THE AUTOREGRESSIVE DISTRIBUTED LAG EXTENSIONS AND APPLICATIONS

SAM CHUNG YAN

UNIVERSITI SAINS MALAYSIA

2022

THE AUTOREGRESSIVE DISTRIBUTED LAG EXTENSIONS AND APPLICATIONS

by

SAM CHUNG YAN

**Thesis submitted in fulfilment of the requirements
for the Degree of
Doctor of Philosophy**

July 2022

ACKNOWLEDGEMENT

First and foremost, I would like to thank my supervisor, Associate Prof. Dr. Goh Soo Khoon, from the Centre for Policy Research & International Studies (CenPRIS), Universiti Sains Malaysia. Dr. Goh as a supervisor she also acts like a family member to me. She always put me in the first place to think about how and what she could do to support me, especially the financial help and the study progression. The main problem in my Ph.D. study was financial support. When I started my PhD study, I did not have any financial support. It was a big problem for me because my family could not afford any financial support to further my study. To ensure I am free of this headache problem, she tried her best to find any financial aids for me, including applying USM fellowship and getting me any research assistant posts under her various research grants. There was once a period she even used her money to pay my living fee. I will never forget what she did to me throughout my studies, and I do not know how to express my gratitude to her. All I can do is sincerely saying, thank you.

I also thank my co-supervisor, Emeritus Prof. Dr. Robert McNown, from University of Colorado at Boulder, Colorado, USA. I am fortunate to get his supervision and continuous support. It’s never an easy job to receive questions and respond to them through emails exchange. However, this is the way we had our communication throughout the years. I would like to take this opportunity to thank Prof McNown and his wife, who hosted Dr. Goh and me to stay in their house for intensive research in 2018 summer. That was a precious experience I had during my visit to the US. With his expertise and support, I finish my Ph.D. study on time.

I would like to give special thanks to my friends Peter See Yew Boon and Teh Khai Loon. Throughout my postgraduate study, they provided me with a safe and comfortable room to live in during my study in Penang. They understood my situation very well and did not request any rental fee or electric bill from me so that I could focus on my study without giving any financial burden to me. It was the happiest moment with their company. I did not feel lonely during my postgraduate study.

My PhD journey was challenging. I faced many challenges in the journey, including financial problems, peer pressure, and an indeterministic future. Fortunately, I withstood all the challenges and went through all the obstacles with my determination and the support from my family, my supervisors, and friends. Without them, my PhD journey would never have come to an end. Lastly, I would like to express my gratitude and sincerely thank you to all of them. Much appreciated. I believe it is going to be a good start in the next stage of my life.

TABLE OF CONTENTS

ACKNOWLEDGEMENT ……………………………………………………………….. **ii**

TABLE OF CONTENTS ………………………………………………………… **iii**

LIST OF TABLES ………………………………………………………………….. **vii**

LIST OF FIGURES ……………………………………………………….………… **ix**

LIST OF ABBREVIATIONS ………………………………………………………... **x**

ABSTRAK ……………………………………………………………………………. **xi**

ABSTRACT ………………………………………………………………………… **xiii**

CHAPTER 1 INTRODUCTION ……………………………………………………. **1**

1.1 Motivation ……………………………………………………………….……….... 1

1.2 Problem Statement ………………………………………………..………….. 3

1.3 Objectives of the Study ……………………………………………………… 5

1.4 Research Questions …………………………………………………………….. 6

1.5 Significance of the Study ………………………………………………………….. 7

1.6 Outline of the Thesis ……………………………………………………………… 7

CHAPTER 2 AN AUGMENTED AUTOREGRESSIVE DISTRIBUTED LAG BOUNDS TEST FOR COINTEGRATION ………………………… **10**

2.1 Introduction ………………………………………………………………… 10

2.2 The ARDL Bounds Testing Procedures ………………………………………. 14

2.2.1 The System of the ARDL Framework ………………………………… 14

2.2.2 The Hypothesis Testing ……………………………………………….. 16

2.2.3 The Lower, Intermediate and Upper Bounds Critical Values ………… 17

2.2.4 Degenerate Lagged Independent Variables …………………………… 19

2.2.5 The $I(1)$ Dependent Variable to Rule Out the Case of Degenerate Independent Variables ……………………………………………… 22

2.2.6 The Bootstrap ARDL Test …………………………………………….. 24

2.3 The Theorems, Distribution, and Simulation Setup for the Lagged Level Independent Variables Test …………………………………………………… 26

2.3.1 The Assumptions of ARDL Bounds Approach ……………………….. 26

2.3.2 The Asymptotic Distribution of Lagged Level Independent Variables Statistic ……………………………………………………... 29

2.3.3 Simulation Setup and Tables of Critical Values 33

2.4 Empirical Application: Revisiting the US, the UK, and France Spend-Tax Correlations 39

2.4.1 Empirical Model and Methodology 40

2.4.2 Data and Unit Root Tests 41

2.4.3 The Empirical Model and Methodology 43

2.4.4 Empirical Results and Discussions 49

2.5 Conclusion 51

CHAPTER 3 A MULTIVARIATE AUTOREGRESSIVE DISTRIBUTED LAG FRAMEWORK FOR UNIT ROOT TESTING 52

3.1 Introduction 52

3.2 The Development of Unit Root Test from Univariate to Multivariate 56

3.3 The Model Framework, Statistical Distributions, and Testing Procedures 61

3.3.1 The ARDL Model and Assumptions 61

3.3.2 Possible Cases Encountered and Its Description 63

3.3.3 The Null Hypotheses and Limiting Distributions for the Tests in ARDL Unit Root Test 66

3.4 Methodology 70

3.4.1 Data Generating Process (DGP) 70

3.4.2 Residual Bootstrap 72

3.4.2(a) Estimate the Optimal ARDL Model 73

3.4.2(b) Bootstrap *t*-Test for Lagged Level Dependent Variable Coefficient 73

3.4.2(c) Bootstrap *F*-Test for Lagged Level Independent Variables Coefficients 75

3.4.3 Size and Power Analysis 78

3.5 Simulation Results and Discussion 81

3.5.1 Size and Power Analysis for DGPs without Contemporaneous Correlation 81

3.5.2 Size and Power Analysis for DGPs with Contemporaneous Correlation 88

3.6 Empirical Application 94

3.7 Conclusion 99

CHAPTER 4 A RE-EXAMINATION OF TAYLOR RULE STUDIES ………… **101**

4.1 Introduction ……………………………………………………………….. 101

4.2 Econometric Challenges in Taylor Rule Empirical Studies …………………. 105

4.2.1 Unit Root Test and Cointegration Issues …………………………….. 105

4.2.2 Unbalanced Regression and ARDL Cointegration Test ……………… 108

4.2.3 Model Misspecification Issues ……………………………………… 110

4.2.4 The Inappropriate Usage of Robust Standard Error in the Taylor Rule for Residual Autocorrelation and Heteroscedasticity Problems .. 113

4.2.5 The ARDL Specification to Correct the Model Misspecification Issue ……………………………………………………………….. 114

4.3 Methodology ……………………………………………………………… 116

4.3.1 Procedures in Re-Examining the Recent Empirical Studies …………. 116

4.3.2 The Augmented ARDL Bounds Test and the Level Relationship Estimates ………………………………………………………….. 117

4.4 Re-Examining Taylor Rule Empirical Studies – A Discussion of Econometric Problems and their Resolutions ……………………………... 120

4.4.1 Empirical Study #1: Forecast Uncertainty and the Taylor Rule, by Bauer and Neuenkirch (2017) ……………………………………. 120

4.4.1(a) Bauer and Neuenkirch (2017)'s Empirical Models and Data Description ………………………………………… 121

4.4.1(b) Bauer and Neuenkirch (2017)'s Problems in the Estimating Procedures of the Paper ……………………… 123

4.4.1(c) Re-Examining Bauer and Neuenkirch (2017) Study using ARDL Approach ………………………………….. 126

4.4.2 Empirical Study #2: Stock Market and Exchange Rate Information in the Taylor Rule: Evidence from OECD Countires, by Heimonen et al. (2017) ………………….……………………………………….. 132

4.4.2(a) Heimonen et al. (2017)'s Empirical Models and Data Description ………………………………………… 133

4.4.2(b) Problems in the Estimating Procedure of Heimonen et al. (2017) Paper …………………………………….. 134

4.4.2(c) Re-Examining Heimonen et al. (2017)'s Study using ARDL Approach ………………………………………… 140

4.4.3 Empirical Study #3: The Taylor Rule, Wealth Effects and the Exchange Rate, by Wang et al. (2016) ……………………………….. 149

4.4.3(a) Wang et al. (2016)'s Empirical Models and Data Description 149

4.4.3(b) Problems in the Estimating Procedure of Wang et al. (2016) Paper 151

4.4.3(c) Re-Examining Wang et al. (2016)'s Study using ARDL Approach 153

4.4.3(d) Re-Examining the Augmented Model Estimation Using ARDL Approach 158

4.5 Conclusion 163

CHAPTER 5 CONCLUSION 166

5.1 Summary of the Thesis 166

5.2 Lessons and Implications of the Study 170

5.3 Recommendations for Future Research 172

REFERENCES 174

APPENDICES

LIST OF TABLES

		Page
Table 2.1	Case I (No intercept no trend)	35
Table 2.2	Case III (Unrestricted intercepts and no trend)	36
Table 2.3	Case V (Unrestricted intercepts; unrestricted trends)	37
Table 2.4	Unit root tests	43
Table 2.5	Diagnostic checking for the optimal models	47
Table 2.6	Model specification and cointegration tests	48
Table 2.7	Summary of the types of upper bounds critical values	49
Table 3.1	Size and power analysis at 5% nominal level for y_t with no-intercept, T=100, N=1000, B=1000	82
Table 3.2	Size and power analysis at 5% nominal level for y_t with intercept, T=100, N=1000, B=1000	85
Table 3.3	Size and power analysis at 5% nominal level for y_t with intercept and trend, T=100, N=1000, B=1000	87
Table 3.4	Size and power analysis at 5% nominal level for y_t with no-intercept, T=100, N=1000, B=1000, $\rho = 0.3$	88
Table 3.5	Size and power analysis at 5% nominal level for y_t with intercept, T=100, N=1000, B=1000, $\rho = 0.3$	90
Table 3.6	Size and power analysis at 5% nominal level for y_t with intercept and trend, T=100, N=1000, B=1000, $\rho = 0.3$	91
Table 3.7	Univariate unit root tests	97
Table 3.8	Multivariate ARDL unit root test for RINT	98
Table 4.1	Taylor rule estimates: pre-crisis period (From Bauer and Neuenkirch, 2017, page 106)	123
Table 4.2	Taylor rule estimates: full sample (From Bauer and Neuenkirch, 2017, page 107)	124
Table 4.3	Unit root tests results	128
Table 4.4	Estimates of the optimal model ARDL(4,4,3,5,2) for the cointegration test	130

Table 4.5 Unit root tests statistics provided by the authors (see Heimonen et al. 2017, p. 7-8) …………………………………………………………. 136
Table 4.6 Results from the estimation of the linear Taylor rule using real-time data (From Heimonen et al., 2017, p. 10, Table 1) ………………….. 139
Table 4.7 Unit root tests results ………………………………………………… 142
Table 4.8 Estimates of the optimal model ARDL(7,0,0,11,3,2) for the cointegration test ………………………………………………….. 144
Table 4.9 Dummies for ARDL(7,0,0,11,3,2) …………………………………. 145
Table 4.10 Level estimates for the case of Germany by the authors …………….. 146
Table 4.11 GLS estimates for Regression (4.23) ……………………………… 147
Table 4.12 Dummies for level estimates of Regression (4.23) …………………… 147
Table 4.13 Unit root tests results ………………………………………………… 155
Table 4.14 Estimates of the optimal model ARDL(5,7,6,7) for the cointegration Test …………………………………………………………………... 157
Table 4.15 Unit root tests results ………………………………………………… 159
Table 4.16 Estimates of the optimal model ARDL(5,7,6,7,8,1) for the Cointegration test ………………………………………………….. 161
Table 4.17 Normalized level relation estimates for interest rate …………………. 162

LIST OF FIGURES

Page

Figure 2.1 Time series plots of LREXP and LRREV from period 1972 to 2015 for (a) the US, (b) the UK, (c) France ………………………….. 42

Figure 3.1 Time series plots for Malaysia and the US inflation rates ……………. 96

Figure 4.1 Time series plots for CBRATE (interest rate), CPI_GAP (expected inflation gap), GDP_GAP (expected output gap), CPI_SD (uncertainty of inflation expectation), GDP_SD (uncertainty of growth expectation) ………………………………………………. 127

Figure 4.2 Time series plots for i_t (nominal 3-month interest rate), π_t (actual inflation), $\tilde{\pi}_t$ (deviation of annual CPI-inflation from its trend value), $\overline{y}_t$ (deviation of real economic activity from its trend value), q_t (real effective exchange rates), d_t (dividend yield) of Germany …………………………………………………. 141

Figure 4.3 Time series plots for variables INT (interest rate), INF (inflation), Y (output gap), and REX (real exchange rate) for Australia ……………………………………………………….... 154

Figure 4.4 Time series plots for variables *fw* (financial wealth) and *hw* (housing wealth) for Australia ……………………………………. 158

LIST OF ABBREVIATIONS

The following table describes the significance of various abbreviations and acronyms used throughout the thesis. The page on which each one is defined or first used is also given. Non-standard acronyms that are used in some places to abbreviate the names of certain white matter structures are not in this list.

Abbreviation	Meaning	Page
ADF	augmented Dickey-Fuller	4
AIC	Akaike information selection	44
ARDL	autoregressive distributed lag	1
ARMA	autoregressive moving averate	52
CADF	covariate augmented Dickey-Fuller	7
DF	Dickey-Fuller	80
DF-B	Dickey-Fuller breakpoint	115
DF-GLS	Elliot-Rothenberg-Stock test	55
DGP	data-generating process	7
DOLS	dynamic ordinary least squares	150
ECM	error correction model	18
GLS	generalized least squares	4
GMM	generalized method of moments	4
HAC	heteroscedasticity-and-autocorrelation-consistent	123
KPSS	Kwiatkowski-Phillips-Schmidt-Shin	55
LM	Lagrange multiplier	115
MAIC	modified Akaike information selection	40
MSG	McNown et al. (2018)	9
NP	Ng-Perron	95
OLS	ordinary least squares	2
PP	Phillips and Perron	40
PSS	Pesaran et al. (2001)	9
RESET	regression equation specification error test	44
VAR	vector autoregressive	25
VECM	vector equilibrium correction model	13
ZA	Zivot-Andrews	115

LANJUTAN DAN APLIKASI AUTOREGRESIF SUSULAN TERTABUR

ABSTRAK

Kerangka Autoregresif Susulan Tertabur (ARDL) adalah kerangka yang mantap dengan kedua-dua fleksibel susunan lat pembolehubah bersandar dan bebas untuk menjelaskan dinamik pembolehubah bergerak balas. Tesis ini mengembangkan pendekatan ARDL untuk memberi praktikal aplikasi dalam analisi ekonomi. Perkembangan ini termasuk memberikan sempadan nilai kritikal pengujian tambahan pada pekali lat tahap pembolehubah bebas kepada ujian kointegrasi ARDL, membuat inovasi kerangka ARDL untuk ujian multivariat akar unit, dan mencadangkan penggunaan ARDL model dalam pengkajian peraturan Taylor. Pengguna yang tidak mahir pengaturcaraan boleh menggunakan nilai kritikal yang disediakan untuk menjalankan prosedur ujian sempadan yang biasa diketahui untuk mengkaji kointegrasi yang dicadangkan oleh McNown et al. (2018). Model ujian multivariat akar unit yang dinamik dan fleksibel ini mengandungi maklumat siri masa yang berkaitan membolehkan ia memperolehi lebih kuasa statistik daripada ujian univariat akar unit. Selain itu, kelemahan dalam kerangka covariate Dickey-Fuller yang mengabaikan kemungkinan hubungan kointegrasi dapat dilindungi oleh ARDL model supaya mengelakkan kehilangan kuasa atas sebab model misspesifikasi. Kebanyakkan kajian empirikal peraturan Taylor tidak mematuhi prosedur ekonometrik yang betul seperti isu ujian akar unit dan kointegrasi, dan pemeriksaan diagnostik. Banyak bukti menunjukkan persamaan peraturan Taylor adalah tidak seimbang yang mempunyai pembolehubah yang berbeza integrasi. Oleh itu, estimasi tahap tanpa kointegrasi yang dilaporkan dalam kajian empirikal mungkin memberikan keputusan palsu. Robust standard error terdapat

digunakan biasa dalam kajian empirikal tetapi kaedah ini tidak mampu menangani autokorelasi residual, terutamanya lat pembolehubah bersandar didapati dalam model itu. Kebanyakkan isu statistik dalam peraturan Taylor adalah disebabkan oleh misspecifikasi model. Namun, kaedah ARDL dapat menangani semua masalah ini dengan cekap. Perkembahan kaedah ARDL dalam tesis ini telah memberikan sumbangan yang penting kepada kesusasteraan ekonometrik dan analisi ujian empirical.

THE AUTOREGRESSIVE DISTRIBUTED LAG

EXTENSIONS AND APPLICATIONS

ABSTRACT

The autoregressive distributed lag (ARDL) framework is comprehensive with both flexible lag orders of dependent and independent variables to explain the dynamic of the responding variable. This thesis extends the ARDL methods to provide practical applications in economic analysis. The extensions include providing bounds of critical values of the additional testing on lagged level independent variables coefficients for the ARDL cointegration test, innovating the ARDL framework for a multivariate unit root test, and proposing ARDL model for Taylor rule studies. Users who are unfamiliar with programming could use the provided critical values to perform the familiar bounds testing procedure to run the cointegration test proposed by McNown et al. (2018). The dynamic and flexible model in the multivariate ARDL unit root test with related time series information helps to gain more statistical power than the univariate framework unit root tests. Besides that, the limitation in the covariate Dickey-Fuller framework that rules out the possibility of cointegration is covered by the ARDL model to avoid power loss caused by model misspecification. Many Taylor rule empirical studies do not follow proper econometric procedures such as unit root, cointegration, and diagnostic tests. Besides, many pieces of evidence show that the Taylor rule regression is unbalanced with mixed integration order variables. Therefore, level estimates without cointegration reported in the studies could give spurious results. The robust standard error is commonly used in empirical studies, but it is helpless to deal with the residual autocorrelation problem, especially if the model includes lagged dependent variable. Most of the statistical issues

in the Taylor rule are caused by model misspecification. Nevertheless, the ARDL method could efficiently address all of them. The extensions of the ARDL methods in this thesis show important contributions to the econometric literature and empirical studies analysis.

CHAPTER 1

INTRODUCTION

1.1 MOTIVATION

In statistics and econometrics, there are two fundamental models to explain the dynamics of a random variable or relationship between variables: the autoregressive model and the distributed lag model. The autoregressive model represents a random variable with a linear relationship on its own lagged (previous) values, and the distributed lag is a regression consisting of a linear relationship with both current and lagged values of the explanatory variables in predicting the current values of a dependent variable. The former is a simple model and does not include any explanatory variables information to estimate the current value of the process. For the latter, it lacks past information of the observed time series data. Both models have their own strengths and drawbacks. Nevertheless, it is possible to combine both models into one to yield a comprehensive model, namely the autoregressive distributed lag (ARDL). The ARDL with richer information framework consists of its own lagged values and both current and lagged values of explanatory variables to explain the dynamics of the variable.

The ARDL model is common and it plays a vital role in economic analysis because the dynamic characteristics of an economic variable could not be solely or fully explained by itself. The behaviour of an economic variable could be affected by the change in another economic variable. Moreover, the change in a variable does not always reflect immediately. The consequences of economic decisions that result in changes in an economic variable can last a long time. For instance, monetary and fiscal policy changes may take half a year to make a noticeable effect. Then, it may take a year or eighteen months for the policy effects to work through the economy (Peter, 2021). Because the ARDL model contains both dependent and explanatory variables, its rich information yields higher estimating power with lower variances, thus testing power. Besides, it provides robust results in small sample size (Narayan, 2004). Furthermore, the ARDL model is flexible by allowing different lag lengths in both dependent and independent variables, unlike the VAR model that restricting all its variables' lag length to be the same. This flexible lag length in the ARDL model further enhancing the model's fitness in explaining the level effect of the dependent variable. Besides, with appropriate modification of the orders of the ARDL model, the residuals serial correlation and the problem of endogeneous regressors can be corrected simultaneously (see the discussion in Pesaran and Shin (1999), page 14 to 15). In addition, we can also test the relationship between the dependent variable and explanatory variables. This ARDL model uses the conventional estimator, ordinary lest squares (OLS), to estimate parameters, and it is a single equation that is easy to use.

The ARDL regression is famous and has been used in econometrics for decades because of its simplicity and practical implication. One of the prominent representatives of the ARDL implication is the ARDL bounds testing method by Pesaran and Shin (1999) and Pesaran et al. (2001). Pesaran et al. (2001)'s ARDL bounds testing approach is a method of examining the cointegrating relationship between variables regardless of its regressors' integration order being $I(0)$ or $I(1)$. After establishing a cointegration relationship, one can analyze the level relationship introduced by Pesaran and Shin (1999). Pesaran et al. (2001)'s paper titled 'Bounds testing approaches to the analysis of level relationships' first published in Journal of Applied Econometrics volume 16 issue 3, page 289-326. This paper is famous and widely cited by researchers over the years. Up to the date of February 2020, according to Wiley Online Library, this paper has been cited 3694 times and 4593 from Researchgate. These citations amount is huge, and this number still in an increasing trend up to the date now.

Nevertheless, the application of the ARDL framework is not limited to the bounds procedures cointegration test only. There are other approaches also using the ARDL framework. For instance, nonlinear ARDL test by Shin et al. (2014), nonlinear panel ARDL test by Campbell and Thompson (2008), smooth transition autoregressive distributed lag (STARDL) model by Bildirici and Ersin (2018), etc. Although there is number of ARDL applications, none of them is used for unit root testing. It is possible to extend the inclusive ARDL framework for the unit root test as it covers more data-generating process environments than the existing unit root test. It helps to enhance the fitness of the model and reduce the variances within the model, therefore improving the statistical power in testing the unit root of a variable. Besides, the ARDL model is an efficient model for empirical studies. The single ARDL model easily passes all the statistical assumptions, giving robust estimates with the conventional OLS estimator and running unit root and cointegration tests with the same model. It makes the analysis procedures easy and simple. Thus, this thesis tries to employ the ARDL framework for the fundamental issue in the field of econometrics, including extension for a new test and demonstrate the usefulness of the model in the empirical study analysis. Hence, contribute more robust inferences and testing techniques in the time series econometrics.

1.2 PROBLEM STATEMENT

The comprehensive ARDL framework that contains both dependent and independent variables with flexible lag orders is useful and potentially for extensions for more applications. Recently, an innovation is made on the ARDL bounds testing done by McNown et al. (2018). They propose additional testing to the bounds test using bootstrap method to relax the bounds test restriction. It improves the ARDL bounds test to detect cointegration status more efficiently and overcome the inconclusive inferences. However, bootstrap ARDL could be too complicated by involving complicated computation and programming for those unfamiliar with coding. Thus, we propose an alternative method. We provide the limiting distribution and the critical values for the additional test to retain the bounds procedures instead of using bootstrap to generate the critical value.

Besides the cointegration test, inspired by Hansen (1995) and Pesaran et al. (2001) methodology, the ARDL framework can also be used for unit root test. There is not many multivariate unit root test in the existing literature. Hansen (1995) shows the importance of including the related explanatory variables into the univariate framework unit root test. He advocates the low power problem in the univariate unit root tests is because its framework ignores important variables. His simulation experiments show that the augmented Dickey-Fuller (ADF) test gains enormous power if related variables are included in the equation. Adding the related variables into the equation helps to reduce the variances in the model and hence, improve the estimation power. Therefore, the root of the problem is due to the model misspecification. However, Hansen's model rules out the possibility of cointegration among the variables. The comprehensive ARDL model is inclusive including the cointegration relationship, which could help to cover the weaknesses of Hansen's model. Thus, we develop a new multivariate unit root test with an ARDL framework that covers a wider range of environments.

Lastly, although the studies of the Taylor rule have been some time, the empirical studies still face many econometric challenges. Those challenges include unit root and cointegration issues and violation of statistical assumptions, especially the autocorrelation assumption. A model fails to reveal its variables' and system's stationarity, as well as to fulfil the necessary statistical assumptions, its estimation is highly questionable. Many pieces of evidence indicate the Taylor regression is an unbalanced regression (see, for example, Enders et al., 2010; Bunzel and Enders, 2010; Osterholm, 2005; Siklos and Wohar, 2006). Cointegration test is required to avoid spurious estimates. For the statistical assumptions, the common practice done by the researchers is to use alternative estimators like GLS, GMM, and robust standard error to deal with the problems. However, the notorious autocorrelation problem in Taylor rule studies is mainly due to model misspecification. To overcome the underlying issues effectively, the ARDL framework is appropriate. The flexible specification of the ARDL model that allows different lag length in the variables to maximize the fitness to the data to explain the dynamic relationship governing the interest rate to meet all the necessary statistical assumptions. Moreover, the ARDL equation also provides the framework to test for cointegration when there are mixed integration orders variables. If there is a cointegration relationship, we can further estimate the long run relationship from the optimal ARDL model.

1.3 OBJECTIVES OF THE STUDY

There are several objectives in this thesis. These objectives are related to the problem statement discussed above, which include the derivation of the limiting distribution of the additional test suggested by McNown et al. (2018), the development of a new multivariate unit root test with ARDL based, and propose using the ARDL methods for the Taylor rule empirical studies. This thesis motivates researchers to apply the ARDL methods to give a more precise cointegration test without falling into a fake cointegration trap, to get a more reliable unit root testing with lower inference error by the ARDL unit root test, and to estimates the empirical Taylor rule efficiently without using additional methods to overcome the statistical assumptions void. To be specific, the main objectives are as follows:

(i) To derive limiting distribution and critical values of the *F*-test for joint significance of lagged level independent variables coefficients, which can be used as an additional ARDL cointegration test.

(ii) To develop a new multivariate unit root test based on the ARDL framework.

(iii) To discuss the inappropriate practices in the Taylor rule empirical studies and to apply ARDL methods for efficient estimation.

1.4 RESEARCH QUESTIONS

The research objectives of this thesis are stated above. To ensure we achieve our aims through the research, here are the questions we need to answer:

(i) What is the limiting distribution and critical values of the *F*-test for joint significance of lagged level independent variables coefficients which used as an additional ARDL cointegration test?

(ii) How to apply a new multivariate unit root test based on the ARDL framework?

(iii) What are the inappropriate practices in the Taylor rule empirical studies and how the ARDL methods are applied for the efficient estimation.

1.5 SIGNIFICANCE OF THE STUDY

This thesis has provided several contributions. The first contribution is conducting the additional test proposed by McNown et al. (2018) back to the familiar bounds procedure without the complicated bootstrap method. This eases the testing procedure and user friendly for those who are not familiar with programming. The second contribution is adding a high-performance unit root test with low statistical errors to the literature. The third contribution is pointing out the inappropriate practices in the Taylor rule empirical studies and demonstrating the ARDL methods efficiently resolve all the statistical challenges in the studies. Specifically, the contributions are listed as follows:

(i) Use bounds procedure for the additional test in ARDL cointegration test by the tables of the additional test's critical values to avoid falling into fake cointegration.

(ii) Deliver a new high performance multivariate unit root testing with low Type I and Type II errors and covering a wider range of data-generating process environments.

(iii) Warn the consequences of the inappropriate practices in the Taylor rule empirical studies and demonstrate the ARDL methods are efficient in the estimation by fulfilling all the statistical assumptions without any robust testing and correction.

1.6 OUTLINE OF THE THESIS

This thesis is written in the form of paper-in-chapter and each chapter responds to each objective listed above. The breakdown of the thesis is the Introduction, Objective 1, Objective 2, Objective 3, and the General Conclusion of the thesis. The contents of the chapters include innovation and improvement of the existing approaches related to the ARDL framework, delivering new limiting distributions of the test statistic, and proposing alternative procedures which help to resolve most challenges faced by the empirical studies.

Chapter 2 begins with the innovation of the well-known Pesaran et al. (2001) ARDL bounds testing procedure for cointegration, namely the augmented ARDL bounds testing. At the beginning of the chapter, we discuss the background of the Pesaran et al. ARDL bounds test and the possibility of the degenerate cases, the fake cointegration. McNown et al. (2018) pointed out the importance of the third test for the lagged level independent variable and the existing testing, the overall *F*-test and *t*-test for lagged level dependent variable, to confirm the exact cointegration status. However, Pesaran et al. do not provide the distribution theory and critical values for the third test. McNown et al. used the bootstrap method to generate bootstrap critical value for the third test. This chapter derives the distribution theory for the third test and presents its limiting distribution, hence, tables of critical values. Without relying upon complicated computation and programming, it is possible to run the third test with the conventional procedure by providing critical values. At the end of the chapter, an empirical example is included to demonstrate how the augmented ARDL bounds test is conducted in testing cointegration.

In Chapter 3, we discuss the new unit root testing based on the multivariate ARDL framework. First, we discuss the development of the multivariate unit root tests and how the ARDL framework improves the existing multivariate unit root test based on Covariate Augmented Dickey-Fuller (CADF). The main advantage of the ARDL framework over the CADF is the inclusion of the cointegration relationship. The ARDL test covers a wider range of data-generating processes (DGP) to impose a valid common factor and avoid model misspecification. Hence, further improve the power of the test in testing the unit root of a process. In this chapter, we discuss the procedures to carry out the multivariate ARDL test and interpret the outcomes given by the test. Several sets of experiments for size and power study are conducted to prove the reliability of the test. At the end of the chapter, an empirical example using the ARDL test is demonstrated.

In Chapter 4, we re-examine the Taylor rule empirical studies and propose to use the ARDL method to resolve the underlying challenges of the studies. In the beginning, we first discuss the inappropriate practices used in the existing empirical studies of Taylor rule, which include skipping the examination of the time series unit root properties, cointegration test was not conducted even though there are variables are potentially unit rooted, autocorrelation problem and the abuse of the usage of robust standard error. Without proper execution of the econometric procedures, invalid or misleading results could have resulted. However, we found that the root of the causes related to the model misspecification. The models in the empirical studies were too restrictive. Thus, we propose the flexible ARDL model and related methods for the Taylor rule empirical studies. Three recently published papers are selected for the re-examination to demonstrate the misdoings and to show how the ARDL methods efficiently overcome the issues. Lastly, in the last chapter we make a general conclusion of the thesis in responding to the Objectives 1, 2, and 3.

CHAPTER 2

AN AUGMENTED AUTOREGRESSIVE DISTRIBUTED LAG BOUNDS TEST FOR COINTEGRATION

2.1 INTRODUCTION

In the early 2000s, Pesaran et al. (2001), henceforth PSS, introduced a cointegration testing approach called the autoregressive distributed lag (ARDL) bounds test. This approach became popular as it breaks the traditional restriction of cointegration tests in that the tested variables must be non-stationary and all the variables are integrated of the same order. Some researchers favour this approach as many of the applications involve economic variables of mixed or unknown order of integration. The conventional cointegration testing restriction, as in the Engle-Granger test (1987) or the Johansen test (1991, 1995), raises problems in conducting cointegration analysis involving mixed orders of variables. In these cases, researchers may either transform the variables into a stationary form, precluding a finding of cointegration, or inappropriately drop some variables.

Nevertheless, PSS made some assumptions in developing the bounds testing approach. These include the exogeneity of the explanatory variables, the dependent variable must be $I(1)$, and the absence of degenerate cases. However, as pointed out by McNown et al. (2018), MSG henceforth, these assumptions were sometimes ignored by researchers, possibly leading to misleading conclusions (see Goh & McNown, 2015; Goh et al., 2017a, Goh et al., 2017b).

PSS introduced two tests for cointegration: the overall *F*-test on all the lagged level variables and the *t*-test on the lagged level of the dependent variable. But these tests need to work with the assumption of $I(1)$ dependent variable to rule out the possibilities of degenerate cases to make a valid conclusion. Degenerate cases, as pointed out by PSS imply non-cointegration. The degenerate cases arise when either the lagged level of the dependent variable or lagged level(s) of the independent variable(s) in the error correction term are found to be insignificant. The case with insignificance of lagged level of the dependent variable is known as the *degenerate lagged dependent variable case* (named degenerate #2 in MSG) while the case of insignificant lagged levels of the independent variables is known as the *degenerate lagged independent variable(s) case* (called degenerate case #1 in MSG). This incomplete error correction term does not close the residual gap between the dependent and independent variable(s) and thus, cointegration does not hold.

In conducting the bounds test, the significance of the overall *F*-test suggests the lagged level of the variables are jointly significant. However, this significance of the *F*-test may arise solely from either the lagged level of the dependent variable or the lagged level of the independent variable(s) alone. Therefore, the *t*-test for the lagged level of the dependent variable is needed to rule out the *degenerate lagged dependent variable case*. The additional assumption that the dependent variable is $I(1)$, rules out the occurrence of the *degenerate lagged independent variable(s) case*. The reason is that if only the lagged level dependent variable is significant, then the ARDL equation reduces to a (generalized) Dickey-Fuller equation. The significance of this lagged dependent variable term implies that the dependent variable is $I(0)$.

Rather than assuming the dependent variable to be $I(1)$, MSG introduced an additional test to examine the significance of the lagged levels of independent variable(s).

The advantage of the additional test is to overcome the reliance on the assumption of an $I(1)$ dependent variable to rule out the degenerate case. This reduces the risk of false conclusions based on standard unit root tests that have low power. This is also consistent with the spirit of ARDL bounds testing that makes minimal assumptions about the orders of integration of the variables in the analysis.

By combining this new test with the two tests presented by PSS, we gain a complete picture of the cointegration status of the system. If all the three tests (overall F-test on lagged level variables, t-test on the lagged level of the dependent variable, and F-test on the lagged levels of the independent variable(s)) are found to be significant, we can conclude that there is cointegration. If the overall F-test and the t-test on the lagged dependent variable are found to be significant but not the test on the lagged independent variable(s), this indicates that it is a degenerate lagged independent variable(s) case. Another possibility is when the overall F-test and the t-test on the lagged level of the independent variable(s) are found to be significant but not for the t-test on the lagged level of the dependent variable. This falls into the case of a degenerate lagged dependent variable. Either of the degenerate cases will imply a case of no cointegration; therefore all three tests must be applied to reach a valid conclusion.

To conduct the test on the lagged levels of the independent variable(s), MSG used a bootstrap procedure to generate its critical values. However, this involves programming and computation that are not convenient or user friendly. Based on theorems that establish the limiting distributions of this test statistic, we provide tables of critical values for the test on the lagged levels of the independent variables. By providing the tables of critical values, this eases the implementation of the test, so that it becomes accessible to a broader range of researchers. The contribution of this chapter will be of use to those wishing to apply the widely employed ARDL procedure with the three tests combined to arrive at a clear conclusion on the status of cointegration. At the end of the chapter, we illustrate the use of the generated critical values in the augmented ARDL bounds testing framework by revisiting the spend-tax relationship for the US, the UK and France.

This chapter is organized as follows. Section 2.2 discusses the ARDL bounds testing procedures and how the bounds testing concept works in testing the cointegration. In Section 2.3, the theorems and distribution of the additional *F*-test for the lagged level independent variable(s) are derived. The simulation setup to generate the critical values discussed in this section too. Section 2.4 illustrates the use of the generated critical values in the augmented ARDL bounds testing framework by revisiting the spend-tax relationship for the US, the UK and France. Lastly, the conclusion of the study in Section 2.5.

2.2 THE ARDL BOUNDS TESTING PROCEDURES

The PSS ARDL bounds testing approach is popular and widely adopted by the researchers. However, some may not fully understand the theoretical concepts and assumptions behind the method. Hence, this method is misused. Those misappropriations include without assuming the dependent variable being $I(1)$ and carry out a single test, the overall F-test, to infer the cointegration status. This action could fall into the degenerate cases, the fake cointegration cases, and thus given misleading results. This section will go into a detail discussion about the bounds testing approach's theoretical framework, hypotheses, assumptions, and the possibilities of the degenerate cases.

2.2.1 The System of the ARDL Framework

Let us consider (1 + k)-vector variables, $\mathbf{z}_t$, and it can be partitioned into $\left(y_t, \mathbf{x}_t\right)$. These $\mathbf{z}_t \approx \left(y_t, \mathbf{x}_t\right)$ variables form a vector equilibrium correction model of order p or VECM(p). The VECM can be written as

$$\Delta\mathbf{z}_t = \mathbf{c}_0 + \mathbf{c}_1 t + \mathbf{\Pi}\mathbf{z}_{t-1} + \sum_{i=1}^{p-1}\mathbf{\Gamma}_i^{'}\Delta\mathbf{z}_{t-i} + \boldsymbol{\omega}^{'}\Delta\mathbf{x}_t + \mathbf{u}_t\,. \tag{2.1}$$

Alternatively, the system can be represented in the form of

$$\Delta y_t = c_0 + c_1 t + \pi_{yy} y_{t-1} + \boldsymbol{\pi}_{yx}\mathbf{x}_{t-1} + \sum_{i=1}^{p-1}\boldsymbol{\psi}_i^{'}\Delta\mathbf{z}_{t-i} + \boldsymbol{\omega}_y^{'}\Delta\mathbf{x}_t + u_t\,, \tag{2.2}$$

$$\Delta\mathbf{x}_t = \mathbf{c}_{x0} + \mathbf{c}_{x1} t + \boldsymbol{\pi}_{xy} y_{t-1} + \mathbf{\Pi}_{xx}\mathbf{x}_{t-1} + \sum_{i=1}^{p-1}\mathbf{\Gamma}_{xi}^{'}\Delta\mathbf{z}_{t-i} + \boldsymbol{\omega}_x^{'}\Delta\mathbf{x}_t + \mathbf{u}_{xt}\,, \tag{2.3}$$

where c_0 and $\mathbf{c}_{x0}$ are coefficients of intercepts, c_1 and $\mathbf{c}_{x1}$ are coefficients of deterministic trends, π_{yy}, $\boldsymbol{\pi}_{yx}$, $\boldsymbol{\pi}_{xy}$ and $\mathbf{\Pi}_{xx}$ are coefficients of lagged levels of variables,

$\boldsymbol{\psi}$ and $\boldsymbol{\Gamma}_x$ are coefficients of lagged differenced of variables, $\boldsymbol{\omega}_y$ and $\boldsymbol{\omega}_x$ are coefficients of lagged independent variables, and u_t, $\mathbf{u}_{xt}$ are error processes. PSS assume *k*-vector of $\boldsymbol{\pi}_{xy} = \mathbf{0}$, indicate no feedback effect running from level of y_t into the *k*-vector of $\mathbf{x}_t$ in the system. Therefore, the y_{t-1} is absent in (2.3). The $\mathbf{x}_t$ serves as *forcing variables* to y_t or $\mathbf{x}_t$ is weakly exogenous. By imposing the restriction, the VECM above will reduce into vector *conditional* equilibrium correction model

$$\Delta y_t = c_0 + c_1 t + \pi_{yy} y_{t-1} + \boldsymbol{\pi}_{yx.x}\mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\psi}_i^{'}\Delta\mathbf{z}_{t-i} + \boldsymbol{\omega}^{'}\Delta\mathbf{x}_t + u_t \text{ [1]}, \tag{2.4}$$

$$\Delta\mathbf{x}_t = \mathbf{a}_{x0} + \mathbf{a}_{x1}t + \boldsymbol{\Pi}_{xx}\mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_{xi}^{'}\Delta\mathbf{z}_{t-i} + \mathbf{u}_{xt}\,. \tag{2.5}$$

Through the implementation of $\boldsymbol{\pi}_{xy} = \mathbf{0}$, now the system is restricted with at most one conditional level relationship between y_t and $\mathbf{x}_t$. Thus, the ARDL bounds test is applied only when our interest is to see the level relationship between y_t and $\mathbf{x}_t$. Interaction level effects among the variables are not allowed under the assumption restriction. That said, for instance, there is a trivariate case with y_t, x_t and z_t. Assuming x_t and z_t can have impacts to y_t in long-run but it supposes that the y_t does not has its impact to either x_t or z_t. Thus, we can only formulate a single ARDL equation using y_t as dependent variable and treating x_t and z_t as regressors. Multiple equations for x_t and z_t that served as dependent variables are not allowed. However, this restriction does not extend to its lagged changes or the short-run estimation. Nevertheless, findings from MSG found that this exogeneity

[1] Note that this single-equation conditional ECM model is commonly referred as ARDL model in the literature. We use the term 'ARDL' to infer the ECM model throughout the thesis, unless there is a need for further clarification.

assumption has minor effect to the estimation power. Estimation still valid even the exogeneity assumption is voided.

2.2.2 The Hypothesis Testing

In carrying out the ARDL bounds procedures for cointegration test, PSS provide tables of critical values with bounds, namely, the lower bound critical values and upper bound critical values for the two tests: overall *F*-test and lagged dependent variable *t*-test. Given the single ARDL regression (2.4),

$$\Delta y_t = c_0 + c_1 t + \pi_{yy} y_{t-1} + \boldsymbol{\pi}_{yx.x} \mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\psi}_i' \Delta \mathbf{z}_{t-i} + \boldsymbol{\omega}' \Delta \mathbf{x}_t + u_t \,, \tag{2.4}$$

the null and alternative hypotheses for the *F* test are given by

$$H_0^F = H_0^{\pi_{yy}} \cap H_0^{\pi_{yx.x}} \qquad \text{vs} \qquad H_1^F = H_1^{\pi_{yy}} \cup H_1^{\pi_{yx.x}} \,,$$

where $H_0^{\pi_{yy}} : \pi_{yy} = 0,\ H_1^{\pi_{yy}} : \pi_{yy} \neq 0$ and $H_0^{\pi_{yx.x}} : \boldsymbol{\pi}_{yx.x} = \mathbf{0}',\ H_1^{\pi_{yx.x}} : \boldsymbol{\pi}_{yx.x} \neq \mathbf{0}'$, and the null and alternative hypotheses for the *t*-test are given by

$$H_0^t : \pi_{yy} = 0 \qquad \text{vs} \qquad H_1^t : \pi_{yy} < 0 \,.$$

The significance of the overall *F*-test rejects the null of $H_0^F = H_0^{\pi_{yy}} \cap H_0^{\pi_{yx.x}}$ and accepts its alternative $H_1^F = H_1^{\pi_{yy}} \cup H_1^{\pi_{yx.x}}$, indicate the coefficients of the lagged level variables from (2.4) are jointly significant. For the *t*-test, the significance of the *t*-test indicates it rejects $H_0^t : \pi_{yy} = 0$ and accepts its alternative $H_1^t : \pi_{yy} < 0$, indicate the coefficient of the lagged level dependent variable is significant.

According to PSS, if one found that the overall *F*-statistic and the lagged dependent variable *t*-statistic falls outside their corresponding upper bound critical values, given the dependent variable is $I(1)$, cointegration can be concluded irrespective the regressors be purely $I(0)$, purely $I(1)$ or mutually cointegrated; On the other hand, if any test statistics fall inside the lower bound critical value, no-cointegration can be concluded irrespective the regressors be purely $I(0)$, purely $I(1)$ or mutually cointegrated; If the test statistic falls between the bounds, inconclusive inference is made and further actions need to be taken in order to make a conclusion.

2.2.3 The Lower, Intermediate and Upper Bounds Critical Values

The asymptotic distributions of the overall *F*-test and the lagged dependent variable *t*-test are non-standard. These distributions depend on the dimension and cointegration rank of the forcing variables $\mathbf{x}_t$, k and r, respectively. The distributions of the test statistics are mixture of standard stationary distribution and non-standard Dickey-Fuller unit-root distribution (see Pesaran et al., p. 298). For better understanding, the dimension k refers to the number of independent variables $\mathbf{x}_t$ while the cointegration rank refers to the number of $I(0)$, $I(1)$ variables in the $\mathbf{x}_t$. The cointegration rank is zero, $r = 0$, when the underlying regressors $\mathbf{x}_t$ are purely $I(1)$; Full cointegration rank, $r = k$, when the regressors $\mathbf{x}_t$ are purely $I(0)$; Reduced rank or mutually cointegrated, $0 < r < k$, when there is a mixture of $I(0)$, $I(1)$ regressors $\mathbf{x}_t$. Note, the null distributions for purely $I(0)$ regressors remain non-standard because under the null the dependent variable is still an $I(1)$.

An important corollary about the distributions of the test statistics discovered by PSS. That is, **the distribution with smaller cointegrating rank tends to shift further away (to the right for the *F*-statistic and to the left for the *t*-statistic) from the null value than the distribution with higher cointegrating rank**. Therefore, the size of the critical values for the null distribution of the tests with smaller cointegrating rank in the regressors is always greater than the one with a higher cointegrating rank:

$$CV_{\text{purely I(1)}} > CV_{\text{mixture of I(0), I(1)}} > CV_{\text{purely I(0)}},$$

where $CV_{\text{purely I(1)}}$ is the null distribution critical value with purely $I(1)$ regressors, $CV_{\text{mixture of I(0), I(1)}}$ is of mixture of $I(1)$ and $I(0)$ regressors, and $CV_{\text{purely I(0)}}$ is with purely $I(0)$ regressors.[2] The critical values with purely $I(1)$ regressors yield the greatest (absolute) critical values for the overall F-test and lagged dependent variable t-test, while the critical values with purely $I(0)$ regressors produce the smallest (absolute) critical values for the F- and t-tests. Hence, two polar cases appear. One polar $\mathbf{x}_t$ is purely $I(1)$, and another polar $\mathbf{x}_t$ is purely $I(0)$. If a test statistic falls outside the upper bound critical value, without knowing the integration/cointegration status of the underlying regressors, the test is significant regardless the underlying regressors are purely $I(0)$, purely $I(1)$ or mutually cointegrated; If the test statistic falls inside the lower bound critical value, the test is insignificant regardless the underlying regressors are purely $I(0)$, purely $I(1)$ or mutually cointegrated. Since PSS only provides the two polar critical values, if any test statistics falls between the bounds, the inference is inconclusive. The knowledge of the order of the integration of the underlying variables is required before conclusive inferences can be made. That is, to determine whether the test is significant, we need to know the

[2] See the statement given by PSS in page 298, last line in the last paragraph. The statement implicitly indicate the mentioned corollary.

cointegration rank r or the number of $I(0)$, $I(1)$ variables in the $\mathbf{x}_t$ and obtain its corresponding distribution's critical value. This particular critical value can be computed via stochastic simulations with the combination of k and cointegration rank r (see Pesaran et al., footnote 11, p. 299).

2.2.4 Degenerate Cases – The Fake Cointegration

There can be a few outcomes when carrying out the ARDL bounds test, i.e. cointegration, no-cointegration or degenerate cases. Degenerate cases, are fake cointegration or no-cointegration. This was first discussed in PSS paper on page 295, and MSG further the discussion. Degenerate cases arise when an incomplete error correction term is found from ARDL (or error correction) framework. This can be either the lagged level dependent variable or lagged level independent variables insignificant from the ARDL model.

Let us consider the single unrestricted intercepts and trends ECM model or equation (2.4) again:

$$\Delta y_t = c_0 + c_1 t + \pi_{yy} y_{t-1} + \boldsymbol{\pi}_{yx.x} \mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\psi}_i' \Delta \mathbf{z}_{t-i} + \boldsymbol{\omega}' \Delta \mathbf{x}_t + u_t \,. \tag{2.4}$$

The terms π_{yy} and $\boldsymbol{\pi}_{\mathbf{yx.x}}$ indicate the coefficients to the lagged level of dependent variable, y_{t-1} and lagged level of independent variable(s), $\mathbf{x}_{t-1}$, respectively. Both terms together can be referred as error correction term with representation of

$$\Delta y_t = c_0 + c_1 t - \alpha \varepsilon_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\psi}_i' \Delta \mathbf{z}_{t-i} + \boldsymbol{\omega}' \Delta \mathbf{x}_t + u_t \,, \tag{2.6}$$

where $\varepsilon_{t-1} = y_{t-1} - \alpha_0 - \alpha_1 t - \boldsymbol{\pi}_x \mathbf{x}_{t-1}$ and α is known to be the speed of adjustment. There are four possible cases with the combinations of the significance of π_{yy} and $\boldsymbol{\pi}_{\mathbf{yx.x}}$.

Case 1: $\pi_{yy} = 0$ and $\boldsymbol{\pi}_{\mathbf{yx.x}} = 0$ [**No-cointegration**].

In this case, both the coefficients π_{yy} and $\boldsymbol{\pi}_{\mathbf{yx.x}}$ are zeroes and the equation (2.4) is reduced to the simple ARDL equation in first differences,

$$\Delta y_t = c_0 + c_1 t + \sum_{i=1}^{p-1} \boldsymbol{\psi}_i' \Delta \mathbf{z}_{t-i} + \boldsymbol{\omega}' \Delta \mathbf{x}_t + u_t \ .$$

The zero values of π_{yy} and $\boldsymbol{\pi}_{\mathbf{yx.x}}$ indicate the error correction term do not exist. y_t and $\mathbf{x}_t$ are not cointegrated, and the deviation of the errors between y_t and $\mathbf{x}_t$ will not adjust back to equilibrium.

Case 2: $\pi_{yy} < 0$ and $\boldsymbol{\pi}_{\mathbf{yx.x}} \neq 0$ [**Cointegration**].

Case 2 describes that both the coefficients π_{yy} and $\boldsymbol{\pi}_{\mathbf{yx.x}}$ are non-zeroes. Since both coefficients are non-zeroes and given π_{yy} is negative, the error correction term is complete and linearly stationary. The deviation among the variables will be adjusted back to the equilibrium. Thus, y_t and $\mathbf{x}_t$ are cointegrated. The π_{yy} is equivalent to the speed of adjustment, -α, in equation (2.6).

Case 3: $\pi_{yy} = 0$ and $\boldsymbol{\pi}_{\mathbf{yx.x}} \neq 0$ [**Degenerate lagged dependent variable**].

Case 3 describes the coefficient of lagged dependent variable is zero, but the vector of lagged independent variables coefficients is non-zero. The equation (2.4) reduces to

$$\Delta y_t = c_0 + c_1 t + \boldsymbol{\pi}_{yx.x}\mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\psi}_i^{'}\Delta\mathbf{z}_{t-i} + \boldsymbol{\omega}^{'}\Delta\mathbf{x}_t + u_t \,.$$

The error correction term is incomplete due to the missing of the lagged level dependent variable. The deviation among the variables is disequilibrium and thus, y_t and $\mathbf{x}_t$ are not cointegrated. PSS and MSG describe this scenario as *degenerate lagged dependent variable case* and it is a fake cointegration. The *degenerate lagged dependent variable case* will mislead one to conclude cointegration if he only carries out the overall *F*-test on joint significance of π_{yy} and $\boldsymbol{\pi}_{\mathbf{yx.x}}$ during the test. This is because the significance of the $\boldsymbol{\pi}_{\mathbf{yx.x}}$ will cause the overall *F*-statistic significant too. Therefore, it is inappropriate to make conclusion based on the overall *F*-test only. This warning was given by MSG.

Case 4: $\pi_{yy} < 0$ and $\boldsymbol{\pi}_{\mathbf{yx.x}} = 0$ [**Degenerate lagged independent variables**].

Similar to Case 3, but Case 4 describes only lagged dependent variable coefficient is non-zero while the vector of lagged independent variables coefficients are zeroes. Again, Case 4 is a fake cointegration due to its incomplete error correction term. MSG name this case as *degenerate lagged independent variable case*. The missing $\boldsymbol{\pi}_{\mathbf{yx.x}}$ in equation (2.6) reduces to

$$\Delta y_t = c_0 + c_1 t + \pi_{yy} y_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\psi}_i^{'}\Delta\mathbf{z}_{t-i} + \boldsymbol{\omega}^{'}\Delta\mathbf{x}_t + u_t \,. \qquad (2.7)$$

Note that the equation (2.7) is a generalized Dickey-Fuller equation describing y_t is a stationary process. Careful attention should be given to Case 4, because one could fall into this degenerate case and mistakenly conclude cointegration even though he found that both the overall *F*-test and the lagged dependent variable *t*-test are significant. To prevent falling into the *degenerate lagged independent variable case*, one could assume y_t an *I*(1) process.

2.2.5 The *I*(1) Dependent Variable to Rule Out the Case of Degenerate Lagged Independent Variables

To establish the cointegrating relationship between y_t and $\mathbf{x}_t$ using the ARDL bounds testing procedure, PSS suggest the overall *F*-test to test the joint significance of the lagged level variables and the *t*-test to test for the significance of the lagged level of dependent variable.

Let us consider again the single unrestricted intercepts and trends ECM model or equation (2.4):

$$\Delta y_t = c_0 + c_1 t + \pi_{yy} y_{t-1} + \boldsymbol{\pi}_{yx.x} \mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\psi}_i' \Delta \mathbf{z}_{t-i} + \boldsymbol{\omega}' \Delta \mathbf{x}_t + u_t \,, \tag{2.4}$$

The significant overall *F*-test with the upper bound critical value indicate the lagged level variables coefficients from the ARDL model that is, π_{yy} and $\boldsymbol{\pi}_{\mathbf{yx.x}}$, are jointly significant regardless of the underlying regressors $\mathbf{x}_t$ are either purely *I*(0), purely *I*(1) or mutually cointegrated. However, the overall *F*-test alone is insufficient to confirm the cointegrating relationship between the y_t and $\mathbf{x}_t$ as the significance of the overall *F*-test could come solely from the significance of the π_{yy} or $\boldsymbol{\pi}_{\mathbf{yx.x}}$ alone (see McNown et al., 2018).

Therefore, PSS consider the Banerjee et al. (1998) procedure for further testing. Similarly, the Banerjee approach considers the single-equation (2.4) and run a *t*-test on the lagged level dependent variable coefficient, π_{yy}, for cointegration test. According to Banerjee methodology, if the π_{yy} is tested significant, then we confirm cointegration and there is a long-run relationship between y_t and $\mathbf{x}_t$. Banerjee et al. (1998) explain that this is because the equation (2.4) can be rearranged and described as

$$\Delta y_t = c_1 + c_2 t + \pi_{yy}\left(y_{t-1} - \boldsymbol{\beta}'_{yx}\mathbf{x}_{t-1}\right) + \sum_{i=1}^{p-1}\phi_i \Delta y_{t-i} + \sum_{j=1}^{q-1}\phi'_j \Delta \mathbf{x}_{t-i} + u_t, \tag{2.8}$$

where π_{yy} is then represents error correction adjustment coefficient with cointegrating vector $\left(1, \ -\boldsymbol{\beta}_{yx}\right)$, and $\boldsymbol{\pi}_{yx.x} = -\pi_{yy}\boldsymbol{\beta}_{yx}$. Thus, the significance of π_{yy} implies the existence of a long-run relationship between y_t and $\mathbf{x}_t$. However, Banerjee approach only considers cases where its underlying regressors are purely *I*(1). To meet the objective of avoiding the pre-testing issue, PSS therefore extend the Benerjee *t*-test to bounds procedure. Like the overall *F*-test, there are two polar critical values. The upper bound critical value is where its underlying regressors are purely *I*(1), and the lower bound critical value is where its underlying regressors are purely *I*(0). If the *t*-statistic falls outside the upper bound critical value, indicate π_{yy} is significant regardless the underlying regressors are either purely *I*(0), purely *I*(1) or mutually cointegrated; If the *t*-statistic falls inside the lower bound critical value, indicate π_{yy} is insignificant regardless the underlying regressors are either purely *I*(0), purely *I*(1) or mutually cointegrated.

Nevertheless, it is still insufficient to establish a cointegration relationship using Banerjee's procedure for the ARDL bounds test. Note that the proposition of the significant π_{yy} implies the error adjustment coefficient and thus, the existence of error correction term, valid only if $\boldsymbol{\beta}_{yx}^{'} \neq 0$. When $\boldsymbol{\beta}_{yx}^{'} = \mathbf{0}$, the cointegrating vector is $(1,\ \mathbf{0})$. Hence, the significant π_{yy} does not establish a cointegrating relationship. However, it falls to Case 4, the *case of degenerate lagged independent variable* with equation (2.7):

$$\Delta y_t = c_0 + c_1 t + \pi_{yy} y_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\psi}_i^{'} \Delta \mathbf{z}_{t-i} + \boldsymbol{\omega}^{'} \Delta \mathbf{x}_t + u_t . \tag{2.7}$$

As noted, equation (2.7) is a generalized Dickey-Fuller equation describing y_t is *I*(0). The assumption of *I*(1) y_t helps to prevent one fall into this degenerate case when both the overall *F*-test and lagged dependent variable *t*-test are significant. However, we want to stress that the lagged dependent variable *t*-test to establish cointegration is sufficient in Banerjee's approach because the approach assumes all the testing variables are *I*(1) including the dependent variable. The *case of degenerate lagged independent variable* is not possible to occur. Therefore, the ARDL bounds test requires the dependent variable to be *I*(1) to give a conclusive inference when both the overall *F*-test and lagged dependent variable *t*-test are significant.

2.2.6 The Bootstrap ARDL Test

The idea of the test on lagged independent variables was first initiated by MSG. The stud''s objective is to investigate the performances of the ARDL bounds test under various environments, including violating the assumption of exogeneity independent variables. It is found that the ARDL framework is suffering from some estimation weaknesses. In the literature, it is well documented that bootstrap methods help to improve the estimation, especially for small sample size testing. And bootstrap methods on cointegration tests are commonly used, such as Harris and Judge (1998), Seo (2006),