breaks. *Renewable and Sustainable Energy Reviews 70*, 715–728.

Perron, P. (1989). The Great crash, the oil price shock and the unit root hypothesis. *Econometrica*, *57*, 1361-1401.

Phillips, P. C. B., & Ouliaris, S. (1990). Asymptotic properties of residual based tests for cointegration. *Econometrica*, *58*, 165–193. doi:10.2307/2938339.

Pierdzioch, C. (2010). Periodically collapsing bubbles in the German stock market, 1876–1913. *Applied Economics Letters*, *17*(9), 907-908. doi:10.1080/17446540802552324

Pierdzioch, C., Risse, M., & Rohloff, S. (2015). Cointegration of the prices of gold and silver: RALS-based evidence. *Finance Research Letters*, *15*, 133-137.

Rahbek, A., Hansen, E., & Dennis, G. J. (2002). ARCH innovation and their impact on cointegration rank testing. Working Paper; *Department of Statistics and Operation Research, University of Copenhagen*.

Saez, M., & Kunst, R. M. (1995) ARCH patterns in cointegrated systems. Working Paper; *Economics Working Paper 110; Department of Economics and Business, Universitat Pompeu Fabra*.

Schmidt, P. & Phillips, P. C. B. (1992). LM tests for a unit root in the presence of deterministic trends. *Oxford Bulletin of Economics and Statistics*, *54*, 257–87.

Tam, P. S. (2012). Size properties of lagrange multiplier cointegration tests in the presence of structural breaks. *Applied Economics Letters*, *19*, 1061-1064.

Taylor, M. P. & Peel, D. A. (1998). Periodically collapsing stock price bubbles: a robust test. *Economics Letters*, *61*, 221–8.

Westerlund, J., & Edgerton, D. L. (2007). New improved tests for cointegration with structural breaks. *Journal of Time Series Analysis*, *28*(2).

Westerlund, J., & Narayan, P. K. (2013). Testing the efficient market hypothesis in conditionally heteroskedastic futures markets. *Journal of Futures Markets*, *33*, 1024-1045. doi: 10.1002/fut.21624

Zivot, E., & Andrews, D. W. K. (1992). Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of Business and Economic Statistics*, *10*, 251-270.

**APPENDICES**

**A.1: Appendix 1.**

***Result 1.*** Derivation for restricted MLE of $\tilde{\alpha}$

$$\begin{aligned}
\frac{\partial lnL}{\partial \alpha} &= -\frac{1}{2\sigma^2}\left(\frac{\partial}{\partial \alpha}(S_1)^2 + \sum_{t=2}^{T}\frac{\partial}{\partial \alpha}(\Delta S_t)^2\right) \\
&= -\frac{1}{2\sigma^2}\left(\frac{\partial}{\partial \alpha}(y_1 - \tilde{\alpha} - \tau - x_1'\beta)^2 + \sum_{t=2}^{T}\frac{\partial}{\partial \alpha}(\Delta y_t - \tau - \Delta x_t'\beta)^2\right)_{\tilde{\alpha}=\alpha+z_0} \\
&= -\frac{1}{\sigma^2}\Big((y_1 - \tilde{\alpha} - \tau - x_1'\beta)(-1)\Big) = S_1
\end{aligned}$$

So

$$\hat{\alpha}_{restricted} = S_1 \tag{A.1}$$

***Result 2.*** The restricted MLE of $\tau$

$$\begin{aligned}
\frac{\partial lnL}{\partial \tau} &= -\frac{1}{2\sigma^2}\left(\frac{\partial}{\partial \tau}(y_1 - \tilde{\alpha} - \tau - x_1'\beta)^2 + \sum_{t=2}^{T}\frac{\partial}{\partial \tau}(\Delta y_t - \tau - \Delta x_t'\beta)^2\right)_{\tilde{\alpha}=\alpha+z_0} \\
&= -\frac{1}{\sigma^2}\left((y_1 - \tilde{\alpha} - \tau - x_1'\beta)(-1) + \sum_{t=2}^{T}(\Delta y_t - \tau - \Delta x_t'\beta)(-1)\right) \\
&= S_1 + \sum_{t=2}^{T}\Delta S_t \\
&= S_1 + (S_2 - S_1) + (S_3 - S_2) + (S_4 - S_3) + \dots + (S_T - S_{T-1}) = S_T
\end{aligned}$$

So

$$\tau = S_T. \tag{A.2}$$

***Result 3.*** Derivation of restricted MLE of $\beta$

$$\begin{aligned}\frac{\partial lnL}{\partial\beta} &= -\frac{1}{2\sigma^2}\left(\frac{\partial}{\partial\beta}(y_1-\tilde{\alpha}-\tau-x_1'\beta)^2+\sum_{t=2}^{T}\frac{\partial}{\partial\beta}(\Delta y_t-\tau-\Delta x_t'\beta)^2\right)_{\tilde{\alpha}=\alpha+z_0}\\ &= -\frac{1}{\sigma^2}\left((y_1-\tilde{\alpha}-\tau-x_1'\beta)(-x_1')+\sum_{t=2}^{T}(\Delta y_t-\tau-\Delta x_t'\beta)(-\Delta x_t')\right)\\ &= S_1x_1'+\sum_{t=2}^{T}\Delta S_t\Delta x_t'\end{aligned}$$

So

$$\hat{\beta}_{restricted} = S_1x_1'+\sum_{t=2}^{T}\Delta S_t\Delta x_t' \tag{A.3}$$

***Result 4.*** $z_{t-1}=S_{t-1}+z_0$

*Proof Result 4.*

$$S_{t-1}=y_{t-1}-\alpha-\tau(t-1)-x_t'\beta-z_0$$

From equation 1, $z_t = y_t - \alpha - \tau t - x_t'\beta$ and also $z_{t-1}=y_{t-1}-\alpha-\tau(t-1)-x_{t-1}'\beta$.

Therefore,

$$S_{t-1}=z_{t-1}-z_0 \text{ and } z_{t-1}=S_{t-1}+z_0 \tag{A.4}$$

***Result 5.*** $\hat{S}_1=0$ and $\hat{S}_T=0$

*Proof Result 5.*

$$\hat{S}_1=y_1-\hat{\alpha}-\hat{\tau}-x_1'\hat{\beta}$$

whereby Appendix 1, *Result 1*, $\hat{S}_1$ can be rewritten as

$$\hat{S}_1=y_1-(y_1-\hat{\tau}-x_1'\hat{\beta})-\hat{\tau}-x_1'\hat{\beta}=0.$$

So

$$\hat{S}_T=y_T-\hat{\alpha}-\hat{\tau}T-x_T'\hat{\beta}=y_T-(y_1-\hat{\tau}-x_1'\hat{\beta})-\hat{\tau}T-x_T'\hat{\beta}$$

$$\hat{S}_T=y_T-y_1-\hat{\tau}(T-1)-(x_T-x_1)'\hat{\beta} \tag{A.5}$$

Using reduced for of SSE,

$$\left( (y_1 - \tilde{\alpha} - \tau - x_1'\beta) + \sum_{t=2}^{T} (\Delta y_t - \tau - \Delta x_t'\beta) \right)_{\tilde{\alpha}=\alpha+z_0}$$

$$(y_1 - \tilde{\alpha} - \tau - x_1'\beta) + \sum_{t=2}^{T} \Delta y_t - \tau(T-1) - \sum_{t=2}^{T} \Delta x_t'\beta$$

$$y_1 - \tilde{\alpha} - x_1'\beta + \sum_{t=2}^{T} \Delta y_t - \sum_{t=2}^{T} \Delta x_t'\beta = \tau + \tau(T-1)$$

Since first three terms are negligible in the above equation, the equation becomes to

$$\sum_{t=2}^{T} \Delta y_t - \sum_{t=2}^{T} \Delta x_t'\beta = \tau T$$

When the both side is divided by $T$, $\Delta y$ and $\Delta x$ denote the sample averages of $\Delta y_t$ and $\Delta x_t$ respectively, then this expression becomes

$$\Delta y - \Delta x\hat{\beta} = \hat{\tau} \tag{A.6}$$

The equation $\hat{S}_T = y_T - y_1 - \hat{\tau}(T-1) - (x_T - x_1)'\hat{\beta}$ can be also written as adapted version with $\Delta y$ and $\Delta x$.

Since

$$\sum_{t=2}^{T} \Delta y_t = (y_2 - y_1) + (y_3 - y_2) + (y_4 - y_3) + ... + (y_T - y_{T-1}) = y_T - y_1. \tag{A.7}$$

The equation becomes $\hat{S}_T = \sum_{t=2}^{T} \Delta y_t - \hat{\tau}(T-1) - \sum_{t=2}^{T} \Delta x_t'\hat{\beta}$

$$\hat{S}_T = \sum_{t=2}^{T} \Delta y_t - (\Delta y - \Delta x'\hat{\beta})(T-1) - \sum_{t=2}^{T} \Delta x_t'\hat{\beta}$$

$$\hat{S}_T = (T-1)\Delta y - (T-1)\Delta y - (T-1)\Delta x\hat{\beta} - (T-1)\Delta x\hat{\beta} = 0 \tag{A.8}$$

***Result 6.*** Obtaining Hessian matrix

$$\left(\frac{\partial^2 logL}{\partial\gamma\partial\gamma'}\right) = -\frac{1}{2\hat{\sigma}^2}\sum_{t=1}^{T}\frac{\partial^2 e_t^2}{\partial\gamma\partial\gamma'}$$

Direct calculations yield the following second-order partial derivatives when evaluated at the restricted maximum likelihood estimators

$$\frac{1}{2}\sum_{t=1}^{T}\frac{\partial^2 e_t^2}{\partial\rho\partial\tilde{\alpha}} = z_0 \tag{A.9}$$

$$\begin{aligned}\frac{1}{2}\sum_{t=1}^{T}\frac{\partial^2 e_t^2}{\partial\rho\partial\tau} = \frac{1}{\sigma^2}\Bigg[& z_0(y_1-\alpha-\tau-x_1'-z_0)\\ &+\sum_{t=2}^{T}(\Delta y_t-\tau-\Delta x_t')(y_{t-1}-\alpha-\tau(t-1)-x_{t-1}'\beta-z_{t-2})\Bigg]\end{aligned}$$

$$\begin{aligned}&=\frac{1}{\sigma^2}\left(-z_0+\sum_{t=2}^{T}-(t-1)\Delta\hat{S}_t-\hat{S}_{t-1}\right)\\ &=\frac{1}{\sigma^2}\left(z_0+\sum_{t=2}^{T}\left(\Delta\hat{S}_t(t-1)+\hat{S}_{t-1}\right)\right)\\ &=\frac{1}{\sigma^2}\left(z_0+(\hat{S}_T-\hat{S}_1)(t-1)+\sum_{t=2}^{T}\hat{S}_{t-1}\right)\\ &=\frac{1}{\sigma^2}\left(z_0+\sum_{t=2}^{T}\hat{S}_{t-1}\right)\end{aligned} \tag{A.10}$$

where the last equality holds because, $\hat{S}_T = \hat{S}_1 = 0$ which follows by writing

$$\begin{aligned}\hat{S}_1 &= y_1-\hat{\alpha}-\hat{\tau}-x_1'\hat{\beta}\\ &= y_1-(y_1-\hat{\tau}-x_1'\hat{\beta})-\hat{\tau}-x_1'\hat{\beta}\end{aligned}$$

Moreover, we have

$$
\begin{aligned}
\frac{1}{2}\sum_{t=1}^{T}\frac{\partial^2 e_t^2}{\partial\rho\partial\beta} &= -\frac{1}{\sigma^2}\left((-x_1z_0)+\sum_{t=2}^{T}(-x_{t-1})\Delta S_t+(\Delta x_t')S_{t-1}\right)\\
&= \frac{1}{\sigma^2}\left(x_1z_0+\sum_{t=2}^{T}x_{t-1}'\Delta S_t+\sum_{t=2}^{T}\Delta x_t' S_{t-1}\right)\\
&= \frac{1}{\sigma^2}\left(x_1z_0+\sum_{t=2}^{T}\Delta x_t' z_{t-1}+\sum_{t=2}^{T}\Delta x_t' z_0+\sum_{t=2}^{T}x_{t-1}'\Delta S_{t-1}\right)\\
&= x_1z_0+(x_2z_0-x_1z_0)+(x_3z_0-x_2z_0)+...+(x_Tz_0-x_{T-1}z_0)\\
&= x_Tz_0+\sum_{t=2}^{T}x_{t-1}'\Delta S_t+\sum_{t=2}^{T}\Delta x_t' S_{t-1}
\end{aligned}
\tag{A.11}
$$

***Result 7.*** $\sum_{t=2}^{T}\hat{S}_{t-1}\Delta\hat{S}_t = -\frac{1}{2}\Big(\sum_{t=2}^{T}\Delta\hat{S}_t^2\Big)$

*Proof Result 7.*

If two terms equal to each other, $2\sum_{t=2}^{T}\hat{S}_{t-1}\Delta\hat{S}_t+\sum_{t=2}^{T}\Delta\hat{S}_t^2$ must be equal to zero.

$$
\begin{aligned}
&= \sum_{t=2}^{T}\Big(\hat{S}_t-\hat{S}_{t-1}\Big)^2+2\sum_{t=2}^{T}\hat{S}_{t-1}\Big(\hat{S}_t-\hat{S}_{t-1}\Big)\\
&= \sum_{t=2}^{T}\hat{S}_t^2-2\sum_{t=2}^{T}\hat{S}_t\hat{S}_{t-1}+\sum_{t=2}^{T}\hat{S}_{t-1}^2+2\sum_{t=2}^{T}\hat{S}_{t-1}\hat{S}_t-2\sum_{t=2}^{T}\hat{S}_{t-1}^2\\
&= \sum_{t=2}^{T}\hat{S}_t^2-\sum_{t=2}^{T}\hat{S}_{t-1}^2=\hat{S}_T^2-\hat{S}_1^2=0
\end{aligned}
\tag{A.12}
$$

## A.2: Appendix 2.

### *A.2.1 Brownian Motion*

A standard brownian motion or standard wiener process are example of continuous timed process in the range of $(W(.))$, $[0, 1]$.

For each $t \in [0, 1]$, $W(t)$ has three properties below

i) Continuous path

ii) Stationary, independent increments,

iii) $W(t) \approx N(0, t)$ for all $t \geq 0$

### *A.2.2 Functional Central Limit Theorem*

According to central limit theorem, $u_t$ as being independent random variables having the same distribution with a zero mean and $\sigma^2$ variance, the asymptotic distribution of $\bar{u}_t$ sample mean will have a normal distribution.

$$\bar{u}_t = T^{-1} \sum_{t=1}^{T} a_t \qquad \sqrt{T}\bar{u}_t \xrightarrow{d} N(0, \sigma^2)$$

Let the step function $X_T(r)$

$$X_T(r) := T^{-1} \sum_{t=1}^{[T_r]} a_t \qquad r \in [0, 1]$$

$$X_T(r) = \begin{cases} 0 & 0 \leq r < \frac{1}{T} \\ \\ \frac{a_1}{T} & \frac{1}{T} \leq r < \frac{2}{T} \\ \\ \frac{(a_1+a_2+a_3+...+a_t)}{T} & r = 1 \end{cases}$$

Later, for all values of $r$, as with respect to central limit theorem, it is obvious that it would be;

$$\sqrt{T}X_T(r) = \frac{1}{\sqrt{T}}\sum_{t=1}^{[T_r]} a_t = \frac{\sqrt{[T_r]}}{\sqrt{T}}\frac{1}{\sqrt{[T_r]}}\sum_{t=1}^{[T_r]} a_t \qquad \frac{1}{\sqrt{[T_r]}}\sum_{t=1}^{[T_r]} a_t \xrightarrow{L} N(0,\sigma^2)$$

since $\frac{\sqrt{[T_r]}}{\sqrt{T}} = \sqrt{r}$

$$\sqrt{T}X_T(r) \xrightarrow{d} N(0, r\sigma^2)$$

$$\sqrt{T}\Big(\frac{X_T(r)}{\sigma}\Big) \xrightarrow{d} N(0, r)$$

Similarly, let $r_2 > r_1$, the asymptotic distribution, mean and variance of the difference of the two step function for $[T_{r_2}]$ and $[T_{r_1}]$ sample size would fit the normal distribution as

$$\sqrt{T}\Big(\frac{X_T(r_2) - X_T(r_1)}{\sigma}\Big) \xrightarrow{L} N(0, r_2 - r_1)$$

If $r = 0$ $X_T(r) = 0$ for each $r_2 > r_1$ value, $\sqrt{T}\Big(\frac{X_T(.)}{\sigma}\Big) := W(.)$ defined by Brownian motion.

The Brownian motion provides the possibility of applying the central limit theorem in a more general way. Using the functional central limit theorem, traditional central limit theorem for $r = 1$ can easily be reached. In other words;

$$X_T(1) = T^{-1} \sum_{t=1}^{[T_r]} a_t \quad r = 1$$

$$\sqrt{T}\Big(\frac{X_T(1)}{\sigma}\Big) = \left(\frac{T^{-1}\sum_{t=1}^{T} a_t}{\sigma}\right)\sqrt{T} = \frac{1}{\sqrt{T}\sigma}\sum_{t=1}^{T} a_t \xrightarrow{d} W(1) \approx N(0,1)$$

### *A.2.3 Continuous Mapping Theorem*

Let $\{X_T\}_{t=1}^{\infty}$ random variables array be a $g : R \to R$ defined continuous function and let $X$ be any random variables. According to continuous mapping theorem, while $T \to \infty$, if $X_T \xrightarrow{d} X$ approximates $g(X_T) \xrightarrow{d} g(X)$ will approximate. Accordingly, $\sqrt{T}(X_T(.)) = \sigma W(.)$, $W(r) \sim N(0,r)$ and $\sqrt{T}(X_T(r)) \xrightarrow{d} \sigma W(r)$. Also if $S_T(.)$ stochastic function of $X$ random variable $\Big(\int_0^1 S(r)dr\Big)$. It is obvious that

$$S_T(r) = \Big(\sqrt{T}X_T(r)\Big)^2 \qquad S_T(r) \xrightarrow{d} \sigma^2[W(r)]^2.$$

**Lemma 1 (Sung K. Ahn (1993))**

Let the $e_t$ be independent random variables with $E(e_t) = 0$, $Var(e_t) = \sigma^2$ and $sup_t E\Big(|e_t^{2+\delta}|\Big) < \infty$ for some $\delta > 0$. We define

$$S_t = \sum_{j=1}^{t} u_j, \qquad u_t = \sum_{j=0}^{\infty} \psi_j e_{t-j},$$

where the $\psi_0 = 1$ and they decay exponentially as $j$ increases.

$$u_t = \psi_0 e_t + \psi_1 e_{t-1} + \psi_2 e_{t-2} + ... + \psi_j e_{t-j} \qquad \Psi = \sum_{j=0}^{\infty} \psi_j$$

$$
\begin{aligned}
\sum_{j=1}^{[T_r]} u_j &= \sum_{t=1}^{[T_r]} \sum_{j=0}^{\infty} \psi_j e_{t-j} = \sum_{t=1}^{[T_r]} (\psi_0 e_t + \psi_1 e_{t-1} + ... + \psi_j e_{t-j}) \\
&= \sum_{t=1}^{[T_r]} \psi_0 e_1 + \sum_{t=1}^{[T_r]} \psi_1 e_{t-1} + \sum_{t=1}^{[T_r]} \psi_2 e_{t-1} + ... \\
&= \sum_{t=1}^{[T_r]} e_1 + \psi_1 \sum_{t=1}^{[T_r]} e_1 + \psi_2 \sum_{t=1}^{[T_r]} e_1 + ... \\
&= \Psi \sum_{t=1}^{[T_r]} e_T + (\psi_1 + \psi_2 + ...) + o_p(1)
\end{aligned}
$$

$$
T^{-1/2} S_{[T_r]} - \frac{[T_r]}{T} \frac{1}{\sqrt{T}} S_T
$$

$$
\begin{aligned}
T T^{-1/2} S_{[T_r]} - \frac{[T_r]}{T} T \frac{1}{\sqrt{T}} S_T &= \sqrt{T} S_{[T_r]} - \frac{[T_r]}{T} \sqrt{T} S_T \\
&= \sigma W(r) - r\sigma W(1) \\
&= \sigma \Big( W(r) - rW(1) \Big) = \sigma V(r)
\end{aligned}
$$

as $T \rightarrow \infty$ approximate to Brownian Bridge

### *A.2.4 Lindeberg Levy Central Limit Theorem*

Probably the most famous CLT is due to Lindeberg and Levy. Consider the simple linear model with a single fixed regressor.

Let $X_1, ..., X_n$ be an iid sample with $E[X_i] = \mu$ and $var(X_i) = \sigma^2 < \infty$. Then

$$
Y_n = \sqrt{n} \left( \frac{\bar{X} - \mu}{\sigma} \right) \xrightarrow{d} Z \sim N(0,1) \text{ as } n \rightarrow \infty
$$

$$y_i = x_i\beta + \epsilon_i,$$

where $x_i$ is fixed and $\epsilon_i$ is iid $(0, \sigma^2)$. The least squares estimator is

$$\hat{\beta} = \left(\sum_{i=1}^{n} x_i^2\right)^{-1} \sum_{i=1}^{n} x_i y_i = \beta + \left(\sum_{i=1}^{n} x_i^2\right)^{-1} \sum_{i=1}^{n} x_i \epsilon_i,$$

and

$$\sqrt{n}(\hat{\beta} - \beta) = \left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right)^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i \epsilon_i.$$

The CLT needs to be applied to the random variable $w_i = x_i\epsilon_i$. However, even though $\epsilon_i$ is iid, $w_i$ is not iid since $var(w_i) = x_i^2\sigma^2$ and, thus, varies with $x_i$.

The Lindeberg-Feller CLT is applicable for the linear regression model with fixed regressors.