NEW UNIT ROOT TESTS TO DECREASE SPURIOUS RESULTS WITH APPLICATIONS IN FINANCE AND TEMPERATURE ANOMALIES

by

Edward Herranz
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
In Partial Fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Computational Sciences and Informatics

Committee:

____________________	Dr. James Gentle, Dissertation Director
____________________	Dr. George Wang, Committee Member
____________________	Dr. Igor Griva, Committee Member
____________________	Dr. Kirk Borne, Committee Member
____________________	Dr. Jessica Lin, Committee Member
____________________	Dr. Kevin Curtin, Acting Department Chair
____________________	Dr. Donna M. Fox, Associate Dean, Office of Student Affairs & Special Programs, College of Science
____________________	Dr. Peggy Agouris, Dean, College of Science

Date: ____________________

Spring Semester 2016
George Mason University
Fairfax, VA

New Unit Root Tests to Decrease Spurious Results with Applications in Finance and Temperature Anomalies

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

By

Edward Herranz
Master of Financial Engineering
Universidad de Alcalá, 2009
Master of Science
MIT, 1992
Bachelor of Science
MIT, 1990

Director: Dr. James Gentle, University Professor
Department of Computational and Data Sciences

Spring Semester 2016
George Mason University
Fairfax, VA

Dedication

This thesis work is dedicated to my wife Noralisa for all of her love and support; without her I would have never been able to come this far. This work is also dedicated to my daughter Gabriela; now we'll have more time to go to the park.

Acknowledgments

First I'd like to thank Dr. James Gentle, my dissertation chair, for his guidance, many fruitful discussions and for providing advice and encouragement to finish this long journey, as well as for teaching interesting courses that motivated me to pursue research ideas. I would like to thank my committee members: Dr. George Wang for providing direction in my study of econometric and time series topics; Dr. Kirk Borne for making me think of applications outside of finance; Dr. Igor Griva for pushing me to tie all of my research together and give practical applications and Dr. Jessica Lin for encouraging me to publish and provide helpful tips on improving my presentations.

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Symbols

1 Indicator function **1**{*condition*}. If the *condition* is met returns 1, otherwise 0. 35, 36, 79, 86, 88, 118, 120, 135, 137

ADF Augmented Dickey Fuller unit root test. The null hypothesis is that the tested time series follows an I(1) process, and the alternative hypothesis is that it follows an I(0) process. ix–xi, xiv, xv, 7, 8, 11, 31, 33, 52, 53, 61, 69, 71–74, 76, 77, 86, 90–93, 95, 98, 99, 102–106, 108–122, 125–129, 131–134, 137, 139–141, 150–153, 155, 158–161, 165–167, 175, 177, 182, 187, 190, 193, 200, 201, 207, 208, 219, 220

α Significance level used in statistical test. We consider 1%,5% and 10%. ix–xiv, 63, 71, 72, 75, 76, 83, 84, 87, 96–98, 100, 101, 104, 105, 107–109, 111, 112, 114, 116–118, 120, 122–124, 126, 129, 130, 132, 133, 135–138, 143–152

AR :AR(n): Auto-regressive process of order n. ix–xv, 1, 2, 5, 6, 8, 24, 27, 31, 35, 39, 52, 53, 61, 63, 64, 66, 69–72, 75, 83, 86, 88, 91–94, 96–98, 100, 102–120, 122–127, 130, 132, 133, 135, 136, 138–140, 144–153, 174, 175, 207, 215, 220, 221

ARMA :ARMA(n,p): Auto-regressive process of order n, and moving average process of order p. xi, 94, 101, 102, 221

BBC Bec, Ben Salem and Carrasco unit root test. The null hypothesis is that the tested time series follows an I(1) process, and the alternative hypothesis is that it follows an I(0) process under a three-regime self-exciting threshold autoregressive (SETAR) model.. 47, 130–135, 141, 200

CCBS A cross-currency basis swap. xv, xvii, xviii, 171, 172, 179–195

CDS A credit default swap. xv, 173, 174, 188–191

CV Critical value of a test statistic of a statistical test. xii, 76, 77, 87, 114, 144–152, 154

DGP The data generating process (DGP.) It refers to a statistical model that is an approximation of the true DGP. 15, 19, 21, 26–29, 86, 113, 135, 136, 138, 163

DLNN Deep Learning Neural Networks. A machine learning classification tool based on Neural Networks with more than 1 hidden layer and usually a neuron dropout feature during training to decrease overfitting. xiv, xv, xvii, 52, 155–163, 165–169, 220, 222

ECM Error correcting model implied by a two variable cointegrating relationship. 39

ERS Elliott Rothenberg Stock unit root test. The null hypothesis is that the tested time series follows an I(1) process, and the alternative hypothesis is that it follows an I(0) process. x, 32, 52, 53, 74–76, 92, 93, 95, 96, 98, 99, 101–104, 106, 108–112, 115–117, 119–122, 125, 126, 128, 129, 131–134, 139–141, 159–161, 165–167, 219, 220

$H_0^{I(1)}$ Null hypothesis of a unit root test. i.e. an AR(1) process that is I(1). ix–xv, xxi, 7, 69, 71, 72, 75, 76, 83, 84, 87, 95–105, 107–126, 129, 130, 132, 133, 135–138, 150–152, 155, 168, 186, 190, 200–205

$H_1^{\textbf{Explosive AR}}$ One possible alternative hypothesis of a unit root test where the process is an explosive AR(1) process with $\phi_1 > 1$. 8, 95, 99, 102–104, 106, 108, 110, 111, 115, 116, 119–121, 123, 124

$H_1^{I(0)}$ One possible alternative hypothesis of a unit root test where the process is an AR(1) I(0) process with $|\phi_1| < 1$. xxi, 8, 95, 96, 99, 102–104, 106, 109, 110, 112, 116, 117, 119, 121, 124, 125, 155, 202

HBPADF New procedure to test a time series to see if it is composed of various I(0) or I(1) segments. First the Bai-Perron algorithm is used to determine changes in any of intercept, linear trend an/or AR multiplier ϕ_1. Then we use the ADF unit root test on each segment to determine if it is likely nonstationary or stationary. xvi, 52, 53, 85, 88, 94, 139, 141, 204, 207, 208, 220, 221

HBPZA A new unit root test proposed in this thesis that allows structural breaks in intercept and trend in the null hypothesis based on first using the Bai-Perron algorithm to infer structural breaks in the series and then using the Zivot Andrews unit root test in each section, and returning a test statistic of the average of the ZA statistics. x, xvi, 52, 53, 78–84, 122–125, 140, 202–204, 220–222

I :I(n): Integrated process of order n. xii, xiii, xix–xxi, 6–8, 24, 31, 32, 37, 38, 62, 63, 65, 75, 80, 85, 86, 113, 126–129, 140, 141, 143, 154, 157, 158, 162, 164, 167, 175, 193

KPSS Kwiatkowski–Phillips–Schmidt–Shin (KPSS) stationarity test. The null hypothesis is that the tested time series follows an I(0) process, and the alternative hypothesis is that it follows a nonstationary I(1) process. 159–161, 165–167

l Length of time series. Most often 500 and 1000 are considered; assuming daily business time series this corresponds to 2 years or 4 years in the United States. ix–xv, 26, 27, 64, 66, 69–72, 75, 76, 83, 84, 87, 96–98, 100, 101, 104, 105, 107–109, 111–114, 116–118, 120, 122–126, 129, 130, 132, 133, 135–138, 144–154, 175

lagged-series New unit root test proposed in this thesis based on testing a series and its lagged version for cointegration. ix, x, xii–xv, 23, 52, 53, 61–63, 65, 68, 69, 71–76, 92, 93, 95–102, 104–122, 125, 126, 129–135, 140, 141, 150–153, 155, 159–161, 165–168, 175, 177, 182, 186, 190, 193, 219–221

m Number of simulations. All simulations are performed using the **R** language. Unless otherwise noted, the seed used for random number generation is 12345. ix–xv, 26, 27, 64, 66, 69–72, 75, 76, 83, 84, 87, 96–98, 100, 101, 104, 105, 107–109, 111–114, 116–118, 120, 122–126, 129, 130, 132, 133, 135–138, 144–152, 154, 175

ϕ_1 First multiplier of an auto regressive process as in $Y(t) = \phi_1 Y(t-1) + ...$ x, xi, xxii, 7, 8, 24, 25, 27–29, 35–37, 46, 51, 52, 61, 64–66, 68–76, 78, 82–84, 91–117, 119–127, 143, 158–160, 162–166, 202, 204, 207, 219

ϕ_2 Second multiplier of an auto regressive process as in $Y(t) = \phi_1 Y(t-1) + \phi_2 Y(t-2) + ...$ 24, 125, 126

PP Phillips Perron unit root test. The null hypothesis is that the tested time series follows an I(1) process, and the alternative hypothesis is that it follows an I(0) process. 36, 158–161, 165–167

S Sample standard deviation. 64, 154, 214, 215

s Seed value used for random number generation used in the **R** command set.seed(s). Unless otherwise noted, the seed used for random number generation is 12345. ix–xiv, 64, 66, 69–72, 75, 96–98, 100, 101, 104, 105, 107–109, 111–114, 116–118, 120, 122–124, 130, 132, 133, 144–149, 154

SVM Support Vector Machine. A machine learning classification tool that classifies data by finding separating hyperplanes between the classes while maximizing a margin between them. 155, 157–161, 165–169

Γ Test statistic of a statistical test. 87, 135–138, 150–154, 201–203, 205

TS NIST (2005) defines a time series (TS) as an ordered sequence of values of a variable at equally spaced time intervals. ix, xii–xiv, xvi, xvii, 1, 4, 6, 9, 11, 13–15, 17–20, 22, 24–26, 32, 38, 39, 41, 46, 49, 51, 52, 62–64, 66–68, 73, 77–82, 85, 90, 92, 95, 113, 122, 129, 137, 139, 141–154, 157, 163, 170, 174, 175, 189, 192, 196, 200–207, 209, 214, 219

URT Unit Root Test with a null hypothesis of $H_0^{I(1)}$ and an alternative hypothesis of $H_1^{I(0)}$. x–xii, xiv–xvi, 6, 31, 37, 52, 53, 61–63, 65, 69, 71, 73–78, 83, 86, 91–93, 95–101, 103–124, 127–129, 135, 139–141, 143, 155, 159–163, 165–168, 182, 193, 200–202, 204, 205, 207, 219–221

VAR Vector auto-regressive model. 5, 39, 40, 194

VECM Vector error correcting model implied by a multivariate cointegrating relationship. 39

ZA Zivot Andrews unit root test. The null hypothesis is that the tested time series follows an I(1) process, and the alternative hypothesis is that it follows an I(0) process. 36, 37, 52, 76–78, 80–82, 92, 93, 95, 98, 99, 102–106, 108–112, 115–122, 139, 140, 201, 202, 220

Abstract

NEW UNIT ROOT TESTS TO DECREASE SPURIOUS RESULTS WITH APPLICATIONS IN FINANCE AND TEMPERATURE ANOMALIES

Edward Herranz, PhD

George Mason University, 2016

Dissertation Director: Dr. James Gentle

Simulation studies show that when testing for cointegration with pairs of independent explosive($\phi_1 > 1$) AR(1) time series almost invariably lead to spurious cointegrating relationships. A new unit root test, the lagged-series test, is proposed with similar power to the ADF test for non-explosive AR(1) series but higher power in the explosive case. The lagged-series unit root test can be combined with other unit root tests such as the Elliot-Rothenberg-Stock tests and the Zivot-Andrews test, as well as the ADF test to improve the statistical power in the explosive case. A new unit root test, the Hybrid Bai-Perron Zivot-Andrews test, is proposed which allows for structural breaks in intercept and linear trend under the null hypothesis and compares favorably in some cases to the Lee-Stratizich unit root test. A new testing procedure to check for stationary to nonstationary shifts in a time series, referred to as the Hybrid Bai-Perron ADF procedure, is proposed and tested. It is shown that different unit root test related statistics can be combined using deep learning neural networks and results in techniques that outperform individual unit root tests in various simulation studies. Simulation based studies of the ADF, ERS-Ptest, ERS-DFGLS, the Zivot-Andrews, and the new lagged-series unit root tests, under various model configurations were made and compared.

Findings are consistent with various behaviors covered in the literature such as these tests are sensitive to the starting value of the AR(1) process, and when there are structural breaks these tests hardly ever reject the null hypothesis of a unit root. The covered interest rate parity formula is expressed as three linear terms and tested for cointegration; statistical evidence is provided showing when cross-currency swap basis spreads are added to one term the cointegration relationship always strengthens. Statistical evidence is provided that shows likely cointegration relationships between bank credit default swap spreads and cross currency basis swap spreads, indicating that bank credit risk is related to cross currency basis swap spreads; also statistical evidence is provided that there are cointegration relationships between bank credit default swap spreads and spot FX, indicating that bank credit risk affects the FX Spot rate. Statistical evidence that USD-JPY cross-currency basis swaps spreads Granger cause JPY fixed-floating interest rate swaps. A possible explanation may be USD based entities issuing JPY fixed debt and hedging it fully or partially with USD-JPY cross currency basis swaps. Various analyses to check for unit roots in zonal temperature anomaly time series were done. A number of models were fitted to these zonal temperature anomalies including a 3 regime SETAR model, and it is shown that better fits are always achieved when including linear trends. Estimated SETAR models for the Southern Hemisphere temperature anomalies are more likely to be stationary than the Northern Hemisphere, which includes an explosive AR(1) middle regime.

Chapter 1: Introduction: Unit Root Nonstationarity and Cointegration

This thesis is divided into sections describing related time series (TS) theory and explanatory examples, the new unit root tests, related financial and economic theory and and applications exploring relationships between forward foreign exchange rates, interest rate swap rates, cross currency basis swap rates and credit default swap rates, as well as an analysis of hemispheric temperature anomalies. NIST (2005) defines a TS as an ordered sequence of values of a variable at equally spaced time intervals. I will focus on discrete time models involving one or more univariate TS. These are mostly commonly auto-regressive AR(1) models, but other AR orders are considered. I use the R Core Team (2015) language and various add-on statistical packages packages, as well as implemented some new tests of my own, to conduct many Monte Carlo simulation experiments, as well as a number of empirical tests. The notation and definitions in this theory are based on Shumway,R.H. and Stoffer, D.S. (2000), Pfaff, B. (2008), Tsay, Ruey S. (2014), Hill, R. Carter and Griffiths, William E. and Lim, Guay C. (2011) and Stock, James H. and Watson, Mark W. (2011). I will provide more specific references at the relevant sections.

1.1 Types of Time Series Statistical Models

NIST (2005) defines a time series (TS) as an ordered sequence of values of a variable at equally spaced time intervals.

We define a TS $\{x_t\}$ as

$$\{x_t\} = \{x_1, x_2, ..., x_n\} \tag{1.1}$$

We will also use the equivalent notation:

$$\{x(t)\} = \{x(1), x(2), ..., x(n)\} \tag{1.2}$$

And we define the time-lag operator as

$$L^k(x_t) = x_{t-k} \tag{1.3}$$

Throughout this thesis we focus on discrete time models, although we make reference to continuous time models used to develop the theory of Brownian motion that is used to derive the asymptotic behavior of some of the test statistics used. As explained in Shumway,R.H. and Stoffer, D.S. (2000) there are two broad categories of time series models: those in the time domain, and those in the frequency domain. I will only discuss time domain models in this thesis. Within the time domain the most commonly used models are the auto-regressive models (AR), the moving average models (MA) and a combination of the two referred to as the auto-regressive moving-average (ARMA) models. An AR(p) model is defined in terms of its lagged values $x(t)$ and the innovations ϵ_t as :

$$x(t) = \sum_{i=1}^{p} \phi_i x(t-i) + \epsilon_t \tag{1.4}$$

It is commonly assumed that the innovations are Gaussian white noise series $\epsilon_t \sim N(0, \sigma^2)$; $\text{cor}(\epsilon_t, \epsilon_{t-1}) = 0$.

An MA(q) model is defined in terms of its current and past innovations ϵ_t as :

$$x(t) = \sum_{i=1}^{q} \theta_i \epsilon_{t-i} \tag{1.5}$$

An ARMA(p,q) model is defined in terms of its lagged values $x(t)$ and its current and past innovations ϵ_t as :

$$x(t) = \sum_{i=1}^{p} \phi_i x(t-i) + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} + \epsilon_t \tag{1.6}$$

We define the difference operator as follows:

$$\Delta(x(t)) = (x(t) - x(t-1)) \tag{1.7}$$

And differencing can be applied d times to a series:

$$\Delta^d(x(t)) = \Delta(...(\Delta(\Delta(x(t))))) \tag{1.8}$$

An ARIMA(p,d,q) model is a model of a time series $x(t)$ where we first difference the series d-times, resulting in $\Delta^d(x(t))$, and then we build an ARMA(p,q) model from the differenced series. Here the "I" stands for integrated.

Other common time series models in the time domain, are the auto-regressive-conditional-heteroskedasticity (ARCH) and general-ARCH (GARCH) models used to model processes with non-constant variances. GARCH models are very commonly used to model financial asset returns–as these exhibit heteroskedastic behavior. Shumway,R.H. and Stoffer, D.S. (2000, p. 286) define a GARCH(m,r) model as:

$$\begin{aligned} y_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= \alpha_0 + \sum_{j=1}^{m} \alpha_j y_{t-j}^2 + \sum_{j=1}^{r} \beta_j \sigma_{t-j}^2 \end{aligned} \tag{1.9}$$

where $\epsilon_t \sim$ i.i.d.$N(0, 1)$

The previous univariate ARMA and also GARCH models can be extended to to multivariate form; these are the VARMA and MGARCH models. All of these models can be considered sub-classes of a more general type of model know as the state-space model, also known as tge dynamic linear model(DLM.) Shumway,R.H. and Stoffer, D.S. (2000, p. 319) define a state-space model first by the **state equation** as a vector autoregression:

$$\boldsymbol{x_t} = \Phi \boldsymbol{x_{t-1}} + \boldsymbol{w_t} \tag{1.10}$$

where $\boldsymbol{x_t}$ is a $p \times 1$ state vector. $\boldsymbol{w_t}$ is a p i.i.d. zero-mean Gaussian vector and it is assumed that the model starts with a given initial vector $\boldsymbol{x_0}$ and has mean $\boldsymbol{\mu_0}$ and $p \times p$ covariance matrix Σ_0. The state-space model assumes that $\boldsymbol{x_t}$ is not directly observable and adds the observation equation:

$$\boldsymbol{y_t} = \boldsymbol{A_t x_t} + \boldsymbol{v_t} \tag{1.11}$$

where $\boldsymbol{A_t}$ is a $q \times p$ observation matrix and $\boldsymbol{v_t}$ is a $p \times 1$ white noise vector. In this thesis much of the consideration is for univariate time series models, however I also consider two-variate time series models in various cases.

Another type of TS models are nonparemtric models. The literature in this area keeps increasing, as in Zhang, Ting and Wu, Wei Biao (2015) where nonparametric time-varying time series model estimation is presented. The key idea behind nonparametric models is for them to be based on the data and be model-free, making as few assumptions as possible.

1.2 Linear Regression Model Estimation Via Least Squares

The following exposition of the estimation of linear regression models using least squares is based on Hastie, Trevor and Tibshirani, Robert and Friedman, Jerome (2001, p. 43-44). Consider the linear regression model expressed in matrix form:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon} \tag{1.12}$$

where $\boldsymbol{y}$ denotes an N-vector of the dependent variable, $\boldsymbol{X}$ denotes an $N \times (p+1)$ matrix of the independent variables. The +1 allows for constants in the model, by making the first element of each row to be 1. $\boldsymbol{\epsilon}$ is a N-vector of errors (or innovations.) $\boldsymbol{\beta}$ is an $(p+1)$-vector of coefficients to be estimated. Consider the residual sum of squares as a function of $\boldsymbol{\beta}$:

$$\mathrm{RSS}\left(\boldsymbol{\beta}\right) = \left(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}\right)^T \left(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}\right) \tag{1.13}$$

We differentiate this quadratic function with respect to the $\boldsymbol{\beta}$ parameter:

$$\frac{\partial \mathrm{RSS}}{\partial \boldsymbol{\beta}} = -2\boldsymbol{X}^T \left(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}\right) \tag{1.14}$$

And we differentiate the previous result one more time:

$$\frac{\partial^2 \mathrm{RSS}}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T} = 2\boldsymbol{X}^T\boldsymbol{X} \tag{1.15}$$

Given the second derivative is always positive, and if we assume that $\boldsymbol{X}$ if of full rank then the minimum of this parabolic function is obtained by setting the first derivative to zero and solving for $\boldsymbol{\beta}$:

$$\boldsymbol{X}^T \left(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}\right) = 0 \tag{1.16}$$

$$\hat{\boldsymbol{\beta}} = \left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1} \boldsymbol{X}^T\boldsymbol{y} \tag{1.17}$$

1.3 Conditional Least Squares for AR(p) Models

As presented by Levine, Michael (2012) an AR(p) model can be rewritten as a VAR(1) model:

$$y_t = \boldsymbol{Y_{t-1}}\boldsymbol{\phi} + \boldsymbol{\epsilon_t} \tag{1.18}$$

where $\boldsymbol{\phi} = \{\phi_1, ..., \phi_p\}^T$ and $\mathbf{Y_{t-1}} = \{y_{t-1}, ..., y_{t-p}\}$ and using least squares as presented before in Equation (1.17) the estimate for the auto-regressive coefficient vector is:

$$\hat{\boldsymbol{\phi}} = \left(\mathbf{Y}_{\boldsymbol{t-1}}^T \mathbf{Y}_{\boldsymbol{t-1}}\right)^{-1} \mathbf{Y}_{\boldsymbol{t-1}}^T \boldsymbol{y_t} \tag{1.19}$$

Equation 1.19 is commonly known as the conditional least squares coefficient estimate of an AR(p) model.

1.4 Stationary Time Series

A weakly stationary TS process is defined as having a constant mean, and an autocovariance function $\gamma(s,t)$ that depends on s and t only through their difference $|s-t|$. Throughout this document a stationary TS is intended to mean a weakly stationary TS. A strictly stationary TS is one where the joint probability distribution does not change with time. Each element x_t of a strictly stationary TS follows the same distribution; there is no standard test for strict stationarity as it is not possible to infer a distribution based on a single element. In any case, strict stationarity is too strict and rarely observed in practice. A non stationary TS can have a time varying mean as well as a time varying covariance function.

1.5 Unit Roots

Unit root tests (URTs) of a TS address the null hypothesis that the series is *unit root* nonstationary; that is, the hypothesis the series is an integrated order 1 process, I(1). The alternative hypothesis is that the TS is weakly stationary. A process that is integrated order n, I(n), is a process that needs to be differenced n times to become weakly stationary. A weakly stationary process is referred to as an I(0) process. This thesis develops new unit root tests (URTs) that can reduce false positives for near-integrated processes. These new unit root tests can also help reduce false positives when it comes to analyzing cointegration relationships between two or more TS, since if any of the variables tested for cointegration are I(0) then the test can often result in a spurious cointegration relationship.

1.5.1 Unit Root Test Hypotheses

We largely focus on AR(1) models, with intercept and linear trend. The ADF unit root test handles AR(p) models by rewriting them as an AR(1) model with additional components of differences; after essentially "taking out" these additional components in a helper regression it is then able to use the critical values of the Dickey-Fuller that were designed with an AR(1) model.

Null Unit Root Test Hypothesis

Unless otherwise noted, we define the null hypothesis as an AR(1) model $x(t) = \phi_1 x(t-1) + \epsilon_t$ where $\phi_1 = 1$. We use the symbol $H_0^{I(1)}$ to denote the null hypothesis. This corresponds to a nonstationary I(1) process. This is unusual in statistics where we would normally consider the multiplier to be 0 under the null as in a t-test for the coefficient of a linear regression model. The reason $\phi_1 = 1$ is picked as the null is because many time series in Finance and Economics are unit roots, so setting the null as I(1) helps us increase the confidence that if we reject the null it is not I(1), and we can choose the significance level to be as precise as needed.

We also consider the extended null hypothesis AR(1) model to deal with an intercept and a linear trend as follows, in a one-dimensional state-space model like representation:

$$\begin{aligned} x(t) &= \phi_1 x(t-1) + \epsilon_t \\ y(t) &= \beta_0 + \beta_1 t + x(t) \\ \phi_1 &= 1 \end{aligned} \tag{1.20}$$

In this thesis we also consider enhancing the unit root test null hypothesis to allow for structural breaks/changepoints in the deterministic coefficients β_0 and β_1.

If instead we had allowed the trend in the AR formula under the null we will see that this leads to quadratic time trends in the levels. Consider:

$$x(t) = \phi_1 x(t-1) + \beta_0 + \beta_1 t + \epsilon_t$$
$$\phi_1 = 1 \tag{1.21}$$

If we solve the recursion, using Faulhaber's formula we would find the following expression in the levels of $x(t)$ with a quadratic time trend component:

$$x(t) = \beta_0 t + \beta_1 \frac{t(t+1)}{2} + \sum_{i=1}^{t} \epsilon_i \tag{1.22}$$

Alternative Unit Root Test Hypothesis

Unless otherwise specified, we consider the alternative to be AR(1) with $\phi_1 \neq 1$. When $|\phi_1| < 1$ this corresponds to a stationary I(0) process. Most of the standard unit root tests, such as the ADF test consider the alternative hypothesis to be $|\phi_1| < 1$. However in the case with an explosive AR(1) process with $\phi_1 > 1$ this is a highly nonstationary process. We use the symbol $H_1^{I(0)}$ to denote the stationary possibility of the alternative hypothesis. And we use the symbol $H_1^{\text{Explosive AR}}$ to denote the stationary possibility of the alternative hypothesis. When testing for cointegration with AR(1) explosive series can quite easily lead to a spurious cointegration result as will be explained later. This is why I think it is beneficial to have unit root tests with power against the explosive alternative hypothesis.

When allowing an intercept and a linear trend in the time series being tested for unit roots we consider the same one-dimensional state-space model like representation as under the null:

$$x(t) = \phi_1 x(t-1) + \epsilon_t$$

$$y(t) = \beta_0 + \beta_1 t + x(t) \tag{1.23}$$

$$\phi_1 \neq 1$$

We also will consider the unit root test alternative hypothesis to allow for structural breaks/changepoints in the deterministic coefficients β_0 and β_1 later in this thesis.

1.6 Introductory Examples

1.6.1 A Spurious Regression Model of a Stock and a Foreign Exchange Rate

Consider Google stock prices and the US dollar/Brazil real exchange rate (expressed as number of reals per 1 dollar) for the period between 2014-03-27 and 2016-01-08. Stock prices and foreign exchange rates are always positive and can be modeled as stochastic processes. We take the logarithm of these TS as the models studied here do not have a restriction on the sign of the variable. Also a non-linear function can be made more linear by taking the logarithm; for example consider a parabolic function: $f(x) = x^2$; $\log(f(x)) = 2log(x)$. See Figure 1.1 for the TS of log Google prices. This figure displays a nonstationary pattern. The concept of stationarity of a stochastic process will be defined more rigorously in subsequent sections–it is related to the constancy of the mean, variance and the auto-covariance of the process.

The models we will study involve discrete time variables. However, a very common continuous time model of financial asset prices such as these is geometric Brownian motion (GBM):

$$S_t = S_0 \exp^{f(t)+\sigma W_t} \tag{1.24}$$

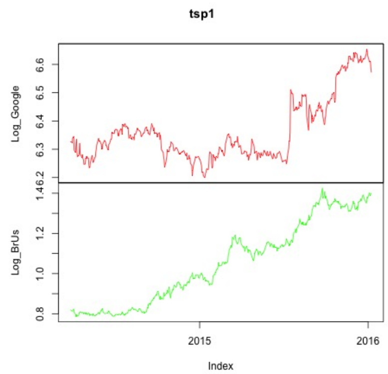


Figure 1.1: log(Google),log(Brazil/USD).

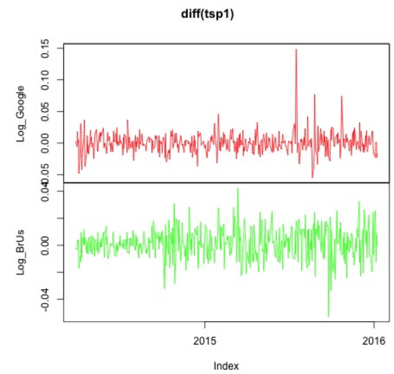


Figure 1.2: diff(log(Google)),diff(log(Brazil/USD)).

where S_0 is the initial asset price/exchange rate, $f(t)$ is a deterministic function of time and W_t represents Brownian motion. GBM is always positive however $\log(S_t) = log(S_0) + f(t) + \sigma W_t$ can fluctuate in sign. As detailed in Tsay, Ruey S. (2014, p. 268), Brownian motion (Wiener process) can be constructed as a re-scaled random walk. More specifically, a re-scaled sum of mean-zero, variance-one random variables ϵ_i as follows. Define $r \in [0, 1]$, the the limit of the sum converges in distribution to a Wiener process:

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nr \rfloor} \epsilon_i \Rightarrow_D W(r) \tag{1.25}$$

The Wiener process has the following properties:

- $W(0) = 0$
- $E(W(t)) = 0$
- $E(W(t)W(s)) = min(t, s)$ for $s, t \geq 0$
- Given $0 \leq t \leq s$, we have $W(s) - W(t)$ is independent of $W(u)$ where $u \leq t$ and has a Gaussian distribution of $N(0, s - t)$

If we difference the S_t process defined in Equation (1.24) as $\log(S_t) - \log(S_{t-1}) = f(t) - f(t-1) + \sigma(W_t - W_{t-1})$ we expect to obtain a deterministic part, which could be a constant, plus independent random increments, $\frac{\epsilon_i}{\sqrt{n}}$. And essentially this is what we see in Figure 1.2. This figure appears to be a process that is much more stationary than Figure 1.1. It seems plausible that the mean is close to zero in Figure 1.2. The variance does not appear to be constant, and that is indeed a common feature of most TS of financial asset values or rates. We will not be focusing on this aspect within this study. As a side note, we will see later that the test statistics for the Dickey-Fuller unit root test, as well as those for the Johansen cointegration test are based on functionals of Wiener processes(Brownian Motion.)

We estimate, using the lm() function of the R language, a simple linear model between the log TS based on the ordinary least squares (OLS) method:

$$\log(\text{Google}(t)) = c + \beta \log(\text{BrUs}(t)) + \epsilon_t \tag{1.26}$$

And derive the coefficient estimates $\hat{c}$ and $\hat{\beta}$. The results are detailed in Table 1.1. The coefficient estimates obtained are very significant, and even the R-squared of 0.419 is not small.

Unfortunately this regression is most likely a spurious one, that happens often with non-stationary variables such as these that follow random walks, or a related process referred to as a unit root. It would be hard to conceive that somehow Google's profits are closely linked to the US dollar/Brazil real exchange rate.

We can perform a statistical test of non-stationarity on each of the TS being analyzed; one of the most famous and commonly used such tests is the Augmented Dickey-Fuller (ADF) test, which will be discussed in more detail later; this is what is referred to in the literature as a unit-root test. The null hypothesis of this test is that the TS is non-stationary (has a unit root) and the alternative is that the TS is stationary. We see that for $\log(\text{Google}(t))$ we cannot reject the null hypothesis of a unit root/random walk, as well

Table 1.1: $\log(\text{Google}) = \hat{c} + \hat{\beta}\log(\text{BrUs})$

	Dependent variable:
	Log_Google
Log_BrUs	0.351***
	(0.020)
Constant	5.992***
	(0.021)
Observations	446
R^2	0.419
Adjusted R^2	0.418
Residual Std. Error	0.084 (df = 444)
F Statistic	320.153*** (df = 1; 444)
Note:	*p<0.1; **p<0.05; ***p<0.01

as for the log of the exchange rate:

Listing 1.1: Unit Root Tests of Log TS

```
> adf.test(tsp1$Log_Google)
        Augmented Dickey-Fuller Test
data:  tsp1$Log_Google
Dickey-Fuller = -1.6311, Lag order = 7, p-value = 0.7337
alternative hypothesis: stationary
 > adf.test(tsp1$Log_BrUs)
        Augmented Dickey-Fuller Test
data:  tsp1$Log_BrUs
Dickey-Fuller = -3.025, Lag order = 7, p-value = 0.1444
alternative hypothesis: stationary
```

So what can be done now? One of the most powerful and useful concepts in Econometrics/Time Series research is that of cointegration. It was originally proposed by Engle, R.

F. and Granger,C. W. J. (1987). This concept will be more precisely defined later, but we will describe it here as a number of non-stationary TS are cointegrated if a linear combination of the series results in a stationary process. There are various ways of testing for cointegration; perhaps the simplest is to test the residuals of the regression between the pair of TS using a unit root test to see if the null hypothesis of a unit root can be rejected or not. This method is only good for two-variable equations; in this example we will use the Johansen test of cointegration which can be used for any number of variables:

Listing 1.2: Johansen Cointegration Test of Log TS

```
######################
# Johansen-Procedure #
######################
Test type: maximal eigenvalue statistic (lambda max) , with linear trend
Eigenvalues (lambda):
[1] 0.0132568305 0.0007475134
Values of test statistic and critical values of test:
          test 10pct  5pct  1pct
r <= 1 | 0.33   6.50   8.18 11.65
r = 0  | 5.93  12.91  14.90 19.19
Eigenvectors, normalised to first column:
(These are the cointegration relations)
               Log_Google.l2 Log_BrUs.l2
Log_Google.l2      1.0000000    1.000000
Log_BrUs.l2       -0.4643095    1.137981
Weights W:
(This is the loading matrix)
               Log_Google.l2  Log_BrUs.l2
Log_Google.d   -0.020232932 0.0004519718
Log_BrUs.d      0.005646056 0.0009019347
```

The Johansen test has multiple related hypothesis tests. The first null-hypothesis of the Johansen cointegration test is that the series are not cointegrated. This can be seen in the line $r = 0$. The test statistic is 5.93 in this case and we cannot reject the null hypothesis. The second test has the null hypothesis that the cointegration order is less than or equal to

1 which we can see in line $r \leq 1$ and we cannot reject this one either. So in short this sample Johansen test provides statistical evidence that log(Google(t)) and log(Brazil/USD(t)) are not cointegrated; this implies there is no reason to think these TS are closely related.

In this example we have seen that using linear regression with variables that follow random walks(unit roots) yielded a spurious relationship. Then we saw that if we checked for cointegration there was no real evidence of a significant relationship between the analyzed variables.

Table 1.2: diff(log(Google)) diff(log(BrUs))

	Dependent variable:
	Diff_Log_Google
Diff_Log_BrUs	−0.096
	(0.069)
Constant	0.001
	(0.001)
Observations	445
R^2	0.004
Adjusted R^2	0.002
Residual Std. Error	0.016 (df = 443)
F Statistic	1.953 (df = 1; 443)
Note:	*p<0.1; **p<0.05; ***p<0.01

Is there a way to consistently use linear regression while avoiding the pitfalls involving random walks? One approach is to run a regression on the differenced TS; we can see that a differenced random walk results in a TS of the random increments: Given $x(t) = \sum_{i=0}^{t} \epsilon_i$ where $\mathrm{E}(\epsilon_i) = 0, \mathrm{Var}(\epsilon_i) = a$, then for the differences we obtain: $x(t) - x(t-1) = \epsilon_t$. For our previous example we regress the first differences of the logs of the variables, and the summary of the regression is in Table 1.2. We see now that the fitted coefficients for this regression are no longer statistically significant, and the R-squared is practically zero.

Maddala, G.S. and Kim, In-Moo (1998, p. 33) survey literature on the consequences of misspecification of models. They consider processes that are stationary around a deterministic trend (TSP) and processes that are stationary after first differences (DSP) and discuss how to distinguish between them. If the TS is DSP but is treated as TSP this is a problem of under-differencing. If the TS is TSP but is treated as DSP this is a case of over-differencing. They refer to literature providing cases where misleading results can be obtained, but also point to solutions such as adjusting for the serial correlations of errors.

1.6.2 Differencing TS Can Hide Relationships

Does the previous example suggest that if we always regress differenced TS when we think we are dealing with random walks we avoid all the potential pitfalls? Unfortunately one can miss out on relationships when we difference. Consider the following DGP (1.27):

$$\begin{aligned}
x(t) &= x(t-1) + \epsilon_1(t) \ ; \ \epsilon_1(t) \sim N(0,1) \\
y(t) &= x(t-\alpha) + \epsilon_2(t) \ ; \ \epsilon_2(t) \sim N(0,1) \\
&\alpha \text{ one of } \{1,2,3,4,5\} \text{ with equal probability} \\
x(0) = y(0) = y(1) = y(2) &= y(3) = y(4) = y(5) = 0 \\
\operatorname{cor}(\epsilon_1(t), \epsilon_2(t)) &= 0
\end{aligned} \tag{1.27}$$

First notice that $x(t)$ is a random walk, and that $y(t)$ is just a random lagged $(t-\alpha)$ value of $x(t)$ plus an innovation term which is not cumulated over time steps. So $y(t)$ is also a random walk with additional noise. So we can see that these two TS are closely related. Since they are both random walks we should be wary of deriving conclusions from a linear regression analysis of the level TS. However in this case if we run a regression on the differenced TS it will not show any significant relationship. Let us try with a simulation.

The data for this test can be replicated using the R language as follows:

Listing 1.3: Random Lag Simulation

```
library(zoo)
set.seed( 12345 )
l<-1000
x <- cumsum(rnorm(l))
y <- rep(0, l)
for( i in 6:l  )
{
        lag_ <- sample(1:5,1)
        y[i] <- x[i-lag_] + rnorm(1)
}
x <- zoo(x)
y <- zoo(y)
```

First we find statistical evidence that $x(t)$ and $y(t)$ are nonstationary:

Listing 1.4: ADF Test of Randomly Lagged Series

```
> adf.test(x)
Dickey-Fuller = -1.8271, Lag order = 9, p-value = 0.6515
alternative hypothesis: stationary
> adf.test(y)
Dickey-Fuller = -1.7453, Lag order = 9, p-value = 0.6861
alternative hypothesis: stationary
```

Now we estimate the following linear regression model:

$$\Delta(x(t)) = \hat{\beta}_0 + \hat{\beta}_1 \Delta(y(t)) + \epsilon_t \tag{1.28}$$

And from the regression results summarized in Table 1.3 we see that there are no statistically significant coefficients and the R-squared is zero. So that could lead us to the incorrect conclusion that these two TS are unrelated.

Table 1.3: $\Delta x = \beta_0 + \beta_1 \Delta y$

	Dependent variable:
	Δx
Δy	-0.00002
	(0.017)
Constant	0.046
	(0.032)
Observations	999
R^2	0.000
Adjusted R^2	-0.001
Residual Std. Error	1.000 (df = 997)
F Statistic	0.00000 (df = 1; 997)
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

As in the previous example given we believe our individual variables are random walks(unit roots) we should perform a test to see if they are cointegrated, and we will use the Johansen test again:

Listing 1.5: Johansen Test of x and y

```
> summary( ca.jo( tsx ) )
######################
# Johansen-Procedure #
######################
Test type: maximal eigenvalue statistic (lambda max) , with linear trend
Eigenvalues (lambda):
[1] 0.332242393 0.005952608
Values of teststatistic and critical values of test:
          test 10pct  5pct  1pct
r <= 1 |   5.96  6.50  8.18 11.65
r = 0  | 403.02 12.91 14.90 19.19
Eigenvectors, normalised to first column:
(These are the cointegration relations)
           x.l2       y.l2
x.l2  1.0000000  1.0000000
y.l2 -0.9967988 -0.2076108
Weights W:
(This is the loading matrix)
           x.l2          y.l2
x.d -0.007177517 -0.0057276936
y.d  0.637051425 -0.0002333098
```

In the Test Result Listing 1.5 we can see that the test statistic for the case of no cointegration ($r = 0$) is 403.02 so we can easily reject the null hypothesis of no cointegration. We cannot reject the second null hypothesis of a cointegration of order 1 ($r \leq 1$). From these results we conclude that x and y are likely cointegrated.

1.6.3 Two Cases of Spurious Cointegration with Near Unit Roots

The previous example showed us that if we are building a model involving TS that appear to be random walks, we should always perform a cointegration test, instead of relying on building a linear regression model based on the differenced TS. Cointegration tests become problematic if they are performed on TS which are not random walks, i.e. are stationary, as they will indicate that the series are cointegrated. This is particularly troublesome when the TS we analyze are close to random walks, and the unit root tests will fail to reject the null hypothesis of nonstationarity. First we will define what is meant by near unit roots.

A simple autoregressive model involving a single lag, AR(1) with standard normal innovations is detailed in Equation (1.29):

$$x_t = \phi_1 x_{t-1} + \epsilon_t \ ; \ \epsilon_t \sim N(0,1) \tag{1.29}$$

The simple unit root case (random walk) consists of when $\phi_1 = 1$ and given this multiplier value we can see the recursive definition would collapse to $x_t = \sum_{i=1}^{n} \epsilon_i$ which is the standard model of a random walk. A near unit root is the case where $\phi_1 <> 1$ but it is close to 1. We will look at two practical examples one where it is less than 1, and one where phi_1 is larger than one, referred to as the explosive case.

A case with an AR(1) Processes with $\phi_1 < 1$

The DGP used for this example is that in Equation (1.29) with $phi_1 = 0.94$ and the data for this test can be replicated using the R language as follows:

Listing 1.6: Two Near Unit Root Series, ts1 and ts2

```
library(zoo)
set.seed(12345)
l <- 200
ts1 <- rep(0,l)
ts2 <- rep(0,l)
ts1[1] <- rnorm(1)
ts2[1] <- rnorm(1)
for(i in 2:l)
{
        ts1[i] <- 0.94 * ts1[i-1] + rnorm(1)
        ts2[i] <- 0.94 * ts2[i-1] + rnorm(1)
}
ts1 <- zoo( ts1)
ts2 <- zoo( ts2)
```

The Augmented Dickey Fuller unit root tests provide statistical evidence that the ts1 and ts2 TS are nonstationary as can be seen in the following results:

Listing 1.7: ADF Tests of ts1 and ts2

```
> adf.test( ts1 )
Dickey-Fuller = -2.9404, Lag order = 5, p-value = 0.1822
alternative hypothesis: stationary
> adf.test( ts2 )
Dickey-Fuller = -2.6176, Lag order = 5, p-value = 0.3174
alternative hypothesis: stationary
```

Now we perform a cointegration test of ts1 and ts2 as detailed in Test Results Listing 1.8:

Listing 1.8: Johansen Test for near unit roots ts1 and ts2

```
######################
# Johansen-Procedure #
######################

Test type: maximal eigenvalue statistic (lambda max) , with linear trend
Eigenvalues (lambda):
[1] 0.07508604 0.03716230
Values of teststatistic and critical values of test:
          test 10pct  5pct  1pct
r <= 1 |  7.50  6.50  8.18 11.65
r = 0  | 15.45 12.91 14.90 19.19
Eigenvectors, normalised to first column:
(These are the cointegration relations)
            ts1.l2    ts2.l2
ts1.l2   1.0000000 1.000000
ts2.l2  -0.3582286 2.523358
Weights W:
(This is the loading matrix)
            ts1.l2        ts2.l2
ts1.d  -0.13853693  -0.006762077
ts2.d   0.04118509  -0.028669928
```

As can be seen in Listing 1.8 the first null hypothesis of no cointegration (a cointegration rank of $r = 0$) can be rejected at the 5% significance level, and the second null hypothesis of a co-integration rank of 1 cannot be rejected with the same significance level. We could conclude that ts1 and ts2 based on this test are cointegrated. The problem here lies in the unit root tests did not reject the null hypothesis of the unit root as they should.

A case with an Explosive Processes (AR(1) with $\phi_1 > 1$)

The DGP used for this example is that in Equation (1.29) with $phi_1 = 1.01$ and the data for this test can be replicated using the R language using the same code as in Listing 1.6 but replacing the 0.94 with 1.01.

Just as before we test the underlying TS ts1 and ts2 for unit roots, and the null hypothesis of nonstationarity cannot be rejected:

Listing 1.9: ADF Tests for Explosive ts1 and ts2

```
> adf.test( ts1 )
Dickey−Fuller = 0.67457, Lag order = 5, p−value = 0.99
alternative hypothesis: stationary
> adf.test( ts2 )
Dickey−Fuller = 0.30167, Lag order = 5, p−value = 0.99
alternative hypothesis: stationary
```

Now we perform a cointegration test of explosive series ts1 and ts2 as detailed in Test Results Listing 1.10:

Listing 1.10: Johansen Test for ts1 and ts2

```
######################
# Johansen-Procedure #
######################
Test type: maximal eigenvalue statistic (lambda max) , with linear trend
Eigenvalues (lambda):
[1] 0.08025426 0.03045790
Values of teststatistic and critical values of test:
          test 10pct  5pct  1pct
r <= 1 |  6.12  6.50  8.18 11.65
r = 0  | 16.56 12.91 14.90 19.19
Eigenvectors, normalised to first column:
(These are the cointegration relations)
          ts1.l2     ts2.l2
ts1.l2  1.000000   1.000000
ts2.l2 -9.027077  -1.353337
Weights W:
(This is the loading matrix)
             ts1.l2       ts2.l2
ts1.d -0.001868595  -0.01847405
ts2.d -0.001105020   0.03229835
```

In Listing 1.8 the first null hypothesis of no cointegration (a cointegration rank of $r = 0$) can be rejected at the 5% significance level, and the second null hypothesis of a co-integration rank of 1 cannot be rejected with the same significance level. We could conclude that ts1 and ts2 based on this test are cointegrated, just as we did on the previous test case with $\phi_1 = 0.94$. The problem again is that the unit root tests did not reject the null hypothesis.

A contribution of this thesis is a new unit root test (the lagged-series unit root test) that can be used alone or in conjunction with other unit root tests to reduce the chance of spurious cointegration, mainly in the explosive case.

1.7 Integrated Time Series

A standard way to convert a non stationary TS into a stationary one is via successive differencing; a TS is integrated order d, labeled I(d), if it has to be differenced d times to become stationary. This is the definition of integration order of a process given by Engle, R. F. and Granger,C. W. J. (1987) in their seminal paper.

The simplest case of an I(1) TS can be described using an auto-regressive process AR(1): $X(t) = \phi_1 X(t-1) + \epsilon$ where $\phi_1 = 1$ and ϵ is a zero-mean random innovation. This process is a random walk with a 0 deterministic drift. We see that this process can be re-written as $X(t) = X(0) + \sum_{i=1}^{n} \epsilon_i$. It can be described as having a long memory, as all innovations ϵ_i are equally weighted. The summation of innovations is commonly referred to as a stochastic trend. If $\phi_1 < 1$ then older values would be increasingly weighted less and less.

Consider an auto regressive process of order p:

$$x(t) = \phi_1 x(t-1) + \phi_2 x(t-2) + ... + \phi_p x(t-p) + \epsilon(t) \tag{1.30}$$

The characteristic polynomial of this AR(p) process is defined as

$$\phi(z) = 1 - \phi(1)z - ... - \phi(p)z^p \tag{1.31}$$

A unit root is defined as an auto-regressive process that has 1 as a valid root of the characteristic polynomial equation as in Chan, Ngai Hang (2010, p. 29). TS with unit roots are non-stationary processes. In the case of an AR(1) process if $abs(\phi_1) = 1$ there will be a unit root. As detailed in Chan, Ngai Hang (2010) if AR(p) process has all of its characteristic polynomial roots with an absolute value greater than one, then such as process is defined to be causal, and will also be stationary. It is possible for an AR(1) process with $|\phi_1| > 1$ to be re-written in a form that would imply it is stationary, even though it is non-convergent however it requires expressing the model in terms of future values which would make it an unusable model. For example for an AR(1) model as detailed in Chan, Ngai Hang (2010, p. 27):

$$x(t) = \frac{x(t+1)}{\phi_1} - \frac{\epsilon(t+1)}{\phi_1} \tag{1.32}$$

There are various statistical tests to check for the presence of unit roots. In this paper we review the Augmented Dickey Fuller test, the Elliot, Rothenberg and Stock tests, and propose new tests.

1.8 Spurious Regressions

Finding linear relationships involving TS that are integrated order 1, $I(1)$, or higher can result in "spurious" or "nonsense" regressions. As Pfaff, B. (2008, p. 74) points out, a spurious regression can have a high unadjusted $R^2 = 1 - \frac{\sum_{i=1}^{n} \hat{\epsilon_i}^2}{\sum_{i=1}^{n}(y(i)-\bar{y})^2}$ The denominator in the second term of the formula becomes very large since extreme values on both sides of the non-constant mean are weighted heavily due to the nonstationarity of the process, so in the limiting case $R^2 \to 1$.

1.8.1 Linear Relationships Between Time Series

Many financial TS are random walks, which are non-stationary processes. Consider the simple Gaussian random walk model expressed as an auto regressive process:

$$X(t) = X(t-1) + \epsilon_t \text{ where } X(0) = 0 \text{ and } \epsilon_t \sim N(0,1) \tag{1.33}$$

This discrete model can be rewritten as follows:

$$X(t) = \sum_{i=1}^{t} \epsilon_i, \text{ for some } t > 0 \tag{1.34}$$

So it is easy to see that the variance of this process is ever increasing as it is the sum of the variances; in this case it would be t.

We consider the following trivial models of linearly related $X(t)$ and $(Y(t)$ series:

$$
\begin{aligned}
X(t) &= gX(t-1) + \epsilon_t \\
X(0) &= 0 \\
\epsilon_t &\sim N(0,1) \\
Y(t) &= hX(t) + \gamma_t \\
\gamma_t &\sim N(0,1) \\
\text{cov}(\epsilon_t, \gamma_t) &= 0
\end{aligned}
\tag{1.35}
$$

We perform simulations where we pick the multiplier g and then we select at random h from a uniform distribution $[-10, 10]$ and derive the $\{X(t)\}$ and $\{Y(t)\}$ TS. The we regress $Y(t) = \beta X(t) + \nu_t$. The results of the regression (R^2 and F-stat p-value) are summarized in Table 1.4. In this case we can see that all regressions are significant as the R^2 are close to 100%, and all of the F-stat p-values are close to zero, rejecting the null-hypothesis of $\beta = 0$. These results are expected given the DGP (1.35) used to generate X and Y.

Table 1.4: Regressing $Y(t)$ on $X(t)$ as Defined in Equation (1.35) with $l = 1000$ and $m = 1000$

	Adj R^2	Fstat p-value
g	Q:0.5	Q:0.50
1.01	1.0000	0.00
1.00	0.9998	0.00
0.90	0.9922	0.00
0.50	0.9714	0.00
0.20	0.9638	0.00
0.00	0.9624	0.00

Now we consider the following trivial models of linearly independent $X(t)$ and $Y(t)$

series:

$$
\begin{aligned}
X(t) &= gX(t-1) + \epsilon_t \\
X(0) = 0 \; &; \; \epsilon_t \sim N(0,1) \\
Y(t) &= hY(t) + \gamma_t \\
\gamma_t \sim N(0,1) \; &; \; \mathrm{cov}(\epsilon_t, \gamma_t) = 0
\end{aligned}
\tag{1.36}
$$

Table 1.5: Regressing $Y(t)$ on $X(t)$ as Defined in Equation (1.36) with $l = 1000$ and $m = 1000$

g	h	Adj R^2 Q:0.5	Fstat p-value Q:0.5
1.01	1.01	1.0000	0.00
1.00	1.00	0.4045	0.00
0.90	0.90	0.034	0.04
0.00	1.00	-0.006	0.50
0.50	1.00	0.003	0.25
0.50	0.50	-0.002	0.38
0.20	0.20	-0.005	0.47
0.00	0.00	-0.006	0.50

Just as before we regress $Y(t) = \beta X(t) + \nu_t$ with the results of the DGP (1.36) in a simulation. Table 1.5 summarizes the regression results, and we see that in the cases of $g = 1.01, h = 1.01$ and $g = 1, h = 1$ the regressions appear to be significant even though $Y(t)$ and $X(t)$ are independent of each other. Even in the case of $g = 0.9, h = 0.9$ with an F-stat p-value 0.04 leads us to reject the null hypothesis of $\beta = 0$ at the 0.05 significance level. Granger, C. W. J. (2000, p. 13) confirm similar findings: that spurious regressions can happen for non I(1) processes (with an AR(1) ϕ_1 multiplier of 0.9.)

1.9 Dickey Fuller Unit Root Test

The original simplest version of the Dickey Fuller test considered the standard AR(1) model with Gaussian white noise, see Dickey, D. A. and Fuller, W.A. (1981):

$$y_t = \phi_1 y_{t-1} + \epsilon_t \ ; \ \epsilon_t \sim N(0,1) \ ; \ \mathrm{cor}(\epsilon_t, \epsilon_{t-1}) = 0 \ ; \ y_0 = 0 \tag{1.37}$$

As detailed in Patterson, Kerry (2010, p. 208) the two main test statistics for the Dickey Fuller test are the normalized bias $\hat{\delta}$ and the t-type statistic $\hat{\tau}$:

$$\hat{\delta} = T(\hat{\phi}_1 - 1) \tag{1.38}$$

$$\hat{\tau} = \frac{(\hat{\phi}_1 - 1)}{\hat{\sigma}_{\hat{\phi}}} \tag{1.39}$$

$\hat{\phi}_1$ is the standard coefficient estimate as derived using ordinary least squares:

$$\hat{\phi}_1 = \frac{\sum_{t=1}^{T} y_t y_{t-1}}{\sum_{t=1}^{T} y_{t-1}^2} \tag{1.40}$$

$$\hat{\phi}_1 - 1 = \frac{\sum_{t=1}^{T} y_{t-1}(y_t - y_{t-1})}{\sum_{t=1}^{T} y_{t-1}^2} = \frac{\sum_{t=1}^{T} y_{t-1}\epsilon_t}{\sum_{t=1}^{T} y_{t-1}^2} \tag{1.41}$$

The last step in Equation (1.41) is based on $\epsilon_t = y_t - y_{t-1}$.

Given these previous equations we can now define the normalized bias as a function of the observations y_t:

$$\hat{\delta} = T\frac{\sum_{t=1}^{T} y_{t-1}\epsilon_t}{\sum_{t=1}^{T} y_{t-1}^2} = \frac{\sum_{t=1}^{T} y_{t-1}\epsilon_t / T}{\sum_{t=1}^{T} y_{t-1}^2 / T^2} \tag{1.42}$$

As detailed in Table 7.1 of Patterson, Kerry (2010, p. 197) for the DGP (1.37) following sample quantities have a limiting form expressed as functionals of Brownian motion:

$$\sum_{t=1}^{T} y_{t-1}\epsilon_t/T \underset{D}{\Rightarrow} \sigma^2 \int_0^1 W_t dW_t \tag{1.43}$$

$$\sum_{t=1}^{T} y_{t-1}^2/T^2 \underset{D}{\Rightarrow} \sigma^2 \int_0^1 W_t^2 dr \tag{1.44}$$

These quantities can be simulated to determine critical values for these test statistics.

The Dickey-Fuller test, as all Unit Root tests consider, has a null hypothesis of $\phi_1 = 1$ and the alternative hypothesis is $|\phi_1| < 1$. The case of a negative unit root $\phi_1 = -1$ leads to a distribution of $\left(\hat{\phi}_1 + 1\right)$ that is the mirror image of the standard Dickey-Fuller distribution as detailed in Choi, In (2010, p. 55). When $\phi_1 > 1$ this leads to an explosive case. As derived in Phillips,Peter C.B. and Magadalinos,Tassos (2005) the asymptotics of the estimation error of ϕ_1 is this case under certain assumptions is Cauchy distributed.

We can extend the DGP (1.37) to have a constant and a linear trend. Then using OLS we can removed these extra terms. This will affect the critical values of the test statistics. Zivot, Eric (2005) suggests a simulation approach based in R. This approach is extended for the trend and drift terms here for the normalized bias (*NB) and t-type (*DF) test statistics; the *R* code to perform these simulations is detailed in Listing 1.11.

Listing 1.11: **NB and DF Test Statistic Simulations**

```
library(fUnitRoots)
library(dynlm)
library(zoo)

wiener_bar <- function(nobs)
{
        e <- rnorm(nobs)
        drift <- runif(1,-20,20)
```

```
        #Model is y_t = C + D * t + phi_1 * y_(t-1) + epsilon
        #H0: we assume that phi_1 = 1 and D=0
        y <- cumsum(e) + cumsum( rep(drift,nobs) )
        tsy <- zoo(y)
        l <- length( tsy[,1] )
        u1 <- 1:l
        model1 <- dynlm(formula = tsy~ u1    )
        D <- model1$coefficients['u1']
        C <- model1$coefficients['(Intercept)']
        #demean and detrend
        y <- y - rep(C,nobs) - D*u1
        ym1 <- y[1:(nobs-1)]
        intW2 <- nobs^(-2) * sum(ym1^2)
        intWdW <- nobs^(-1) * sum(ym1*e[2:nobs])
        ans <- list(intW2=intW2,intWdW=intWdW)
        ans
}
 set.seed(12345)
 nobs <- 1000
 nsim <- 10000
 NB <- rep(0,nsim)
 DF <- rep(0,nsim)
 for (i in 1:nsim)
 {
                BN.moments <- wiener_bar(nobs)
                NB[i]   <- BN.moments$intWdW/BN.moments$intW2
                DF[i] <- BN.moments$intWdW/sqrt(BN.moments$intW2)
}
quantile(DF, probs=c(0.01,0.05,0.1))
qunitroot(c(0.01,0.05,0.10), trend='ct', statistic='t', N=10000)
quantile(NB, probs=c(0.01,0.05,0.1))
qunitroot(c(0.01,0.05,0.10), trend="ct", statistic="n", N=10000)
```

And results of the simulation code from Listing 1.11 and comparison with the critical values to those from the R qunitroot() function from the fUnitRoots R package as reported by MacKinnon, J.G. (1996) is detailed in Results Listing 1.12:

Listing 1.12: Critical Values of Simulated DF Statistics

```
> quantile(DF, probs=c(0.01,0.05,0.1))
        1%           5%          10%
-3.980987  -3.406626  -3.125696
> qunitroot(c(0.01,0.05,0.10), trend='ct', statistic='t', N=10000)
[1]  -3.958799  -3.410295  -3.126858
> quantile(NB, probs=c(0.01,0.05,0.1))
        1%           5%          10%
-29.29784  -21.59540  -18.13030
> qunitroot(c(0.01,0.05,0.10), trend="ct", statistic="n", N=10000)
[1]  -29.32316  -21.68862  -18.23051
```

We see in the Simulation Results Listing 1.12 that the simulated values and the ones computed by MacKinnon, J.G. (1996) as reported in the qunitroot() function of the fUnitRoots R package are close.f

Said, S. E. and Dickey, D. A. (1984) extended the Dickey Fuller URT for ARMA models and not just AR models; this is known as the Augmented Dickey Fuller(ADF) unit root tests and is one of the most commonly used URTs in the literature (Choi, In, 2010, p. 33). The ADF test regression is fitted using OLS:

$$\triangle y_t = \alpha + \delta t + \beta y_{t-1} + \sum_{i=1}^{n} \gamma_i \triangle y_{t-i} + \epsilon_t \tag{1.45}$$

where $\triangle$ is the difference operator and ϵ_t represent 0-mean white-noise innovations. Under the null hypothesis y_t is considered to be I(1) which is equivalent to $\triangle y_t$ being I(0) in which case β would be zero. The test statistic is the standard regression t-statistic $t_\beta = \frac{\hat{\beta}}{s.e.(\hat{\beta})}$. A normalized bias test statistic is used as well.

The tests' authors derived the critical values for these test statistics as they follow a non-standard distribution; they are the same as the standard Dickey-Fuller critical values which depend on the form of the deterministic components. The lagged differences allow correcting for serial correlation in the innovations. For the details see Said, S. E. and Dickey, D. A. (1984).

1.10 Elliott, Rothenberg and Stock Unit Root Tests

See Pfaff, B. (2008, p. 93) for a detailed overview of the Elliott, Rothenberg and Stock (ERS) unit root tests; the following overview is taken from there. To increase the power of a unit root test under the null hypothesis of a unit root, Elliott, Rothenberg and Stock [1996] proposed a local to unity detrending of the TS. The assumed generating process the series y_t is as follows:

$$y_t = d_t + u_t \qquad u_t = au_{t-1} + v_t \qquad d_t = \hat{\beta}' z_t \tag{1.46}$$

where z_t is a deterministic $(q \times 1)$ vector and ν_t is a stationary zero-mean process. If $a = 1$ the y_t is I(1), but if $|a| < 1$ then y_t is I(0).

The authors developed feasible point-optimal tests, which take serial correlation of the error term into account. The feasible point-optimal test statistic is defined as:

$$P_T = \frac{S(a = \bar{a}) - \bar{a}S(a = 1)}{\hat{\omega}^2} \tag{1.47}$$

where $S(a = \bar{a})$ and $S(a = 1)$ are the sums of squared residuals from a least-squares regression of y_a on Z_a with $y_a = (y_1, y_2 - ay_1, ..., y_T - ay_{T-1})$ and $\boldsymbol{Z}_a = (z_1, z_2 - az_1, ..., z_T - az_{T-1})$ and y_a is a T-dimensional vector and $\boldsymbol{Z_a}$ is a $(T \times q)$ matrix. The estimator for the variance of the error process ν_t is:

$$\hat{\omega} = \frac{\hat{\sigma_\nu^2}}{(1 - \sum_{i=1}^{p} \hat{\alpha}_i)^2} \tag{1.48}$$

where $\hat{\sigma_\nu^2}$ and $\hat{\alpha_i}$ for $i = 1, ..., p$ are from the OLS regression:

$$\Delta y_t = \alpha_0 + \alpha_1 \Delta y_{t-1} + ... + \alpha_{p+1} \Delta y_{t-p} + \nu_t \tag{1.49}$$

And $\bar{a} = 1 + \bar{c}/T$ where $\bar{c}$ is a constant, set to -7 in the case of a constant or -13.5 in the case of a linear trend. The second test type is denoted as the DF-GLS test, which is a modified ADF-type test applied to the detrended data without the intercept, which is the t-statistic for testing $\alpha_0 = 0$ in the regression:

$$\Delta y_t^d = \alpha_0 y_{t-1}^d + \alpha_1 \Delta y_{t-1}^d + ... + \alpha_p \Delta y_{t-p}^d + \epsilon_t \tag{1.50}$$

where y_t^d are residuals from $y_t - \hat{\beta}' z_t$

Both of these tests are available in the `urca` package in the ur.ers() function.

1.10.1 Phillips Perron Unit Root Test

Phillips, Peter C. B. and Perron, Pierre (1988) developed a unit root test that corrects for serial correlation and heteroskedasticity of the innovations. Pfaff, B. (2008, p. 95) reviews the Phillips-Perron (PP) procedure details. The PP test regression is fitted using OLS. The first case is with out a linear-trend:

$$y_t = \mu + \alpha y_{t-1} + \epsilon_t \tag{1.51}$$

$$Z(\hat{\alpha}) = T(\hat{\alpha} - 1) - \frac{\hat{\lambda}}{\bar{m}_{yy}} \tag{1.52}$$

$$Z(\tau_{\hat{\alpha}}) = \frac{\hat{s}t_{\hat{\alpha}}}{\hat{\sigma}_{Tl}} - \frac{\hat{\lambda}' \hat{\sigma}_{Tl}}{\bar{m}_{yy}^{\frac{1}{2}}} \tag{1.53}$$

$$Z(\tau_{\hat{\mu}}) = \frac{\hat{s}t_{\hat{\mu}}}{\hat{\sigma}_{Tl}} - \frac{\hat{\lambda}'\hat{\sigma}_{Tl}m_y}{\bar{m}_{yy}^{\frac{1}{2}}m_{yy}^{\frac{1}{2}}} \tag{1.54}$$

with $\bar{m}_{yy} = T^{-2}\sum(y_t - \bar{y})^2$, $m_{yy} = T^{-2}\sum(y_t^2)$, $m_y = T^{-3/2}\sum(y_t)$, and $\hat{\lambda} = 0.5(\hat{\sigma}_{Tl}^2 - \hat{s}^2)$, where $\hat{s}^2$ is the sample variance of the residuals, $\hat{\lambda}' = \hat{\lambda}/\hat{\sigma}_{Tl}^2$, and $t_{\hat{\alpha}}$, $t_{\hat{\mu}}$ are the student t ratios of $\hat{\alpha}$ $\hat{\mu}$ respectively. The long run variance $\hat{\sigma}_{Tl}^2$ is estimated as follows:

$$\hat{\sigma}_{Tl}^2 = T^{-1}\sum_{t=1}^{T}\hat{\epsilon}_t^2 + 2T^{-1}\sum_{s=1}^{l}w_{sl}\sum_{t=s+1}^{T}\hat{\epsilon}_t\hat{\epsilon}_{t-s} \tag{1.55}$$

where $w_{sl} = 1 - s/(l+1)$

The case of adding a linear trend is now considered:

$$y_t = \mu + \beta(t - \frac{T}{2}) + \alpha y_{t-1} + \epsilon_t \tag{1.56}$$

In this case the test statistics used are:

$$Z(\hat{\alpha}) = T(\hat{\alpha} - 1) - \frac{\hat{\lambda}}{M} \tag{1.57}$$

$$Z(\tau_{\hat{\alpha}}) = \frac{\hat{s}t_{\hat{\alpha}}}{\hat{\sigma}_{Tl}} - \frac{\hat{\lambda}'\hat{\sigma}_{Tl}}{M^{\frac{1}{2}}} \tag{1.58}$$

$$Z(\tau_{\hat{\mu}}) = \frac{\hat{s}t_{\hat{\mu}}}{\hat{\sigma}Tl} - \frac{\hat{\lambda}'\hat{\sigma}_{Tl}m_y}{M^{\frac{1}{2}}(M + m_y^2)^{\frac{1}{2}}} \tag{1.59}$$

$$Z(\tau_{\hat{\beta}}) = \frac{\hat{s}t_{\hat{\beta}}}{\hat{\sigma}Tl} - \frac{\hat{\lambda}'\hat{\sigma}_{Tl}(0.5m_y - m_{ty})}{(\frac{M}{12})^{\frac{1}{2}}m_{yy}^{\frac{1}{2}}} \tag{1.60}$$

where m_y, $\bar{m}_{yy}$, $\hat{\lambda}$, $\hat{\lambda}'$ and $\hat{\sigma}_{Tl}$ are as defined previously and $m_{ty} = T^{-5/2}\sum ty_t$ and $t_{\hat{\mu}}$,$t_{\hat{\beta}}$,$t_{\hat{\alpha}}$ are the student t ratios of $\hat{\mu}$,$\hat{\beta}$,$\hat{\alpha}$ respectively. The scalar M is defined as $M = (1 - T^{-2})m_{yy} - 12m_{ty}^2 + 12(1 + T^{-1})m_{ty}m_y - (4 + 6T^{-1} + 2T^{-2})m_y^2$. The critical values used are the same as for the Dickey Fuller unit root test.

1.11 Structural Breaks and Unit Root Tests

1.11.1 Perron(1989) Unit Root Test

Perron, Pierre (1989) proposed unit root tests for processes with a linear trend with a break in intercept with the unit root as a null hypothesis, and the break model without the unit root as an alternative hypothesis (Perron, Pierre, 1989, p. 4-5):

$$\begin{aligned} H_0 :\ & y_t = y_{t-1} + \mu_1 + (\mu_2 - \mu_1)\mathbf{1}\{t > T_B\} + \beta t + e_t \\ H_1 :\ & y_t = \mu_1 + (\mu_2 - \mu_1)\mathbf{1}\{t > T_B\} + \beta t + e_t \end{aligned} \tag{1.61}$$

where the innovations e_t are stationary ARMA processes:

$$A(L)e_t = B(L)v_t \ ;\ v_t \text{ i.i.d } (0, \sigma^2) \tag{1.62}$$

with $A(L)$ and $B(L)$ p^{th} and q^{th} order polynomials, so the innovation series $\{e_t\}$ is taken to be an ARMA(p,q) process.

And the AR(1) multiplier of $y_{t-1}$is estimated using OLS ($\hat{\phi_1}$). As presented by Choi, In (2010, p. 60) $\hat{\phi_1}$ is asymptotically biased towards 1, but is consistent for Model (1.61), which makes this test have low statistical power.

Perron, Pierre (1989) also considers tests for a process with breaks in intercept in trend and intercept:

$$H_0: \; y_t = y_{t-1} + \mu_1 + (\mu_2 - \mu_1)\mathbf{1}\{t > T_B\} + (t - T_B)(\beta_2 - \beta_1)\mathbf{1}\{t > T_B\} + \beta_1 t + e_t$$

$$H_1: \; y_t = \mu_1 + (\mu_2 - \mu_1)\mathbf{1}\{t > T_B\} + (t - T_B)(\beta_2 - \beta_1)\mathbf{1}\{t > T_B\} + \beta_1 t + e_t \tag{1.63}$$

Choi, In (2010, p. 60) explains that $\hat{\phi}_1 \underset{\longrightarrow}{p} 1$ for Model (1.63) implying that the unit root tests for this model are inconsistent: a stationary process with a breaking trend **cannot** be distinguished from a unit root process with drift.

1.11.2 Zivot Andrews Unit Root Test

Andrews, Donald and Zivot, Eric (1992) developed a unit root test (ZA) that can handle structural breaks. The test statistic of the ZA test is the Student t ratio just as was the case with the PP test. As detailed in Pfaff, B. (2008, p. 110):

$$t_{\hat{\alpha}^i}[\hat{\lambda}^i_{\text{inf}}] = \inf_{\lambda \in \Delta} t_{\hat{\alpha}^i}(\lambda) \text{ for } i = A, B, C \tag{1.64}$$

where Δ is a closed subset of $(0, 1)$. Depending on the model, the test statistic is inferred from one of these three regression models:

$$y_t = \hat{\mu}^A + \hat{\theta}^A DU_t(\hat{\lambda}) + \hat{\beta}^A t + \hat{\alpha}^A t y_{t-1} + \sum_{i=1}^{k} \hat{c}_i^A \Delta y_{t-i} + \hat{\epsilon}_t \tag{1.65}$$

$$y_t = \hat{\mu}^B + \hat{\gamma}^B DT_t^*(\hat{\lambda}) + \hat{\beta}^B t + \hat{\alpha}^B t y_{t-1} + \sum_{i=1}^{k} \hat{c}_i^B \Delta y_{t-i} + \hat{\epsilon}_t \tag{1.66}$$

$$y_t = \hat{\mu}^C + \hat{\theta}^C DU_t(\hat{\lambda}) + \hat{\beta}^C t + \hat{\alpha}^C t y_{t-1} + \hat{\gamma}^B DT_t^*(\hat{\lambda}) + \sum_{i=1}^{k} \hat{c}_i^C \Delta y_{t-i} + \hat{\epsilon}_t \tag{1.67}$$

where $DU_t(\lambda) = 1$ if $t > T\lambda$ and 0 otherwise, and $DT_t^*(\lambda) = t - \lambda T$ for $t > T\lambda$ and 0 otherwise. The null hypothesis of the ZA URT does not allow structural breaks. Only breaks are allowed in the alternative hypothesis. Glynn, J. and Perera, N. and Verma, R. (2007) criticize this since if there are breaks under the null ($\phi_1 = 1$) we can mistakenly conclude that the series is stationary (with breaks.)

1.12 Kwiatkowski–Phillips–Schmidt–Shin (KPSS) Stationarity Test

The KPSS test proposed by Kwiatkowski, D. and Phillips, P. C. B. and Schmidt, P. and Shin, Y. (1992) has the null hypothesis of stationarity around a deterministic trend and the alternative of a unit root. The following definitions and explanations of the KPSS test are taken from Pfaff, B. (2008, p. 103). The underlying model considered in the KPSS test is:

$$y_t = \xi t + r_t + \epsilon_t \tag{1.68}$$

$$r_t = r_{t-1} + u_t \tag{1.69}$$

where r_t is a random walk and the innovations are assumed to be i.i.d. $(0, \sigma_u^2)$. r_0 is assumed to be a constant level. If $\xi <> 0$ then the model will contain a time trend component in addition to the level. Under the null hypothesis y_t is assumed to be I(0) so that $sigma_u^2 = 0$ and $r_t = r_0 =$ constant.

The test statistic is derived as follows. First y_t is regressed on a constant or a constant plus a linear trend. The the partial sums S_t of the residuals $\hat{\epsilon}_t$ are computed:

$$S_t = \sum_{i=1}^{t} \hat{\epsilon}_i, t = 1, ..., T \tag{1.70}$$

The test statistic LM is computed as:

$$LM = \frac{\sum_{t=1}^{T} S_t^2}{\hat{\sigma_\epsilon}^2} \tag{1.71}$$

where $\hat{\sigma_\epsilon}^2$ is an estimate of the error variance of the residuals. The authors suggest using a Bartlett window $w(s,l) = 1 - s/(l+1)$ as an optimal weighting function for estimating the long run variance of the errors:

$$\hat{\sigma_\epsilon}^2 = T^{-1} \sum_{t=1}^{T} \hat{\epsilon_t}^2 + 2T - 1 \sum_{s=1}^{l} + \frac{s}{l+1} \sum_{t=s+1}^{T} \hat{\epsilon_t} \epsilon_{\hat{t-1}} \tag{1.72}$$

The critical values of the test statistic are given in Kwiatkowski, D. and Phillips, P. C. B. and Schmidt, P. and Shin, Y. (1992).

1.13 Cointegration

Linear relationships involving integrated non-stationary TS are meaningful only if the TS are cointegrated. There are various definitions of cointegration; we define a cointegrating relationship between two or more TS each having unit roots (I(1)) if a linear combination exists that is stationary, i.e. I(0). When there are only two TS to test for cointegration a common approach is two use the two-step Engle Granger procedure, where the first step consists of using least squares to derive a linear relationship between the two variables, and the second step consists of using a unit root test on the residuals of the first step's regression. When there are more than two TS to check for cointegration, there are multiple possible cointegrating relationships, and the Engle Granger two step methodology is not sufficiently flexible. In this case the most common approach is to use the Johansen cointegration test.

1.13.1 Error Correcting Model

Assume that x_t and y_t are cointegrated and define $s_t = y_t - \beta x_t$ as the error in the equilibrium relationship at time t. Consider the so called error correction model (ECM):

$$\Delta y_t = \lambda_0 + \gamma_0 \Delta x_t + \rho \Delta y_{t-1} + \gamma_1 \Delta x_{t-1} + \delta s_t \epsilon_t$$

$$\Delta y_t = \lambda_0 + \gamma_0 \Delta x_t + \rho \Delta y_{t-1} + \gamma_1 \Delta x_{t-1} + \delta (y_{t-1} - \beta x_{t-1}) + \epsilon_t \tag{1.73}$$

This model factors in an equilibrium relationship between the non-stationary level variables x and y. The error correction term s_t is an adjustment that is a linear function of the disequilibrium between both variables. This type of ECM can be expanded to include any number of variables in what is referred to as a vector error correcting model (VECM.) See Engle, R. F. and Granger,C. W. J. (1987).

1.13.2 Johansen Cointegration Test

A vector autoregression (VAR) model is a linear model involving multiple TS. VAR models are a multivariate extension of the univariate AR models. The most common way to test for multiple cointegrating relationships is the approach developed by Johansen, S. (1988) and Johansen, S. and Juselius, K. (1990), which is based on the estimation of a p^{th}–order VAR in the k variables. The VAR in the k–vector y is:

$$y_t = \Pi_1 y_{t-1} + \Pi_2 y_{t-2} + \cdots + \Pi_p y_{t-p} + \Psi D_t + \epsilon_t \tag{1.74}$$

where D_t is a d–vector of deterministic terms, such as a constant, time trend and seasonal dummies if necessary. The VAR may be expressed as an ECM:

$$\Delta y_t = \Pi y_{t-2} + \Gamma_1 \Delta y_{t-1} + \Gamma_2 \Delta y_{t-2} + \cdots + \Gamma_{p-1} \Delta y_{t-(p-1)} + \Psi D_t + \epsilon_t \tag{1.75}$$

Equation (1.75) here can be augmented to include exogenous variable differences as Johansen, S. and Juselius, K. (1990) did, but we do not include them here.

The following matrix Π will be the key element to analyze the rank of the system to estimate the number of possible cointegrating relationships if any:

$$\Pi = -\left(I - \Pi_1 - \Pi_2 - ... - \Pi_p\right) \tag{1.76}$$

No assumption is made about the rank of Π. In the decomposition $\Pi = \alpha\beta'$, α and β are k x k matrices. We seek to determine whether any columns of β (that is, rows of β') are statistically indistinguishable from zero vectors. The existence of r cointegrating vectors reduces the rank of Π by $k - r$: that is, If there were r cointegrating relationships between the given variables, then there would be r nonzero eigenvalues in the dynamic system, and $k - r$ zero eigenvalues. Consider the decomposition $\Pi = \alpha\beta'$ where α and β are both k x r matrices. If the cointegration rank was $r = k$, then the VAR is stationary in the levels. However if the cointegration rank were zero no cointegration relationship would exist.

The Johansen methodology provides inference on the number of nonzero eigenvalues, or cointegrating relationships, by setting up an eigenvalue problem derived from the levels and differences of the k variables. The eigenvalues are ordered, from largest to smallest. The space spanned by the r largest eigenvalues is the r–dimensional cointegrating space. If $r = 1$, β is k x 1, and is the eigenvector corresponding to the largest eigenvalue. If $r = 2$, β is k x 2; the first column is as before,and the second column is the eigenvector corresponding to the second largest eigenvalue.

One of the key ideas behind the Johansen approach is given by the Frisch–Waugh–Lovell theorem. Consider the regression:

$$y_1 = \beta_1 x_1 + \beta_2 x_2 + \epsilon \tag{1.77}$$

The idea is we can "take out" x_2 to estimate the coefficient on x_1 as follows. We regress

y_1 on x_2 and get residuals r_1:

$$y_1 = \beta_3 x_2 + \epsilon_1 \tag{1.78}$$

$$r_1 = y_1 - \hat{\beta}_3 x_2 \tag{1.79}$$

We regress x_1 on x_2 and obtain residuals r_2:

$$x_1 = \beta_4 x_2 + \epsilon_2 \tag{1.80}$$

$$r_2 = x_1 - \hat{\beta}_4 x_2 \tag{1.81}$$

Finally we regress r_1 on r_2 without an intercept.

$$r_1 = \beta_5 r_2 + \epsilon_3 \tag{1.82}$$

The estimated coefficient on r_2 is equivalent to the β_1 coefficient that would have been estimated from Equation (1.77) using OLS: $\hat{\beta_5} = \hat{\beta_1}$.

As we will see shortly to estimate Equation (1.75) we will separate it out to two separate regressions in Equations (1.86), and (1.86) which reflect the regressors $\Gamma_a \Delta y_{t-a}$ and Πy_{t-2}.

Following is a detailed overview of the Johansen cointegration test as performed in the `ca.jo()` function in the `urca` R package.

Define $\boldsymbol{x_t} = (x_1(t), \ldots, x_k(t))^\intercal$ as a vector of k TS to be tested for a cointegrating relationship, with n observations of each. As Pfaff, B. (2008, p. 81)explains, the Johansen cointegration approach is based on *canonical correlation* analysis as a way to reduce the information content of the n observations in the k dimensional space, to a lower dimensional space.

For this purpose $2k$ regressions are estimated using ordinary-least-squares. $\Delta \boldsymbol{x_t}$ is regressed on lagged differences of $\boldsymbol{x_t}$, and the residuals are labeled $\boldsymbol{R}_0$.

Then $\boldsymbol{x}_{t-p}$ is regressed on lagged differences of $\boldsymbol{x_t}$, and the residuals are labeled $\boldsymbol{R}_1$. In more detail:

$\boldsymbol{Z}_0$ and $\boldsymbol{Z}_k$ are $n \times k$ matrices, and $\boldsymbol{Z_1}$ is a $n \times (k+1)$ matrix. The first column of $\boldsymbol{Z}_1$ are all 1 entries.

$$\boldsymbol{Z}_0 = \underset{((n-k)\times k)}{\begin{bmatrix} \Delta(x_1(k+1)) & ... & \Delta(x_k(k+1)) \\ \Delta(x_1(k+2)) & ... & \Delta(x_k(k+2)) \\ ... & ... & ... \\ \Delta(x_1(n)) & ... & \Delta(x_k(n)) \end{bmatrix}} \tag{1.83}$$

$$\boldsymbol{Z}_1 = \underset{((n-k)\times(k+1)}{\begin{bmatrix} 1 & \Delta(x_1(k)) & ... & \Delta(x_k(k)) \\ 1 & \Delta(x_1(k+1)) & ... & \Delta(x_k(k+1)) \\ ... & ... & ... & ... \\ 1 & \Delta(x_1(n-1)) & ... & \Delta(x_k(n-1)) \end{bmatrix}} \tag{1.84}$$

$$\boldsymbol{Z}_k = \underset{((n-k)\times k)}{\begin{bmatrix} x_1(1) & x_k(1) \\ x_1(2) & x_k(2) \\ ... & ... \\ x_1(n-k) & x_k(n-k) \end{bmatrix}} \tag{1.85}$$

$$\boldsymbol{Z}_0 = \boldsymbol{Z}_1\boldsymbol{\beta}_1 + \boldsymbol{\epsilon}_1 \tag{1.86}$$

$$\boldsymbol{Z}_k = \boldsymbol{Z}_1\boldsymbol{\beta}_2 + \boldsymbol{\epsilon}_2 \tag{1.87}$$

$$\boldsymbol{R}_0 = \boldsymbol{Z}_0 - \boldsymbol{Z}_1\widehat{\boldsymbol{\beta}}_1 \tag{1.88}$$

$$\boldsymbol{R}_k = \boldsymbol{Z}_k - \boldsymbol{Z}_1\widehat{\boldsymbol{\beta}}_2 \tag{1.89}$$

$$\boldsymbol{M_{00}} = \frac{\boldsymbol{Z}_0^\intercal \boldsymbol{Z}_0}{n} \tag{1.90}$$

$$\boldsymbol{M_{11}} = \frac{\boldsymbol{Z}_1^\intercal \boldsymbol{Z}_1}{n} \tag{1.91}$$

$$\boldsymbol{M_{kk}} = \frac{\boldsymbol{Z}_k^\intercal \boldsymbol{Z}_k}{n} \tag{1.92}$$

$$\boldsymbol{M_{01}} = \frac{\boldsymbol{Z}_0^\intercal \boldsymbol{Z}_1}{n} \tag{1.93}$$

$$\boldsymbol{M_{0k}} = \frac{\boldsymbol{Z}_0^\intercal \boldsymbol{Z}_k}{n} \tag{1.94}$$

$$\boldsymbol{M_{k0}} = \frac{\boldsymbol{Z}_k^\intercal \boldsymbol{Z}_0}{n} \tag{1.95}$$

$$\boldsymbol{M_{10}} = \frac{\boldsymbol{Z}_1^\intercal \boldsymbol{Z}_0}{n} \tag{1.96}$$

$$\boldsymbol{M_{1k}} = \frac{\boldsymbol{Z}_1^\intercal \boldsymbol{Z}_k}{n} \tag{1.97}$$

$$\boldsymbol{M_{k1}} = \frac{\boldsymbol{Z}_k^\intercal \boldsymbol{Z}_1}{n} \tag{1.98}$$

$$\boldsymbol{R_0} = \boldsymbol{Z}_0 - \left(\boldsymbol{M}_{01}\boldsymbol{M}_{11}^{-1}\boldsymbol{Z}_1^\intercal\right)^\intercal \tag{1.99}$$

$$\boldsymbol{R_k} = \boldsymbol{Z}_k - \left(\boldsymbol{M}_{k1}\boldsymbol{M}_{11}^{-1}\boldsymbol{Z}_1^\intercal\right)^\intercal \tag{1.100}$$

$$\boldsymbol{R_0} = (R_{01}, ..., R_{0n}) \tag{1.101}$$

$$\boldsymbol{R_k} = (R_{k1}, ..., R_{kn}) \tag{1.102}$$

The residual vectors $\boldsymbol{R_0}$ and $\boldsymbol{R_0}$ are used to compute the product moment matrices:

$$\widehat{\boldsymbol{S}}_{ij} = \frac{1}{n}\sum_{t=1}^{n} \boldsymbol{R}_{it}\boldsymbol{R}_{jt}^\intercal \tag{1.103}$$

The four product moment matrices are computed in `urca` as follows:

$$\boldsymbol{S_{00}} = \boldsymbol{M}_{00} - \left(\boldsymbol{M}_{01}\boldsymbol{M}_{11}^{-1}\boldsymbol{M}_{10}\right) \tag{1.104}$$

$$\boldsymbol{S_{0k}} = \boldsymbol{M}_{0k} - \left(\boldsymbol{M}_{01}\boldsymbol{M}_{11}^{-1}\boldsymbol{M}_{1k}\right) \tag{1.105}$$

$$\boldsymbol{S_{k0}} = \boldsymbol{M}_{k0} - \left(\boldsymbol{M}_{k1}\boldsymbol{M}_{11}^{-1}\boldsymbol{M}_{10}\right) \tag{1.106}$$

$$\boldsymbol{S_{kk}} = \boldsymbol{M}_{kk} - \left(\boldsymbol{M}_{k1}\boldsymbol{M}_{11}^{-1}\boldsymbol{M}_{1k}\right) \tag{1.107}$$

$$\boldsymbol{S_{kk}} \approx \boldsymbol{C} = \boldsymbol{L}\boldsymbol{L}^{\intercal} \tag{1.108}$$

And finally the matrix $\boldsymbol{E}$ will be used to decompose into eigenvalues that will be used by the Johansen test statistics:

$$\boldsymbol{E} = \boldsymbol{C}^{-1}\boldsymbol{S}_{k0}\boldsymbol{S}_{00}^{-1}\boldsymbol{S}_{0k}\left(\boldsymbol{C}^{-1}\right)^{\intercal} \tag{1.109}$$

Johansen defined two statistics to determine the cointegration rank of $\boldsymbol{\Pi}$ by computing the eigenvalues of matrix $\boldsymbol{E}$: first, the trace statistic,

$$\text{trace} = -n \sum_{i=r+1}^{k} \ln(1-\lambda_i) \tag{1.110}$$

which allows for the test of the hypothesis $H(r)$: the rank of $\boldsymbol{\Pi}$ is r, v.s. the alternative hypothesis that the rank of Π is k. A large value of the trace statistic provides evidence against hypothesis $H(r)$: for example, if with $r = 1$, the value of the trace statistic is greater than the appropriate critical value then we would reject the hypothesis of cointegration with rank $r = 1$ in favor of a cointegration with rank $r > 1$. The test can then be repeated for $r = 2$, $r = 3$, ... and so forth.

The λ_{max} test statistic may be used as well :

$$\lambda_{max} = -n \ln(1 - \lambda_{r+1}) \tag{1.111}$$

This test allows for the comparison of the null hypothesis of cointegration rank of r against the alternative hypothesis of a cointegration rank of $r+1$. This test also may then be repeated for larger values of r until the null hypothesis can no longer be rejected.

The distributions of trace and λ_{max} statistics are nonstandard, and depend on the deterministic parameters in D_t. Critical values have been tabulated by the authors of the Johansen approach and others Johansen, S. and Juselius, K. (1990). The Johansen test used in this paper was implemented as part of the `urca` R software package written by Pfaff, B. (2008).

The critical values for these test statistics are determined using simulation. As Johansen, S. and Juselius, K. (1990, p. 179) explains the asymptotic distribution for the trace statistic is based on the trace of the following stochastic matrix where $\boldsymbol{W}$ denotes an $(n-r)$ dimensional Brownian motion

$$\int_0^1 (\boldsymbol{dW})\boldsymbol{F}^T \left(\int_0^1 \boldsymbol{F}\boldsymbol{F}^T dr\right)^{-1} \int_0^1 \boldsymbol{F}(\boldsymbol{dW})^T \tag{1.112}$$

$$\boldsymbol{F} = (F_1, ..., F_{n-r}) \tag{1.113}$$

$$F_i(t) = W_i(t) - \int_0^t W_i(t) ds \tag{1.114}$$

The asymptotic distribution for the maximum eigenvalue test is given by the maximum eigenvalue of the stochastic matrix Equation (1.112). For the case when the model allows for a mean and/or a trend the multivariate Brownian motions in the previous formulas should be replaced by the demeaned and/or detrended Brownian motions; this will result in a different set of critical values for these test statistics.

1.14 Threshold Cointegration

In their seminal work, Balke, Nathan S. and Fomby, Thomas B. (1997) introduced the concept of threshold cointegration.

$$y_t + \alpha x_t = z_t \text{ , where } z_t = \phi^{(i)} z_{t-1} + \epsilon_t \tag{1.115}$$

$$y_t + \beta x_t = B_t \text{ , where } B_t = B_{t-1} + \eta_t \tag{1.116}$$

where B_t is the common stochastic trend of x_t and y_t and ϵ_t and η_t are zero mean i.i.d. random variables, and $\phi^{(i)}$ is defined as follows:

$$\phi^{(i)} = \begin{cases} \phi_1, \text{ with } |\phi_1| < 1, & \text{if } |z_{t-1}| \leq \theta \\ 1, & \text{if } |z_{t-1}| > \theta \end{cases} \tag{1.117}$$

Equation (1.115) represents an equilibrium relationship, while Equation (1.116) represents a common stochastic trend, i.e., a nonstationary relationship. If the equilibrium error is less than the threshold then x_t and y_t do not revert to an equilibrium, however if the error is greater than the threshold then the two TS variables equilibrate.

This formulation of threshold cointegration can be equivalently expressed as a threshold error correcting model:

$$\Delta y_t = \gamma_1^{(i)} \ z_t + \upsilon_{1t} \tag{1.118}$$

$$\Delta x_t = \gamma_2^{(i)} \ z_t + \upsilon_{2t} \tag{1.119}$$

Balke, Nathan S. and Fomby, Thomas B. (1997) also consider a more general equilibrium model of z_t based on a self exciting threshold auto regressive framework (SETAR) such as the following with a low, middle and high regimes:

$$z_t = \begin{cases} \mu_h + \phi_h z_{t-1} + \epsilon_t, & \text{if } z_{t-1} > \theta_H \\ \mu_m + \phi_m z_{t-1} + \epsilon_t, & \text{if } \theta_L \leq z_{t-1} \leq \theta_H \\ \mu_l + \phi_l z_{t-1} + \epsilon_t, & \text{if } z_{t-1} < \theta_L \end{cases} \tag{1.120}$$

As Stigler,Matthieu (2010) points out the common assumed case where Equation (1.120) is stable is when $\phi_h < 1$ and $\phi_l < 1$.

Seo, B. (2006) developed a test for the linear no cointegration null hypothesis against threshold cointegration in a threshold vector error correction model. They used a sup-Wald type test and derived its null asymptotic distribution. They proposed using a residual-based bootstrap and showed with Monte Carlo simulations that the bootstrap corrected size distortions of the asymptotic distribution in finite samples, and that it had greater power against the threshold cointegration alternative than conventional cointegration tests.

Stigler,Matthieu (2010) posits that SETAR models are so complex that their probability distribution and moments are only known for simple cases. Estimation for more than one threshold is not easy. He points out that in a cointegration model with threshold effects there is no known procedure to test for stationarity with an unknown cointegrating vector. A test to determine 2 versus 3 regimes only works in a restricted case.

Bec,Frederique and Ben Salem,Melika and Carrasco,Marine (2004) proposed the BBC unit root test which considers a three-regime self-exciting threshold autoregressive (SETAR) model.

1.15 Long Memory Processes

The following definitions and explanations are taken from Pfaff, B. (2008, p. 63-68). A series $\{x_t\}$ is a stationary and invertible ARFIMA(p,d,q) process if it can be written as

$$\Delta^d x_t = z_t \tag{1.121}$$

where $\{z_t\}_{t=-\infty}^{\infty}$ is an ARMA(p,q) process such that $z_t = \phi_p(L)^{-1}\theta_q(L)\epsilon_t$ and both polynomials have their roots outside of the unit root circle where ϵ_t is a 0-mean i.i.d. random variable with variance σ^2 and $d \in (-0.5, 0.5]$.

When $0 < d < 0.5$ the process has long memory, meaning the autocorrelation of $\{x_t\}$ is highly persistent so the sum of the correlations for different lags becomes infinite, unlike standard ARMA models where the correlation decays to zero over a relatively short amount of time. When $-0.5 < d < 0$ the sum of the autocorrelations for different lags is more or less constant, and are often labeled as having "intermediate memory." As long as $d > 0.5$ the process has an invertible moving average representation. It can be shown that the auto-correlation function of long memory processes declines hyperbolically instead of exponentially as it does with ARMA processes. The speed of the decay depends on the parameter d.

As Pfaff, B. (2008, p. 63) points out, Hurst, H. E. (1951) provided a commonly used approach for detecting the presence of long-term memory. Hurst proposed the rescaled range statistic(R/S.) This descriptive measure is defined as

$$R/S = \frac{1}{s_T}\left(\max_{1\le k\le T}\sum_{j=1}^{k}(y_j - \bar{y}) - \min_{1\le k\le T}\sum_{j=1}^{k}(y_j - \bar{y})\right), \tag{1.122}$$

where s_T is the standard deviation estimator

$$s_T = \left(\frac{1}{T} \sum_{j=1}^{T} (y_j - \bar{y})^2 \right)^{\frac{1}{2}} \tag{1.123}$$

Hurst[1951] showed that the probability limit

$$\operatorname*{plim}_{T \to \infty} \left(T^{-H} \frac{R/S}{s_T} \right) = \text{constant} \tag{1.124}$$

The Hurst coefficient (H) can be estimated as:

$$\hat{H} = \frac{\log R/S}{\log T} \tag{1.125}$$

A short-memory process has a value of H = 0.5 , and estimated values greater than 0.5 are indicative of long-memory behavior. The estimated differencing parameter $\hat{d} = \hat{H} - \frac{1}{2}$.

Granger, C. W. J. (2000) discuss "stylized facts" of financial asset return TS where their autocorrelations remain positive and significant for very long lags (2000); even though there is considerable literature that uses this as a justification for modeling the returns as a fractionally integrated I(d) process, the author argues that an adequate alternative explanation is a non-fractionally integrated, or an I(0) model with occasional structural breaks.

1.16 Cointegration Under Structural Breaks

When testing for cointegration of variables that experience structural breaks the standard tests cannot be used as the asymptotic distributions of the test statistics change as explained by Giles, David E. and Godwin, Ryan T. (2011). These authors provide a new set of critical values to use with the standard Johansen cointegration test.

Lutkepohl, H. and Saikkonen, P. and Trenkler, C. (2004) propose a different methodology for testing cointegration for variables with structural shifts in level(a change in the mean);first the breakpoint is estimated and the effect of the structural shift in the deterministic terms is removed and then the standard Johansen cointegration procedure can be used with the standard critical values.

1.17 Estimation of Breakpoints

As Zeileis,Achim and Leisch,Friedrich and Hornik,Kurt and Kleiber,Christian (2002) point out the foundation for estimating breaks was given by Bai, J. (1994) for one break and later extended to multiple breaks by Bai, J. and Perron, P. (1998). In this thesis we use the *breakpoints()* function of the *strucchange* R package which implements the multiple simultaneous breakpoint estimation procedure in Bai, J. and Perron, P. (2003). This implementation was proposed in Zeileis,Achim and Kleiber,Christian and Kramer,Walter and Hornik,Kurt (2003).

We follow the explanation given by Zeileis,Achim and Leisch,Friedrich and Hornik,Kurt and Kleiber,Christian (2002, p. 12). The algorithm for computing the optimal breakpoints given the number of breaks is based on a dynamic programming approach. It assumes the underlying linear model with a dependent one dimensional variable y_t, a $p \times 1$ covariates vector $\boldsymbol{x_t}$ with corresponding coefficient vector $\boldsymbol{\beta}$, and the innovations ϵ_t:

$$y_t = \boldsymbol{x_t}\boldsymbol{\beta} + \boldsymbol{\epsilon_t} \tag{1.126}$$

If there are m change points in the coefficient this implies $m+1$ regimes. Equation (1.126) can be rewritten as:

$$y_t = \boldsymbol{x_t}\boldsymbol{\beta_j} + \boldsymbol{\epsilon_t} \ (j = T_{j-1} + 1, ..., T_j, j = 1, ..., m + 1) \tag{1.127}$$

The underlying idea is that of the Bellman Principle of Optimality that posits solving a problem by dividing it into independent optimally solvable sub-problems, whose solutions can be combined to solve the larger problem. The sub-problem solutions are typically stored in memory ("memoized".) In this case a triangular residual sum of squares (RSS) matrix is computed and stored in memory which can be reused over and over again to derive the residual sum of squares for a segment starting at observation t and ending at t' with $t < t'$. This approach is considerably faster than the brute force approach of computing the RSS for all possible sub-segments. The Bai, J. and Perron, P. (2003) algorithm uses only $O(T^2)$ least squares operations for any number of change points m. The brute force approach would require $O(T^m)$ least squares operations. If considering breaks does not lower the RSS, then the procedure will return NA.

There are other methods to infer changepoints and break dates as Stock, James H. and Watson, Mark W. (2011, p.557-560) explain: when the break dates are known dummy variables can be added to enhance the model with date dependent coefficients, and these can be tested with F-test statistics with the null hypothesis being that these additional coefficients are zero. This is known as a Chow type test. When the break-dates are not known then Chow-tests can be performed across all possible dates and the one that results in the best fit can be returned. This modified Chow-test is called the Quandt likelihood ratio (QLR) statistic (Quandt,Richard E., 1960), and is also known as a *sup-Wald* statistic. Since the QLR statistic is the maximum of many F-statistics it has a different distribution from an F-statistic.

1.18 Summary of Research Contributions

This thesis contributes the following novel research findings:

- I showed via simulation studies that when testing for cointegration of two slightly explosive TS ($\phi_1 > 1$) results almost invariably in spurious cointegration; this is not mentioned in the literature. None of the standard established unit root tests analyzed in this thesis such as the ADF URT reject the null hypothesis of I(1) in the case of $\phi_1 > 1$. See Tables 3.36, 3.37 and 3.38.

- I developed a new unit root test (URT)–the lagged-series URT which has similar statistical power to the Augmented Dickey Fuller (ADF) test when the auto-regressive multiplier $\phi_1 < 1$ but exceeds the power of the ADF test when $\phi_1 > 1$[1]. We show empirically via simulation studies that a valid[2] cointegration relationship between a TS and its lag implies an I(1) TS. This cointegration test will never result in a case of no possible cointegrating relationships for a reasonable lag.
- The new lagged-series test does not reject I(2) series similarly to the ADF and ERS URTs. However the new lagged-series URT rejects at least 65% of I(3) tests and 85% of I(4) tests; by comparison the ADF URT only rejected 5% of the I(3) and I(4) tests.
- I combined the new lagged-series (URT) with other unit root tests such as ERS and ZA tests which improve the power of these tests when the AR(1) auto-regressive multiplier $\phi_1 > 1$.
- I developed and tested the HBPZA unit root test which allows for structural breaks in intercept and linear trend under the null hypothesis and compared it to an existing implementation of the Lee-Stratizich URT that also allows breaks in the null. The new test performs better than the Lee-Stratizich[3] in a number of situations.
- I developed the HBPADF testing procedure that allows discerning I(0)-I(1) shifts within the evaluated time series, and performed simulation studies on it.
- I combined URTs with deep learning neural networks (DLNNs) which outperform individual tests when we consider the net error across null and alternative hypotheses.

[1] Chandra, Suresh K. and Janhavi, J.V. (2008) developed a modified ADF URT that is supposed to reject explosive unit roots. I did not have access to this test so I cannot make any comparisons with the lagged-series test.

[2] When testing a TS with its lag if it has two cointegrating relationships this implies that the series is I(0). Two variables can only have one valid cointegration relationship.

[3] I am in contact with Johannes Lips,the author of the R implementation of the Lee-Stratizich URT. This test appears to reject much more than expected under the null hypothesis, so it is not certain if this is a bug in the software or an issue with the procedure itself.

- I performed simulation based studies of the ADF, ERS-Ptest, ERS-DFGLS, the Zivot-Andrews, and the new lagged-series URTs, under various Model configurations. Findings are consistent with the literature that points out these tests are sensitive to the starting value of the unit root. Also when there are structural breaks in the null hypothesis the standard unit root tests hardly ever reject the null hypothesis of a unit root.

- I used a linear form of the covered interest rate parity (CIP) formula and showed that if cross-currency swap basis spreads are added to one of the 3 terms the cointegration relationship always strengthens.

- I showed that there are likely cointegration relationships between bank credit default swap spreads and cross currency basis swap spreads. This would provide evidence bank credit risk is related to cross currency basis swap spreads. I showed that there are cointegration relationships between bank credit default swap spreads and spot FX. This would indicate that bank credit risk affects the FX Spot rate.

- I analyzed interest rate and cross currency swap liquidity and ensured the series are not unit roots, and provided statistical evidence that USD-JPY cross-currency basis swaps Granger cause JPY fixed-floating IR swaps. A possible explanation may be USD based entities issuing JPY fixed debt and hedging it fully or partially with USD-JPY cross currency basis swaps.

- I used various analyses to check for unit roots in zonal temperature anomaly time series including under structural breaks in the null hypothesis using the new HBPZA test as well as the HBPADF testing procedure.A number of model were fitted to the temperature anomaly data, including a 3 regime SETAR model and showed that better fits are always achieved when adding linear trends.

- I showed that the estimated 3-regime SETAR models for the Southern Hemisphere temperature anomalies are more likely to be stationary than the Northern Hemisphere, which includes an AR explosive ($\phi_1 > 1$) middle regime.

1.18.1 Building Linear Models with Time Series

As is shown throughout this thesis there are certain cases when building linear regression models or cointegration models of time series that may appear to yield meaningful results, however they are spurious. The figures in this section consist of flowcharts that detail how to build a two or more linear model of time series variables to avoid as much as possible spurious relationships.

Consider the following AR(1) model with constant and a linear trend of a time series y_t:

$$y_t = x_t + \beta_0 + \beta_1 t \ ; \ x_t = \phi_1 x_{t-1} + \epsilon_t \tag{1.128}$$

We assume that we do not know what type of time series variables we are dealing with a priori but they can each be one of the following:

- A stationary I(0) variable: $|\phi_1| < 1$ in Equation (1.128)
- A unit root I(1) variable: $\phi_1 = 1$ in Equation (1.128)
- An explosive variable: $\phi_1 > 1$ in Equation (1.128)
- An I(n) TS variable z_t where $n > 1$: $z_t = w_t + \beta_0 + \beta_1 t \ ; \ w_t = \sum_{i_1=1}^{t} \sum_{i_2=1}^{t} \cdots \sum_{i_n=1}^{t} y_{i_n}$ where x_i is defined as in Equation (1.128)

Some of the major factors that contribute to spurious and/or sub-optimal linear models are as follows:

- A unit root test should be run on each time series variable to determine if they are either I(0) (stationary) or I(1) (unit root non-stationary.) The tests are not perfect and type I and type II errors can occur. In addition the series may be of another type besides I(1) or I(0) such as I(n) or explosive.
- Linear regression models involving I(1) time series variables produce spurious results. If all of the time series analyzed are I(1) and they are cointegrated then an error

correcting model can be estimated. Otherwise a linear regression on the differenced time series should be used.

- When testing for cointegration with explosive time series this almost invariably results in spurious cointegration. Most unit root tests cannot distinguish between an I(1) variable and an explosive variable. The lagged-series unit root test developed in this thesis can distinguish between I(1) and explosive processes as long at the series are not too short.

- Standard unit root tests such as the Augmented-Dickey-Fuller test cannot distinguish between an I(0) process with structural breaks in constant(β_0) and/or linear trend(β_1) and an I(1) process.

- The Zivot-Andrews unit root test can distinguish between an I(0) process with structural breaks in constant(β_0) and/or linear trend(β_1) and an I(1) process; however this test does not allow structural breaks under the null hypothesis I(1)–which could lead to a user accepting that an I(1) process with breaks is actually I(0) with breaks. In this thesis the HBPZA unit root test was developed that allows for breaks under the null of I(1).

- Standard unit root tests assume that an entire time series being tested can be either I(1) or I(0). It is possible that the series could have a strucutural break in the auto-regressive coefficient ϕ_1 such that it could be composed of alternating I(0)/I(1) segments in no particular order, and of any possible length, and this series could also exhibit structural breaks in the β_0 and β_1 coefficients. In this thesis the HBPADF testing procedure was developed to address this case.

- Another type of model that captures structural breaks is a 3-regime SETAR (threshold) model where the active regime is determined by checking the previous level of the variable to see if it belongs to a regime specific range. This type of model can have a non stationary middle regime while the outer regimes are stationary.

The suggested approach to build linear models of two or more time series variables is as follows in this order:

1. Figure 1.3 details how to estimate if each time-series is either I(0),I(1), Explosive or I(n) with $n > 1$ under the possibility of change points in all coefficients of Equation (1.128): β_0, β_1 and ϕ_1.

2. The second step is detailed in Figure 1.4 and shows how to estimate if each time-series is either I(0),I(1), Explosive or I(n) with $n > 1$ without change points. The first step also calls to this second step because either there was no evidence of change-points in the prior step, or because the breakpoints were "removed" after estimating them via the usage of dummy variables in the regressions.

3. Once all of the individual variables have been analyzed in the previous steps it is time to attempt to build a linear model. If all of the variables are of the same type, either all I(0) or all I(1) and there is no evidence of structural breaks then the procedure detailed in Figure 1.5 should be followed. If there is evidence of breakpoints in any of the series then the procedure detailed in Figure 1.6 should be followed instead.

4. This thesis does not consider the cases of how to build models when there are mixed orders of integration I(0), I(1), I(2). There is an expanding literature in this area but it was not considered in this thesis.

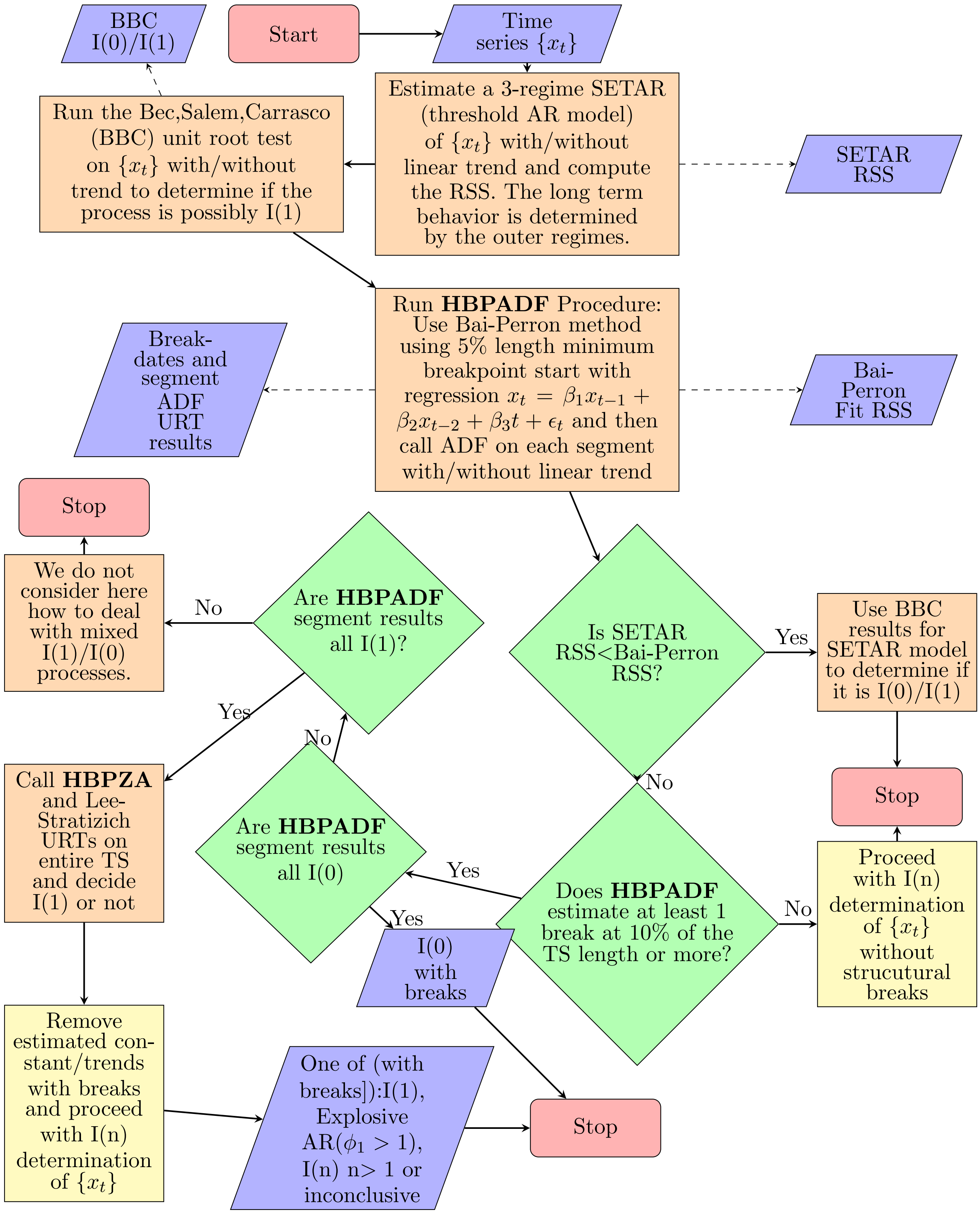


Figure 1.3: Determining a TS's Order of Integration/Explosiveness With/Without Structural Breaks

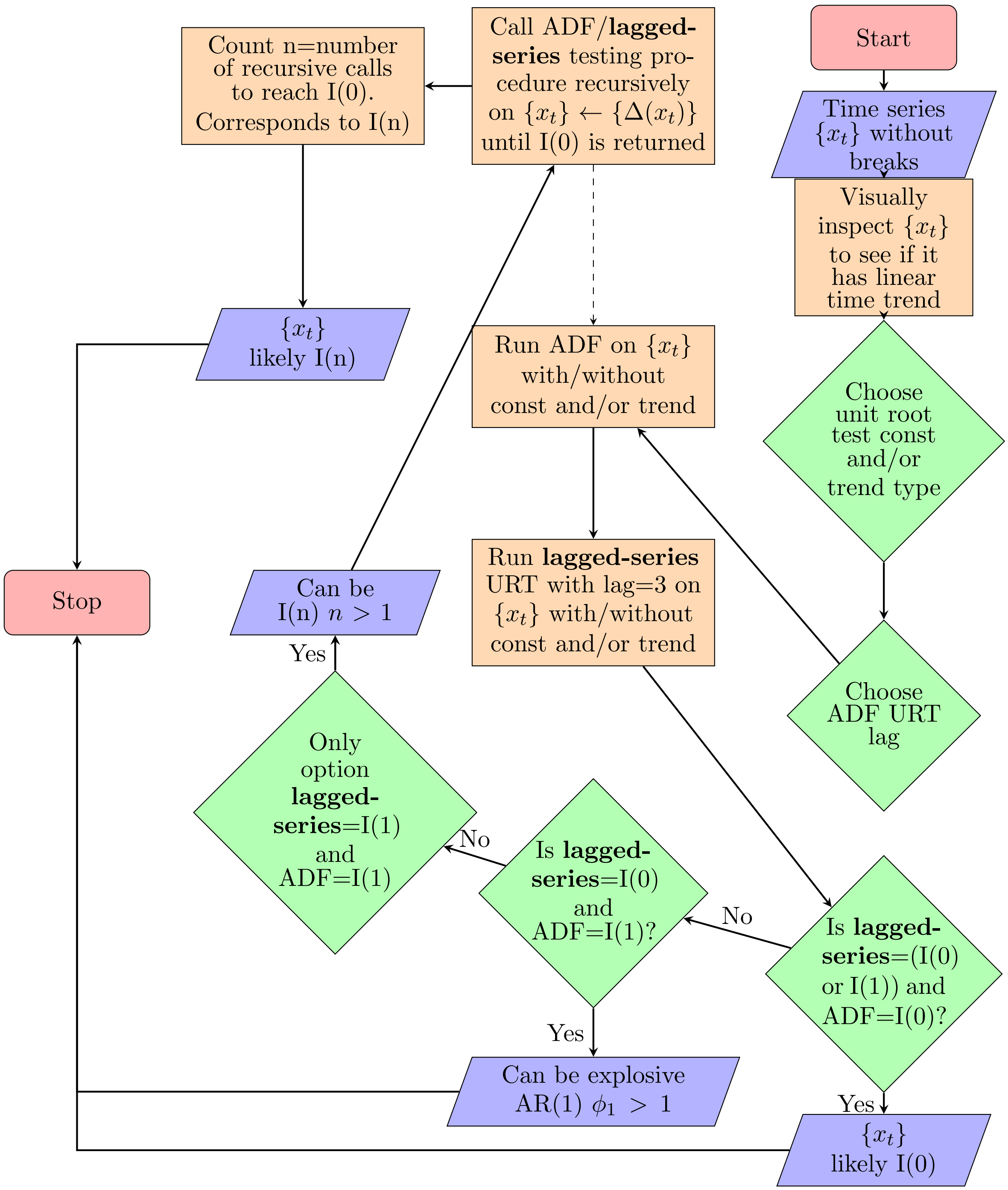


Figure 1.4: Determining a TS's Order of Integration/Explosiveness Without Structural Breaks

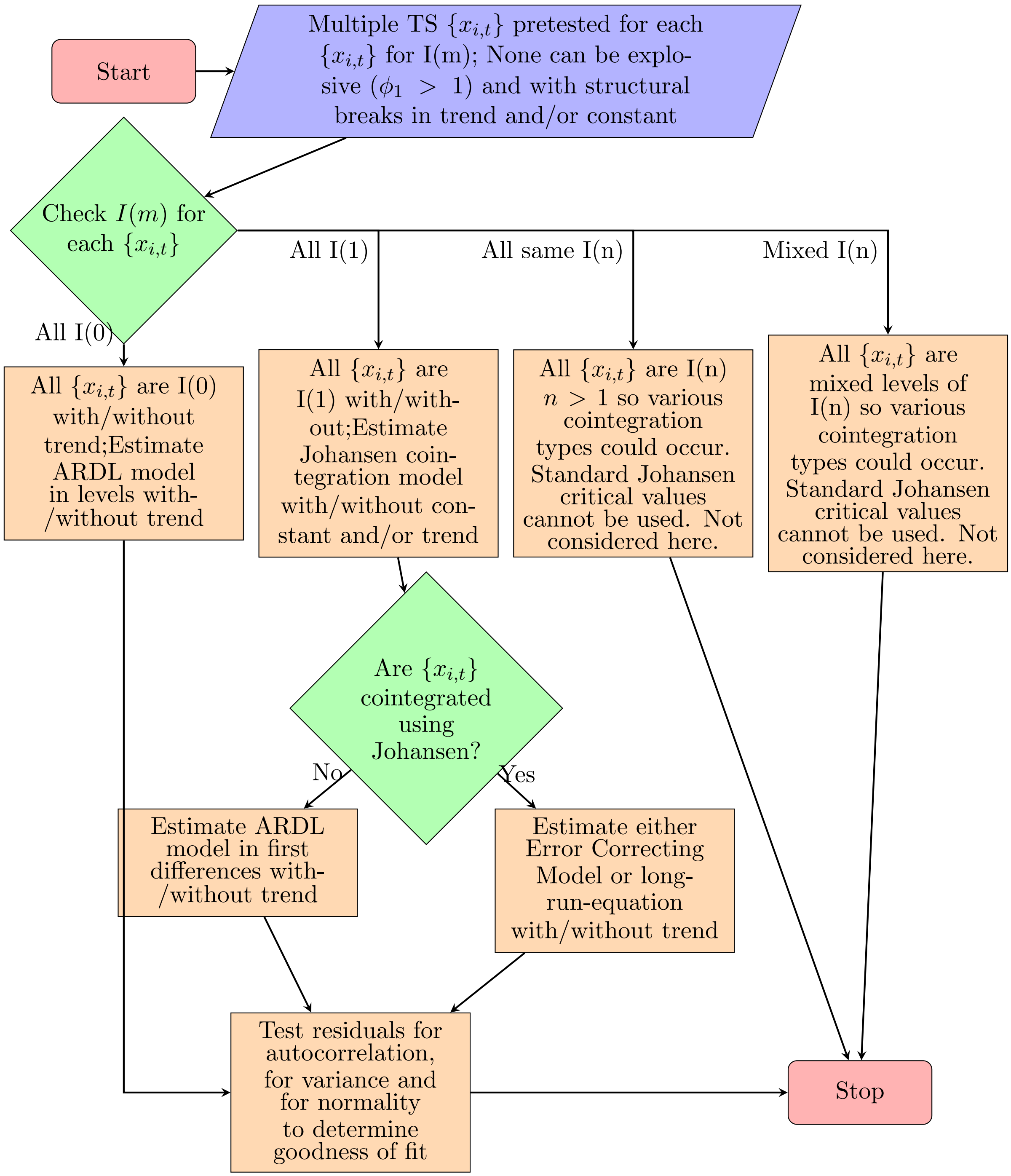


Figure 1.5: Regressions with I(n) TS Variables Without Structural Breaks Based on Hill et al. (2011, Fig. 12.4)

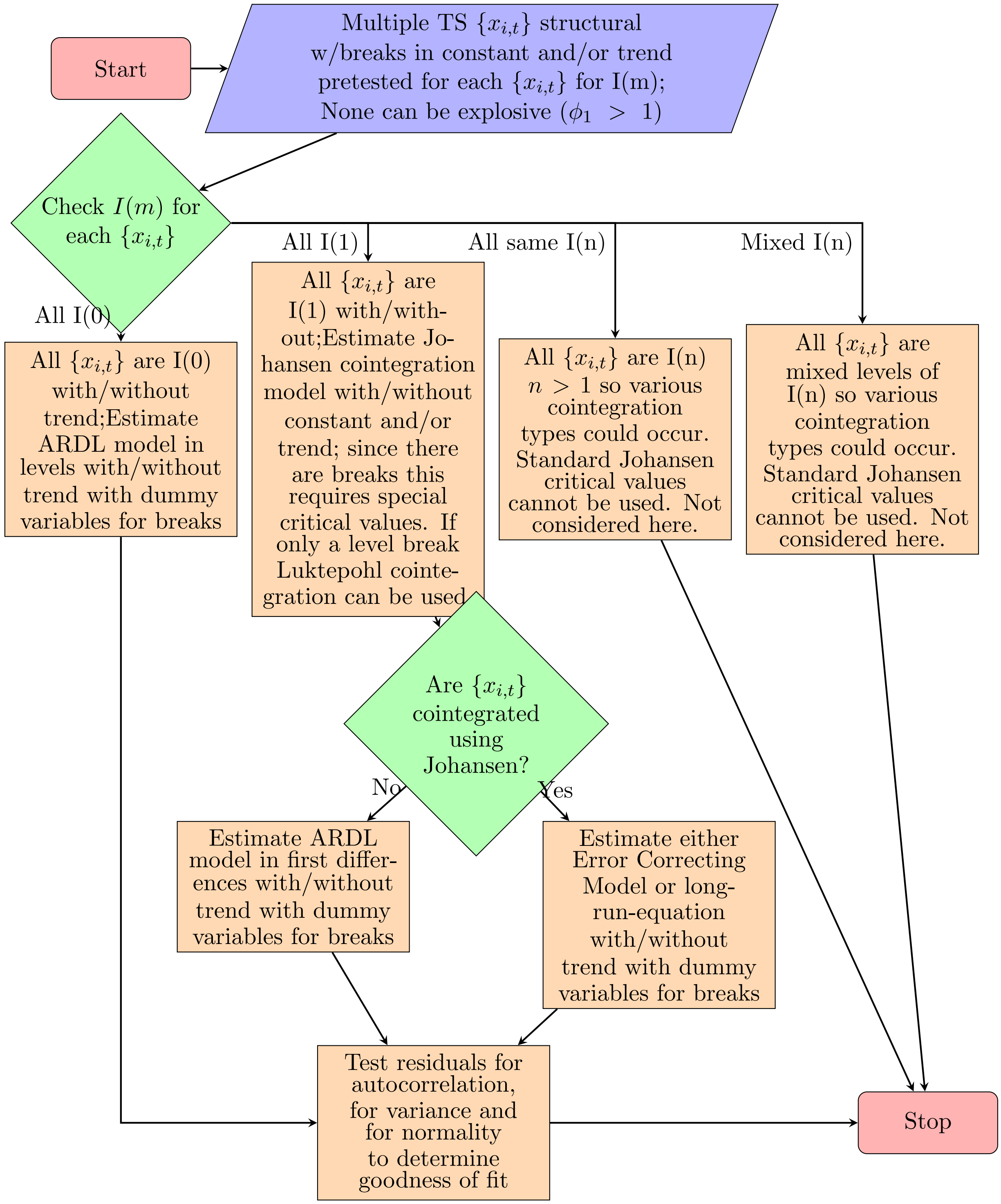


Figure 1.6: Regressions with I(n) TS Variables Without Structural Breaks Based on Hill et al. (2011, Fig. 12.4)

Chapter 2: New Unit Root Tests

2.1 The lagged-series Unit Root Test

I propose a new URT where we will show empirically performs better than the other analyzed tests in the case when $\phi_1 > 1$ in the AR(1) process, which I refer to throughout this thesis an an explosive AR(1) process. Because this test uses statistics computed from the lagged series I call it the lagged-series unit root test. Chandra, Suresh K. and Janhavi, J.V. (2008) proposed a modified ADF unit root test to better handle explosive AR(1) processes. I did not have access to an implementation of this URT for comparison purposes.

To give an intuition of why the explosive case is not easy to deal with we refer to the work of Phillips,Peter C.B. and Magadalinos,Tassos (2005) where they consider the following AR(1) framework:

$$x(t) = \phi_L x(t-1) + \epsilon_t \tag{2.1}$$

where the auto-regressive multiplier is assumed to have the following structure where L is the time-series length and c is a constant:

$$\phi_L = \frac{c}{k_L} \text{ ; } k_L = O(L) \tag{2.2}$$

They find the following asymptotic properties of the difference between the estimated and actual ϕ_L coefficient is as follows for the non-explosive case of $c < 0$ the scaled error made in the estimation is Gaussian with variance of $-2c$:

$$\sqrt{Lk_L}\left(\hat{\phi_L} - phi_L\right) \Rightarrow N(0, -2c) \text{ for } c < 0 \tag{2.3}$$

and for the explosive case of $c > 0$ the scaled estimation error is as follows:

$$\frac{k_L \phi_L^L}{2c}\left(\hat{\phi_L} - phi_L\right) \Rightarrow C \text{ for } c > 0 \tag{2.4}$$

where C denotes a Cauchy variate, which is problematic since the mean and variance of the Cauchy distribution do not exist.

2.1.1 Calculation

I propose the new lagged-series URT where we test for the number of cointegrating relationships of a time series (TS) with the lagged same TS using the Johansen cointegration test as explained earlier. *If the time-series is a unit root, I(1), we expect only one valid cointegrating relationship but if there are two cointegrating relationships the process is I(0).* Given two variables, only one valid cointegration relationship is possible,

We test the cointegration of the TS $\{\boldsymbol{x}_t^*\}$:

$$\{x_t^*\} = \{x_{k+1}, ..., x_n\} \tag{2.5}$$

And the lagged TS $\{L^k(\boldsymbol{x}_t)\}$:

$$\{L^k(x_t)\} = \{x_1, x_2, ..., x_{n-k}\} \tag{2.6}$$

The null hypothesis of the lagged-series test is that the series being tested, $X(t)$, has a unit root. i.e. that it is an I(1) process. The alternative hypothesis is that the TS is stationary, i.e. it is an I(0) process. This version of the URT is for a driftless TS; later we will consider the case with an intercept an d a linear trend. The Johansen cointegration test has different critical values depending on whether the estimated relationship is assumed to have a constant, a linear trend or is driftless. Here we use the critical values corresponding to the driftless case.

The lagged-series URT is conducted as follows:

- We run the Johansen cointegration test of the TS $\{x_t\}$ and the k-lagged TS $\{L^k(x_t)\}$
- The Johansen test will compute a primary and secondary test statistic when run with two variables; the first test statistic is for the null hypothesis of a cointegration rank of 0. The second test statistic is for a null-hypothesis of a cointegration rank of less than or equal to 1.
- We examine the *second* cointegration test statistic of a rank of less than or equal to 1. If this test statistic is less than or equal to the critical value for the specified significance level α then we do not reject the null of I(1). Otherwise, if the statistic is greater than the critical value we reject the null of I(1) and accept the alternative of I(0).

I explain why this lagged-series URT works. Consider the vector error correcting model:

$$\Delta y_t = \Pi y_{t-1} + \Gamma_1 \Delta y_{t-1} + \Gamma_2 \Delta y_{t-2} + \cdots + \Gamma_{p-1} \Delta y_{t-(p-1)} + \Psi D_t + \epsilon_t \tag{2.7}$$

$$\Pi = -\left(I - \Pi_1 - \Pi_2 - ... - \Pi_p\right) \tag{2.8}$$

I performed a simulation that is summarized in Table 2.1. I ran 100 simulations of an AR(1) process of length 1000 and tested it for cointegration with itself lagged by 3. This cointegration test has two eigenvalues since there are two variables. I computed the mean and standard deviations of the first and second eigenvalues.

- When the process is stationary, I(0), then Π will have a rank of 2. In this case $\Pi_1 = \Pi_2 = ... = \Pi_p = 0$ so $\Pi = -I$, therefore we expect there two be two cointegrating relationships in this trivial case. We see when the multiplier is 0 that both eigen values are near 0.3 and the standard deviation is much smaller.

- When the process is a unit root, I(1), then Π will have a rank of 1. X_t and X_{t-k} will never be too far apart, thus we expect to see a cointegrating relationship. We see when the multiplier is 1 that the first eigenvalue is near 0.5 and the second one is near 0. The standard deviations are small.

- When the multiplier of the AR(1) process is from 0.25 to 0.725 we observe similar behavior as with the 0 multiplier case, however we see that the magnitude of the second eigenvalue continues to decrease as the multiplier increases.

- When the process is an explosive AR(1) with a multiplier of 1.01 the second eigenvalue is larger than with a multiplier of 1.

Table 2.1: Mean and Standard Deviation of Johansen Cointegration Eigenvalues of AR(1) Process with Itself lagged by 3 with a TS with $l = 1000$,$m = 100$ and $s = 1234$

ϕ_1	$\overline{\lambda_1}$	$\overline{\lambda_2}$	S_{λ_1}	S_{λ_2}
0.000	0.343	0.326	0.018	0.016
0.250	0.343	0.233	0.014	0.017
0.500	0.377	0.153	0.016	0.013
0.750	0.436	0.090	0.018	0.010
1.000	0.501	0.003	0.016	0.002
1.010	0.696	0.183	0.008	0.012

To see more intuitively why this test works consider these equivalent AR(1) processes:

$$y_t = \phi_1 y_{t-1} + \epsilon_t \tag{2.9}$$

$$y_{t-3} = \phi_1 y_{t-4} + \epsilon_{t-3} \tag{2.10}$$

And we consider the TS aligned with itself lagged by 3–these are the two TS used in the co-integration test:

$$\begin{cases} y_t, y_{t+1}, ..., y_{t+n} \\ y_{t-3}, y_{t-2}, ..., y_{t+n-3} \end{cases} \tag{2.11}$$

We can now write y_t as a function of y_{t-3}:

$$y_t = \phi_1^3 y_{t-3} + \phi_1^2 \epsilon_{t-2} + \phi_1 \epsilon_{t-1} + \epsilon \tag{2.12}$$

If $\phi_1 = 0$ then the pairing would be between y_t and ϵ_t, and $E(y_t) = 0$. Since ϵ_t is an I(0) process this represents the nonsense case of two possible cointegration cases; the Π matrix with a rank of 2 as discussed earlier.If $\phi_1 = 1$ then $y_t = y_{t-3} + \epsilon_{t-2} + \epsilon_{t-1} + \epsilon$ and therefore $E(y_t|y_{t-3}) = y_{t-3}$ and in this case we would expect there to be a valid cointegrating relationship between y_t and y_{t-3}. If we were to consider an in between ϕ_1 value such as $\phi_1 = 0.8$ the relationship between the series and the lagged value would be $y_t = 0.512 y_{t-3} + 0.64 \epsilon_{t-2} + 0.8 \epsilon_{t-1} + \epsilon$ and the relationship with the lagged value would be weaker.

I performed another similar simulation, however this time the lag was 30 instead of 3. The results are summarized in Table 2.2. Overall the behavior is similar to what we see in Table 2.1. The main difference is that in the case of a multiplier of 1, the unit root case, the mean first eigenvalue is much smaller than when a lag of 3 was used. This simulation would seem to indicate that the lagged-series URT can be useful under various lags. Under certain cases, such as when testing I(3) and I(4) processes for unit roots can lead to errors in matrix computations when using a short lag like 3, however if the lag is increased to 30 these errors disappear. In Table 2.2 we see the mean and standard deviations of the eigen values using a lag of 30, and overall we see a similar pattern to when the lag is 3—however the magnitudes are lower in the case of the unit root (multiplier of 1.)

The Johansen Critical Values for no cointegration ($r = 0$) and one cointegrating relationship between two series ($r \leq 1$) for the maximum eigenvalue test statistic are detailed

Table 2.2: Mean and Standard Deviation of Johansen Cointegration Eigenvalues of AR(1) Process with Itself lagged by 30 with a TS with $l = 1000$,$m = 100$ and $s = 1234$

ϕ_1	$\overline{\lambda_1}$	$\overline{\lambda_2}$	S_{λ_1}	S_{λ_2}
0.000	0.344	0.325	0.017	0.018
0.250	0.283	0.266	0.015	0.016
0.500	0.210	0.195	0.013	0.013
0.750	0.121	0.108	0.010	0.010
1.000	0.036	0.003	0.006	0.002
1.010	0.510	0.019	0.021	0.004

in Table 2.3.

Table 2.3: Critical Values for Johansen Maximum Eigenvalue Test Statistic

Coint Rank	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
r $\leq$ 1	6.50	8.18	11.65
r = 0	12.91	14.90	19.19

We simulate the trace and maximum eigenvalue test statistics from the asymptotic Brownian motion functionals in Equations (1.112), (1.113) and (1.114) for the two variable case with cointegration rank 0 and 1. See the R code Listing 2.1 for details.

Listing 2.1: Simulation of Johansen Brownian Functional for 2 TS

```
> quantile(DF, probs=c(0.01,0.05,0.1))
N <- 400
# time increment
W1 <- numeric(N+1)
W2 <- numeric(N+1)
# initialization of the vector W approximating
# Wiener process
t <- seq(0,T, length=N+1)
```

```
simulations <- 8000
JohansenMaxEigenval <- rep(0,simulations )
JohansenTrace <- rep(0, simulations)
r0stat <- rep(0, simulations)
for (sim in 1:simulations )
{
        dW1 <- rnorm(N)
        dW2 <- rnorm(N)
        W1 <- c(0, cumsum( dW1 ))
        W2 <- c(0, cumsum( dW2 ))
        W1 <- W1 - mean(W1)
        W2 <- W2 - mean(W2)
        intWdW <- matrix(rep(0,4), nrow=2,ncol=2)
        intWWdr  <- matrix(rep(0,4), nrow=2,ncol=2)
        intW1dW1 <- 0
        intW1W1dr  <- 0
        W <- matrix( c(W1[1:N], W2[1:N]), nrow=N, ncol=2 )
        dW <- matrix( c(dW1[1:N], dW2[1:N]), nrow=N, ncol=2 )
        for( i in 1:N )
        {
                intWdW  <- intWdW  + W[i,] %*% t(dW[i,])
                intWWdr <- intWWdr + W[i,] %*% t(W[i,])
                intW1dW1 <- intW1dW1   + W[i,1]*dW[i,1]
                intW1W1dr <- intW1W1dr + W[i,1]*W[i,1]
        }
        stochasticMatrix <- t(intWdW ) %*% solve(intWWdr) %*% intWdW
        JohansenTrace[sim] <- sum(diag(stochasticMatrix))
        JohansenMaxEigenval[sim] <- max(eigen(stochasticMatrix)$values)
        r0stat[sim] <- intW1dW1 * intW1dW1 / intW1W1dr
}
quantile(r0stat,probs=c(0.90,0.95,0.99))
quantile(JohansenTrace,probs=c(0.90,0.95,0.99))
quantile(JohansenMaxEigenval,probs=c(0.90,0.95,0.99)
```

The results of the Brownian functional simulation for the maximum eigenvalue test can be found in Table 2.4 and we see that the results are close to the published critical values in Table 2.3. The results will depend on the choice of time series length, number of simulations and the specific random number generation functionality. I chose a time series length of 400 as it the same as that in Johansen, S. and Juselius, K. (1990). I chose 8000 simulations instead of the 6000 the authors used originally, as this produced closer results to the published critical values.

Table 2.4: Johansen Critical Values from Brownian Functional Simulations

Coint Rank	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
r <= 1	6.51	8.11	11.55
r = 0	12.90	14.75	19.27

I performed simulations of the basic lagged-series approach of testing for cointegration of a TS of length 1000, with it's lagged version with a lag of 3 in Table 2.5. We see that for the I(1) case with $\phi_1 = 1$ for one cointegrating relationship ($r \leq 1$) the **0.9, 0.95 and 0.99 quantiles** are **6.55, 8.24, 12.08** with a **TS length of 1000**. When we use a **TS of length 100** instead as detailed Table 2.7, the corresponding quantiles are close to the previous one; they are: **7, 8.64, 12.04**. These quantiles for both $l = 1000$ and $l = 100$ series are close to the Brownian functional **simulation** derived quantile values of **6.51, 8.11, 11.55** in Table 2.4 where the assumed time series length was $l = 400$, and they are close to the published Johansen critical values in Table 2.3 of **6.50, 8.18, 11.65** used by the ca.jo() function of the urca *R* package (Pfaff, B., 2008). *I use the fact that the (0,9, 0.95 and 0.99) quantiles of the simulated test statistics under the null are close to the published Johansen critical values for a cointegration rank less than or equal to 1 as empirical evidence that we can use the published Johansen critical values for the lagged-series URT.*

In addition we see that if we **increase the lag to 30** the **simulation results** as detailed in Table 2.6 are still very close for at least the 0.9 and 0.95 quantiles and larger for the 0.99 quantile: **6.61, 8.18, 13.07**. *So I argue that the asymptotic behavior of the lagged-series test statistic under the null hypothesis of a unit root for the $r \leq 1$ case is the same as that of the Johansen cointegration test for reasonable lags.*

Table 2.5: Quantiles of Maximum Eigenvalue Test Statistics Run on Various AR(1) Processes and their lagged version(lag=-3) with $l = 1000$, $m = 1000$ and $s = 12345$

	$r = 0$	$r = 0$	$r = 0$	$r \leq 1$	$r \leq 1$	$r \leq 1$
ϕ_1	Q0.9	Q0.95	Q0.99	Q0.9	Q0.95	Q0.99
1.009	1199.76	1203.59	1213.79	222.54	227.87	240.45
1.008	1198.59	1202.45	1211.22	222.20	227.59	239.97
1.005	1160.58	1170.60	1181.11	206.13	215.38	229.24
1.004	1033.03	1063.55	1103.57	170.28	181.39	202.71
1.003	780.70	812.66	874.07	66.58	86.60	112.84
1.002	734.70	744.67	767.12	7.28	10.02	16.68
1.001	732.50	743.67	766.19	5.72	7.14	10.24
1.000	732.68	743.40	766.53	6.55	8.24	12.08
0.990	729.39	739.11	762.40	11.58	13.06	16.29
0.970	719.11	729.60	753.43	22.34	24.23	28.65
0.950	708.21	718.77	744.23	32.53	34.41	39.67
0.900	682.48	693.15	719.65	54.85	58.46	63.91
0.500	502.79	511.58	529.99	186.71	193.20	208.26
0.000	460.19	475.81	496.30	422.25	432.01	446.87

Table 2.6: Quantiles of Maximum Eigenvalue Test Statistics Run on Various AR(1) Processes and their lagged version(lag=-30) with $l = 1000$, $m = 1000$ and $s = 12345$

	$r = 0$	$r = 0$	$r = 0$	$r \leq 1$	$r \leq 1$	$r \leq 1$
ϕ_1	Q0.9	Q0.95	Q0.99	Q0.9	Q0.95	Q0.99
1.009	737.49	750.01	785.66	24.38	26.23	29.57
1.008	736.11	749.59	785.89	24.30	26.19	29.44
1.005	670.24	696.67	732.75	23.72	25.17	29.05
1.004	497.37	548.23	607.40	22.80	24.37	28.36
1.003	156.88	194.11	287.18	19.09	21.44	25.64
1.002	46.33	50.82	64.18	6.93	9.00	14.39
1.001	42.91	46.53	53.26	5.85	7.43	9.96
1.000	43.13	46.18	53.35	6.61	8.18	13.07
0.990	38.30	40.68	44.92	11.19	13.01	17.68
0.970	36.64	38.95	42.65	21.14	23.34	28.59
0.950	42.74	44.97	49.86	31.60	34.07	39.16
0.900	67.30	70.27	77.06	57.34	60.57	65.43
0.500	252.49	261.09	275.52	232.89	239.45	252.53
0.000	443.51	453.11	467.62	414.72	424.12	442.51

Table 2.7: Quantiles of Maximum Eigenvalue Test Statistics Run on Various AR(1) Processes And their lagged version(lag=-3) with $l = 100$, $m = 1000$ and $s = 12345$

ϕ_1	$r = 0$ Q0.90	$r = 0$ Q0.95	$r = 0$ Q0.99	$r \leq 1$ Q0.9	$r \leq 1$ Q0.95	$r \leq 1$ Q0.99
1.009	79.71	84.78	95.35	6.08	7.71	11.50
1.008	79.82	84.77	95.24	6.35	7.64	11.56
1.005	79.88	84.74	95.52	6.45	8.59	11.66
1.004	79.89	84.73	95.54	6.46	8.59	11.83
1.003	79.89	84.73	95.56	6.65	8.70	12.31
1.002	79.90	84.72	95.57	6.90	8.91	12.14
1.001	79.91	84.82	95.58	7.03	8.71	12.34
1.000	79.90	84.80	95.58	7.00	8.64	12.04
0.990	79.76	84.72	95.31	7.09	8.65	12.39
0.970	79.18	84.13	94.51	7.51	8.97	12.79
0.950	78.53	83.28	93.43	8.29	9.92	13.22
0.900	76.18	80.63	91.30	10.41	12.14	15.64
0.500	57.57	61.98	72.54	25.27	27.73	34.52
0.000	58.85	64.15	78.75	45.52	48.19	55.34

Table 2.8: Proportion of Failures to Reject $H_0^{I(1)}$ lagged-series and ADF Tests Run on Various AR(1) Processes with $l = 1000$, $m = 1000$ and $s = 12345$ and $\alpha = 0.01$

ϕ_1	ADF	lagged-series
1.010	1.000	0.000
1.005	0.999	0.054
1.000	0.988	0.983
0.990	0.685	0.904
0.980	0.100	0.553
0.970	0.003	0.142
0.960	0.000	0.016
0.950	0.000	0.000
0.900	0.000	0.000
0.000	0.000	0.000

Table 2.8 shows the proportions of failures to reject the null hypothesis of a unit root ($H_0^{I(1)}$) for various cases of ϕ_1 and for the significance level of 0.01 for both the new lagged-series URT and the ADF URT using the critical values for no trend and no intercept using the adfTest() function of the *fUnitRoots R* package (Wuertz, Diethelm and many others, 2013). We see that for the case of $\phi_1 = 1$ both lagged-series and ADF reject the null hypothesis around 1% of the cases as expected. The lagged-series URT rejects much more than the ADF test when $\phi_1 > 1$; for example when $\phi_1 = 1.01$ lagged-series rejects 100% of the test cases, yet ADF rejects none. When $\phi_1 < 1$ the statistical power of the ADF URT is larger than that of the new lagged-series test–that is ADF rejects the null hypothesis more often than the lagged-series in this scenario.

Table 2.9: Proportion of Failures to Reject $H_0^{I(1)}$ llagged-series and ADF Tests Run on Various AR(1) Processes with $l = 1000$, $m = 1000$ and $s = 12345$ and $\alpha = 0.05$

ϕ_1	ADF	lagged-series
1.010	1.000	0.000
1.005	0.999	0.042
1.000	0.946	0.948
0.990	0.262	0.670
0.980	0.004	0.195
0.970	0.000	0.014
0.960	0.000	0.000
0.950	0.000	0.000
0.900	0.000	0.000
0.000	0.000	0.000

Table 2.9 shows the proportions of failures to reject the null hypothesis of a unit root ($H_0^{I(1)}$) for various cases of ϕ_1 and for the significance level of 0.05 for both the new lagged-series URT and the ADF URT using the critical values for no trend and no intercept. We see that for the case of $\phi_1 = 1$ both lagged-series and ADF reject the null hypothesis around 0.05 of the cases as expected. Just as before, the lagged-series URT rejects much more than the ADF test when $\phi_1 > 1$; for example when $\phi_1 = 1.01$ lagged-series rejects 100% of the

test cases, yet ADF rejects none. Just as in the previous case, when $\phi_1 < 1$ the statistical power of the ADF URT is larger than that of the new lagged-series test.

Table 2.10: Proportion of Failures to Reject $H_0^{I(1)}$ lagged-series and ADF Tests Run on Various AR(1) Processes with $l = 1000$, $m = 1000$ and $s = 12345$ and $\alpha = 0.10$

ϕ_1	ADF	lagged-series
1.010	1.000	0.000
1.005	0.995	0.040
1.000	0.899	0.899
0.990	0.075	0.491
0.980	0.001	0.061
0.970	0.000	0.003
0.960	0.000	0.000
0.950	0.000	0.000
0.900	0.000	0.000
0.000	0.000	0.000

Table 2.10 shows the proportions of failures to reject the null hypothesis of a unit root ($H_0^{I(1)}$) for various cases of ϕ_1 and for the significance level of 0.10 for both the new lagged-series URT and the ADF URT using the critical values for no trend and no intercept. We see that for the case of $\phi_1 = 1$ both lagged-series and ADF reject the null hypothesis around 10% of the cases as expected. Just as with previous significance level tests, the lagged-series URT rejects much more than the ADF test when $\phi_1 > 1$; for example when $\phi_1 = 1.01$ llagged-series rejects 100% of the test cases, yet ADF rejects none. As in the previous two cases for significance levels of 0.05 and 0.01, when $\phi_1 < 1$ the statistical power of the ADF URT is larger than that of the new lagged-series test.

Even though the statistical power of the lagged-series test is less than that of the ADF test we will see later in a test of spurious cointegration of two TS when the lagged-series URT is used to pretest the series before performing a cointegration test, the results are not significantly worse than when the ADF URT is used for pretesting when $\phi_1 < 1$. For detaild see Tables reftbl:UnitRootCointegrationTest1000:01, reftbl:UnitRootCointegrationTest1000:05

and reftbl:UnitRootCointegrationTest1000:10. Section 3.1 contains multiple simulation studies of the lagged-series URT under various model specifications.

2.2 The Lagged-Series Unit Root Test with Constants and/or Linear Trends

In Section 2.1 the new unit root test was proposed for the case when there is no drift. Here a modification is considered to handle a constant and/or a linear trend. We consider an intercept and a linear trend that is added to the auto-regressive process as follows:

$$X_t = \phi_1 X_{t-1} + \epsilon_t \tag{2.13}$$

$$Y_t = X_t + \alpha + \beta t \tag{2.14}$$

Now we re-order the terms

$$X_t = Y_t - \alpha - \beta t \tag{2.15}$$

$$(Y_t - \alpha - \beta t) = \phi_1(Y_{t-1} - \alpha - \beta(t-1)) + \epsilon_t \tag{2.16}$$

$$Y_t = \phi_1 Y_{t-1} + (1-\phi_1)\alpha + (t - \phi_1 t + \phi_1)\beta + \epsilon_t \tag{2.17}$$

This expression can be rewritten as:

$$Y_t = \phi_1 Y_{t-1} + C + Dt + \epsilon_t \tag{2.18}$$

where $C = (1-\phi_1)\alpha + \phi_1\beta$ and $D = \beta(1-\phi_1)$

Now we will run a regression based on first differences to determine ϕ_1^a, C and D terms

:

$$\Delta Y_t = Y_t - Y_{t-1} = \phi_1^a Y_{t-1} + C + Dt + \epsilon_t \tag{2.19}$$

where $\phi_1^a = (\phi_1 - 1)$

And we can finally back out the original auto-regressive model terms ϕ_1, α and β as follows: $\phi_1 = \phi_1^a + 1$, $\beta = D/(1 - \phi_1)$ and $\alpha = (C - \beta\phi_1)/(1 - \phi_1)$.

The Johansen cointegration test has different critical values depending on whether the estimated relationship is assumed to have a constant, a linear trend or is driftless. Here we use the critical values corresponding to the critical values for the trend case. In the driftless version of the test we use the Johansen driftless critical values. Table 2.11 shows some tests of this procedure compared with existing URTs. When $\phi_1 = 1$ all the tests reject the null close to 5% of the test cases as expected given the 0.05 significance level in this simulation test done with 500 simulations. When $\phi_1 > 1$ we see that the lagged-series rejects the null much more often than the other tests; when $\phi_1 = 1.005$ the lagged-series test rejects 95% of the tests while the other URTs in this simulation study reject none. When $\phi_1 < 1$ the lagged-series and ADF tests are very similar, and the ERS tests have higher statistical power in this case.

Section 3.1 contains multiple simulation studies of the lagged-series URT with intercept and linear-trend under various model specifications.

2.3 The ERS-PTest-lagged-series and ZA-lagged-series URTs

I propose a new URT which consists of mixing the Elliott, Rothenberg & Stock test together with the lagged-series Test proposed in the previous section, as detailed in Algorithm 1. This algorithm returns *TRUE* if the hypothesis of a unit root (I(1)) cannot be rejected, or *FALSE* if it is rejected. Using simulation studies I selected a critical value of 20.09299 for the lagged-series test statistic that does not interfere with the null hypothesis of a unit root AR(1) $\phi_1 = 1$ yet it rejects the null when $\phi_1 > 1$; these simulations where performed with a time series of length $l = 1000$. When I combine this lagged-series URT with the ERS-Ptest

Table 2.11: Proportion of Failures to Reject $H_0^{I(1)}$ of URTs on AR(1) Processes with Intercept and Linear Trend with $l = 1000$, $m = 500$, and $\alpha = 0.05$ and $s = 12345$

ϕ_1	ADF	lagged-series	ERS-Ptest	ERS-DFGLS
1.0050	1.00	0.05	1.00	1.00
1.0025	0.97	0.70	0.96	0.97
1.0000	0.94	0.94	0.95	0.95
0.9950	0.92	0.92	0.86	0.89
0.9900	0.82	0.83	0.63	0.70
0.9800	0.45	0.42	0.15	0.20
0.9700	0.15	0.12	0.00	0.01
0.9600	0.02	0.02	0.00	0.00
0.9500	0.00	0.00	0.00	0.00
0.9000	0.00	0.00	0.00	0.00
0.7000	0.00	0.00	0.00	0.00
0.5000	0.00	0.00	0.00	0.00
0.3000	0.00	0.00	0.00	0.00
0.1000	0.00	0.00	0.00	0.00
0.0000	0.00	0.00	0.00	0.00

URT under alternative hypotheses it results in a increase of statistical power relative to the ERS-Ptest as can be seen in Table 2.12. The computations to combine the two tests are detailed in Algorithm 1.

Algorithm 1 ERS-lagged-series URT

$C \leftarrow 20.09299$

$x0 \leftarrow$ LaggedSeriesUnitRootTest(tsx, lag, cvalLevel)@Statistic
$x1 \leftarrow$ ur.ers(tsx,type='P-test',model='trend')@Statistic
if $x1 >$ CV and $x0 < C$ **then**

  return(TRUE)
else

  return(FALSE)
end if

Table 2.12: Proportions of Failures to Reject $H_0^{I(1)}$ ERS-lagged-series URT on a Series with $l = 1000$,$m = 5000$ and $\alpha = 0.10$ Compared to ADF,ERS

ϕ_1	ADF	lagged-series	ERS-Ptest	ERS-DFGLS	ERS-Ptest-lagged-series
1.010	1.0000	0.0040	1.0000	0.0652	0.0008
1.005	0.9974	0.0466	0.9968	0.9968	0.0728
1.000	0.9022	0.8894	0.8970	0.8968	0.8968
0.990	0.6998	0.5064	0.5082	0.5108	0.5082
0.980	0.2862	0.0612	0.0570	0.0662	0.0570
0.970	0.0636	0.0030	0.0020	0.0036	0.0020
0.960	0.0078	0.0000	0.0000	0.0000	0.0000
0.950	0.0080	0.0000	0.0000	0.0000	0.0000
0.900	0.0000	0.0000	0.0000	0.0000	0.0000
0.000	0.0000	0.0000	0.0000	0.0000	0.0000

In a similar fashion the Zivot Andrews (ZA) URT is combined with the lagged-series URT. The details are in Algorithm 2. The deterministic component can incorporate either an intercept, time trend, or both. The same approach could be used to combine the lagged-series URT and the ADF URT. In Section 3.1 there are various simulation experiments performed using this ZA-lagged-seriesURT under various model specifications.

Algorithm 2 ZA-lagged-series URT

```
C ← 20.09299

x0 ← LaggedSeriesUnitRootTest( tsx, lag, cvalLevel)@Statistic
x1 ← ur.za(tsx,model='trend')@Statistic
if x1 > CV and x0 < C  then

  return(TRUE)
else

  return(FALSE)
end if
```

2.4 The Hybrid Bai-Perron Zivot-Andrews (HBPZA) Unit Root Test

URTs assume the null hypothesis that there is at least one unit root present in the TS being tested. One of the most widely used URTs is the Augmented Dickey Fuller (ADF) test.

Perron, Pierre (1990) enhanced the Augmented Dickey-Fuller tests to allow for a known structural break. Given the arbitrariness in determining the break, Andrews, Donald and Zivot, Eric (1992) proposed URTs where the break point was determined 'endogenously' from the data; these are the ZA URTs. Unlike Perron, Pierre (1990)'s URT null hypothesis, these ZA endogenous tests assume there are no breaks under the unit root null. Not allowing breaks in the unit root null can bias these tests to provide evidence of stationarity with breaks (Lee, J. and Strazicich, M.C. (2001)). The one-break Lee, J. and Strazicich, M.C. (2001) procedure and the two-break Lee, J. and Strazicich, M.C. (2003) procedure allows for the breaks to be determined endogenously from the data and breaks are allowed under both the null and the alternative hypothesis.

The ZA URT is detailed in Section 1.11.2. Under the null hypothesis of this test only a unit root process with a deterministic component; no strucutural breaks are allowed under the null.

The ADF URT and many others fail to reject a false null hypothesis of a unit root under the presence of structural changes in intercept and/or linear trend, as can be seen in Tables 3.20, 3.21, and 3.22 in Section 3.1. We can see that the ZA URT and related ZA-lagged-series URT are able to reject false null hypotheses ($\phi_1 \neq 1$) much more significantly than the other tests, however they do not perform well in the valid unit root case ($\phi_1 = 1$) and reject much more than they should. This is because the ZA test does not consider structural breaks in the null hypothesis.

2.4.1 Calculation

I propose a new URT that allows structural breaks under the null hypothesis, which we refer to here are the Bai-Perron-Zivot-Andrews (HBPZA) URT.

1. Given a TS, we first use the Bai-Perron break-point estimation procedure in Bai, J. and Perron, P. (2003) using Regression Model (2.20) to detect changes in the coefficients. This divides the TS into $k+1$ segments given a total of k breakpoints.

2. For each segment within the TS we compute the ZA URT statistic.

3. A final test statistic is computed by weighing each sub-test statistic by the segment length.

4. If there are more than 3 breakpoints estimated, only the first three breakpoints are used to determine 4 segments.

The Regression Model (2.20) is used to detect breakpoints in the coefficients β_0, β_1 and β_2 using the Bai, J. and Perron, P. (2003) procedure:

$$x_t = \beta_0 + \beta_1 t + \beta_2 x_{t-1} + \epsilon_t \tag{2.20}$$

Model (2.21) was used to simulate TS under the null hypothesis (H_0) of a unit root with 1 structural break at time T_B to derive the critical values of the test statistic (via simulation):

$$
\begin{aligned}
y_t = \phi_1 y_{t-1} + \mu_1 + (\mu_2 - \mu_1)\mathbf{1}\{t > T_B\} + (t - T_B)(\beta_2 - \beta 1_1)\mathbf{1}\{t > T_B\} + \beta t + e_t \\
\mu_1 \sim \text{unif}(-10, 10) \; ; \; \mu_2 \sim \text{unif}(-10, 10) \\
\beta_1 \sim \text{unif}(-3, 3) \; ; \; \beta_2 \sim \text{unif}(-3, 3) \\
H_0 : \phi_1 = 1 \\
H_A : \phi_1 \neq 1
\end{aligned}
\tag{2.21}
$$

Model (2.22) was used to simulate TS under the null hypothesis (H_0) of a unit root with 2 structural breaks to derive the critical values of the test statistic (via simulation):

$$
\begin{aligned}
y_t = \phi_1 y_{t-1} + \mu_1 + (\mu_2 - \mu_1)\mathbf{1}\{t > T_B\} + (\mu_3 - \mu_2)\mathbf{1}\{t > T_C\} + \\
+ (t - T_B)(\beta_2 - \beta_1)\mathbf{1}\{t > T_B\} + (t - T_C)(\beta_3 - \beta_2)\mathbf{1}\{t > T_C\} + \beta t + e_t \\
\mu_1 \sim \text{unif}(-10, 10) \; ; \; \mu_2 \sim \text{unif}(-10, 10) \; ; \; \mu_3 \sim \text{unif}(-10, 10) \\
\beta_1 \sim \text{unif}(-3, 3) \; ; \; \beta_2 \sim \text{unif}(-3, 3) \; ; \; \beta_3 \sim \text{unif}(-3, 3) \\
H_0 : \phi_1 = 1 \\
H_A : \phi_1 \neq 1
\end{aligned}
\tag{2.22}
$$

Only 300 simulations for each number of breaks (0,1,2) were chosen to derive the critical values as each one takes at least 30 seconds to run on a desktop machine running an Intel Core I7 processor. The HBPZA critical values for 0, 1 and 2 breakpoints under the null hypothesis of a unit root are detailed in Table 2.13.

Table 2.13: HBPZA Unit Root Critical values from Simulations

Breaks	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
2	-6.021853	-4.414791	-4.151678
1	-6.069573	-4.631409	-4.336683
0	-4.837634	-4.376655	-4.178676

Table 2.14: Zivot-Andrews Critical values with Intercept and Trend

$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$
-5.57	-5.08	-4.82

Table 2.14 details the published ZA critical values with 1 breakpoint used in the urzaTest() function of the *fUnitRoots R package* (Wuertz, Diethelm and many others, 2013). Since the HBPZA test statistic is a estimated break-segment weighted ZA test statistic, there is some validity for comparing the HBPZA and ZA test statistic values. We simulate I(1) TS and compute the quantiles of the ZA test statistics using the R Code in Listing 2.2:

Listing 2.2: R Code to Compute Quantiles of Simulated ZA Test Statistics with an I(1) TS without Breaks

```
library(fUnitRoots)
set.seed(12345)
s <- 500
zaTests <- rep(0,s)
for (i in 1:s )
{
  zaTests[i] <- urzaTest(cumsum(rnorm(1000)),doplot=FALSE,
    model = 'trend')@test$test@teststat
}
quantile( zaTests, probs=c(0.01, 0.05, 0.10))
```

The results of running the R code in Listing 2.2 are as detailed in Results Listing 2.2:

Listing 2.3: Quantiles of Simulated ZA Test Statistics with an I(1) TS without Breaks

```
> quantile( zaTests, probs=c(0.01, 0.05, 0.10))
       1%         5%        10%
-4.909521  -4.430225  -4.183176
```

We see that the **(0.01, 0.05,0.10)** quantiles in Results Listing 2.3 for 0 breaks of **(-4.909521, -4.430225, -4.183176)** are close to the critical values in Table 2.14 for 0 breaks of **(-4.837634, -4.376655, -4.178676)**; this gives us some comfort that the Bai, J. and Perron, P. (2003) estimation procedure error when there are no breaks does not significantly distort the ZA statistic. In addition we see that the (0.01, 0.05, 0.10) critical values of the HBPZA test with 1 break **(-6.069573, -4.631409, -4.336683)** in Table 2.13 are not quite different from the published ZA critical values of **(-5.57, -5.08, -4.82)** in Table 2.14.

Now we perform a simulation test computing the ZA test statistic (0.01, 0.05, 0.10) quantiles when run on an I(0) process without breaks using the *R code* in Listing 2.4:

Listing 2.4: R Code to Compute Quantiles of Simulated ZA Test Statistics with an I(0) TS without Breaks

```
set.seed(12345)
s <- 500
zaTests <- rep(0,s)
for (i in 1:s )
{
         zaTests[i] <- urzaTest(rnorm(1000),
       doplot=FALSE, model = 'trend')@test$test@teststat
}

quantile( zaTests, probs=c(0.01, 0.05, 0.10))
```

Listing 2.5: Quantiles of Simulated ZA Test Statistics with an I(0) TS without Breaks

```
> quantile( zaTests, probs=c(0.01, 0.05, 0.10))
        1%          5%         10%
−20.17250  −19.51512  −19.27001
```

We see that **the values for the I(0) test** in Results Listing 2.5 **(-20.17250, -19.51512, -19.2700) are to the left of the I(1) test** results in Listing 2.3 of **(-4.909521, -4.430225, -4.183176)**.

Table 2.15: HBPZA Quantiles with $\phi_1 = 0.95$ from Simulations

Breaks	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
2	-6.076275	-5.545940	-5.328129
1	-6.302416	-5.830326	-5.531494
0	-6.974804	-6.579809	-6.375370

Table 2.15 contains the HBPZA test statistic (0.01, 0.05, 0.10) quantiles for 0,1 and 2 breaks in intercept and linear trend under the alternative hypothesis where the AR(1) multiplier $\phi_1 = 0.95$ which is somewhat close to a unit root using just as before 300 simulations for each number of breaks (0,1 and 2). We see that every value in Table 2.15 with $\phi_1 = 0.95$ is to the left of the corresponding value on the critical value Table 2.13 with $\phi_1 = 1$. This provides some evidence that the HBPZA test can distinguish to some extent between $\phi = 1$ and $\phi = 0.95$.

Table 2.16: Proportions of Failures to Reject $H_0^{I(1)}$ on a Series with Randomly chosen intercepts(-10,10) and trends(-10,10) $l = 1000$,$m = 200$ and $\alpha = 0.05$

ϕ_1	HBPZA	Breaks
1.000	0.96	2
0.950	0.63	2
0.900	0.09	2
1.000	0.94	1
0.950	0.46	1
0.900	0.00	1
1.000	0.98	0
0.950	0.02	0
0.900	0.00	0

Using Models (2.21) and (2.22) we simulated under various sets of ϕ_1 and ran the HBPZA and for 0,2 and 2 breaks in intercept and linear trend with a significance level $\alpha = 0.05$. The results for a time series length of $l = 1000$ are in Table 2.16. We see that in the null hypothesis case for 2 breaks it rejects 4% of the cases, for 1 break it rejects 6% of the cases, and for 0 breaks it rejects 2% of the cases. The number of simulations is only $m = 200$ for each ϕ_1 as these tests are quite slow–in the order of 30 seconds per simulation on a Pentium Core I7. For the cases when $\phi_1 < 1$ the HBPZA test rejects 37% of the test cases when $\phi_1 = 0.95$ for 2 breaks and rejects 54% for 1 break and 98% of the cases when there are no breaks.

The results of the HBPZA URT results on simulations with a time series of length

Table 2.17: Proportions of Failures to Reject $H_0^{I(1)}$ on a Series with Randomly chosen intercepts[-10,10] and trends[-10,10] $l = 100$,$m = 1000$ and $\alpha = 0.05$

ϕ_1	HBPZA	Breaks
1.000	0.92	1
0.950	0.92	1
0.900	0.91	1
0.800	0.85	1
0.500	0.41	1
0.000	0.01	1
1.000	0.96	0
0.950	0.95	0
0.900	0.92	0
0.800	0.77	0
0.500	0.09	0
0.000	0.00	0

$l = 100$ with 1 break and 0 breaks are summarized in Table 2.17. We see that the results are worse in this case than the previous one with $l = 1000$. The null hypothesis is rejected in 8% of the cases which is higher than it should be for a significance level of 0.05 When there are zero breaks the null hypothesis is rejected in 4% of the cases. In the case of $\phi_1 < 1$ for example for $\phi_1 = 0.8$ under 2 breaks only 15% of the test cases are rejected, and this rejection rate is 23% when there are no breaks.

We see that these HBPZA tests are sensitive to the time series length; the critical values were derived via simulation using $l = 1000$ and we could derive them for the time series length we are interested in testing.

2.5 The Hybrid Bai-Perron-ADF I(0)/I(1) Test Procedure (HBPADF)

Kim (2003) showed that standard unit root tests are not consistent against processes displaying a shift from stationarity (I(0)) to nonstationarity (I(1)) and vice versa. Therefore new methods are required for differentiating between processes that are I(0) or I(1) for the entire period v.s. those with a shift from I(0) to I(1) or vice versa.

Kejriwal, Mohitosh and Perron, Pierre and Zhou, Jing (2013) proposed hybrid testing procedures that allows ruling out of stable stationary processes or ones that are subject to only stationary changes under the null, helping the researcher in interpreting a rejection as emanating from a switch between an I(1) and an I(0) regime. The authors use a combination of their own developed unit root tests together with a test with a null of no breaks vs an alternative with one or more breaks given by Bai, J. and Perron, P. (1998). The calculation of the test statistics and the asymptotic critical values are done using the dynamic programming algorithm proposed in Perron and Qu (2006, Journal of Econometrics 134, 373–399).

Kejriwal, Mohitosh and Perron, Pierre and Zhou, Jing (2013) used the approach of detrending the data when trends are included prior to using their hybrid testing procedure and they suggest using a sequential procedure developed by Kejriwal, Mohitosh and Perron, Pierre and Zhou, Jing (2013) that is robust to whether the errors are I(1) or I(0). The smallest time series length the authors consider is 150.

I propose a testing procedure referred to as the HBPADF procedure whereby given a TS to consider if the series follows one of the following four possibilities:

- The entire series is stationary (I(0))
- The entire series is nonstationary (I(1))
- The series consists of a stationary segment (I(0)) followed by a nonstationary segment (I(1))
- The series consists of a nonstationary segment (I(1)) followed by a stationary segment (I(0))

The idea for this type of testing procedure was suggested by my adviser, Dr. James Gentle. The proposed approach is to use the Bai, J. and Perron, P. (1998) methodology of estimating structural break date/s based on finding the model specification that minimizes the RSS via a dynamic programming approach, and then to use the ADF URT to test each section to determine if it is likely I(0) or I(1). To do this we use the following Regression Model (2.23) with the Bai, J. and Perron, P. (1998) procedure which will estimate structural breaks in the coefficients $\beta_0, \beta_1, \beta_2$:

$$x(t) = \beta_0 + \beta_1 t + \beta_2 x(t-1) + \epsilon_t \tag{2.23}$$

If the Bai, J. and Perron, P. (1998) structural date methodology does not find a break date, then we run the ADF test on the entire time series–and both segments would have the same result.

If we assume that if any structural changes in linear trend and/or intercept happen they do so at the same time as a structural break in AR(1) autoregressive multiplier ϕ_1 then the ADF and the Bai, J. and Perron, P. (1998) methodology were highly accurate, then the ADF test would fit the correct linear trend and intercept in each segment; in any case experiments show that the Bai, J. and Perron, P. (1998) approach can handle changes in β_2 across both I(0) and I(1) regimes, as well as when there are changes in the other coefficients.

$$y_t = \phi_1^A y_{t-1}\mathbf{1}\{t \leq T_B\} + \phi_1^B y_{t-1}\mathbf{1}\{t > T_B\} \tag{2.24}$$

Simulations were performed using the DGP (2.24) which has breaks only in the AR ϕ_1 multiplier with time series of length $l = 1000$ and are summarized in Table (2.18). We see that the approach works well to distinguish between $\phi_1 = 1$ vs. $\phi_1 = 0.9$ when the break happens in the middle of the time series (break proportion=0.5.) When the break happens in the first quarter of the time series (break proportion=0.25) the approach works well to distinguish between an I(1) to I(0) with $\phi_1 = 0.9$ transition, but it is not good to distinguish between an I(0) to I(1) change.

Table 2.18: Proportions of Failures to Reject $H_0^{I(1)}$ on a AR(1) Series with a Break in ϕ_1 $l = 1000$,$m = 100$ and $\alpha = 0.05$

Break Proportion	ϕ_1^A	ϕ_1^B	$\overline{\Gamma_{\text{ADF}}^{\text{A}} > CV_\alpha}$	$\overline{\Gamma_{\text{ADF}}^{\text{B}} > CV_\alpha}$	$\overline{\Gamma_{\text{ADF}}^{\text{All}} > CV_\alpha}$
0.50	0.000	0.000	0.00	0.00	0.00
0.50	1.000	1.000	0.93	0.93	0.94
0.50	1.000	0.500	0.93	0.04	0.84
0.50	0.900	1.000	0.07	0.90	0.86
0.50	1.000	0.900	0.89	0.05	0.82
0.50	1.000	0.950	0.82	0.30	0.81
0.50	1.000	0.980	0.84	0.69	0.86
0.25	0.000	0.000	0.00	0.00	0.00
0.25	1.000	1.000	0.93	0.93	0.94
0.25	1.000	0.500	0.97	0.02	0.61
0.25	0.900	1.000	0.70	0.91	0.94
0.25	1.000	0.900	0.84	0.00	0.43
0.25	1.000	0.950	0.54	0.06	0.42
0.25	1.000	0.980	0.73	0.66	0.72

Table 2.19: Proportions of Failures to Reject $H_0^{I(1)}$ on a AR(1) Series with a Break in ϕ_1, intercept and linear trend with $l = 1000$,$m = 100$ and $\alpha = 0.05$

Break Proportion	ϕ_1^A	ϕ_1^B	μ^A	μ^B	μ_t^A	μ_t^B	$\overline{\Gamma_{\text{ADF}}^{\text{A}} > CV_\alpha}$	$\overline{\Gamma_{\text{ADF}}^{\text{B}} > CV_\alpha}$	$\overline{\Gamma_{\text{ADF}}^{\text{All}} > CV_\alpha}$
0.50	0.000	0.000	10	-30	5	-2	0.00	0.00	1.00
0.50	1.000	1.000	10	-30	5	-2	0.95	0.97	1.00
0.50	1.000	0.500	10	-30	5	-2	0.96	0.00	1.00
0.50	0.900	1.000	10	-30	5	-2	0.04	0.97	1.00
0.50	1.000	0.900	10	-30	5	-2	0.95	0.00	1.00
0.50	1.000	0.950	10	-30	5	-2	0.95	0.12	1.00
0.50	1.000	0.980	10	-30	5	-2	0.95	0.65	1.00
0.25	0.000	0.000	10	-30	5	-2	0.00	0.00	0.03
0.25	1.000	1.000	10	-30	5	-2	0.98	0.96	0.02
0.25	1.000	0.500	10	-30	5	-2	0.98	0.00	0.10
0.25	0.900	1.000	10	-30	5	-2	0.53	0.96	0.00
0.25	1.000	0.900	10	-30	5	-2	0.98	0.00	0.00
0.25	1.000	0.950	10	-30	5	-2	0.98	0.01	0.01
0.25	1.000	0.980	10	-30	5	-2	0.98	0.60	0.01

$$
\begin{aligned}
y_t = \phi_1^A y_{t-1}\mathbf{1}\{t \leq T_B\} + \phi_1^B y_{t-1}\mathbf{1}\{t > T_B\} + \\
+\mu_1 + (\mu_2 - \mu_1)\mathbf{1}\{t > T_B\} + \\
+(t - T_B)(\beta_2 - \beta 1_1)\mathbf{1}\{t > T_B\} + \beta t + e_t \\
\mu_1 = 10 \; ; \; \mu_2 = -30 \; ; \; \beta_1 = 5 \; ; \; \beta_2 = -2
\end{aligned}
\tag{2.25}
$$

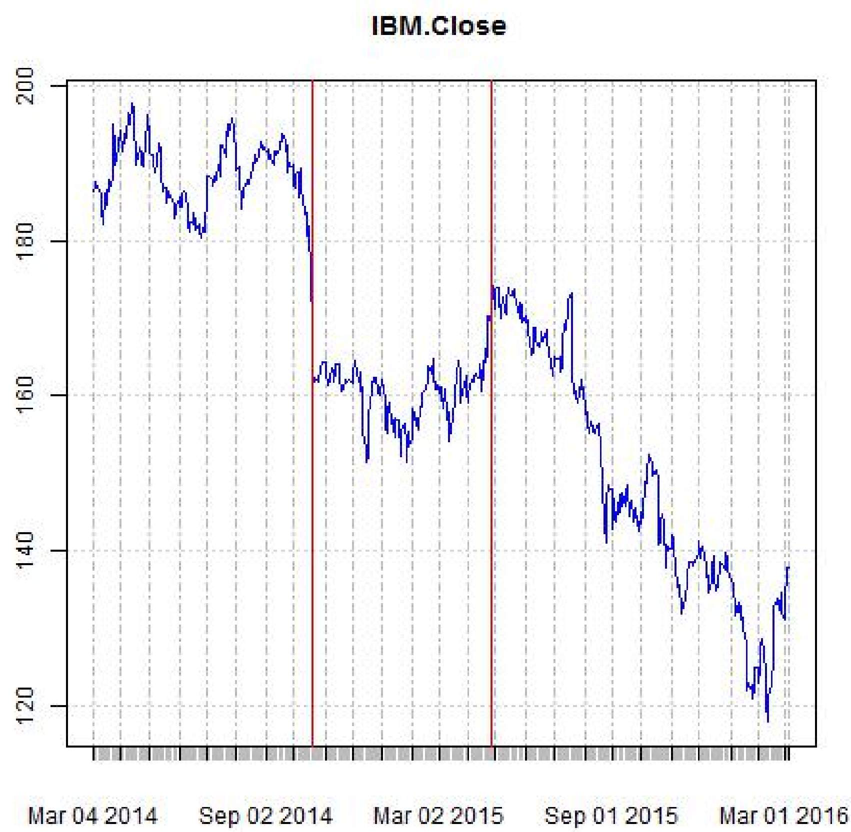
Table 2.19 summarizes the results of the new testing procedure run on simulations were performed with changes in the AR ϕ_1 multiplier as well as in the intercept and linear trend with time series of length $l = 1000$. Just as in the previous trendless case we see that the approach works well to distinguish between $\phi_1 = 1$ vs. $\phi_1 = 0.9$ when the break happens in the middle of the time series (break proportion=0.5.) Also as before when the break happens in the first quarter of the time series (break proportion=0.25) the approach works well to distinguish between an I(1) to I(0) with $\phi_1 = 0.9$ transition, but it is not good to distinguish between an I(0) to I(1), however it performs somewhat better than in the trendless case.

See Section 3.2 for additional simulation tests of the HBPADF procedure using more simulations, as well as tests on a time series of length $l = 100$.

2.5.1 Running the HBPADF I(0)/I(1) Test Procedure on IBM Stock Prices

We run the HBPADFf I(0)/I(1) Test Procedure on IBM Stock Prices from '2014-03-04' until '2016-03-04'. The procedure estimates two breakpoints in the series as can be see in Figure 2.1 as detailed in Results Listing 2.6:

Figure 2.1: Estimated Changepoints in IBM Prices.



Listing 2.6: IBM Estimated Break Dates

```
> IBM_$IBM.Close[bp1$breakpoints]
           IBM.Close
2014-10-20    169.10
2015-04-27    170.73
```

The R code used to provide the estimated changepoints is provided in Listing 2.7.

Listing 2.7: R code to Estimate IBM Breakpoints

```
library(quantmod)
library(strucchange)
library(tseries)
getSymbols('IBM', src='yahoo')
IBM_ <- window( IBM,start='2014-03-04', end='2016-03-04')
ts1 <- IBM_$IBM.Close
l <- length(ts1)
ts_dat <- as.data.frame(merge(ts1, diff(ts1), lag(ts1, k = -1), zoo(1:l)))
ts_dat[,'date'] <- time(ts1)
colnames(ts_dat) <- c('ts1','dts1','ts1l1','trend1','date')
bp1 <- breakpoints(ts1 ~ ts1l1+ trend1, data = ts_dat,h=0.05)
```

The two estimated breakpoints imply 3 segments in the original IBM TS. We will independently test each segment for unit roots using the ADF procedure. The results are provided in Listing 2.8.

Listing 2.8: ADF URTs on IBM Segments

```
> adf.test(IBM_$IBM.Close[1:bp1$breakpoints[1]])
Dickey-Fuller = -1.2919, Lag order = 5, p-value =  0.8717
alternative hypothesis: stationary
> adf.test(IBM_$IBM.Close[bp1$breakpoints[1]:bp1$breakpoints[2]])
Dickey-Fuller = -1.7957, Lag order = 5, p-value = 0.6615
alternative hypothesis: stationary
> adf.test(IBM_$IBM.Close[bp1$breakpoints[2]:l])
Dickey-Fuller = -2.4495, Lag order = 5, p-value = 0.3874
alternative hypothesis: stationary
```

We see in Listing 2.8 that the null hypothesis of a unit root cannot be rejected for any of the three segments, which leads us to conclude they are likely nonstationary: I(1),I(1),I(1). This finding is consistent with the common assumption that asset prices follow a random walk. Even though breakpoints are estimated using the Bai, J. and Perron, P. (2003) procedure the ADF tests show that it is likely the AR(1) multiplier ϕ_1 is 1, even if the drift has a changepoint. Visually we see that there is a level-shift between the first and second segments, and there is likely a linear trend starting in the third component.

Chapter 3: Simulation Studies of Unit Root and Cointegration Tests

3.1 Simulation Studies of Unit Root Tests

This chapter describes the results of various Monte Carlo simulation studies of existing unit root tests (URTs), as well as the new lagged-series and Zivot Andrews-lagged-series URT on AR time series (TS). The URT used were: ADF, ERS Ptest, ERS DFGLS, ZA, lagged-series and ZA-lagged-series. The ADF and ERS tests were chosen as these are often the most widely used under the assumption of no structural breaks. The ZA URT was chosen as this is one of the preferred tests under the assumption of the existence of a structural break. These tests are described in Sections 1.45, 1.10 and 1.11.2. The purpose of these tests was to compare their accuracy with AR(1) processes with $0 < \phi_1 < 1.1$ under various types of deterministic intercepts and components, various assumptions of the distributional characteristics of the innovations, as well as under structural breaks. These tests offer different versions depending on the assumed deterministic form under; these are generally a constant and/or a linear trend. To simplify the number of combinations to test, there are two general configurations. In the case where there is 0 drift and 0 linear trend we select the following versions of the URT:

- lagged-series with no constant/drift.
- ZA-lagged-series with intercept term on Zivot Andrews and lagged-series with no drift.
- ADF with no deterministic component.
- ZA with an intercept.
- ERS-Ptest with a constant.

- ERS-DFGLS with a constant.

In the case where there is either a non zero drift or a non-zero linear trend we select the following versions of the URTs:

- lagged-series with constant and trend.
- ZA-lagged-series with drift and trend term on Zivot Andrews and lagged-series with constant and trend.
- ADF with constant and trend.
- ZA with constant and trend.
- ERS-Ptest with a trend.
- ERS-DFGLS with a trend.

Unless specified the initial value of the simulated series is $x(0) \leftarrow N(0,1)$***.***

More specifically the test scenarios consist of simulations of:

- driftless AR(1) processes where the ϕ_1 multiplier is chosen at random with independent identically distributed Gaussian innovations.
- AR(1) processes with a deterministic intercept and linear trend where the ϕ_1 multiplier is chosen at random with independent identically distributed Gaussian innovations.
- AR(1) processes with a deterministic linear trend where the ϕ_1 multiplier is chosen at random with independent identically distributed Gaussian innovations.
- AR(1) processes with deterministic intercept where the ϕ_1 multiplier is chosen at random with independent identically distributed Gaussian innovations.
- driftless AR(1) processes where the ϕ_1 multiplier is chosen at random with pairwise correlated Gaussian innovations.

- driftless AR(1) processes where the ϕ_1 multiplier is chosen at random with independent identically distributed Beta innovations.

- driftless AR(1) processes where the ϕ_1 multiplier is chosen at random with independent identically distributed Gaussian innovations with random standard deviations picked as the absolute value of Gaussian random variables.

- AR(1) processes with one structural break in intercept and linear trend where the ϕ_1 multiplier is chosen at random with independent identically distributed Gaussian innovations.

- AR(1) processes with with one structural break in linear trend where the ϕ_1 multiplier is chosen at random with independent identically distributed Gaussian innovations.

- AR(1) processes with with one structural break in intercept where the ϕ_1 multiplier is chosen at random with independent identically distributed Gaussian innovations.

- AR(1) with negative $\phi_1 < 1$

- ARMA(1,3) processes

- AR(1) with time series of length 50

- Various AR(2) processes.

- I(2), I(3) and I(4) processes.

- The new Hybrid Bai-Perron-ADF (HBPADF) testing procedure to detect changes in I(0)-I(1) sections within a time series.

3.1.1 Unit Root Tests on Driftless AR(1) TS

Table 3.1 summarizes the results of simulations performed with Model (3.1). This model consists of driftless AR(1) processes with independent identically distributed Gaussian innovations for various values of ϕ_1.

The results can be summarized as follows:

- Data simulated under the alternative hypothesis where $\phi_1 > 1$ ($H_1^{\text{Explosive AR}}$): ERS, ADF and ZA tests do not significantly reject the null hypothesis of even though ideally they should in this case, however the lagged-series and ZA-lagged-series tests significantly reject the null hypothesis of a unit-root in these cases.

- Data simulated under the null Hypothesis where $\phi_1 = 1$ ($H_0^{I(1)}$): The URTs reject the null hypothesis as expected close to the significance level of 5%: ADF 5%, ERS-Ptest 4%, ERS-DFGLS 4% and ZA 6%, lagged-series 5%, ZA-lagged-series 6%

- Data simulated under the alternative hypothesis where $\phi_1 < 1$ ($H_1^{I(0)}$): The null hypothesis is rejected at a proportion of 5% or less as follows from best to worst:
  - Both ERS tests: $\phi_1 \leq 0.98$
  - ADF test: $\phi_1 \leq 0.98$
  - lagged-series test: $\phi_1 \leq 0.97$
  - ZA test: $\phi_1 \leq 0.95$
  - ZA-lagged-series test: $\phi_1 \leq 0.95$

$$\begin{aligned} x(t) = \phi_1 x(t-1) + \epsilon(t) \ ; \ x(0) = 0 \\ \epsilon(t) \sim N(0,1) \\ \operatorname{cor}(\epsilon(t), \epsilon(t-1)) = 0 \ ; \ t = 1, ..., l \end{aligned} \tag{3.1}$$

We can see that these URTs are sensitive to the initial starting value of the TS in Table 3.2 where $x(0) = 20$. Choi, In (2010, p. 52) points out that the initial value of the time series affect the statistical power of the of the URTs. He states that for large initial values the power decreases; that is given the alternative hypothesis is true, the URTs do

Table 3.1: Proportions of Failures to Reject $H_0^{I(1)}$ of URT on Driftless AR(1) Processes with $l = 1000$, $m = 1000$, and $\alpha = 5\%$ and $s = 12345$

ϕ_1	ADF	lagged-series	ERS-Ptest	ERS-DFGLS	ZA-lagged-series	ZA
1.0050	1.00	0.04	1.00	1.00	0.07	1.00
1.0025	0.99	0.67	0.99	0.99	0.85	0.97
1.0000	0.95	0.95	0.95	0.95	0.94	0.94
0.9950	0.70	0.89	0.69	0.68	0.91	0.91
0.9900	0.25	0.71	0.26	0.26	0.86	0.86
0.9800	0.00	0.17	0.00	0.00	0.67	0.67
0.9700	0.00	0.01	0.00	0.00	0.32	0.33
0.9600	0.00	0.00	0.00	0.00	0.10	0.11
0.9500	0.00	0.00	0.00	0.00	0.02	0.02
0.9000	0.00	0.00	0.00	0.00	0.00	0.00
0.7000	0.00	0.00	0.00	0.00	0.00	0.00
0.5000	0.00	0.00	0.00	0.00	0.00	0.00
0.3000	0.00	0.00	0.00	0.00	0.00	0.00
0.1000	0.00	0.00	0.00	0.00	0.00	0.00
0.0000	0.00	0.00	0.00	0.00	0.00	0.00

not reject the null hypothesis of a unit root as often as they should. Muller, Ulrich K. and Elliott, Graham (2003) find that the optimal URT for when the initial value is large in the Dickey-Fuller t-ratio, however when the initial value is small the optimal test is the ERS-DFGLS test. In Table 3.2 we see that for a starting value of 20 the ERS tests have quite low power as they do not reject any cases when $\phi_1 < 1$. The lagged-series URT performs slightly better than ADF when $\phi_1 < 1$ and much better than ADF when $\phi_1 > 1$.

In Table 3.3 we see that for a starting value of -10 the ERS tests have quite low power as they do not reject any cases when $\phi_1 < 1$, just as was the case with the previous example with $x(0) = 20$. The lagged-series URT performs slightly better than ADF when $\phi_1 < 1$ and much better than ADF when $\phi_1 > 1$. However we see that all tested URTs reject the null hypothesis of $\phi_1 = 1$ at more than the expected 5% level.

In Table 3.4 shows simulations of Model(3.1) using negative values of ϕ_1. The results for data simulated under the alternative hypothesis where $\phi_1 < 0$ ($H_1^{I(0)}$) can be summarized

Table 3.2: Proportions of Failures to Reject $H_0^{I(1)}$ of URTs on Driftless AR(1) Processes with $l = 1000$, $m = 1000$, and $\alpha = 5\%$, $s = 12345$ and initial value of 20

ϕ_1	ADF	lagged-series	ERS-Ptest	ERS-DFGLS	ZA-lagged-series	ZA
1.010	1.00	0.00	1.00	0.00	0.00	1.00
1.005	1.00	0.01	1.00	1.00	0.01	1.00
1.000	0.94	0.95	0.96	0.96	0.94	0.95
0.990	0.78	0.75	0.99	0.98	0.90	0.90
0.980	0.29	0.21	1.00	1.00	0.60	0.64
0.970	0.04	0.01	1.00	1.00	0.18	0.22
0.960	0.00	0.00	1.00	1.00	0.02	0.03
0.950	0.00	0.00	1.00	1.00	0.00	0.00
0.900	0.00	0.00	1.00	1.00	0.00	0.00
0.000	0.00	0.00	1.00	1.00	0.00	0.00

Table 3.3: Proportions of Failures to Reject $H_0^{I(1)}$ of URTs on AR(1) Processes with $X(0) = -10$ with $l = 1000$, $m = 200$, and $\alpha = 5\%$ and $s = 12345$

ϕ_1	ADF	lagged-series	ERS-Ptest	ERS-DFGLS	ZA-lagged-series
1.010	1.00	0.00	1.00	1.00	0.00
1.005	0.99	0.04	1.00	1.00	0.06
1.000	0.90	0.91	0.93	0.91	0.93
0.990	0.16	0.55	0.93	0.93	0.81
0.980	0.00	0.10	0.98	0.97	0.61
0.970	0.00	0.00	0.98	0.99	0.22
0.960	0.00	0.00	0.99	0.99	0.04
0.950	0.00	0.00	1.00	1.00	0.00
0.900	0.00	0.00	1.00	1.00	0.00
0.000	0.00	0.00	0.99	1.00	0.00

Table 3.4: Proportions of Failures to Reject $H_0^{I(1)}$ of URTs on Driftless AR(1) Processes with $\phi_1 < 0$ and $l = 1000$, $m = 1000$, and $\alpha = 5\%$, $s = 12345$

ϕ_1	ADF	lagged-series	ERS-Ptest	ERS-DFGLS	ZA-lagged-series	ZA
1.0000	0.95	0.95	0.95	0.95	0.94	0.94
-0.9950	0.00	0.00	0.68	0.05	0.00	0.00
-0.9900	0.00	0.00	0.26	0.06	0.00	0.00
-0.9800	0.00	0.00	0.01	0.06	0.00	0.00
-0.9700	0.00	0.00	0.00	0.05	0.00	0.00
-0.9600	0.00	0.00	0.00	0.05	0.00	0.00
-0.9500	0.00	0.00	0.00	0.05	0.00	0.00
-0.9000	0.00	0.00	0.00	0.05	0.00	0.00
-0.7000	0.00	0.00	0.00	0.02	0.00	0.00
-0.5000	0.00	0.00	0.00	0.01	0.00	0.00
-0.3000	0.00	0.00	0.00	0.01	0.00	0.00
-0.1000	0.00	0.00	0.00	0.01	0.00	0.00
0.0000	0.00	0.00	0.00	0.00	0.00	0.00

as follows: The null hypothesis is rejected as follows from best to worst:

- ADF, lagged-series, ZA and ZA-lagged-series reject the null for all tested negative values of ϕ_1 clearly outperforming the ERS tests.
- ERS-DFGLS: The null hypothesis rejection rate is not higher than 95% of the test cases when $-0.995 \leq \phi_1 \leq -0.9$
- ERS-Ptest: This test fails to significantly reject the null for $\phi_1 = \{-0.995, -0.99\}$ only 32% and 74% of the test cases respectively.

3.1.2 URTs on AR(1) TS with Deterministic Intercept and a Linear Trend

Table 3.5 summarizes the results of simulations performed with Model (3.2). This model consists of AR(1) processes with a deterministic intercept and a linear trend with independent identically distributed Gaussian innovations for various values of ϕ_1.

The results can be summarized as follows:

- Data simulated under the alternative hypothesis where $\phi_1 > 1$ ($H_1^{\text{Explosive AR}}$): The ERS, ADF and ZA tests do not significantly reject the null hypothesis of even though ideally they should in this case. However the two lagged-series based tests reject 95% of the time for $\phi_1 = 1.005$.

- Data simulated under the null Hypothesis where $\phi_1 = 1$ ($H_0^{I(1)}$): The URTs reject the null hypothesis as expected close to the significance level of 5%: ADF, and lagged-series tests at 6%, ERS-Ptest 5%, ERS-DFGLS 5% and ZA 6%

- Data simulated under the alternative hypothesis where $\phi_1 < 1$ ($H_1^{I(0)}$): The null hypothesis is rejected at a rate of 5% or less as follows from best to worst:
  - Both ERS tests: $\phi_1 \leq 0.97$
  - ADF and lagged-series tests: $\phi_1 \leq 0.96$
  - ZA and ZA-lagged-series tests: $\phi_1 \leq 0.95$

$$\begin{aligned} y(t) &= \phi_1 x(t) + \alpha + \beta t \\ x(t) &= x(t-1) + \epsilon(t) \ ; \ x(0) = 0 \\ \epsilon(t) &\sim N(0,1) \ ; \ \text{cor}(\epsilon(t), \epsilon(t-1)) = 0 \ ; \ t = 1, ..., l \\ \alpha &\sim unif(-1,1) \ ; \ \beta \sim unif(-1,1) \end{aligned} \tag{3.2}$$

In Table 3.6 shows simulations of Model (3.2) using negative values of ϕ_1. The results for data simulated under the alternative hypothesis where $\phi_1 < 0$ ($H_1^{I(0)}$) can be summarized as follows: The null hypothesis is rejected as follows from best to worst:

- ADF, lagged-series, ZA and ZA-lagged-series reject the null for all tested negative values of ϕ_1 clearly outperforming the ERS tests.

Table 3.5: Proportions of Failures to Reject $H_0^{I(1)}$ of URTs on AR(1) Processes with Intercept and Linear Trend with $l = 1000$, $m = 500$, and $\alpha = 5\%$ and $s = 12345$

ϕ_1	ADF	lagged-series	ERS-Ptest	ERS-DFGLS	ZA-lagged-series	ZA
1.0050	1.00	0.05	1.00	1.00	0.07	1.00
1.0025	0.97	0.70	0.96	0.97	0.81	0.96
1.0000	0.94	0.94	0.95	0.95	0.94	0.94
0.9950	0.92	0.92	0.86	0.89	0.93	0.93
0.9900	0.82	0.83	0.63	0.70	0.89	0.89
0.9800	0.45	0.42	0.15	0.20	0.69	0.70
0.9700	0.15	0.12	0.00	0.01	0.34	0.36
0.9600	0.02	0.02	0.00	0.00	0.11	0.12
0.9500	0.00	0.00	0.00	0.00	0.02	0.02
0.9000	0.00	0.00	0.00	0.00	0.00	0.00
0.7000	0.00	0.00	0.00	0.00	0.00	0.00
0.5000	0.00	0.00	0.00	0.00	0.00	0.00
0.3000	0.00	0.00	0.00	0.00	0.00	0.00
0.1000	0.00	0.00	0.00	0.00	0.00	0.00
0.0000	0.00	0.00	0.00	0.00	0.00	0.00

Table 3.6: Proportions of Failures to Reject $H_0^{I(1)}$ of URTs on Driftless AR(1) Processes with $\phi_1 < 0$ and $l = 1000$, $m = 1000$, and $\alpha = 5\%$, $s = 12345$

ϕ_1	ADF	lagged-series	ERS-Ptest	ERS-DFGLS	ZA-lagged-series	ZA
1.0000	0.95	0.93	0.94	0.95	0.94	0.94
-0.9950	0.00	0.00	1.00	0.34	0.00	0.00
-0.9900	0.00	0.00	1.00	0.19	0.00	0.00
-0.9800	0.00	0.00	1.00	0.08	0.00	0.00
-0.9700	0.00	0.00	1.00	0.03	0.00	0.00
-0.9600	0.00	0.00	1.00	0.02	0.00	0.00
-0.9500	0.00	0.00	1.00	0.01	0.00	0.00
-0.9000	0.00	0.00	0.87	0.01	0.00	0.00
-0.7000	0.00	0.00	0.23	0.01	0.00	0.00
-0.5000	0.00	0.00	0.01	0.00	0.00	0.00
-0.3000	0.00	0.00	0.00	0.00	0.00	0.00
-0.1000	0.00	0.00	0.00	0.00	0.00	0.00
0.0000	0.00	0.00	0.00	0.00	0.00	0.00

- ERS-DFGLS: The null hypothesis rejection rate is not higher than 95% of the test cases when $-0.995 \leq \phi_1 \leq -0.98$
- ERS-Ptest: This test fails does not reject at all the null for $-0.995 \leq \phi_1 \leq -0.95$ and only 13% for $\phi_1 = -0.9$ and 77% of the test cases when $\phi_1 = -0.7$.

3.1.3 URTs on ARMA(1,3) TS with Deterministic Intercept and a Linear Trend

$$
\begin{gathered}
y_t = \phi_1 x_t + \alpha + \beta t \\
x_t = x_{t-1} + \epsilon_t + 0.5\epsilon_{t-1} + 0.4\epsilon_{t-2} - 0.2\epsilon_{t-3} \ ; \ x_0 = 0 \\
\epsilon_t \sim N(0,1) \ ; \ \mathrm{cor}(\epsilon_t, \epsilon_{t-1}) = 0 \ ; \ t = 1, ..., l \\
\alpha \sim unif(-1,1) \ ; \ \beta \sim unif(-1,1)
\end{gathered}
\tag{3.3}
$$

Table 3.7: Proportions of Failures to Reject $H_0^{I(1)}$ of URTs on ARMA(1,3) Processes having Intercept and Trend with $\phi_1 < 0$ and $l = 1000$, $m = 1000$, and $\alpha = 5\%$, $s = 12345$

ϕ_1	ADF	lagged-series	ERS-Ptest	ERS-DFGLS	ZA-lagged-series	ZA
1.0050	1.00	0.06	1.00	1.00	0.07	1.00
1.0025	0.98	0.72	0.96	0.97	0.80	0.92
1.0000	0.96	0.91	0.93	0.93	0.87	0.87
0.9950	0.95	0.88	0.84	0.84	0.82	0.82
0.9900	0.88	0.76	0.63	0.63	0.72	0.72
0.9800	0.57	0.31	0.13	0.14	0.41	0.41
0.9700	0.22	0.07	0.01	0.01	0.14	0.15
0.9600	0.06	0.01	0.00	0.00	0.03	0.03
0.9500	0.01	0.00	0.00	0.00	0.00	0.00
0.9000	0.00	0.00	0.00	0.00	0.00	0.00
0.7000	0.00	0.00	0.00	0.00	0.00	0.00
0.5000	0.00	0.00	0.00	0.00	0.00	0.00
0.3000	0.00	0.00	0.00	0.00	0.00	0.00
0.1000	0.00	0.00	0.00	0.00	0.00	0.00
0.0000	0.00	0.00	0.00	0.00	0.00	0.00

Table 3.7 summarizes the results of simulations performed with Model (3.3). This model consists of ARMA(1,3) processes with a deterministic intercept and a linear trend with independent identically distributed Gaussian innovations for various values of ϕ_1.

The results can be summarized as follows:

- Data simulated under the alternative hypothesis where $\phi_1 > 1$ ($H_1^{\text{Explosive AR}}$): The ERS, ADF and ZA tests do not significantly reject the null hypothesis of even though ideally they should in this case. However the two lagged-series based tests reject at least 93% of the test cases for $\phi_1 = 1.005$.

- Data simulated under the null hypothesis $\phi_1 = 1$ ($H_0^{I(1)}$): Only the ADF rejects $\leq 5\%$ of the test cases as expected with a significance level $\alpha = 0.05$. The lagged-series test rejects 9% of the test cases, the ERS tests reject 7% of the cases and the ZA based tests reject 13% of the test cases.

- Data simulated under the alternative hypothesis where $\phi_1 < 1$ ($H_1^{I(0)}$): The null hypothesis is rejected at a rate of 5% or less as follows from best to worst:
  - Both ERS tests: $\phi_1 \leq 0.97$
  - lagged-series, ZA and ZA lagged-series tests: $\phi_1 \leq 0.96$
  - ADF and lagged-series tests: $\phi_1 \leq 0.95$

Given that the null hypothesis is rejected more often than it should, a solution to better handle this case would be to use the standard critical values for $\alpha = 0.01$ as the critical value for $\alpha = 0.05$, and the standard critical value for $\alpha = 0.05$ as the critical value for $\alpha = 0.10$.

3.1.4 URTs on AR(1) TS with a Deterministic Linear Trend

Table 3.8 summarizes the results of simulations performed with Model (3.4). This model consists of AR(1) processes with a deterministic linear trend with independent identically distributed Gaussian innovations for various values of ϕ_1.

The results can be summarized as follows:

- Data simulated under the alternative hypothesis where $\phi_1 > 1$ ($H_1^{\text{Explosive AR}}$): The ERS, ADF and ZA tests do not significantly reject the null hypothesis of even though ideally they should in this case. The Lagges-Series URT rejects 27% of the cases when $\phi_1 = 1.005$ when none of the other standard tests reject more than 4%(the ZA-lagged-series test rejects 16% of the cases when $\phi_1 = 1.005$.)

- Data simulated under the null Hypothesis where $\phi_1 = 1$ ($H_0^{I(1)}$): The URTs reject the null hypothesis as expected close to the significance level of 5%: ADF 5%, ERS-Ptest 6%, ERS-DFGLS 4% and ZA 4%

- Data simulated under the alternative hypothesis where $\phi_1 < 1$ ($H_1^{I(0)}$): The null hypothesis is rejected at a rate of 5% of the test cases or less as follows from best to worst:

  - All tests: $\phi_1 \leq 0.90$

$$
\begin{aligned}
y(t) &= x(t) + \beta t \\
x(t) &= \phi_1 x(t-1) + \epsilon(t) \ ; \ x(0) = 0 \\
\epsilon(t) \sim N(0,1) \ ; \ \mathrm{cor}(\epsilon(t), \epsilon(t-1)) &= 0 \ ; \ t = 1, ..., l \\
\beta &\sim unif(-1, 1)
\end{aligned}
\tag{3.4}
$$

3.1.5 URTs on AR(1) TS with a Deterministic Intercept

Table 3.9 summarizes the results of simulations performed with Model (3.5). This model consists of AR(1) processes with a deterministic intercept with independent identically distributed Gaussian innovations for various values of ϕ_1.

Table 3.8: Proportions of Failures to Reject $H_0^{I(1)}$ of URTs on AR(1) Processes with a Linear Trend using $l = 1000$, $m = 500$, and $\alpha = 5\%$ and $s = 12345$

ϕ_1	ADF	lagged-series	ERS-Ptest	ERS-DFGLS	ZA-lagged-series	ZA
1.0050	0.98	0.73	0.98	0.98	0.84	0.96
1.0025	0.95	0.94	0.94	0.95	0.94	0.94
1.0000	0.95	0.95	0.94	0.96	0.95	0.95
0.9950	0.95	0.95	0.93	0.94	0.94	0.94
0.9900	0.93	0.93	0.88	0.90	0.93	0.93
0.9800	0.83	0.84	0.68	0.72	0.89	0.90
0.9700	0.69	0.69	0.39	0.46	0.83	0.83
0.9600	0.52	0.49	0.16	0.22	0.70	0.72
0.9500	0.35	0.29	0.06	0.10	0.56	0.57
0.9000	0.02	0.00	0.00	0.00	0.02	0.02
0.7000	0.00	0.00	0.00	0.00	0.00	0.00
0.5000	0.00	0.00	0.00	0.00	0.00	0.00
0.3000	0.00	0.00	0.00	0.00	0.00	0.00
0.1000	0.00	0.00	0.01	0.00	0.00	0.00
0.0000	0.00	0.00	0.01	0.00	0.00	0.00

The results can be summarized as follows:

- Data simulated under the alternative hypothesis where $\phi_1 > 1$ ($H_1^{\text{Explosive AR}}$): The ERS, ADF and ZA tests do not significantly reject the null hypothesis of even though ideally they should in this case. However the lagged-series with $\phi_1 = 1.005$ rejects in 95% of the cases.

- Data simulated under the null hypothesis where $\phi_1 = 1$ ($H_0^{I(1)}$): The URTs reject the null hypothesis as expected close to the significance level of 5%: ADF 6%, ERS-Ptest 5%, lagged-series and ZA-lagged-series 6%, ERS-DFGLS 5% and ZA 6%

- Data simulated under the alternative hypothesis where $\phi_1 < 1$ ($H_1^{I(0)}$): The null hypothesis is rejected at a rate of 5% of the test cases or less as follows from best to worst:

  - Both ERS tests: $\phi_1 \leq 0.97$

– ADF, lagged-series: $\phi_1 \leq 0.96$

– ZA and ZA-lagged-series test: $\phi_1 \leq 0.95$

$$y(t) = x(t) + \alpha$$

$$x(t) = \phi_1 x(t-1) + \epsilon(t)$$

$$x(0) = 0 \tag{3.5}$$

$$\epsilon(t) \sim N(0,1) \; ; \; \mathrm{cor}(\epsilon(t), \epsilon(t-1)) = 0 \; ; \; t = 1, ..., l$$

$$\alpha \sim unif(-1, 1)$$

Table 3.9: Proportions of Failures to Reject $H_0^{I(1)}$ of URTs on AR(1) Processes with an Intercept using $l = 1000$, $m = 500$, and $\alpha = 5\%$ and $s = 12345$

ϕ_1	ADF	lagged-series	ERS-Ptest	ERS-DFGLS	ZA-lagged-series	ZA
1.0050	1.00	0.05	1.00	1.00	0.07	1.00
1.0025	0.97	0.70	0.96	0.97	0.81	0.96
1.0000	0.94	0.94	0.95	0.95	0.94	0.94
0.9950	0.92	0.92	0.88	0.89	0.93	0.93
0.9900	0.82	0.83	0.67	0.70	0.89	0.89
0.9800	0.45	0.42	0.17	0.20	0.69	0.70
0.9700	0.15	0.12	0.00	0.01	0.34	0.36
0.9600	0.02	0.02	0.00	0.00	0.11	0.12
0.9500	0.00	0.00	0.00	0.00	0.02	0.02
0.9000	0.00	0.00	0.00	0.00	0.00	0.00
0.7000	0.00	0.00	0.00	0.00	0.00	0.00
0.5000	0.00	0.00	0.00	0.00	0.00	0.00
0.3000	0.00	0.00	0.00	0.00	0.00	0.00
0.1000	0.00	0.00	0.00	0.00	0.00	0.00
0.0000	0.00	0.00	0.00	0.00	0.00	0.00

3.1.6 URTs on Driftless AR(1) TS with Consecutive Pairwise Correlated Innovations

Tables 3.10,3.11 and 3.12 summarize the results of simulations performed with Model (3.6), for values of $\rho = 0.9999$, $\rho = 0.8$, and $\rho = -0.8$. This model consists of AR(1) processes with a deterministic intercept with consecutive pairwise correlated Gaussian innovations for various values of ϕ_1.

$$\begin{gathered} x(t) = \phi_1 x(t-1) + \epsilon(t) \ ; \ x(0) = 0 \\ \epsilon(t) \sim N(0,1) \ ; \ \mathrm{cor}(\epsilon(t), \epsilon(t-1)) = \rho \ ; \ t \leftarrow 1, ..., l \end{gathered} \tag{3.6}$$

The results can be summarized as follows. For $\mathrm{cor}(\epsilon(t), \epsilon(t-1)) = 0.9999$, summarized in Table 3.10 the null hypothesis of a unit root is barely ever rejected. This is an extreme case of highly correlated innovations but we can see how they certainly affect the results. Masuda, Junya and Ohtani, Kazuhiro (2008, p. 361) also conclude that with AR(1) time series lengths of up to 100 with serially correlated errors with high correlations, the null hypothesis of a unit root is never rejected.

For the case of $\mathrm{cor}(\epsilon(t), \epsilon(t-1)) = -0.8$ summarized in Table 3.11:

- Data simulated under the alternative hypothesis where $\phi_1 > 1$ ($H_1^{\text{Explosive AR}}$): All ERS, ADF, lagged-series, ZA and ZA-lagged-series URTs do not significantly reject the null hypothesis of even though ideally they should in this case.

- Data simulated under the alternative hypothesis where $\phi_1 < 1$ ($H_1^{I(0)}$): The null hypothesis fails to be rejected at a rate of 5% of the test cases or less as follows from best to worst:

  - ADF test: $\phi_1 \leq 0.97$

  - ZA, lagged-series and ZA-lagged-series tests: $\phi_1 \leq 0.90$

Table 3.10: Proportions of Failures to Reject $H_0^{I(1)}$ of URTs on Driftless AR(1) Processes with Consecutive Innovations Correlated by 0.9999 with $l = 1000$, $m = 500$, and $\alpha = 5\%$ and $s = 12345$

ϕ_1	ADF	lagged-series	ERS-Ptest	ERS-DFGLS	ZA-lagged-series	ZA
1.0050	1.00	0.86	0.99	1.00	1.00	1.00
1.0025	0.99	0.84	0.96	0.99	0.99	0.99
1.0000	0.95	0.90	0.92	0.94	0.93	0.93
0.9950	0.90	0.90	0.94	0.90	0.93	0.93
0.9900	0.90	0.91	0.95	0.89	0.92	0.92
0.9800	0.90	0.92	0.97	0.89	0.92	0.92
0.9700	0.90	0.92	0.97	0.89	0.92	0.92
0.9600	0.90	0.93	0.97	0.89	0.93	0.93
0.9500	0.90	0.92	0.97	0.89	0.93	0.93
0.9000	0.90	0.92	0.97	0.89	0.93	0.93
0.7000	0.90	0.92	0.95	0.89	0.94	0.94
0.5000	0.89	0.93	0.94	0.92	0.94	0.94
0.3000	0.89	0.93	0.93	0.92	0.94	0.94
0.1000	0.88	0.93	0.92	0.93	0.94	0.94
0.0000	0.88	0.93	0.93	0.93	0.94	0.94

- ERS-Ptest: $\phi_1 \leq 0.7$
- ERS-DFGLS test: never rejects less than 8% of the cases and for the range $0.9 \leq \phi_1 \leq 0$ as ϕ_1 decreases the rejection rate of the null hypothesis decreases as well making this the worst performing test in this case.

Table 3.11: Proportions of Failures to Reject $H_0^{I(1)}$ of URTs on Driftless AR(1) Processes with Consecutive Innovations Correlated by -0.8 with $l = 500$, $m = 1000$, and $\alpha = 5\%$ and $s = 12345$

ϕ_1	ADF	lagged-series	ERS-Ptest	ERS-DFGLS	ZA-lagged-series	ZA
1.0050	0.99	0.65	0.99	0.99	0.85	0.99
1.0025	0.98	0.97	0.98	0.98	0.96	0.96
1.0000	0.95	0.95	0.96	0.95	0.95	0.95
0.9950	0.86	0.93	0.87	0.84	0.94	0.94
0.9900	0.68	0.89	0.72	0.69	0.94	0.94
0.9800	0.24	0.70	0.43	0.39	0.88	0.88
0.9700	0.03	0.41	0.23	0.22	0.81	0.81
0.9600	0.00	0.18	0.15	0.14	0.68	0.68
0.9500	0.00	0.06	0.10	0.11	0.50	0.50
0.9000	0.00	0.00	0.06	0.08	0.01	0.01
0.7000	0.00	0.00	0.03	0.12	0.00	0.00
0.5000	0.00	0.00	0.01	0.19	0.00	0.00
0.3000	0.00	0.00	0.01	0.27	0.00	0.00
0.1000	0.00	0.00	0.00	0.32	0.00	0.00
0.0000	0.00	0.00	0.00	0.35	0.00	0.00

For $\text{cor}(\epsilon(t), \epsilon(t-1)) = 0.8$ summarized in Table 3.12:

- Data simulated under the alternative hypothesis where $\phi_1 > 1$ ($H_1^{\text{Explosive AR}}$): The ERS, ADF and ZA tests do not significantly reject the null hypothesis of even though ideally they should in this case. The lagged-series test rejects 94% of the tests cases $\phi_1 = 1.005$ which is significantly more than the other tests.
- Data simulated under the null hypothesis where $\phi_1 = 1$ ($H_0^{I(1)}$): The URTs reject the null hypothesis as expected close to the significance level of 5%: ADF, ERS-Ptest, ERS-DFGLS, and lagged-series at 5% and ZA and ZA-lagged-series at 6%.

- Data simulated under the alternative hypothesis where $\phi_1 < 1$ ($H_1^{I(0)}$): The null hypothesis fails to be rejected at a rate of 5% of the test cases or less as follows from best to worst:
  - ADF test and both ERS tests: $\phi_1 \leq 0.98$
  - lagged-series test: $\phi_1 \leq 0.97$
  - ZA and ZA-lagged-series test: $\phi_1 \leq 0.90$

The results with a correlation of 0.8 are slightly worse in the sense of lower statistical power than the results with a 0 pairwise correlation detailed in Table 3.1.

Table 3.12: Proportions of Failures to Reject $H_0^{I(1)}$ of URTs on Driftless AR(1) Processes with Consecutive Innovations Correlated by 0.8 with $l = 1000$, $m = 500$, and $\alpha = 5\%$ and $s = 12345$

ϕ_1	ADF	lagged-series	ERS-Ptest	ERS-DFGLS	ZA-lagged-series	ZA
1.0050	1.00	0.06	1.00	1.00	0.08	1.00
1.0025	0.99	0.67	0.99	0.98	0.85	0.97
1.0000	0.95	0.95	0.95	0.95	0.94	0.94
0.9950	0.70	0.90	0.69	0.69	0.91	0.91
0.9900	0.27	0.76	0.26	0.28	0.87	0.88
0.9800	0.01	0.24	0.01	0.01	0.71	0.71
0.9700	0.00	0.04	0.00	0.00	0.44	0.44
0.9600	0.00	0.00	0.00	0.00	0.20	0.20
0.9500	0.00	0.00	0.00	0.00	0.06	0.07
0.9000	0.00	0.00	0.00	0.00	0.00	0.00
0.7000	0.00	0.00	0.00	0.00	0.00	0.00
0.5000	0.00	0.00	0.00	0.00	0.00	0.00
0.3000	0.00	0.00	0.00	0.00	0.00	0.00
0.1000	0.00	0.00	0.00	0.00	0.00	0.00
0.0000	0.00	0.00	0.00	0.00	0.00	0.00

3.1.7 URTs on Driftless AR(1) TS with Beta Distributed Innovations

Table 3.13 summarizes the results of simulations performed with Model (3.7). This model consists of AR(1) processes with a deterministic intercept with independent identically distributed Beta(5,2) innovations for various values of ϕ_1.

$$\begin{aligned} x(t) = \phi_1 x(t-1) + \epsilon(t) \ ; \ x(0) = 0 \\ \epsilon(t) \sim B(5,2) \ ; \ \text{cor}(\epsilon(t), \epsilon(t-1)) = 0 \ ; \ t = 1, ..., l \end{aligned} \tag{3.7}$$

The results can be summarized as follows:

- Data simulated under the alternative hypothesis where $\phi_1 > 1$ ($H_1^{\text{Explosive AR}}$): The ERS, ADF and ZA tests do not significantly reject the null hypothesis of even though ideally they should in this case. However the lagged-series URT rejects 95% , and the ZA-lagged-series 93% of the tests when $\phi_1 = 1.005$.

- Data simulated under the null hypothesis where $\phi_1 = 1$ ($H_0^{I(1)}$): The unit root tests reject the null hypothesis as expected close to the significance level of 5%: ADF 5%, ERS-Ptest 4%, ERS-DFGLS 4% and ZA, lagged-series and the ZA-lagged-series URTs 6%.

- Data simulated under the alternative hypothesis where $\phi_1 < 1$ ($H_1^{I(0)}$): The null hypothesis fails to be rejected at a rate of 5% of the test cases or less as follows from best to worst:

  - ADF and both ERS tests: $\phi_1 \leq 0.98$
  - lagged-series test: $\phi_1 \leq 0.97$
  - ZA and ZA-lagged-series test: $\phi_1 \leq 0.95$

3.1.8 URTs on Driftless AR(1) TS with Gaussian Innovations with Random Standard Deviations

Table 3.14 summarizes the results of simulations performed with Model (3.8). This model consists of AR(1) processes with independent identically distributed Gaussian innovations

Table 3.13: Proportions of Failures to Reject $H_0^{I(1)}$ of Unit Root Tests on Driftless AR(1) Processes with Beta(5,2) Innovations with $l = 1000$, $m = 1000$, and $\alpha = 5\%$ and $s = 12345$

ϕ_1	ADF	lagged-series	ERS-Ptest	ERS-DFGLS	ZA-lagged-series	ZA
1.0050	1.00	0.05	1.00	1.00	0.07	1.00
1.0025	0.99	0.69	0.99	0.99	0.85	0.98
1.0000	0.94	0.94	0.94	0.94	0.94	0.94
0.9950	0.69	0.88	0.68	0.68	0.91	0.91
0.9900	0.25	0.71	0.25	0.27	0.87	0.87
0.9800	0.00	0.19	0.00	0.00	0.66	0.66
0.9700	0.00	0.01	0.00	0.00	0.35	0.35
0.9600	0.00	0.00	0.00	0.00	0.10	0.11
0.9500	0.00	0.00	0.00	0.00	0.02	0.02
0.9000	0.00	0.00	0.00	0.00	0.00	0.00
0.7000	0.00	0.00	0.00	0.00	0.00	0.00
0.5000	0.00	0.00	0.00	0.00	0.00	0.00
0.3000	0.00	0.00	0.00	0.00	0.00	0.00
0.1000	0.00	0.00	0.00	0.00	0.00	0.00
0.0000	0.00	0.00	0.00	0.00	0.00	0.00

with a random standard deviation parameter that is the absolute value of a Gaussian distributed variable for various values of ϕ_1.

$$
\begin{gathered}
x(t) = x(t-1) + \epsilon(t) \; ; \; x(0) = 0 \\
\epsilon(t) \sim N(0, a) \; ; \; a \sim |N(0,1)| \; ; \; t = 1, ..., l \\
\mathrm{cor}(\epsilon(t), \epsilon(t-1)) = 0 \; ;
\end{gathered}
\tag{3.8}
$$

The results can be summarized as follows:

- Data simulated under the alternative hypothesis where $\phi_1 > 1$ ($H_1^{\text{Explosive AR}}$): The ERS, ADF and ZA tests do not significantly reject the null hypothesis of even though ideally they should in this case. However the lagged-series URT rejects 95% and the ZA-lagged-series URT rejects 92% when $\phi_1 = 1.005$.

Table 3.14: Proportions of Failures to Reject $H_0^{I(1)}$ of URTs on Driftless AR(1) Processes with N(0,—N(0,1)—) Innovations with $l = 1000$, $m = 1000$, and $\alpha = 5\%$ and $s = 12345$

ϕ_1	ADF	lagged-series	ERS-Ptest	ERS-DFGLS	ZA-lagged-series	ZA
1.0050	1.00	0.05	1.00	1.00	0.08	1.00
1.0025	0.99	0.68	0.99	0.98	0.83	0.97
1.0000	0.95	0.95	0.94	0.94	0.93	0.93
0.9950	0.67	0.86	0.67	0.66	0.93	0.93
0.9900	0.26	0.69	0.26	0.26	0.89	0.89
0.9800	0.00	0.18	0.01	0.01	0.69	0.69
0.9700	0.00	0.01	0.00	0.00	0.37	0.38
0.9600	0.00	0.00	0.00	0.00	0.11	0.12
0.9500	0.00	0.00	0.00	0.00	0.01	0.01
0.9000	0.00	0.00	0.00	0.00	0.00	0.00
0.7000	0.00	0.00	0.00	0.00	0.00	0.00
0.5000	0.00	0.00	0.00	0.01	0.00	0.00
0.3000	0.00	0.00	0.00	0.01	0.00	0.00
0.1000	0.00	0.00	0.00	0.01	0.00	0.00
0.0000	0.00	0.00	0.00	0.01	0.00	0.00

- Data simulated under the null hypothesis where $\phi_1 = 1$ ($H_0^{I(1)}$): The URTs reject the null hypothesis as expected close to the significance level of 5%: ADF 5%, ERS-Ptest 6%, ERS-DFGLS 6% and ZA 7%

- Data simulated under the alternative hypothesis where $\phi_1 < 1$ ($H_1^{I(0)}$): The null hypothesis fails to be rejected at a rate of 5% of the test cases or less as follows from best to worst:

  - ADF and both ERS tests: $\phi_1 \leq 0.98$
  - lagged-series test: $\phi_1 \leq 0.96$
  - ZA and ZA-lagged-series tests: $\phi_1 \leq 0.95$

3.1.9 URTs with ADF and Lagged-Series on a Time Series of Length 50

We use DGP 3.9 to perform simulation studies on short TS of length 50 with the lagged-series and the ADF URTs.

$$
\begin{aligned}
y(t) &= \phi_1 x(t) + \alpha + \beta t \\
x(t) &= x(t-1) + \epsilon(t) \ ; \ x(0) = 0 \\
\epsilon(t) \sim N(0,1) \ ; \ \mathrm{cor}(\epsilon(t), \epsilon(t-1)) &= 0 \ ; \ t = 1, ..., l \\
\alpha \sim unif(-1000, 1000) \ ; \ \beta &\sim unif(-100, 100)
\end{aligned}
\tag{3.9}
$$

Table 3.15: Quantiles of $\hat{\phi_1}$ Estimated Via OLS on AR(1) Processes with Intercept and Trend with $l = 50$, $m = 5000$ and $s = 12345$

ϕ_1	Q:0	Q:0.25	Q:0.50	Q:0.75	Q:1
1	0.21	0.75	0.82	0.88	1.07
0	-0.47	-0.13	-0.04	0.05	0.45

Table 3.15 shows the quantiles of the OLS (conditional-least-squares) estimated AR(1) multiplier ϕ_1 using DGP 3.9 for two cases when $\phi_1 = 1$ (I(1)) and $\phi_1 = 0$ (I(0)). We see that there is considerable variance in both cases; this is expected since it is a very short time series.

Table 3.16: Proportions Rejections of $H_0^{I(1)}$ with ADF URT AR(1) Processes with Intercept and Trend with $l = 50$, $m = 5000$ and $s = 12345$

ϕ_1	0.01	0.05	0.10
1	0.01	0.05	0.09
0	0.24	0.56	0.73

Table 3.16 shows the proportions to reject the null hypothesis of a unit root using the ADFURT for two scenarios of ϕ_1. We see that this test rejects as expected under the null of $\phi_1 = 1$. In the alternative hypothesis case of $\phi_1 = 0$ADF does not reject the null as much as ideally; even with $\alpha = 0.10$ it does not reject in 27% of the test cases.

Table 3.17: Proportions Rejections of $H_0^{I(1)}$ with lagged-series URT AR(1) Processes with Intercept and Trend with $l = 50$, $m = 5000$ and $s = 12345$

ϕ_1	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$
1	0.02	0.08	0.14
0	0.65	0.93	0.98

Table 3.17 shows the proportions to reject the null hypothesis of a unit root using the lagged-series URT. We see that this test in this circumstance rejects more than it should under the null of $\phi_1 = 1$. In the $\phi_1 = 0$ it rejects considerably more than the ADF case.

Table 3.18: Proportions Rejections of $H_0^{I(1)}$ with lagged-series URT with adjusted CVs on AR(1) Processes with Intercept and Trend with $l = 50$, $m = 5000$ and $s = 12345$

ϕ_1	$\alpha = 0.05$	$\alpha = 0.10$
1	0.02	0.08
0	0.65	0.93

We recommend that when using the lagged-series URT if the time series length is short, less than 300, to use the 0.01 standard CV as the 0.05 CV, and the standard 0.05 CV as the 0.10 CV. Table 3.18 shows the results of doing this. We see that the lagged-series used this way has a higher statistical-power than the ADF test: when $\phi_1 = 0$ with $\alpha = 0.05$ it rejects 65% of the test cases as opposed to 56% of the cases with ADF and with $\alpha = 0.10$ it rejects 93% of the test cases as opposed to 73% of the cases with ADF .

3.1.10 URTs on AR(1) TS with Structural Breaks in Intercept

Tables 3.19, and 3.20 summarize the results of simulations performed with Models (3.10). These models consist of AR(1) processes with one structural break in deterministic intercept where the break time is determined at random and with independent identically distributed Gaussian innovations for various values of ϕ_1. There are two sets of tests, the first one has an intercept of -1 before the structural break and an intercept of 1 after the break ($\beta_0 = 1.$).

$$
\begin{aligned}
y(t) &= x(t) + \alpha_t \\
x(t) = \phi_1 x(t-1) + \epsilon(t) \;;\; x(0) &= 0 \;;\; \epsilon(t) \sim N(0,1) \\
\mathrm{cor}(\epsilon(t), \epsilon(t-1)) &= 0 \\
\alpha_t &= \begin{cases} \beta_0, & \text{if } t \geq t_u \\ -\beta_0, & \text{otherwise} \end{cases} \\
t_u \sim \mathrm{unif}(3, l-2) \;;\; t &= 1, ..., l
\end{aligned}
\tag{3.10}
$$

The simulation results detailed in Table 3.19 for the intercept pair of -1 before the break and 1 after the break, can be summarized as follows:

- Data simulated under the alternative hypothesis where $\phi_1 > 1$ ($H_1^{\text{Explosive AR}}$): The ERS, ADF and ZA tests do not significantly reject the null hypothesis of even though ideally they should in this case. However, The lagged-series URT rejects 95% and the ZA-lagged-series URT rejects 93% of the cases when $\phi_1 = 1.005$.

- Data simulated under the null hypothesis where $\phi_1 = 1$ ($H_0^{I(1)}$): The URTs reject the null hypothesis as expected close to the significance level of 5%: ADF 5%, ERS-Ptest 6%, ERS-DFGLS 6%, lagged-series and ZA-lagged-series 6% and ZA 5%.

Table 3.19: Proportions of Failures to Reject $H_0^{I(1)}$ of URTs on AR(1) Processes with one Structural Break in Intercept (-1,1) with $l = 1000$, $m = 1000$, and $\alpha = 5\%$ and $s = 12345$

ϕ_1	ADF	lagged-series	ERS-Ptest	ERS-DFGLS	ZA-lagged-series	ZA
1.0050	1.00	0.05	1.00	1.00	0.07	1.00
1.0025	0.98	0.75	0.98	0.98	0.84	0.97
1.0000	0.95	0.94	0.95	0.95	0.94	0.95
0.9950	0.92	0.93	0.90	0.90	0.94	0.94
0.9900	0.85	0.86	0.71	0.73	0.91	0.91
0.9800	0.49	0.48	0.18	0.23	0.71	0.71
0.9700	0.16	0.14	0.01	0.01	0.41	0.42
0.9600	0.03	0.01	0.00	0.00	0.13	0.15
0.9500	0.01	0.00	0.00	0.00	0.02	0.03
0.9000	0.00	0.00	0.00	0.00	0.00	0.00
0.7000	0.00	0.00	0.00	0.00	0.00	0.00
0.5000	0.00	0.00	0.00	0.00	0.00	0.00
0.3000	0.00	0.00	0.00	0.00	0.00	0.00
0.1000	0.00	0.00	0.00	0.00	0.00	0.00
0.0000	0.00	0.00	0.00	0.00	0.00	0.00

- Data simulated under the alternative hypothesis where $\phi_1 < 1$ ($H_1^{I(0)}$): The null hypothesis fails to be rejected at a rate of 5% of the test cases or less as follows from best to worst:
  - Both ERS tests: $\phi_1 \leq 0.97$
  - ADF and lagged-series tests: $\phi_1 \leq 0.96$
  - ZA and ZA-lagged-series tests: $\phi_1 \leq 0.95$

The simulation results for the intercept pair of -10 before the break and 10 after the break ($\beta_0 = 10$) detailed in Table 3.20 can be summarized as:

- Data simulated under the alternative hypothesis where $\phi_1 > 1$ ($H_1^{\text{Explosive AR}}$): ERS-Ptest, ADF and ZA tests do not significantly reject the null hypothesis of even though ideally they should in this case. However, The lagged-series URT rejects 94% and the ZA-lagged-series URT rejects 93% of the cases when $\phi_1 = 1.005$. The ERS-DFGLS

Table 3.20: Proportions of Failures to Reject $H_0^{I(1)}$ of URTs on AR(1) Processes with one Structural Break in Intercept (-10,10) with $l = 1000$, $m = 200$, and $\alpha = 5\%$ and $s = 12345$

ϕ_1	ADF	lagged-series	ERS-Ptest	ERS-DFGLS	ZA-lagged-series	ZA
1.010	1.00	0.00	1.00	0.10	0.01	1.00
1.005	1.00	0.06	1.00	1.00	0.07	0.99
1.000	0.92	0.92	0.94	0.93	0.71	0.71
0.990	0.92	0.92	0.93	0.92	0.51	0.51
0.980	0.88	0.88	0.87	0.86	0.19	0.19
0.970	0.85	0.85	0.83	0.83	0.03	0.03
0.960	0.84	0.83	0.79	0.81	0.00	0.00
0.950	0.84	0.82	0.76	0.81	0.00	0.00
0.900	0.89	0.82	0.67	0.83	0.00	0.00
0.000	0.93	0.91	1.00	1.00	0.00	0.00

rejects the null 90% of the cases when $\phi_1 = 1.01$.

- Data simulated under the null hypothesis where $\phi_1 = 1$ ($H_0^{I(1)}$): The URTs reject the null hypothesis more than the significance level of 5%: ADF and lagged-series at 8%, ERS-Ptest 6%, ERS-DFGLS 7%, and ZA-lagged-series and ZA 29%.

- Data simulated under the alternative hypothesis where $\phi_1 < 1$ ($H_1^{I(0)}$): The null hypothesis never fails to be rejected at a rate of 5% of the test cases or less for the ADF, lagged-series and ERS tests. However for $\phi_1 < 0.97$ the ZA and ZA-lagged-series URTs fully reject the null hypothesis of a unit root.

This set of tests cases is the worst performing so far, and only the ZA test has reasonable behavior under alternative hypotheses but performs poorly under the case of the null hypothesis of a unit root (with breaks.) This result is understandable as the Zivot Andrews (ZA) URT does not allow breaks under the null; it only allows breaks under the alternative hypothesis. Authors, as surveyed by Choi, In (2010, p. 60), note that a stationary AR process with breaks in intercepts and/or linear trends can be indistinguishable from a unit root with a linear trend. The seminal work in unit root testing under structural breaks was done by Perron, Pierre (1989).

3.1.11 URTs on AR(1) TS with Structural Breaks in Linear Trend

Table 3.21 summarizes the results of simulations performed with Model (3.11). This model consists of AR(1) processes with one structural break in intercept with independent identically distributed Gaussian innovations for various values of ϕ_1. The break time is determined at random; it is uniformly distributed.

$$\begin{aligned} y(t) &= x(t) + \beta_1 t + (\beta_2 - \beta_1)(t - t_u)\mathbf{1}\{t > t_u\} \\ x(t) &= \phi_1 x(t-1) + \epsilon(t) \ ; \ x(0) = 0 \ ; \ \epsilon(t) \sim N(0,1) \\ \mathrm{cor}(\epsilon(t), \epsilon(t-1)) &= 0 \ ; \ \beta_1 = -1 \ ; \ \beta_2 = 1 \\ t_u &\sim \mathrm{unif}(3, l-2) \ ; \ t = 1, ..., l \end{aligned} \tag{3.11}$$

Table 3.21: Proportions of Failures to Reject $H_0^{I(1)}$ of URTs on AR(1) Processes with one Structural Break in Linear Trend (-1,1) with $l = 1000$, $m = 500$, and $\alpha = 5\%$ and $s = 12345$

ϕ_1	ADF	lagged-series	ERS-Ptest	ERS-DFGLS	lagged-series-ZA	ZA
1.010	1.00	0.02	1.00	0.87	0.01	0.99
1.005	1.00	0.82	0.99	0.99	0.12	0.25
1.000	0.94	0.95	0.93	0.93	0.01	0.01
0.990	0.93	0.94	0.91	0.91	0.01	0.01
0.980	0.92	0.93	0.90	0.90	0.01	0.01
0.970	0.91	0.92	0.90	0.90	0.01	0.01
0.960	0.91	0.92	0.90	0.90	0.00	0.00
0.950	0.91	0.92	0.90	0.90	0.00	0.00
0.900	0.91	0.92	0.90	0.90	0.00	0.00
0.000	0.91	0.92	0.90	0.90	0.00	0.00

The results can be summarized as follows:

- Data simulated under the alternative hypothesis where $\phi_1 > 1$ ($H_1^{\text{Explosive AR}}$): The ERS, ADF and ZA tests do not significantly reject the null hypothesis of even though ideally they should in this case. The lagged-series URT rejects 98% of the test cases, and the ZA-lagged-series URT rejects 99% of the cases when $\phi_1 = 1.01$.

- Data simulated under the null hypothesis where $\phi_1 = 1$ ($H_0^{I(1)}$): Three of the four URTs reject the null hypothesis as expected close to the significance level of 5%: ADF 6%, lagged-series 5%, ERS-Ptest 7%, ERS-DFGLS 7%. However the ZA URT and ZA-lagged-series URT have a rejection rate of **99%** , which is quite poor.

- Data simulated under the alternative hypothesis where $\phi_1 < 1$ ($H_1^{I(0)}$): The null hypothesis is never significantly rejected with te ADF, both ERS and lagged-series URTs regardless of the value of ϕ_1. All of these URTs perform quite poorly in the case of a structural break of a linear trend. The ZA and the ZA-lagged-series close to fully reject all test cases which makes them work very well under the alternative hypothesis,

These results are similar yet worse than those in Table 3.20 with breaks in intercept. Just as before, the results are understandable as the Zivot Andrews (ZA) URT does not allow breaks under the null; it only allows breaks under the alternative hypothesis.

3.1.12 URTs on AR(1) TS with Structural Breaks in Intercept and Linear Trend

Tables 3.22 and 3.23 summarize the results of simulations performed with Model (3.12). This model consists of AR(1) processes with one structural break in intercept and one structural break in linear trend with independent identically distributed Gaussian innovations for various values of ϕ_1. The break time is determined at random; it is uniformly distributed, and the same break time is used for the intercept change and the linear trend change. Is it tested with two different set of intercept and trend levels.

$$y(t) = x(t) + \alpha_1 + (\alpha_2 - \alpha_1)\mathbf{1}\{t > t_u\} + \beta_1 t + (\beta_2 - \beta_1)(t - t_u)\mathbf{1}\{t > t_u\}$$

$$x(t) = \phi_1 x(t-1) + \epsilon(t) \; ; \; x(0) = 0 \; ; \; \epsilon(t) \sim N(0,1) \; ; \; \text{cor}(\epsilon(t), \epsilon(t-1)) = 0 \tag{3.12}$$

$$t_u \sim \text{unif}(3, l-2) \; ; \; t = 1, ..., l$$

Table 3.22: Proportions of Failures to Reject $H_0^{I(1)}$ of URTs on AR(1) Processes with one Structural Break in Intercept U(-1,1) and Linear Trend U(-1,1) with $l = 1000$, $m = 500$, and $\alpha = 5\%$ and $s = 12345$

ϕ_1	ADF	lagged-series	ERS-Ptest	ERS-DFGLS	ZA-lagged-series	ZA
1.010	1.00	0.01	1.00	0.48	0.00	0.99
1.005	1.00	0.44	0.99	0.99	0.10	0.57
1.000	0.97	0.97	0.97	0.98	0.13	0.13
0.990	0.94	0.95	0.91	0.93	0.09	0.09
0.980	0.88	0.89	0.85	0.86	0.05	0.05
0.970	0.85	0.85	0.83	0.84	0.03	0.03
0.960	0.84	0.84	0.82	0.83	0.01	0.01
0.950	0.84	0.84	0.82	0.83	0.00	0.00
0.900	0.84	0.83	0.83	0.83	0.00	0.00
0.000	0.88	0.87	0.86	0.87	0.00	0.00

The results can be summarized as follows. For Model (3.12) with the smaller intercepts and trends we pick them uniformly random for each simulation as follows: $\alpha_1, \alpha_2, \beta_1, \beta_2 \sim \text{unif}(-1, 1)$:

- Data simulated under the alternative hypothesis where $\phi_1 > 1$ ($H_1^{\text{Explosive AR}}$): The ERS, ADF and ZA tests do not significantly reject the null hypothesis of even though ideally they should in this case. However the lagged-series test rejects 99% of the test cases and the ZA-lagged-series test rejects 100% of the cases when $\phi_1 = 1.01$.

- Data simulated under the null hypothesis where $\phi_1 = 1$ ($H_0^{I(1)}$): Four of the unit root tests reject the null hypothesis as expected close to the significance level of 5%: ADF 3%, ERS-Ptest 3%, lagged-series 3% and ERS-DFGLS 2%. However the ZA and the ZA-lagged-series tests have a rejection rate of **87%** under the null. The ZA related tests are the worst performing test under the null hypothesis.

- Data simulated under the alternative hypothesis where $\phi_1 < 1$ ($H_1^{I(0)}$): The null hypothesis is never significantly rejected with the ADF, the lagged-series and both ERS tests regardless of the value of ϕ_1. The ZA and the ZA-lagged-series tests reject the null hypothesis 95% or more of the cases when $\phi_1 \leq 0.98$.

Similar to previous structural change cases, the results are understandable as the Zivot Andrews (ZA) URT does not allow breaks under the null; it only allows breaks under the alternative hypothesis. And when using standard unit root tests such as the ADF it is not possible to distinguish a stationary process with structural breaks in the linear trend from a unit root with a drift(trend) as in Choi, In (2010, p. 60).

For Model (3.12) with the larger deterministically chosen intercepts $\alpha_1 = -10$, $\alpha_2 = 10$ and trends $\beta_1 = 10$ and $\beta_2 = -10$ the results can be summarized as follows:

- Data simulated under the alternative hypothesis where $\phi_1 > 1$ ($H_1^{\text{Explosive AR}}$): The ERS, ADF and ZA tests do not significantly reject the null hypothesis of even though ideally they should in this case. However the lagged-series test rejects 88% of the test cases and the ZA-lagged-series test rejects 97% of the cases when $\phi_1 = 1.01$.

- Data simulated under the null hypothesis where $\phi_1 = 1$ ($H_0^{I(1)}$): Four of the unit root tests reject the null hypothesis is not close to the significance level of 5%: ADF 8%, ERS-Ptest 10%, lagged-series 8% and ERS-DFGLS 10%. The ZA and the ZA-lagged-series tests have a rejection rate of **100%** under the null. The ZA related tests are the worst performing test under the null hypothesis.

- Data simulated under the alternative hypothesis where $\phi_1 < 1$ ($H_1^{I(0)}$): The null hypothesis is never significantly rejected with the ADF, the lagged-series and both

ERS tests regardless of the value of ϕ_1. The ZA and the ZA-lagged-series tests reject the null hypothesis 100% of the cases when $\phi_1 \leq 1$.

Table 3.23: Proportions of Failures to Reject $H_0^{I(1)}$ of URTs on AR(1) Processes with one Structural Break in Intercept(10,-10) and Linear Trend (-10,10) with $l = 1000$, $m = 500$, and $\alpha = 5\%$ and $s = 12345$

ϕ_1	ADF	lagged-series	ERS-Ptest	ERS-DFGLS	ZA-lagged-series	ZA
1.010	1.00	0.12	1.00	0.98	0.03	0.92
1.005	0.96	0.94	0.97	0.97	0.02	0.05
1.000	0.92	0.92	0.90	0.90	0.00	0.00
0.990	0.92	0.92	0.90	0.90	0.00	0.00
0.980	0.91	0.92	0.90	0.90	0.00	0.00
0.970	0.91	0.92	0.90	0.90	0.00	0.00
0.960	0.91	0.92	0.90	0.90	0.00	0.00
0.950	0.91	0.92	0.90	0.90	0.00	0.00
0.900	0.91	0.92	0.90	0.90	0.00	0.00
0.000	0.91	0.92	0.90	0.90	0.00	0.00

Just as in the previous case, the results are understandable as the Zivot Andrews (ZA) URT does not allow breaks under the null; it only allows breaks under the alternative hypothesis. And the standard unit root tests such as the ADF cannot distinguish a stationary process with structural breaks in the linear trend from a unit root with a drift(trend) as in Choi, In (2010, p. 60).

3.1.13 Comparison of Lee Strazicich and HBPZA URT under Breaks in Intercept and Trend

The Lee Strazicich (LS) URT (Lee, J. and Strazicich, M.C., 2003) and the Hybrid Bai-Perron Zivot-Andrews (HBPZA) URT proposed in this thesis both allow strucutural breaks of the intercept and/or linear trend under the null hypothesis of a unit root. Simulations were preformed with AR(1) TS with one structural break using Model (3.12) with $\alpha_1 = 50, \alpha_2 = 1000, \beta_1 = 1, \beta_2 = 3$ under various levels of ϕ_1. The break times were 25% and

50% of the time series length. Two sets of tests were performed, one with time series of length 100 summarized in Table 3.24, and one with time series length 500 summarized in Table 3.25.

Table 3.24: Proportions of Failures to Reject $H_0^{I(1)}$ of URTs on AR(1) Processes with 1 break in intercept and trend with $l = 100$, $m = 500$, and $\alpha = 5\%$ and $s = 123456$

ϕ_1	HBPZA	LS	Breaks	Breakpoint Proportion	α_1	α_2	β_1	β_2
1.01	0.97	0.08	1	0.25	50	1000	1	3
1.00	0.96	0.05	1	0.25	50	1000	1	3
0.97	0.95	0.07	1	0.25	50	1000	1	3
0.95	0.96	0.06	1	0.25	50	1000	1	3
0.90	0.95	0.05	1	0.25	50	1000	1	3
0.50	0.51	0.04	1	0.25	50	1000	1	3
0.25	0.03	0.02	1	0.25	50	1000	1	3
0.00	0.00	0.01	1	0.25	50	1000	1	3
1.01	0.95	0.51	1	0.50	50	1000	1	3
1.00	0.95	0.49	1	0.50	50	1000	1	3
0.97	0.96	0.50	1	0.50	50	1000	1	3
0.95	0.96	0.51	1	0.50	50	1000	1	3
0.90	0.95	0.49	1	0.50	50	1000	1	3
0.50	0.37	0.33	1	0.50	50	1000	1	3
0.25	0.04	0.24	1	0.50	50	1000	1	3
0.00	0.00	0.20	1	0.50	50	1000	1	3

The summary of results with time series length 100 is as follows:

- Data simulated under the alternative hypothesis where $\phi_1 > 1$ ($H_1^{\text{Explosive AR}}$): The HBPZA test does not significantly reject the null hypothesis of a unit root even though ideally they should in this case. However the Lee-Strazicich(LS) test rejects 92% of the test cases when $\phi_1 = 1.01$ and the break time is 25; when the break time is 50 it only rejects 49% of the test cases.

- Data simulated under the null hypothesis where $\phi_1 = 1$ ($H_0^{I(1)}$): The HBPZA test rejects the null hypothesis close to the significance level of 5% for both break times of 50 and 25. However the Lee-Srazicich test rejects **95%** of the test cases under the null when the break time is 25 and 51% when the break time is 50.

- Data simulated under the alternative hypothesis where $\phi_1 < 1$ ($H_1^{I(0)}$): When the break time is 25 the Lee-Strazicich significantly outperforms the HBPZA test in this scenario up to $\phi_1 = 0.25$ when both tests become similar. However when the break time is 50 the Lee-Strazicich test does not reject as much as before in this scenario and the HBPZA performs better than with a break time of 25.

Table 3.25: Proportions of Failures to Reject $H_0^{I(1)}$ of URTs on AR(1) Processes with 1 break in intercept and trend with $l = 500$, $m = 500$, and $\alpha = 5\%$ and $s = 123456$

ϕ_1	HBPZA	LS	Breaks	Breakpoint Proportion	α_1	α_2	β_1	β_2
1.01	0.13	0.45	1	0.25	50	1000	1	3
1.00	0.99	0.00	1	0.25	50	1000	1	3
0.97	0.99	0.00	1	0.25	50	1000	1	3
0.95	0.96	0.00	1	0.25	50	1000	1	3
0.90	0.73	0.00	1	0.25	50	1000	1	3
0.50	0.00	0.00	1	0.25	50	1000	1	3
0.25	0.00	0.00	1	0.25	50	1000	1	3
0.00	0.00	0.00	1	0.25	50	1000	1	3
1.01	0.54	0.10	1	0.50	50	1000	1	3
1.00	1.00	0.64	1	0.50	50	1000	1	3
0.97	0.99	0.68	1	0.50	50	1000	1	3
0.95	0.98	0.68	1	0.50	50	1000	1	3
0.90	0.88	0.37	1	0.50	50	1000	1	3
0.50	0.00	0.00	1	0.50	50	1000	1	3
0.25	0.00	0.00	1	0.50	50	1000	1	3
0.00	0.00	0.00	1	0.50	50	1000	1	3

The summary of results with time series length 500 is as follows:

- Data simulated under the alternative hypothesis where $\phi_1 > 1$ ($H_1^{\text{Explosive AR}}$): The HBPZA test 87% of the test cases and the Lee-Strazicich(LS) test rejects 55% of the test cases when $\phi_1 = 1.01$ when the break time is 125. When the break time is 250 the HBPZA rejects 46% of the test cases, and the Lee-Strazicich rejects 90% of the test cases when $\phi_1 = 1.01$.

- Data simulated under the null hypothesis where $\phi_1 = 1$ ($H_0^{I(1)}$): The HBPZA test rejects the null hypothesis less than the significance level of 5% for both break times of 125 and 250. However the Lee-Srazicich test rejects **100%** of the test cases under the null when the break time is 50 and 51% when the break time is 50.

- Data simulated under the alternative hypothesis where $\phi_1 < 1$ ($H_1^{I(0)}$): When the break time is 125 the Lee-Strazicich significantly outperforms the HBPZA test in this scenario up to $\phi_1 = 0.5$ when both tests become similar. However when the break time is 250 the Lee-Strazicich test does not reject as much as before in this scenario.

3.1.14 Testing with AR(2) Processes

I tested the lagged-series, ERS Ptest and DFGLS unit root tests with AR(2) processes with $l = 500$, $m = 1000$ for the Model $x(t) = \phi_1 x(t-1) + \phi_2 x(t-2) + \epsilon_t$ where $x(0) \sim N(0,1)$; $x(1) \sim N(0,1)$; $\epsilon_t \sim N(0,1)$ and with independent innovations. The results are summarized on Table 3.26. The AR characteristic polynomial Equation (1.31) needs to be solved to determine if these AR(2) processes have 1, 2, or no unit roots, where a unit root is a root with a modulus of 1. The modulus of the two roots are in the last two columns of Table 3.26. We see that the lagged-series Unit test out-performs the cases where the multipliers are $\phi_1 = 1, \phi_2 = 0.01$ and $\phi_1 = 1, \phi_2 = 0.005$ which do not contain a unit root. In the case of $\phi_1 = 1, \phi_2 = -1$ ERS-DFGLS outperforms the others, however it still does not make for a valid unit root test under this scenario as it rejects 68% of the test cases and it should only reject 5% of the cases. In any case the lagged-series Test always produces better results than the ADF test, and sometimes better and sometimes worse than the ERS DF-GLS and Ptest tests for the evaluated AR(2) cases.

Table 3.26: Proportions of Failures to Reject $H_0^{I(1)}$ with lagged-series and Other Tests on AR(2) Processes with $l = 500$, $m = 1000$ and $\alpha = 5\%$

ϕ_1	ϕ_2	ADF	lagged-series	ERS PTest	ERS DFGLS	Root1 Modulus	Root2 Modulus
2.000	-1.000	0.934	0.943	0.780	0.951	1.000	1.000
1.897	-0.900	0.78	0.625	0.540	0.632	1.038	1.071
1.000	-1.000	0.000	0.000	0.013	0.320	1.000	1.000
1.000	-0.050	0.310	0.062	0.039	0.075	1.056	18.944
1.000	0.010	1.000	0.047	0.998	0.998	0.990	100.990
1.000	0.005	0.984	0.686	0.977	0.981	0.995	200.995

3.1.15 URTs on I(2),I(3) and I(4) Processes

We will develop how to derive I(2), I(3) and I(4) processes for the purpose of simulation and we will write them in auto-regressive form first to consider how these processes compare to I(1) processes. First we define a random walk AR(1) process in recursive form in Equation (3.13). This is an I(1) process:

$$\begin{aligned} x(t) &= x(t-1) + \epsilon(t) \\ x(0) &= 0 \end{aligned} \tag{3.13}$$

Now we define an I(2) process in recursive form in Equation (3.14), which reuses Equation (3.13).

$$\begin{aligned} y(t) &= y(t-1) + x(t-1) + \epsilon(t) \\ y(0) &= 0 \end{aligned} \tag{3.14}$$

We define an I(3) process in recursive form in Equation (3.15), which reuses Equations (3.14) and (3.13).

$$\begin{aligned} z(t) &= z(t-1) + y(t-1) + x(t-1) + \epsilon(t) \\ z(0) &= 0 \end{aligned} \tag{3.15}$$

We see that all I(1), I(2) and I(3) processes embed a unit root plus additional additive terms that weigh past values even higher. However these processes are not as potentially ill-behaved/non-convergent as explosive AR(1) processes with $\phi_1 > 1$. So we would not be surprised if URTs such as ADF are used with I(2), I(2) and I(4) series that they would not reject the null hypothesis of a unit root since they do not reject with an explosive AR(1) process ($\phi_1 > 1$). We see this behavior in the simulations performed in Chapter 3 these higher integrated processes.

In Equation (3.16) we see that if we difference Equation (3.13) once we end up with an I(0) series.

$$\Delta x = x(t) - x(t-1) = \epsilon(t) \tag{3.16}$$

In Equation (3.17) we see that if we difference Equation (3.14) twice we end up with a stationary process, i.e., I(0).

$$\begin{aligned} \Delta y &= y(t) - y(t-1) = x(t-1) + \epsilon(t) \\ \Delta\Delta y &= x(t-1) + \epsilon(t) - x(t-2) - \epsilon(t-1) \\ &= \epsilon(t-1) + \epsilon(t) - \epsilon(t-1) = \epsilon(t) \end{aligned} \tag{3.17}$$

In Equation (3.18) we see that if we difference Equation (3.15) thrice we end up with an I(0) series.

$$
\begin{aligned}
\Delta z &= z(t) - z(t-1) = y(t-1) + x(t-1) + \epsilon(t) \\
\Delta\Delta z &= y(t-1) + x(t-1) + \epsilon(t) - y(t-2) - x(t-2) - \epsilon(t-1) \\
&= x(t-2) + \epsilon(t-1) + \epsilon(t-1) + \epsilon(t) - \epsilon(t-1) = x(t-2) + \epsilon(t-1) + \epsilon(t) \\
&= x(t-2) + x(t-1) - x(t-2) + \epsilon(t) = \\
&\quad x(t-1) + \epsilon(t) \\
\Delta\Delta\Delta z &= x(t-1) + \epsilon(t) - x(t-2) - \epsilon(t-1) \\
&= \epsilon(t-1) + \epsilon(t) - \epsilon(t-1) \\
&= \epsilon(t)
\end{aligned}
\tag{3.18}
$$

In R we can generate these integrated order n processes of length 1000 as follows:

- For I(1): `cumsum(rnorm(1000))`
- For I(2): `cumsum(cumsum(rnorm(1000)))`
- For I(3): `cumsum(cumsum(cumsum(rnorm(1000))))`
- For I(4): `cumsum(cumsum(cumsum(cumsum(rnorm(1000)))))`

The `cumsum()` function creates a series that has the cumulative sum of the elements of initial series up to that each element.

Table 3.27 shows simulation tests of the proportions of failures to reject the null hypothesis of a unit root with I(1), I(2), I(3) and I(4) processes using various URTs with a time series length of 1000. Table 3.28 is similar but summarizes the results using a time series length of 500 instead. We see that the ADF, ERS-DFGLS and ERS-ptest never reject more than 6% of the cases. The lagged-series URT rejects 65.4% of cases for I(3) processes and close to 86% of cases for I(4) processes.

Table 3.27: Proportions of Failures to Reject $H_0^{I(1)}$ with lagged-series and Other Tests on I(1),I(2),I(3) and I(4) TS with $l = 1000$, $m = 2000$ and $\alpha = 0.05$ with a lagged-series lag of 30

I(x)	ADF	lagged-series	ERS-Ptest	ERS-DFGLS	ERS-lagged-series
I(1)	0.9460	0.9405	0.9505	0.9495	0.9505
I(2)	0.9370	0.9375	0.7805	0.9440	0.7805
I(3)	0.9470	0.3245	0.8755	0.9550	0.4675
I(4)	0.9435	0.1115	0.9695	0.9720	0.2015

Table 3.28: Proportions of Failures to Reject $H_0^{I(1)}$ with lagged-series and Other Tests on I(1),I(2),I(3) and I(4) TS with $l = 500$, $m = 2000$ and $\alpha = 0.05$ with a lagged-series lag of 30

I(x)	ADF	lagged-series	ERS-Ptest	ERS-DFGLS	ERS-lagged-series
I(1)	0.9570	0.9435	0.9465	0.9530	0.9460
I(2)	0.9385	0.9255	0.7760	0.9520	0.7750
I(3)	0.9605	0.3460	0.8600	0.9555	0.4830
I(4)	0.9425	0.1380	0.9705	0.9675	0.2430

3.1.16 Threshold Unit Root Tests

The three regime threshold Model (3.19) was used for simulations under various combinations of the auto-regressive multipliers ϕ_1^H, ϕ_1^M and ϕ_1^L corresponding to the high(H), middle(M) and low(L) regimes. Two different sets of symmetric thresholds were used ($\theta_H = 2$, $\theta_L = -2$) and ($\theta_H = 20$,$\theta_L = -20$). Threshold models are complex to fit as there are many factors that affect the quality of the fit including the amount of data in each regime, and the number of parameters that need to be fit, and the length of the TS. If the threshold parameters are known, then conditional least squares can be used to fit the model.

$$z_t = \begin{cases} \phi_1^H z_{t-1} + \epsilon_t, & \text{if } z_{t-1} > \theta_H \\ \phi_1^M z_{t-1} + \epsilon_t, & \text{if } \theta_L \leq z_{t-1} \leq \theta_H \\ \phi_1^L z_{t-1} + \epsilon_t, & \text{if } z_{t-1} < \theta_L \end{cases} \tag{3.19}$$

Stigler,Matthieu (2010) points out the common assumed case where Model (3.19) is stationary is when $\phi_h < 1$ and $\phi_l < 1$. The AR multiplier ϕ_1^M in the middle regime does not matter much as long as the high and/or low regimes are being activated. If the realization of the process were to always stay in the middle regime because the threshold values are never reached, then for all practical purposes this regime would be the important one to consider. Simulations were performed under different combinations of AR multipliers for each high(ϕ_1^H), middle(ϕ_1^M) and low(ϕ_1^L) regimes, for two sets of symmetric thresholds ($\theta_L = -2, \theta_H = 2$) and ($\theta_L = -20, \theta_H = 20$) and tests were done for time series of length 1000, 300, and 100.

Table 3.29: Proportions of Failures to Reject $H_0^{I(1)}$ of Unit Root Tests on Various Threshold AR(1) Processes with $l = 1000$, $m = 500$, and $\alpha = 0.05$ and $s = 12345$

θ^H	θ^L	ϕ_1^H	ϕ_1^M	ϕ_1^L	ADF	lagged-series	ERS-Ptest	ERS-DFGLS	BBC
2	-2	1.00	1.00	1.00	0.94	0.94	0.93	0.93	0.94
2	-2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	-2	1.00	0.90	1.00	0.91	0.93	0.91	0.91	0.92
2	-2	1.00	1.10	1.00	0.96	0.94	0.95	0.96	0.93
2	-2	1.01	1.10	1.01	1.00	0.00	1.00	1.00	0.00
2	-2	0.90	1.00	0.90	0.00	0.00	0.00	0.00	0.00
2	-2	0.90	1.10	0.90	0.00	0.00	0.00	0.00	0.00
20	-20	1.00	0.90	1.00	0.00	0.00	0.00	0.00	0.00
20	-20	1.00	1.10	1.00	1.00	0.34	1.00	1.00	0.41
20	-20	1.01	1.10	1.01	1.00	0.00	1.00	1.00	0.00
20	-20	0.90	1.00	0.90	0.91	0.83	0.91	0.91	0.48
20	-20	0.90	1.10	0.90	1.00	0.00	1.00	1.00	0.00

The results for tests with time series of length 1000 are in Table 3.29 and can be summarized as follows:

- Trivial unit root($\phi_1^H = \phi_1^M = \phi_1^L = 1$): As expected the tests reject close to the 5% rate of the test cases as dictated by the significance level of 0.05: ADF 6%, lagged-series 6%, ERS-Ptest 7%, ERS-DFGLS 7% and BBC 6%.

- Unit root($\phi_1^H = \phi_1^L = 1, \phi_1^M <> 1$): When the threshold range is ($\theta_L = -2, \theta_H = 2$) all the tests perform relatively well and reject only close to 5% of the test cases as expected, however the ERS tests perform worse when $\phi_1^M = 0.9$ and reject 9% of the cases. When the threshold range is ($\theta_L = -20, \theta_H = 20$) the results are significantly worse where all of the tests fully reject the null of a unit root when $\phi_1^M = 0.9$ most likely because the process stays in the middle regime. When $\phi_1^M = 1.1$ the ERS tests do not reject the null root hypothesis at all, and the lagged-series and the BBC test reject much more, 66% and 59% respectively.

- Stationary process($\phi_1^H = \phi_1^L = 0.9$) When the threshold range is ($\theta_L = -2, \theta_H = 2$) all the tests perform quite well and reject the null hypothesis of a unit root 100% of the cases. When the threshold range is ($\theta_L = -20, \theta_H = 20$) when $\phi_1^M = 1$ all of the tests do not significantly reject the null of a unit root; the BBC test is the best with a 52% rejection rate. When the middle regime is explosive ($\phi_1^M = 1.1$) both the lagged-series and BBC tests fully reject all tests for a unit root, and the ADF and ERS tests do not reject any. This is the scenario and the fully explosive scenario is where we see the largest contrast between tests.

- Fully explosive process($\phi_1^H = 1.01, \phi_1^M = 1.1, \phi_1^L = 1.01$): Both the lagged-series and BBC tests fully reject all tests for a unit root, and the ADF and ERS tests do not reject any for both threshold ranges.

The results for tests with time series of length 300 are in Table 3.30 and can be summarized as follows:

Table 3.30: Proportions of Failures to Reject $H_0^{I(1)}$ of Unit Root Tests on Various Threshold AR(1) Processes with $l = 300$, $m = 500$, and $\alpha = 0.05$ and $s = 12345$

θ^H	θ^L	ϕ_1^H	ϕ_1^M	ϕ_1^L	ADF	lagged-series	ERS-Ptest	ERS-DFGLS	BBC
2	-2	1.00	1.00	1.00	0.95	0.95	0.94	0.94	0.95
2	-2	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.00
2	-2	1.00	0.90	1.00	0.93	0.95	0.92	0.91	0.93
2	-2	1.00	1.10	1.00	0.96	0.94	0.96	0.96	0.92
2	-2	1.01	1.10	1.01	1.00	0.58	0.99	0.99	0.82
2	-2	0.90	1.00	0.90	0.00	0.08	0.00	0.01	0.64
2	-2	0.90	1.10	0.90	0.00	0.16	0.01	0.01	0.49
20	-20	1.00	0.90	1.00	0.00	0.05	0.00	0.01	0.70
20	-20	1.00	1.10	1.00	1.00	0.33	1.00	1.00	0.55
20	-20	1.01	1.10	1.01	1.00	0.06	1.00	1.00	0.11
20	-20	0.90	1.00	0.90	0.95	0.94	0.94	0.94	0.80
20	-20	0.90	1.10	0.90	1.00	0.08	1.00	1.00	0.00

- Trivial unit root($\phi_1^H = \phi_1^M = \phi_1^L = 1$): As expected the tests reject close to the 5% rate of test cases as dictated by the significance level of 0.05: ADF 5%, lagged-series 5%, ERS-Ptest 6%, ERS-DFGLS 6% and BBC 5%.

- Unit root($\phi_1^H = \phi_1^L = 1, \phi_1^M <> 1$): When the threshold range is $(\theta_L = -2, \theta_H = 2)$ all the tests perform relatively well and reject only close to 5% of the test cases as expected, however the ERS tests perform worse when $\phi_1^M = 0.9$ and reject 8% to 9% of the cases. When the threshold range is $(\theta_L = -20, \theta_H = 20)$ with $\phi_1^M = 0.9$ the results are significantly worse where the ERS reject 100% to 99% of the cases, the lagged-series tests reject 95% of the tests of a unit root. However the BBC only rejects 30% of the cases; the process spends considerable more time in the middle regime than with the smaller thresholds. When $\phi_1^M = 1.1$ the ERS tests do not reject the null root hypothesis at all, and the lagged-series and the BBC test reject much more, 67% and 45% respectively.

- Stationary process($\phi_1^H = \phi_1^L = 0.9$) When the threshold range is $(\theta_L = -2, \theta_H = 2)$ all the ADF and ERS perform very well and reject the null hypothesis of a unit root at least 99% of the cases. The lagged-series tests rejects at least 84% of the tests in this scenario. However the BBC test does not perform as well and rejects at most 36% of the cases. When the threshold range is ($\theta_L = -20, \theta_H = 20$) when $\phi_1^M = 1$ all of the tests do not significantly reject the null of a unit root; the BBC test is the best with a 20% rejection rate. When the middle regime is explosive ($\phi_1^M = 1.1$) the BBC tests fully reject all tests for a unit root, the lagged-series rejects 92% of the cases and and the ADF and ERS tests do not reject any. This is the scenario and the fully explosive scenario is where we see the largest contrast between tests.

- Fully explosive process($\phi_1^H = 1.01, \phi_1^M = 1.1, \phi_1^L = 1.01$): The ADF and ERS tests at most reject 1% of the test cases for a unit root for both threshold ranges. The lagged-series tests rejects 42% of the tests cases for a unit root and the BBC rejects 18% when the threshold range is ($\theta_L = -2, \theta_H = 2$). When the threshold range is ($\theta_L = -20, \theta_H = 20$) the lagged-series tests rejects 94% of the tests cases for the null of a unit root and the BBC rejects 89%.

Table 3.31: Proportions of Failures to Reject $H_0^{I(1)}$ of Unit Root Tests on Various Threshold AR(1) Processes with $l = 100$, $m = 500$, and $\alpha = 0.05$ and $s = 12345$

θ^H	θ^L	ϕ_1^H	ϕ_1^M	ϕ_1^L	ADF	lagged-series	ERS-Ptest	ERS-DFGLS	BBC
2.0	-2.0	1.00	1.00	1.00	0.96	0.95	0.93	0.93	0.95
2.0	-2.0	0.00	0.00	0.00	0.00	0.00	0.02	0.21	0.01
2.0	-2.0	1.00	0.90	1.00	0.91	0.95	0.90	0.90	0.95
2.0	-2.0	1.00	1.10	1.00	0.97	0.94	0.95	0.95	0.95
2.0	-2.0	1.01	1.10	1.01	0.98	0.96	0.97	0.97	0.95
2.0	-2.0	0.90	1.00	0.90	0.39	0.82	0.39	0.51	0.90
2.0	-2.0	0.90	1.10	0.90	0.52	0.84	0.49	0.59	0.86
20.0	-20.0	1.00	0.90	1.00	0.27	0.76	0.28	0.40	0.93
20.0	-20.0	1.00	1.10	1.00	1.00	0.67	1.00	1.00	0.52
20.0	-20.0	1.01	1.10	1.01	1.00	0.94	1.00	1.00	0.82
20.0	-20.0	0.90	1.00	0.90	0.96	0.95	0.93	0.93	0.95
20.0	-20.0	0.90	1.10	0.90	1.00	0.49	1.00	1.00	0.00

The results for tests with time series of length 100 are in Table 3.31 and can be summarized as follows:

- Trivial unit root($\phi_1^H = \phi_1^M = \phi_1^L = 1$): As expected the tests reject close to the 5% rate of test cases as dictated by the significance level of 0.05: ADF 4%, lagged-series 5%, ERS-Ptest 7%, ERS-DFGLS 7% and BBC 5%.

- Unit root($\phi_1^H = \phi_1^L = 1, \phi_1^M <> 1$): When the threshold range is ($\theta_L = -2, \theta_H = 2$) all the tests perform relatively well and reject only close to 5% of the test cases as expected, however the ERS tests perform worse when $\phi_1^M = 0.9$ and reject 10% of the cases. When the threshold range is ($\theta_L = -20, \theta_H = 20$) with $\phi_1^M = 0.9$ the results are significantly worse where the ERS reject 60% to 72% of the cases, the lagged-series tests reject 24% of the tests of a unit root. However the BBC only rejects 7% of the cases; the process spends considerable more time in the middle regime than with the smaller thresholds. When $\phi_1^M = 1.1$ the ERS tests do not reject the null root hypothesis at all, and the lagged-series and the BBC test reject much more, 33% and 48% respectively.

- Stationary process($\phi_1^H = \phi_1^L = 0.9$): When the threshold range is ($\theta_L = -2, \theta_H = 2$) none of the tests sufficiently reject, the ERS-Ptest performs best here rejecting 61% of the cases when $\phi_1^M = 1$ and 51% of the cases when $\phi_1^M = 1.1$. When the threshold range is ($\theta_L = -20, \theta_H = 20$) when $\phi_1^M = 1$ all of the tests do not significantly reject the null of a unit root. When the middle regime is explosive ($\phi_1^M = 1.1$) the BBC tests fully reject all tests for a unit root, the lagged-series rejects 51% of the cases and and the ADF and ERS tests do not reject any. This is the scenario where we see the largest contrast between tests with length 100.

- Fully explosive process($\phi_1^H = 1.01, \phi_1^M = 1.1, \phi_1^L = 1.01$): All tests hardly reject the null hypothesis of a unit root.

3.1.17 Summary of Results with Threshold Simulations

For the given threshold simulations URTs performed with the given URTs there is no single test that outperforms the others in every case. It is clear that the time series length is a significant factor in these tests, as well as of course the threshold settings–if the process does not have sufficient data in all thresholds then it is impossible to expect the tests to work across all thresholds. We see that the lagged-series and the BBC tests deal best with explosive regimes.

3.2 Simulations with the Hybrid Bai-Perron ADF I(0)/I(1) (HBPADF) Testing Procedure

$$y_t = \phi_1^A y_{t-1}\mathbf{1}\{t \leq T_B\} + \phi_1^B y_{t-1}\mathbf{1}\{t > T_B\} \tag{3.20}$$

Table 3.32: Proportions of Failures to Reject $H_0^{I(1)}$ on a AR(1) Series with a Break in ϕ_1 $l = 1000$,$m = 1000$ and $\alpha = 0.05$

Break Proportion	ϕ_1^A	ϕ_1^B	$\overline{\Gamma^{\text{A}}_{\text{ADF}}}$	$\overline{\Gamma^{\text{B}}_{\text{ADF}}}$	$\overline{\Gamma^{\text{All}}_{\text{ADF}}}$
0.50	0.000	0.000	0.00	0.00	0.00
0.50	1.000	1.000	0.93	0.92	0.94
0.50	1.000	0.500	0.94	0.03	0.87
0.50	0.900	1.000	0.09	0.89	0.85
0.50	1.000	0.900	0.91	0.05	0.83
0.50	1.000	0.950	0.84	0.31	0.83
0.50	1.000	0.980	0.90	0.74	0.90
0.25	0.000	0.000	0.00	0.00	0.00
0.25	1.000	1.000	0.93	0.92	0.94
0.25	1.000	0.500	0.96	0.01	0.65
0.25	0.900	1.000	0.68	0.90	0.94
0.25	1.000	0.900	0.83	0.01	0.45
0.25	1.000	0.950	0.58	0.11	0.47
0.25	1.000	0.980	0.72	0.62	0.72

Simulations were performed using the DGP (3.20) which has breaks only in the AR ϕ_1

multiplier with time series of length $l = 1000$ and are summarized in Table 3.32. We see that the approach works well to distinguish between $\phi_1 = 1$ vs. $\phi_1 = 0.9$ when the break happens in the middle of the time series (break proportion=0.5), however the I(1) to I(0) transition is better detected than the I(0) to I(1) transition. When the break happens in the first quarter of the time series (break proportion=0.25) the approach works much better to distinguish between an I(1) to I(0) with $\phi_1 = 0.9$ transition, but it is not good to distinguish between an I(0) to I(1) change. The results when the break happens in the first quarter of the time series are not as good, especially much worse results with I(0) to I(1) changes than with I(1) to I(0) changes.

Table 3.33: Proportions of Failures to Reject $H_0^{I(1)}$ on a AR(1) Series with a Break in ϕ_1 $l = 100$,$m = 1000$ and $\alpha = 0.05$

Break Proportion	ϕ_1^A	ϕ_1^B	$\overline{\Gamma^{\mathrm{A}}_{\mathrm{ADF}}}$	$\overline{\Gamma^{\mathrm{B}}_{\mathrm{ADF}}}$	$\overline{\Gamma^{\mathrm{All}}_{\mathrm{ADF}}}$
0.50	0.000	0.000	0.06	0.06	0.06
0.50	1.000	1.000	0.94	0.94	0.96
0.50	1.000	0.500	0.93	0.65	0.91
0.50	0.900	1.000	0.90	0.93	0.93
0.50	1.000	0.900	0.94	0.93	0.95
0.50	1.000	0.950	0.95	0.95	0.97
0.50	1.000	0.980	0.94	0.95	0.96
0.25	0.000	0.000	0.06	0.06	0.06
0.25	1.000	1.000	0.94	0.94	0.96
0.25	1.000	0.500	0.79	0.43	0.74
0.25	0.900	1.000	0.94	0.95	0.96
0.25	1.000	0.900	0.91	0.88	0.91
0.25	1.000	0.950	0.93	0.92	0.94
0.25	1.000	0.980	0.95	0.94	0.95

Simulations performed using the DGP (3.20) which has breaks only in the AR ϕ_1 multiplier with a now shorter time series of length $l = 100$ are summarized in Table 3.33. We see that the approach does not work well to distinguish between $\phi_1 = 1$ vs. $\phi_1 = 0.9$. If we consider the case where the series transitions from a unit root ($\phi_1 = 1$) to $\phi_1 = 0.5$ with a break in the middle of the series, the proportions of the failures to reject the null of a unit root are not good (0.95, 0.65) so it rejects only 35% of cases on the second segment which is I(0). However if we were to compare this to running the ADF test on the entire TS as is displayed on the last column of the table (0.91) we would only reject in 9% of the cases.

$$
\begin{aligned}
y_t = \phi_1^A y_{t-1} \mathbf{1}\{t \le T_B\} + \phi_1^B y_{t-1} \mathbf{1}\{t > T_B\} + \\
+\mu_1 + (\mu_2 - \mu_1)\mathbf{1}\{t > T_B\} + \\
+(t - T_B)(\beta_2 - \beta 1_1)\mathbf{1}\{t > T_B\} + \beta t + e_t \\
\mu_1 = 10 \ ; \ \mu_2 = -30 \ ; \ \beta_1 = 5 \ ; \ \beta_2 = -2
\end{aligned}
\tag{3.21}
$$

Table 3.34: Proportions of Failures to Reject $H_0^{I(1)}$ on a AR(1) Series with a Break in ϕ_1, intercept and linear trend with $l = 1000$,$m = 1000$ and $\alpha = 0.05$

Break Proportion	ϕ_1^A	ϕ_1^B	μ^A	μ^B	μ_t^A	μ_t^B	$\overline{\Gamma_{\text{ADF}}^{\text{A}}}$	$\overline{\Gamma_{\text{ADF}}^{\text{B}}}$	$\overline{\Gamma_{\text{ADF}}^{\text{All}}}$
0.50	0.000	0.000	10	-30	5	-2	0.00	0.00	1.00
0.50	1.000	1.000	10	-30	5	-2	0.96	0.96	1.00
0.50	1.000	0.500	10	-30	5	-2	0.96	0.00	1.00
0.50	0.900	1.000	10	-30	5	-2	0.01	0.96	1.00
0.50	1.000	0.900	10	-30	5	-2	0.96	0.00	1.00
0.50	1.000	0.950	10	-30	5	-2	0.96	0.14	1.00
0.50	1.000	0.980	10	-30	5	-2	0.96	0.67	1.00
0.25	0.000	0.000	10	-30	5	-2	0.00	0.00	0.03
0.25	1.000	1.000	10	-30	5	-2	0.96	0.95	0.04
0.25	1.000	0.500	10	-30	5	-2	0.96	0.00	0.15
0.25	0.900	1.000	10	-30	5	-2	0.47	0.95	0.02
0.25	1.000	0.900	10	-30	5	-2	0.96	0.00	0.03
0.25	1.000	0.950	10	-30	5	-2	0.96	0.02	0.04
0.25	1.000	0.980	10	-30	5	-2	0.96	0.58	0.04

Table 3.34 summarizes the results of the new testing procedure run on simulations were

performed with changes in the AR ϕ_1 multiplier as well as in the intercept and linear trend with time series of length $l = 1000$ using DGP (3.21). We see that the approach works very well to distinguish between $\phi_1 = 1$ vs. $\phi_1 = 0.9$ when the break happens in the middle of the time series (break proportion=0.5) as well as the transition from $\phi_1 = 0.9$ to $\phi_1 = 1$. When the break happens in the first quarter of the time series (break proportion=0.25) the approach works well to distinguish between an I(1) to I(0) with $\phi_1 = 0.9$ transition, but it is not good to distinguish between an I(0) to I(1).

Table 3.35: Proportions of Failures to Reject $H_0^{I(1)}$ on a AR(1) Series with a Break in ϕ_1, intercept and linear trend with $l = 100$,$m = 1000$ and $\alpha = 0.05$

Break Proportion	ϕ_1^A	ϕ_1^B	μ^A	μ^B	μ_t^A	μ_t^B	$\overline{\Gamma^{\text{A}}_{\text{ADF}}}$	$\overline{\Gamma^{\text{B}}_{\text{ADF}}}$	$\overline{\Gamma^{\text{All}}_{\text{ADF}}}$
0.50	0.000	0.000	10	-30	5	-2	0.49	0.21	1.00
0.50	1.000	1.000	10	-30	5	-2	0.96	0.95	1.00
0.50	1.000	0.500	10	-30	5	-2	0.96	0.52	1.00
0.50	0.900	1.000	10	-30	5	-2	0.94	0.95	1.00
0.50	1.000	0.900	10	-30	5	-2	0.96	0.92	1.00
0.50	1.000	0.950	10	-30	5	-2	0.96	0.95	1.00
0.50	1.000	0.980	10	-30	5	-2	0.96	0.95	1.00
0.25	0.000	0.000	10	-30	5	-2	0.83	0.09	0.00
0.25	1.000	1.000	10	-30	5	-2	0.95	0.95	0.00
0.25	1.000	0.500	10	-30	5	-2	0.95	0.33	0.00
0.25	0.900	1.000	10	-30	5	-2	0.95	0.95	0.00
0.25	1.000	0.900	10	-30	5	-2	0.95	0.90	0.00
0.25	1.000	0.950	10	-30	5	-2	0.95	0.95	0.00
0.25	1.000	0.980	10	-30	5	-2	0.95	0.94	0.00

Table 3.35 summarizes the results of the new testing procedure run on simulations were performed with changes in the AR ϕ_1 multiplier as well as in the intercept and linear trend with time series of length $l = 100$ using DGP (3.21). We see that the approach does not work well to distinguish between $\phi_1 = 1$ vs. $\phi_1 = 0.9$. If we consider the case where the series transitions from a unit root ($\phi_1 = 1$) to $\phi_1 = 0.5$ with a break in the middle of the series, the proportions of the failures to reject the null of a unit root are not good (0.96,

0.52) so it rejects only 48% of cases on the second segment which is I(0). However if we were to compare this to running the ADF test on the entire TS as is displayed on the last column of the table (1.0) we would not reject at all in this case.

3.2.1 Summary of BPADF Tests

We have seen that this new HBPADF testing procedure is sensitive to the time series length and the location of the structural break. When the series is of length 1000 and the break occurs in the middle the procedure works well; it is not as good when the break happens in the first quarter. When the time series length is reduced to 100 the results are worse, however when compared to using a single unit root test on the entire time series, this approach is still significantly more accurate.

3.2.2 Summary of Unit Root Tests

The ADF,ERS-Ptest, ERS-DFGLS and ZA URTs (URT) simulation studies detailed in this chapter can be summarized as follows:

- None of these tests reject the null hypothesis of a unit root ($\phi_1 = 1$) when the simulated data follow an explosive AR process ($\phi_1 > 1$)
- When there are no structural breaks in intercept and/or linear trend in the simulated data:
  - Under a significance level of 0.05 all of these tests when run with simulated data under the null hypothesis reject the null hypothesis close to 5% rate of test cases as expected.
  - This is true regardless of whether the innovations are Gaussian, or Beta distributed, or Gaussian with Gaussian distributed standard deviations, or even if the Gaussian innovations are significantly positively or negatively pairwise correlated (-0.8, +0.8 or +0.9999).

- When $\phi_1 < 1$ generally the ERS tests are the best performers in terms of rejecting the null hypothesis, followed by the ADF test and finally the ZA test is the worst performing when there are no structural breaks. The best performing test in these scenarios is the ERS-Ptest.
- When $\phi_1 < 1$ and the innovations are Gaussian and largely negatively pairwise correlated (-0.8) the tests do not perform as well as with 0 or positively pairwise correlated. For instance for ERS-Ptest the null hypothesis is rejected with simulated data with $\phi_1 \leq 0.97$ at a rate of 5% or less with a 0.8 correlation, yet with a correlation of -0.8 the null hypothesis is rejected at a rate of 5% or less with $\phi_1 \leq 0.90$.

- When there are structural breaks in intercept in the simulated data if the intercepts are small (-1,1) then all URTs perform reasonably well. However if the intercepts are large (-10,10) then the only test that produces somewhat reasonable results is ZA even though it is far from ideal–it will reject the null hypothesis at a rate of 26% when tested on data simulated with $\phi_1 = 1$.
- When there are structural breaks in linear trend in the simulated data none of the tests perform satisfactorily.
- When there are structural breaks in intercept and linear trend in the simulated data none of the tests perform satisfactorily either.
- HBPZA and Lee-Stratizich URTs with breaks in the null: The Lee-Strazich tests do not perform well under the null hypotheses of unit roots tested–however they reject alternative hypothesis at higher rates than the HBPZA test. The HBPZA test performs well under the null hypotheses considered.
- AR(2): The lagged-series URT performs sometimes better and sometimes worse than the ADF and ERS URTs on simulations with AR(2) processes.
- I(2),I(3),I(4): The ADF and ERS URTs can distinguish between I(1), and I(2), I(3)

and I(4) processes. However the lagged-series URT rejects the null hypothesis of a unit root much more than the other tests for I(3) and I(4) series.

- URTs with a Threshold model. For the tested ADF, ERS, BBC and lagged-series URTs there is no one that outperforms the other in every case. The lagged-series and BBC tests outperform the others when there are explosive ($\phi_1 > 1$ regimes.)

- The new HBPADF I(0)/I(1) testing procedure is sensitive to the time series length and the location of the structural break. Even when the time series are short, this approach is still significantly more accurate than using a single URT on the entire series.

3.3 Simulation Studies of Cointegration Tests

Figure 3.1 is a flowchart on how to perform regressions of two or more TS variables. This Figure is an extension of Hill, R. Carter and Griffiths, William E. and Lim, Guay C. (2011, Fig. 12.4); The referred author only considers nonstationary TS but here we expand the approach to consider stationary ones as well. The assumption in this diagram is that all of the series being analyzed are either I(0) or I(1). We do not consider here if the series are of mixed integration order, or if the series are of higher integration orders than 1. One approach to deal with these cases would be to difference every series as many times required to obtain an I(0) series and then regress all of the sufficiently differenced series; however we note that every time we difference time series we are loosing information and there can be cases as the example provided in Section 1.6.2. We can see that determining if a series is I(1) or I(0) is quite important as it will dictate which methodology should be used to infer a valid linear regression model. Regression models with independent I(1) variables very often lead to spurious results as was seen in Table 1.5, and instead we should pursue a model with first differenced variables. If the I(1) variables are cointegrated then we can estimate a short run model(Error Correcting Model) a long run model in the levels of the series.

However if we believe that we have variables that are I(1), but they area near-integrated

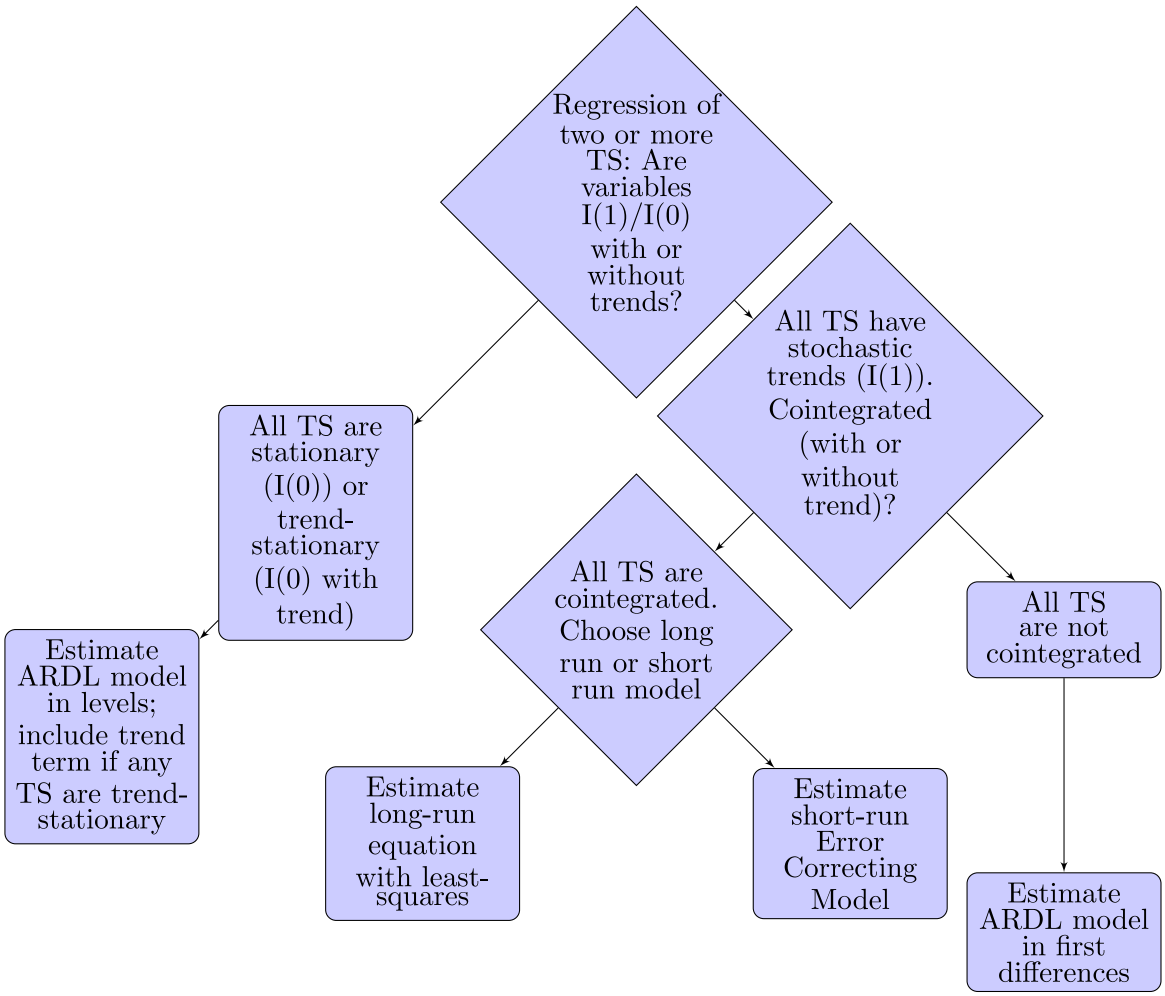


Figure 3.1: Regressions with I(0)/I(1) TS Variables Based on Hill, R. Carter and Griffiths, William E. and Lim, Guay C. (2011, Fig. 12.4)

($\phi_1 = 1-\epsilon$ for a small positive *epsilon*) or they are explosive $\phi_1 > 1$ this can lead to spurious cointegration results. In this section we perform simulation studies where we simulate two independent TS and first check to see if they are I(1) using a URT; if the test does not reject the null hypothesis of a unit root for both TS then we perform a cointegration test using the Johansen procedure. Finally we check for spurious cointegration results.

3.3.1 Spurious Cointegration

$$\begin{aligned} x_t &= \phi_1^a x_{t-1} + \sigma_1 \\ y_t &= \phi_1^b y_{t-1} + \sigma_2 \\ \sigma_1 &\sim N(0,1) \\ \sigma_2 &\sim N(0,1) \end{aligned} \tag{3.22}$$

Spurious Cointegration Without Pretesting for Unit Roots

We use DGP (3.22), where x_t and y_t are independent, to perform simulation tests of spurious cointegration where the Johansen cointegration test is performed on the two series, without considering if they are unit roots or not.

Tables 3.36, 3.37 and 3.38 show these results of spurious cointegration for α=0.01,α=0.05 and α=0.10 respectively with time series of length 1000.

We can see that in Tables 3.36, 3.37 and 3.38 for the case of $\phi_1^a = 1$ and $\phi_1^b = 1$ the failure to reject the null hypothesis of no-cointegration is the significance level.

We can see that in Tables 3.36, 3.37 and 3.38 for the case of $\phi_1^a > 1$ and/or $\phi_1^b > 1$ the failure to reject the null hypothesis of no-cointegration are quite high–showing evidence of cointegration in 90% to 100% of the cases which is entirely spurious. The same is true if $\phi_1^a < 0.9$ and/or $\phi_1^b < 0.9$

Now we use DGP (3.22) to run simulations with time series of length 100. Tables 3.36, 3.37 and 3.38 show these results of spurious cointegration for α=0.01,α=0.05 and α=0.10

Table 3.36: Proportions of Failures to Reject H_0^{Johansen} $r = 0$ of Pairs of AR(1) TS with $m = 5000$, l=1000 and $\alpha = 0.01$ as Defined in Model (3.22) and s=12345

ϕ_1^a	ϕ_1^b	$\overline{\Gamma_{\text{co}} > CV_{\text{co}}^{\alpha}}$
1.010	1.010	1.00
1.000	1.020	1.00
1.000	1.010	1.00
1.000	1.005	0.93
1.000	1.000	0.01
0.990	0.990	0.04
1.000	0.960	0.81
1.000	0.970	0.46
1.000	0.980	0.13
1.000	0.990	0.03
1.000	0.900	1.00
1.000	0.200	1.00
0.900	0.900	1.00
0.500	0.500	1.00
0.200	0.200	1.00

Table 3.37: Proportions of Failures to Reject $\mathrm{H}_0^{\mathrm{Johansen}}$ $r = 0$ of Pairs of AR(1) TS with $m = 5000$, l=1000 and $\alpha = 0.05$ as Defined in Model (3.22) and s=12345

ϕ_1^a	ϕ_1^b	$\overline{\Gamma_{\mathrm{co}} > CV_{\mathrm{co}}^{\alpha}}$
1.010	1.010	1.00
1.000	1.020	1.00
1.000	1.010	1.00
1.000	1.005	0.94
1.000	1.000	0.05
0.990	0.990	0.17
1.000	0.960	0.97
1.000	0.970	0.80
1.000	0.980	0.42
1.000	0.990	0.12
1.000	0.900	1.00
1.000	0.200	1.00
0.900	0.900	1.00
0.500	0.500	1.00
0.200	0.200	1.00

Table 3.38: Proportions of Failures to Reject $\mathrm{H}_0^{\text{Johansen}}$ $r = 0$ of Pairs of AR(1) TS with $m = 5000$, l=1000 and $\alpha = 0.10$ as Defined in Model (3.22) and s=12345

ϕ_1^a	ϕ_1^b	$\overline{\Gamma_{\text{co}} > CV_{\text{co}}^{\alpha}}$
1.010	1.010	1.00
1.000	1.020	1.00
1.000	1.010	1.00
1.000	1.005	0.95
1.000	1.000	0.10
0.990	0.990	0.31
1.000	0.960	0.99
1.000	0.970	0.92
1.000	0.980	0.60
1.000	0.990	0.22
1.000	0.900	1.00
1.000	0.200	1.00
0.900	0.900	1.00
0.500	0.500	1.00
0.200	0.200	1.00

Table 3.39: Proportions of Failures to Reject H_0^{Johansen} $r = 0$ of Pairs of AR(1) TS with $m = 5000$, l=100 and $\alpha = 0.01$ as Defined in Model (3.22) and s=12345

ϕ_1^a	ϕ_1^b	$\overline{\Gamma_{\text{co}} > CV_{\text{co}}^{\alpha}}$
1.010	1.010	0.01
1.000	1.020	0.02
1.000	1.010	0.01
1.000	1.005	0.01
1.000	1.000	0.01
0.990	0.990	0.01
1.000	0.960	0.02
1.000	0.970	0.02
1.000	0.980	0.01
1.000	0.990	0.01
1.000	0.900	0.04
1.000	0.200	1.00
0.900	0.900	0.06
0.500	0.500	0.99
0.200	0.200	1.00

Table 3.40: Proportions of Failures to Reject H_0^{Johansen} $r = 0$ of Pairs of AR(1) TS with $m = 5000$, l=100 and $\alpha = 0.05$ as Defined in Model (3.22) and s=12345

ϕ_1^a	ϕ_1^b	$\overline{\Gamma_{\text{co}} > CV_{\text{co}}^{\alpha}}$
1.010	1.010	0.05
1.000	1.020	0.08
1.000	1.010	0.06
1.000	1.005	0.06
1.000	1.000	0.06
0.990	0.990	0.07
1.000	0.960	0.07
1.000	0.970	0.07
1.000	0.980	0.06
1.000	0.990	0.06
1.000	0.900	0.13
1.000	0.200	1.00
0.900	0.900	0.19
0.500	0.500	1.00
0.200	0.200	1.00

respectively. For the case of $\phi_1^a = 1$ and $\phi_1^b = 1$ the failure to reject the null hypothesis of no-cointegration is larger than the significance level for glsalpha=0.05 where it is 0.06 and for α=0.10 where it is 0.12.

Table 3.41: Proportions of Failures to Reject $\mathrm{H}_0^{\mathrm{Johansen}}$ $r = 0$ of Pairs of AR(1) TS with $m = 5000$, l=100 and $\alpha = 0.10$ as Defined in Model (3.22) and s=12345

ϕ_1^a	ϕ_1^b	$\overline{\Gamma_{\mathrm{co}} > CV_{\mathrm{co}}^{\alpha}}$
1.010	1.010	0.10
1.000	1.020	0.14
1.000	1.010	0.11
1.000	1.005	0.12
1.000	1.000	0.12
0.990	0.990	0.13
1.000	0.960	0.13
1.000	0.970	0.13
1.000	0.980	0.12
1.000	0.990	0.12
1.000	0.900	0.23
1.000	0.200	1.00
0.900	0.900	0.33
0.500	0.500	1.00
0.200	0.200	1.00

We can see that in Tables 3.39, 3.40 and 3.41 for the case of $\phi_1^a > 1$ and/or $\phi_1^b > 1$ the failure to reject the null hypothesis of no-cointegration is not much higher than the significance level; this is because in 100 time steps the AR(1) processes have not had much of a chance to "explode" yet. However if $\phi_1^a < 0.5$ and/or $\phi_1^b < 0.5$ this results in a very high failure rate to reject the null hypothesis of a unit root.

Spurious Cointegration Pretesting for Unit Roots

Model (3.22), where x_t and y_t are independent, is used to perform simulation tests of spurious cointegration where the Johansen cointegration test is performed only if both series cannot reject the null hypothesis of a unit root test.

Table 3.42: Proportions of Failures to Reject $H_0^{I(1)}$ with ADF and lagged-series Tests Together with Cointegration of Pairs of AR(1) TS with $m = 5000$, l=1000 and $\alpha = 0.01$ as Defined in Model (3.22)

ϕ_1^a	ϕ_1^b	$\overline{\left(\Gamma_{\text{la}} < CV_{\text{la}}^{\alpha}\right) \wedge \left(\Gamma_{\text{co}} > CV_{\text{co}}^{\alpha}\right)}$	$\overline{\left(\Gamma_{\text{ad}} > CV_{\text{ad}}^{\alpha}\right) \wedge \left(\Gamma_{\text{co}} > CV_{\text{co}}^{\alpha}\right)}$
1.01	1.00	0.00	0.99
1.01	1.01	0.00	1.00
1.01	0.97	0.00	0.00
1.00	1.00	0.01	0.01
0.99	0.99	0.01	0.01
1.00	0.96	0.00	0.00
1.00	0.97	0.01	0.00
1.00	0.98	0.02	0.00
1.00	0.99	0.01	0.01

Table 3.43: Proportions of Failures to Reject $H_0^{I(1)}$ with ADF and lagged-series Tests Together with Cointegration of Pairs of AR(1) TS with $m = 5000$, l=1000 and $\alpha = 0.05$ as Defined in Model (3.22)

ϕ_1^a	ϕ_1^b	$\overline{\left(\Gamma_{\text{la}} < CV_{\text{la}}^{\alpha}\right) \wedge \left(\Gamma_{\text{co}} > CV_{\text{co}}^{\alpha}\right)}$	$\overline{\left(\Gamma_{\text{ad}} > CV_{\text{ad}}^{\alpha}\right) \wedge \left(\Gamma_{\text{co}} > CV_{\text{co}}^{\alpha}\right)}$
1.01	1.00	0.00	0.95
1.01	1.01	0.00	1.00
1.01	0.97	0.00	0.00
1.00	1.00	0.03	0.04
0.99	0.99	0.02	0.00
1.00	0.96	0.00	0.00
1.00	0.97	0.00	0.00
1.00	0.98	0.02	0.00
1.00	0.99	0.03	0.01

Table 3.44: Proportions of Failures to Reject $H_0^{I(1)}$ with ADF and lagged-series Tests Together with Cointegration of Pairs of AR(1) TS with $m = 5000$, l=1000 and $\alpha = 0.10$ as Defined in Model (3.22)

ϕ_1^a	ϕ_1^b	$\overline{(\Gamma_{\text{la}} < CV_{\text{la}}^{\alpha}) \wedge (\Gamma_{\text{co}} > CV_{\text{co}}^{\alpha})}$	$\overline{(\Gamma_{\text{ad}} > CV_{\text{ad}}^{\alpha}) \wedge (\Gamma_{\text{co}} > CV_{\text{co}}^{\alpha})}$
1.01	1.00	0.00	0.90
1.01	1.01	0.00	1.00
1.01	0.97	0.00	0.00
1.00	1.00	0.05	0.08
0.99	0.99	0.02	0.00
1.00	0.96	0.00	0.00
1.00	0.97	0.00	0.00
1.00	0.98	0.01	0.00
1.00	0.99	0.04	0.01

Table 3.45: Proportions of Failures to Reject $H_0^{I(1)}$ with ADF and lagged-series Tests Together with Cointegration of Pairs of AR(1) TS with $m = 5000$, l=100 and $\alpha = 0.01$ as Defined in Model (3.22)

ϕ_1^a	ϕ_1^b	$\overline{(\Gamma_{\text{la}} < CV_{\text{la}}^{\alpha}) \wedge (\Gamma_{\text{co}} > CV_{\text{co}}^{\alpha})}$	$\overline{(\Gamma_{\text{ad}} > CV_{\text{ad}}^{\alpha}) \wedge (\Gamma_{\text{co}} > CV_{\text{co}}^{\alpha})}$
1.01	1.00	0.01	0.01
1.01	1.01	0.01	0.01
1.01	0.97	0.01	0.01
1.00	1.00	0.01	0.01
0.99	0.99	0.01	0.01
1.00	0.96	0.01	0.01
1.00	0.97	0.02	0.02
1.00	0.98	0.01	0.01
1.00	0.99	0.01	0.01

Table 3.46: Proportions of Failures to Reject $H_0^{I(1)}$ with ADF and lagged-series Tests Together with Cointegration of Pairs of AR(1) TS with $m = 5000$, l=100 and $\alpha = 0.05$ as Defined in Model (3.22)

ϕ_1^a	ϕ_1^b	$\overline{\left(\Gamma_{\text{la}} < CV_{\text{la}}^{\alpha}\right) \wedge \left(\Gamma_{\text{co}} > CV_{\text{co}}^{\alpha}\right)}$	$\overline{\left(\Gamma_{\text{ad}} > CV_{\text{ad}}^{\alpha}\right) \wedge \left(\Gamma_{\text{co}} > CV_{\text{co}}^{\alpha}\right)}$
1.01	1.00	0.04	0.05
1.01	1.01	0.04	0.05
1.01	0.97	0.04	0.04
1.00	1.00	0.04	0.05
0.99	0.99	0.05	0.05
1.00	0.96	0.05	0.04
1.00	0.97	0.04	0.05
1.00	0.98	0.04	0.05
1.00	0.99	0.04	0.05

Table 3.47: Proportions of Failures to Reject $H_0^{I(1)}$ with ADF and lagged-series Tests Together with Cointegration of Pairs of AR(1) TS with $m = 5000$, l=100 and $\alpha = 0.10$ as Defined in Model (3.22)

ϕ_1^a	ϕ_1^b	$\overline{\left(\Gamma_{\text{la}} < CV_{\text{la}}^{\alpha}\right) \wedge \left(\Gamma_{\text{co}} > CV_{\text{co}}^{\alpha}\right)}$	$\overline{\left(\Gamma_{\text{ad}} > CV_{\text{ad}}^{\alpha}\right) \wedge \left(\Gamma_{\text{co}} > CV_{\text{co}}^{\alpha}\right)}$
1.01	1.00	0.07	0.09
1.01	1.01	0.07	0.09
1.01	0.97	0.06	0.06
1.00	1.00	0.07	0.10
0.99	0.99	0.08	0.09
1.00	0.96	0.07	0.05
1.00	0.97	0.07	0.06
1.00	0.98	0.07	0.08
1.00	0.99	0.07	0.09

Tables 3.42, 3.43, 3.44 summarize simulations tests where we simulate two AR(1) processes of length 1000 with multipliers ϕ_1^a and ϕ_1^b, then we check if the appropriate unit root test statistic for lagged-series (Γ_{la}) and ADF (Γ_{ad}) with the given significance level and also test for cointegration using the Johansen cointegration test (Γ_{co}) with significance levels of 0.01, 0.05 and 0.10. CV^{α} represents the critical value of relevant test.

We see in Tables 3.42, 3.43, 3.44 of TS of length $l = 1000$ that when we use the ADF test if either $\phi_1^b > 1$ or$\phi_1^a > 1$ this leads to spurious cointegration in almost 100% of the cases. In contrast if we use the new lagged-series test in this case the spurious cointegration is fully avoided.

We see in Tables 3.45, 3.46, 3.47 of TS of length 100 that when we use the ADF test if either $\phi_1^b > 1$ or$\phi_1^a > 1$ this does not lead to much spurious cointegration; since the series are short they were not able to "explode". If we use the new lagged-series test the results are slightly improved.

3.3.2 Testing Pairs of Cointegrated Variables under Threshold Effects using Johansen

Model (3.23) is used to simulate pairs of cointegrated pairs of time series $x(t)$ and $y(t)$ with a threshold error series $z(t)$ as defined in Equation (3.19).

$$\begin{aligned} x(t) = x(t-1) + \epsilon_t \ ; \ t = 1, ..., l \ ; \ x(0) = \epsilon_0 \ ; \ \epsilon_t \sim N(0,1) \\ y(t) = \beta_1 x(t) + z(t) \ ; \ \beta_1 \sim \text{unif}(-20, 20) \end{aligned} \tag{3.23}$$

Table 3.48 summarizes the proportions of cointegrated pairs x(t) and y(t) using Johansen for various simulations performed under two sets of threshold variables ($\theta^L = -2, \theta^H = 2$) and ($\theta^L = -20, \theta^H = 20$) and different combinations of the auto-regressive multipliers for the high(ϕ_1^H), middle(ϕ_1^M) and low(ϕ_1^L) regimes. The column labeled $\overline{\frac{\hat{\beta}_1 - \beta_1}{\beta_1}}$ lists the average proportion error of the estimated cointegration multiplier and the column labeled $S_{\frac{\hat{\beta_1}-\beta_1}{\beta_1}}$ displays the standard deviation of the proportion error of the cointegration multiplier.

Table 3.48: Proportions of Two Cointegrated TS under Threshold Noise with m=1000, l=1000, s=12345

θ^L	θ^H	ϕ_1^H	ϕ_1^M	ϕ_1^L	$\overline{\frac{\hat{\beta_1}-\beta_1}{\beta_1}}$	$S_{\frac{\hat{\beta_1}-\beta_1}{\beta_1}}$	$\overline{\Gamma_{\text{co}} > CV_{\text{co}}^{0.10}}$	$\overline{\Gamma_{\text{co}} > CV_{\text{co}}^{0.05}}$	$\overline{\Gamma_{\text{co}} > CV_{\text{co}}^{0.01}}$
-2	2	1.000	1.000	1.000	5.23	97.94	0.11	0.06	0.01
-2	2	0.000	0.000	0.000	0.00	0.01	1.00	1.00	1.00
-2	2	1.000	0.900	1.000	1.27	9.26	0.10	0.05	0.01
-2	2	1.000	1.100	1.000	2.16	22.21	0.11	0.05	0.01
-2	2	1.009	1.100	1.009	0.12	0.59	**1.00**	**1.00**	**1.00**
-2	2	0.900	1.000	0.900	0.01	0.03	1.00	1.00	1.00
-2	2	0.900	1.100	0.900	0.03	0.56	1.00	1.00	1.00
-20	20	1.000	0.900	1.000	0.01	0.05	**1.00**	**1.00**	**1.00**
-20	20	1.000	1.100	1.000	1.92	32.65	**0.62**	**0.49**	**0.28**
-20	20	1.009	1.100	1.009	0.17	0.80	**1.00**	**1.00**	**1.00**
-20	20	0.900	1.000	0.900	0.60	6.58	0.17	0.10	0.03
-20	20	0.900	1.100	0.900	0.02	0.14	1.00	1.00	1.00

The low and high regimes determine in the long run if $z(t)$ is a stationary I(0) or nonstationary I(1) process, or a nonstationary explosive process ($\phi_1 > 1$). The largest estimation errors and standard deviations of the estimated β occur in all the cases when ($\phi_1^H = 1, \phi_1^L = 1$) except when ($\phi_1^H = 1, \phi_1^M = 0.9, \phi_1^L = 1$) when ($\theta^L = -20, \theta^H = 20$) because this process spends most of the time in the middle regime, so it is stationary for practical purposes since the other nonstationary regimes are not reached.

When ($\phi_1^H = 1.009, \phi_1^L = 1.009$) $z(t)$ is not stationary however the Johansen test always returns that $x(t)$ and $y(t)$ are cointegrated even though this is just a spurious cointegration. This is also the problem when ($\phi_1^H = 1, \phi_1^M = 1.1, \phi_1^L = 1$)and ($\theta^L = -20, \theta^H = 20$) since the $z(t)$ process spends a considerable amount of time in the middle regime which is explosive which results in spurious cointegration. These spurious cointegration cases with non-threshold setups.

Chapter 4: Combining Unit Root Tests with Machine Learning Techniques

4.1 Combining Unit Root Tests with SVM and DLNN Techniques

I propose two new types of unit root test based on combining multiple unit root tests, including the new the new lagged-series unit root test in this thesis, one with support vector machines(SVM) and one with deep learning neural networks(DLNN.) We train and test these configurations under two scenarios: one where there are intercepts and linear trends allowed, and the other case where there are intercepts allowed with strucutural breaks in the intercept.

The SVM and DLNN were trained on the following 3 input parameters that where provided for each simulation:

- The test statistic of the Zivot-Andrews URT.
- A binary value corresponding to the lagged-series URT that is 0 if the test accepts the alternative hypothesis of $H_1^{I(0)}$ for $\alpha = 0.05$, and 1 if the test fails to reject the null-hypothesis of $H_0^{I(1)}$.
- The p-value of the ADF URT.

And the expected output was provided as well for each training simulation:

- The output of 1 for an I(1) process and 0 for an I(0) process.

During the testing phase the 4 inputs were provided to the SVM and DLNN and the estimated output was compared to the actual output to determine accuracy.

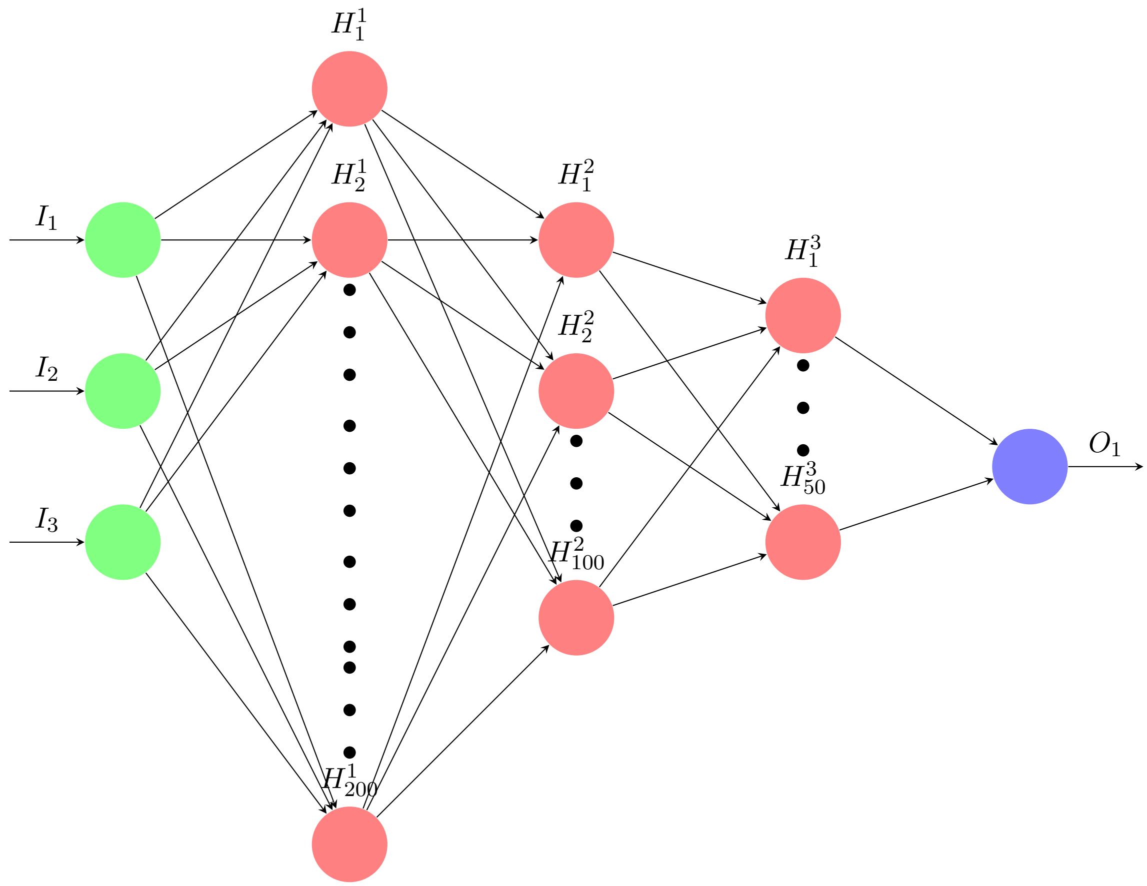


Figure 4.1: DLNN Architecture

The DLNN architecture, which is depicted in Figure 4.1, and settings were as follows:

- An input layer with 3 inputs at the very bottom as previously detailed.
- Three hidden layers from the bottom-up of sizes 200, 100 and 50 nodes.
- A top output layer with 1 node.
- All of the nodes between two layers are fully inter-connected with each other.
- The neural activation function of tanh with dropout.
- Hidden layer dropout ratios of 0.1, 0.01 and 0.01 for the layers with 200, 100 and 50 nodes respectively.
- The training data class counts are balanced via over/under-sampling for improved predictive accuracy. This was done since in each of the training cases there are 3000 I(1) TS and 7000 I(0) series.
- The number of passes of the training set (epochs) was set to 2000. This number can be increased for improving accuracy.

The DLNN software used is documented in H2O.ai Team (2016).

The settings for the SVM were:

- cost=1000. These are the cost of constraints violation; it is the “C”-constant of the regularization term in the Lagrange formulation.
- gamma=50; Gamma is the free parameter of the Gaussian radial basis function.
- The output of 1 for an I(1) process and 0 for an I(0) process.

The SVM software used is detailed in Meyer,David and Dimitriadou,Evgenia and Hornik,Kurt ,Weingessel,Andreas and Leisch,Friedrich (2015).

4.2 Combined Unit Root Tests with SVM and DLNN under Intercepts and Linear Trends

$$y(t) = \phi_1 x(t) + \alpha + \beta t$$

$$x(t) = x(t-1) + \epsilon(t) \ ; \ x(0) = 0 \tag{4.1}$$

$$\epsilon(t) \sim N(0,1) \ ; \ \mathrm{cor}(\epsilon(t), \epsilon(t-1)) = 0 \ ; \ t = 1, ..., l$$

The **training data** for this case without structural breaks was generated using the following randomly chosen intercepts and linear trends:

$$\alpha \sim \mathrm{unif}(-10, 10) + 5$$

$$\beta \sim \mathrm{unif}(-10, 10) + 5 \tag{4.2}$$

$$\mathrm{seed} \leftarrow 12345$$

The **testing data** for this case without structural breaks was generated using the following randomly chosen intercepts and linear trends:

$$\alpha \sim \mathrm{unif}(-1000, 1000) + 100$$

$$\beta \sim \mathrm{unif}(-1000, 1000) + 100 \tag{4.3}$$

$$\mathrm{seed} \leftarrow 54321$$

The tested values of ϕ_1 were 0, 0.1, 0.3, 0.5, 0.7, 0.9, 0.95, 0.960, 0.970, 0.980, 0.99, 0.995, 1, 1, 1, 1, 1, 1, 1.0025 and 1.005. There were 500 cases for each value of ϕ_1. The value of 1 was repeated 6 times to increase the number of unit root cases. That is a total of 7000 I(0) cases and 3000 I(1) cases. We will compare the SVM and DLNN methods to the standard ADF, Philips-Perron (PP) .

Table 4.1: Proportions of URT Training Errors under Intercept and Linear Trends

ϕ_1	DLNN	lagged-series	ERS-Ptest	ADF	PP	SVM	KPSS
0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.94
0.10	0.00	0.00	1.00	0.00	0.00	0.00	0.94
0.30	0.00	0.00	1.00	0.00	0.00	0.00	0.92
0.50	0.00	0.00	1.00	0.00	0.00	0.00	0.90
0.70	0.00	0.00	1.00	0.00	0.00	0.00	0.83
0.90	0.00	0.00	1.00	0.00	0.00	0.00	0.47
0.95	0.00	0.00	1.00	0.00	0.00	0.00	0.11
0.96	0.00	0.01	1.00	0.00	0.00	0.00	0.08
0.97	0.06	0.09	1.00	0.06	0.05	0.05	0.01
0.98	0.46	0.53	1.00	0.46	0.43	0.45	0.02
0.99	0.84	0.86	1.00	0.84	0.84	0.83	0.00
0.99	0.93	0.95	1.00	0.93	0.94	0.93	0.00
1.00	0.05	0.06	0.00	0.05	0.05	0.05	1.00
1.00	0.05	0.05	0.00	0.04	0.05	0.06	1.00
1.00	0.04	0.05	0.00	0.04	0.04	0.05	1.00
1.00	0.05	0.06	0.00	0.05	0.05	0.06	1.00
1.00	0.05	0.06	0.00	0.05	0.06	0.07	1.00
1.00	0.06	0.06	0.00	0.06	0.06	0.07	1.00
1.0025	0.75	0.74	1.00	0.98	0.98	0.64	0.00
1.005	0.04	0.04	0.99	1.00	1.00	0.04	0.00

Table 4.1 shows the proportion of training errors of the DLNN, and SVM machine learning techniques as well as other unit root tests and the KPSS stationarity test.

Table 4.2: Proportions of URT Testing Errors under Intercept and Linear Trends

ϕ_1	DLNN	lagged-series	ERS-Ptest	ADF	PP	SVM	KPSS
0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.93
0.10	0.00	0.00	1.00	0.00	0.00	0.00	0.94
0.30	0.00	0.00	1.00	0.00	0.00	0.00	0.94
0.50	0.00	0.00	1.00	0.00	0.00	0.00	0.90
0.70	0.00	0.00	1.00	0.00	0.00	0.00	0.83
0.90	0.00	0.00	1.00	0.00	0.00	0.00	0.39
0.95	0.00	0.00	1.00	0.00	0.00	0.00	0.09
0.96	0.01	0.01	1.00	0.01	0.00	0.01	0.07
0.97	0.07	0.11	1.00	0.07	0.06	0.07	0.05
0.98	0.43	0.48	1.00	0.43	0.39	0.43	0.01
0.99	0.82	0.84	1.00	0.82	0.80	0.81	0.00
0.99	0.93	0.95	1.00	0.93	0.93	0.93	0.00
1.00	0.05	0.04	0.00	0.04	0.05	0.07	1.00
1.00	0.04	0.05	0.00	0.04	0.05	0.05	1.00
1.00	0.06	0.07	0.00	0.06	0.07	0.07	1.00
1.00	0.05	0.05	0.00	0.05	0.05	0.06	1.00
1.00	0.06	0.06	0.00	0.06	0.06	0.06	1.00
1.00	0.04	0.05	0.00	0.04	0.04	0.04	1.00
1.0025	0.75	0.74	1.00	0.98	0.98	0.65	0.00
1.005	0.05	0.05	1.00	1.00	1.00	0.04	0.00

Table 4.3 shows the total training error proportions of the various techniques across null and alternative hypotheses. We see that DLNN, and SVM machine achieve the lowest error of 0.17. However we see in Table 4.1 that the I(1) errors ($\phi_1 = 1$) are higher for the SVM than for the DLNN. This is probably due to the DLNN technique over-samples the I(1) test cases to balance the classes and the SVM technique used does not.

Table 4.2 shows the proportion of testing errors of the DLNN, and SVM machine learning techniques as well as other unit root tests and the KPSS stationarity test.

Table 4.4 shows the total testing error proportions of the various techniques across null and alternative hypotheses. We see that SVM achieves the lowest error of 0.16, and DLNN

Table 4.3: Proportions of Total URT Training Errors under Intercept and Linear Trends

URT	Total Error Proportion
DLNN	0.17
SVM	0.17
lagged-series	0.18
ADF	0.23
PP	0.23
ERS-Ptest	0.70
KPSS	0.56

Table 4.4: Proportions of Total URT Testing Errors under Intercept and Linear Trends

URT	Total Error Proportion
SVM	0.16
DeepLearning	0.17
lagged-series	0.18
PP	0.22
ADF	0.23
KPSS	0.56
ERS-Ptest	0.70

the second lowest error of 0.17. However just as was the case during training we can see the same in testing in Table 4.2 where the I(1) errors ($\phi_1 = 1$) are higher for the SVM than for the DLNN.

Table 4.5: DLNN URT Training Confusion Proportion Matrix under Intercept and Linear Trends

	I(0)	I(1)
I(0)	0.78	0.22
I(1)	0.05	0.95

Table 4.5 shows the total training confusion proportion matrix for DLNN. We see there are 5% of the I(1) cases are considered I(0), which makes this machine learning technique a reasonable approach consistent with $\alpha = 0.05$.

Table 4.6: DLNN URT Testing Confusion Proportion Matrix under Intercept and Linear Trends

	I(0)	I(1)
I(0)	0.78	0.22
I(1)	0.05	0.95

Table 4.6 shows the total testing confusion proportion matrix for DLNN. We see there are 5% of the I(1) cases are considered I(0), which makes this machine learning technique a reasonable approach consistent with $\alpha = 0.05$.

Table 4.7: Correct DLNN URT Training Classifications under Intercept and Linear Trends

Correct Proportion	Incorrect Proportion
0.82	0.17

Table 4.7 shows that 82% of the training cases are classified correctly with DLNN.

Table 4.8: Correct DLNN URT Testing Classifications under Intercept and Linear Trends

Correct Proportion	Incorrect Proportion
0.83	0.17

Table 4.8 shows that 83% of the testing cases are classified correctly with DLNN.

4.3 Combined Unit Root Tests with SVM and DLNN Techniques under Structural Breaks in Intercept

$$y(t) = x(t) + \alpha_t + \beta t$$

$$x(t) = \phi_1 x(t-1) + \epsilon(t) \ ; \ x(0) = 0 \ ; \ \epsilon(t) \sim N(0,1)$$

$$\mathrm{cor}(\epsilon(t), \epsilon(t-1)) = 0$$

$$\alpha_t = \begin{cases} \alpha_1, & \text{if } t \geq t_u \\ \alpha_2, & \text{otherwise} \end{cases} \tag{4.4}$$

$$t_u \sim \mathrm{unif}(3, l-2) \ ; \ t = 1, ..., l$$

$$\alpha_1 = \mathrm{unif}(-A+B, B+A)$$

$$\alpha_2 = \mathrm{unif}(-A+B, B+A)$$

$$\beta = \mathrm{unif}(-A+B, B+A)$$

We use DGP (4.4) to simulate TS with breaks in level.

The break-time is randomly picked between the $0.25l$ and $0.75l$, where $l = 1000$ is the length of the time series.

The **training data** for this case with structural breaks was generated using the following randomly chosen intercepts and linear trends:

$$
\begin{aligned}
\alpha_1 &\sim \text{unif}(-1000, 1000) + 100 \\
\beta_1 &\sim \text{unif}(-1000, 1000) + 100 \\
\alpha_2 &\sim \text{unif}(-1000, 1000) + 100 \\
\text{seed} &\leftarrow 12345
\end{aligned}
\tag{4.5}
$$

The **testing data** for this case with structural breaks was generated using the following randomly chosen intercepts and linear trends:

$$
\begin{aligned}
\alpha_1 &\sim \text{unif}(-10, 10) + 5 \\
\beta_1 &\sim \text{unif}(-10, 10) + 5 \\
\alpha_2 &\sim \text{unif}(-10, 10) + 5 \\
\beta_2 &\sim \text{unif}(-10, 10) + 5 \\
\text{seed} &\leftarrow 54321
\end{aligned}
\tag{4.6}
$$

The tested values of ϕ_1 were 0, 0.1, 0.3, 0.5, 0.7, 0.9, 0.95, 0.960, 0.970, 0.980, 0.99, 0.995, 1, 1, 1, 1, 1, 1, 1.0025, 1.005. There were 500 cases for each value of ϕ_1. The value of 1 was repeated 6 times to increase the number of unit root cases. That is a total of 7000 I(0) cases and 3000 I(1) cases.

Table 4.9 show the proportions of training errors of the various techniques.

Table 4.10 show the proportions of testing errors of the various techniques.

Table 4.11 shows SVM and KPSS and then DLNN with the lowest training errors. We

Table 4.9: Proportions of URT Training Errors with Intercept Changepoints

ϕ_1	DLNN	lagged-series	ERS-Ptest	ADF	PP	SVM	KPSS
0.00	0.14	1.00	1.00	0.99	0.98	0.00	0.00
0.10	0.13	0.99	1.00	0.98	0.98	0.00	0.00
0.30	0.14	0.99	1.00	0.97	0.98	0.00	0.00
0.50	0.14	0.99	1.00	0.99	0.99	0.00	0.00
0.70	0.19	0.99	1.00	0.99	0.99	0.00	0.00
0.90	0.32	0.98	1.00	0.98	0.98	0.00	0.00
0.95	0.41	0.99	1.00	0.99	0.99	0.00	0.00
0.96	0.45	0.99	1.00	0.98	0.98	0.00	0.00
0.97	0.53	0.99	1.00	0.99	0.99	0.00	0.00
0.98	0.55	0.99	1.00	1.00	0.99	0.00	0.00
0.99	0.67	1.00	1.00	1.00	1.00	0.00	0.00
0.99	0.74	1.00	1.00	1.00	1.00	0.00	0.00
1.00	0.21	0.00	0.00	0.00	0.00	1.00	1.00
1.00	0.24	0.00	0.00	0.00	0.00	1.00	1.00
1.00	0.19	0.00	0.00	0.00	0.00	1.00	1.00
1.00	0.20	0.00	0.00	0.00	0.00	1.00	1.00
1.00	0.18	0.00	0.00	0.00	0.00	1.00	1.00
1.00	0.22	0.00	0.00	0.00	0.00	1.00	1.00
1.0025	0.91	0.99	1.00	1.00	1.00	0.00	0.00
1.005	0.69	0.69	1.00	1.00	1.00	0.00	0.00

Table 4.10: Proportions of URT Testing Errors with Intercept Changepoints

ϕ_1	DLNN	lagged-series	ERS-Ptest	ADF	PP	SVM	KPSS
0.00	0.05	0.34	1.00	0.05	0.03	0.00	0.03
0.10	0.03	0.27	1.00	0.03	0.03	0.00	0.05
0.30	0.03	0.24	1.00	0.03	0.04	0.00	0.05
0.50	0.01	0.20	1.00	0.01	0.06	0.00	0.05
0.70	0.01	0.10	1.00	0.01	0.05	0.00	0.09
0.90	0.02	0.06	1.00	0.02	0.03	0.00	0.09
0.95	0.07	0.12	1.00	0.07	0.08	0.00	0.05
0.96	0.13	0.19	1.00	0.13	0.13	0.00	0.02
0.97	0.24	0.33	1.00	0.26	0.24	0.00	0.03
0.98	0.53	0.60	1.00	0.54	0.51	0.00	0.01
0.99	0.84	0.87	1.00	0.86	0.83	0.00	0.00
0.99	0.90	0.91	1.00	0.91	0.90	0.00	0.00
1.00	0.05	0.04	0.00	0.03	0.04	1.00	1.00
1.00	0.05	0.04	0.00	0.05	0.05	1.00	1.00
1.00	0.07	0.06	0.00	0.06	0.07	1.00	1.00
1.00	0.06	0.05	0.00	0.04	0.05	1.00	1.00
1.00	0.06	0.05	0.00	0.05	0.06	1.00	1.00
1.00	0.06	0.05	0.00	0.05	0.04	1.00	1.00
1.0025	0.78	0.78	1.00	0.98	0.98	0.00	0.00
1.005	0.04	0.04	0.99	1.00	1.00	0.00	0.00

Table 4.11: Proportions of URT Total Training Errors with Intercept Changepoints

URT	Total Error Proportion
SVM	0.300
KPSS	0.300
DLNN	0.362
lagged-series	0.680
ADF	0.693
PP	0.693
ERS-Ptest	0.700

see in Table 4.9 that SVM and KPSS pretty much return the same answer no matter what– that the series is stationary. The DLNN tecnique is the only one that returns different values.

Table 4.12: Proportions of URT Total Testing Errors with Intercept Changepoints

URT	Total Error Proportion
DLNN	0.201
ADF	0.259
PP	0.261
lagged-series	0.268
SVM	0.300
KPSS	0.323
ERS-Ptest	0.700

Table 4.12 shows DLNN with the lowest training errors.

Table 4.13: DLNN URT Training Confusion Proportion Matrix with Intercept Changepoints

	I(0)	I(1)
I(0)	0.57	0.43
I(1)	0.21	0.79

Table 4.13 shows that the I(1) cases reported as I(0) happens 21% of the cases.

Table 4.14: DLNN URT Testing Confusion Proportion Matrix with Intercept Changepoints

	I(0)	I(1)
I(0)	0.74	0.26
I(1)	0.06	0.94

Table 4.14 when the drifts are smaller shows that the I(1) cases reported as I(0) happens 6% of the cases; this is close to a significance level of $\alpha = 0.05$.

Table 4.15: DLNN URT Correct Training Classifications with Intercept Changepoints

Correct Proportion	Incorrect Proportion
0.64	0.36

Table 4.15 shows that DLNN classifies 64%of the cases correctly.

Table 4.16: DLNN URT Correct Testing Classifications with Intercept Changepoints

Correct Proportion	Incorrect Proportion
DLNN 0.80	0.20

Table 4.16 shows that DLNN classifies 80%of the cases correctly.

4.4 Summary of SVM and DLNN Unit Root Tests

The unit root tests derived from combining the standard unit root tests as well as the new lagged-series URT using DLNN outperform all of the others when testing, not necessarily when training. The DLNN procedure is able to preserve more or less the test significance under the null hypothesis of $H_0^{I(1)}$ as the classical statistical based tests, and the overall error across null and alternative hypotheses is the smallest among all of the tests. In the structural break case, we train the DLNN with a very wide range of possible intercepts and trends, and it does not achieve the target 0.05 significance level – it is closer to 0.10. However when we test with a more reasonable range of trends, then the desired significance level for the null hypothesis is reached. This indicates that the DLNN architecture has derived a worthwhile combination of inputs, and has not simply built a model that overfits the training data. The SVM based test, while it has one of the smallest overall errors across all null and alternative hypotheses, it has high level of errors under the null hypothesis. We also see that the SVM methodology has the smallest training errors compared to all of the other methodologies, and these errors become much larger during the testing phase. This leads us to believe that the SVM model has a tendency to overfit I also believe that a significant reason for the performance of the DLNN architecture is the dropout methodology which reduces overfitting by randomly dropping out a number of the units with their connections with associated weights during training. Another issue is that the DLNN technique over-samples the I(1) training cases, as there are less than I(0) cases, to balance the classes. The SVM tested did not implement this feature.

This approach of combining multiple statistical tests using machine learning techniques could be employed for other hypothesis testing problems where there is no single test that outperforms the others. One example would be testing the equality of two population variances.

Chapter 5: Applications In Financial Time Series(TS)

5.1 Related Economic and Financial Theory

5.1.1 Covered Interest Rate Parity and Cross Currency Basis Swaps

The following overview of covered interest rate parity (CIP), and relationship with cross currency basis swaps is derived from Mazzi, Biagio (2013) as well as many detailed personal discussions with the author.

The CIP formula relates the price of the FX forward ($F_T^{X/Y}$) for maturity T with the spot FX ($F_0^{X/Y}$) and the discount factors from the current date to time T of both currencies (D_T^X, D_T^Y) as follows:

$$F_T^{X/Y} = F_0^{X/Y} \frac{D_T^Y}{D_T^X} \tag{5.1}$$

$F_T^{X/Y}$ is the number of units of currency Y one obtains for one unit of currency X at time T (the X/Y forward rate at time T). $F_0^{X/Y}$ is the current value of that exchange (the spot X/Y rate). D_T^Y and D_T^X are the discount factors at time T in currencies Y and X respectively. Equation ((5.1)) a very generic way of describing the CIP relationship as it does not specify how to compute the discount factors, which can be derived in various ways.

In the econometric community the currency basis is the value needed to restore the covered interest rate parity when it does not hold, that is that value b_F such as expressed in the following formula:

$$F_T^{X/Y} = F_0^{X/Y} \frac{(1 + R_X + b_F)^T}{(1 + R_Y)^T} \tag{5.2}$$

In Equation (5.2) all symbols have the same meaning as in Equation (5.1). R_X and R_Y are some interest rates for currencies X and Y. If b_F is zero we have a covered interest rate parity. Another term for b_F is "CIP deviation".

In the Fixed Income trading community by X-currency basis one means the number b_C that when paid on top of a (usually three months) $X-$London Interbank Offered Rate (LIBOR) in exchange for a flat $Y-$LIBOR (usually three months USD-LIBOR[1]), produces a par value in the swap. This instrument is called a cross-currency basis swap (CCBS) and formally we have

$$
\begin{aligned}
N^X \left\{ -1 + \sum_{i=1}^{T} \left(L_i^X + b_C \right) D_i^X \Delta_i + D_T^X \right\} \qquad\qquad (5.3) \\
= N^Y \left\{ -1 + \sum_{i}^{T} L_i^Y D_i^Y \Delta_i + D_T^Y \right\}
\end{aligned}
$$

where L_i^X and L_i^Y are the 3 month LIBOR rates in currencies X and Y setting at time T_{i-1} and paying at time T_i, D_i^X and D_i^Y are the discount factors at time T_i in currencies X and Y and Δ_i is the day count fraction. At the beginning and at the end there is an exchange of notional. A cross-currency basis swap compares the *relative* worth, so to speak, of two floating LIBOR rates and estimate it through the basis: by exchanging the notionals and applying therefore the covered interest rate parity, we eliminate the *absolute* difference between the two rates.

Overnight Indexed Swap Rates (OIS) are (typically geometric) averages of daily rates set by Central Banks. Since the 2009 financial crisis OIS rates have replaced LIBOR rates as the new proxy for risk free rates for various reasons. One reason is the market views payments with a lower reset frequency (say 3 months) resulting in a higher credit risk than

[1]There are some exceptions where the currency basis in Y is not paid against USD (for example for some Eastern European currencies where it is paid against EUR) or when it is not paid in the Y-leg of the swap (as in Chilean Pesos or Mexican Pesos) but usually this swap set up holds as general.

payments with higher reset frequency such as daily. Another reason for the preference of OIS over LIBOR is that OIS rates are the interest rates paid on posted collateral.

USD OIS is based on daily Fed Funds, and EONIA is based on daily European Central Bank rates. The majority of interest rate swaps, including cross currency basis swaps (CCBS) are still linked to LIBOR rates; this is probably because it would be too complex and would cause too much uncertainty to switch existing obligations to a new index. The market practice has evolved into using multiple curves to value trades where the coupon curves (such as 3 month LIBOR) is bootstrapped separately from the curve used for discounting (OIS.)

LIBOR rates as well as OIS rates are used for un-collateralized transactions between top rated private banks and Central Banks, which embed non-trivial credit risk as even top banks can default as proven in the 2009 financial crisis. A mystery and inconsistency of fixed income financial markets is given OIS and LIBOR rates are used in risky transactions as detailed before, why are these rates most widely used for over-the-counter (OTC) collateralized transactions, which bear hardly any credit risk at all.

The two values b_F and b_C are closely related in a formal way: a one payment cross-currency basis swap as given by Equation (5.3) could be made to look quite similar to Equation (5.2). Because of this both quantities would deserve to be called currency basis if by basis we define (see Mazzi, Biagio (2013)) a numerical value assigned to a financial anomaly: in the first case, given by Equation (5.2), the deviation from the covered interest rate parity and in the second, given by Equation (5.3), the assumption that a stream of floating rates discounted with the same rate should be worth par.

The first important difference is that whereas b_F is quantity that we can somehow imply (see Baba, Naohiko and Packer, Frank (2008) as one example out of a vast literature), b_C is a traded quantity. Not only b_C is a traded quantity but we know exactly, by the terms of the cross-currency basis swap, to what type of rates the basis is applied to. Admittedly there is no unique way of finding the discount factors. In order to imply b_F we need to make an assumptions about R_X and R_Y appearing in Equation (5.2): what kind of rates are they? It is difficult to say, but anecdotal evidence from the trading world (see

Mazzi, Biagio (2013)) would suggest that these rates used to be considered lower than the interbank lending rates (LIBOR): this anecdotal evidence seems to be supported by the work of Mancini Griffoli, Tommaso and Ranaldo, Angelo (2010) and McAndrews, James J. (2009). As with any implied quantity, the issue is far from clear and in the literature there are examples where LIBOR type rates (Coffey,N. and Hrung, W.B. and Sarkar, A. (2009)) were applied to estimate the deviation from the covered interest rate parity. As long as the choice is consistent, i.e. the rate in the two currencies is of the same nature (both interbank rates, both swap rates, both OIS rates, etc.), the impact of the choice should not be very large.

The literature on covered interest rate parity and in particular the deviation from it is vast and covers economic theory and finance. It starts with the seminal work of Frenkel, Jacob A. and Levich, Richard M. (1975) and Frenkel, Jacob A. and Levich, Richard M. (1977) and the work of Bilson, John F. O. (1981) who first draw the attention to the concept of "forward premium puzzle". It continues with works concentrating on the connection between interest rate derivatives and CIP deviation such as the one of Fletcher,Donna J. and Sultan, Jahangir (1997), Popper, Helen (1993) or Takezawa, Nobuya (1995) although the focus is on fixed for floating interest rate swaps rather than basis swaps. In the wake of the recent financial crisis the focus has moved to the relationship between CIP deviation and credit such as in the works of Baba, Naohiko and Packer, Frank (2008) and Baba, Naohiko and Packer, Frank (2009) or Genberg, Hans and Hui, C. H. and Wong, Alfred and Chung, T. K. (2009).

5.1.2 Bank Credit Default Swaps

A credit default swap (CDS) is an over-the-counter (OTC) financial contract where the buyer of the CDS pays premiums to the seller, typically on a regularly recurring basis, and has a reference bond or loan. If the bond issuer(or loan borrower) defaults (or triggers some other credit event) on the bond (or loan), then the CDS buyer would deliver one of the reference obligor's defaulted bonds (or loans) to the CDS seller in exchange for par, i.e.

the full bond (or loan) notional amount; this is the case of physical settlement of the CDS contract.

If it so happens that there are no available defaulted bonds (or loans) in the market, then an auction would be conducted to determine the bond's (or loan's) recovery rate and then the CDS contract would be cash settled. CDS are often used as a form of insurance by bond buyers (or loan creditors.)

5.1.3 Granger Causality

Consider an auto-regressive linear model:

$$X(t) = \beta_1 X(t-1) + ... + \beta_p X(t-p) + \epsilon(t) \tag{5.4}$$

And now consider extending the model with lagged versions of another predictor:

$$X(t) = \beta_1 X(t-1) + ... + \beta_p X(t-p) + \gamma_1 Y(t-1) + ... + \gamma_q X(t-q) + \epsilon(t) \tag{5.5}$$

We say that Y Granger-causes X if it is a useful predictor in Model (5.5). This can be tested using an F-test where the null hypothesis is that γ_i are zero. See Stock, James H. and Watson, Mark W. (2011, p. 538) for a more detailed overview.

5.2 FX Rates with Explosive Periods

We analyze two daily FX rate TS from 2012-09-10 until 2015-09-10, Brazilian Reals per 1 US Dollar (DEXBZUS) and Malaysian Ringgits per 1 US Dollar (DEXMAUS.) The data was obtained from the Federal Reserve Economic Data - FRED - St. Louis Fed. We see in Figures 5.1 and 5.2 that these series are very steep.

First as we perform a simulation test with AR(1) models to infer the multiplier in $X(t) = \beta X(t-1) + \epsilon_t$ by regressing $X(t)$ on $X(t-1)$ as summarized in Table 5.1. We see that as the actual multipliers become higher than 1 the variability of the estimated

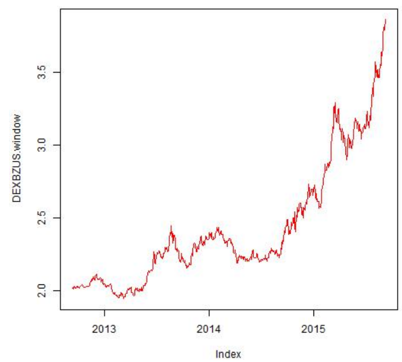


Figure 5.1: Brazilian Reals per 1 US Dollar.

Figure 5.2: Malaysian Ringgits per 1 US Dollar.

multipliers decreases significantly; this would give us confidence that this regression would help us find explosive AR(1) processes.

Table 5.1: Quantiles of Inferred AR(1) Multipliers of Series with $l = 500$ and $m = 500$ with Seed 1234

β	$\hat{\beta}$ Q0	$\hat{\beta}$ Q0.25	$\hat{\beta}$ Q0.50	$\hat{\beta}$ Q0.75	$\hat{\beta}$ Q1
1.050	1.050	1.050	1.050	1.050	1.050
1.020	1.020	1.020	1.020	1.020	1.020
1.010	0.966	1.010	1.010	1.010	1.013
1.000	0.944	0.986	0.991	0.995	1.004
0.990	0.928	0.974	0.982	0.987	0.998

So we perform the regression of $X(t)$ on $X(t-1)$ on the DEXBZUS and DEXMAUS TS and the results are summarized in Tables 5.2 and 5.3. In both cases the inferred multiplier is larger than 1.

Now we run the lagged-series and ADF unit root tests without linear trends and without constants on DEXBZUS and DEXMAUS and we can see that the ADF test does not reject I(1), however the lagged-series test does (for DEXBZUS at $\alpha = 0.10$, and for DEXMAUS

Table 5.2: DEXBZUS Auto-Regression Results

	Dependent variable:
	DEXBZUS.window
L(DEXBZUS.window, 1)	1.006^{***}
	(0.002)
Constant	-0.011^{**}
	(0.005)
Observations	752
R^2	0.997
Adjusted R^2	0.997
Residual Std. Error	0.024 (df = 750)
F Statistic	$241{,}712.100^{***}$ (df = 1; 750)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 5.3: DEXMAUS Auto-Regression Results

	Dependent variable:
	DEXMAUS.window
L(DEXMAUS.window, 1)	1.008^{***}
	(0.002)
Constant	-0.023^{***}
	(0.008)
Observations	752
R^2	0.996
Adjusted R^2	0.996
Residual Std. Error	0.016 (df = 750)
F Statistic	$198{,}162.900^{***}$ (df = 1; 750)
Note:	*p<0.1; **p<0.05; ***p<0.01

at $\alpha = 0.05$). The results are in Tables 5.4 and 5.5.

Table 5.4: Unit Root Tests on DEXBZUS

lagged-series I(1)	lagged-series Test Stat	lagged-series Crit Val	ADF I(1)	ADF P-value	α
FALSE	6.675	6.500	TRUE	0.990	0.100

Table 5.5: Unit Root Tests on DEXMAUS

lagged-series I(1)	lagged-series Test Stat	lagged-series Crit Val	ADF I(1)	ADF P-value	α
FALSE	14.153	8.180	TRUE	0.990	0.050

Now we perform a cointegration test of DEXBZUS and DEXMAUS which provides evidence for a cointegration relationship–however we would have not considered cointegration had we been using the lagged-series unit root test. The results of the cointegration test are in Results Listing 5.1. The fact that we have 2 variables and both the $r = 0$ and $r <= 1$ rank hypotheses can both be rejected at 5% significance level should make us suspicious of this cointegrating relationship in any case, and it is likely spurious.

Listing 5.1: Johansen Cointegration Test of DEXBZUS and DEXMAUS

```
######################
# Johansen-Procedure #
######################
Test type: maximal eigenvalue statistic (lambda max) , with linear trend
Eigenvalues (lambda):
[1] 0.02331532 0.01712828
Values of teststatistic and critical values of test:
          test 10pct  5pct  1pct
r <= 1 | 12.97  6.50  8.18 11.65
r = 0  | 17.72 12.91 14.90 19.19
Eigenvectors, normalised to first column:
(These are the cointegration relations)
           DEXBZUS.12 DEXMAUS.12
DEXBZUS.12  1.0000000   1.000000
DEXMAUS.12 -0.8202336  -1.719604
Weights W:
(This is the loading matrix)
           DEXBZUS.12   DEXMAUS.12
DEXBZUS.d 0.009335439 -0.025177956
DEXMAUS.d 0.010763129  0.007381374
```

;

5.3 Cross Currency Basis Swaps and Covered Interest Rate Parity

We define a spot FX rate at time[2] 0 as the number x in currency c1, per 1 unit of currency c2: $FX(0) = \frac{x_{c1}}{1_{c2}}$. We denote the discount factor for amounts in currency c at time $t = T$ as $DF_c(T)$ which present values those amounts to today $t = 0$. In Equation (5.6) I start from the covered interest rate parity formula and I re-arrange it as three separate terms, involving interest rates (R_c) in both currencies.

[2] Usually spot FX has an implicit number of settlement days which varies for currency pairs but is often 2 business days. Here we assume it is 0.

$$\mathrm{FX}(t) = \mathrm{FX}(0)\frac{\mathrm{DF}_{c2}(t)}{\mathrm{DF}_{c1}(t)}$$

$$\mathrm{FX}(t) = \mathrm{FX}(0)\frac{\left(1+\frac{R_{c1}}{f}\right)^{ft}}{\left(1+\frac{R_{c2}}{f}\right)^{ft}} \tag{5.6}$$

$$\frac{\mathrm{FX}(t)}{\mathrm{FX}(0)}\left(1+\frac{R_{c2}}{f}\right)^{ft} = \left(1+\frac{R_{c1}}{f}\right)^{ft}$$

Care must be take when applying this equation as FX rates have a preferred quoting convention; some are x XYZ per 1 USD and others are x USD per 1 XYZ: examples are $\frac{c1=JPY}{c2=USD}$ and $\frac{c1=USD}{c2=GBP}$.

I use the quarterly compounded ($f = 4$) swap rates in both currencies.

The three linear terms to test for cointegration are :

$$\begin{gathered}\log\left(\frac{\mathrm{FX}(t)}{\mathrm{FX}(0)}\right) + (ft)\log\left(1+\frac{R_{c2}}{f}\right) = (ft)\log\left(1+\frac{R_{c1}}{f}\right) \\ A + B = C\end{gathered} \tag{5.7}$$

In Table 5.6 we show the estimated coefficients of the A,B,C terms in Equation 5.7 with and without cross-currency basis. If covered interest rate parity were to hold perfectly then we would expect the coefficients for the three terms to be 1 for A, 1 for B and -1 for C. We see in general this is closer when we add the cross-currency basis to the interest rates, vs when we do not.

In the following Figures 5.3-5.20 we compare the XYZ interest rate inferred from USD swap rates to XYZ Swap Rate and also to XYZ Swap rate + CCBS Spread and we see that in every case when we add the CCBS Spread we get closer to the inferred rate.

Table 5.6: Estimated Coefficients for A,B and C Terms

Currency2	Term	A_{CCBS}	B_{CCBS}	C_{CCBS}	A	B	C
JPY	1	1.00	0.98	-1.04	1.00	1.54	-3.59
JPY	5	1.00	1.02	-1.13	1.00	0.77	-1.25
JPY	10	1.00	1.04	-1.18	1.00	0.77	-1.25
GBP	1	1.00	0.81	-0.86	1.00	0.57	-0.08
GBP	5	1.00	0.87	-1.23	1.00	0.04	0.43
GBP	10	1.00	0.85	-1.40	1.00	0.04	0.43
EUR	1	1.00	1.05	-1.58	1.00	1.04	-1.57
EUR	5	1.00	1.00	-1.59	1.00	0.93	-1.29
EUR	10	1.00	0.99	-1.70	1.00	0.93	-1.29
AUD	1	1.00	1.00	-0.99	1.00	0.97	-0.98
AUD	5	1.00	1.06	-1.69	1.00	1.04	-1.73
AUD	10	1.00	0.94	-1.67	1.00	1.04	-1.73
NZD	1	1.00	1.08	-1.65	1.00	1.06	-1.65
NZD	5	1.00	1.07	-1.78	1.00	1.03	-1.81
NZD	10	1.00	0.89	-1.54	1.00	1.03	-1.81
CHF	1	1.00	1.00	-1.27	1.00	1.01	-1.36
CHF	5	1.00	1.01	-1.21	1.00	0.87	-1.35
CHF	10	1.00	1.15	-1.48	1.00	0.87	-1.35

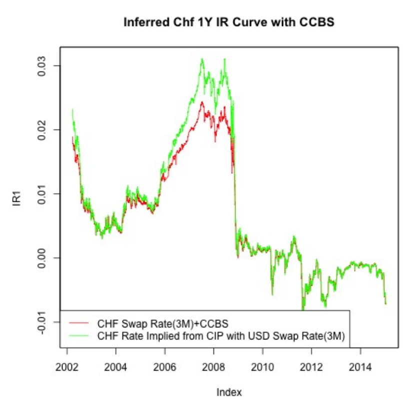


Figure 5.3: Inferred CHF 1Y Swap with CCBS Spread.

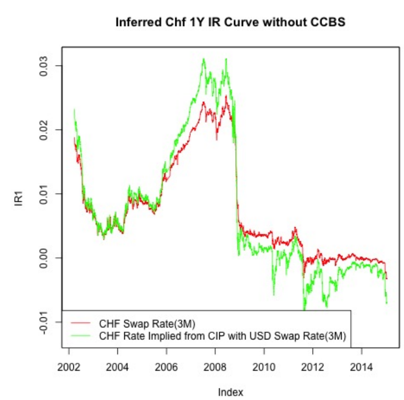


Figure 5.4: Inferred CHF 1Y Swap without CCBS Spread.

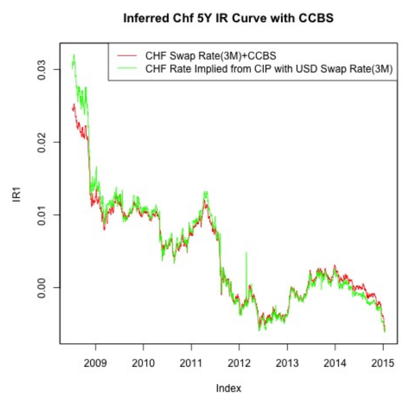


Figure 5.5: Inferred CHF 5Y Swap with CCBS Spread.

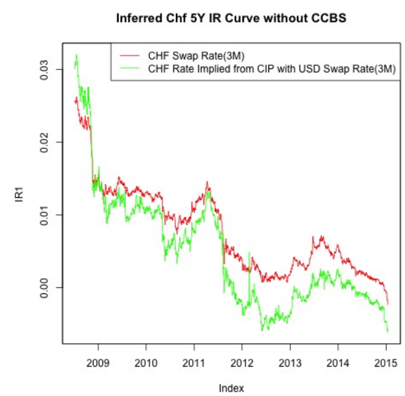


Figure 5.6: Inferred CHF 5Y Swap without CCBS Spread.

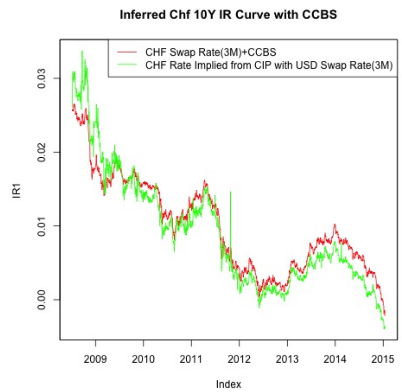


Figure 5.7: Inferred CHF 10Y Swap with CCBS Spread.

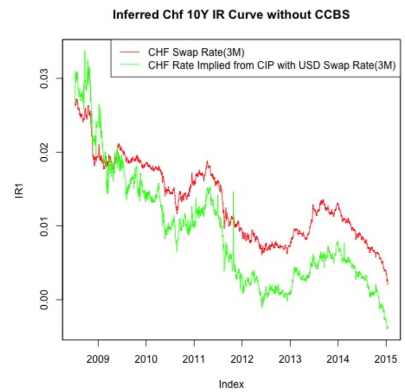


Figure 5.8: Inferred CHF 10Y Swap without CCBS Spread.

In the following Table 5.9 we see that the covered interest rate parity (CIP) relationship between the three terms detailed earlier always strengthens when we add CCBS spread to the XYZ swap rates.

First we check the 3 terms mentioned earlier in Equation (5.7) with with and without unit CCBS spread for unit roots with a linear trend and with a constant term. The terms are T1,T2 and T3 without the cross currency basis and T1 CCBSS,T2 CCBSS,T3 CCBSS with the cross currency basis. We see that in the majority of cases the terms are unit roots for a significance level of 0.01 as can be seen in Tables 5.7 using the lagged-series URT and 5.8 using the ADF URT.

In Table 5.9 we see that the cointegration relationship (with a linear trend) between the three CIP terms detailed earlier in Equation (5.7) always strengthens when we add CCBS spread to the XYZ swap rates.

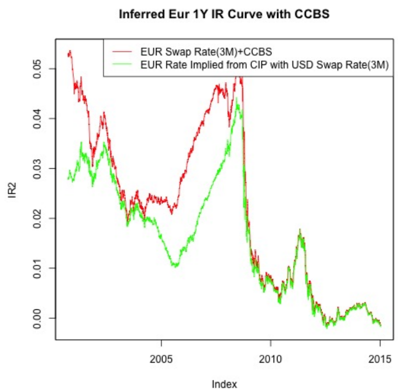


Figure 5.9: Inferred EUR 1Y Swap with CCBS Spread.

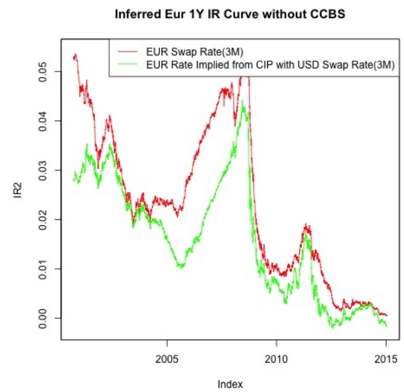


Figure 5.10: Inferred EUR 1Y Swap without CCBS Spread.

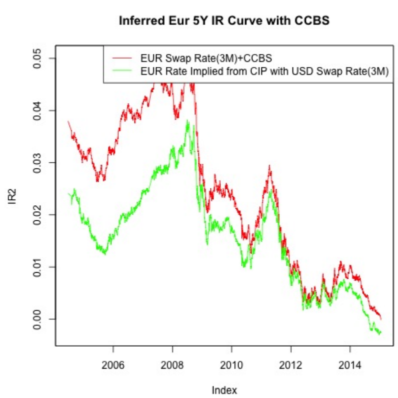


Figure 5.11: Inferred EUR 5Y Swap with CCBS Spread.

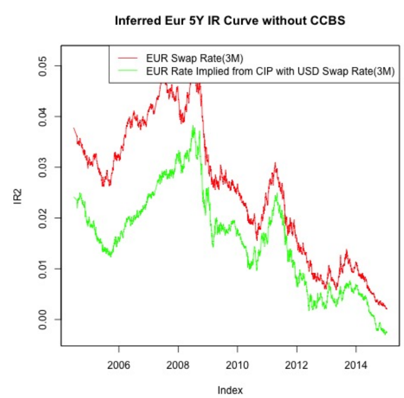


Figure 5.12: Inferred EUR 5Y Swap without CCBS Spread.

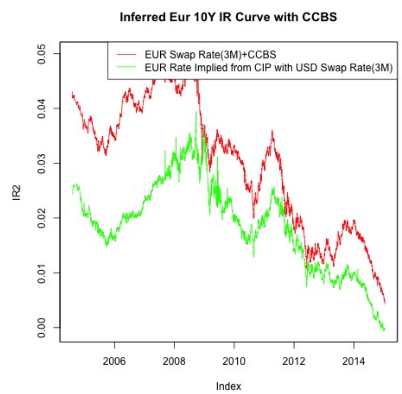


Figure 5.13: Inferred EUR 10Y Swap with CCBS Spread.

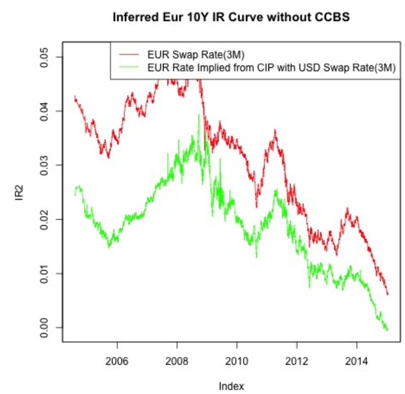


Figure 5.14: Inferred EUR 10Y Swap without CCBS Spread.

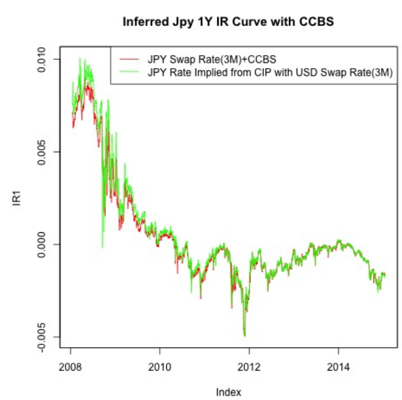


Figure 5.15: Inferred JPY 1Y Swap with CCBS Spread.

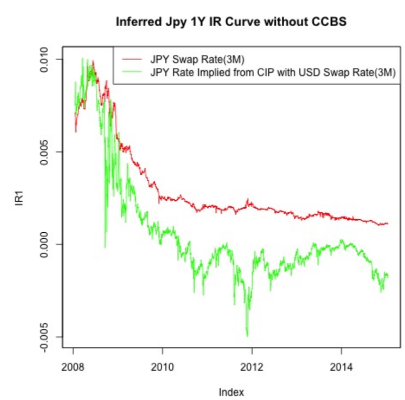


Figure 5.16: Inferred JPY 1Y Swap without CCBS Spread.

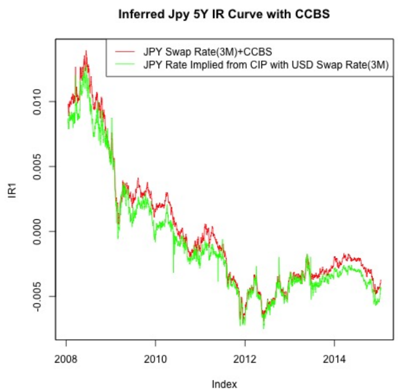


Figure 5.17: Inferred JPY 5Y Swap with CCBS Spread.

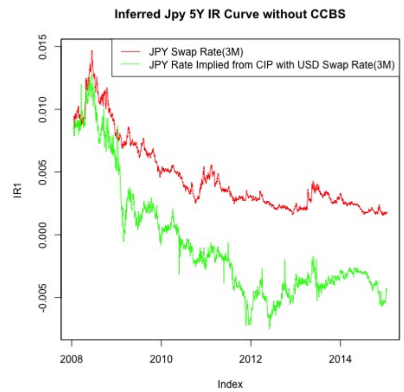


Figure 5.18: Inferred JPY 5Y Swap without CCBS Spread.

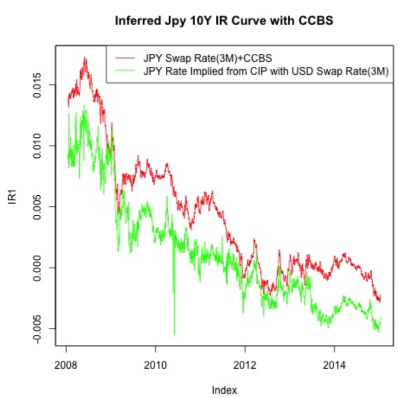


Figure 5.19: Inferred JPY 10Y Swap with CCBS Spread.

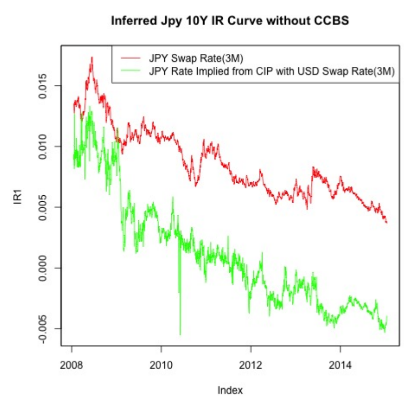


Figure 5.20: Inferred JPY 10Y Swap without CCBS Spread.

Table 5.7: **lagged-series** Unit Root Test Failures to Reject $H_0^{I(1)}$ of the 3 Terms in Equation (5.7) with and without CCBS Spread for 1,5 and 10 Years for $\alpha = 0.01$

Currency2	**Term Years**	**A CCBSS**	**B CCBSS**	**C CCBSS**	**A**	**B**	**C**
JPY	1	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
JPY	5	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
JPY	10	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
GBP	1	TRUE	**FALSE**	TRUE	TRUE	TRUE	TRUE
GBP	5	**FALSE**	TRUE	TRUE	**FALSE**	TRUE	TRUE
GBP	10	**FALSE**	TRUE	TRUE	**FALSE**	TRUE	TRUE
EUR	1	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
EUR	5	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
EUR	10	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
AUD	1	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
AUD	5	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
AUD	10	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
NZD	1	**FALSE**	TRUE	TRUE	**FALSE**	TRUE	TRUE
NZD	5	**FALSE**	TRUE	TRUE	**FALSE**	TRUE	TRUE
NZD	10	**FALSE**	TRUE	TRUE	**FALSE**	TRUE	TRUE
CHF	1	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
CHF	5	**FALSE**	TRUE	TRUE	**FALSE**	TRUE	TRUE
CHF	10	**FALSE**	TRUE	TRUE	**FALSE**	TRUE	TRUE

Table 5.8: **ADF** Unit Root Test Pvalues of the 3 Terms in Equation (5.7) with and without CCBS Spread for 1,5 and 10 Years

Currency2	Term Years	A CCBSS	B CCBSS	C CCBSS	A	B	C
JPY	1	0.08	0.42	0.21	0.08	0.42	0.90
JPY	5	0.06	0.51	0.76	0.06	0.51	0.56
JPY	10	0.10	0.40	0.43	0.10	0.40	0.07
GBP	1	0.31	0.50	0.70	0.31	0.61	0.70
GBP	5	0.02	0.67	0.56	0.02	0.73	0.56
GBP	10	**0.01**	0.54	0.42	**0.01**	0.54	0.42
EUR	1	0.59	0.84	0.80	0.59	0.91	0.80
EUR	5	0.77	0.70	0.50	0.77	0.72	0.50
EUR	10	0.60	0.69	0.30	0.60	0.74	0.30
AUD	1	0.81	0.68	0.95	0.81	0.74	0.95
AUD	5	0.38	0.47	0.47	0.38	0.54	0.47
AUD	10	0.09	0.24	0.09	0.09	0.26	0.09
NZD	1	0.10	0.77	0.90	0.10	0.80	0.90
NZD	5	**0.01**	0.55	0.48	**0.01**	0.61	0.48
NZD	10	**0.01**	0.66	0.38	**0.01**	0.71	0.38
CHF	1	0.70	0.96	0.85	0.70	0.96	0.94
CHF	5	0.02	0.36	0.25	0.02	0.36	0.12
CHF	10	0.02	0.35	0.50	0.02	0.35	0.46

Table 5.9: **Cointegration Statistic** of 3 Terms in Equation (5.7) for $r = 0$

Currency2	Term Years	With CCBBS	Without CCBBS
JPY	1	514.28	65.67
JPY	5	247.12	25.07
JPY	10	320.91	45.08
GBP	1	67.38	71.38
GBP	5	81.62	22.41
GBP	10	183.29	85.80
EUR	1	56.12	53.52
EUR	5	45.83	26.54
EUR	10	208.38	130.37
AUD	1	898.64	322.35
AUD	5	268.32	251.44
AUD	10	229.53	216.91
NZD	1	100.92	70.72
NZD	5	402.99	362.63
NZD	10	387.88	215.06
CHF	1	228.93	58.40
CHF	5	89.93	50.12
CHF	10	107.55	35.43

5.4 Bank Credit Default Swap and Cross Currency Basis Swap Spreads

See Sections 5.1.2 for an explanation of Credit Default Swaps and 5.1.1 for an introduction to cross-currency basis swap spreads (CCBS spreads.) First we compute average Bank Credit Default Swap (CDS) spreads on senior unsubordinated debt for Banks in the United States, Europe and Japan for 1 year, 5 year and 10 year maturities. I detail the market data sources.

Data Sources:

United States' Banks (US): Citigroup, Inc., Bank of America and JP Morgan Chase & Co.

- 1Y Bloomberg tickers: CINC.CDS.USD.SR.1Y.D14.Corp, BOFA.CDS.USD.SR.1Y.D14.Corp, JPMCC.CDS.USD.SR.1Y.D14.Corp

- 5Y Bloomberg tickers: CINC.CDS.USD.SR.5Y.D14.Corp, BOFA.CDS.USD.SR.5Y.D14.Corp, JPMCC.CDS.USD.SR.5Y.D14.Corp
- 10Y Bloomberg tickers: CINC.CDS.USD.SR.10Y.D14.Corp, BOFA.CDS.USD.SR.10Y.D14.Corp, JPMCC.CDS.USD.SR.10Y.D14.Corp

European Banks (EU): BNP Paribas, HSBC Bank, Le Crédit Lyonnais, S.A., and Deutsche Bank.

- 1Y Bloomberg tickers: BNP.CDS.EUR.SR.1Y.D14.Corp, HSBC.BK.CDS.EUR.SR.1Y.D14.Corp, LCL.SA.CDS.EUR.SR.1Y.D14.Corp, DB.CDS.EUR.SR.1Y.D14.Corp
- 5Y Bloomberg tickers: BNP.CDS.EUR.SR.5Y.D14.Corp, LCL.SA.CDS.EUR.SR.5Y.D14.Corp, DB.CDS.EUR.SR.5Y.D14.Corp, HSBC.BK.CDS.EUR.SR.5Y.D14.Corp
- 10Y Bloomberg tickers: BNP.CDS.EUR.SR.10Y.D14.Corp, HSBC.BK.CDS.EUR.SR.10Y.D14.Corp, LCL.SA.CDS.EUR.SR.10Y.D14.Corp, DB.CDS.EUR.SR.10Y.D14.Corp

Japanese Banks (JP): Sumitomo Mitsui Banking Corporation and Nomura Holdings.

- 1Y Bloomberg tickers: SMBC.CDS.JPY.SR.1Y.D14.Corp
- 5Y Bloomberg tickers: SMBC.CDS.JPY.SR.5Y.D14.Corp, NOMURAH.CDS.JPY.SR.5Y.D14.Corp
- 10Y Bloomberg tickers: SMBC.CDS.JPY.SR.10Y.D14.Corp

First we show that all of these CDS TS are likely unit-roots as evidenced in Table 5.10. We test for unit roots without linear trend and without constants.

Table 5.11 shows evidence there are cointegrating relationships between European Bank CDS spreads and EUR-USD cross currency basis swaps as well as other Bank CDS and EUR-USD combinations. We test for cointegration without a linear trend.

Table 5.12 shows evidence there are cointegrating relationships between European Bank CDS spreads, American Bank CDS spreads and EUR-USD CCBS as well as many other Two Bank Country CDS and CCBS combinations.

Table 5.10: Unit Root Tests of Country Bank CDS

Bank Country	Maturity	lagged-series Failure to Reject $H_0^{I(1)}$	ADF P-value	Start	End
US	1	TRUE	0.033	2003-05-15	2015-03-20
US	5	TRUE	0.071	2002-09-11	2015-03-20
US	10	TRUE	0.191	2005-01-12	2015-03-20
EU	1	TRUE	0.172	2004-03-15	2015-02-27
EU	5	TRUE	0.143	2002-07-22	2015-03-20
EU	10	TRUE	0.266	2004-03-12	2015-02-27
JP	1	TRUE	0.086	2008-06-18	2015-02-10
JP	5	TRUE	0.411	2005-04-08	2015-03-20
JP	10	TRUE	0.287	2004-02-02	2015-02-10
US	1	TRUE	0.135	2010-01-05	2015-03-20
US	5	TRUE	0.334	2010-01-01	2015-03-20
US	10	TRUE	0.389	2010-01-01	2015-03-20
EU	1	TRUE	0.224	2010-01-01	2015-02-27
EU	5	TRUE	0.315	2010-01-01	2015-03-20
EU	10	TRUE	0.387	2010-01-01	2015-02-27
JP	1	TRUE	0.116	2010-01-04	2015-02-10
JP	5	TRUE	0.378	2010-01-04	2015-03-20
JP	10	TRUE	0.383	2010-01-01	2015-02-10

Table 5.11: Johansen Cointegration Tests Between a Country Bank CDS Rate and CCBS. CV(10%,5%,1%)=(12.91, 14.9, 19.19).

Cointegration Variables	Γ_{Johansen}	Start	End	Cointegration Vector
EuCds5y, EurUsdCcbs5y	33.68	2002-07-22	2015-01-14	1.00, 3.64
EuCds5y, EurUsdCcbs5y	24.34	2010-01-01	2015-01-14	1.00, 3.65
UsCds5y, EurUsdCcbs5y	41.94	2002-09-11	2015-01-14	1.00, 4.44
UsCds5y, EurUsdCcbs5y	22.74	2010-01-01	2015-01-14	1.00, 4.30
EuCds5y, JpyUsdCcbs5y	24.98	2002-07-22	2015-01-14	1.00, 1.95
EuCds5y, JpyUsdCcbs5y	14.49	2010-01-01	2015-01-14	1.00, 3.41

Table 5.12: Various Cointegration Tests Between Two Countries' Bank CDS Rates and a CCBS. CV(10%,5%,1%)=(18.9, 21.07, 25.75).

Cointegration Variables	Γ_{Johansen}	Start	End	Cointegration Vector
EuCds1y, UsCds1y, EurUsdCcbs1y	40.61	2004-03-15	2014-09-15	1.00, -0.14, 2.10
EuCds1y, UsCds1y, EurUsdCcbs1y	32.50	2010-01-05	2014-09-15	1.00, 0.28, 3.58
EuCds5y, UsCds5y, EurUsdCcbs5y	52.37	2002-09-11	2015-01-14	1.00, 1.05, 8.18
EuCds5y, UsCds5y, EurUsdCcbs5y	28.45	2010-01-01	2015-01-14	1.00, 0.30, 4.95
EuCds10y, UsCds10y, EurUsdCcbs10y	43.38	2005-01-12	2014-04-09	1.00, 11.62, 74.21
EuCds10y, UsCds10y, EurUsdCcbs10y	18.75	2010-01-01	2014-04-09	1.00, -14.90, -55.51
JpCds1y, UsCds1y, JpyUsdCcbs1y	40.77	2008-06-18	2015-01-14	1.00, -0.02, 4.77
JpCds1y, UsCds1y, JpyUsdCcbs1y	25.49	2010-01-05	2015-01-14	1.00, -0.50, -1.22
JpCds5y, UsCds5y, JpyUsdCcbs5y	28.96	2005-04-08	2015-01-14	1.00, -1.11, 0.23
JpCds5y, UsCds5y, JpyUsdCcbs5y	37.74	2010-01-04	2015-01-14	1.00, -1.22, 0.02
JpCds10y, UsCds10y, JpyUsdCcbs10y	37.45	2005-01-21	2015-01-14	1.00, -1.01, -0.46
JpCds10y, UsCds10y, JpyUsdCcbs10y	44.32	2010-01-01	2015-01-14	1.00, -1.13, -0.43

Table 5.13 shows evidence there are cointegrating relationships between Japanese Bank CDS spreads, American Bank CDS spreads and JPY-USD spot FX rates as well as many other two Bank Country CDS and FX combinations.

Table 5.13: Various Cointegration Tests Between Two Country Bank CDS Rates and Spot FX. CV(10%,5%,1%)=(18.9, 21.07, 25.75).

Cointegration Variables	Γ_{Johansen}	Start	End	Coint Vector
JpCds5y, UsCds5y, UsdJpy	32.26	2005-04-08	2015-01-14	1.00, -1.19, 0.05
JpCds5y, UsCds5y, UsdJpy	36.47	2010-01-04	2015-01-14	1.00, -1.29, -0.39
EuCds5y, UsCds5y, EurUsd	20.18	2002-09-11	2015-01-14	1.00, -0.97, 163.96
EuCds5y, UsCds5y, EurUsd	16.82	2010-01-01	2015-01-14	1.00, -0.84, 59.88

5.5 Interest Rate Swap and Cross-Currency Basis Swap Liquidity

Figure 5.21 shows the daily liquidity of USD-JPY 1 year *cross currency basis swap* (CCBS) liquidity as the sum of the JPY leg notionals. Figure 5.22 shows the daily liquidity of JPY 1 year *interest rate swap* liquidity as the sum of the JPY leg notionals. These TS do not appear to be unit roots. In any case we will perform This data was obtained from the Depository Trust & Clearing Corporation (DTCC) swap data repository (SDR), referred to as the DTCC Data Repository (DDR). On December 31, DDR began accepting data from swap dealers for over-the-counter (OTC) trades as outlined by the Dodd Frank Act (DFA) and the Commodity Futures Trading Commission's (CFTC) real-time and regulatory reporting rules.

The Press Release announcing the availability of the DDR can be found here: `http://www.dtcc.com/news/2013/january/03/swap-data-repository-real-time.aspx`. This is an excerpt:

> "DDR is now publishing real-time price information. Since the December 31 swap dealer reporting deadline, DDR has disseminated more than 10,000 records, which represent the vast majority of the reportable OTC derivatives market. Reports are available through file transfers, RSS feeds and internet access to a ticker page, Excel and search functions on DRR's website,`https://rtdata.dtcc.com/gtr/dashboard.do`"

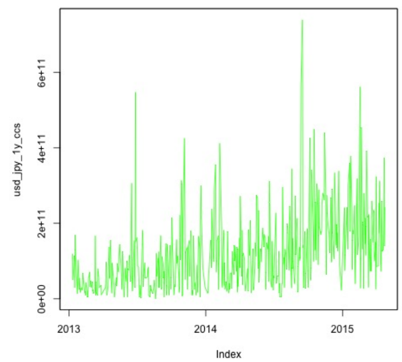


Figure 5.21: USD JPY 1Y CCBS Liquidity in Net USD Notional per Day.

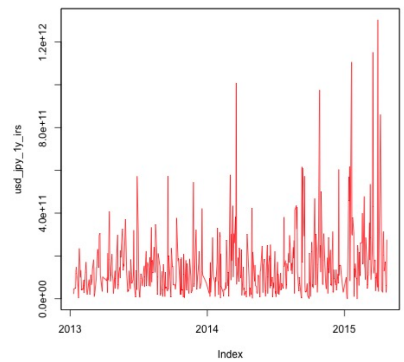


Figure 5.22: JPY 1Y Fixed-Float Swap Liquidity in Net JPY Notional per Day.

Listing 5.2: URTs for $\ln(\mathrm{CCBS(t)}^{usd/jpy})$ and $\ln(\mathrm{IRS}(t)^{\mathrm{jpy}})$

```
> adf.test(log_usd_jpy_1y_irs)
Dickey-Fuller = -7.1271, Lag order = 7, p-value = 0.01
alternative hypothesis: stationary
> adf.test(log_usd_jpy_1y_ccs)
Dickey-Fuller = -6.6781, Lag order = 7, p-value = 0.01
alternative hypothesis: stationary
```

Listing 5.2 displays the results of ADF URTs on the two series being tested, $\ln(\mathrm{CCBS(t)}^{usd/jpy})$ and $\ln(\mathrm{IRS}(t)^{\mathrm{jpy}})$, and we can see that in both cases we can reject the null hypothesis of a unit-root, so we can assume they are stationary I(0) processes. If instead of using ADF we were to use the lagged-series URT we reach the same conclusions.

Listing 5.3: VAR Model Lag Selection for ln(CCBS(t)$^{usd/jpy}$) and ln(IRS(t)$^{\text{jpy}}$)

```
>irs_ccs_log <- merge(log_jpy_1y_irs,log_usd_jpy_1y_ccs)
>colnames(irs_ccs_log) <- c('log_jpy_1y_irs','log_usd_jpy_1y_ccs')
>VARselect(irs_ccs_log, type='both')
$selection
AIC(n)  HQ(n)  SC(n) FPE(n)
     5      1      1      5
> v1 <- vars::VAR(irs_ccs_log,p=1,type='both')
```

To perform the selection of the lag for the VAR model to use for Granger causality testing I used the VARselect() function of the R vars package. Listing 5.3 show the results and I picked 1 based on the Hannan-Quinn criterion (HQ) and Schwarz-Bayes criterion (SC), and since it is a much simpler model than using a lag of 5, as suggested by the Akaike Information Criterion (AIC) and Akaike's Final Prediction Error (FPE) criterion.

Listing 5.4: Granger Causality Tests of ln(CCBS(t)$^{usd/jpy}$) and ln(IRS(t)$^{\text{jpy}}$)

```
> causality(v1, cause = "log_usd_jpy_1y_ccs", boot=TRUE, boot.runs= 100000 )$Granger
        Granger causality H0: log_usd_jpy_1y_ccs do not Granger-cause log_jpy_1y_irs
data:  VAR object v1
F-Test = 6.3908, boot.runs = 1e+05, p-value = 0.00893
> causality(v1, cause = "log_jpy_1y_irs", boot=TRUE, boot.runs= 100000 )
$Granger
        Granger causality H0: log_jpy_1y_irs do not Granger-cause log_usd_jpy_1y_ccs
data:  VAR object v1
F-Test = 2.2413, boot.runs = 1e+05, p-value = 0.1114
$Instant
        H0: No instantaneous causality between: log_jpy_1y_irs and log_usd_jpy_1y_ccs
data:  VAR object v1
Chi-squared = 0.040932, df = 1, p-value = 0.8397
```

Listing 5.4 shows the results of two Granger-causality tests on 1 year USD-JPY CCBSs and 1 year JPY interest rate swaps, as well as an instantaneous causality test between them.

- The first Granger causality test rejects the null and accepts that $\ln(\text{CCBS(t)}^{usd/jpy})$ Granger-causes the 1 year $\ln(\text{IRS}(t)^{\text{jpy}})$
- and the second causality test does not reject the null so the 1 year $\ln(\text{IRS}(t)^{\text{jpy}})$ does not Granger-cause $\ln(\text{CCBS(t)}^{usd/jpy})$
- the Instant causality test between $\ln(\text{CCBS(t)}^{usd/jpy})$ and $\ln(\text{IRS}(t)^{\text{jpy}})$ is rejected

These tests provide statistical evidence that while there is no instant causality (correlation) between $\ln(\text{IRS}(t)^{\text{jpy}})$ and $\ln(\text{CCBS(t)}^{usd/jpy})$, and that trading activity in JPY-USD cross currency basis swaps appears to Granger Cause trading activity in JPY interest rate swaps, and not the other way around. One potential explanation for this is that often when a USD-JPY cross currency swap is traded a JPY floating-fixed interest rate swap is traded in conjunction. This can happen if a USD based entity would issue fixed-rate debt denominated in JPY (with a very low interest rate) and engages in a JPY-USD cross currency swap to hedge some or all of the FX risk. However when a JPY fixed-floating swap is traded, it is not necessarily the case that a related cross-currency swap would be traded.

Chapter 6: Unit Roots in Other Fields

The bulk of the literature on unit roots and cointegration spans mostly the social sciences, largely focusing on economic and financial applications. However there are other fields where they may be proving to be useful tools.

Kipinski, L. and Konig, R and Sieluzycki, C. and Kordecki, W. (2011) used Phillips-Perron unit root tests as well as other statistical tests to investigate the stationarity of magnetoof (MEG) or electroencephalography (EEG) time series (TS) , and they found they are largely stationary. However, Koruek, Mehmet and Ozkaya, Ata (2010) analyzed electroencephalogram (EEG) TS before, during and after seizures, and they determined that short interictal (between seizures) EEG series were nonstationary and could be modeled as an ARIMA unit root process.

There have been various studies on testing worldwide zonal temperature anomalies for unit roots under various assumptions of strucutural changes as in Coggin, T. Daniel (2012), Romilly, Peter (2005) and Ivanov, Martin A. and Evtimov, Stilian N. (2010). In the next section we will use some of the existing unit root tests as well as some that were developed in this thesis.

6.1 Zonal Temperature Anomalies

I analyzed multiple TS of global zonal temperature anomalies provided by NASA/Goddard Institute for Space Studies(GISS). The details of the data source are:

- Annual mean Land-Ocean Temperature Index in .01 degrees Celsius for selected zonal means.

- Sources: GHCN-v3 1880-09/2015 + SST: 1880-09/2015 ERSST v4 using elimination of outliers and homogeneity adjustment.
- The data an be found in `http://data.giss.nasa.gov/gistemp/graphs_v3/Fig.B.txt`.

Details of this data are explained in the Frequently Asked Questions section of the GISS Surface Temperature Analysis (GISTEMP) website: `http://data.giss.nasa.gov/gistemp/FAQ.html#q103`

We quote verbatim the explanation of the Land-Ocean Temperature index which refers to temperature anomalies over land, water and see-ice:

> Q. What is L-OTI, the Land-Ocean Temperature Index? A. Weather stations reporting surface air temperatures (SATs) are positioned on land, which covers only one third of the planet; the rest is covered by oceans where SAT reports are rare. However, water temperatures (SSTs, sea surface temperatures) are available from ship and buoy reports, and more recently there are also SST estimates derived from satellite data. Whereas SATs and SSTs may be very different (since air warms and cools much faster than water), their anomalies are very similar (if the water temperature is 5 degrees above normal, the air right above the water is also likely to be about 5 degrees warmer than normal). This is not true in the presence of sea ice, since in that case water temperature will stay at the freezing level. This allows us to use SST anomalies as proxies for SAT anomalies in regions without sea ice. L-OTI maps show SAT anomalies over land and sea ice, and show SST anomalies over (ice-free) water.

The definition of a meteorological year is from December 1 to November 30, as explained by GISS:

> Q. What is a meteorological year? A. When comparing seasonal temperatures, it is convenient to use "meteorological seasons" based on temperature

and defined as groupings of whole months. Thus, Dec-Jan-Feb is the Northern Hemisphere meteorological winter, Mar-Apr-May is N.H. meteorological spring, Jun-Jul-Aug is N.H. meteorological summer and Sep-Oct-Nov is N.H. meteorological autumn. String these four seasons together and you have the meteorological year that begins on Dec. 1 and ends on Nov. 30.

Instead of temperatures, anomalies are used as explained by GISS:

> Q. What are temperature anomalies (and why prefer them to absolute temperatures)? A. Temperature anomalies indicate how much warmer or colder it is than normal for a particular place and time. For the GISS analysis, normal always means the average over the 30-year period 1951-1980 for that place and time of year. This base period is specific to GISS, not universal. But note that trends do not depend on the choice of the base period: If the absolute temperature at a specific location is 2 degrees higher than a year ago, so is the corresponding temperature anomaly, no matter what base period is selected, since the normal temperature used as base point is the same for both years.
>
> Note that regional mean anomalies (in particular global anomalies) are not computed from the current absolute mean and the 1951-80 mean for that region, but from station temperature anomalies. Finding absolute regional means encounters significant difficulties that create large uncertainties. This is why the GISS analysis deals with anomalies rather than absolute temperatures. For a more detailed discussion of that topic, please see "The Elusive Absolute Temperature".

As GISS explains the reason for using the 1951-1980 base period is for consistency purposes, but it is not the only standard:

> Q. Why stick with the 1951-1980 base period? A. The primary focus of the GISS analysis are long-term temperature changes over many decades and centuries, and a fixed base period makes the anomalies consistent over time.

However, organizations like the NWS who are more focused on current weather conditions work with a time frame of days, weeks, or at most a few years. In that situation it makes sense to move the base period occasionally, i.e., to pick a new "normal" so that roughly half the data of interest are above normal and half below.

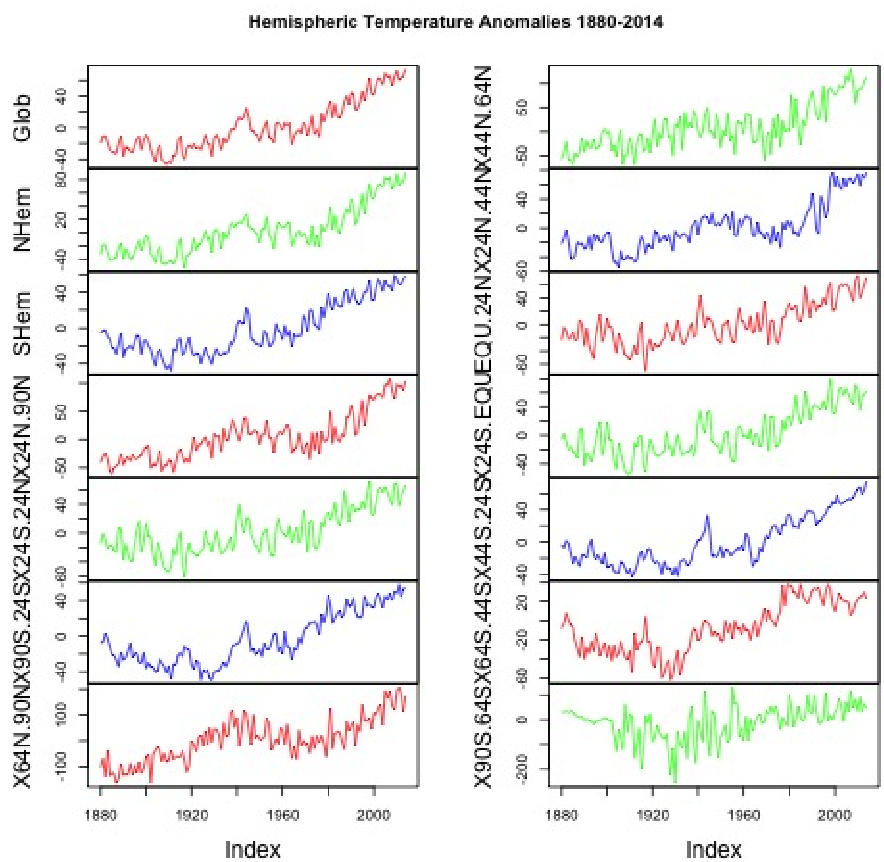


Figure 6.1: Zonal Temperature Anomalies.

Table 6.1: ADF P-value For Zonal Temperature Anomalies 1880-2014

ADF P-value	Zone TS	Reject $H_0^{I(1)}$ $\alpha = 0.01$
0.74	Glob	FALSE
0.92	NHem	FALSE
0.33	SHem	FALSE
0.94	X24N.90N	FALSE
0.36	X24S.24N	FALSE
0.21	X90S.24S	FALSE
0.81	X64N.90N	FALSE
0.65	X44N.64N	FALSE
0.85	X24N.44N	FALSE
0.39	EQU.24N	FALSE
0.30	X24S.EQU	FALSE
0.57	X44S.24S	FALSE
0.05	X64S.44S	FALSE
0.06	X90S.64S	FALSE

6.1.1 Unit Root Tests of Temperature Anomalies

Table 6.1 displays the p-values of the ADF URT on the zonal temperature anomaly yearly TS being analyzed for the entire period of 1880 to 2014. We see that the null-hypothesis of a unit root ($H_0^{I(1)}$) is not rejected in any of the cases. If we look back to the URT simulation results in the case of structural breaks under the null, and alternative hypotheses as detailed in Tables 3.22 and 3.23 we see there that the ADF URT case barely rejects.

Table 6.2: BBC Critical Values

$\alpha = 0.10$ $Q : 0.90$	$\alpha = 0.05$ $Q : 0.95$	$\alpha = 0.01$ $Q : 0.99$
15.77	17.90	22.23

Table 6.3 displays the BBC results of the URT under a three-regime threshold (SETAR)

Table 6.3: BBC Test Statistics For Zonal Temperature Anomalies 1880-2014

Γ_{BBC}	Zone TS	Reject $H_0^{I(1)}$ $\alpha = 0.01$
5.53	Glob	FALSE
5.90	NHem	FALSE
9.59	SHem	FALSE
4.03	X24N.90N	FALSE
12.62	X24S.24N	FALSE
9.97	X90S.24S	FALSE
2.82	X64N.90N	FALSE
8.00	X44N.64N	FALSE
4.11	X24N.44N	FALSE
11.79	EQU.24N	FALSE
10.71	X24S.EQU	FALSE
8.73	X44S.24S	FALSE
18.09	**X64S.44S**	**TRUE**
32.07	**X90S.64S**	**TRUE**

model. Other than the last two TS on the table the results are the same as with the ADF URT in Table 6.1, where the null-hypothesis of a unit root ($H_0^{I(1)}$) is not rejected.

Table 6.4: Zivot Andrews Critical Values

$\alpha = 0.01$ $Q : 0.01$	$\alpha = 0.05$ $Q : 0.05$	$\alpha = 0.10$ $Q : 0.10$
-5.34	-4.80	-4.58

Table 6.5 displays the results of the the ZA URT on the zonal temperature anomaly yearly TS being analyzed. We see that the null-hypothesis of a unit root ($H_0^{I(1)}$) is rejected in most of the cases. If we look back to the URT simulation results in the case of structural breaks under the null, and alternative hypotheses as detailed in Tables 3.22 and 3.23 we see there that for the ZA URT rejects almost always for the alternative hypothesis case of

Table 6.5: Zivot Andrews Test Statistics For Zonal Temperature Anomalies 1880-2014

Γ_{ZA}	Zone TS	Reject $H_0^{I(1)}$ $\alpha = 0.01$
-5.54	Glob	TRUE
-6.39	NHem	TRUE
-5.86	SHem	TRUE
-7.23	X24N.90N	TRUE
-7.34	X24S.24N	TRUE
-5.31	**X90S.24S**	**FALSE**
-9.86	X64N.90N	TRUE
-9.91	X44N.64N	TRUE
-6.42	X24N.44N	TRUE
-7.17	EQU.24N	TRUE
-7.49	X24S.EQU	TRUE
-4.39	**X44S.24S**	**FALSE**
-5.61	X64S.44S	TRUE
-10.21	X90S.64S	TRUE

$\phi_1 < 0$ ($H_1^{I(0)}$), and it also rejects considerably under the null-hypothesis of $\phi_1 = 1$ ($H_0^{I(1)}$). This is because the ZA test did not consider breaks under the null and only under the alternative hypotheses.

These URT results leads us to testing the zonal temperature anomalies using URTs that allow structural breaks under the null hypothesis, namely the Lee Strazicich, and the newly developed HBPZA URTs.

Table 6.6: HBPZA Critical Values

$\alpha = 0.01$ $Q : 0.01$	$\alpha = 0.05$ $Q : 0.05$	$\alpha = 0.10$ $Q : 0.10$
-6.07	-4.63	-4.34

Table 6.7 displays the results of the the newly developed HBPZA URT on the zonal temperature anomaly yearly TS being analyzed that allows structural breaks under the

Table 6.7: HBPZA Test Statistics For Zonal Temperature Anomalies 1880-2014

Γ_{HBPZA}	Zone TS	Reject $H_0^{I(1)}$ with Breaks $\alpha = 0.01$
-4.96	Glob	FALSE
-4.81	NHem	FALSE
-5.29	SHem	FALSE
-4.94	**X24N.90N**	**FALSE**
-6.51	X24S.24N	TRUE
-4.50	X90S.24S	FALSE
-5.42	X64N.90N	FALSE
-5.53	X44N.64N	FALSE
-4.51	X24N.44N	FALSE
-7.87	EQU.24N	TRUE
-6.73	X24S.EQU	TRUE
-4.78	X44S.24S	FALSE
-3.93	**X64S.44S**	**FALSE**
-7.03	X90S.64S	TRUE

null. We see that the null-hypothesis of a unit root ($H_0^{I(1)}$) is rejected in 4 out of 14 TS. If we look back to the URT simulation results in the case of structural breaks under the null, and alternative hypotheses of the HBPZA and Lee Strazicich URTs as detailed in Tables 3.24 we see that these tests perform better that the standard URTs even on a time series length of 100. The TS X64S.44S and X24N.90N and in the subset of series where the tests fail to reject the null hypothesis of a unit root.

Table 6.8: Lee Strazicich URT Critical Values

λ	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$
0.10	-5.11	-4.50	-4.21
0.20	-5.07	-4.47	-4.20
0.30	-5.15	-4.45	-4.18
0.40	-5.05	-4.50	-4.18
0.50	-5.11	-4.51	-4.17

Table 6.9 displays the results of the Lee Strazicih URT on the zonal temperature anomaly yearly TS being analyzed that allows structural breaks under the null. We see that the null-hypothesis of a unit root ($H_0^{I(1)}$) is rejected in 6 out of 14 TS. The zonal TSs X64S.44S and X24N.90N are also in the subset of series where the Lee Strazicih URTs fail to reject the null hypothesis of a unit root, just as they were when the HBPZA URT was used.

6.1.2 Structural Breaks in Temperature Anomalies

We use the Bai, J. and Perron, P. (1998) methodology that was used to develop the HBPADF testing methodology previously, to test here for structural breaks in the zonal temperature anomalies. First the Regression Model (6.1) is employed to test for breaks in the coefficients with an intercept:

$$x(t) = \phi_1 x(t) + \beta_0 + \epsilon(t) \tag{6.1}$$

Under the null hypothesis of a unit root ($\phi_1 = 1$) Model (6.1) results in a linear time

Table 6.9: Lee Strazicich URT Statistics For Zonal Temperature Anomalies 1880-2014

Γ_{LS}	Zone TS	Reject $H_0^{I(1)}$ with Breaks $\alpha = 0.05, \lambda = 0.5$
-3.91	Glob	FALSE
-4.14	NHem	FALSE
-5.10	SHem	FALSE
-3.98	**X24N.90N**	**FALSE**
-5.31	X24S.24N	TRUE
-4.68	X90S.24S	FALSE
-5.92	X64N.90N	TRUE
-5.42	X44N.64N	TRUE
-5.04	X24N.44N	FALSE
-5.19	EQU.24N	TRUE
-5.38	X24S.EQU	TRUE
-4.58	X44S.24S	FALSE
-3.88	**X64S.44S**	**FALSE**
-10.33	X90S.64S	TRUE

Table 6.10: Estimated Breakpoints for Zonal Temperature Anomalies 1880-2014 without a linear trend

Break	Zone TS
1918	X24N.90N
1918	X64N.90N
1918	X44N.64N
1918	X64S.44S
1928	NHem
1932	Glob
1934	X64S.44S
1935	X24N.44N
1965	X44S.24S
1968	X90S.64S
1973	X64S.44S
1974	EQU.24N
1975	Glob
1975	SHem
1975	X24S.24N
1975	X90S.24S
1975	X24S.EQU
1984	NHem
1986	X24N.90N
1986	X44N.64N
1993	X64N.90N
1995	Glob
1996	X24N.44N

trend in levels. Table 6.10 displays the breaks inferred with Regression Model (6.1). We see that most zones have an estimate of two break periods.

We also use the Regression Model (6.2) to test for breaks which adds a linear time trend in the AR formula:

$$x(t) = \phi_1 x(t) + \beta_0 + \beta_1 t + \epsilon(t) \tag{6.2}$$

Table 6.11: Estimated Bai-Perron Breakpoints for Zonal Temperature Anomalies 1880-2014 with AR(1) with a Linear Trend and Intercept

Break	Zone TS
1929	X90S.64S
1931	X64S.44S
1935	X64N.90N
1943	X90S.24S
1943	X44S.24S
1944	SHem
1944	X24S.EQU
1962	Glob
1962	**NHem**
1962	X24N.90N
1962	X24S.24N
1962	X44N.64N
1970	X64N.90N
1992	X24N.44N
NA	EQU.24N

Under the null hypothesis of a unit root ($\phi_1 = 1$) Model (6.2) results in a quadratic time trend in levels. Table 6.11 displays the breaks inferred with Regression Model (6.1). We see that most zones have an estimate of one break period, except for EQU.24N which has none.

We perform the HBPADF URT testing procedure where we use the ADF URT on each of the two sections. A number of the zones fail to reject the null-hypothesis of a unit root in both time periods; these include X64S.44S and X24N.90N, which also showed up when we did not include a time trend.

Table 6.12: HBPADF Estimated AR Break Models for Zonal Temperature Anomalies 1880-2014 with a time trend

Zone	Break Year	Break Prop	ADF Pvalue$_A$	$\beta_{0,A}$	$\phi_{1,A}$	$\beta_{1,A}$	ADF Pvalue$_B$	$\beta_{0,B}$	$\phi_{1,B}$	$\beta_{1,B}$
X90S.64S	1929	0.37	0.02	37.6	0.1	-2.1	0.01	-136.1	0.0	1.5
X64S.44S	1931	0.38	**0.34**	-8.9	0.4	-0.4	**0.34**	-18.1	0.7	0.2
X64N.90N	1935	0.41	**0.06**	-132.6	0.1	2.8	**0.79**	-49.9	0.5	0.8
X90S.24S	1943	0.47	0.78	-5.5	0.8	0.1	0.07	-44.6	0.5	0.5
X44S.24S	1943	0.47	0.72	-5.7	0.8	0.1	0.08	-33.9	0.7	0.4
SHem	1944	0.48	0.90	-6.2	0.8	0.1	0.04	-61.7	0.4	0.7
X24S.EQU	1944	0.48	0.12	-11.7	0.5	0.1	0.01	-72.5	0.3	0.9
Glob	1962	0.61	0.33	-10.8	0.7	0.1	0.05	-139.5	0.1	1.5
NHem	1962	0.61	**0.44**	-16.8	0.6	0.3	**0.24**	-148.9	0.3	1.6
X24N.90N	1962	0.61	**0.52**	-26.9	0.5	0.5	**0.16**	-197.5	0.2	2.1
X24S.24N	1962	0.61	0.24	-11.3	0.5	0.1	0.02	-112.5	0.1	1.3
X44N.64N	1962	0.61	0.40	-40.8	0.1	0.7	0.02	-247.8	0.0	2.7
X24N.44N	1992	0.83	**0.42**	-11.9	0.6	0.1	**0.66**	-156.0	0.4	1.6
EQU.24N	NA	NA	0.39	-16.4	0.6	0.3	NA	NA	NA	NA

Ivanov, Martin A. and Evtimov, Stilian N. (2010) provide statistical arguments for a change in trend in the Northern Hemisphere in 1963. We can see in Table 6.11 that our estimated changepoint for the Northern Hemisphere(NHem) was in 1962. The authors' provided these data source details:

> The first is the HADCRUT3 (Brohan P. and Kennedy JJ. and Harris I. and Tett SFB. and Jones PD. (2006)) version of combined land and marine data, the second is the CRUTEM3 (Brohan P. and Kennedy JJ. and Harris I. and Tett SFB. and Jones PD. (2006)) series of land air data and the third series is HADSST2 (Rayner NA. and Brohan P. and Parker DE. and Folland CK. and Kennedy JJ. and Vanicek M. and Ansell TJ. and Tett SFB. (2006)) for marine data. The anomalies are against the 1961–1990 climatology and span the 1850–2007 period.

6.1.3 Cointegration of X24N.90N and X64S.44S Temperature Anomalies

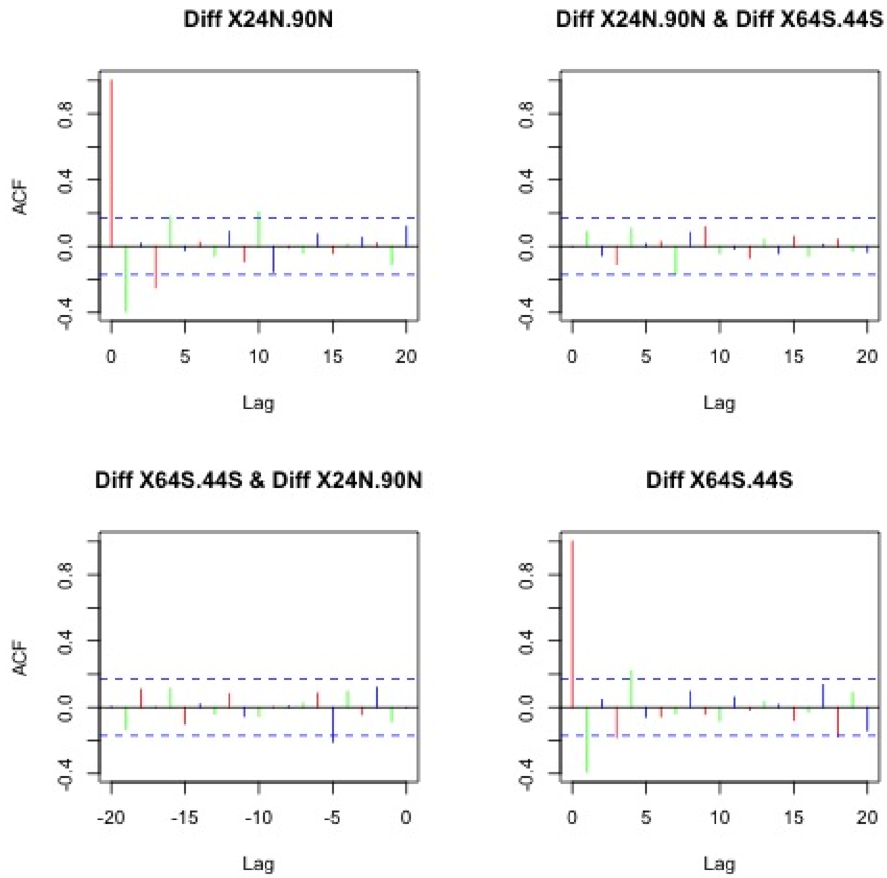


Figure 6.2: Correlation of Differences of X24N.90N and Differences of X64S.44S

Figure 6.2 shows that the differences of the X24N.90N and X64S.44S TS are not cross-correlated. There seems to be some autocorrelation with X24N.90N differences with a lag of 1 and 3, and with X64S.44S differences for a lag of 1.

Listing 6.1: Cointegration Test of X24N.90N, X64S.44S with no trend,nor constant

```
######################
# Johansen-Procedure #
######################
Test type: maximal eigenvalue statistic (lambda max) , with linear trend
Eigenvalues (lambda):
[1] 0.037395208 0.002030704
Values of teststatistic and critical values of test:
          test 10pct  5pct  1pct
r <= 1 | 0.27   6.50   8.18 11.65
r = 0  | 4.99  12.91  14.90 19.19
Eigenvectors, normalised to first column:
(These are the cointegration relations)
             X24N.90N.l5 X64S.44S.l5
X24N.90N.l5     1.000000   1.0000000
X64S.44S.l5    -2.803177  -0.2642928
Weights W:
(This is the loading matrix)
             X24N.90N.l5 X64S.44S.l5
X24N.90N.d -0.04004946 0.016427154
X64S.44S.d  0.03017643 0.009176225
```

Results Listing 6.1 shows that there is no evidence of cointegration using the Johansen cointegration test without a linear trend.

Listing 6.2: Cointegration Test of X24N.90N, X64S.44S with a linear trend

```
#####################
# Johansen-Procedure #
#####################
Test type: maximal eigenvalue statistic (lambda max) , with linear trend
in cointegration
Eigenvalues (lambda):
[1] 1.156881e-01 2.513105e-02 6.541002e-18
Values of teststatistic and critical values of test:
          test 10pct  5pct  1pct
r <= 1 |  3.33 10.49 12.25 16.26
r = 0  | 16.11 16.85 18.96 23.65
Eigenvectors, normalised to first column:
(These are the cointegration relations)
            X24N.90N.l5 X64S.44S.l5   trend.l5
X24N.90N.l5    1.000000   1.0000000  1.0000000
X64S.44S.l5    2.405250  -1.1594729 -0.1883196
trend.l5      -2.394985  -0.9555919 -0.2977645
Weights W:
(This is the loading matrix)
            X24N.90N.l5  X64S.44S.l5      trend.l5
X24N.90N.d -0.01828637 -0.067173804 -1.001798e-15
X64S.44S.d -0.09748640  0.006931988 -4.774705e-16
```

Results Listing 6.2 shows that there is no evidence of cointegration using the Johansen cointegration test with a linear trend.

Listing 6.3: Lutkepohl Cointegration Test of X24N.90N, X64S.44S with a Linear Trend

```
######################
# Johansen-Procedure #
######################
Test type: trace statistic , with linear trend in shift correction
Eigenvalues (lambda):
[1] 0.1163060 0.0289185
Values of teststatistic and critical values of test:
          test 10pct  5pct  1pct
r <= 1 |  3.73  5.42  6.79 10.04
r = 0  | 18.15 13.78 15.83 19.85
Eigenvectors, normalised to first column:
(These are the cointegration relations)
            X24N.90N.l1 X64S.44S.l1
X24N.90N.l1    1.000000   1.0000000
X64S.44S.l1    1.778241  -0.3498645
Weights W:
(This is the loading matrix)
         X24N.90N.l1 X64S.44S.l1
X24N.90N -0.07876826 -0.14403678
X64S.44S -0.13769160  0.04069908
```

Results Listing 6.3 shows that there is evidence of cointegration using the Lutkepohl cointegration test with a linear trend with a significance level of $\alpha = 0.05$.

Listing 6.4: Lutkepohl Cointegration Test of X24N.90N, X64S.44S without a Linear Trend

```
######################
# Johansen-Procedure #
######################
Test type: trace statistic , without linear trend in shift correction
Eigenvalues (lambda):
[1] 0.097385138 0.006518081
Values of teststatistic and critical values of test:
          test 10pct  5pct  1pct
r <= 1 |  0.85  3.00  4.12  6.89
r = 0  | 13.02 10.45 12.28 16.42
Eigenvectors, normalised to first column:
(These are the cointegration relations)
            X24N.90N.l1 X64S.44S.l1
X24N.90N.l1     1.00000   1.0000000
X64S.44S.l1   -47.33876  -0.3765719
Weights W:
(This is the loading matrix)
          X24N.90N.l1 X64S.44S.l1
X24N.90N -0.005076601 -0.03503179
X64S.44S  0.003196657 -0.02415221
```

Results Listing 6.4 shows that there is evidence of cointegration using the Lutkepohl cointegration test without a linear trend with a significance level of $\alpha = 0.05$.

The Lutkepohl cointegration test allows for structural breaks in the level, but not in the linear trend. Other earlier analyses may indicate potential breaks in trend–so it is unclear what we can conclude.

6.2 Three Regime SETAR Models of NHem and SHem

We use the auto.arima() function of the R package to build the best possible fit models to the Northern Hemisphere and the Southern Hemisphere temperature anomalies. We also build

3-regime SETAR models of these TS using the setar() function of the tsDyn R package (Di Narzo, Antonio and Di Narzo, Fabio and Aznarte,Jose Luis and Stigler,Matthieu, 2009). For known thresholds the SETAR model can be fit using conditional least squares; the setar() function performs a grid search to find the thresholds that minimize the RMSE. We consider for both ARIMA and SETAR models versions with intercepts and linear trends and without any deterministic components.

We allow for a constant and linear trend in the auto-regressive process which is equivalent to a quadratic trend [finish.]

Table 6.13: Model Fits of NHem and SHem TS

Zone	Model Type	Drift	Shapiro p-value	Box Pierce p-value	S_{resid}
NHem	ARIMA(1,1,1)	TRUE	0.42	0.86	13.14
NHem	SETAR 3-regimes	TRUE	0.02	0.68	11.91
SHem	ARIMA(1,1,1)	TRUE	0.61	0.81	9.55
SHem	SETAR 3-regimes	TRUE	0.62	0.26	8.92
NHem	ARIMA(0,1,2)	FALSE	0.32	0.77	13.15
NHem	SETAR 3-regimes	FALSE	0.61	0.01	13.98
SHem	ARIMA(1,1,1)	FALSE	0.64	0.87	9.57
SHem	SETAR 3-regimes	FALSE	0.66	0.04	10.06

Table 6.13 shows model various model fits of the NHem and SHem TS. These include the following information regarding the residuals:

- the p-value of the Shapiro-Wilk test of normality; the null hypothesis is that the residuals are Gaussian
- The Box-pierce test of independence of values in a time series; the null hypothesis is that the residuals are independent of each other
- the standard deviation of the residuals (S_{resid})

We see that:

- The versions of the models with time trends and intercepts(drift) always produce the best results.
- The SETAR 3-regime models provide the best fits in terms of lowest RMSE and lowest standard deviation of residuals (S_{resid}).
- The fits are better for the Southern Hemisphere than they are for the Northern Hemisphere.

Listing 6.5: NHem 3 Regime SETAR Model

```
Non linear autoregressive model
SETAR model ( 3 regimes)
Coefficients:
Low regime:
   const.L      trend.L      phiL.1
-5.8130703   0.2855119   0.8795019
Mid regime:
    const.M       trend.M        phiM.1
-19.1155054    0.2297229    1.0047473
High regime:
       const.H          trend.H           phiH.1
-2.530446e+02   2.580310e+00   7.563802e-03
Threshold:
-Variable: Z(t) = + (1) X(t)
-Value: -16  29
Proportion of points in low regime: 34.07%        Middle regime: 48.15%
High regime: 17.78%
```

Listing 6.5 shows the fitted 3-regime SETAR model of the Northern Hemisphere temperature anomalies (NHem.) We see that there is a middle AR(1) regime that is explosive. Whyte, JM and Metcalfe, AV (2011) proposed a similar model they refer to as MTAR using temperature anomaly data from Burgundy, France. The constants are negative but the linear trends are positive indicating a diverging relationship further exacerbated by the explosive middle regime.

Listing 6.6: SHem 3 Regime SETAR Model

```
Non linear autoregressive model
SETAR model ( 3 regimes)
Coefficients:
Low regime:
   const.L     trend.L      phiL.1
-8.3633052   0.1050809   0.7485235
Mid regime:
    const.M      trend.M       phiM.1
-47.7776797    0.6255392    0.3539460
High regime:
    const.H      trend.H       phiH.1
-58.4964850    1.0803433   -0.5868101
Threshold:
-Variable: Z(t) = + (1) X(t)
-Value: 10 35
Proportion of points in low regime: 68.15%       Middle regime: 15.56%
High regime: 16.3%
```

Listing 6.6 shows the fitted 3-regime SETAR model of the Southern Hemisphere temperature anomalies (SHem.) None of the regimes are explosive and appear to all be stationary. The high regime has a negative $\phi_1 < 0$ value which encourages an oscillatory behavior in that regime. The constants are negative but the linear trends are positive indicating a diverging relationship.

We simulate 200 stepwise paths starting from the beginning of the SHem series and using the standard deviation of the residuals as the standard deviation of the error terms, and plot them in Figure 6.4. We can see that this model is fairly well behaved.

We simulate 200 step-wise paths starting from the beginning of the NHem series and using the standard deviation of the residuals as the standard deviation of the error terms, and plot them in Figure 6.6. We can see that this model is less stationary than the one of

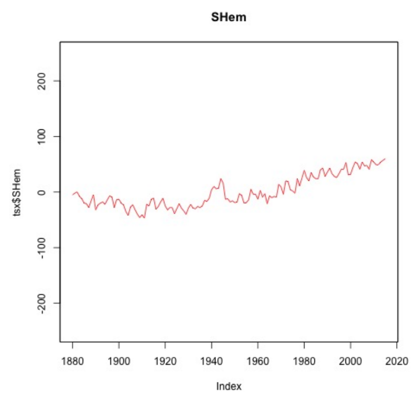


Figure 6.3: SHem Original Data

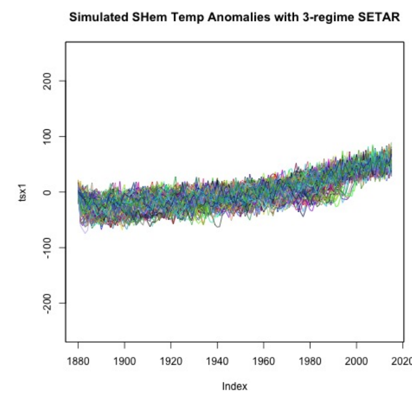


Figure 6.4: 200 SHem SETAR Simulations with $\sigma = 8.916936$

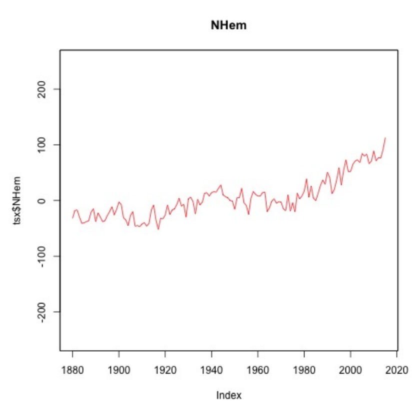


Figure 6.5: NHem Original Data

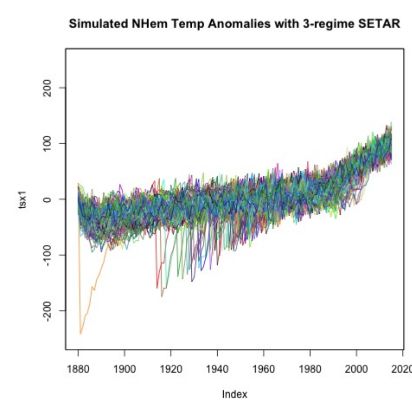


Figure 6.6: 200 NHem SETAR Simulations with $\sigma = 11.91102$

the SHem series.

A likely reason for the differences between hemispheric models is in the Southern Hemisphere, the ratio of land to water is one-third that in the Northern Hemisphere (Mielke, HW, 1989). Oceans help stabilize temperatures. Another factor may be that only 13% of the world's population lives in the Southern Hemisphere (Warren, Stephen G, 2015).

Chapter 7: Conclusions

As Choi, In (2010, p. 57) states, no unit root test is dominant, and the author recommends trying as many tests as possible and seeing if there are differences. In this thesis I proposed new tests with certain advantages over existing ones, I compared existing and new tests under various scenarios and had practical applications in areas of Finance and Temperature Anomalies. More specifically these were the research contributions of this thesis:

- I showed via simulation studies that when testing for cointegration of two slightly explosive TS ($\phi_1 > 1$) results almost invariably in spurious cointegration; this is not mentioned in the literature. None of the standard established unit root tests analyzed in this thesis such as the ADF URT reject the null hypothesis of I(1) in the case of $\phi_1 > 1$. See Tables 3.36, 3.37 and 3.38.

- I developed a new unit root test (URT)–the lagged-series URT which has similar statistical power to the Augmented Dickey Fuller (ADF) test when the auto-regressive multiplier $\phi_1 < 1$ but exceeds the power of the ADF test when $\phi_1 > 1$[1]. We show empirically via simulation studies that a valid[2] cointegration relationship between a TS and its lag implies an I(1) TS. This cointegration test will never result in a case of no possible cointegrating relationships for a reasonable lag.

- The new lagged-series test does not reject I(2) series similarly to the ADF and ERS URTs. However the new lagged-series URT rejects at least 65% of I(3) tests and 85% of I(4) tests; by comparison the ADF URT only rejected 5% of the I(3) and I(4) tests.

[1] Chandra, Suresh K. and Janhavi, J.V. (2008) developed a modified ADF URT that is supposed to reject explosive unit roots. I did not have access to this test so I cannot make any comparisons with the lagged-series test.

[2] When testing a TS with its lag if it has two cointegrating relationships this implies that the series is I(0). Two variables can only have one valid cointegration relationship.

- I combined the new lagged-series (URT) with other unit root tests such as ERS and ZA tests which improve the power of these tests when the AR(1) auto-regressive multiplier $\phi_1 > 1$.

- I developed and tested the HBPZA unit root test which allows for structural breaks in intercept and linear trend under the null hypothesis and compared it to an existing implementation of the Lee-Stratizich URT that also allows breaks in the null. The new test performs better than the Lee-Stratizich[3] in a number of situations.

- I developed the HBPADF testing procedure that allows discerning I(0)-I(1) shifts within the evaluated time series, and performed simulation studies on it.

- I combined URTs with deep learning neural networks (DLNNs) which outperform individual tests when we consider the net error across null and alternative hypotheses.

- I performed simulation based studies of the ADF, ERS-Ptest, ERS-DFGLS, the Zivot-Andrews, and the new lagged-series URTs, under various Model configurations. Findings are consistent with the literature that points out these tests are sensitive to the starting value of the unit root. Also when there are structural breaks in the null hypothesis the standard unit root tests hardly ever reject the null hypothesis of a unit root.

- I used a linear form of the covered interest rate parity (CIP) formula and showed that if cross-currency swap basis spreads are added to one of the 3 terms the cointegration relationship always strengthens.

- I showed that there are likely cointegration relationships between bank credit default swap spreads and cross currency basis swap spreads. This would provide evidence bank credit risk is related to cross currency basis swap spreads. I showed that there

[3] I am in contact with Johannes Lips,the author of the R implementation of the Lee-Stratizich URT. This test appears to reject much more than expected under the null hypothesis, so it is not certain if this is a bug in the software or an issue with the procedure itself.

are cointegration relationships between bank credit default swap spreads and spot FX. This would indicate that bank credit risk affects the FX Spot rate.

- I analyzed interest rate and cross currency swap liquidity and ensured the series are not unit roots, and provided statistical evidence that USD-JPY cross-currency basis swaps Granger cause JPY fixed-floating IR swaps. A possible explanation may be USD based entities issuing JPY fixed debt and hedging it fully or partially with USD-JPY cross currency basis swaps.

- I used various analyses to check for unit roots in zonal temperature anomaly time series including under structural breaks in the null hypothesis using the new HBPZA test as well as the HBPADF testing procedure.A number of model were fitted to the temperature anomaly data, including a 3 regime SETAR model and showed that better fits are always achieved when adding linear trends.

- I showed that the estimated 3-regime SETAR models for the Southern Hemisphere temperature anomalies are more likely to be stationary than the Northern Hemisphere, which includes an AR explosive ($\phi_1 > 1$) middle regime.

7.1 Suggested Future Work

- The critical values for the HBPZA test were not derived using sufficient simulations due to how slow the test is; only 300 were used. This needs to be redone with 5000 simulations.

- Critical values were derived for the lagged-series and HBPZA URTs using simulation based arguments. It would be helpful to mathematically attempt to derive their asymptotic behavior under certain assumptions.

- The new lagged-series and HBPZA URTs can be further enhanced to deal better with short time series, as well as with ARMA time series.

- Further experiments combining unit root tests with deep learning neural network architectures that allow for structural breaks in linear trend under the null hypothesis of a unit root such as the HBPZA test. These tests typically are very computationally intensive, so this experiment would take considerable computational effort and time. But once the DLNN is trained it is very fast to run.

- We would like to investigate how to fit a 3-regime SETAR model that allows for structural breaks in intercept and trend for fixed periods of time, along with the breaks inherent in transitioning between regimes.

- The procedure of combining multiple statistical tests using machine learning techniques which was used in Chapter 4 could be could be employed for other hypothesis testing problems where there is no single test that outperforms the others. One example would be testing the equality of two population variances.

Bibliography

Andrews, Donald and Zivot, Eric. Further evidence on the great crash, the oil price shock, and the unit-root hypothesis. *Journal of Business and Economic Statistics*, 10:251–270, 1992.

Baba, Naohiko and Packer, Frank. Interpreting deviations from covered interest parity during the financial market turmoil of 2007-08. Technical Report 267, BIS working paper, 2008.

Baba, Naohiko and Packer, Frank. From turmoil to crisis: dislocation in the fx swap market before and after the failure of lehman brothers. Technical Report 285, BIS working paper, 2009.

Bai, J. Least squares estimation of a shift in linear processes. *Journal of Time Series Analysis*, 15:453–472, 1994.

Bai, J. and Perron, P. Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, 18:1–22, 2003.

Bai, J. and Perron, P. Estimating and testing linear models with multiple structural changes. *Econometrica*, 66:47–78, 1998.

Balke, Nathan S. and Fomby, Thomas B. Threshold cointegration. *International Economic Review*, 38(3):627–45, August 1997.

Bec,Frederique and Ben Salem,Melika and Carrasco,Marine. Detecting Mean Reversion in Real Exchange Rates from a Multiple Regime STAR Model. RCER Working Papers 509, University of Rochester - Center for Economic Research (RCER), September 2004. URL `https://ideas.repec.org/p/roc/rocher/509.html`.

Bilson, John F. O. . The 'speculative efficiency' hypothesis. *The Journal of Business*, 54: 435–451, 1981.

Brohan P. and Kennedy JJ. and Harris I. and Tett SFB. and Jones PD. Uncertainty estimates in regional and global observed temperature changes: a new dataset from 1850. *Journal of Geophysical Research*, 111(D12):D12106.1–D12106.21., 2006.

Chan, Ngai Hang. *Time Series Applications to Finance with R and S-Plus*. John Wiley & Sons, Inc., Hoboken, New Jersey, 2010.

Chandra, Suresh K. and Janhavi, J.V. Unit root tests for time series in the presence of an explosive root. Technical Report 27, MADRAS SCHOOL OF ECONOMICS Working Paper., 2008.

Choi, In. *Almost All About Unit Roots*. Cambridge University Press, New York, New York, USA, 2010.

Coffey,N. and Hrung, W.B. and Sarkar, A. Capital constraints, counterparty risk, and deviations from covered interest rate parity. (393), 2009.

Coggin, T. Daniel. Using econometric methods to test for trends in the hadcrut3 global and hemispheric data. *International journal of climatology*, 32(2):315–320, 2012.

Di Narzo, Antonio and Di Narzo, Fabio and Aznarte,Jose Luis and Stigler,Matthieu. *tsDyn: Time series analysis based on dynamical systems theory*, 2009. URL `http://stat.ethz.ch/CRAN/web/packages/tsDyn/vignettes/tsDyn.pdf`. R package version 0.7.

Dickey, D. A. and Fuller, W.A. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, 49(4):1057–1072, 1981.

Engle, R. F. and Granger,C. W. J. Co-integration and error-correction: Representation, estimation, and testing. *Econometrica*, 55(2):251–276, 1987.

Fletcher,Donna J. and Sultan, Jahangir. Cross currency swap rates and deviations from interest rate parity. *Journal of Financial Engineering*, 6:47–69, 1997.

Frenkel, Jacob A. and Levich, Richard M. Covered interest arbitrage: unexploited profits. *Journal of Political Economy*, 83:325–338, 1975.

Frenkel, Jacob A. and Levich, Richard M. Transnational costs and interest arbitrage: tranquil versus turbulent periods. *Journal of Political Economy*, 85:1209–1226, 1977.

Genberg, Hans and Hui, C. H. and Wong, Alfred and Chung, T. K. The link between fx swaps and currency strength during the credit crisis of 2007-2008. 2009.

Giles, David E. and Godwin, Ryan T. . Testing for multivariate cointegration in the presence of structural breaks: p-values and critical values. Technical Report EWP1110, Department of Economics, University of Victoria, Econometrics Working Paper, 2011.

Glynn, J. and Perera, N. and Verma, R. Unit root tests and structural breaks: A survey with applications. *Revista De Métodos Cuantitativos Para La Economía Y La Empresa*, 55(3):63–79, 2007.

Granger, C. W. J. Current perspectives on long memory processes, 2000. ISSN 1018-161X. Zsfassung in chines. Sprache.

H2O.ai Team. H2o github repository for deep learning documentation. https://github.com/h2oai/h2o/tree/master/docs/deeplearning/DeepLearningRVignetteDemo, 2016.

Hastie, Trevor and Tibshirani, Robert and Friedman, Jerome. *The Elements of Statistical Learning*. Springer Series in Statistics. Springer New York Inc., New York, NY, USA, 2001.

Hill, R. Carter and Griffiths, William E. and Lim, Guay C. *Principles of Econometrics*. John Wiley and Sons, Hoboken, NJ, 4th edition, 2011.

Hurst, H. E. . Long term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116:770–779, 1951.

Ivanov, Martin A. and Evtimov, Stilian N. 1963: The break point of the northern hemisphere temperature trend during the twentieth century. *INTERNATIONAL JOURNAL OF CLIMATOLOGY*, 30(11):61–70, 2010.

Johansen, S. . Statistical analysis of cointegration vectors, journal of economic dynamics and control. *Journal of Economic Dynamics and Control*, 12:231–254, 1988.

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

Kejriwal, Mohitosh and Perron, Pierre and Zhou, Jing. Wald tests for detecting multiple structural changes in persistence. *Econometric Theory*, 29(02):289–323, April 2013. URL `https://ideas.repec.org/a/cup/etheor/v29y2013i02p289-323_00.html`.

Jae-Young Kim. Inference on segmented cointegration. *Econometric Theory*, null:620–639, 8 2003. ISSN 1469-4360. doi: 10.1017/S0266466603194078. URL `http://journals.cambridge.org/article_S0266466603194078`.

Kipinski, L. and Konig, R and Sieluzycki, C. and Kordecki, W. Application of modern tests for stationarity to single-trial meg data. *Biological Cybernetics*, 105(3):183–195, 2011.

Koruek, Mehmet and Ozkaya, Ata. A new method to estimate short-run and long-run interaction mechanisms in interictal state. *Digital Signal Processing*, 20(2):347–358, 2010.

Kwiatkowski, D. and Phillips, P. C. B. and Schmidt, P. and Shin, Y. Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54 (1–3):159–178, 1992.

Lee, J. and Strazicich, M.C. Break point estimation and spurious rejections with endogenous unit root tests. *Review of Economics and Statistics*, 63:535–558, 2001.

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

Levine, Michael. Time series and applications. Statistics 520 Class Notes, 2012.

Lutkepohl, H. and Saikkonnen, P. and Trenkler, C. Testing for the cointegrating rank of a var process with level shift at unknown time. *Econometrica*, 72(2):647–662, 2004.

MacKinnon, J.G. Numerical distribution functions for unit root and cointegration tests. *Journal of Applied Econometrics*, 11:601–618, 1996.

Maddala, G.S. and Kim, In-Moo. *Unit Roots, Cointegration and Structural Change*. Cambridge University Press, New York, New York, USA, 1998.

Mancini Griffoli, Tommaso and Ranaldo, Angelo. Limits to arbitrage during the crisis: funding liquidity constraints and covered interest parity. Technical report, Swiss National Bank Working Papers, 2010.

Masuda, Junya and Ohtani, Kazuhiro. Exact distribution and critical values of a unit root test when error terms are serially correlated. *Applied Economics Letters*, 15(5): 359 – 362, 2008. ISSN 13504851. URL `http://mutex.gmu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=31411663&site=ehost-live&scope=site`.

Mazzi, Biagio. *Treasury finance and development banking. A guide to credit, debt and risk*. John Wiley & Sons Inc., Hoboken, New Jersey, USA, 2013.

McAndrews, James J. . Segmentation in the us dollar money markets during the financial crisis. Technical report, FRBNY working paper, 2009.

Meyer,David and Dimitriadou,Evgenia and Hornik,Kurt ,Weingessel,Andreas and Leisch,Friedrich. *e1071: Misc Functions of the Department of Statistics- Probability Theory Group (Formerly: E1071)- TU Wien*, 2015. URL `https://CRAN.R-project.org/package=e1071`. R package version 1.6-7.

Mielke, HW. Fire as an old and new agent of change. In *Patterns of Life*, pages 241–261. Springer, 1989.

Muller, Ulrich K. and Elliott, Graham. Tests for unit roots and the initial condition. *Econometrica*, 71(4):1269–1286, 2003. ISSN 1468-0262. doi: 10.1111/1468-0262.00447. URL `http://dx.doi.org/10.1111/1468-0262.00447`.

NIST. Nist/sematech e-handbook of statistical methods. *http://www.itl.nist.gov/div898/handbook/*, 2005.

Patterson, Kerry. *A Primer for Unit Root Testing*. Palgrave Macmillan, Basingstoke, Hampshire, England, 2010.

Perron, Pierre. The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis. *Econometrica*, 57(6):1361–1401, November 1989. URL `https://ideas.repec.org/a/ecm/emetrp/v57y1989i6p1361-1401.html`.

Perron, Pierre. Testing for a unit root in a time series with a changing mean. *Journal of Business and Economic Statistics*, 8:153–162, 1990.

Pfaff, B. *Analysis of Integrated and Cointegrated Time Series with R*. Springer, New York, 2 edition, 2008.

Phillips, Peter C. B. and Perron, Pierre. Testing for a unit root in time series regression. *Biometrika*, 75:335–346, 1988.

Phillips,Peter C.B. and Magadalinos,Tassos . Limit Theory for Moderate Deviations from a Unit Root under Weak Dependence. Cowles Foundation Discussion Papers 1517, Cowles Foundation for Research in Economics, Yale University, June 2005. URL `https://ideas.repec.org/p/cwl/cwldpp/1517.html`.

Popper, Helen. Long-term covered interest parity: evidence from currency swaps. *Journal of International Money and Finance*, 12:439–448, 1993.

Quandt,Richard E. Tests of the hypothesis that a linear regression system obeys two separate regimes. *j-J-AM-STAT-ASSOC*, 55(290):324–330, June 1960. ISSN 0162-1459 (print), 1537-274X (electronic). URL `http://www.jstor.org/stable/2281745`.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2015. URL `https://www.R-project.org/`.

Rayner NA. and Brohan P. and Parker DE. and Folland CK. and Kennedy JJ. and Vanicek M. and Ansell TJ. and Tett SFB. Improved analyses of changes and uncertainties in sea surface temperature measured in situ since the mid-nineteenth century: The hadsst2 dataset. *Journal of Climate*, 19(3):446–469, 2006.

Romilly, Peter. Time series modelling of global mean temperature for managerial decision-making. *Journal of environmental management*, 76(1):61–70, 2005.

Said, S. E. and Dickey, D. A. Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, 71(3):599–607, 1984.

Seo, B. Bootstrap testing for the null of no cointegration in a threshold vector error correction model. *Journal of Econometrics*, 127(1):129–150, 2006.

Shumway,R.H. and Stoffer, D.S. *Time Series Analysis and Its Applications*. Springer, New York, 2000.

Stigler,Matthieu. *Threshold cointegration: overview and implementation in R*, 2010. URL `http://stat.ethz.ch/CRAN/web/packages/tsDyn/vignettes/ThCointOverview.pdf`.

Stock, James H. and Watson, Mark W. *Introduction to Econometrics, Third Edition*. Adisson-Wesley, Boston, MA, 2011.

Takezawa, Nobuya. Currency swaps and long-term covered interest parity. *Economics Letters*, 49:181–185, 1995.

Tsay, Ruey S. *Multivariate Time Series Analysis*. John Wiley & Sons, Inc., Hoboken, New Jersey, 2014.

Warren, Stephen G. Can human populations be stabilized? *Earth's Future*, 3(2):82–94, 2015.

Whyte, JM and Metcalfe, AV. A proposed mildly explosive/self-exciting threshold autoregressive model applied to climatic time series. In *19th International Congress on Modelling and Simulation (2011, Perth, Australia) MODSIM2011*, 2011.

Wuertz, Diethelm and many others. *fUnitRoots: Trends and Unit Roots*, 2013. URL `https://CRAN.R-project.org/package=fUnitRoots`. R package version 3010.78.

Zeileis,Achim and Kleiber,Christian and Kramer,Walter and Hornik,Kurt. Testing and dating of structural changes in practice. *Computational Statistics & Data Analysis*, 44: 109–123, 2003.

Zeileis,Achim and Leisch,Friedrich and Hornik,Kurt and Kleiber,Christian . strucchange: An r package for testing for structural change in linear regression models. *Journal of Statistical Software*, 7(2):1–38, 2002. URL `http://www.jstatsoft.org/v07/i02/`.

Zhang, Ting and Wu, Wei Biao. Time-varying nonlinear regression models: Nonparametric estimation and model selection. *Ann. Statist.*, 43(2):741–768, 04 2015. doi: 10.1214/14-AOS1299. URL `http://dx.doi.org/10.1214/14-AOS1299`.

Zivot, Eric. Unit root tests. Econ584 Class Notes, 2005.

Curriculum Vitae

Ed Herranz received his bachelor of science in mathematics with computer science from MIT in 1990 and his master of science in 1992 from the MIT Media Lab where he was a research assistant in the Advanced Human Interface Group working on merging speech recognition, hand gesture tracking and eye tracking funded by DARPA. He also received a master in financial engineering from the Universidad de Alcalá in 2009 in Madrid, Spain. Ed has worked as an analyst in Wall Street, including Salomon Brothers, as well as a programmer and financial engineer for financial software firms, such as Misys and Reval. He has been working at the International Bank for Reconstruction and Development as a senior financial officer in the Structured Finance Group of the Finance and Accounting Unit for the last 5 years. Prior to that Ed was a modeling application manager at Freddie Mac for 2 years in the Market Research and Models Group.