

2002

Testing for nonlinearities in time series with an application to exchange rates

Yan Zhang
Iowa State University

Follow this and additional works at: <https://lib.dr.iastate.edu/rtd>

 Part of the [Economics Commons](#)

Recommended Citation

Zhang, Yan, "Testing for nonlinearities in time series with an application to exchange rates " (2002). *Retrospective Theses and Dissertations*. 964.
<https://lib.dr.iastate.edu/rtd/964>

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

**ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600**

UMI[®]

Testing for nonlinearities in time series with an application to exchange rates

by

Yan Zhang

**A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY**

Major: Economics

Program of Study Committee:

Barry Falk, Major Professor

Helle Bunzel

Harvey Lapan

Peter Orazem

Yuhong Yang

Iowa State University

Ames, Iowa

2002

Copyright © Yan Zhang, 2002. All rights reserved.

UMI Number: 3051506

UMI[®]

UMI Microform 3051506

Copyright 2002 by ProQuest Information and Learning Company.

**All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.**

ProQuest Information and Learning Company

300 North Zeeb Road

P.O. Box 1346

Ann Arbor, MI 48106-1346

**Graduate College
Iowa State University**

**This is to certify that the doctoral dissertation of
Yan Zhang
has met the dissertation requirements of Iowa State University**

Signature was redacted for privacy.

Major Professor

Signature was redacted for privacy.

For the Major Program

TABLE OF CONTENTS

LIST OF FIGURES	v
LIST OF TABLES	vi
CHAPTER 1. INTRODUCTION	1
CHAPTER 2. LITERATURE REVIEW	4
Why is Nonlinear Time Series Model?	4
What is Nonlinear Time Series Model?	8
The Bilinear (BL) Model	11
The Exponential Autoregressive (EAR) Model	14
The Threshold Autoregressive (TAR) Model	16
The Smooth Transition Autoregressive (STAR) Model	21
The Autoregressive Conditional Heteroskedastic (ARCH) Model	27
CHAPTER 3. TESTING NONLINEARITIES IN TIME SERIES	30
Introduction	30
Lagrange Multiplier Tests	31
The BDS Test	34
Bootstrap Method and Monte Carlo Simulation	35
Models and Tests	37
The BL Model and BLT Statistic	38
The EAR Model and EART Statistic	38
The TAR Model	39
The STAR Model and START Statistic	39
The ARCH Model and ARCHT Statistic	41
Power Comparison of LM Test using Monte Carlo Method	42
Simulating BL Models	43

Simulating EAR Models	51
Simulating TAR Models	59
Simulating STAR Models	59
Simulating ARCH Models	62
Size and Power Comparison using Bootstrap Method	63
Discussion	64
CHAPTER 4. EMPIRICAL STUDY: AN APPLICATION TO EXCHANGE RATES	76
Introduction	76
Data Analysis	79
Linear Estimation	80
Nonlinear Exploration	83
Nonlinear Estimation	90
Forecasting	93
BIBLIOGRAPHY	102
ACKNOWLEDGMENTS	113

LIST OF FIGURES

Figure 1.1 General linearity test (left) and specific linearity test (right)	2
Figure 2.1 Comparison of transition function g for ESTAR and LSTAR models	23
Figure 3.1 Empirical power curves of BLT(1,1) against BL(1,1,1)	46
Figure 3.2 Empirical power curves of BLT(2,2) against BL(1,1,1)	47
Figure 3.3 Empirical power curves of EART(1) against BL(1,1,1)	48
Figure 3.4 Empirical power curves of START(1) against BL(1,1,1)	49
Figure 3.5 Empirical power curves of ARCHT(1) against BL(1,1,1)	50
Figure 3.6 Empirical power curves of BLT(1,1) against EAR(1)	54
Figure 3.7 Empirical power curves of BLT(2,2) against EAR(1)	55
Figure 3.8 Empirical power curves of EART(1) against EAR(1)	56
Figure 3.9 Empirical power curves of START(1) against EAR(1)	57
Figure 3.10 Empirical power curves of ARCHT(1) against EAR(1)	58
Figure 4.1 Monthly average exchange rates (1973:1-2001:3)	80

LIST OF TABLES

Table 3.1 Empirical powers of LM tests against BL(1,1,1) using ACV	43
Table 3.2 Empirical power curves of LM tests against BL(1,1,1)	45
Table 3.3 Empirical powers of LM tests against EAR(1) using ACV	52
Table 3.4 Empirical power curves of LM tests against EAR(1)	53
Table 3.5 Empirical powers of LM tests against TAR(2;1,1) using ACV	60
Table 3.6 Empirical powers of LM tests against ESTAR(1) using ACV	61
Table 3.7 Empirical powers of LM tests against LSTAR(1) using ACV	62
Table 3.8 Empirical powers of LM tests against ARCH(1,1) using ACV	62
Table 3.9 Empirical sizes of LM tests against AR(1) using ACV and BCV	65
Table 3.10 Empirical powers of LM tests against BL(1,1,1) using ACV and BCV	66
Table 3.11 Empirical powers of LM tests against EAR(1) using ACV and BCV	67
Table 3.12 Empirical powers of LM tests against TAR(2;1,1) using ACV and BCV	68
Table 3.13 Empirical powers of LM tests against ESTAR(1) using ACV and BCV	69
Table 3.14 Empirical powers of LM tests against LSTAR(1) using ACV and BCV	70
Table 3.15 Empirical powers of LM tests against ARCH(1,1) using ACV and BCV	71
Table 3.16 Rank of power performance of LM tests	72
Table 4.1 Unit root test	81
Table 4.2 Summary statistics of transformed foreign exchange rates (1973:2-2001:3)	81
Table 4.3 Linear estimation	82
Table 4.4 Summary statistics of residuals from linear estimations (1974:2-2001:3)	83
Table 4.5 BDS statistics of residuals from linear estimations	84
Table 4.6 Small sample distribution of BDS statistic	85
Table 4.7 ESTART statistics	87
Table 4.8 LSTART statistics	88
Table 4.9 Nonlinear estimation	92

Table 4.10 MSPEs for in-sample forecasts	95
Table 4.11 MSPEs for out-of-sample forecasts	96
Table 4.12 D-M statistics for in-sample forecasts with RW as the base model	98
Table 4.13 D-M statistics for out-of-sample forecasts with RW as the base model	99
Table 4.14 D-M statistics for in-sample forecasts with RWD as the base model	100
Table 4.15 D-M statistics for out-of-sample forecasts with RWD as the base model	101

CHAPTER 1

INTRODUCTION

Looking through the enormous literature about nonlinear economic time series, we find that there are basically two types of linearity tests: the general linearity tests and the specific linearity tests. Although both tests have the linear null, the general linearity tests do not specify the form of the alternative model, while the specific linearity tests do. The generality of the general linearity tests has made them a popular tool for preliminary investigation and diagnostic checking in time series analysis. However, the general linearity tests do not indicate what kind of nonlinear parametric form the series should follow if they reject the null. On the contrary, the specific linearity tests reject the null in favor of the alternative model, which is well specified. But this is built upon the strong assumption that the series is generated by the alternative process if not by the linear process. And problems arise when the true data generating process (DGP) is a nonlinear series other than the nonlinearity specified by the alternative. An illustration of the difference between general linearity test and specific linearity test is provided in Figure 1.1. If the specific linearity test has low power against other nonlinearities, it could incorrectly place the true DGP in the half sphere of the linear series, to which the true DGP does not belong. Some researchers try to justify the assumption of specific linearity test by comparing the alternative model with other nonlinear models in estimation and forecasting, but success is not always achieved.

This paper is inspired by the work of Luukkonen et al. [77] and Saikkonen and Luukkonen [106]. Given the idea that, as Luukkonen, et al. [77] say, “a test designed with a specific non-linear alternative in mind may also have power against other non-linear models” and “the LM tests may then be regarded as linearity tests against incorrect non-linear models”, Luukkonen et al. [77] conduct a power properties study of six tests (the two LM tests with bilinear alternatives, the LM test with exponential autoregressive

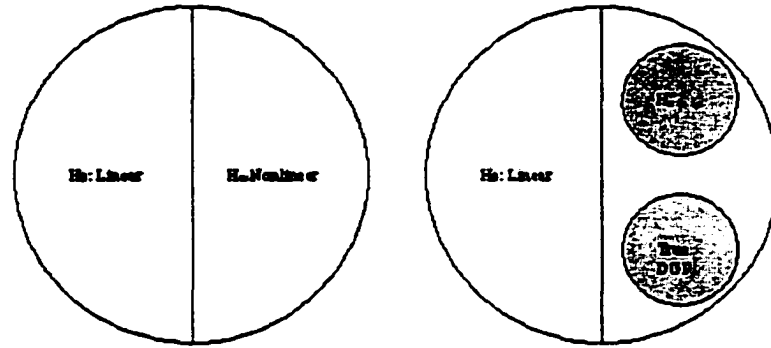


Figure 1.1: General linearity test (left) and specific linearity test (right)

alternative, the McLeod & Li's test, the Keenan's test, and the Tsay's test) against four nonlinear models (the ARCH, BL, EAR and SETAR models) using both asymptotic theory and Monte Carlo simulation. They come to the conclusion that "the power of LM tests against 'incorrect' alternatives varies widely depending on the parameters of the data generating process" and none of the discussed tests can be used as a tool for determining the type of non-linearity in univariate time series models. Saikkonen and Luukkonen [106] perform a simulation experiment to investigate the size and power properties of five tests (the two LM tests with bilinear alternatives, the LM test with exponential autoregressive alternative, the McLeod & Li's test, and the Keenan's test) for series simulated from AR, BL, and EAR models. They find that none of the tests can serve as an overall test for the two types of nonlinearities considered.

Despite the results of Luukkonen et al. [77] and Saikkonen and Luukkonen [106], it is still possible that there exists a test which has generally good powers against nonlinearities in time series, especially after quite a long time of development in time series analysis. By utilizing the power of a specific linearity test against "incorrect" nonlinear models, the paper considers applying the linearity test constructed specifically to general linearity testing. As Luukkonen et al. [77] and Saikkonen and Luukkonen [106] do, the LM test is discussed as the specific linearity test. To evaluate its performance as a general linearity test, a popular general linearity test, the BDS test, is introduced for comparison. Both power and size properties of the tests are investigated using Monte Carlo simulation and Bootstrap methods. The bootstrap method is introduced to see the effect of factors such as sample size, autoregressive coefficient,

etc. on the power and size of the tests. The data generating process are the nonlinear models introduced in Chapter Two. Though the powers of the LM tests vary according to the data generating process, as Luukkonen et al. [77] and Saikkonen and Luukkonen [106] find, we are able to find that the START, the LM test with the STAR model as the alternative, has generally good powers for all the nonlinear models discussed. Comparison of the START with the BDS test confirms that.

The rest of the paper is organized as follows. From the perspectives of specification, testing, estimation and forecasting, Chapter Two provides a literature review about the six nonlinear models: the BL, EAR, TAR, ESTAR, LSTAR and ARCH models. Chapter Three "Testing Nonlinearities in Time Series" conducts a size and power properties study of five LM tests (the two BLTs, the EART, the START, the ARCHT) and the BDS test against time series simulated from the AR and the five nonlinear models introduced in Chapter Two. An empirical study of four major countries' foreign exchange rate series is provided in Chapter Four. Linearity tests in the form of START and BDS are applied to the data before they are estimated as nonlinear ESTAR and BL models. Meese and Rogoff's myth that nonlinear models cannot do better than the random walk model is also discussed for this specific data.

CHAPTER 2

LITERATURE REVIEW

This chapter is a literature review of five nonlinear time series models with an effort to provide an overall view about specification, testing, estimation and forecasting for each model. But before that, let's first discuss two questions that are frequently brought up about the nonlinear time series model.

2.1 Why is Nonlinear Time Series Model?

Although the parametric time series analysis starts with the linear time-series modeling, in recent years, more and more nonlinear time series models have been proposed and the linear models no longer dominate the time series model analysis.

The birth of the nonlinear time series models can be attributed to the research of frequency-domain phenomena in the natural sciences. In the frequency-domain, issues such as limit cycles, jump resonance, amplitude-frequency dependency, have been considered. These problems are beyond the ability of the linear time series models because:

1. With the linearity assumption, the stationary solution to time series converges to a constant point as time goes to infinity. This constant point is called the "limit point". So it is hard to use linear time series models to describe limit cycles.
2. The joint distribution of linear models is symmetric, therefore they may not be suitable for strongly asymmetric data.
3. Linear models are not ideally suited for data exhibiting sudden bursts of very large amplitude at irregular time epochs.

In the economics field, similar nonlinearities have been found in many important economic variables, such as unemployment rate, production variables like GNP, GDP, and industrial production, consumption, price index, nominal and real exchange rates, stock prices, etc.

The evidence of nonlinearities is particularly compelling for the unemployment variable. By constructing a test under the framework of a second-order Markov process, Neftci [86] finds that the postwar quarterly US unemployment variables are characterized by sudden jumps and slow drops. Rothman [104] applies a first-order Markov procedure to the US unemployment variables and finds evidence of asymmetry in aggregate unemployment rates. Moreover, he identifies the cyclical behavior of the unemployment rate in the manufacturing sector as the primary source of asymmetry in the aggregate unemployment rate. With the introduction of the BDS test, Brock and Sayers [19] are able to detect the nonlinearities in US unemployment rate. Burgess [21] presents evidence of nonlinear dynamics in the UK employment variable by using a multivariate, structural model of employment. Besides the steepness of the business cycle that is shown by the unemployment variables, Sichel [110] constructs a test and shows that the unemployment rates also demonstrate the deepness character of the business cycle. Peel and Speight [94] model the unemployment rate data of Germany, Japan, UK and US by the ARCH, SETAR and TAR-X models and they find that there are nonlinear structures in the unemployment rates of all the countries investigated except that of Japan. Rothman [105] examines the nonlinearities in the US unemployment rates from the perspective of forecasting. The out-of-sample forecasting performances of six nonlinear models are compared with that of the linear model. He finds that the nonlinear models improve on the linear model for out-of-sample forecasting once the unemployment rates are transformed to stationarity.

Nonlinearity is also found in output variable such as GNP, GDP, and industrial production. Applying the Markov regime switching autoregressive model to the postwar US real GNP, Hamilton [58] concludes that there is a recurrent periodic shift from a positive growth rate to a negative growth rate. Unlike the Markov switching model, whose movements are regulated by an unobserved stochastic process, the movements of the SETAR model depend on the past realization of the process. The latter has also been used widely in the discussion of output variables, for example, Tiao and Tsay [118], Potter [102], Clements and Krolzig [24], Clements and Smith [25]. The model proposed by Beaudry and Koop [9] can be regarded as a modification of the TAR model, so is the floor and ceiling model by Pesaran and Potter [96]. Both models confirm the existence of asymmetries in the responses of output to positive and

negative shocks. French and Sichel [49] model the conditional mean and variance of US real GNP by an asymmetric EGARCH model. They find strong evidence of asymmetry and persistence in the variance of real GNP and the conditional variance is highest around business cycle troughs. Weiss [127] uses a simple ARCH model and found ARCH structures in US industrial production. By applying the smooth transition autoregressive model, Teräsvirta and Anderson [116] find evidence of asymmetric dynamics of the industrial production for most countries they discuss. Peel and Speight [95] test for the presence of output mean and variance nonlinearities in international industrial production and US and UK sectoral production using ARMA-GQARCH, BL and joint BL-GQARCH models and find evidence in favor of nonlinearities.

Nonlinear dynamics are not rare in financial series such as exchange rates, stock prices, and interest rates. Nelson [87] proposes the nonlinear EGARCH model to estimate the risk premium on the CRSP (Center for Research in Security Prices) value-weighted market index from 1962 to 1987. Cao and Tsay [23] explore the use of TAR model in describing monthly stock volatility series. It is suggested that the stock volatility exhibits significant lower-order serial correlation when the volatility is large, indicating certain volatility clustering in stock returns. Engle and Ng [46] measure and test the asymmetric impact of news on stock return volatility by various ARCH models. Hsieh [65] finds that there is substantial nonlinearity in a multiplicative rather than additive form in daily changes of five major foreign exchange rates. Byers and Peel [22] discover that spot exchange rates exhibit nonlinearity in either of, or both, mean and variance by using a bilinear quadratic ARCH model. Engel [42] applies the Markov switching model to 18 quarterly exchange rates series and finds that the Markov model beats the random walk model in predicting the direction of exchange rate changes. Kräger and Kugler [71] show that moderate and large exchange rate changes have different behavior by estimating the TAR model to five weekly dollar exchange rate series. Michael et al. [83], Taylor and Peel [113], Baum et al. [8] study the nonlinear adjustment to purchasing power parity by the ESTAR model. Balke and Fomby [6] find that various short-term interest rates exhibit threshold cointegration. The unit-root tests by Enders and Granger [41] show that it is an asymmetric process in the form of TAR and especially M-TAR for the interest

rates to move toward the long-run equilibrium.

Possible economic explanations for the nonlinear behaviors of those economic variables are as follows.

First, the business cycle asymmetry hypothesis. The idea that the business cycle is asymmetric can be traced back at least sixty years to the work of Mitchell [84] and Keynes [69]. The idea of the so-called Mitchell-Keynes business cycle hypothesis can be expressed clearly by Keynes' own words: "the substitution of a downward for an upward tendency often takes place suddenly and violently, whereas there is, as a rule, no such sharp turning point when an upward is substituted for a downward tendency". If this theory holds, then the indicators of the business cycle, such as GNP, GDP, industrial production, inventory, price index, unemployment, consumption, investment, etc., should be characterized by the same asymmetry associated closely with the business cycle.

While the business cycle asymmetry hypothesis can explain the nonlinear behavior of some macro variables, the nonlinearities in other variables, such as the exchange rates and the stock prices can be approached microeconomically.

Heterogeneity of participants in the financial market is often cited as a major source of nonlinear dynamics of financial series. De Grauwe et al. [29] prove that the interaction between the fundamentalists and the chartists can generate chaotic exchange rate dynamics for a wide range of parameter values. Brock and Hommes [17] propose an asset pricing model, under which market participants have heterogeneous expectations and move between different beliefs of future prices according to a "performance" indicator. This model is able to generate various nonlinear dynamics for exchange rates. Other sources of heterogeneity come from the traders' objectives (Peters [97], Guillaume et al. [55]), such as investment horizons, geographical location, and various types of risk profiles and institutional constraints.

There is an emerging theoretical literature on the persistence of the deviation of the real exchange rates from the purchasing power parity in the presence of market frictions. Dumas [36], Sercu et al. [108] develop equilibrium models of real exchange rate determination which take into account transaction costs and suggest that the exchange rate deviations from equilibrium may be controlled by nonlinear adjustment such that the speed of reversion towards equilibrium increases with the size of the deviation

from equilibrium when there are nonlinear factors governing the costs of arbitrage. Flood et al. [47], Rose and Svensson [103], propose the target zone modeling of managed floating of exchange rates. The Imperfectness of exchange market has also been used to explain the nonlinearities in nominal exchange rates. Hsieh [66] considers standard rational expectations monetary model of exchange rate determination with stochastic intervention rules.

2.2 What is Nonlinear Time Series Model?

What is a nonlinear time series model? To put it in an easy way, it can be said that any time series model that is not linear, or not of ARMA type, is a nonlinear time series model. While mathematically, it is hard to incorporate all the nonlinear time series models in one simple expression. Here let's first look at a seemingly unrelated topic, the Wold's theorem [131], which can be stated as the following.

If $\{x_t\}$ is a zero mean non-deterministic second order stationary process, it can be expressed as

$$x_t = \sum_{i=1}^{\infty} a_i v_{t-i} + v_t, \quad (2.1)$$

where $\sum_{i=1}^{\infty} a_i^2 < \infty$, $\{v_t\}$ is a process uncorrelated with $\{x_t\}$, and the sequences $\{a_t\}$ and $\{v_t\}$ are uniquely determined with v_t being the limit of linear combinations of x_s , $s \leq t$.

Basically, the Wold's theorem, which is also called the Wold's decomposition, states that a second order stationary process has a one-sided infinite order moving average representation in terms of an uncorrelated process $\{v_t\}$, but it does not mean that the process necessarily conforms to a linear model.

The basic objective in time series model building can be described as follows. Given a zero mean stationary time series $\{y_t\}$, the goal is to seek a function $h(\cdot)$ such that

$$h(y_t, y_{t-1}, \dots) = \varepsilon_t, \quad (2.2)$$

where $\{\varepsilon_t\}$ is an independently identically distributed (IID) process, and to be consistent, the assumption for $\{\varepsilon_t\}$ will remain the same throughout the paper unless stated otherwise. Note that sometimes¹ $\{\varepsilon_t\}$ is assumed to be normally independently identically distributed (NIID), which is a more restrictive

¹For example, while testing series by LM test.

condition than the IID assumption. The class of linear models is defined by restricting $h(\cdot)$ to be a linear function of y_t, y_{t-1}, \dots , like

$$H(B)y_t = \varepsilon_t, \quad (2.3)$$

with $H(B) = \sum_{i=0}^{\infty} b_i B^i$, B being called the backward shift operator. The equation (2.3) can be expanded as

$$y_t + \sum_{i=1}^{\infty} b_i y_{t-i} = \varepsilon_t. \quad (2.4)$$

If $H(z) \neq 0, |z| < 1$, (2.3) can be transformed as

$$y_t = H^{-1}(B)\varepsilon_t = \Gamma(B)\varepsilon_t \quad (2.5)$$

with $\Gamma(B) = \sum_{i=0}^{\infty} c_i B^i$, or

$$y_t = \sum_{i=1}^{\infty} c_i \varepsilon_{t-i} + \varepsilon_t. \quad (2.6)$$

Equation (2.1) and (2.6) look very similar except that according to the Wold's theorem a stationary series can be expressed in a linear combination of uncorrelated series, while a linear series requires a linear combination of independent processes, which is a more restrictive assumption.

If $H(B)$ can be written in the form

$$H(B) = \frac{1 + \phi'_1 B + \dots + \phi'_p B^p}{1 + \theta_1 B + \dots + \theta_q B^q},$$

then the resulting model for y_t becomes

$$y_t = \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t \quad (2.7)$$

with $\phi_i = -\phi'_i, i = 1, \dots, p$. The model in the form of (2.7) is the well-known autoregressive moving average (ARMA) model with the autoregressive order p and moving average order q . The autoregressive model of order p (AR(p)),

$$y_t = \sum_{i=1}^p \phi_i y_{t-i} + \varepsilon_t, \quad (2.8)$$

is obtained from (2.7) by letting $\theta_1 = \theta_2 = \dots = \theta_q = 0$. If the autoregressive coefficients $\{\phi_i\}$ are all zeros, the ARMA(p, q) model is reduced to the moving average model MA(q),

$$y_t = \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t. \quad (2.9)$$

To define the nonlinear time series model, let's again look at the equation (2.2). Obviously, the nonlinearity could come from the fact that $h(\cdot)$ is not a linear function of y_t, y_{t-1}, \dots , but it's also possible that the deviation of ε_t from the IID assumption causes the nonlinearity of y_t . The source of the nonlinearity divides the two classes of nonlinear time series models this paper is going to discuss.

A simple class of nonlinear time series models is the nonlinear autoregressive (NLAR) models which has a general form as

$$y_t = f(\cdot) + g(\cdot) + \varepsilon_t, \quad (2.10)$$

where $f(\cdot)$ is an AR process of $\{y_t\}$, $g(\cdot)$ is a nonlinear function of y_{t-i} and ε_{t-j} , lagged terms of the dependent variable y_t and the disturbance variable ε_t . The NLAR model can be easily extended to NLARMA model by defining $f(\cdot)$ as an ARMA process. However, as we know that an ARMA process can be approximated by an AR process with long enough lags under quite mild conditions, the NLAR model instead of the NLARMA model will be discussed in this dissertation. Introduction of lagged error terms ε_{t-j} into the nonlinear $g(\cdot)$ function can be looked upon as a parsimonious way to express the nonlinearity in $g(y_t, y_{t-1}, \dots)$. So modeling a series as the NLAR model can be boiled down to the task of finding a nonlinear function $g(\cdot)$ for the pre-whitened series $z_t = y_t - f(\cdot)$ such that $g(z_{t-i}, \varepsilon_{t-j}) = \varepsilon_t$.

With a defined form of $g(\cdot)$, the NLAR model can be classified into more specific nonlinear time series models. In this paper, the bilinear (BL) model, the exponential autoregressive (EAR) model, the threshold autoregressive (TAR) model, and the smooth transition autoregressive (STAR) model will be discussed for the NLAR class.

The NLAR model describes the nonlinearity in time series by the model mean. However, there is another group of nonlinear model that captures the nonlinearity by the disturbance term. The representative is the autoregressive conditional heteroskedastic (ARCH) model.

2.3 The Bilinear (BL) Model

Similar to Wold's [131] work of presenting a stationary series as a limit of linear combination of some uncorrelated process, Volterra [126] has shown that under certain conditions the process $\{y_t\}$ can be written as

$$y_t = \sum_{i=1}^{\infty} \left[\sum_{u_1} \cdots \sum_{u_i} g_i(u_1, u_2, \dots, u_i) \prod_{j=1}^i e_{t-u_j} \right]. \quad (2.11)$$

This is known as a Volterra expansion and the kernels $g_i(u_1, u_2, \dots, u_i)$ are known as Volterra kernels.

However, it is difficult to estimate a nonlinear time series process in the form of its Volterra expansion. Mohler [85] introduces the class of bilinear models which are later discussed in detail by Granger and Anderson [51, 52], Subba Rao and Gabr [112].

The general representation of bilinear model $BL(p, m, k)$ is

$$y_t = \mu + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{i=1}^m \sum_{j=1}^k c_{ij} y_{t-i} \varepsilon_{t-j} + \varepsilon_t, \quad (2.12)$$

where $c_{mk} \neq 0$. Based on the $AR(p)$ model, the $BL(p, m, k)$ model adds the cross-product terms of y_{t-i} and ε_{t-j} to account for the nonlinearity, $i \in [1, 2, \dots, m]$ and $j \in [1, 2, \dots, k]$. Therefore, if all the bilinear coefficients c_{ij} are zeros, the $BL(p, m, k)$ model becomes an $AR(p)$. Brockett [20] shows that with suitable choice of the model parameters, the bilinear model can approximate to an arbitrary degree of accuracy any "well-behaved" Volterra series relationship over a finite time interval. And compared with the Volterra series, the bilinear model is more parsimonious.

Applications of the bilinear model include Maravall [79], Subba Rao and Gabr [112], and Terdik [117], among others.

In his study of the Spanish currency from 1974 to 1980, Maravall compares the bilinear model with the linear MA model for estimation and forecasting. His introduction of the bilinear model is based on the diagnostic checking of the estimated residuals from the MA model: although first order diagnostic checking of the residuals series reveals no conflict with the assumption that it is a white noise process, the ACFs of squared residuals series display pattern of autocorrelations for low-order and seasonal lags, which means the squared residuals series is not a white noise process let alone being independent. Therefore, the residual process is not independent and may be nonlinear since the square of an independent series

should be independent too. Estimating the currency series as the bilinear model is able to decrease the residual variance and increase the forecastability. And the ACFs of the estimated residuals from the bilinear model indicate there are no nonlinearities left. In conclusion, Maravall mentions that the benefit of the bilinear model is observed especially at periods of atypical behavior or sequences of outliers. In his empirical analysis, the improvement of the forecast errors is concentrated over the last two months, when linear forecasts are over-estimating owing to an unexpected drop in currency. Actually, the bilinear model can be regarded as an outlier filter in the sense that it becomes operative and smooths outliers when atypical behavior sets in and is mostly inoperative during a normal regime. Maravall also illustrates that the bilinear model can capture a special type of nonstationarity that cannot be captured by linear model.

Subba Rao and Gabr [112] study the unemployment variable of west Germany by looking at the monthly unemployment rate from January 1948 to May 1980. Test confirms that there are nonlinearities in the unemployment series. Estimation and one-step-ahead forecasts show that the bilinear model is better than the linear AR model.

Terdik [117] estimates two financial series in the form of bilinear model: the daily closing S&P 500 index between 04/21/1982 to 10/02/1998, and the daily closing prices of IBM stock prices for the period 01/02/1968-05/29/1998. For the S&P 500 index, the bilinear model is significantly better than an AR(1) model according to estimated variance, which decreases from 25.2099 to 1.4912.

Subba Rao and Gabr [112] also propose an algorithm to estimate a parsimonious BL($p, 0; m, k$) model

$$x_t + \sum_{i=1}^p a_i x_{t-i} = \alpha + \sum_{i=1}^m \sum_{j=1}^k b_{ij} x_{t-i} e_{t-j} + e_t.$$

The steps are:

1. Estimate a full AR(p) model, by using the AIC to select p for $0 \leq p \leq \gamma$, where $\gamma \geq m, k$. Record the mean sum of squares of the residuals as $\hat{\sigma}_e^2(1)$ and the AIC as AIC(1). Then the best subset AR model is fit based on the full model using the algorithm described by Haggan and Oyetunji [56]. The

resulting model is

$$x_t + a_{k_1} x_{t-k_1} + a_{k_2} x_{t-k_2} + \cdots a_{k_l} x_{t-k_l} = e_t,$$

where $1 \leq k_1 \leq k_2 \leq \cdots \leq k_l \leq p$, and the mean sum of squares of the residuals is $\hat{\sigma}_e^2(2)$ and the AIC is AIC(2).

2. Fit the bilinear model

$$x_t + a_{k_1} x_{t-k_1} + a_{k_2} x_{t-k_2} + \cdots a_{k_l} x_{t-k_l} = a + b_{ij} x_{t-i} e_{t-j} + e_t,$$

for all $(i, j) \in T_1$, where $T_1 = \{(i, j); i, j = 1, 2, \dots, \xi, p \leq \xi \leq \gamma\}$. Choose the combination (r_1, l_1) by the AIC and record it as $SBL(k_l, 1)$, which is

$$x_t + a_{k_1} x_{t-k_1} + a_{k_2} x_{t-k_2} + \cdots a_{k_l} x_{t-k_l} = a + b_{r_1 l_1} x_{t-r_1} e_{t-l_1} + e_t,$$

with $\hat{\sigma}_e^2(SBL(k_l, 1))$ and $AIC(SBL(k_l, 1))$.

3. Estimate $SBL(k_l, 2)$ by

$$x_t + a_{k_1} x_{t-k_1} + a_{k_2} x_{t-k_2} + \cdots a_{k_l} x_{t-k_l} = a + b_{r_1 l_1} x_{t-r_1} e_{t-l_1} + b_{r_s l_2} x_{t-r_s} e_{t-l_2} + e_t,$$

where $(r, s) \in T_2 = T_1 - (r_1, l_1)$. Again choose the $SBL(k_l, 2)$ model with (r_2, l_2) by AIC, which is in the form

$$x_t + a_{k_1} x_{t-k_1} + a_{k_2} x_{t-k_2} + \cdots a_{k_l} x_{t-k_l} = a + b_{r_1 l_1} x_{t-r_1} e_{t-l_1} + b_{r_2 l_2} x_{t-r_2} e_{t-l_2} + e_t.$$

Same estimation procedure will continue until a bilinear model $SBL(k_l, m)$ with minimum AIC is found, which can be written as

$$x_t + a_{k_1} x_{t-k_1} + a_{k_2} x_{t-k_2} + \cdots a_{k_l} x_{t-k_l} = a + \sum_{j=1}^m b_{r_j l_j} x_{t-r_j} e_{t-l_j} + e_t.$$

4. Once the order of the bilinear model is found, the model parameters can be estimated by the Newton-Raphson iteration method. Initial parameter values come from the estimates of previous step. For example, when estimate the $SBL(k_l, 1)$, it autoregressive coefficients can take the initial value equal to the estimates of the subset AR model, and the bilinear coefficient $(r_1, l_1) = (0, 0)$.

An LM test against the bilinear alternative is presented by Weiss [128], and Saikkonen and Luukkonen [106]. Chapter Three will discuss the test in detail.

2.4 The Exponential Autoregressive (EAR) Model

The literature about the EAR model is concentrated in the late 70s and early 80s. Haggan and Ozaki [57] discuss the EAR model generally and provide an estimation procedure. Ozaki and Oda [92] and Ozaki [91] propose an EAR(2) model. Lawrance and Lewis [72, 73] introduce the EMA(1) model and the EARMA(p, q) model, which extend the EAR model by allowing for nonlinear MA terms.

The general EAR($p; d$) model can be expressed as

$$y_t = \mu + \sum_{i=1}^p \phi_i y_{t-i} + [\exp(-\gamma y_{t-d}^2) - 1] \sum_{i=1}^p \theta_i y_{t-i} + \varepsilon_t, \quad (2.13)$$

where y_{t-d} is the transition variable, d is the delay parameter that varies between 1 and p , γ serves as the transition parameter that determines the speed of transition.

We can see that the EAR(p) model is actually an AR(p) model plus some nonlinear terms in the form of $[\exp(-\gamma y_{t-d}^2) - 1] y_{t-i}$, $\forall i$. The term $[\exp(-\gamma y_{t-d}^2) - 1]$ is bounded between -1 and 0 and is symmetric around 0 for $y_{t-d} \in (-\infty, +\infty)$. When $[\exp(-\gamma y_{t-d}^2) - 1] = 0$, the EAR model is reduced to an AR(p) model

$$y_t = \mu + \sum_{i=1}^p \phi_i y_{t-i} + \varepsilon_t; \quad (2.14)$$

when $[\exp(-\gamma y_{t-d}^2) - 1] = -1$, the EAR model is reduced to another AR(p) process

$$y_t = \mu + \sum_{i=1}^p (\phi_i - \theta_i) y_{t-i} + \varepsilon_t; \quad (2.15)$$

if $[\exp(-\gamma y_{t-d}^2) - 1] = h$ ($-1 < h < 0$), y_t is an AR(p) process like

$$y_t = \mu + \sum_{i=1}^p (\phi_i + h \times \theta_i) y_{t-i} + \varepsilon_t, \quad (2.16)$$

which combines the two extremes of (2.14) and (2.15).

The value of $[\exp(-\gamma y_{t-d}^2) - 1]$ depends on two terms, the transition parameter γ and the transition variable y_{t-d} . In order to have a meaningful model, the value of γ should be from the set $(0, +\infty)$. Given that y_{t-d} takes some intermediate values, when $\gamma = 0$, the EAR model becomes an AR model with the form (2.14), when $\gamma \rightarrow +\infty$, the EAR model is reduced to another AR model like (2.15), when γ takes intermediate values, we get an EAR model. The value of $[\exp(-\gamma y_{t-d}^2) - 1]$ will also change with t thus

with y_{t-d} . When $y_{t-d} = 0$, $\exp(0) = 1$, y_t is in the form (2.14); when $y_{t-d} \rightarrow \pm\infty$, $\exp(-\infty) = 0$, y_t can be expressed as (2.15); when y_{t-d} takes intermediate values, y_t has the form (2.16). Depending on different values taken by y_{t-d} , the EAR model provides a smooth transition from one regime to another regime. And the model exhibits similar behaviors when it moves from the outer regimes to the middle regime.

Saikkonen and Luukkonen [106] construct an LM test with the EAR model as the alternative. It will be introduced in Chapter Three.

In Haggan and Ozaki's paper [57], they suggest that the EAR model can incorporate both amplitude-dependent frequency and limit cycle behavior by making the coefficients of an autoregressive model amplitude-dependent so that it becomes

$$x_t = \left(\phi_1 + \pi_1 e^{-\gamma x_{t-1}^2} \right) x_{t-1} + \cdots + \left(\phi_p + \pi_p e^{-\gamma x_{t-p}^2} \right) x_{t-p} + e_t. \quad (2.17)$$

In order to capture limit cycle behavior, the EAR(p) model (2.17) should satisfy

- (i). the roots of $\lambda^p - \sum_{i=1}^p \phi_i \lambda^{p-i} = 0$ lie inside the unit circle;
- (ii). the roots of $\lambda^p - \sum_{i=1}^p (\phi_i + \pi_i) \lambda^{p-i} = 0$ do not all lie inside the unit circle;
- (iii). $(1 - \sum_{i=1}^p \phi_i) / \sum_{i=1}^p \pi_i > 1$ or $(1 - \sum_{i=1}^p \phi_i) / \sum_{i=1}^p \pi_i < 0$.

When the absolute value of x_{t-1} is large, the value $e^{-\gamma x_{t-1}^2} \sum_{i=1}^p \pi_i x_{t-i}$ tends to be trivial and the polynomials $\lambda^p - \sum_{i=1}^p \phi_i \lambda^{p-i}$ applies. So satisfying condition (i) means the system damps down to zero for large deviation of x_{t-1} from 0; similarly, condition (ii) restricts the system to explode when the absolute value of x_{t-1} is small. And condition (iii) guarantees that there is no single solution to the model.

For an EAR(p) model, it is a nonlinear optimization procedure to estimate the order p and the coefficients $\{\gamma, \phi_i, \pi_i; i = 1, \dots, p\}$ simultaneously. However, Haggan and Ozaki [57] avoid doing nonlinear optimization in early 1980s by fixing the parameter γ at one of a grid of values and estimating the order p and the corresponding coefficients ϕ_i and π_i . By doing this, the estimation becomes a linear regression of x_t on x_s and $e^{-\gamma x_{t-s}^2} x_s$, for $s < t$. The range of γ will be chosen such that $e^{-\gamma x_{t-s}^2}$ is different from zero and one for most values of x_{t-s} . The order of the exponential autoregressive model

p will be decided by AIC for each fixed γ , and then AIC is used again to find the best model over all γ values.

By applying the EAR model to the Canadian lynx data, Haggan and Ozaki [57] shows that there is obvious periodic behavior for the Canadian lynx series.

However, recently, there is a tendency that more STAR model instead of EAR model is used to model nonlinearities in time series . Perhaps it is due to the fact that the STAR model is a more general model which nests the EAR model and the single TAR model as its special cases.

2.5 The Threshold Autoregressive (TAR) Model

The TAR model was first proposed by Tong [119] and then developed by Tong and Lim [122], Tong [120, 121].

The TAR model can be regarded as a “piecewise-linear” approximation to the nonlinearity. For the $TAR(k; p, d)$ model, it has the following form

$$y_t = \mu^{(j)} + \sum_{i=1}^p \phi_i^{(j)} y_{t-i} + \varepsilon_t^{(j)}, \quad (2.18)$$

where $r_j \leq x^{(j)} < r_{j-1}$ ($j = 1, \dots, k$), k is the number of regimes separated by $(k - 1)$ nontrivial thresholds r_j ($j = 1, \dots, k - 1$), and p denotes the AR order. The quantity d is a positive integer, which is the threshold lag, or delay parameter. The variable x is the threshold variable. When the threshold variable is the past realization of the process $\{y_t\}$, the TAR model is self-exciting and called SETAR model. Unless specified, the TAR model discussed in the paper is of SETAR type.

The TAR model partitions the one-dimensional Euclidean space into k regimes and follows a linear AR model in each regime. The individual AR models are different from each other, thus is the TAR model nonlinear on the whole, and is asymmetric and adjusts abruptly and suddenly from one regime to another. According to Tong [121], the key features of TAR model include time-irreversibility, asymmetric limit cycle and jump phenomenon.

Depending on the values of k , p and d , the TAR model can have various forms. A simple form of

TAR model is the two-regime TAR(2; p, d), which can be expressed as

$$y_t = I_t \mu_1 + (1 - I_t) \mu_2 + I_t \sum_{i=1}^p \phi_{1i} y_{t-i} + (1 - I_t) \sum_{i=1}^p \phi_{2i} y_{t-i} + \varepsilon_t, \quad (2.19)$$

where $\phi_{1i} \neq \phi_{2i}$ for at least one value of i , I_t is the Heaviside indicator function such that

$$I_t = \begin{cases} 1 & \text{if } y_{t-d} \geq 0 \\ 0 & \text{if } y_{t-d} < 0 \end{cases}. \quad (2.20)$$

Here we have two regimes, in each regime there is an AR(p) process, the linear attractor is 0 that is the long-run equilibrium value of the sequence $\{y_t\}$. To be more complicated, the linear attractor can take some value other than 0, like a nonzero constant or a trend term, and the disturbance terms for two regimes can be different too.

Enders and Granger [41] propose a momentum threshold autoregressive (M-TAR) model by using Δy_{t-d} , change in the lagged dependent variable, as the transition variable instead of y_{t-d} in the indicator function. For the model M-TAR(2; p, d), the indicator function of the model now becomes

$$I_t = \begin{cases} 1 & \text{if } \Delta y_{t-d} \geq 0 \\ 0 & \text{if } \Delta y_{t-d} < 0 \end{cases}. \quad (2.21)$$

By introducing the M-TAR model, we can explain the “sharpness” or “steepness” in business cycles. Sichel [110] discusses the distinction between “sharpness” and “deepness”. Sharpness occurs when contractions are steeper than expansions and deepness occurs when troughs are more pronounced than peaks. Ordinary TAR model can capture the “deepness” by the different values of ϕ_1 and ϕ_2 . For example, in the TAR(2; 1, 1) model, if $|\phi_1| < |\phi_2|$, the negative phase of $\{y_t\}$ will be more persistent than the positive phase. While using M-TAR model we can express the “sharpness” by the values of ϕ_1 and ϕ_2 . In the example, if $|\phi_1| < |\phi_2|$, the M-TAR model will exhibit substantial decay for negative Δy_{t-1} , but little decay for positive Δy_{t-1} .

The TAR model has been widely used in the analysis of economic variables such as GNP (Tiao and Tsay [118], Potter [102], Beaudry and Koop [9], Pesaran and Potter [96]), unemployment (Peel and Speight [94]), and exchange rates (Kräger and Kugler [71], Pippenger and Goering [101]), to mention but a few.

Potter [102] fits a TAR(2;5,5) model on the seasonally adjusted quarterly real GNP of US measured at 1982 prices for the period 1947:1-1990:4. Tiao and Tsay [118] estimate a TAR(2;2,2) model and a four-regime refinement with the US GNP for the same period. Although the two estimated TAR model are not exactly the same, they both have a large negative coefficient on the second lag in the lower regime, indicating that the postwar US economy moves quickly out of recession. Tiao and Tsay [118] also go ahead to evaluate the forecasting performance of the TAR model against a linear AR(2) model by the Monte Carlo simulation method. They find that although the overall forecasting performance of the TAR model is only marginally better than that of the AR model, the TAR model is markedly superior if considering forecasts made in the lower regime. The reason behind this is that the majority of the data points fall into the upper regime, so the linear AR(2) model, which is determined by those points to a large extent, resembles high similarity to the TAR model in this regime. Therefore, the nonlinear model forecasts made in the upper regime will be close to that of the linear model. But when the forecasts origin from the lower regime, only the TAR model can detect the regime change and express the change by a different process. So it is here that the TAR model can gain relative to the linear model.

Beaudry and Koop [9] introduce a variable called current depth of a recession (CDR), which is the gap between the current level of output and the economy's historical maximum level, as the threshold variable in their TAR fitting of the quarterly real GNP of the postwar US. They find negative innovations to GNP are much less persistent than positive ones. Moreover, the effect of a recession on the output forecasting is negligible after only eight to twelve quarters while the effect of a positive shock is persistent and increasing over time. Pesaran and Potter [96] propose a floor and ceiling model for US output, which is considered as an extension to Beaudry and Koop's model.

With the transition variable being the past growth rates of industrial production, Peel and Speight [94] apply the TAR model to the unemployment rates for Germany, Japan, UK and US. The estimated TAR models yield substantial reductions in residual variance over the linear counterparts, and have a threshold near zero, which clearly indicates the asymmetric behavior of unemployment during expansion and contraction of the countries' economy.

Kräger and Kugler [71] find that the TAR model can be used to describe the behavior of five weekly dollar exchange rate series: the French franc, the Italian lira, the Japanese yen, the Swiss franc and the German mark against the US dollar. The estimated TAR models for the first four exchange rates share the common characteristics: they are TAR models of three regimes; the process of the upper and the lower regimes are similar with the AR order equal 0 and the middle regime is an AR(3) process; and the estimated standard deviations of the outer regimes are larger than that of the middle regime. Such kind of TAR models supports Kräger and Kugler's argument that authority intervenes the exchange market when the depreciations and appreciations are large but adopts a Laissez-faire attitude when the changes are small. Pippenger and Goering [101] prove that the TAR model can be used to describe the log changes of 13 countries' monthly exchange rates. Moreover, they find that the TAR model produces better forecasts than the random walk model in terms of MSPE for both in-sample and out-of-sample forecasting.

Literature emphasizes the forecasting performance of the TAR model include Clements and Smith [25, 26, 27] and Clements and Krolzig [24].

Clements and Smith [25] compare the forecasting performance of TAR model with the AR model for US GNP using five methods: a Monte Carlo simulation method, a bootstrap method, a normal forecast errors method, a dynamic estimation method, and a naive method. They conclude that the Monte Carlo simulation method is generally good and whether a TAR model can forecast better than an AR model out of sample depends on the historical epoch and the regime where the forecast is made, which confirms the results by Tiao and Tsay [118]. They also conduct a Monte Carlo study of the forecasting performance of the proposed empirical TAR models on exchange rates, GNP, GDP, and savings ratio by several quantitative and qualitative criteria (Clements and Smith [26]), and a forecast evaluation of TAR model against the random walk model for exchange rates based on traditional measures such as MSPE as well as generalized impulse response functions (Clements and Smith [27]). Both again emphasize that "state of nature" has a very important effect on the forecasting performance of TAR model.

Tsay [124] and Hansen [59, 60, 61] discuss intensively about the testing of TAR model.

Tsay [124] proposes a procedure to test a TAR model that combines the nonlinearity tests of Keenan [68], Tsay [123], and Petrucci and Davies [98]. The rationale of his test is that: within each threshold regime, the predictive residuals are white noise asymptotically and orthogonal to the regressors. However, the orthogonality between the predictive residuals and the regressors is destroyed once the observations enter another regime. So the F statistic of the regression of the predictive residuals on the regressors can be used to test for threshold nonlinearity.

Based on his test, Tsay suggest to follow the following steps to identify a TAR model:

1. Select the order of $AR(p)$ and the set S of all possible threshold lags d . Tsay prefers to use the PACF to identify the order of an AR process. For a give p , the set S of all possible threshold lags d is assured to be $\{1, \dots, p\}$. When there is seasonality, the set S can also include seasonal lags.

2. Arrange the data cases according to the threshold variable y_{t-d} . Then fit arranged autoregressions for a given p and every element d of S to get the predictive residuals. Perform the threshold nonlinearity test $\widehat{F}(p, v)$ by regressing the standardized predictive residuals on the regressors. If nonlinearity is detected, select the delay parameter d_p . The delay parameter d_p is chosen by $\max_{v \in S} \{ \widehat{F}(p, v) \}$ with respect to v .

3. For given p and d_p , locate the threshold values by using the scatter plots of the threshold variable against the predictive residuals, the standardized predictive residuals, or the t ratios of recursive estimates of an AR coefficient.

4. Refine the AR order and threshold values, if necessary, in each regime by using linear autoregression techniques. This step may rely on information criteria such as the AIC.

Tsay applies the testing and estimation procedures to the classic sunspot and the Canadian lynx data. Both are estimated as a three-regime TAR model. His testing technique is also used by Tiao and Tsay [118], Potter [102].

According to Hansen, the testing of nonlinearity in the context of TAR model falls into the class of tests with nuisance parameters that are not identified under the null hypotheses. Specifically, the variables of the alternative TAR model such as the lag of the threshold variable, the number of regimes, are

not included in the parameters of the null linear model, therefore they are nuisances in the linearity test. Due to the existence of the nuisance parameters, although the linearity test is easy to construct, there is no theoretical knowledge as for what kind of distribution the test will follow exactly or asymptotically. Hansen [59] shows that a transformation based upon a conditional probability measure yields an asymptotic distribution free of nuisance parameters and that this transformation can be easily approximated via simulation. He calculates the test statistics over a grid of values for the nuisance parameters and then takes the supremum, average or exponential transformation of the statistics. Application of the result to Potter's study of US GNP [102] suggests that the finding of a threshold effect in US GNP growth rates can be possibly explained by sampling variation. Application of Hansen's testing can be seen in Clements and Smith [27], Pesaran and Potter [96]. Hansen [60] finds that if the threshold effect (the difference in slopes between the two regimes) becomes small as the sample size increases, the asymptotic distribution of the threshold estimator is free of nuisance parameters. Similarly, the likelihood ratio statistic for testing hypotheses concerning the unknown threshold is asymptotically free of nuisance parameters. These asymptotic distributions are nonstandard, but are available in closed form, so critical values are readily available. The US unemployment rate is tested and statistically significant threshold effects are found. In his paper "Testing for Linearity" [61], Hansen discusses the problem of testing for linearity and the number of regimes for TAR model. The asymptotic and bootstrap distributions of the test statistic are calculated by simulation. Special attention is paid to the problem of conditional heteroskedasticity in the error while drawing inference on the conditional mean. The method is applied to annual sunspot means and monthly US industrial production. Both series follow a TAR(2) processes.

2.6 The Smooth Transition Autoregressive (STAR) Model

The idea of smooth transition can be traced back to Bacon and Watts [4]. The full development of the STAR model should be attributed to Teräsvirta and his coauthors (Luukkonen, Saikkonen and Teräsvirta [78], Granger and Teräsvirta [53], Teräsvirta [115], Teräsvirta and Anderson [116], Eitrheim and Teräsvirta [40]).

The STAR model can be generalized as

$$y_t = \mu_1 + \sum_{i=1}^p \phi_{1i} y_{t-i} + \left(\mu_2 + \sum_{i=1}^p \phi_{2i} y_{t-i} \right) \times g(y_{t-d} - c; \gamma) + \varepsilon_t, \quad (2.22)$$

where p is the autoregressive order, $g(y_{t-d} - c; \gamma)$ is the continuous transition function, y_{t-d} is the transition variable, d is the delay parameter, γ is the transition parameter with a positive value, c is the mid-point of the transition regime.

If $g(y_{t-d} - c; \gamma)$ is of the exponential form

$$\left\{ 1 - \exp \left[-\gamma (y_{t-d} - c)^2 \right] \right\}, \quad (2.23)$$

the model is the exponential smooth transition autoregressive ESTAR($p; d$) model. Note that when $\gamma \rightarrow 0$ or ∞ , the ESTAR model will become linear. The term $g(y_{t-d} - c)$ in the ESTAR model is bounded between zero and unity and is symmetric around zero for $(y_{t-d} - c) \in (-\infty, +\infty)$. When $y_{t-d} - c = 0$, $g(0) = 0$, y_t is an AR(p); when $(y_{t-d} - c) \rightarrow \pm\infty$, $g(\pm\infty) = 1$, $\{y_t\}$ is another AR(p); when $(y_{t-d} - c)$ takes intermediate values, $\{y_t\}$ is an AR(p) which is the combination of the two aforementioned. According to Teräsvirta and Anderson [116], the ESTAR model implies that the contraction and expansion have rather similar dynamic structures, whereas the middle ground can have different dynamics. An ESTAR model can therefore represent an economy that returns from high growth towards more “normal” growth in much the same fashion as it accelerates from low or negative growth towards the middle ground. Note that the EAR model is a special case of ESTAR with $\mu_2 = 0$ and $c = 0$.

When $g(y_{t-d} - c)$ is of the logistic form

$$\left[\{1 + \exp[-\gamma^* (y_{t-d} - c^*)]\}^{-1} - 0.5 \right], \quad (2.24)$$

the model is called the logistic smooth transition autoregressive LSTAR($p; d$) model. When $\gamma^* \rightarrow \infty$, $g(y_{t-d} - c^*) = -0.5$ if $y_{t-d} \leq c^*$, and $g(y_{t-d} - c^*) = 0.5$ if $y_{t-d} > c^*$, therefore the LSTAR model becomes a single-threshold autoregressive model. The $g(y_{t-d} - c^*)$ term in the LSTAR model is bounded between $[-0.5, 0.5]$ for $(y_{t-d} - c^*) \in (-\infty, +\infty)$. When $(y_{t-d} - c^*) \rightarrow +\infty$, $g(+\infty) = 0.5$, $\{y_t\}$ is an AR(p); when $(y_{t-d} - c^*) \rightarrow -\infty$, $g(-\infty) = -0.5$, $\{y_t\}$ is another AR(p); when $(y_{t-d} - c^*)$ takes intermediate values, $\{y_t\}$ is a combined AR(p). Different from the ESTAR model, the LSTAR model

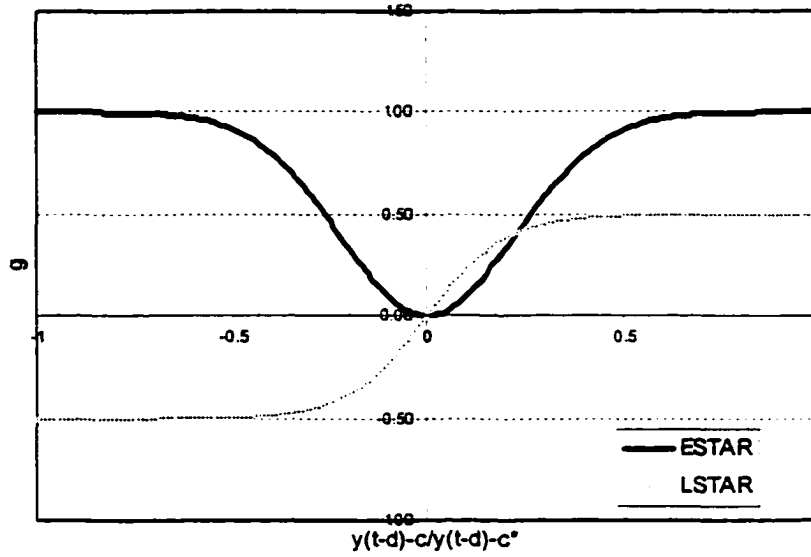


Figure 2.1: Comparison of transition function g for ESTAR and LSTAR models

isn't symmetric. The comparison of function g for the two STAR models are illustrated in Figure 2.1 for $\gamma = \gamma^* = 10$. Applied to the modeling of business cycle indicators, the LSTAR model describes a situation where the contraction and expansion phases of an economy have different dynamics (Teräsvirta and Anderson [116]).

Note that when $\gamma = 0$, the $\text{ESTAR}(p; d)$ model is reduced to an $\text{AR}(p)$ model. Similarly, under $\gamma^* = 0$, the $\text{LSTAR}(p; d)$ model is reduced to an $\text{AR}(p)$ model too.

Teräsvirta and Anderson [116] study the STAR models of the form:

$$y_t = \pi_{10} + \pi'_1 w_t + (\pi_{20} + \pi'_2 w_t) F(y_{t-d}) + u'_t, \quad (2.25)$$

where $u_t \sim \text{NIID}(0, \sigma^2)$, $\pi_j = [\pi_{j1} \ \cdots \ \pi_{jp}]'$, $j = 1, 2$, $w_t = [y_{t-1} \ \cdots \ y_{t-p}]'$ and F is a transition function, which is

$$F(y_{t-d}) = \{1 + \exp[-\gamma(y_{t-d} - c)]\}^{-1}$$

for the logistical STAR model and is

$$F(y_{t-d}) = 1 - \exp[-\gamma(y_{t-d} - c)^2]$$

for the exponential STAR model, with $\gamma > 0$. As for the specification of the STAR model, they suggest a three-step method:

1. Specification of a linear AR model.

This step is the base of the linearity test since the linear AR(p) model is nested in the STAR(p) model. So estimation of the autoregressive order p has a direct effect on the linearity test. If p is over-estimated, the linearity test tends to have lower power; if p is under-estimated, it is possible that the test will incorrectly reject the null.

2. Testing linearity for different values of the delay parameter d , and if it is rejected, determining d .

With a pre-specified value of d , a Lagrange multiplier test of linearity against the STAR alternative can be constructed with the null hypothesis

$$H_0 : \beta_{2j} = \beta_{3j} = \beta_{4j} = 0, \quad j = 1, \dots, p \quad (2.26)$$

against H_1 : “ H_0 is not valid” in the artificial regression

$$y_t = \beta_0 + \beta_1' w_t + \sum_{j=1}^p \beta_{2j} y_{t-j} y_{t-d} + \sum_{j=1}^p \beta_{3j} y_{t-j} y_{t-d}^2 + \sum_{j=1}^p \beta_{4j} y_{t-j} y_{t-d}^3 + v_t, \quad (2.27)$$

which is the third-order expansion of the original STAR model (2.25).

If linearity can be rejected for more than one value of d for $1 \leq d \leq D$, then the value of d can be chosen as

$$\hat{d} = \arg \min p(d),$$

where $p(d)$ is the p -value of the corresponding linearity test. This is based on the argument that the power of the corresponding test against the mis-specified nonlinear model can hardly be expected to be systematically higher than the power of the test for the correctly specified model.

3. Choosing between LSTAR and ESTAR models using a sequence of tests of nested hypothesis.

Based on interpreting the coefficients β_{ij} as functions of the parameters in the original model, the

hypotheses of the sequence of tests are

$$H_{01}:\beta_{4j} = 0, j = 1, \dots, p; \quad (2.28)$$

$$H_{02}:\beta_{3j} = 0 | \beta_{4j} = 0, j = 1, \dots, p; \quad (2.29)$$

$$H_{03}:\beta_{2j} = 0 | \beta_{3j} = \beta_{4j} = 0, j = 1, \dots, p. \quad (2.30)$$

Combinations of rejecting or accepting one or another sequence of hypothesis will lead to different conclusions. If (2.28) is rejected, the chosen model will be an LSTAR model; if (2.28) is accepted and (2.29) is rejected, the ESTAR model will be selected; if (2.28) and (2.29) are accepted but (2.30) is rejected, the model will be LSTAR. However, when both (2.29) and (2.30) are rejected, the conclusion is uncertain.

As for interpretation of the estimated models, Teräsvirta and Anderson look at the roots of the characteristic polynomials and the implied limiting behavior of the series y_t . The roots of the STAR(p) model can be calculated by solving

$$z^p - \sum_{j=1}^p (\hat{\pi}_{1j} + \hat{\pi}_{2j} F) z^{p-j} = 0$$

for $F = 0, 1$. The reason to obtain the roots when F takes extreme values is that they provide good description of the local dynamics of recession and expansion. The behavior of the series y_t is a function of its lagged values and the estimated parameters. Nonetheless, an explicit solution of y_t is usually unavailable. But the limiting behavior of y_t as $t \rightarrow \infty$ given a starting value set is achievable and it can:

- (i). converge to a stable stationary point;
- (ii). display a limit cycle, i.e. the same set of q values $\bar{y}_t, \bar{y}_{t-1}, \dots, \bar{y}_{t-q-1}$ which repeats itself over time independent of the starting values;
- (iii). diverge;
- (iv). demonstrate chaotic behavior. The process does not diverge, but small changes in starting values can generate quite different solution paths.

Eitrheim and Teräsvirta [40] construct a series of LM tests to exam the adequacy of the STAR model. They are the LM test for the hypothesis of no error autocorrelation, the LM test for the hypothesis of

no remaining nonlinearity, and the LM test for the hypothesis of parameter constancy. Small-sample properties of the tests' F statistics are provided by simulation.

Though the applications of the STAR to economic data series have occurred only recently, the literature is extensive and still growing.

A representative is Teräsvirta and Anderson's [116] effort to estimate quarterly OECD and European industrial production series with the STAR model. The LM tests reject linearity for most of the industrial production series. For the estimated STAR models, the characteristic polynomial of the recession regime has at least one complex pair of explosive roots while the expansion regime is stationary, suggesting that business cycles are asymmetric in the sense that the transition from recession to expansion is strong and swift but that from growth to contraction is weak and slow.

Based on the theoretical work of Eitrheim and Teräsvirta [40], Öcal and Osborn [90] apply the STAR model with two additive smooth transition components to the UK consumers' expenditure and industrial production. It is shown that the consumption is characterized by two regimes: recession and expansion and the industrial production can be divided into three phases: recession, normal growth and high growth.

Leybourne and Mizen [76] use the smooth transition analysis to endogenously determine the transition path (speed and midpoint of the transition) of the price series between two monetary policy regimes. The result shows that the transition process is related to the central bank independence and the widespread economy slowdown in early 1990s.

Sarantis [107] tests and models the nonlinearities in the real effective exchange rates of ten major industrial countries using the STAR model. The LM test proposed by Teräsvirta and his coauthors rejects linearities for eight exchange rates. Among them, the effective exchange rates of Germany, France and Belgium are specified as LSTAR models, suggesting different dynamics during the expansion and contraction phases; all the other five exchange rates are modeled as the ESTAR model, implying that the exchange rates move from upper or lower regimes toward the middle ground in a similar fashion. The estimated parameters γ , which controls the transition speed of the STAR model, are small in all countries,

indicating that the transition of the exchange rates from one regime to another is quite slow, which is contrary to the assumption of the TAR model and the Markov regime switching model. He also finds that the estimated parameters c are small though significant, which implies the appreciation and depreciation of the exchange rates are described by the different regimes of the STAR model. Characteristic roots of the estimated STAR model are also calculated to investigate the dynamic behavior of the effective exchange rates. Evidence of cyclical movements and asymmetric dynamics are found. The STAR models also outperform the Markov regime-switching model in out-of-sample forecasting.

Inspired by the idea that adjustment process of deviation from PPP is nonlinear due to the existence of transaction cost in exchange rate market, Michael et al. [83], Baum et al. [8] introduce the ESTAR model to the study of purchasing power parity. Linearity is rejected for every adjustment process, and the estimated ESTAR model depicts that adjustment process behaves like random walk when deviations from PPP are small but is mean-reverting for large deviations from PPP.

2.7 The Autoregressive Conditional Heteroskedastic (ARCH) Model

All of the above mentioned models assume an IID disturbance term $\{\varepsilon_t\}$, so that the nonlinearity is attributed to the model mean rather than to the model variance. However, the ARCH model has a disturbance term whose conditional variance is not constant, which gives rise to nonlinearities.

The ARCH model was first introduced by Engle [44]. For an ARCH(p, m) series, its model specification is linear in the form of an AR(p) as

$$y_t = \mu + \sum_{i=1}^p \phi_i y_{t-i} + v_t. \quad (2.31)$$

However it tries to capture the nonlinearity by the disturbance term v_t , which can be written as

$$v_t = \varepsilon_t \sqrt{h_t}, \quad (2.32)$$

where

$$h_t = \alpha_0 + \sum_{j=1}^m \alpha_j v_{t-j}^2. \quad (2.33)$$

To ensure that $h_t > 0$, it should be satisfied that $\alpha_0 > 0$, $\{\alpha_j\}_{j=1}^m \geq 0$. Also, condition $\sum_{j=1}^m \alpha_j < 1$ is

set to ensure weak stationarity of the series. $\{\varepsilon_t\}$ is an IID process with its variance equal to 1 and is independent of v_{t-j} , $j = 1, \dots, m$.

If h_t is an ARMA(m, n) process, we will have a $\{y_t\}$ series called GARCH(p, m, n) (generalized ARCH model) which was brought forward by Bollerslev [10]. However, we do not go into detail about GARCH model here by the same reason that we discuss NLAR model instead of NLARMA model in this paper.

It can be seen that for an ARCH model, the conditional variance of $\{v_t\}$ is predictable. By a little computation, the conditional variance of $\{v_t\}$ can be expressed by its past values

$$\text{var}(v_t | v_{t-1}, v_{t-2}, \dots) = E[v_t^2 | v_{t-1}, v_{t-2}, \dots] = h_t = \alpha_0 + \sum_{i=1}^m \alpha_i v_{t-i}^2. \quad (2.34)$$

Therefore a large conditional variance tends to be followed by another large conditional variance and small conditional variance by small conditional variance. This feature implies that ARCH model is capable of producing clusters of outliers, which is common in economic and financial time series.

As for the testing of ARCH disturbances, Engle considers the LM test as an ideal candidate. The LM test for the ARCH(p, m) model in the form (2.31), (2.32), and (2.33) will be

$$LM = \frac{1}{2} \left(\sum_{t=1}^T \hat{z}_t \hat{f}_t \right)' \left(\sum_{t=1}^T \hat{z}_t \hat{z}_t' \right)^{-1} \left(\sum_{t=1}^T \hat{z}_t \hat{f}_t \right)$$

with $\hat{z}_t = [1 \quad \hat{v}_{t-1}^2 \quad \dots \quad \hat{v}_{t-m}^2]'$, $\hat{f}_t = \hat{\sigma}_t^{-2} \hat{v}_t^2 - 1 = (T^{-1} \sum_{i=1}^T \hat{v}_i^2)^{-1} \hat{v}_t^2 - 1$, $\hat{v}_{t-i}^2 = \hat{\varepsilon}_{t-i}^2$, $\forall i$ under the null restrictions.

Since information matrix of the ARCH model is block diagonal, the following stepwise estimator proposed by Engle is asymptotically efficient. First, estimate the model mean parameters α by ordinary least squares, and obtain the residuals. Then estimate the conditional variance parameters β from the residuals. Based upon the estimates of α , a near estimate of β can be found. The cycle of estimating α and β sequentially continues until convergence occurs. The iterations are calculated using the scoring algorithm, and the resulted estimates of α and β are maximum likelihood estimates.

Engle applies the ARCH model in the estimation of the variance of inflation for the UK. The quarterly consumer price index is used for the investigation of the variance of inflation for the UK

from 1958:II to 1977:II. It is assumed that price inflation follows wage inflation, so the lagged quarterly index of nominal wage rates is included as a pre-determined variable. The mean of the ARCH model is constructed as the following

$$\dot{p} = \beta_1 \dot{p}_{-1} + \beta_2 \dot{p}_{-4} + \beta_3 \dot{p}_{-5} + \beta_4 (p - w)_{-1} + \beta_5, \quad (2.35)$$

where \dot{p} is the first difference of the log of the quarterly consumer price index p and w is the log of the quarterly index of nominal wage rates.

The fit of the UK inflation in the form (2.35) is reasonable and good. The inclusion of the first, fourth, and fifth lags of the first difference of the endogenous variable in the model accounts for the typical seasonal behavior. Least squares estimation produces less than 1 percent standard error of forecast and t -statistic greater than 3. The estimates of β_2 and β_3 have equal and opposite signs, suggesting that it is the acceleration of inflation one year ago which explains much of the short-run behavior in prices. Serial correlation is checked and rejected by the LM test and Durbin's h statistic.

The LM test for linear ARCH effects shows that a fourth-order ARCH process is significant. In order to meet the non-negativity and stationarity constraints of the ARCH model, a two-parameter variance function is assumed in the form

$$h_t = \alpha_0 + \alpha_1 (0.4\epsilon_{t-1}^2 + 0.3\epsilon_{t-2}^2 + 0.2\epsilon_{t-3}^2 + 0.1\epsilon_{t-4}^2),$$

where a linearly declining set of weights is given to the squared residuals by assuming that agents discount past residuals. The coefficients are estimated by the maximum likelihood method.

Comparing the ARCH model with the least squares model through standard errors, one-step-ahead forecast variances and outliers, Engle shows that the ARCH process is able to improve upon the performance of the linear least squares model and is able to provide more realistic forecast variances.

After the work by Engle, the literature on ARCH expands dramatically, leading to the introduction of GARCH, EGARCH, ARCH-M models, etc. The ARCH effects have been seen especially significant in financial data such as stock return, interest rates, and foreign exchange rates. Surveys include Bollerslev et al. [11, 12], Engle [45], Ghysels et al. [50], and Shephard [109].

CHAPTER 3

TESTING NONLINEARITIES IN TIME SERIES

3.1 Introduction

Linear or Nonlinear?

For a given time series, there are many specific models that we can fit it to. Broadly speaking, these models can be divided into two groups: linear time series models and nonlinear time series models. There are numerous types of linear and nonlinear models. Chapter Two gives a glimpse of some of them. So, which type of model should we choose? And how can we know that the selected model is appropriate?

In this chapter, we will focus on Lagrange multiplier (LM) tests for detecting nonlinearities in time series. The LM statistics can be calculated to test the null of linearity against almost every major nonlinear model. And, for each LM test, only the estimation of the null linear model is required. This means that an LM test constructed for one alternative can be used against a variety of other alternatives. The finite-sample size and power of different LM tests against various nonlinear models will be compared through the Monte Carlo method. Because the asymptotic critical values might not be a good approximation of the true critical values due to factors such as small sample size, the bootstrap method is also used to generate critical values.

The LM test is a “specific” linearity test in the sense that it has a linear time series model (usually an AR(MA) type model) under the null and a specified nonlinear time series model as the alternative. There are other types of linearity tests called “general” linearity tests that do not have a particular nonlinear alternative in mind. Examples include McLeod and Li [80], Keenan [68], Tsay [124], Subba Rao and Gabr [111], Hinich [63], Ashley et al. [3], NEGM [89], White [129, 130]. The most widely used general linearity test is the BDS test. In this chapter, we will also discuss the BDS test and provide a

comparison to the LM test.

3.1.1 Lagrange Multiplier Tests

First suggested by Aitchison and Silvey [1, 2], the Lagrange multiplier (LM) method tests the effect of imposing the first order conditions for a maximum of the likelihood function under the restrictions of the null hypothesis.

Consider a model that has a finite number K of unknown parameters $\theta = (\theta_1, \dots, \theta_K)'$, log likelihood function $L(\theta)$, and the null hypothesis is in the form

$$H_0 : h_j(\theta) = 0$$

with $j = 1, \dots, p$ and $p < K$.

Differentiating the Lagrangian function

$$L(\theta) + \sum_{j=1}^p \lambda_j h_j(\theta)$$

with respect to the unknown parameters θ and Lagrange multipliers λ_j leads to the first order conditions

$$\begin{cases} \hat{D} + \hat{H}\hat{\lambda} = \mathbf{0} \\ h_j(\hat{\theta}) = 0 \end{cases}$$

where $\hat{D} = \left\{ \frac{\partial L(\hat{\theta})}{\partial \theta_i} \right\}_{K \times 1}$, $\hat{H} = \left\{ \frac{\partial h_j(\hat{\theta})}{\partial \theta_i} \right\}_{K \times p}$, $\hat{\lambda} = (\hat{\lambda}_1 \dots \hat{\lambda}_p)'_{p \times 1}$, $i = 1, \dots, K$, $j = 1, \dots, p$.

The logic behind the LM test is that when the null hypothesis is true the restricted estimates $\hat{\theta}$ will tend to be near the unrestricted maximum likelihood estimate so that \hat{D} will be close to the zero vector. Therefore the LM statistic can be calculated as

$$LM = \hat{D}'\hat{\mathcal{J}}^{-1}\hat{D}, \quad (3.1)$$

where $\hat{\mathcal{J}}$ is the Fisher's information matrix evaluated at the restricted estimates $\hat{\theta}$ when the null hypothesis is true. If the regularity condition that the order of differentiation and integration can be interchanged is satisfied, the information matrix can be obtained from the first derivatives of the log likelihood and be written as

$$\mathcal{J} = E \left[-\frac{\partial^2 L}{\partial \theta \partial \theta'} \right] = E \left[\left(\frac{\partial L}{\partial \theta} \right) \left(\frac{\partial L}{\partial \theta} \right)' \right].$$

According to Crowder [28],

$$p \lim \left\{ \mathcal{J}^{-1} \left[-\frac{\partial^2 L}{\partial \theta \partial \theta'} \right] \right\} = I_K,$$

where I_K is the $K \times K$ identity matrix. So, instead of the information matrix \mathcal{J} , $-\frac{\partial^2 L}{\partial \theta \partial \theta'}$ or $\left(\frac{\partial L}{\partial \theta}\right) \left(\frac{\partial L}{\partial \theta}\right)'$ can be used to calculate the LM statistic without affecting the asymptotic properties of the statistic.

We can always apply the LM test to a nested situation, where the null model is a subset of the alternative model, the parameter vector θ can be partitioned into two subsets, and the restrictions are imposed on one of the subsets. Let $\theta_1 = (\theta_1 \cdots \theta_p)'$ denote the vector of parameters of interest and $\theta_2 = (\theta_{p+1} \cdots \theta_K)'$ the vector of nuisance parameters, then the null hypothesis is

$$H_0 : \theta_1 = \theta_{10}$$

or

$$H_0 : h(\theta) = \begin{bmatrix} i_{1 \times p} & 0_{1 \times (K-p)} \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} - \theta_{10} = 0.$$

where the elements of $i_{1 \times p}$ are all ones and the elements of $0_{1 \times (K-p)}$ are all zeros.

Partitioning D^1

$$D = \begin{bmatrix} D_1 \\ D_2 \end{bmatrix} = \begin{bmatrix} D_1 \\ \mathbf{0} \end{bmatrix}$$

and \mathcal{J}

$$\mathcal{J} = \begin{bmatrix} \mathcal{J}_{11} & \mathcal{J}_{12} \\ \mathcal{J}_{21} & \mathcal{J}_{22} \end{bmatrix}$$

will lead to

$$LM = \widehat{D}'_1 \left(\widehat{\mathcal{J}}_{11} - \widehat{\mathcal{J}}_{12} \widehat{\mathcal{J}}_{22}^{-1} \widehat{\mathcal{J}}_{21} \right)^{-1} \widehat{D}_1. \quad (3.2)$$

Suppose for example, the non-linear model under the alternative hypothesis has the form

$$y_t = g(x_t; \theta) + \varepsilon_t,$$

where x_t is one-dimensional, and the ε_t 's are normally, identically and independently distributed as $\text{NIID}(0, \sigma^2)$. It is assumed that $g(\cdot)$ is linear if $\theta_1 = \theta_{10}$, and is nonlinear otherwise.

¹ $D_2 = 0$ since if there is no restrictions on θ_2 , the first derivative vector of the likelihood function with respect to θ_2 will be a zero vector.

Under the above assumptions the log likelihood (omitting constants) will be

$$L = -\frac{1}{2}T \log \sigma^2 - \frac{1}{2}\sigma^{-2} \sum_{t=1}^T \varepsilon_t^2.$$

Setting $z_{1t} = \left(\frac{\partial \varepsilon_t}{\partial \theta_1} \right)_{\text{when } H_0 \text{ is true}}$, $z_{2t} = \left(\frac{\partial \varepsilon_t}{\partial \theta_2} \right)_{\text{when } H_0 \text{ is true}}$, $M_{12} = M'_{21} = \sum_{t=1}^T z_{1t} z'_{2t}$ and $M_{ii} = \sum_{t=1}^T z_{it} z'_{it}$ for $i = 1, 2$, $\sigma^2 = E \left[T^{-1} \sum_{t=1}^T \varepsilon_t^2 \right]$, then the statistic is given by²

$$\text{LM} = \hat{\sigma}^{-2} \left(\sum_{t=1}^T \hat{z}_{1t} \hat{\varepsilon}_t \right)' \left(\widehat{M}_{11} - \widehat{M}_{12} \widehat{M}_{22}^{-1} \widehat{M}_{21} \right)^{-1} \left(\sum_{t=1}^T \hat{z}_{1t} \hat{\varepsilon}_t \right) \quad (3.3)$$

with $\{\hat{\varepsilon}_t\}$ being the residuals obtained when the null is true.

Formula (3.3) can also be interpreted as TR^2 with R^2 being the coefficient of multiple determination in the regression of $\hat{\varepsilon}_t$ on \hat{z}_{1t} and \hat{z}_{2t} (Harvey [62]).

Under the null hypothesis, the LM test statistic is asymptotically distributed as χ^2 with degrees of freedom equal to p , the dimension of θ_1 or the number of coefficient restrictions under the null hypothesis.

Early application of LM tests in time series analysis can be seen in the work of Pagan [93]. Later contributions include Engle [44] who develops an LM test against ARCH model. Weiss [128] and Saikkonen and Luukkonen [106] have considered an LM test for testing a bilinear alternative. Saikkonen and Luukkonen also propose an LM test for the EAR model. Luukkonen, Saikkonen and Teräsvirta [78] and Teräsvirta [115] suggest an LM test for the STAR model. Details about LM tests for various nonlinear models will be discussed in the section "Models and Tests".

²Since

$$D_1 = \left(\frac{\partial L}{\partial \theta_1} \right) = \left(\frac{\partial \left(-\frac{1}{2} \sigma^{-2} \sum_{t=1}^T \varepsilon_t^2 \right)}{\partial \theta_1} \right) = -\sigma^{-2} \left(\sum_{t=1}^T \frac{\partial \varepsilon_t}{\partial \theta_1} \varepsilon_t \right) = -\sigma^{-2} \left(\sum_{t=1}^T z_{1t} \varepsilon_t \right),$$

and

$$\begin{aligned} I_{ij} &= E \left[\left(\frac{\partial L}{\partial \theta_i} \right) \left(\frac{\partial L}{\partial \theta_j} \right)' \right] = \sigma^{-4} E \left[\left(\sum_{t=1}^T \frac{\partial \varepsilon_t}{\partial \theta_i} \varepsilon_t \right) \left(\sum_{t=1}^T \frac{\partial \varepsilon_t}{\partial \theta_j} \varepsilon_t \right)' \right] \\ &= \sigma^{-2} \sum_{t=1}^T \left(\frac{\partial \varepsilon_t}{\partial \theta_i} \right) \left(\frac{\partial \varepsilon_t}{\partial \theta_j} \right)' = \sigma^{-2} \sum_{t=1}^T z_{it} z'_{jt} = \sigma^{-2} M_{ij}, \end{aligned}$$

$i, j = 1, 2$.

3.1.2 The BDS Test

The BDS test gets its name from its original authors William Brock, Davis Dechert and Jose Scheinkman, who develop it in 1987 [15]. But, it is more fully described in the revised paper published by Brock, Dechert, Scheinkman and LeBaron in 1996 [16].

The BDS test is a test for white noise. The alternative hypothesis of the test includes both non-white linear and non-white nonlinear processes. Therefore, the BDS test does not provide a direct test for nonlinearity. Nevertheless, if the linear structure is removed through prior pre-whitening and the BDS test rejects the hypothesis that the residual process is white noise, this could be interpreted as evidence of nonlinearity in the original series.

The BDS test uses the correlation function (also called the correlation integral) as the test statistic, which is calculated as follows.

First, for some integer embedding dimension $m \geq 2$, $x_t^m = (x_t, x_{t-1}, \dots, x_{t-m+1})$, the “ m -histories” of the data are calculated for $t = m, m+1, \dots, n$. Then $c_{m,n}(\epsilon)$ is calculated, which counts the proportion of points in m -dimensional hyperspace that are within a distance ϵ of each other:

$$c_{m,n}(\epsilon) = \frac{2}{(n-m+1)(n-m)} \sum_{s=m}^n \sum_{t=s+1}^n \prod_{j=0}^{m-1} I_{\epsilon}(x_{s-j}, x_{t-j}). \quad (3.4)$$

where $I_{\epsilon} = \begin{cases} 1 & \text{if } \|x_{s-j}, x_{t-j}\| < \epsilon \\ 0 & \text{if } \|x_{s-j}, x_{t-j}\| \geq \epsilon \end{cases}$, $\|\cdot\|$ denotes the superior norm. Then the BDS test statistic can be constructed as³

$$W_{m,n}(\epsilon) = \sqrt{n-m+1} \frac{c_{m,n}(\epsilon) - c_{1,n-m+1}(\epsilon)^m}{\sigma_{m,n}(\epsilon)}, \quad (3.5)$$

where the estimated variance of $c_{m,n}(\epsilon) - c_{1,n-m+1}(\epsilon)^m$ is given by

$$\sigma_{m,n}^2(\epsilon) = 4 \left\{ k^m + 2 \sum_{j=1}^{m-1} k^{m-j} c^{2j} + (m-1)^2 c^{2m} - m^2 k c^{2m-2} \right\}. \quad (3.6)$$

³LeBaron [74], Barnett et al [7] and many others state this equation as

$$W_{m,n}(\epsilon) = \sqrt{n} \frac{c_{m,n}(\epsilon) - c_{1,n}(\epsilon)^m}{\sigma_{m,n}(\epsilon)}.$$

Kanzler [67] suggests that the right formula should be

$$W_{m,n}(\epsilon) = \sqrt{n-m+1} \frac{c_{m,n}(\epsilon) - c_{1,n-m+1}(\epsilon)^m}{\sigma_{m,n}(\epsilon)},$$

arguing that the other formula gives the misleading impression that correlation integral of dimension 1 is to be estimated over the full sample and that the ratio term is to be multiplied by the square root of the full sample size.

The parameter c is the first-dimensional correlation integral

$$c = c_{1,n}(\epsilon), \quad (3.7)$$

and $k(\epsilon)$ is the probability of any triplet of observations lying within distance ϵ of each other and is computed as

$$k(\epsilon) = \frac{2}{n(n-1)(n-2)} \sum_{t=1}^n \sum_{s=t+1}^n \sum_{r=s+1}^n \{I_\epsilon(x_t, x_s) I_\epsilon(x_s, x_r) + I_\epsilon(x_t, x_r) I_\epsilon(x_r, x_s) + I_\epsilon(x_s, x_t) I_\epsilon(x_t, x_r)\}. \quad (3.8)$$

Kanzler [67] shows that $k(\epsilon)$ can be transformed to

$$k(\epsilon) = \frac{1}{n(n-1)(n-2)} \times \left\{ \sum_{t=1}^n \left[\sum_{s=1}^n I_\epsilon(x_t, x_s) \right]^2 - 3 \sum_{s=1}^n \sum_{t=s+1}^n I_\epsilon(x_s, x_t) + 2n \right\} \quad (3.9)$$

to simplify the computational algorithm.

Brock, Dechert, and Scheinkman [15] prove that $W_{m,n}(\epsilon)$ is distributed asymptotically as a standard normal $N(0,1)$ for any m and ϵ under the null hypothesis that $\{x_t\}$ is independently and identically distributed.

The need to choose the values of ϵ (the radius of the hypersphere which determines whether two points are “close” or not) and m (the value of the embedding dimension) can be a complication in using the BDS test. Kanzler [67] illustrates that the BDS statistic appears to be most efficiently estimated if the measure of dimensional distance ϵ is chosen such that the first-order correlation integral estimate c lies around 0.7. Unfortunately, the cost of finding the optimal ϵ is quite high in terms of CPU time and memory. For a near-normal distribution, ϵ being three halves the standard deviation of the actual data will achieve the same efficiency without any extra computational burden. Brock et al. [18] recommend that ϵ is set to between half and three halves the standard deviation of the actual data and m is set in line with the number of observations available (e.g. use only $m \leq 5$ for $T \leq 500$, etc.).

3.1.3 Bootstrap Method and Monte Carlo Simulation

When we are performing hypothesis testing, sometimes we know the exact distribution of the test statistic, but usually the knowledge is less complete. Often large-sample theory is used to justify an

approximating distribution. But a test based on asymptotic results may not give good results for small samples. In other situations, the exact distribution of a test statistic may be known, but the test may not be robust to violations of model assumptions. An alternative way to estimate the sampling distribution of a statistic is by resampling the data itself using jackknife [125] or bootstrap [38] procedures. The latter is receiving more and more attention. Textbooks describing the bootstrap include LePage and Billard [75], Efron and Tibshirani [39] and Hjorth [64].

The bootstrap method might be the simplest technique for simulating the probability distribution of any statistic. Bootstrapping amounts to resampling a record, with replacement, to generate B bootstrap samples, from which one can compute B estimates of a given statistic, leading to an empirical probability distribution of the statistic. Suppose we wish to estimate the cumulative distribution function of a statistic $\hat{\theta}$ that is estimated from a given random sample x_t , $t = 1, \dots, n$, which we denote X . A bootstrap sample X_i is generated by resampling n values from the observed data. That is, each observation x_i is resampled, with replacement, with an equal probability of $1/n$. A total of B bootstrap samples X_i , $i = 1, \dots, B$ is obtained. Each bootstrap sample X_i yields a bootstrap estimate of the statistic θ leading to the B bootstrap estimates $\hat{\theta}_i$, $i = 1, \dots, B$. The empirical distribution of the $\hat{\theta}_i$'s is used to approximate the distribution of $\hat{\theta}$. This practice is valid given the assumption that the elements of X are independently identically distributed. However, time series data are usually not IID, so the resampling algorithm requires modification.

In this chapter, the residual bootstrap method that aims at non-IID type of data is conducted. If there is a parametric model that can fit the data with IID residuals, the bootstrap method can be applied to the residuals. The implementation steps are as follows. First, the time series is estimated by a parametric model and the residuals are obtained. Second, the residuals are resampled. Finally, bootstrap replicates of the time series data are constructed from the estimated model and the bootstrapped residuals.

We use Monte Carlo simulation to compute the size and power of the bootstrap tests. The first step of the Monte Carlo simulation is to specify the parameters and initial conditions of the model

sequence. The next step is to generate from a given distribution pseudo-random numbers. Usually the random numbers are designed to be normally distributed and serially uncorrelated so that they can represent an IID sequence. The third step is to construct the model sequence by using the parameters, initial conditions and random numbers. The fourth step is to repeat steps two and three thousands of times. The reason to do the repetition for so many times is to ensure that the statistical properties of the constructed sequence are in accord with the true distribution. One limitation of the Monte Carlo experiment is that it is specific to the assumptions used to generate the simulated data. If the sample size, the parameters or the initial conditions are changed in the data generating process, a new simulation needs to be performed. However, if the change is small, we can expect the results to hold approximately by continuity.

3.2 Models and Tests

For the linearity tests, the model is linear under the null hypothesis and the alternative can be a nonlinear model that depends on the specific test. We will specify the linear model as the AR model:

$$y_t = \mu + \sum_{i=1}^p \phi_i y_{t-i} + \varepsilon_t, \quad (3.10)$$

where the roots of the polynomial $1 - \sum_{i=1}^p \phi_i z^i = 0$ lie outside the unit circle and $\{\varepsilon_t\} \sim \text{NIID}(0, \sigma^2)$. Note that μ is the intercept of the model; the mean of the stationary series $\{y_t\}$ is $\mu(1 - \sum_{i=1}^p \phi_i)^{-1}$.

Since the null model is $\text{AR}(p)$, we will have $\varepsilon_t = y_t - \mu - \sum_{i=1}^p \phi_i y_{t-i}$ and $\theta_2 = [\mu \ \phi_1 \ \cdots \ \phi_p]'$.

So

$$z_{2t} = -[1 \ y_{t-1} \ \cdots \ y_{t-p}]' \quad (3.11)$$

for each LM test, and z_{1t} will be determined by the specific alternative model. We will consider some major representatives of nonlinear time series models: the bilinear (BL) model, exponential autoregressive (EAR) model, threshold autoregressive (TAR) model, smooth transition autoregressive (STAR) model (including Exponential STAR (ESTAR) and Logistic STAR (LSTAR) model), and autoregressive conditional heteroskedastic (ARCH) model.

3.2.1 The BL Model and BLT Statistic

The BL model adds the cross-product terms of y_{t-i} and ε_{t-j} to a linear model to account for the nonlinearity. For the BL(p, m, k), the model structure can be written as

$$y_t = \mu + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{i=1}^m \sum_{j=1}^k c_{ij} y_{t-i} \varepsilon_{t-j} + \varepsilon_t, \quad (3.12)$$

where $c_{mk} \neq 0$.

With $\theta_1 = [c_{11} \ c_{12} \ \cdots \ c_{mk}]' = \mathbf{0}$ as the null hypothesis, the LM statistic testing the linear null (3.10) against the bilinear model (3.12) can be constructed with

$$\begin{aligned} z_{1t} &= \left(\frac{\partial \varepsilon_t}{\partial \theta_1} \right)_{\theta_1=0} \\ &= \left[\frac{\partial}{\partial \theta_1} \left(y_t - \mu - \sum_{i=1}^p \phi_i y_{t-i} - \sum_{i=1}^m \sum_{j=1}^k c_{ij} y_{t-i} \varepsilon_{t-j} \right) \right]_{\theta_1=0} \\ &= - [y_{t-1} \varepsilon_{t-1} \quad y_{t-1} \varepsilon_{t-2} \quad \cdots \quad y_{t-m} \varepsilon_{t-k}]'. \end{aligned} \quad (3.13)$$

Saikkonen and Luukkonen [106] derive the necessary and sufficient conditions for the BLT(m, k) (the LM test for the BL(p, m, k) model) to be well defined. In this chapter, we will assume $\phi_p \neq 0$ and $k \leq p+1$. With this restriction, the BLT statistic will be distributed asymptotically as a χ_{mk}^2 under the null hypothesis since the dimension of the vector θ_1 is mk .

3.2.2 The EAR Model and EART Statistic

For the EAR model such as

$$y_t = \mu + \sum_{i=1}^p \phi_i y_{t-i} + [\exp(-\gamma y_{t-d}^2) - 1] \sum_{i=1}^p \theta_i y_{t-i} + \varepsilon_t, \quad (3.14)$$

with $\gamma \neq 0$,

$$\begin{aligned} \left(\frac{\partial \varepsilon_t}{\partial \gamma} \right)_{\gamma=0} &= \left\{ \frac{\partial}{\partial \gamma} \left\{ y_t - \mu - \sum_{i=1}^p \phi_i y_{t-i} - [\exp(-\gamma y_{t-d}^2) - 1] \sum_{i=1}^p \theta_i y_{t-i} \right\} \right\}_{\gamma=0} \\ &= \left[\exp(-\gamma y_{t-d}^2) \sum_{i=1}^p \theta_i y_{t-i} y_{t-d}^2 \right]_{\gamma=0} \\ &= \sum_{i=1}^p \theta_i y_{t-i} y_{t-d}^2. \end{aligned}$$

Therefore the $\text{EART}(p)$ — the LM test statistic against $\text{EAR}(p)$ — can be constructed with

$$z_{1t} = - \left[y_{t-1}y_{t-d}^2 \quad y_{t-2}y_{t-d}^2 \quad \cdots \quad y_{t-p}y_{t-d}^2 \right]', \quad (3.15)$$

which has an asymptotic χ_p^2 distribution under linearity.

3.2.3 The TAR Model

Here we consider a simple two-regime TAR model, which can be expressed as

$$y_t = \begin{cases} \mu^{(1)} + \sum_{i=1}^p \phi_i^{(1)} y_{t-i} + \varepsilon_t^{(1)} & \text{if } y_{t-d} \geq c \\ \mu^{(2)} + \sum_{i=1}^p \phi_i^{(2)} y_{t-i} + \varepsilon_t^{(2)} & \text{if } y_{t-d} < c \end{cases}. \quad (3.16)$$

where $\phi_i^{(1)} \neq \phi_i^{(2)}$, for some $i = 1, \dots, p$. There are two regimes and in each regime there is an $\text{AR}(p)$ process.

Note that there is no corresponding LM test statistic for the TAR model. This is because the likelihood function is not differentiable with respect to the threshold parameter c .

3.2.4 The STAR Model and START Statistic

The $\text{ESTAR}(p)$ model has the general form

$$y_t = \mu_1 + \sum_{i=1}^p \phi_{1i} y_{t-i} + \left(\mu_2 + \sum_{i=1}^p \phi_{2i} y_{t-i} \right) \times \left\{ 1 - \exp \left[-\gamma (y_{t-d} - c)^2 \right] \right\} + \varepsilon_t, \quad (3.17)$$

and the basic $\text{LSTAR}(p)$ model can be written as

$$y_t = \mu_1 + \sum_{i=1}^p \phi_{1i} y_{t-i} + \left(\mu_2 + \sum_{i=1}^p \phi_{2i} y_{t-i} \right) \times \left[\left\{ 1 + \exp \left[-\gamma^* (y_{t-d} - c^*) \right] \right\}^{-1} - 0.5 \right] + \varepsilon_t. \quad (3.18)$$

Under the null hypothesis $\gamma = 0$ or $\gamma^* = 0$, both will become a linear AR model.

To construct the LM test for the ESTAR model (ESTART), we have

$$\begin{aligned} \left(\frac{\partial \varepsilon_t}{\partial \gamma} \right)_{\gamma=0} &= \left\{ \frac{\partial}{\partial \gamma} \left\{ - \left(\mu_2 + \sum_{i=1}^p \phi_{2i} y_{t-i} \right) \times \left\{ 1 - \exp \left[-\gamma (y_{t-d} - c)^2 \right] \right\} \right\} \right\}_{\gamma=0} \\ &= \left\{ - \left(\mu_2 + \sum_{i=1}^p \phi_{2i} y_{t-i} \right) \exp \left[-\gamma (y_{t-d} - c)^2 \right] (y_{t-d} - c)^2 \right\}_{\gamma=0} \\ &= - \left(\mu_2 + \sum_{i=1}^p \phi_{2i} y_{t-i} \right) (y_{t-d} - c)^2. \end{aligned}$$

After transformation, the concise form of z_{1t} for ESTAR(p) is

$$z_{1t} = - \left[y_{t-1}y_{t-d} \quad y_{t-2}y_{t-d} \quad \cdots \quad y_{t-p}y_{t-d} \quad y_{t-1}y_{t-d}^2 \quad y_{t-2}y_{t-d}^2 \quad \cdots \quad y_{t-p}y_{t-d}^2 \right]' . \quad (3.19)$$

The ESTAR(p) statistic has an asymptotic χ_{2p}^2 distribution under linearity.

For the LM test for the LSTAR model (LSTART), it can be shown that

$$\begin{aligned} \left(\frac{\partial \varepsilon_t}{\partial \gamma^*} \right)_{\gamma^*=0} &= \left\{ \frac{\partial}{\partial \gamma^*} \left\{ - \left(\mu_2 + \sum_{i=1}^p \phi_{2i} y_{t-i} \right) \times \left[\{1 + \exp[-\gamma^* (y_{t-d} - c^*)]\}^{-1} - 0.5 \right] \right\} \right\}_{\gamma^*=0} \\ &= \left\{ - \left(\mu_2 + \sum_{i=1}^p \phi_{2i} y_{t-i} \right) \frac{\exp[-\gamma^* (y_{t-d} - c^*)] (y_{t-d} - c^*)}{\{1 + \exp[-\gamma^* (y_{t-d} - c^*)]\}^2} \right\}_{\gamma^*=0} \\ &= - \left(\mu_2 + \sum_{i=1}^p \phi_{2i} y_{t-i} \right) \frac{(y_{t-d} - c^*)}{4} . \end{aligned}$$

Therefore, the LM test against LSTAR(p) model can be obtained with

$$z_{1t} = - \left[y_{t-1}y_{t-d} \quad y_{t-2}y_{t-d} \quad \cdots \quad y_{t-p}y_{t-d} \right]' .$$

Teräsvirta [115] shows that the LM tests for LSTAR based on the above regressions may not have much power for alternative with $\mu_2 \neq 0$. He suggests constructing an LM test statistic for an auxiliary regression that is based on third order Taylor expansion but adjusted for small samples. The auxiliary regression is

$$y_t = \mu + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=1}^p \beta_{1j} y_{t-j} y_{t-d} + \beta_2 y_{t-d}^3 + \varepsilon_t \quad (3.20)$$

with null hypothesis $\beta_{11} = \beta_{12} = \cdots = \beta_{1p} = \beta_2 = 0$, so the z_{1t} becomes

$$z_{1t} = - \left[y_{t-1}y_{t-d} \quad y_{t-2}y_{t-d} \quad \cdots \quad y_{t-p}y_{t-d} \quad y_{t-d}^3 \right]' , \quad (3.21)$$

which is equivalent to solving z_{1t} by adding the third order differentiation term $\frac{\partial^3 \varepsilon_t}{\partial \gamma^{*3}}$ to the original formula of z_{1t} so that

$$z_{1t} = \left(\begin{array}{c} \frac{\partial \varepsilon_t}{\partial \gamma^*} \\ \frac{\partial^3 \varepsilon_t}{\partial \gamma^{*3}} \end{array} \right)_{\gamma^*=0} .$$

In the following LM testing of the LSTAR(p) model, vector z_{1t} will be in the form (3.21), therefore the LSTART(p) statistic has an asymptotic χ_{p+1}^2 distribution if the null is true.

3.2.5 The ARCH Model and ARCHT Statistic

A different approach is required to calculate the LM statistics for ARCH(p, m) type models

$$y_t = \mu + \sum_{i=1}^p \phi_i y_{t-i} + v_t \quad (3.22)$$

with

$$v_t = \varepsilon_t \sqrt{h_t} = \varepsilon_t \sqrt{\alpha_0 + \sum_{j=1}^m \alpha_j v_{t-j}^2}, \quad (3.23)$$

where $\{\varepsilon_t\} \sim \text{NIID}(0, 1)$.

As stated in Breusch and Pagan [14], v_t will be distributed as $\text{NIID}(0, h_t = z_t' \alpha)$ with $z_t' = [1 \ v_{t-1}^2 \ \dots \ v_{t-m}^2]$ and $\alpha' = [\alpha_0 \ \alpha_1 \ \dots \ \alpha_m]$. So

$$\frac{\partial h_t}{\partial \alpha} = z_t.$$

The log likelihood function (omitting the constant term) is

$$L(\alpha) = -\frac{1}{2} \sum_{t=1}^T \log h_t - \frac{1}{2} \sum_{t=1}^T h_t^{-1} v_t^2.$$

Therefore,

$$D_1 = \frac{\partial L}{\partial \alpha} = \frac{1}{2} \sum_{t=1}^T h_t^{-1} z_t f_t,$$

and

$$J_{11} = \frac{\partial L}{\partial \alpha \partial \alpha'} = \frac{1}{2} \sum_{t=1}^T h_t^{-2} z_t z_t',$$

where $f_t = h_t^{-1} v_t^2 - 1$.⁴

It also can be proved that

$$\frac{\partial L}{\partial \alpha \partial \beta'} = \frac{\partial L}{\partial \beta \partial \alpha'} = \mathbf{0}$$

⁴The result is obtained as follows:

$$\begin{aligned} \frac{\partial L}{\partial \alpha} &= \frac{\partial \left(-\frac{1}{2} \sum_{t=1}^T \log h_t \right)}{\partial \alpha} + \frac{\partial \left(-\frac{1}{2} \sum_{t=1}^T h_t^{-1} v_t^2 \right)}{\partial \alpha} = -\frac{1}{2} \sum_{t=1}^T h_t^{-1} z_t + \frac{1}{2} \sum_{t=1}^T h_t^{-2} z_t v_t^2 = \frac{1}{2} \sum_{t=1}^T h_t^{-1} z_t f_t, \\ \frac{\partial L}{\partial \alpha \partial \alpha'} &= \frac{\partial \left(-\frac{1}{2} \sum_{t=1}^T h_t^{-1} z_t \right)}{\partial \alpha'} f_t + \left(-\frac{1}{2} \sum_{t=1}^T h_t^{-1} z_t \right) \frac{\partial (f_t)}{\partial \alpha'} = \frac{1}{2} \sum_{t=1}^T h_t^{-2} z_t z_t' f_t + \frac{1}{2} \sum_{t=1}^T h_t^{-1} z_t \left(h_t^{-2} v_t^2 z_t' \right) \\ &= \frac{1}{2} \sum_{t=1}^T h_t^{-2} z_t z_t' \left(f_t + h_t^{-1} v_t^2 \right) = \frac{1}{2} \sum_{t=1}^T h_t^{-2} z_t z_t'. \end{aligned}$$

where $\beta' = [\mu \ \phi_1 \ \dots \ \phi_p]'$ is the coefficient vector of the AR model.

So, referring to section 3.1.1, under the null $\alpha_1 = \alpha_2 = \dots = \alpha_m = 0$, the LM statistic for an ARCH(p, m) model is

$$\text{ARCHT}_{(p,m)} = \frac{1}{2} \left(\sum_{t=1}^T \hat{z}_t \hat{f}_t \right)' \left(\sum_{t=1}^T \hat{z}_t \hat{z}_t' \right)^{-1} \left(\sum_{t=1}^T \hat{z}_t \hat{f}_t \right)$$

with $\hat{z}_t = [1 \ \hat{v}_{t-1}^2 \ \dots \ \hat{v}_{t-m}^2]'$, $\hat{f}_t = \hat{h}_t^{-1} \hat{v}_t^2 - 1 = \left(T^{-1} \sum_{i=1}^T \hat{v}_i^2 \right)^{-1} \hat{v}_t^2 - 1$, $\hat{v}_{t-i}^2 = \hat{\varepsilon}_{t-i}^2$, $\forall i$ under the null restrictions. Asymptotically, the ARCHT(m) (LM test statistic for ARCH(p, m) model) follows a χ_m^2 distribution under the null hypothesis.

3.3 Power Comparison of LM Tests using Monte Carlo Method

In this section, Monte Carlo simulations are used to compare the power of various LM tests for different nonlinear time series data. The null model is restricted to be AR(1). The alternative models will be the bilinear model, the exponential autoregressive model, the threshold model, the smooth transition autoregressive model and the autoregressive conditional heteroskedastic model, which are constructed according to (3.12); (3.14); (3.16); (3.17) and (3.18); and (3.22) and (3.23), respectively. Moreover, the BL, EAR, TAR, STAR models are discussed for the lag parameter $d = 1$, and the ARCH model is limited to cases when $m = 1$.

As for the test statistics, due to the identification condition $k \leq p+1$ of the BLT(m, k) statistic, the BLT(1,1) and the BLT(2,2) statistics will be used; the EART(p) statistic is reduced to the EART(1) statistic; the ESTART(p) and LSTART(p) statistics will be the same for $p = 1$, which we denote as START(1); and the ARCHT(m) statistic becomes ARCHT(1). Therefore, there will be five LM test statistics used in the chapter. The power properties of the five statistics will be investigated by using the asymptotic critical values (ACV). A 5% significance level is used throughout the discussion.

The simulations are performed using the computer language FORTRAN. Subroutines and functions of the NAG FORTRAN library are used extensively in the programming. The number of replications for the Monte Carlo simulation is 10,000. The sample sizes $T = 50, 100, 200$ are used and represent typical lengths of economic time series.

Table 3.1: Empirical powers of LM tests against BL(1,1,1) using ACV

(i). $\mu=2, \phi=-0.9, c=-0.1, \sigma^2=1$			
Statistics	T=50	T=100	T=200
BLT(1,1)	44.3	73.1	94.7
BLT(2,2)	27.3	53.0	84.3
EART(1)	12.9	8.8	9.4
START(1)	22.1	26.1	39.9
ARCHT(1)	7.9	13.8	24.0
(ii). $\mu=1, \phi=0.3, c=-0.2, \sigma^2=1$			
Statistics	T=50	T=100	T=200
BLT(1,1)	39.1	69.9	95.2
BLT(2,2)	23.8	50.3	85.2
EART(1)	36.1	63.0	91.1
START(1)	36.0	61.0	91.1
ARCHT(1)	8.1	18.4	36.4
(iii). $\mu=0, \phi=0.6, c=0.7, \sigma^2=1$			
Statistics	T=50	T=100	T=200
BLT(1,1)	37.0	44.5	50.7
BLT(2,2)	63.3	75.1	83.9
EART(1)	40.7	51.0	62.5
START(1)	65.2	73.8	81.9
ARCHT(1)	60.9	88.7	98.9

- Note: 1) The empirical powers are in %.
 2) The significance level is 5%.
 3) 10,000 independent realizations of the BL(1,1,1) model as equation (3.12).

3.3.1 Simulating BL Models

Three different BL(1, 1, 1) models are specified according to equation (3.12), differing according to the parameter values of μ , ϕ and c . The test statistics for the three series are listed in Table 3.1. For case (i), the empirical power of the EART(1) statistic is smaller when $T = 100$ or $T = 200$ than when $T = 50$, which is quite abnormal since usually the empirical power of a test will increase with sample size. This could simply be due to approximation errors inherent in simulation-based measurements. For c values near zero (case (i) and case (ii)), the BLT tests have higher rejection frequencies than the other tests and the BLT(1, 1) has higher powers than the BLT(2, 2). However, in case (iii) when c is relatively large, the BLT(1, 1) test is the worst, the BLT(2, 2) test is better than BLT(1, 1) and the EART(1). START(1) and ARCHT(1) tests have high empirical powers against a BL(1, 1, 1) sequence.

Guegan and Pham [54] show that

- * For a BL(1, 1, 1) model, the theoretical power of BLT(1, 1) is a function of ϕ , c and σ^2 .
- * The BLT(1, 1) test is a local test in the sense that it is more efficient for testing a BL(1, 1, 1)

alternative near the null hypothesis than far from the null hypothesis.

We reproduce Guegan and Pham's empirical power study of the BLT(1, 1) test against BL(1, 1, 1) with the coefficients $\mu = 0$, $\sigma^2 = 1$ and $T = 200$. The number of replications is increased from 500 to 10,000. The parameter ϕ takes the values $-0.9, -0.6, -0.3, 0.3, 0.6, 0.9$, and the parameter c varies from -0.9 to 0.9 with a step of 0.1 . Also we extend the empirical power study from the BLT(1, 1) to include the BLT(2, 2), EART(1), START(1) and ARCHT(1) tests against BL(1, 1, 1). Table 3.2 tabulates and Figures 3.1, 3.2, 3.3, 3.4, and 3.5 depict the power curves for the five LM tests. We can see that:

- * The two BLT tests reach their highest power at c around ± 0.2 . For c from 0 to ± 0.2 , the power of BLT(1, 1) increases faster than the power of BLT(2, 2). For c larger than ± 0.2 , the power of BLT(1, 1) has lower level and bigger variance (within-curve and between-curve) than that of BLT(2, 2).

- * The power of the ARCHT(1) test is increasing as the absolute value of c gets larger and finally the power of ARCHT(1) test is larger than that of BLT tests, although for smaller values of c , the performance of the ARCHT(1) test is worse than that of the BLT tests.

- * The shape of the empirical power curves of START(1) is quite similar to that of BLT(1, 1) overall. However, the autoregressive coefficient ϕ has an effect on the shape of the graph. When ϕ equals to ± 0.9 , the power is always increasing as the absolute value of c increases, while when ϕ takes other values, the empirical powers reach their maximum at around $c = \pm 0.3$.

Table 3.2: Empirical power curves of LM tests against BL(1,1,1)

$\phi \setminus c$	-0.9	-0.8	-0.7	-0.6	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
BLT(1,1)																			
-0.9	78.7	76.8	76.5	77.4	79.7	85.8	93.8	99.1	87.5	5.0	87.9	99.2	94.2	85.4	80.6	77.9	77.1	77.4	77.9
-0.6	70.5	74.5	81.6	90.5	96.8	99.4	99.9	97.6	57.7	4.8	56.6	97.9	99.9	99.2	96.4	90.6	82.1	74.4	71.5
-0.3	66.4	74.2	85.8	95.2	99.0	99.9	99.9	95.9	49.3	5.0	49.4	96.0	99.9	99.9	99.0	95.4	85.8	74.6	65.9
0.3	50.9	52.3	61.6	77.9	91.6	98.6	99.8	95.7	49.5	4.6	48.9	95.8	99.8	98.7	92.0	77.1	61.8	52.6	51.0
0.6	55.5	52.9	51.8	55.3	69.8	90.0	98.7	98.1	56.0	4.8	57.0	97.8	98.8	89.3	69.6	55.5	52.7	53.3	56.9
0.9	65.5	65.6	64.4	61.0	56.4	54.4	68.4	94.5	88.3	4.5	88.5	94.1	69.1	55.4	56.3	62.1	65.0	66.4	66.5
BLT(2,2)																			
-0.9	94.3	92.7	91.6	90.6	91.4	93.4	96.5	98.4	71.5	5.0	71.6	98.5	96.7	92.9	90.9	91.0	91.7	92.8	94.5
-0.6	90.8	92.3	94.9	97.8	99.4	99.9	99.7	91.6	37.1	4.5	35.9	91.4	99.6	99.8	99.2	97.7	94.7	92.2	91.6
-0.3	90.0	93.4	97.1	99.5	99.9	100.0	99.4	86.7	29.2	4.8	30.3	86.8	99.5	100.0	99.9	99.4	97.1	93.3	89.8
0.3	84.4	85.5	89.3	95.0	98.8	99.9	99.6	87.2	29.7	4.6	28.9	87.9	99.7	99.9	98.7	94.8	89.3	85.6	84.6
0.6	87.8	86.3	85.4	85.9	90.1	97.1	99.6	93.8	35.4	4.7	35.9	93.1	99.6	96.8	91.0	86.2	85.1	86.4	87.8
0.9	94.0	92.6	91.8	89.7	87.6	83.8	87.9	96.6	75.4	4.5	76.4	96.9	87.9	85.2	86.8	90.4	91.9	93.0	93.1
EART(1)																			
-0.9	85.8	85.2	85.0	82.3	78.6	69.9	54.6	32.9	12.2	4.8	11.4	32.9	54.0	69.3	79.0	83.1	85.1	86.1	86.0
-0.6	62.9	60.3	58.0	57.0	57.2	55.1	46.4	28.1	9.4	4.7	9.8	28.6	46.2	55.5	57.2	57.4	58.7	60.2	64.8
-0.3	49.8	48.1	51.3	58.3	64.3	64.9	51.6	27.2	8.9	4.7	8.9	28.0	51.6	64.7	65.2	58.6	51.2	47.3	49.1
0.3	54.6	48.6	40.3	33.3	31.5	40.7	44.3	29.3	10.1	4.7	9.9	28.9	44.2	40.7	31.7	33.0	39.6	49.0	53.6
0.6	71.5	69.0	62.7	51.2	34.1	18.2	18.5	20.8	8.1	4.6	8.8	20.2	18.5	18.2	34.2	51.1	62.0	67.9	71.2
0.9	86.8	88.0	87.4	86.8	83.8	74.9	50.5	11.8	5.2	4.1	4.7	12.2	49.8	75.2	84.5	87.3	88.0	87.6	86.3
ESTAR(1)																			
-0.9	92.6	91.7	90.5	88.9	86.3	82.1	76.6	71.6	39.0	4.7	37.8	70.8	76.5	81.1	86.0	88.8	91.0	92.1	92.8
-0.6	81.9	81.7	83.3	87.6	92.5	96.5	97.5	88.1	38.2	4.9	37.8	88.2	97.6	96.0	92.1	87.2	83.5	81.3	83.5
-0.3	80.4	84.8	91.4	97.3	99.3	99.9	99.6	91.2	38.1	5.1	38.9	91.5	99.6	99.9	99.5	97.2	91.6	84.5	80.4
0.3	81.7	84.1	89.9	95.7	99.2	99.7	99.0	87.7	35.6	4.4	34.6	87.5	99.0	99.7	99.1	95.9	89.6	84.4	81.1
0.6	84.5	82.5	81.9	84.5	90.6	95.8	95.6	81.6	31.9	4.4	32.5	80.2	95.1	95.5	90.5	84.6	81.9	81.6	84.6
0.9	92.8	92.2	90.4	88.6	87.1	84.6	80.4	67.8	31.8	5.2	32.7	68.1	80.6	84.0	87.4	89.0	90.2	92.3	92.6
ARCHT(1)																			
-0.9	99.5	99.4	98.8	98.1	96.6	94.9	90.3	67.4	19.1	3.7	18.8	68.6	89.6	95.0	96.6	98.1	98.9	99.3	99.6
-0.6	99.7	99.7	99.6	99.5	98.7	95.0	78.9	42.4	10.5	4.1	10.0	41.4	78.4	94.7	98.8	99.5	99.6	99.7	99.8
-0.3	100.0	100.0	100.0	99.8	99.2	95.1	75.6	37.0	8.4	3.8	8.9	36.6	75.5	94.7	98.8	99.5	99.6	99.7	99.8
0.3	99.8	99.7	99.6	99.4	98.5	93.8	74.2	35.3	9.2	3.7	8.4	36.8	75.4	94.0	98.6	99.4	99.5	99.8	99.8
0.6	99.6	99.3	98.8	97.9	96.4	91.8	77.7	40.7	9.5	3.8	9.4	40.2	77.6	92.1	96.1	97.8	98.7	99.3	99.7
0.9	99.6	99.5	98.9	98.4	97.2	95.3	90.7	73.3	20.8	3.4	20.6	73.5	90.6	94.9	96.8	98.2	98.9	99.3	99.6

Note: 1) 10,000 independent realizations of the BL(1,1,1) model as equation (3.12) when $\mu = 0$, $\sigma^2 = 1$ and $T = 200$.
 2) For $c = 0$, the null model is true, so the empirical rejection percentage should be close to the test's nominal 5% level.

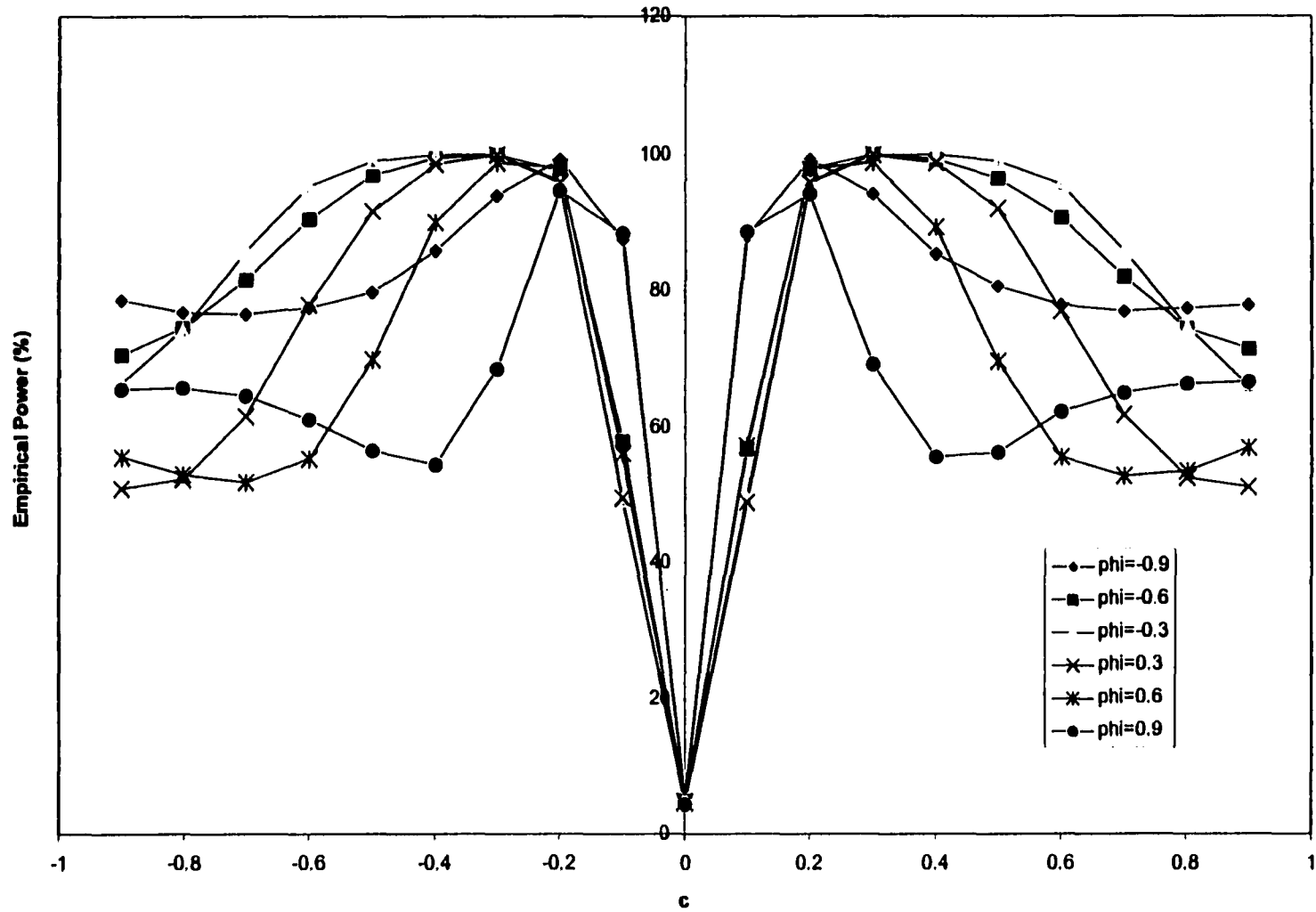


Figure 3.1: Empirical power curves of BLT(1,1) against BL(1,1,1)

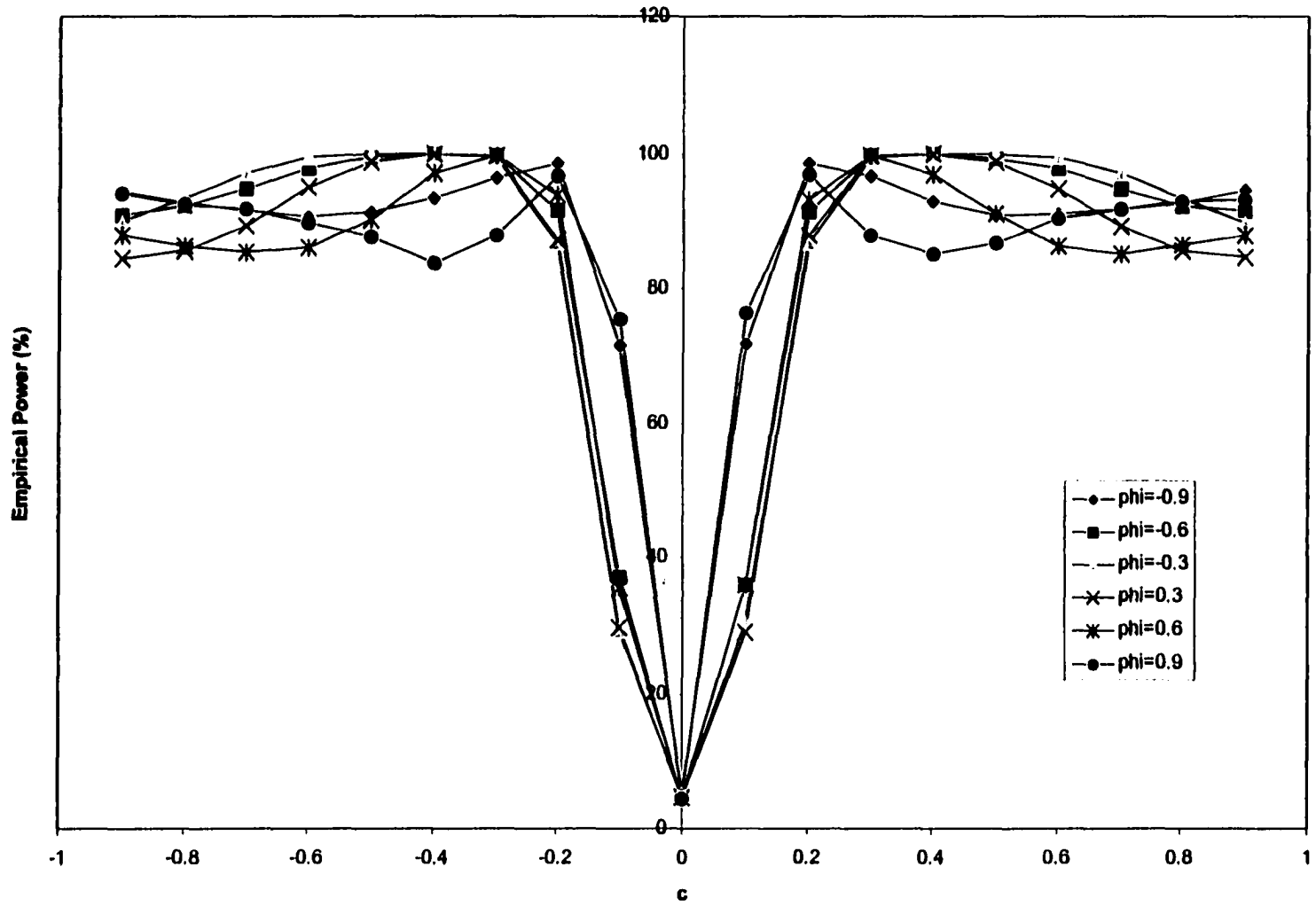


Figure 3.2: Empirical power curves of BLT(2,2) against BL(1,1,1)

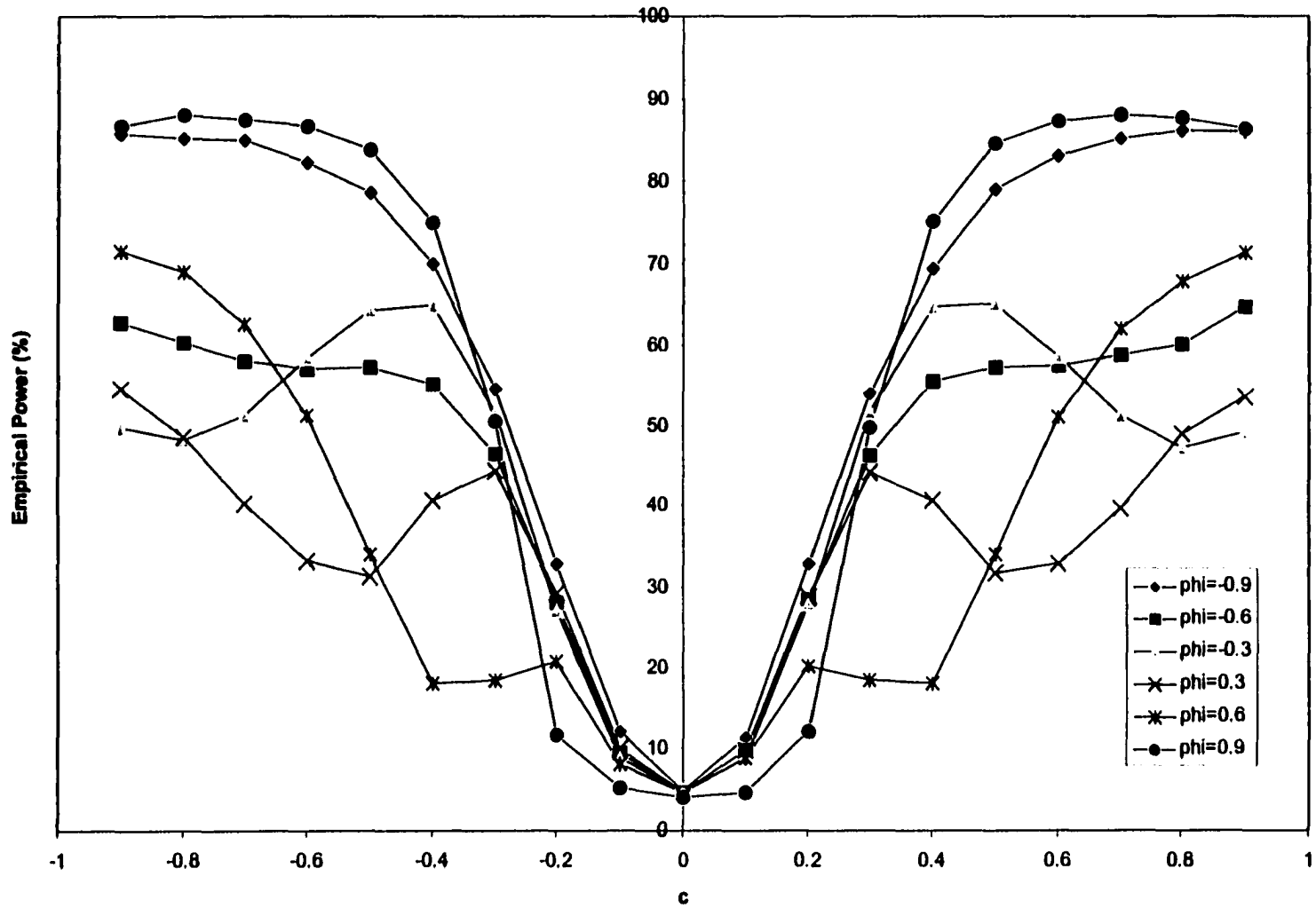


Figure 3.3: Empirical power curves of EART(1) against BL(1,1,1)

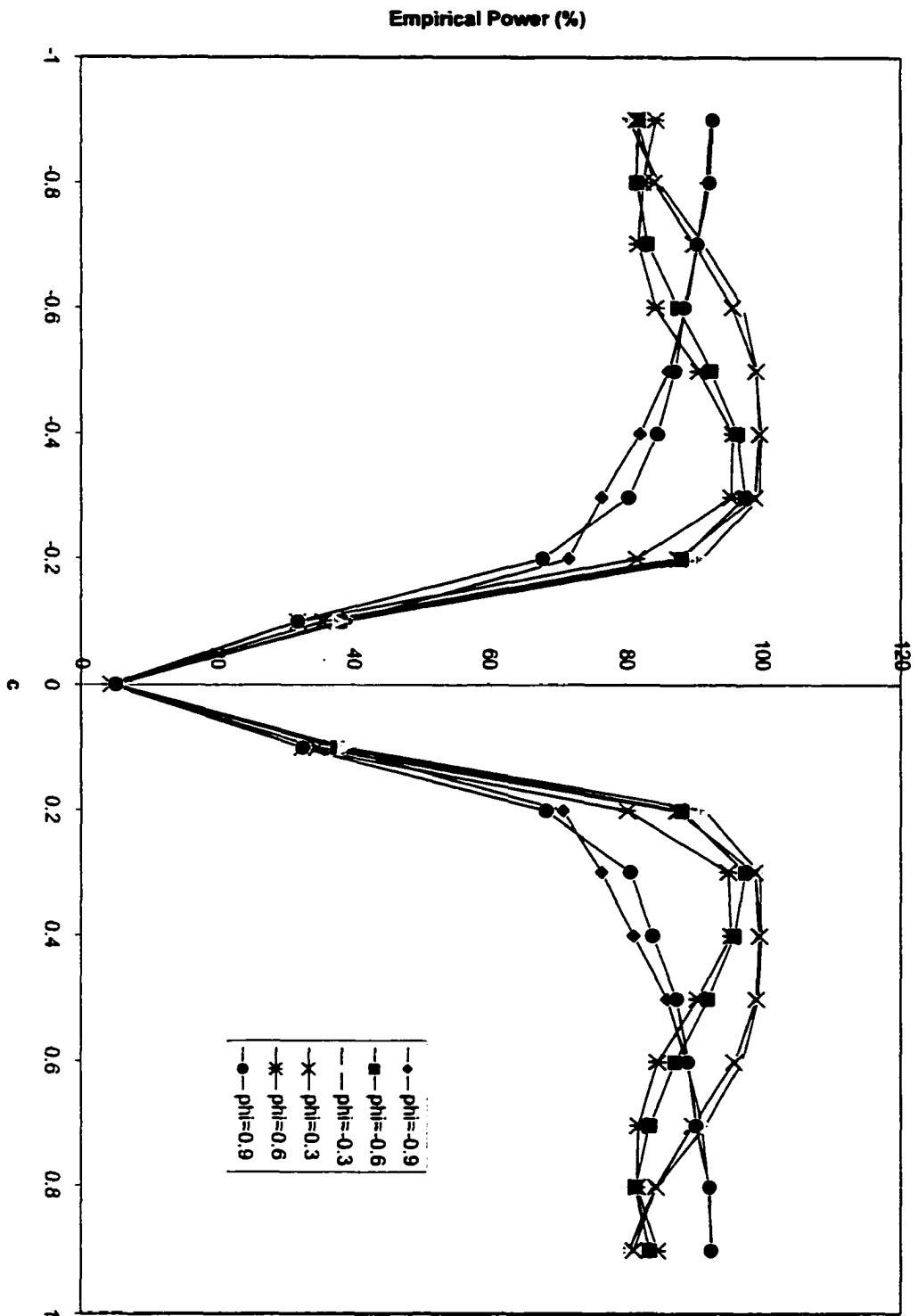


Figure 3.4: Empirical power curves of STARF(1) against BL(1,1,1)

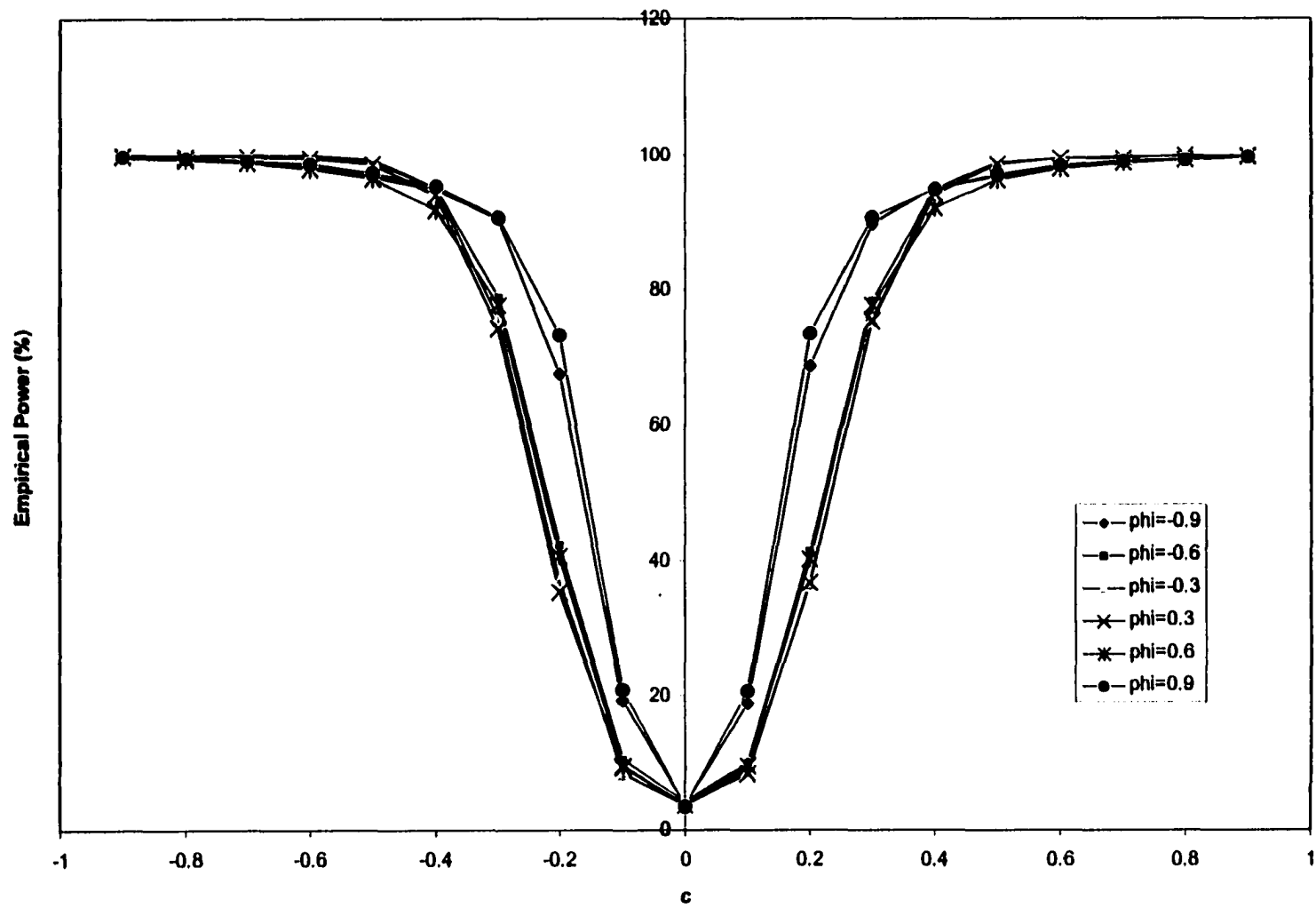


Figure 3.5: Empirical power curves of ARCH(1) against BL(1,1,1)

3.3.2 Simulating EAR Models

The test results for data generated by the EAR(1) model are tabulated in Table 3.3. From the table, we can see that the model coefficients do not have much effect on the test performance, while the value of σ^2 often makes a large difference. The five LM tests reject the null hypothesis more frequently when σ^2 is smaller. One possible explanation could be that the small variance makes it easy for a test to detect the nonlinearity from the model mean. However, the EART(1) almost always has higher power than the other tests, regardless of the value of σ^2 . The START(1) comes next, with empirical powers quite close to those of the EART(1). The BLT test statistics will have good power against the EAR(1) model under some situations. The ARCHT(1) seems to never work well for EAR models with empirical powers very close to 5%, the nominal significance level.

An attempt to plot the empirical power curves of the five LM tests against the EAR(1) model is also made here (see Table 3.4, and Figures 3.6, 3.7, 3.8, 3.9, and 3.10). However an LM test can be invalid for EAR model when $|\phi - \theta| \geq 1$. Also, the plots show that near the region that the LM test does not exist, there tend to be some irregular jumps in the power levels.

The simulated EAR model can be rewritten as

$$y_t = (\phi - \theta)y_{t-1} + \exp(-y_{t-1}^2) \theta y_{t-1} + \varepsilon_t. \quad (3.24)$$

Haggan and Ozaki's conditions for an EAR model like (3.24) to exhibit the limit cycle behavior is:

- (i). $|\phi - \theta| < 1$;
- (ii). $\phi > 1$;
- (iii). $\frac{1-\phi}{\theta} > 0$ or $\frac{1-\phi}{\theta} < -1$.

Since the model is simulated with the parameter ϕ lying within the unit circle, condition (ii) is not satisfied. And when $|\phi - \theta| \geq 1$, the condition (i) is violated. So if $|\phi - \theta| \geq 1$ the simulated model does not exhibit limit cycle. Moreover, for $|\phi - \theta| \geq 1$, though the model is stationary when y_{t-1} is small, it is explosive when y_{t-1} takes large values. So when $|\phi - \theta| \geq 1$ the simulated EAR follows neither a limit cycle nor limit point stationary, it actually diverges. That is probably the reason why there is no LM statistics when $|\phi - \theta| \geq 1$.

Table 3.3: Empirical powers of LM tests against EAR(1) using ACV

(i). $\mu=0, \phi=0.3, \gamma=1, \theta=0.9$				
Statistics	σ^2	T=50	T=100	T=200
BLT(1,1)	0.36	6.9	5.0	5.3
	1	5.4	4.5	4.4
BLT(2,2)	0.36	5.2	5.5	6.3
	1	4.2	4.4	4.7
EART(1)	0.36	26.7	52.0	82.9
	1	12.3	22.2	42.1
START(1)	0.36	22.4	43.5	75.5
	1	12.7	20.1	35.8
ARCHT(1)	0.36	2.7	6.1	11.1
	1	1.9	3.1	4.0
(ii). $\mu=0.3, \phi=0.3, \gamma=1, \theta=0.9$				
Statistics	σ^2	T=50	T=100	T=200
BLT(1,1)	0.36	15.0	23.7	38.9
	1	6.3	6.3	7.9
BLT(2,2)	0.36	9.1	13.8	23.8
	1	5.1	5.1	5.5
EART(1)	0.36	31.9	59.5	87.5
	1	12.9	23.0	41.7
START(1)	0.36	25.4	49.7	80.9
	1	12.3	19.4	35.2
ARCHT(1)	0.36	2.6	5.6	9.5
	1	1.9	2.9	3.8
(iii). $\mu=1, \phi=0.3, \gamma=1, \theta=0.9$				
Statistics	σ^2	T=50	T=100	T=200
BLT(1,1)	0.36	38.4	63.1	88.0
	1	11.6	16.3	26.0
BLT(2,2)	0.36	24.1	46.7	79.2
	1	8.4	9.8	15.8
EART(1)	0.36	42.2	65.7	90.1
	1	15.2	22.8	37.9
START(1)	0.36	42.4	66.0	90.6
	1	13.0	17.9	30.8
ARCHT(1)	0.36	2.5	4.1	6.1
	1	1.8	2.7	3.5

- Note: 1) The empirical powers are in %.
 2) The significance level is 5%.
 3) 10,000 independent realizations of the EAR(1) model in equation (3.14)

Table 3.4: Empirical power curves of LM tests against EAR(1)

$\phi\theta$	-0.9	-0.8	-0.7	-0.6	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
BLT(1,1)																			
-0.9	5.8	5.5	5.4	5.2	5.2	5.3	5.4	5.2	5.4	5.0	4.9	NA	NA	NA	NA	NA	NA	NA	NA
-0.6	6.7	6.4	5.9	5.4	5.3	5.3	5.2	5.1	5.1	5.2	5.0	4.9	4.9	4.8	NA	NA	NA	NA	NA
-0.3	8.2	7.0	6.2	5.7	5.5	5.5	5.4	5.2	5.0	5.2	5.3	5.2	5.0	4.7	4.5	4.4	4.8	NA	NA
0.3	NA	NA	8.3	5.7	5.6	5.4	5.1	5.0	4.9	4.8	5.0	5.3	5.3	5.2	5.3	5.4	5.3	5.2	4.8
0.6	NA	NA	NA	NA	NA	8.3	5.1	5.0	4.9	4.9	5.0	5.2	5.2	5.5	5.8	5.9	5.9	5.8	5.7
0.9	NA	NA	NA	NA	NA	NA	NA	NA	8.4	5.0	5.0	5.1	5.3	5.4	5.6	6.0	6.2	6.5	6.4
BLT(2,2)																			
-0.9	5.2	5.1	5.1	5.2	5.1	5.0	5.0	4.9	4.9	4.8	8.6	NA	NA	NA	NA	NA	NA	NA	NA
-0.6	5.7	5.2	4.9	4.9	4.9	4.8	4.9	4.7	4.6	4.7	4.7	4.6	4.7	8.6	NA	NA	NA	NA	NA
-0.3	6.7	6.0	5.3	5.0	5.0	4.9	4.9	4.8	4.7	4.6	4.5	4.4	4.4	4.4	4.5	4.6	8.5	NA	NA
0.3	NA	NA	10.9	5.4	5.2	5.1	4.8	4.6	4.5	4.5	4.5	4.7	4.8	4.7	4.8	4.6	4.5	4.4	4.3
0.6	NA	NA	NA	NA	NA	11.0	5.1	4.7	4.7	4.4	4.4	4.4	4.5	4.5	4.8	5.0	5.0	5.0	4.9
0.9	NA	NA	NA	NA	NA	NA	NA	NA	11.1	4.9	4.7	4.6	4.5	4.6	4.7	5.0	5.0	5.2	5.3
EART(1)																			
-0.9	52.7	43.7	34.6	26.9	19.4	14.0	9.9	7.1	5.1	4.4	43.0	NA	NA	NA	NA	NA	NA	NA	NA
-0.6	51.3	43.1	35.2	27.9	21.4	15.4	11.0	7.8	5.6	4.3	3.8	3.5	3.5	42.7	NA	NA	NA	NA	NA
-0.3	42.0	37.6	31.6	25.2	19.9	15.0	10.7	7.6	5.8	4.9	4.8	5.4	6.8	7.6	7.4	5.3	42.4	NA	NA
0.3	NA	NA	49.0	6.0	7.1	7.6	6.7	5.5	4.6	4.7	5.7	7.8	11.0	15.0	20.1	26.7	32.8	38.6	42.2
0.6	NA	NA	NA	NA	NA	49.6	4.0	3.6	3.7	4.0	5.1	7.5	10.9	15.1	21.2	28.5	36.4	44.5	52.6
0.9	NA	NA	NA	NA	NA	NA	NA	NA	50.4	4.6	5.0	6.6	9.7	14.3	20.4	27.4	35.7	44.8	53.6
ESTAR(1)																			
-0.9	43.4	34.9	27.4	20.6	15.4	11.3	8.2	6.2	5.0	4.4	47.4	NA	NA	NA	NA	NA	NA	NA	NA
-0.6	41.0	34.4	28.3	22.0	16.6	12.4	9.2	7.0	5.5	4.9	4.5	4.2	4.1	46.7	NA	NA	NA	NA	NA
-0.3	30.3	27.4	23.2	19.3	15.3	11.6	8.9	7.0	5.8	5.1	5.3	6.3	7.1	7.8	7.2	5.2	46.3	NA	NA
0.3	NA	NA	74.1	6.5	5.6	5.5	4.9	4.4	4.4	4.7	5.5	6.8	8.9	12.3	16.7	21.3	26.7	32.2	36.1
0.6	NA	NA	NA	NA	NA	74.7	5.5	3.8	3.6	4.0	4.9	6.6	9.2	12.6	16.7	22.0	28.9	36.1	43.5
0.9	NA	NA	NA	NA	NA	NA	NA	NA	75.2	5.8	5.3	6.5	8.6	11.6	16.1	21.3	28.0	35.7	44.0
ARCHT(1)																			
-0.9	4.3	4.1	4.0	4.0	3.9	3.9	3.8	3.9	4.0	4.1	4.2	NA	NA	NA	NA	NA	NA	NA	NA
-0.6	4.4	4.3	4.1	3.9	3.8	3.8	3.6	3.7	3.7	3.8	3.8	3.9	4.1	4.2	NA	NA	NA	NA	NA
-0.3	4.1	4.2	4.2	4.0	4.0	3.9	3.8	3.8	3.7	3.8	3.8	3.8	3.8	3.9	3.9	4.0	4.3	NA	NA
0.3	NA	NA	4.4	3.7	3.6	3.8	3.7	3.8	3.8	3.8	3.9	4.0	3.9	4.0	4.1	4.1	4.1	4.1	4.1
0.6	NA	NA	NA	NA	NA	4.3	3.8	3.7	3.8	3.8	3.8	3.9	3.9	3.9	3.9	4.0	4.2	4.3	4.3
0.9	NA	NA	NA	NA	NA	NA	NA	NA	4.4	4.0	4.0	3.9	3.9	4.0	4.0	3.9	4.1	4.1	4.3

Note: 1) 10,000 independent realizations of the EAR(1) model in equation (3.14) when $\mu = 0$, $\gamma = d = \sigma^2 = 1$ and $T = 200$.
 2) For $c = 0$, the null model is true, so the empirical rejection percentage should be close to the test's nominal 5% level.

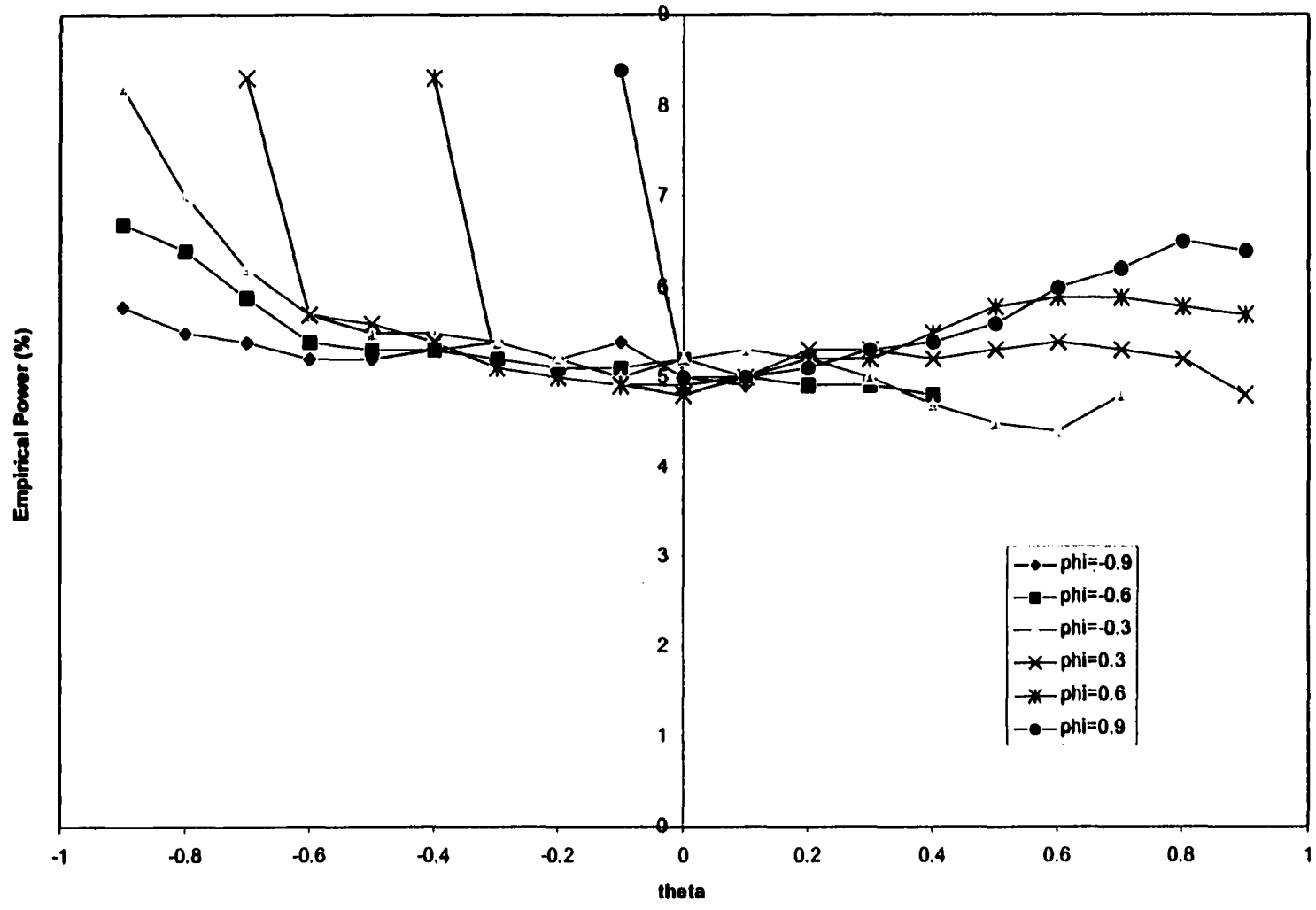


Figure 3.6: Empirical power curves of BLT(1,1) against EAR(1)

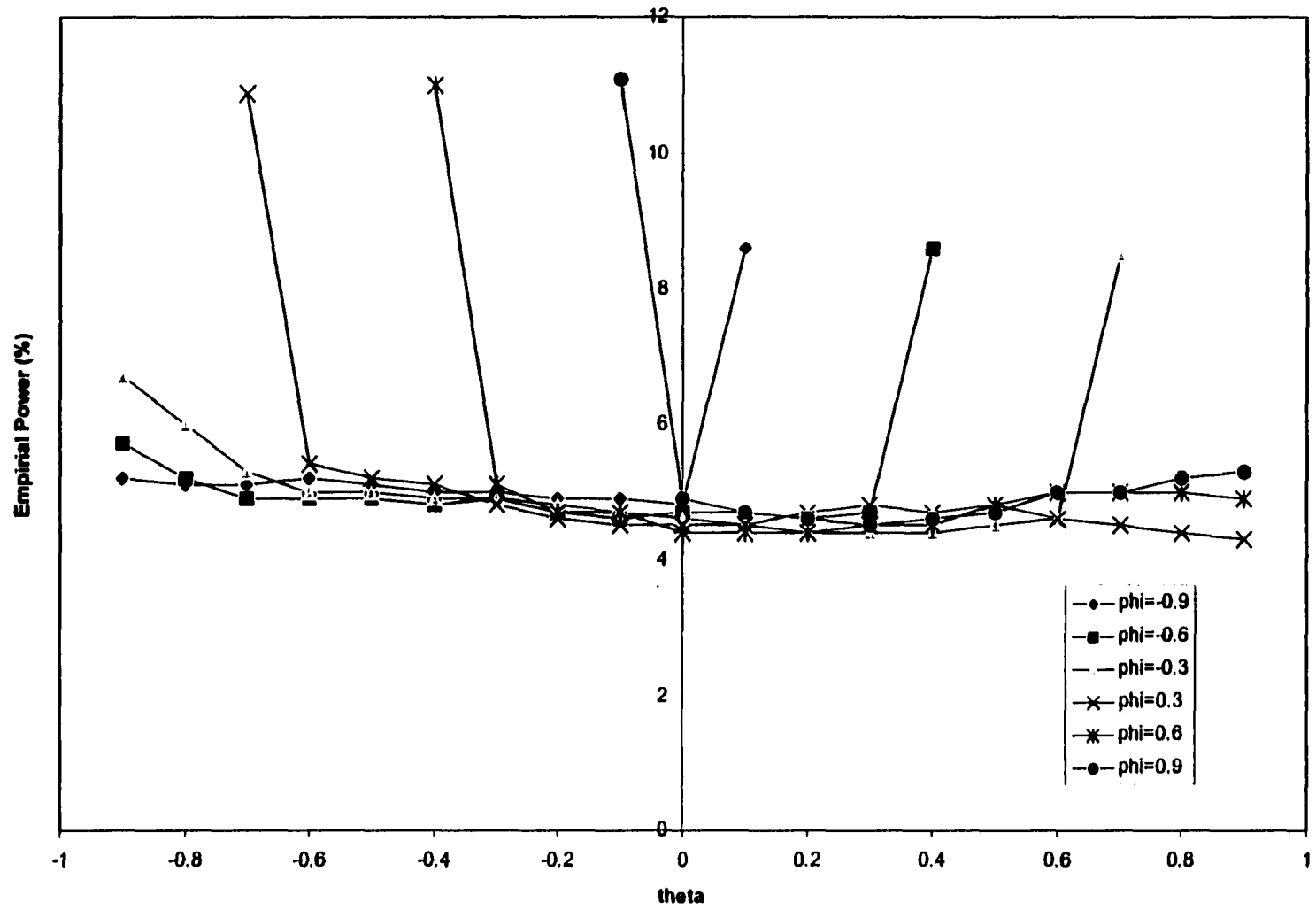


Figure 3.7: Empirical power curves of BLT(2,2) against EAR(1)

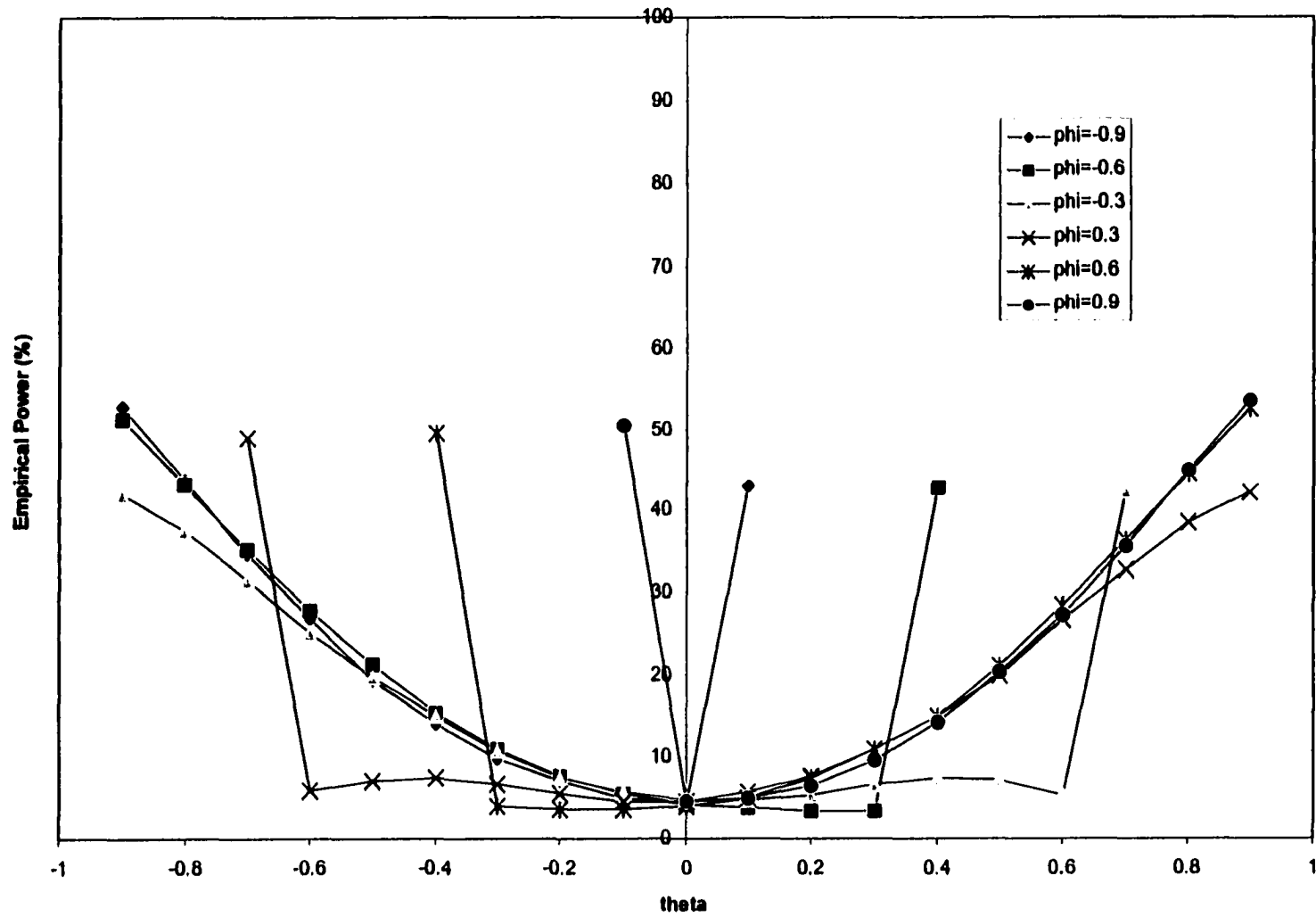


Figure 3.8: Empirical power curves of EART(1) against EAR(1)

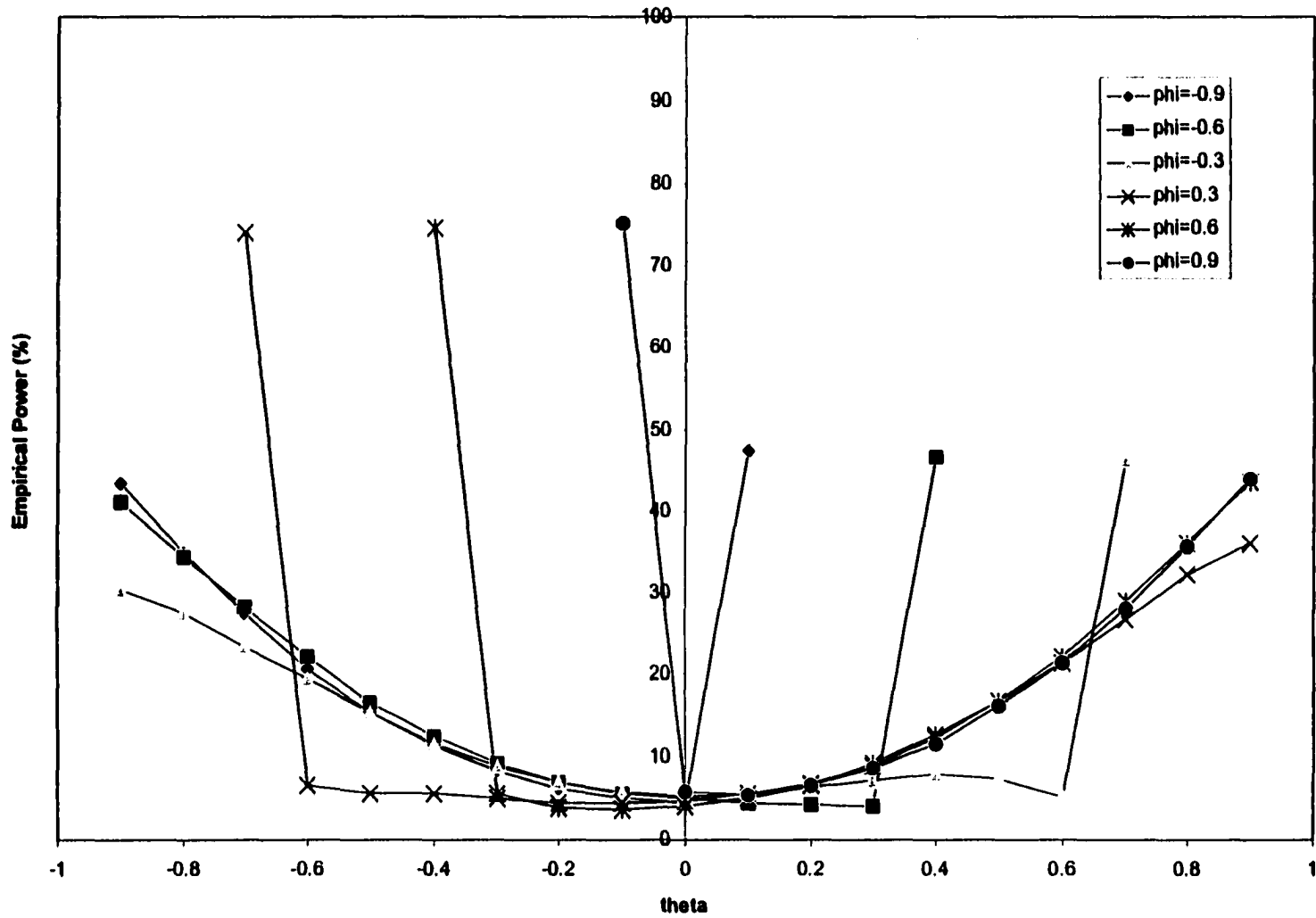


Figure 3.9: Empirical power curves of START(1) against EAR(1)

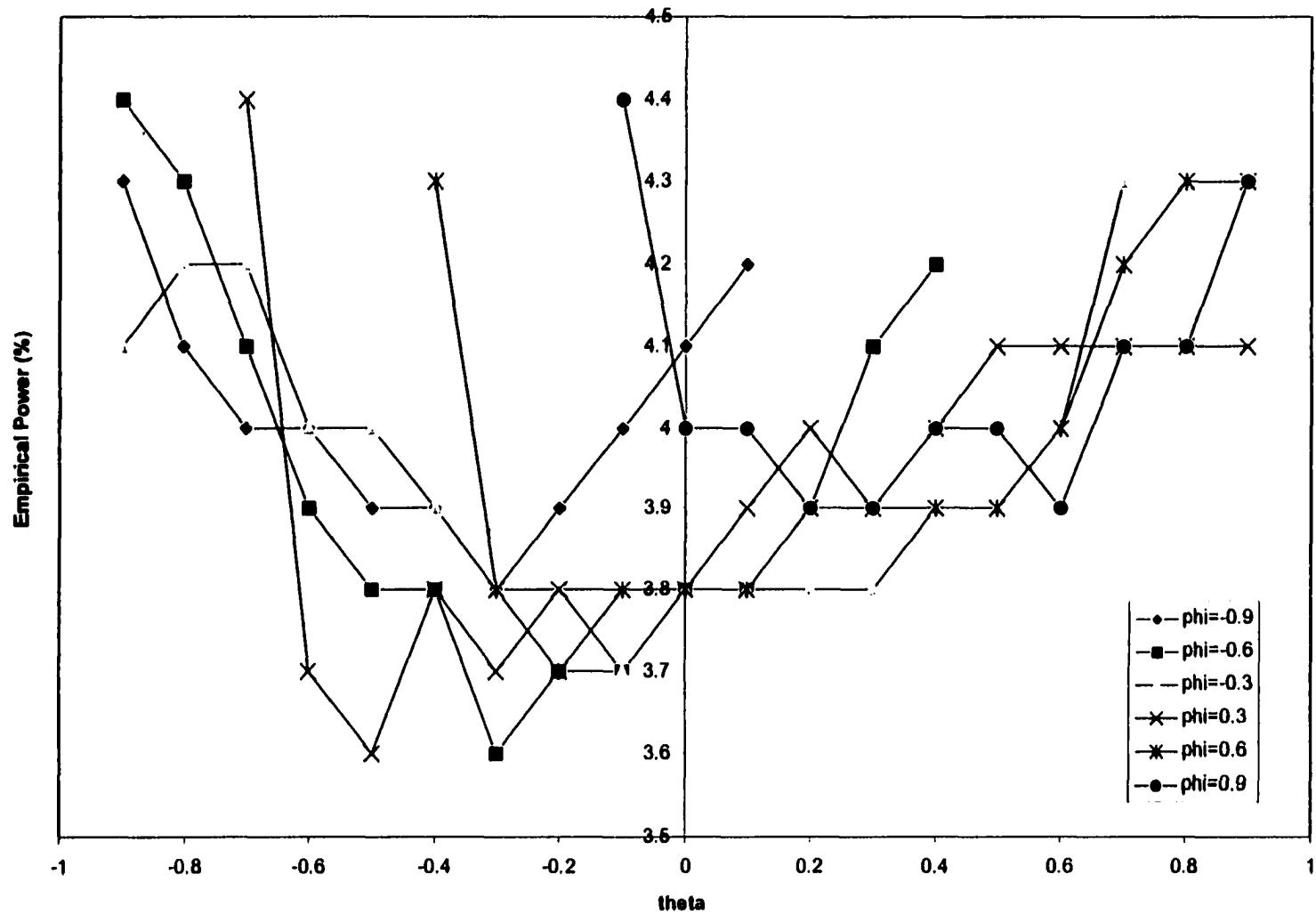


Figure 3.10: Empirical power curves of ARCHT(1) against EAR(1)

3.3.3 Simulating TAR Models

As mentioned before, the TAR model does not have a corresponding LM test. But a comparison among the five test statistics in Table 3.5 still shows some interesting features. First, the intercept plays a role in the test performance. In cases (i), (ii) and (iii), when the intercepts are set to zero, the BLT and the START statistics are better than the EART and ARCHT statistics; in cases (iv) and (v), when the intercepts are not equal to zero, the EART and the START statistics are better than the BLT statistics. The farther apart μ_1 and μ_2 are, the more frequently the null will be rejected by the tests. This can be clearly seen from the comparison of cases (iv) and (v). Second, the larger the difference between ϕ_1 and ϕ_2 , the more frequently the null will be rejected by the tests. We can see from Table 3.5 that the rejection percentages in (i) and (ii) are higher than that in (iii). Third, if the variances of the two regimes are different, the ARCHT test will have high power since it responds to heteroskedasticity of errors. This can be seen by comparing (v) and (vi). Last, note also that for all six cases, the START statistic performs about as well or better than the alternatives.

Empirical power curves are not provided for the TAR model since for the TAR model there are too many coefficients to be plotted in two-dimensional plots. For the same reason, empirical power curves are not illustrated for the STAR and ARCH models.

3.3.4 Simulating STAR Models

Here we discuss two types of STAR models, the ESTAR and the LSTAR models. The results are in Tables 3.6 and 3.7, respectively.

For the ESTAR model, we want to see the importance of the values of the intercepts μ_1 and μ_2 and the coefficient γ of the ESTAR model on the five tests' power performance. As γ increases, the powers of the tests decrease. When it reaches 10,000, there is almost no power for any of the five tests. This finding complies with the theory that ESTAR is approximately a linear model as $\gamma \rightarrow \infty$. As mentioned before the ESTAR model is reduced to an EAR model when $\mu_2 = 0$. We find in cases (i), (ii) and (iii), that the EART is better than the START and in cases (iv), (v), (vi), the START is better than the

Table 3.5: Empirical powers of LM tests against TAR(2;1,1) using ACV

(i). $\mu_1=\mu_2=0, \phi_1=0.3,$ $\phi_2=-0.6, c=0.2, \sigma_1^2=\sigma_2^2=0.49$				(ii). $\mu_1=\mu_2=0, \phi_1=-0.3,$ $\phi_2=0.6, c=0.2, \sigma_1^2=\sigma_2^2=0.49$			
Statistics	T=50	T=100	T=200	Statistics	T=50	T=100	T=200
BLT(1,1)	39.0	67.3	93.3	BLT(1,1)	32.7	62.5	90.0
BLT(2,2)	21.8	45.2	81.3	BLT(2,2)	17.2	41.2	76.4
EART(1)	11.3	15.5	24.2	EART(1)	9.0	15.4	27.5
START(1)	32.8	59.6	90.4	START(1)	29.4	59.4	90.0
ARCHT(1)	4.6	8.5	16.2	ARCHT(1)	4.3	8.4	15.2
(iii). $\mu_1=\mu_2=0, \phi_1=0.3,$ $\phi_2=0.6, c=0.2, \sigma_1^2=\sigma_2^2=0.49$				(iv). $\mu_1=1, \mu_2=0.5, \phi_1=0.3,$ $\phi_2=-0.3, c=0.2, \sigma_1^2=\sigma_2^2=0.49$			
Statistics	T=50	T=100	T=200	Statistics	T=50	T=100	T=200
BLT(1,1)	7.6	13.8	22.6	BLT(1,1)	6.3	6.6	8.9
BLT(2,2)	5.1	8.1	13.9	BLT(2,2)	6.3	6.2	7.6
EART(1)	5.5	5.3	5.6	EART(1)	9.0	9.2	11.5
START(1)	7.2	10.8	19.5	START(1)	42.4	24.8	17.5
ARCHT(1)	2.1	3.6	4.5	ARCHT(1)	2.7	4.1	5.5
(v). $\mu_1=1, \mu_2=-0.5, \phi_1=0.3,$ $\phi_2=-0.3, c=0.2, \sigma_1^2=\sigma_2^2=0.49$				(vi). $\mu_1=1, \mu_2=-0.5, \phi_1=0.3,$ $\phi_2=-0.3, c=0.2, \sigma_1^2=0.49, \sigma_2^2=0.04$			
Statistics	T=50	T=100	T=200	Statistics	T=50	T=100	T=200
BLT(1,1)	10.0	18.0	33.7	BLT(1,1)	10.9	17.3	29.2
BLT(2,2)	8.5	15.4	33.9	BLT(2,2)	14.6	22.4	36.9
EART(1)	19.8	46.4	82.1	EART(1)	15.0	18.1	26.0
START(1)	39.9	52.8	82.0	START(1)	46.8	43.9	50.1
ARCHT(1)	4.3	8.4	12.8	ARCHT(1)	7.0	18.7	39.6

- Note: 1) The empirical powers are in %.
 2) The significance level is 5%.
 3) 10,000 independent realizations of the TAR(2;1,1) model in equation (3.16).

Table 3.6: Empirical powers of LM tests against ESTAR(1) using ACV

(i). $\mu_1=0, \mu_2=0, \phi_1=-0.5,$ $\phi_2=0.6, \gamma=0.1$				(ii). $\mu_1=0, \mu_2=0, \phi_1=-0.5,$ $\phi_2=0.6, \gamma=1$			
Statistics	T=50	T=100	T=200	Statistics	T=50	T=100	T=200
BLT(1,1)	7.0	7.8	7.4	BLT(1,1)	5.4	5.9	5.7
BLT(2,2)	5.3	6.2	6.1	BLT(2,2)	4.4	4.7	4.9
EART(1)	11.4	18.0	31.2	EART(1)	11.1	17.0	27.6
START(1)	10.8	14.5	24.5	START(1)	9.1	12.8	21.6
ARCHT(1)	2.7	4.5	6.5	ARCHT(1)	1.8	3.1	4.0
(iii). $\mu_1=0, \mu_2=0, \phi_1=-0.5,$ $\phi_2=0.6, \gamma=10$				(iv). $\mu_1=0, \mu_2=0, \phi_1=-0.5,$ $\phi_2=0.6, \gamma=10,000$			
Statistics	T=50	T=100	T=200	Statistics	T=50	T=100	T=200
BLT(1,1)	5.1	5.3	5.0	BLT(1,1)	5.1	5.3	5.0
BLT(2,2)	4.4	4.5	4.8	BLT(2,2)	4.4	4.6	4.7
EART(1)	5.5	5.3	5.1	EART(1)	5.6	5.3	4.9
START(1)	5.3	4.9	5.1	START(1)	5.4	4.9	5.2
ARCHT(1)	1.8	3.2	3.8	ARCHT(1)	1.9	3.2	3.8
(v). $\mu_1=1, \mu_2=-1, \phi_1=-0.5,$ $\phi_2=0.6, \gamma=0.1$				(vi). $\mu_1=1, \mu_2=-1, \phi_1=-0.5,$ $\phi_2=0.6, \gamma=1$			
Statistics	T=50	T=100	T=200	Statistics	T=50	T=100	T=200
BLT(1,1)	7.5	8.2	7.9	BLT(1,1)	15.9	27.6	45.9
BLT(2,2)	6.7	6.6	6.5	BLT(2,2)	9.3	16.4	29.6
EART(1)	10.1	12.6	20.1	EART(1)	8.0	8.7	12.5
START(1)	11.9	17.0	28.4	START(1)	31.2	57.8	87.0
ARCHT(1)	2.9	4.8	7.0	ARCHT(1)	2.9	6.3	11.2
(vii). $\mu_1=1, \mu_2=-1, \phi_1=-0.5,$ $\phi_2=0.6, \gamma=10$				(viii). $\mu_1=1, \mu_2=-1, \phi_1=-0.5,$ $\phi_2=0.6, \gamma=10,000$			
Statistics	T=50	T=100	T=200	Statistics	T=50	T=100	T=200
BLT(1,1)	12.0	22.6	39.7	BLT(1,1)	5.3	5.3	5.0
BLT(2,2)	6.8	12.0	22.3	BLT(2,2)	4.2	4.5	4.8
EART(1)	5.0	4.4	4.5	EART(1)	5.5	5.2	4.8
START(1)	12.2	22.4	39.1	START(1)	5.3	4.8	5.1
ARCHT(1)	1.5	2.6	4.2	ARCHT(1)	1.9	3.1	3.8

- Note: 1) The empirical powers are in %.
 2) The significance level is 5%.
 3) 10,000 independent realizations of the ESTAR(1) model in equation (3.17) with $c=0, p=1, d=1$.

EART. The ARCHT performs poorly for the ESTAR model over the full range of coefficients, while the BLT tests have some power when the model is ESTAR model but not for the EAR model. The best candidates for testing the ESTAR model would be the START and EART.

For the LSTAR model, the ARCHT(1) is not good overall with regard to power. The BLT type tests are pretty good, however the START is the best. The effect of γ^* on the test power is contrary to that for the ESTAR model: the powers increase when γ^* goes up. This is because with the LSTAR structure, the model is approaching a TAR model if $\gamma^* \rightarrow \infty$, showing abrupt nonlinearity.

Table 3.7: Empirical powers of LM tests against LSTAR(1) using ACV

(i). $\mu_1=0, \mu_2=0, \phi_1=0.3,$ $\phi_2=0.6, \gamma^*=0.5$				(ii). $\mu_1=0, \mu_2=0, \phi_1=0.3,$ $\phi_2=0.6, \gamma^*=100$			
Statistics	T=50	T=100	T=200	Statistics	T=50	T=100	T=200
BLT(1,1)	9.5	14.9	28.3	BLT(1,1)	17.7	32.5	61.4
BLT(2,2)	5.8	8.5	16.2	BLT(2,2)	9.7	18.2	41.3
EART(1)	6.0	6.5	7.6	EART(1)	6.5	8.6	14.6
START(1)	7.8	11.7	21.7	START(1)	14.8	28.0	56.8
ARCHT(1)	2.2	3.4	5.3	ARCHT(1)	2.6	4.3	7.1

(iii). $\mu_1=1, \mu_2=-1, \phi_1=0.3,$ $\phi_2=0.6, \gamma^*=0.5$				(iv). $\mu_1=1, \mu_2=-1, \phi_1=0.3,$ $\phi_2=0.6, \gamma^*=100$			
Statistics	T=50	T=100	T=200	Statistics	T=50	T=100	T=200
BLT(1,1)	11.1	14.7	24.5	BLT(1,1)	28.5	47.4	77.5
BLT(2,2)	7.6	9.2	15.7	BLT(2,2)	17.2	32.8	65.8
EART(1)	9.4	12.3	21.7	EART(1)	18.8	32.4	59.8
START(1)	27.5	20.6	25.1	START(1)	48.2	62.0	87.3
ARCHT(1)	2.2	3.4	5.2	ARCHT(1)	4.6	9.1	18.0

- Note: 1) The empirical powers are in %.
 2) The significance level is 5%.
 3) 10,000 independent realizations of the LSTAR(1) model in equation (3.18) with $c^*=0, p=1, d=1$.

Table 3.8: Empirical powers of LM tests against ARCH(1,1) using ACV

(i). $\mu=0, \phi=0.4, \alpha_0=1, \alpha_1=0.8$				(ii). $\mu=3, \phi=0.6, \alpha_0=1, \alpha_1=0.8$			
Statistics	T=50	T=100	T=200	Statistics	T=50	T=100	T=200
BLT(1,1)	22.9	31.5	40.0	BLT(1,1)	20.8	25.3	33.7
BLT(2,2)	34.4	49.4	64.1	BLT(2,2)	38.7	49.6	61.7
EART(1)	21.7	28.0	37.8	EART(1)	36.1	30.2	34.7
START(1)	33.7	44.7	56.5	START(1)	85.0	77.2	70.6
ARCHT(1)	56.8	90.6	99.7	ARCHT(1)	57.5	90.8	99.7

- Note: 1) The empirical powers are in %.
 2) The significance level is 5%.
 3) 10,000 independent realizations of the ARCH(1,1) model in equation (3.22) and (3.23) with $\sigma^2=1$.

3.3.5 Simulating ARCH Models

By choosing different coefficient values for the model, we simulate several ARCH processes. The test results are included in Table 3.8. It seems that the choice of the linear coefficients does not affect the test power much. Breusch [13] provides a rationale for this; for a discussion, see Diebold and Pauly [35]. Among the five statistics, the ARCHT(1) consistently has the largest power, which is reasonable since the true data are generated according to the ARCH model.

3.4 Size and Power Comparison using Bootstrap Method

In this section, size and power comparison of the five LM tests will be performed for different nonlinear models using bootstrap methods. For comparison, the general linearity test BDS test is also included.

There are various methods of calculating BDS statistic using different computer languages or software. Dechert compiles his program for the DOS operating system, LeBaron writes his program in C, Kanzler [67] bring forward an algorithm programmed in MATLAB, which provides a fast and correctly sized estimate of the BDS statistic. Here, Kanzler's MATLAB program is adopted for calculating the BDS statistic. So for each Monte Carlo simulated series, five LM tests and one BDS test (with $\epsilon = 1.5\sigma$ and $m = 2$) are performed in MATLAB. Different from FORTRAN, which is a compiled language, MATLAB is an interpreted language, so it takes much longer time to run a job by MATLAB than by FORTRAN. In section 3.3, where the statistics are calculated by FORTRAN, the running time ranges from several seconds to around 30 minutes under a UNIX environment, where the server has a speed of 400 MHZ, 256 MB to 1 GB RAM. For example, running five LM tests for BL(1,1,1) with $T = 100$ takes only a few seconds by FORTRAN in UNIX system, but about 3 days by MATLAB under the same environment. For series with length longer than 100, it is even difficult to produce a result by MATLAB in UNIX. That's because the bottleneck of running MATLAB is the CPU time, not the RAM, or the hard drive. So under the UNIX system with time-sharing characteristic, the excessive CPU requirement by complicated MATLAB programs is hard to satisfy and the programs often end up being idle and eventually killed before producing results. We run the MATLAB program on a PC with 1.7 GHZ speed, 128 MB RDRAM. The situation improves a lot concerning running time. We are able to shorten the running time of LM tests for series with 100 observations from 3 days to 6 hours or so. For $T = 200$, the running time for five LM tests is around 12 hours, for one BDS test is about 23 hours, and for five LM and one BDS tests, it is approximately 24 hours. Comparing the time needed to run six tests and five tests, it seems that simply running five LM tests does not take too long, the majority of the 12 hours to run five LM tests is used by MATLAB to generate data and do the bootstrap. There is no doubt that

the BDS test is a time-consuming process.

Table 3.9 compares the empirical sizes of the five LM tests and the BDS test using asymptotic and bootstrap critical values (BCV) at the 5% significance level. The critical values are good if the empirical sizes are near 5%. It is obvious that for all the tests, the empirical sizes by the bootstrap critical value are always near 5%, while the empirical sizes by the asymptotic critical values significantly deviate from 5% for ARCH(1) and, when the autoregressive coefficient ϕ approaches the unit circle for START(1). All of the empirical sizes by asymptotic critical value for the BDS tests are bigger than 5% for $T = 25$. To get an overall impression of the size performance of the two methods (i.e., asymptotic vs. bootstrap), the proportion of the empirical rejection percentages falling between 4% and 6% are examined. When $T = 100$, 31 out of the 36 bootstrap rejection rates fall within the range [4%, 6%] with a minimum of 3.8% and a maximum of 6.2%. The corresponding rejection rates using asymptotic critical values range from 2.3% to 5.9% with a proportion of $\frac{24}{36}$ falling within [4%, 6%]. When $T = 50$, the proportions are $\frac{29}{36}$ and $\frac{18}{36}$, and the ranges are [1.9%, 6.6%] and [1.4%, 9.1%] for the bootstrap and asymptotic empirical sizes. When $T = 25$, the bootstrap empirical sizes range from 2.5% to 6.3%, among which $\frac{33}{36}$ are in the range [4%, 6%]. The sizes found from the asymptotic critical values are scattered between 1.0% and 16.9%, and only 12 out of 36 are in the range [4%, 6%]. Therefore, the bootstrap method appears to have better size properties than the asymptotic method.

Tables 3.10, 3.11, 3.12, 3.13, 3.14 and 3.15 provide power comparisons of the bootstrap method and the asymptotic method for the LM tests and the BDS test when the simulated series follows the structure of a BL(1,1), EAR(1), TAR(2;1,1), ESTAR(1), LSTAR(1) or ARCH(1,1) model. Two conclusions emerge from these tables. First, the bootstrap powers and the asymptotic powers are quite close to one another in most cases. Second, to the extent that there are differences, the bootstrap power are generally larger than the asymptotic powers.

3.5 Discussion

This chapter has investigated size and power properties of asymptotic and bootstrap versions of five LM tests — the BLT(1,1), BLT(2,2), EART(1), START(1) and ARCH(1) tests— and one general

Table 3.9: Empirical sizes of LM tests against AR(1) using ACV and BCV

(i). $\mu=0, T=100, \sigma^2=1$								
ϕ	Critical Values	-0.9	-0.6	-0.3	0.3	0.6	0.9	
BLT(1,1)	Asymptotic	5.9	5.7	5.1	4.7	5.3	3.8	
	Bootstrap	5.9	5.1	4.8	5.0	5.9	5.1	
BLT(2,2)	Asymptotic	4.6	5.3	4.5	4.5	4.1	3.1	
	Bootstrap	5.7	5.6	4.7	5.2	5.0	3.8	
EART(1)	Asymptotic	3.9	4.1	5.6	5.7	4.5	3.8	
	Bootstrap	4.8	5.2	5.9	6.1	5.7	5.5	
START(1)	Asymptotic	4.5	4.5	5.7	4.3	4.5	4.9	
	Bootstrap	4.8	4.9	5.9	5.2	5.4	3.6	
ARCHT(1)	Asymptotic	2.4	3.4	2.5	2.9	2.9	2.3	
	Bootstrap	4.4	5.2	5.1	5.5	5.3	4.5	
BDS	Asymptotic	3.8	4.2	4.3	4.5	4.5	2.5	
	Bootstrap	5.8	5.4	5.3	6.2	6.1	5.0	
(ii). $\mu=0, T=50, \sigma^2=1$								
ϕ	Critical Values	-0.9	-0.6	-0.3	0.3	0.6	0.9	
BLT(1,1)	Asymptotic	4.3	6.9	5.4	4.1	4.2	3.2	
	Bootstrap	4.4	6.3	5.3	4.9	5.0	3.5	
BLT(2,2)	Asymptotic	4.6	3.8	4.3	3.9	3.4	1.7	
	Bootstrap	6.5	4.8	5.1	5.8	5.3	2.3	
EART(1)	Asymptotic	2.8	3.6	5.2	4.3	4.2	4.2	
	Bootstrap	4.3	5.1	5.5	5.5	5.6	4.9	
START(1)	Asymptotic	4.3	4.8	4.8	3.1	3.3	9.1	
	Bootstrap	4.8	5.2	5.0	3.9	4.9	1.9	
ARCHT(1)	Asymptotic	1.6	2.2	3.1	1.4	1.9	2.2	
	Bootstrap	5.2	4.0	6.6	4.2	5.3	5.7	
BDS	Asymptotic	4.0	5.0	5.4	5.9	6.4	4.7	
	Bootstrap	4.3	4.2	4.7	5.6	5.1	5.2	
(iii). $\mu=0, T=25, \sigma^2=1$								
ϕ	Critical Values	-0.9	-0.6	-0.3	0.3	0.6	0.9	
BLT(1,1)	Asymptotic	4.1	6.7	5.1	3.7	3.5	4.7	
	Bootstrap	4.8	5.4	5.1	4.5	4.5	4.2	
BLT(2,2)	Asymptotic	4.6	4.0	3.9	3.5	2.9	4.6	
	Bootstrap	6.3	5.9	5.3	4.9	4.3	4.3	
EART(1)	Asymptotic	3.6	3.1	4.2	3.9	3.1	5.3	
	Bootstrap	4.1	3.5	5.1	4.2	4.7	4.3	
START(1)	Asymptotic	7.7	4.3	4.0	3.0	3.2	16.9	
	Bootstrap	4.7	4.2	4.5	4.6	4.0	2.5	
ARCHT(1)	Asymptotic	1.1	1.0	1.3	6.0	5.0	1.3	
	Bootstrap	4.3	4.9	5.0	4.6	4.5	5.4	
BDS	Asymptotic	8.9	12.0	10.8	10.6	11.2	9.1	
	Bootstrap	5.3	5.6	4.8	5.6	5.5	5.0	

- Note: 1) The empirical rejection percentages are in %.
 2) The significance level is 5%
 3) For each parameter setting, 1,000 Monte Carlo replications of AR(1) are generated and within each Monte Carlo replication, 1,000 AR(1) series are bootstrapped.

Table 3.10: Empirical powers of LM tests against BL(1,1,1) using ACV and BCV

(i). $\mu=2, \phi=-0.9, c=-0.1, \sigma^2=1$				
Statistics	Critical Values	T=50	T=100	T=200
BLT(1,1)	Asymptotic	39.2	67.4	93.3
	Bootstrap	41.8	69.4	94.2
BLT(2,2)	Asymptotic	19.8	47.3	82.4
	Bootstrap	23.7	51.3	84.8
EART(1)	Asymptotic	5.3	3.9	6.2
	Bootstrap	5.8	5.5	7.7
START(1)	Asymptotic	13.1	19.8	38.3
	Bootstrap	12.6	19.9	40.6
ARCHT(1)	Asymptotic	7.6	14.0	24.4
	Bootstrap	11.6	16.7	27.8
BDS	Asymptotic	9.5	11.3	14.7
	Bootstrap	10.5	16.2	26.1
(ii). $\mu=1, \phi=0.3, c=-0.2, \sigma^2=1$				
Statistics	Critical Values	T=50	T=100	T=200
BLT(1,1)	Asymptotic	39.0	70.5	95.7
	Bootstrap	40.3	71.6	96.1
BLT(2,2)	Asymptotic	22.4	49.4	86.2
	Bootstrap	24.3	51.5	86.7
EART(1)	Asymptotic	34.5	62.9	91.4
	Bootstrap	36.4	64.1	91.6
START(1)	Asymptotic	29.2	61.0	92.3
	Bootstrap	29.8	62.3	92.5
ARCHT(1)	Asymptotic	7.2	16.9	32.5
	Bootstrap	10.7	21.5	36.7
BDS	Asymptotic	18.1	23.9	36.9
	Bootstrap	16.1	26.9	43.4

- Note:
- 1) The empirical powers are in %.
 - 2) The significance level is 5%.
 - 3) For each parameter setting, 1,000 Monte Carlo replications of BL(1,1,1) are generated and within each Monte Carlo replication, 1,000 BL(1,1,1) series are bootstrapped.

Table 3.11: Empirical powers of LM tests against EAR(1) using ACV and BCV

(i). $\mu=0, \phi=0.3, \gamma=1, \theta=0.9, \sigma^2=0.36$				
Statistics	Critical Values	T=50	T=100	T=200
BLT(1,1)	Asymptotic	6.8	7.2	5.5
	Bootstrap	6.5	7.1	5.8
BLT(2,2)	Asymptotic	4.8	5.5	4.8
	Bootstrap	6.3	6.2	5.4
EART(1)	Asymptotic	25.0	54.0	84.0
	Bootstrap	27.6	55.7	84.2
START(1)	Asymptotic	21.3	45.3	77.4
	Bootstrap	22.2	45.6	77.3
ARCHT(1)	Asymptotic	3.1	7.0	11.5
	Bootstrap	6.3	10.2	12.9
BDS	Asymptotic	7.5	8.0	9.8
	Bootstrap	6.7	9.9	13.4
(ii). $\mu=0.3, \phi=0.3, \gamma=1, \theta=0.9, \sigma^2=0.36$				
Statistics	Critical Values	T=50	T=100	T=200
BLT(1,1)	Asymptotic	17.6	22.3	39.9
	Bootstrap	16.7	22.1	39.6
BLT(2,2)	Asymptotic	7.5	13.9	24.7
	Bootstrap	9.3	14.8	25.7
EART(1)	Asymptotic	30.8	57.7	88.3
	Bootstrap	31.6	58.4	88.7
START(1)	Asymptotic	24.4	48.3	81.4
	Bootstrap	24.9	49.0	81.6
ARCHT(1)	Asymptotic	2.3	5.9	10.2
	Bootstrap	6.2	8.3	12.2
BDS	Asymptotic	8.8	9.2	11.9
	Bootstrap	7.9	10.8	17.0
(iii). $\mu=1, \phi=0.3, \gamma=1, \theta=0.9, \sigma^2=0.36$				
Statistics	Critical Values	T=50	T=100	T=200
BLT(1,1)	Asymptotic	40.0	64.9	88.3
	Bootstrap	40.0	64.4	88.2
BLT(2,2)	Asymptotic	23.1	44.5	79.2
	Bootstrap	24.9	46.4	79.5
EART(1)	Asymptotic	42.5	68.6	91.3
	Bootstrap	41.0	67.9	91.1
START(1)	Asymptotic	42.8	67.4	92.6
	Bootstrap	39.9	67.6	92.3
ARCHT(1)	Asymptotic	2.9	4.9	6.1
	Bootstrap	6.4	7.1	7.6
BDS	Asymptotic	10.7	8.8	12.8
	Bootstrap	10.1	11.2	18.2

- Note:
- 1) The empirical powers are in %.
 - 2) The significance level is 5%.
 - 3) For each parameter setting, 1,000 Monte Carlo replications of EAR(1) are generated and within each Monte Carlo replication, 1,000 EAR(1) series are bootstrapped

Table 3.12: Empirical powers of LM tests against TAR(2;1,1) using ACV and BCV

(i). $\mu_1=\mu_2=0, \phi_1=0.3,$ $\phi_2=-0.6, c=0.2, \sigma_1^2=\sigma_2^2=0.49$					(ii). $\mu_1=\mu_2=0, \phi_1=-0.3,$ $\phi_2=0.6, c=0.2, \sigma_1^2=\sigma_2^2=0.49$				
Statistics	CV	50	100	200	Statistics	CV	50	100	200
BLT(1,1)	Asymptotic	39.6	67.6	93.1	BLT(1,1)	Asymptotic	32.3	62.6	91.6
	Bootstrap	39.7	68.3	93.4		Bootstrap	34.7	63.9	92.1
BLT(2,2)	Asymptotic	22.6	44.5	79.1	BLT(2,2)	Asymptotic	17.5	39.3	78.9
	Bootstrap	25.3	46.9	80.6		Bootstrap	21.0	41.8	80.1
EART(1)	Asymptotic	10.8	16.2	22.9	EART(1)	Asymptotic	7.1	13.0	31.1
	Bootstrap	11.7	17.3	23.7		Bootstrap	8.3	15.6	32.7
START(1)	Asymptotic	32.4	62.0	89.4	START(1)	Asymptotic	26.9	58.3	91.2
	Bootstrap	34.2	62.8	90.3		Bootstrap	30.4	62.2	92.0
ARCHT(1)	Asymptotic	5.4	8.9	17.5	ARCHT(1)	Asymptotic	3.4	8.5	12.7
	Bootstrap	8.5	11.9	19.4		Bootstrap	6.9	11.9	15.5
BDS	Asymptotic	14.2	19.1	22.9	BDS	Asymptotic	12.3	13.1	20.6
	Bootstrap	13.0	21.7	29.1		Bootstrap	10.8	14.9	27.8
(iii). $\mu_1=-1, \mu_2=0.5, \phi_1=0.3,$ $\phi_2=-0.3, c=0.2, \sigma_1^2=\sigma_2^2=0.49$					(iv). $\mu_1=1, \mu_2=-0.5, \phi_1=0.3,$ $\phi_2=-0.3, c=0.2, \sigma_1^2=\sigma_2^2=0.49$				
Statistics	CV	50	100	200	Statistics	CV	50	100	200
BLT(1,1)	Asymptotic	12.0	19.6	30.6	BLT(1,1)	Asymptotic	9.9	18.8	34.7
	Bootstrap	11.5	18.8	30.8		Bootstrap	10.6	19.2	35.4
BLT(2,2)	Asymptotic	9.3	16.8	27.4	BLT(2,2)	Asymptotic	5.5	15.5	32.2
	Bootstrap	11.3	18.4	28.4		Bootstrap	6.9	17.6	32.8
EART(1)	Asymptotic	59.9	88.9	99.3	EART(1)	Asymptotic	17.0	44.8	81.5
	Bootstrap	64.8	90.8	99.4		Bootstrap	21.5	49.1	84.0
START(1)	Asymptotic	51.9	85.0	99.2	START(1)	Asymptotic	22.1	46.1	79.9
	Bootstrap	54.2	86.3	99.3		Bootstrap	15.5	43.6	80.7
ARCHT(1)	Asymptotic	3.4	4.9	8.2	ARCHT(1)	Asymptotic	4.4	8.9	12.7
	Bootstrap	7.1	7.8	9.8		Bootstrap	8.5	11.4	14.3
BDS	Asymptotic	5.4	5.4	3.7	BDS	Asymptotic	8.9	9.2	11.4
	Bootstrap	4.6	6.7	6.0		Bootstrap	8.7	11.4	18.3

- Note: 1) The empirical powers are in %.
2) The significance level is 5%.
3) For each parameter setting, 1,000 Monte Carlo replications of TAR(2;1,1) are generated and within each Monte Carlo replication, 1,000 TAR(2;1,1) series are bootstrapped.

Table 3.13: Empirical powers of LM tests against ESTAR(1) using ACV and BCV

(i). $\mu_1=0, \mu_2=0, \phi_1=-0.5, \phi_2=0.6, \gamma=50$				
Statistics	Critical Values	T=50	T=100	T=200
BLT(1,1)	Asymptotic	4.4	4.5	4.9
	Bootstrap	4.7	5.2	5.4
BLT(2,2)	Asymptotic	4.3	4.7	4.1
	Bootstrap	5.8	5.8	4.3
EART(1)	Asymptotic	5.2	4.9	5.7
	Bootstrap	5.8	5.1	6.1
START(1)	Asymptotic	4.3	5.0	4.8
	Bootstrap	5.0	6.1	5.2
ARCHT(1)	Asymptotic	1.0	2.2	4.2
	Bootstrap	3.7	5.5	6.0
BDS	Asymptotic	5.7	4.5	3.4
	Bootstrap	5.1	5.9	5.8
(ii). $\mu_1=1, \mu_2=-1, \phi_1=-0.5, \phi_2=0.6, \gamma=50$				
Statistics	Critical Values	T=50	T=100	T=200
BLT(1,1)	Asymptotic	6.0	8.2	13.0
	Bootstrap	6.9	8.4	13.9
BLT(2,2)	Asymptotic	5.2	4.1	7.4
	Bootstrap	6.1	4.4	7.9
EART(1)	Asymptotic	4.0	4.5	4.6
	Bootstrap	4.9	4.8	4.8
START(1)	Asymptotic	5.6	6.8	11.1
	Bootstrap	6.9	7.8	12.1
ARCHT(1)	Asymptotic	3.4	3.3	5.2
	Bootstrap	4.2	6.4	7.4
BDS	Asymptotic	4.7	2.4	1.6
	Bootstrap	4.0	4.2	2.6

- Note: 1) The empirical powers are in %.
 2) The significance level is 5%.
 3) For each parameter setting, 1,000 Monte Carlo replications of ESTAR(1) are generated and within each Monte Carlo replication, 1,000 ESTAR(1) series are bootstrapped.

Table 3.14: Empirical powers of LM tests against LSTAR(1) using ACV and BCV

(i). $\mu_1=1, \mu_2=-1, \phi_1=-0.5, \phi_2=0.6, \gamma^*=50$				
Statistics	Critical Values	T=50	T=100	T=200
BLT(1,1)	Asymptotic	23.4	41.1	69.8
	Bootstrap	23.3	41.8	69.9
BLT(2,2)	Asymptotic	13.9	27.1	50.5
	Bootstrap	15.9	28.2	52.4
EART(1)	Asymptotic	7.7	9.4	13.7
	Bootstrap	8.5	11.0	15.0
START(1)	Asymptotic	20.7	37.8	65.8
	Bootstrap	21.4	39.3	67.0
ARCHT(1)	Asymptotic	2.7	6.3	7.1
	Bootstrap	5.9	9.8	9.0
BDS	Asymptotic	10.3	9.9	10.0
	Bootstrap	9.0	12.1	14.1
(ii). $\mu_1=1, \mu_2=-1, \phi_1=-0.5, \phi_2=0.6, \gamma^*=50$				
Statistics	Critical Values	T=50	T=100	T=200
BLT(1,1)	Asymptotic	49.7	76.7	97.0
	Bootstrap	48.6	75.8	97.1
BLT(2,2)	Asymptotic	34.0	69.6	96.5
	Bootstrap	36.5	71.8	96.5
EART(1)	Asymptotic	45.5	74.9	94.5
	Bootstrap	46.5	76.5	94.9
START(1)	Asymptotic	50.7	82.3	98.8
	Bootstrap	51.5	82.9	98.9
ARCHT(1)	Asymptotic	3.9	7.6	10.9
	Bootstrap	7.9	10.5	13.4
BDS	Asymptotic	13.2	13.3	15.8
	Bootstrap	12.0	16.3	21.8

- Note: 1) The empirical powers are in %.
 2) The significance level is 5%.
 3) For each parameter setting, 1,000 Monte Carlo replications of LSTAR(1) are generated and within each Monte Carlo replication, 1,000 LSTAR(1) series are bootstrapped.

Table 3.15: Empirical powers of LM tests against ARCH(1,1) using ACV and BCV

(i). $\mu=0, \phi=0.6, \alpha_0=1, \alpha_1=0.8, \sigma^2=1$				
Statistics	Critical Values	T=50	T=100	T=200
BLT(1,1)	Asymptotic	20.1	25.5	37.9
	Bootstrap	23.1	28.0	39.4
BLT(2,2)	Asymptotic	33.2	45.6	60.7
	Bootstrap	36.3	48.0	61.1
EART(1)	Asymptotic	19.4	25.6	33.7
	Bootstrap	22.6	29.6	37.5
START(1)	Asymptotic	27.5	37.4	54.1
	Bootstrap	31.5	39.8	56.3
ARCT(1)	Asymptotic	59.0	91.1	100.0
	Bootstrap	64.0	91.5	99.9
BDS	Asymptotic	64.5	92.3	99.8
	Bootstrap	63.7	93.3	99.8
(ii). $\mu=1, \phi=0.6, \alpha_0=1, \alpha_1=0.8, \sigma^2=1$				
Statistics	Critical Values	T=50	T=100	T=200
BLT(1,1)	Asymptotic	19.4	24.7	37.6
	Bootstrap	20.2	25.3	38.0
BLT(2,2)	Asymptotic	32.4	42.7	59.6
	Bootstrap	33.3	44.8	59.9
EART(1)	Asymptotic	14.8	21.3	33.3
	Bootstrap	17.0	24.6	36.2
START(1)	Asymptotic	34.7	39.5	54.3
	Bootstrap	24.6	37.1	55.7
ARCT(1)	Asymptotic	60.0	90.1	100.0
	Bootstrap	66.9	91.6	99.9
BDS	Asymptotic	60.4	90.1	99.4
	Bootstrap	61.1	92.5	99.7

- Note:
- 1) The empirical powers are in %.
 - 2) The significance level is 5%.
 - 3) For each parameter setting, 1,000 Monte Carlo replications of ARCH(1,1) are generated and within each Monte Carlo replication, 1,000 ARCH(1,1) series are bootstrapped.

Table 3.16: Rank of power performance of LM tests

Model \ Test	BLT(1,1)	BLT(2,2)	EART(1)	START(1)	ARCHT(1)
BL(1,1,1)	3	1	5	2	4
EAR(1)	5	5	1	2	5
TAR(2;1,1)	2	2	2	1	2
ESTAR(1)	3	4	2	1	5
LSTAR(1)	2	4	4	2	5
ARCH(1,1)	2	2	2	2	1

linearity test — the BDS test. The null model in this investigation is the AR(1) model and the alternative models include various parameterization of BL(1,1,1), EAR(1), TAR(2;1,1), ESTAR(1), LSTAR(1) and ARCH(1,1) models.

The main conclusions are:

1. Power comparison of LM tests

We discover that the five LM tests BLT(1,1), BLT(2,2), EART(1), START(1), and ARCHT(1) vary widely in terms of power levels depending on model parameters. The natural thought that the specific LM test corresponding to the nonlinear model will have higher power than other LM tests is proved wrong. For example, for a BL(1,1,1) series with large values of c , BLT(1,1) might not be the test with highest power since it is a local test.

However, among the five LM tests we discussed, the START(1) seems to be the best for the simulated series of BL(1,1,1), EAR(1), TAR(2;1,1), ESTAR(1), LSTAR(1) and ARCH(1,1) with respect to overall power performance. Based on the research we have done, a trial to order the tests by empirical power for different series is listed in Table 3.16. The numbers in the table represents the rank a test will have for a certain type of series. The smaller the number is, the higher the rank. And we can see that START(1) has ranked at least the second for all the series investigated.

(1). The BL(1,1,1) model: $y_t = \mu + \phi y_{t-1} + c y_{t-1} \varepsilon_{t-1} + \varepsilon_t$

Figures 3.1–3.5 show that for the coefficient c near the null value 0, BLT(1,1), BLT(2,2), START(1) are good; when the absolute value of c increases, BLT(2,2) and START(1) are better than BLT(1,1) for high empirical power levels and small variance between power curves. ARCHT(1) works not very well when c is small. EART(1) seems to be the worst. Usually we will think that BLT(1,1) should be the

best test for BLT(1,1,1) model, but it is not.

$$(2). \text{ The EAR(1) model: } y_t = \mu + \phi y_{t-1} + [\exp(-y_{t-1}^2) - 1] \theta y_{t-1} + \varepsilon_t$$

From Figure 3.6–3.10, it's very obvious that the BLT(1,1), BLT(2,2), ARCHT(1) statistics almost have no power against the EAR(1) model: excluding the jump points⁵, their empirical powers are around 5% and not sensitive to the value taken by the model coefficients ϕ and θ . The pattern of the empirical powers of the EART(1) is very similar to that of the START(1), and the level of the empirical powers of the EART(1) is a little bit higher than that of the START(1).

$$(3). \text{ The TAR(2;1,1) model: } y_t = \begin{cases} \mu_1 + \phi_1 y_{t-1} + \varepsilon_{2t} & \text{if } y_{t-1} \geq c \\ \mu_2 + \phi_2 y_{t-1} + \varepsilon_{1t} & \text{if } y_{t-1} < c \end{cases}$$

There are many coefficients in TAR(2;1,1) model: $\mu_1, \mu_2, \phi_1, \phi_2, \sigma_1, \sigma_2$ and c , so empirical power curves are not produced for the TAR model. Still, from Table 3.5, we can get a brief impression about the tests' power performance. ARCHT(1) is not very sensitive to TAR models when $\sigma_1 = \sigma_2$. However, it has the highest power when $\sigma_1 \neq \sigma_2$. START(1) comes next. For cases when $\sigma_1 = \sigma_2$, EART(1) and START(1) are better than the BLT type tests if intercepts are not equal to 0 and START(1) is better than EART(1); if intercepts are equal to 0, the power of BLT(1,1) is highest, the second is START(1), and the rest are in the order BLT(2,2), EART(1) and ARCHT(1). The situation for the TAR model is really complicated, concerning consistently high power, START(1) has a score of 1, and the others have 2.

$$(4). \text{ The ESTAR(1) model: } y_t = \mu_1 + \phi_1 y_{t-1} + (\mu_2 + \phi_2 y_{t-1}) \times \left\{ 1 - \exp \left[-\gamma (y_{t-1})^2 \right] \right\} + \varepsilon_t$$

It is difficult to plot the empirical power curves for the ESTAR model since there are five model coefficients. Special cases are illustrated in Table 3.6. ARCHT(1) seems not to have obvious power against this model. When γ takes very a large value, such as 10,000, the ESTAR approximates the liner model, so all the tests' powers are around 5%. When the two intercepts are equal to zero, the EART(1) and the START(1) are the best two, with EART(1) being a little better than the START(1); the BLT tests' powers are very low. When intercepts are nonzero, START(1) is the best. The order of EART(1), BLT(1,1) and BLT(2,2) changes as γ changes. Throughout all the cases in Table 3.6, BLT(1,1) is better than BLT(2,2).

⁵When the model coefficients $|\phi - \theta| \geq 1$, there will be either a jump in the test statistics or no statistics value at all.

(5). The LSTAR(1) model: $y_t = \mu_1 + \phi_1 y_{t-1} + (\mu_2 + \phi_2 y_{t-1}) \times \left[\{1 + \exp[-\gamma(y_{t-1})]\}^{-1} - 0.5 \right] + \varepsilon_t$

Again, there are no empirical power curves for LSTAR model due to the number of coefficients.

When intercepts are zero, the order is like: BLT(1,1), START(1), BLT(2,2), EART(1) and ARCHT(1);

when intercepts are nonzero, the order is: START(1), BLT(1,1), BLT(2,2) or EART(1), ARCHT(1).

Results are in Table 3.7.

(6). The ARCH(1,1) model: $y_t = \mu + \phi_1 y_{t-1} + v_t$ with $v_t = \varepsilon_t \sqrt{h_t} = \varepsilon_t \sqrt{\alpha_0 + \alpha_1 v_{t-1}^2}$

The ARCHT(1) is best, though the other four have good powers too. See Table 3.8.

Overall, the START(1) can be applied widely since it has relatively high powers for all the nonlinear series. The ARCHT(1) is relatively narrow because it does not have high power for non-ARCH series.

Of course, the above conclusion may not always be true since the research here is still limited.

2. Asymptotic critical values and bootstrap critical values

We use asymptotic critical values in the above power comparison. However, for small sample size, critical values created from the asymptotic distribution may not be good approximations of the true critical values. We consider replacing the asymptotic critical values with the bootstrap critical values. It turns out that the bootstrap method generally has higher empirical power than the asymptotic method, though the differences are rather small. Moreover, the bootstrap method is better than the asymptotic method concerning size. The empirical size obtained by using bootstrap critical values is always near the nominal 5% level, while the empirical size produced by the asymptotic critical values deviate from 5% for small sample size and large absolute values of autocorrelation coefficient ϕ , especially for the START(1). Last, the actual size using asymptotic critical values for ARCHT(1) is systematically below 5%.

In all, we should consider using bootstrap critical values rather than asymptotic critical values when

- * Sample size is less than 200;
- * The autoregressive coefficient is near the unit circle;
- * The linearity test is ARCHT or BDS.

3. LM and BDS tests

The comparison of LM test and BDS test can be approached from the following perspectives:

(1). Generality

At first sight, the advantage of BDS test over an LM test is its generality. But the research here shows that

- * An LM test can have good power against a variety of nonlinear series, sometimes even higher power than the LM test constructed specifically for a nonlinear series with known structure.

- * Though the BDS test is supposed to be a general test, its power usually is not very high. From section 3.4, we observe that except for the ARCH model, the powers of the BDS test are generally below 20%, which puts its power rank 5th or 6th among the six tests discussed.

(2). Construction and running time

It is very time-consuming to calculate the BDS statistic, while it is pretty simple to calculate LM statistics. There are some uncertainties when constructing a BDS test. We need to pre-specify the parameters m and ϵ . So in order to get a complete view of BDS testing results, various BDS statistics with different m and ϵ need to be computed, which puts more stress on the computer.

The building of an LM test requires the parameterization of the alternative model, which is usually unknown in empirical studies. This problem could possibly be handled in two ways:

- * Primary data investigation and the characteristics of a certain type of data could provide some clues. For example, clusters of bursts in high frequency financial data usually implies an ARCH type of data. It is like having the model first and using the test to verify the instinct.

- * We can also run various LM tests on the data. It is most likely that the LM test with the highest power points to the right nonlinear model. Contrary to the first method, this method uses tests to search for a possible nonlinear model.

No matter which method we use, we have to spend time either doing the primary data investigation or running several LM tests. Therefore, in empirical studies, a linearity test is not simply doing a test. It becomes a system of work, requiring time and scrutiny.

CHAPTER 4

EMPIRICAL STUDY: AN APPLICATION TO EXCHANGE RATES

In this chapter, possible nonlinearities in four major countries' foreign exchange rates are explored. Both the START test, which is the champion of competition of various linearity tests in last chapter, and the BDS test are performed on the data. Comparisons of forecasts generated by the ESTAR and BL models are made with forecasts generated by the linear AR model and the random walk model (with and without drift). In-sample and out-of-sample forecasting comparisons are made.

4.1 Introduction

In their 1983 paper, Meese and Rogoff [81] suggest that for the out-of-sample forecasting performance of nominal foreign exchange rates, the seemingly naive random walk model does the best. Other models, including univariate time series models, unconstrained vector autoregression models, or structural models like the Dornbusch-Frankel sticky-price monetary model cannot beat the random walk model as far as out-of-sample forecast performance is concerned.

After that, many efforts have been tried to overturn this conclusion.

As Meese and Rogoff say in their paper [81], "as long as the exchange rate does not exactly follow a random walk, we could expect one of the estimated time series models to prevail in a large enough sample". However, since the floating system of foreign exchange rates came into effect in 1973, the data for foreign exchange rates may still not be long enough even after 18 years. Some (For example, Frankel [48], Edison [37]) are able to reject the random walk model with mixed fixed and floating rate data, but it always raises doubts that whether the rejection is due to the long data or due to the change of the exchange rate system.

Meese and Rogoff [81] measure the out-of-sample accuracy by three statistics: mean error (ME), mean absolute error (MAE), and root mean square error (RMSE) with RMSE as the principal criterion

for comparing forecasts. Many others use the mean square prediction error (MSPE), which is the analogy to the RMSE. Yet some also judge the forecasts by looking at their ability to predict the direction of the changes of foreign exchange rates, besides their ability to minimize the MSPE or RMSE, etc. For example, Engel [42] shows that by measuring predictive ability of the direction of changes, the Markov regime-switching model is better than the random walk model, though it does not improve on the random walk model in achieving minimum MSPE. By using a forecast density evaluation technique, Clements and Smith [27] are able to show the inadequacy of the random walk model versus the SETAR model.

Meese and Rogoff [81] make no attempt to account for possible nonlinearities in the data. Therefore, it is possible that there are nonlinear models which can do better ex-ante forecasting than the random walk model. Diebold and Nason [34] apply the nonparametric locally-weighted regression to exchange rate changes in an autoregressive framework. Meese and Rose [82] allow for nonlinear extensions to structural exchange rate models. Their work indicates some nonlinear dependence in conditional means, however, the out-of-sample forecasting is not improved over the random walk model by accounting for nonlinearities.

Motivated by the conjecture that the managed float system leads to different behavior of moderate and large exchange rate changes (Hsieh [66]), Kräger and Kugler [71] find that the self-exciting threshold autoregressive (SETAR) models can be used to describe the behavior of five weekly dollar exchange rate series from 1980 to 1990. However, the BDS and stability tests suggest that SETAR models may be mis-specified. Similar problems occur with the GARCH model. Pippenger and Goering [101] apply the SETAR model to the logarithm changes of 13 countries' monthly exchange rates. Besides a significant fit, they find that the SETAR model produces better forecasts than the random walk model in terms of MSPE for both in-sample and out-of-sample forecasting, though the Diebold-Mariano (D-M) statistic [33] indicates that the two models generate forecasts of similar accuracy. Clements and Smith [27] consider the forecasting performance of two SETAR exchange rate models by Kräger and Kugler [71]. They argue that whether the SETAR model is better than the random walk model or not depends on the "state of nature" and the forecasting measurement.

By constructing a Markov regime-switching model [58] as a sequence of stochastic, segmented time trends on the percentage changes of quarterly foreign exchange rates of Germany, France and UK from 1973:IV to 1988:I, Engel and Hamilton [43] are able to explain the phenomenon that the value of the dollar appears to move in one direction for long periods of time. The random walk model is formally rejected in favor of the model of long swings, which also provides better forecasts in sample. However, the advantage of the Markov regime-switching model over the random walk model while forecasting out of sample is ambiguous. Engel [42] argues that while the Markov regime-switching model does not generate superior out-of-sample forecasts measured by the MSPE, it does a better job in predicting the direction of the foreign exchange rate changes than the random walk model.

According to Kilian and Taylor [70], the most successful nonlinear empirical exchange rate models embody the smooth transition autoregression model. Several papers have investigated the nonlinearities in the form of an ESTAR model of real exchange rates, including Balke and Fomby [6], Taylor and Peel [113], Taylor et al. [114], and Baum et al. [8]. Possible explanations for the model could be transaction costs, foreign exchange market interventions, uncertainty about the equilibrium level of the exchange rate, etc. The ESTAR model is attractive for describing nonlinearities in exchange rates because it allows a smooth transition between two extreme regimes and symmetric adjustment for deviations above and below the equilibrium level.

This empirical study can be regarded as looking at the foreign exchange rates from the nonlinear perspective. The data start after the breakdown of the Bretton Wood system in 1973. Forecasts will be compared by the traditional MSPE and the MSPE difference will be tested by the D-M test. But possible nonlinearities in the exchange rates will be tested, estimated and forecasted. For the six nonlinear models discussed before, only the ESTAR and the BL model are used here. The ESTAR model is preferred to the TAR model because the latter is built under the assumption that the transition between regimes is abrupt and sudden. For the foreign exchange rate markets, that requires that many of the market participants make the same decision at the same time, which is quite improbable. Besides, the ESTAR model includes the TAR model as a special case. Another smooth transition model, the LSTAR, implies

asymmetric smooth transition from one regime to another, which is deemed to be inappropriate for modeling exchange rate movement. The EAR model is also a special case of the ESTAR and sometimes considered as the precedent for the ESTAR model. The ARCH model has done a good job in estimating and forecasting daily and weekly exchange rate data, but many researchers, like Baillie and Bollerslev [5], find the ARCH type nonlinearities are not a typical pattern when monthly or annual data are tested. Actually in this case, estimation of the data as an ARCH model produces a worse SBC than the linear AR model. Literature about application of the bilinear model to the nominal foreign exchange rate is rare so far. But we find the BL model to be a strong competitor in this empirical study.

4.2 Data Analysis

The data set comprises monthly averages of the Canadian dollar, French franc, British pound, and Japanese yen against the US dollar, which are denoted as x_1 , x_2 , x_3 , x_4 , respectively. Note that subscriptions 1, 2, 3, and 4 will be used to refer to country Canada, France, Britain and Japan in the rest of the chapter. The sample period is from January 1973 to March 2001 for all four series. So in total, there will be 339 observations for each series. The above data are downloaded from the webpage of Pacific Exchange Rate Services (<http://pacific.commerce.ubc.ca/xr/data.html>), which is supported by Professor Werner Antweiler of the University of British Columbia at Vancouver, Canada.

For the data transformation the first step is to take the natural logarithm for the foreign exchange rate series in order to alleviate possible heteroskedasticities and avoid any problems arising from Jensen's inequality (Meese and Rogoff [81]). Also to keep the data in a comfortable magnitude, multiplication by 100 is applied to the logarithmic foreign exchange rates. The resulting patterns are illustrated by Figure 4.1.

The augmented Dickey-Fuller test (ADF) [30, 31] and the Phillips-Perron test (PP) [99, 100] are used to examine if there is any presence of a unit root. As Table 4.1 shows, the foreign exchange rates are $I(1)$ processes. Accordingly, a first difference transformation is applied to the previously transformed data. Also, means are subtracted from the data to make later estimations easier. The means are $c_1 = 0.1315$, $c_2 = 0.1040$, $c_3 = 0.1448$, and $c_4 = -0.2692$. Therefore, the transformed data series

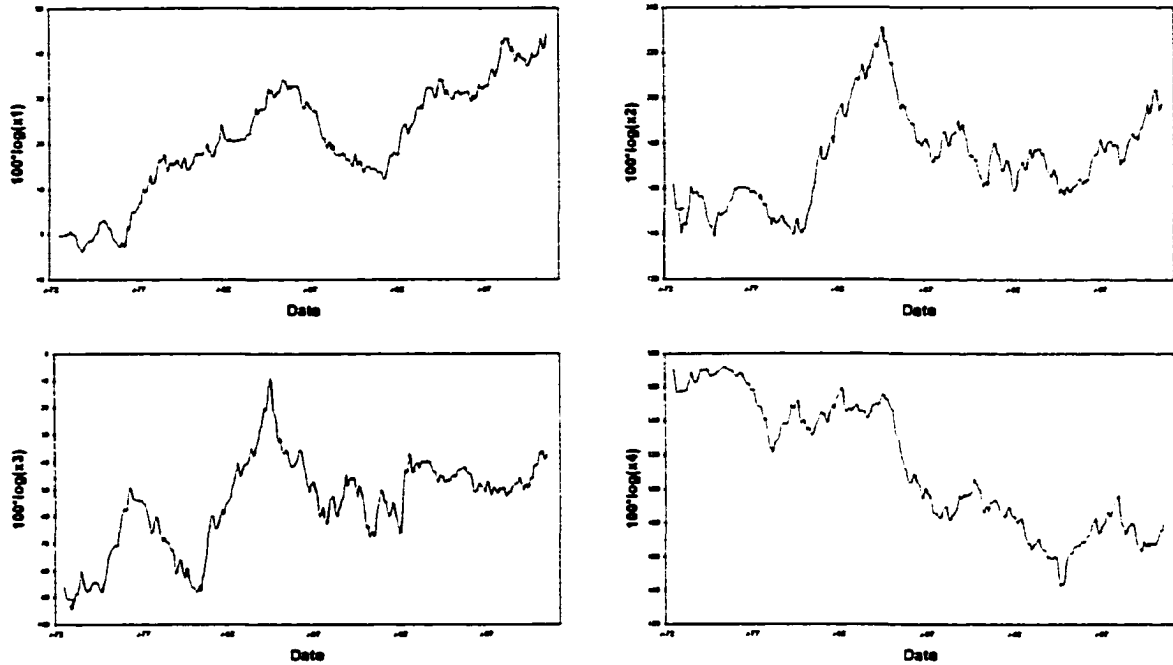


Figure 4.1: Monthly average exchange rates (1973:1-2001:3)

are: $100 \times \{\log x_{1t} - \log x_{1t-1}\} - c_1$, $100 \times \{\log x_{2t} - \log x_{2t-1}\} - c_2$, $100 \times \{\log x_{3t} - \log x_{3t-1}\} - c_3$, $100 \times \{\log x_{4t} - \log x_{4t-1}\} - c_4$. We name these $\{y_{1t}\}$, $\{y_{2t}\}$, $\{y_{3t}\}$, $\{y_{4t}\}$, respectively.

4.3 Linear Estimation

The data properties of $\{y_{1t}\}$, $\{y_{2t}\}$, $\{y_{3t}\}$, and $\{y_{4t}\}$ from February 1973 to March 2001 are listed in Table 4.2¹. We find that there are strong indications of serial correlation for all four series according to the Ljung-Box Q statistics.

We start our parametric modeling of the series with the linear autoregressive model first. To get a parsimonious model, the method of forward selection (Neter et al. [88]) is used here. The possible regressor set for the four exchange rates includes autoregressive lag terms of the dependent variables with a maximum lag of 12 (which is reasonable for monthly data), so there will be 12 candidate regressors². To be consistent, the starting point for estimation of the dependent series $\{y_{1t}\}$, $\{y_{2t}\}$, $\{y_{3t}\}$, and $\{y_{4t}\}$ will be February 1974 and the ending point is March 2001. The regressor selection criterion could be

¹Numbers in parentheses are corresponding p values.

²Since the data has already been demeaned, a constant term is not included in the regressor set.

Table 4.1: Unit root test

Country	Series	ADF test			PP test	
		τ	τ_{μ}	τ_{τ}	τ_{μ}	τ_{τ}
Canada	x_1	1.3396	-0.7383	-1.6384	-0.6576	-1.5344
	dx_1	-15.2276	-15.4241	-15.4018	-15.3664	-15.3437
France	x_2	0.4871	-1.5446	-1.7174	-1.3170	-1.5760
	dx_2	-13.6594	-13.6638	-13.6431	-13.6273	-13.6070
Britain	x_3	-1.5334	-2.4236	-2.6868	-2.0397	-2.2643
	dx_3	-11.9670	-11.9992	-11.9982	-12.5512	-12.5406
Japan	x_4	-1.0899	-1.1808	-2.2668	-1.3022	-1.9088
	dx_4	-12.8216	-12.8586	-12.8450	-12.7817	-12.7689

- Note: 1) The regression equations used by the test statistics τ , τ_{μ} , and τ_{τ} are
 $\Delta y_t = \gamma y_{t-1} + \sum_{i=1}^p \beta_i \Delta y_{t-i} + \varepsilon_t$,
 $\Delta y_t = a_0 + \gamma y_{t-1} + \sum_{i=1}^p \beta_i \Delta y_{t-i} + \varepsilon_t$,
 $\Delta y_t = a_0 + a_1 t + \gamma y_{t-1} + \sum_{i=1}^p \beta_i \Delta y_{t-i} + \varepsilon_t$, respectively.
- 2) The lag of the ADF test is chosen by BIC, the lag of the PP test is fixed at 3.
- 3) Not only the single unit root tests are performed (rows of x_i 's), but also the double unit root tests (rows of dx_i 's), $i = 1, \dots, 5$.
- 4) The critical values are the same for the ADF and PP test. For a sample size of 250, the 5% critical values are -1.95, -2.88 and -3.42 for τ , τ_{μ} , and τ_{τ} .

Table 4.2: Summary statistics of transformed foreign exchange rates (1973:2-2001:3)

Statistics	y_1	y_2	y_3	y_4
Variance	1.0804	7.2594	6.3949	8.3724
Minimum	-2.5089	-7.2344	-9.8841	-10.1554
Maximum	3.2126	8.8577	10.5984	8.4752
Skewness	0.1912 (0.1531)	0.0762 (0.5689)	0.1246 (0.3517)	-0.5608 (0.0000)
Kurtosis	-0.0625 (0.8164)	0.2882 (0.2844)	1.4254 (0.0000)	0.8703 (0.0012)
Q(8)	17.2253 (0.0278)	37.0551 (0.0000)	44.3822 (0.0000)	49.5994 (0.0000)
Q(16)	41.0422 (0.0005)	39.1232 (0.0010)	57.4685 (0.0000)	63.3213 (0.0000)
Q(24)	47.7519 (0.0027)	48.8865 (0.0020)	71.0741 (0.0000)	75.1589 (0.0000)
Q(32)	60.4543 (0.0017)	52.7281 (0.0120)	81.2776 (0.0000)	79.5927 (0.0000)
Q(40)	70.4172 (0.0021)	65.5555 (0.0066)	97.3855 (0.0000)	99.7865 (0.0000)
Q(48)	83.2750 (0.0012)	72.7056 (0.0122)	127.9524 (0.0000)	128.3407 (0.0000)

Table 4.3: Linear estimation

Country	Estimated AR Models	AIC	SBC	RSS
Canada	$y_{1t}=0.1488 \times y_{1t-1} + 0.1373 \times y_{1t-8} + 0.1394 \times y_{1t-10} + \varepsilon_t$ (0.0065) (0.0130) (0.0121)	1905.46	1916.82	339.19
France	$y_{2t}=0.2762 \times y_{2t-1} + \varepsilon_t$ (0.0000)	2481.53	2485.31	2010.05
Britain	$y_{3t}=0.3952 \times y_{3t-1} - 0.1419 \times y_{3t-2} + \varepsilon_t$ (0.0000) (0.0100)	2448.48	2456.06	1805.18
Japan	$y_{4t}=0.3409 \times y_{4t-1} + 0.1409 \times y_{4t-11} + \varepsilon_t$ (0.0000) (0.0063)	2526.05	2533.63	2290.11

the Schwartz Bayesian Criterion (SBC), Akaike Information Criterion (AIC), or Residual Sum Squares (RSS). Here, the SBC is used. The regression procedure for series $\{y_1\}$, for example, is:

1. First, run a do-loop to regress $\{y_{1t}\}$ on each regressor from the regressor set $\{\{y_{1t-1}\}, \{y_{1t-2}\}, \dots, \{y_{1t-11}\}, \{y_{1t-12}\}\}$. Choose the regressor with the lowest SBC (SBC_1), which is $\{y_{1t-1}\}$ for the Canadian dollar;

2. Second, conduct a do-loop by running regressions of $\{y_{1t}\}$ on $\{y_{1t-1}\}$ and one regressor from the remaining regressor set. The term $\{y_{1t-8}\}$ is chosen with the lowest SBC (SBC_2) from the 11 candidates. However, only if SBC_2 is less than SBC_1 can $\{y_{1t-8}\}$ be added to the regression function. Otherwise, the regression function is done at the first step. Also if the introduction of $\{y_{1t-8}\}$ seriously decreases the t statistic of $\{y_{1t-1}\}$, then $\{y_{1t-1}\}$ should be dropped from the regression function. The Canadian dollar series $\{y_{1t}\}$ does not have this problem;

3. The third round do-loop is similar to do-loop two. As a result, the term $\{y_{1t-11}\}$ is added to the regression function and neither $\{y_{1t-1}\}$ nor $\{y_{1t-8}\}$ are dropped;

4. The fourth do-loop produces a minimum SBC (SBC_4) that is larger than the third round minimum SBC (SBC_3), so the selection stops at step three. Therefore the final autoregressive terms of $\{y_{1t}\}$ are $\{y_{1t-1}\}$, $\{y_{1t-8}\}$ and $\{y_{1t-11}\}$.

All the estimated autoregressive models and the relevant values of the AIC, SBC and RSS are listed in Table 4.3. The properties of the estimated residuals (see Table 4.4), which are named $\{\varepsilon_1\}$, $\{\varepsilon_2\}$, $\{\varepsilon_3\}$, $\{\varepsilon_4\}$, are compared with $\{y_1\}$, $\{y_2\}$, $\{y_3\}$, $\{y_4\}$. Through the linear filter, the variance and serial

Table 4.4: Summary statistics of residuals from linear estimation (1974:2-2001:3)

Statistics	ε_1	ε_2	ε_3	ε_4
Variance	1.0436	6.1848	5.5543	7.0465
Minimum	-2.9076	-7.6166	-8.3636	-9.2408
Maximum	3.1751	9.7354	8.2783	7.4692
Mean	0.0102	-0.0012	-0.0116	-0.0019
Skewness	0.0425 (0.7550)	0.0556 (0.6834)	0.0578 (0.6713)	-0.5588 (0.0000)
Kurtosis	-0.0361 (0.8954)	0.8359 (0.0023)	0.9774 (0.0004)	0.8907 (0.0012)
Q(8)	3.1042 (0.9276)	8.1729 (0.4168)	7.4280 (0.4912)	13.4830 (0.0963)
Q(16)	15.4569 (0.4915)	13.3876 (0.6442)	16.1280 (0.4441)	21.0748 (0.1757)
Q(24)	23.7134 (0.4781)	24.6846 (0.4231)	25.6570 (0.3708)	31.2883 (0.1456)
Q(32)	37.8677 (0.2191)	29.2429 (0.6068)	31.6687 (0.4833)	39.4150 (0.1722)
Q(40)	41.6394 (0.3993)	41.6500 (0.3988)	41.4380 (0.4078)	50.9572 (0.1148)
Q(48)	47.6404 (0.4875)	48.9730 (0.4338)	59.3005 (0.1271)	64.3404 (0.0576)

correlation of the data series decrease. None of the Ljung-Box Q statistics for the residuals calculated for lags from 8 to 48 with a step of 8 are significant for $\alpha = 5\%$.

4.4 Nonlinear Exploration

It seems that a linear regression does a good job of modeling the foreign exchange rates of the four countries. But is it possible that some nonlinear time series models might surpass the linear time series models? The BDS test, a general nonlinear diagnostic test is used here to investigate the possible nonlinear structure of the prewhitened data, $\{r_1\}$, $\{r_2\}$, $\{r_3\}$, and $\{r_4\}$.

The BDS tests results with m (correlation integral dimensions) running from 2 to 5 and l (distance in terms of data variance) ranging from 0.5 to 1.5 with a step of 0.25 are listed in Table 4.5. The BDS test is asymptotically distributed as a standard normal distribution $N(0, 1)$. So for significance

Table 4.5: BDS statistics of residuals from linear estimations

m	l	ε_1	ε_2	ε_3	ε_4
2	1.5	2.3741	-2.2403	2.0952	2.1885
3	1.5	3.2238	-1.2567	2.3775	2.2010
4	1.5	3.2724	-0.3699	3.3085	2.1067
5	1.5	3.1292	0.2852	4.0210	2.4171
2	1.25	2.4177	-1.9650	2.0223	1.8505
3	1.25	3.2769	-0.9709	2.4418	2.0135
4	1.25	3.2113	-0.0637	3.4749	2.0664
5	1.25	3.1833	0.4535	4.2155	2.5066
2	1	2.5739	-2.0252	1.6435	1.9031
3	1	3.5246	-0.9610	2.4180	2.2997
4	1	3.2709	0.2395	3.6321	2.3100
5	1	3.4016	0.8171	4.5205	2.5996
2	0.75	2.2461	-1.1463	1.6785	1.9842
3	0.75	3.2703	0.0586	2.2302	2.7874
4	0.75	3.2438	1.2987	3.5470	2.9809
5	0.75	3.4769	2.4009	4.3610	3.1226
2	0.5	1.9053	-1.6638	1.9110	2.7065
3	0.5	3.4186	-1.0700	2.3564	2.8692
4	0.5	4.2320	-0.8349	3.6496	3.2080
5	0.5	4.0929	-0.8837	4.6387	4.2922

level 5%, the asymptotic critical value is ± 1.96 . However, the asymptotic distribution usually requires sample size no less than 500 to have a good approximation of the true distribution. Since our sample size is only 326 for estimation period from 1974:02 to 2001:03, special attention is needed toward the discrepancy between the asymptotic distribution and the true distribution of the BDS statistic. Kanzler [67] conducts Monte-Carlo simulations to tabulate the small sample distribution of the BDS test. The relevant information is listed in Table 4.6.

Examining Tables 4.5 and 4.6, we find that there is evidence of nonlinear structures in the Canadian dollar, British pound and Japanese yen, and almost no evidence of nonlinearities for the French franc. Except for the BDS statistic -2.2403 when $l = 1.5$ and $m = 2$, which is only slightly significant at best, other BDS statistics of French franc are not significant for 95% significance level.

The BDS test shows that there could be nonlinear structures in the foreign exchange rate data of three countries. To compare with the BDS test, the LM test for the alternative STAR model is conducted for the four countries. Although the STAR is derived as a specific test, it was shown in the last chapter that the STAR has quite general power against other nonlinear models besides the STAR model and

Table 4.6: Small sample distribution of BDS statistic

T	m	l	2.5%	97.5%	T	m	l	2.5%	97.5%
250	2	1.5	-2.13	2.22	500	2	1.5	-2.01	2.10
250	3	1.5	-2.12	2.22	500	3	1.5	-2.02	2.09
250	4	1.5	-2.14	2.25	500	4	1.5	-2.00	2.12
250	5	1.5	-2.13	2.25	500	5	1.5	-1.98	2.11
250	2	1.0	-2.09	2.26	500	2	1.0	-2.01	2.15
250	3	1.0	-2.11	2.30	500	3	1.0	-2.00	2.15
250	4	1.0	-2.09	2.35	500	4	1.0	-1.99	2.20
250	5	1.0	-2.10	2.44	500	5	1.0	-1.96	2.26
250	2	0.5	-2.54	2.81	500	2	0.5	-2.22	2.46
250	3	0.5	-2.86	3.13	500	3	0.5	-2.38	2.64
250	4	0.5	-3.29	3.75	500	4	0.5	-2.62	2.92
250	5	0.5	-4.11	4.78	500	5	0.5	-3.09	3.55

it usually performs better than the BDS test. So here the START is used as a general linearity test to detect the nonlinearities in the exchange rates data rather than as a specific test testing for STAR type nonlinearities.

One problem arises here. In Chapter Three, we limit the order and the lag parameter of the alternative ESTAR and LSTAR models to be 1, so the corresponding LM tests for ESTAR(1) and LSTAR(1) are the same, which we name START(1). However, in this chapter, we relax the restriction and allow the order to vary between 1 to 12 and the lag take values from 1 to 5. Therefore, we will have different ESTAR and LSTAR, and the information from Chapter Three can not tell us which START test, ESTAR or LSTAR, has the better general powers. So we will calculate both the ESTAR and LSTAR statistics for the exchange rate data in this chapter.

The ESTAR for the alternative single-term ESTAR model

$$y_t = \sum_{i=1}^p \phi_i y_{t-i} + \left\{ 1 - \exp \left[-\gamma (y_{t-d} - c)^2 \right] \right\} \times \theta_j y_{t-j} + \varepsilon_t \quad (4.1)$$

has the z_{1t} term

$$z_{1t} = - \left[y_{t-j} y_{t-d} \quad y_{t-j} y_{t-d}^2 \right]' \quad (4.2)$$

So the ESTAR will have a χ^2 distribution asymptotically with the degrees of freedom equal 2.

The LSTAR for the alternative single-term LSTAR model

$$y_t = \sum_{i=1}^p \phi_i y_{t-i} + \left[\{ 1 + \exp \left[-\gamma (y_{t-d} - c) \right] \}^{-1} - 0.5 \right] \times \theta_j y_{t-j} + \varepsilon_t \quad (4.3)$$

has the z_{1t} term

$$z_{1t} = - [y_{t-j}y_{t-d} \quad y_{t-d}^3]' . \quad (4.4)$$

So the asymptotic distribution of LSTART is also χ_2^2 .

The ESTART statistics when model order p runs from 1 to 12 for transition variables being lagged dependent variables with $d = 1, 2, 3, 4,$ and 5 are listed in Table 4.7. The highest ESTART statistic for each series is underlined. The results are encouraging. For Canada, the highest statistic 9.487 happens for a transition variable y_{1t-2} with $p = 1$; the French franc exhibits nonlinearities with the highest ESTART 9.376 for $p = 6$ and $d = 3$; with the transition variable being y_{3t-3} , the ESTART(9) equals 20.688, indicating that the British pound has significant nonlinearities; Japan has the largest ESTART 12.001 with $p = 2$ and $d = 1$. Asymptotically, the ESTART statistics for a specific value of p and d , whose alternative is a single-term ESTAR model, follows a chi-square distribution with degrees of freedom equal 2, which has a critical value of 5.99 for the 95% significance level. Although it is not the exact asymptotic distribution of the $\text{sup}(\text{ESTART})$ test we use here, a statistic value of $\text{sup}(\text{ESTART})$ much larger than the critical value for ESTART makes us believe that the four foreign exchange rate series do exhibit nonlinear dynamics. The LSTART statistics for the four foreign exchange rates are tabulated in Table 4.8 with the $\text{sup}(\text{LSTART})$ underlined. All the $\text{sup}(\text{LSTART})$ statistics are significantly larger than 5.99, the critical value of χ_2^2 at the 95% significance level. Therefore, we come to the conclusion that both the ESTART and LSTART detect nonlinearities in the four exchange rate series.

Comparing the conclusions drawn from the BDS test and ESTART tests, the BDS test shows evidence of nonlinear structures for the foreign exchange rates of Canada, Britain, and Japan and almost no evidence of nonlinearities for the French franc. However, the ESTART statistics indicate nonlinear behaviors for all four countries' exchange rates. Given the better performance of ESTART against BDS test, it is believed that nonlinearities do exist in the exchange rates of the four countries. In the next section, further investigation of the nonlinearity is conducted: specific nonlinear models such as the exponential smooth transition autoregressive (ESTAR) model and the bilinear (BL) model will be discussed for the foreign exchange rates data.

Table 4.7: ESTART statistics

d\p	1	2	3	4	5	6	7	8	9	10	11	12
Canada												
1	2.432	6.745	2.442	0.523	1.269	1.203	0.511	3.261	1.663	1.541	3.337	0.270
2	<u>9.487</u>	0.979	0.282	0.397	1.706	2.132	3.320	1.688	0.327	0.715	5.622	5.357
3	1.822	0.859	0.404	5.506	4.070	1.070	0.059	5.723	1.852	0.267	0.358	3.883
4	1.148	1.263	2.178	0.019	1.846	3.028	0.450	0.060	0.950	1.891	4.781	1.607
5	3.044	2.538	1.864	2.043	1.416	0.342	0.637	0.503	1.571	9.051	1.161	2.795
France												
1	0.585	1.961	3.409	1.387	1.369	1.311	2.761	2.310	2.105	2.460	1.551	1.916
2	1.589	0.763	0.555	1.844	4.219	1.673	4.159	4.012	0.702	0.062	0.222	1.030
3	3.296	1.121	7.634	1.743	3.066	<u>9.376</u>	3.095	2.251	1.540	5.016	0.451	2.274
4	0.210	1.712	1.073	0.253	1.843	0.299	5.204	0.935	2.995	2.281	1.599	3.024
5	0.179	0.317	1.742	1.598	0.152	1.110	1.699	3.622	4.595	0.470	1.524	1.046
Britain												
1	2.304	0.986	1.669	1.616	2.503	2.468	1.778	7.099	5.611	1.527	3.433	2.079
2	3.294	0.264	0.661	0.333	1.517	4.356	3.127	0.460	0.810	0.047	2.733	0.704
3	3.633	2.242	10.730	6.834	9.013	0.397	7.498	3.143	<u>20.688</u>	16.036	0.065	3.464
4	1.215	3.963	7.675	0.843	1.265	3.743	0.253	0.217	<u>1.650</u>	2.795	5.035	2.339
5	1.675	1.823	10.004	6.053	4.085	8.178	1.074	1.599	1.309	3.320	4.684	0.196
Japan												
1	1.034	<u>12.001</u>	3.496	2.657	1.554	3.886	3.839	3.252	0.354	5.501	0.905	0.393
2	10.801	3.462	0.398	0.373	1.421	0.147	2.232	5.688	2.142	2.016	4.351	1.660
3	4.134	0.380	0.633	1.783	4.709	3.430	1.681	0.414	1.064	2.056	0.431	2.224
4	0.146	1.813	3.040	0.007	5.408	1.296	0.534	2.249	1.119	1.961	0.916	7.716
5	0.857	1.158	2.967	1.098	3.971	2.576	1.883	7.909	6.480	2.783	0.386	1.316

Table 4.8: LSTART statistics

d\p	1	2	3	4	5	6	7	8	9	10	11	12
Canada												
1	2.432	1.082	2.446	1.119	1.562	1.640	1.317	3.564	1.244	0.984	0.979	1.087
2	1.238	0.979	0.723	1.257	2.275	2.976	4.434	2.328	0.737	0.851	1.652	1.966
3	2.199	0.372	0.404	2.060	2.156	0.372	0.497	0.389	1.533	0.372	0.643	3.192
4	0.270	0.404	1.335	0.019	1.314	2.598	0.049	0.063	0.012	0.589	0.300	0.184
5	2.530	2.956	2.784	3.119	1.416	1.901	2.201	1.905	1.499	<u>9.793</u>	2.300	3.608
France												
1	0.585	0.399	3.144	0.215	0.215	0.932	1.830	0.509	2.151	1.695	0.305	0.505
2	0.762	0.760	1.025	1.873	0.774	2.068	4.288	1.122	1.002	0.663	0.858	1.529
3	<u>10.668</u>	8.569	7.487	8.943	7.698	10.318	10.369	7.675	7.916	8.436	7.276	9.124
4	0.219	1.791	1.119	0.252	0.994	0.301	4.786	1.083	2.823	0.237	0.680	2.963
5	0.000	0.074	1.654	0.702	0.151	0.993	0.826	0.075	1.081	0.455	1.481	0.032
Britain												
1	2.304	0.182	1.648	0.148	0.000	1.054	0.597	1.212	5.462	1.574	2.461	1.975
2	0.177	0.272	0.040	0.322	1.041	2.605	3.178	0.318	0.083	0.066	0.963	0.508
3	11.802	12.708	11.272	11.277	<u>17.870</u>	11.194	11.468	11.336	11.795	11.466	11.169	11.268
4	0.564	0.746	0.357	0.843	0.954	3.353	0.357	0.352	0.579	1.536	0.982	1.945
5	2.015	2.619	9.945	4.037	4.089	7.926	2.009	2.284	1.967	4.359	4.811	2.107
Japan												
1	1.034	<u>7.162</u>	3.900	0.841	1.596	1.284	4.005	2.915	1.073	0.881	1.201	1.172
2	6.321	3.399	4.913	3.427	4.220	3.549	3.751	4.500	4.915	4.192	5.241	4.146
3	3.096	0.753	0.629	1.942	2.766	3.990	2.202	0.925	1.972	0.696	0.624	1.891
4	0.055	0.365	1.940	0.007	1.228	1.296	0.071	0.010	1.101	0.130	0.133	3.472
5	1.051	0.919	2.782	1.132	3.940	2.278	2.540	0.614	6.607	1.548	0.788	1.234

But before the specific modeling, some general problems need to be addressed.

The first issue is how to treat the relationship between the linearity test and nonlinear estimation. The usual way is to perform the test and then estimate the model according to the information provided by the test. For example, Teräsvirta and Anderson [116] conduct an LM type test of linearity against the STAR alternative on quarterly logarithmic production indices for 13 countries and Europe. After the test rejects linearity for most of the series, they go ahead to estimate the data by STAR model. But is the nonlinear model specified as the alternative model the best nonlinear model when linearity is rejected? Here, we do not take this for granted, and will conduct the LM linearity test and the nonlinear model selection separately to see if they really match.

This naturally raises the question: what do we mean by the best model or the better model? In linear estimation, criteria such as AIC, SBC are widely used to pick the model. Here, the SBC is used for the model selection within AR models and across nonlinear models. The reason the SBC is preferred to the AIC is because the AIC tends to choose a less parsimonious model than the SBC does. It's commonly agreed that the RSS is not a good criterion to use in model selection because adding more parameters can always produce a smaller RSS. However, neither the AIC nor the SBC work very well when used to compare different types of models. The major reason is:

The SBC or AIC can be affected by a lot of factors external to the model. For instance, the number of regressors in the regression, which could change as the estimation method changes. For example, when estimating an EAR model, if a grid search method is used instead of nonlinear least squares (NLS) regression, the estimated models could be very similar, while the NLS method will have one more regressor than the grid search method. So it is possible that sometimes the similar estimated models will have different SBC and AIC values. Also, the magnitudes of the SBC, AIC and RSS will be affected if we transform the data, even though nothing real changes.

In Teräsvirta and Anderson's paper [116], the ratio of residual variances has been used to compare the AR model with the STAR model. We can see that comparing the residual variance ratio is nothing more than comparing RSS. And it is obviously unfair to compare an AR(6) model with at most 7

regressors with an LSTAR model with 13 parameters, and prefer the latter because the ratio of residual variances is less than one (p S125).

Therefore, in this chapter, the RSS and AIC will not be used as a model selection criterion, though the RSS and the AIC (adjusted) will be displayed in relevant tables. The SBC will be used but will be adjusted according to model type and estimation method. For example, in the ESTAR estimation of Canadian dollar, there are 5 regressors by the NLS: 3 for the AR terms, 1 for the ESTAR term, and 1 for the estimation of γ . In the calculation of the adjusted SBC and AIC, the number of regressors will be recorded as 4, omitting 1 parameter for γ .

4.5 Nonlinear Estimation

First, we estimate the data series as an ESTAR model in the following form³:

$$y_t = \sum_{i=1}^p \phi_i y_{t-i} + \left\{ 1 - \exp \left[-\gamma (y_{t-d} - c)^2 \right] \right\} \sum_{j=1}^q \theta_j y_{t-j} + \varepsilon_t. \quad (4.5)$$

Note that the autoregressive terms $\sum_{i=1}^p \phi_i y_{t-i}$ are those selected by the forward stepwise regression method in the linear estimation part. So an LM test with the null hypothesis $\gamma = 0$ can be easily constructed, with the alternative model as equation (4.5) and the null model actually the estimated autoregressive model. In equation (4.5), $\gamma \in (0, +\infty)$ controls the speed of the nonlinear adjustment in the form of $\left\{ 1 - \exp \left[-\gamma (y_{t-d} - c)^2 \right] \right\} \sum_{j=1}^q \theta_j y_{t-j}$. The variable y_{t-d} is the transition variable with d as the lag parameter. Here, we will choose the optimal value of d by SBC from the integers 1, 2, 3, 4, and 5.

In this paper, the estimation procedures will be conducted by the RATS (regression analysis of time series) software. The estimation methodology for ESTAR model is straightforward. Nonlinear least squares regressions of the dependent variable on the autoregressive terms and the possible exponential autoregressive terms are conducted, which under the RATS software, are conducted by the NLS

³Since the data is demeaned, (4.5) is used instead of

$$y_t = \mu_1 + \sum_{i=1}^p \phi_i y_{t-i} + \left\{ 1 - \exp \left[-\gamma (y_{t-d} - c)^2 \right] \right\} \left(\mu_2 + \sum_{j=1}^q \theta_j y_{t-j} \right) + \varepsilon_t$$

for estimation.

(nonlinear least squares) procedure using a Gauss-Newton algorithm with numerical partial derivatives. Then those ESTAR terms that produce the minimum SBC value are chosen. However, care must be taken with regard to the value of γ . Theoretically, γ can take any value from 0 to $+\infty$. However, if γ is too large the software treats the value of $\gamma(y_{t-d} - c)^2$ as a missing number. So, in the empirical programming, the upper limit of γ is obtained such that $\gamma(y_{t-d} - c)^2$ can be expressed by the software.

Since there are lagged error terms in the bilinear cross-product, maximum likelihood estimation is employed to estimate the bilinear process. The bilinear terms are chosen one by one by the forward selection method with SBC as the criterion. And for RATS, there are three algorithms available to calculate the maximum of a likelihood function: the Simplex algorithm, the Berndt-Hall-Hall-Hausman (BHHH) algorithm and the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm. The simplex algorithm is a search algorithm which does not require derivatives. The major disadvantage is that it can not provide standard errors for the estimated parameters. Both the BHHH and the BFGS algorithms assume that the likelihood function is twice differentiable. The BHHH algorithm is based on the result that the expected second derivatives matrix of the likelihood function is the covariance matrix of the first derivatives vectors of the likelihood function. The BFGS algorithm is a slight modification of the better known Davidon-Fletcher-Powell (DFP) algorithm. It is extremely effective and is among the most widely used gradient methods. So this paper will use the BFGS algorithm.

The estimated ESTAR and BL models are listed in Table 4.9. All the countries have lower values of SBC and AIC when their exchange rates are estimated as nonlinear ESTAR and BL models instead of the linear AR model.

The forward selection procedure with SBC as the criterion picks up the ESTAR model with the highest ESTAR statistic for Canada, France and Japan. The British pound is modeled as an ESTAR model with ESTAR equal 16.036, which is the second largest ESTAR statistic. Why there is such discrepancy between the test and the estimation? Estimation details of the ESTAR model for Great Britain are examined. The ESTAR model when $p = 9$ and $d = 3$ has a convergence problem. The final ESTAR model has the second largest ESTAR statistic 16.036, which is still very significant.

Table 4.9: Nonlinear estimation

Estimated ESTAR and BL Models					AIC	SBC	RSS
Canada							
$y_{1t} = 0.5247 \times y_{1t-1} + 0.1185 \times y_{1t-8} + 0.1223 \times y_{1t-10} - 0.5741 y_{1t-1} \times \{1 - \exp[-2.0775 \times (y_{1t-2} + 0.1580)^2]\} + \varepsilon_t$					1892.35	1907.50	323.83
(0.0001)	(0.0301)	(0.0258)	(0.0001)	(0.1685)	(0.2433)		
$y_{1t} = 0.1446 \times y_{1t-1} + 0.1461 \times y_{1t-8} + 0.1474 \times y_{1t-10} + 0.1653 \times y_{1t-10} \times \varepsilon_{1t-5} + \varepsilon_t$					1898.38	1913.52	329.87
(0.0040)	(0.0083)	(0.0084)	(0.0025)				
France							
$y_{2t} = 0.2561 \times y_{2t-1} + 1.0379 \times y_{2t-6} \times \{1 - \exp[-0.0117 \times (y_{2t-3} - 0.2561)^2]\} + \varepsilon_t$					2474.82	2482.40	1957.08
(0.0000)	(0.7117)	(0.7691)	(0.2875)				
$y_{2t} = 0.2766 \times y_{2t-1} - 0.0528 \times y_{2t-4} \times \varepsilon_{2t-7} + \varepsilon_t$					2477.15	2484.72	1971.08
(0.0000)	(0.0086)						
Britain							
$y_{3t} = 0.3950 \times y_{3t-1} - 0.1343 \times y_{3t-2} - 1.5055 \times y_{3t-10} \times \{1 - \exp[-0.0093 \times (y_{3t-3} + 0.1783)^2]\} + \varepsilon_t$					2437.09	2448.45	1732.49
(0.0000)	(0.0143)	(0.5474)	(0.6413)	(0.8348)			
$y_{3t} = 0.4054 \times y_{3t-1} - 0.1069 \times y_{3t-2} + 0.0915 \times y_{3t-3} \times \varepsilon_{3t-5} - 0.0731 \times y_{3t-7} \times \varepsilon_{3t-12} - 0.0591 \times y_{3t-5} \times \varepsilon_{3t-6} + \varepsilon_t$					2422.97	2441.91	1638.87
(0.0000)	(0.0180)	(0.0002)	(0.0001)	(0.0003)			
Japan							
$y_{4t} = 0.4625 \times y_{4t-1} + 0.1509 \times y_{4t-11} - 0.4522 y_{4t-2} \times \{1 - \exp[-0.0844 \times (y_{4t-1} - 0.5515)^2]\} + \varepsilon_t$					2508.94	2520.30	2159.72
(0.0000)	(0.0029)	(0.0043)	(0.2543)	(0.4527)			
$y_{4t} = 0.3942 \times y_{4t-1} + 0.1338 \times y_{4t-11} + 0.0688 \times y_{4t-1} \times \varepsilon_{4t-2} + 0.0706 \times y_{4t-1} \times \varepsilon_{4t-7} - 0.0530 \times y_{4t-5} \times \varepsilon_{4t-9} + \varepsilon_t$					2504.53	2523.46	2104.70
(0.0000)	(0.0102)	(0.0001)	(0.0003)	(0.0021)			

Teräsvirta [115] mentions that it is sometimes difficult for a nonlinear optimization procedure to converge while trying to estimate the ESTAR model parameters simultaneously. The reason is that the estimator of γ tends to be very heavily negatively correlated with that of θ . He suggests to standardize the nonlinear part $\left\{1 - \exp\left[-\gamma(y_{t-d} - c)^2\right]\right\}$ by dividing it by the variance of the dependent variable y . Haggan and Ozaki [57] propose a grid search method by keeping γ fixed and using a grid of values to estimate it. We tried both methods for British pound. There was no convergence using Teräsvirta's method, while for grid search, the model picks up either the starting value or the ending value of γ , which implies a problem of convergence again. Maybe we have to accept the fact that the ESTAR model is not sufficiently well specified for the British pound data.

4.6 Forecasting

In an effort to find a counter example of Meese and Rogoff's conclusion [81], and to show the possible superiority of nonlinear models over linear models of monthly nominal exchange rates, the in-sample and out-of-sample mean square prediction error for the nonlinear ESTAR and BL model are calculated and compared to those of the random walk without drift (RW) model, random walk with drift (RWD) model and linear AR model. The MSPE are calculated for forecasts from 1-step ahead up to 12-step ahead, which is appropriate for monthly exchange rate data. Since our linear and nonlinear models are estimated with the transformed data $y_{it} = 100 \times \{\log x_{it} - \log x_{it-1}\} - c_i$ with $i = 1, 2, 3, 4$, while the random walk specifications are based on the undifferentiated data, a transformation is needed so that the MSPEs are comparable. Suppose $e_{it} = 100 \times \log x_{it}$, then for the random walk model without drift, the k -step ahead forecasts at time t will be

$$\hat{e}_{i,t+k|t} = e_{i,t}.$$

No matter how far away the forecast in the future, the random walk model always predicts the forecasts will be the same as today's exchange rate. If a drift term d_i is included in the random walk model, the forecasts are calculated by

$$\hat{e}_{i,t+k|t} = e_{i,t} + kd_i.$$

For the AR, ESTAR and BL model, if the k -step-ahead forecasts are expressed as $\widehat{y}_{i,t+k|t}$, then the forecasts of the magnified logarithm of the exchange rates are

$$\widehat{e}_{i,t+k|t} = e_{i,t} + (\widehat{y}_{i,t+1|t} + d_i) + (\widehat{y}_{i,t+2|t} + d_i) + \cdots + (\widehat{y}_{i,t+k|t} + d_i).$$

For the out-of-sample forecasts, the last five years are set aside for forecasting, and the data from 1973:2 to 1996:3 are used for model estimation. Correspondent changes are made to the data means c_i , $i = 1, 2, 3, 4$ and 5 in the data analysis procedure. The MSPEs of the in-sample and out-of-sample forecasts are shown in Tables 4.11 and 4.10, respectively. For each step, the minimum MSPE for the five models is underlined.

When considering in-sample forecasts, the nonlinear models seems to do a relatively better job than the random walk models and the linear model. For Canada, the ESTAR model has 8 minimum MSPEs out of 12 MSPEs; the ESTAR model has the minimum MSPE for all 12 forecasts predicting French francs; when forecasting the British pound, the BL model is the best for 1-step and 2-step ahead forecasting, while the ESTAR model is the best for forecasts from 4-step ahead up to 10-step ahead; for Japan, the BL model produces the minimum MSPE at all forecast horizons.

However, the nonlinear models are not always the best for ex-ante forecasts. For Canada, the random walk model with drift has a better forecast than any other model in 9 out of 12 cases. For the remaining three countries, the nonlinear models pick up most of the minimum MSPE: it is the ESTAR model for France and the BL model for Britain and Japan.

We also provide the Diebold-Mariano (D-M) [33] test statistics which test the hypothesis that the median of the squared forecast errors difference is zero:

$$\text{Median}(d_t) = \text{Median}(e_{A_t}^2 - e_{B_t}^2) = 0,$$

where A and B are two forecasts, $e_{A_t}^2$ and $e_{B_t}^2$ are the squared forecast errors for them at time t .

The D-M test statistic can be calculated according to the formula

Table 4.10: MSPEs for in-sample forecasts

Model\Step	1	2	3	4	5	6	7	8	9	10	11	12
Canada												
RW	1.129	2.627	4.065	5.590	7.264	8.907	10.550	12.311	14.428	16.724	19.439	22.423
RWD	1.110	2.551	3.896	5.285	6.774	8.200	<u>9.599</u>	<u>11.084</u>	12.894	14.844	17.169	19.742
AR	1.041	2.122	<u>3.331</u>	4.842	6.520	8.270	<u>10.023</u>	<u>11.573</u>	<u>11.230</u>	12.200	11.259	12.689
BL	1.012	2.116	<u>3.403</u>	5.133	6.819	8.723	10.612	12.273	11.919	13.006	12.008	13.377
ESTAR	<u>0.993</u>	<u>2.078</u>	<u>3.358</u>	<u>4.817</u>	<u>6.467</u>	<u>8.156</u>	9.941	11.526	11.398	<u>11.962</u>	<u>10.989</u>	<u>11.603</u>
France												
RW	6.707	17.392	28.768	41.866	55.319	69.628	84.696	100.171	116.553	133.601	152.035	170.885
RWD	6.694	17.343	28.653	41.651	54.956	69.079	83.920	99.148	115.272	132.031	150.102	168.555
AR	6.166	12.854	21.920	34.096	46.701	60.514	75.042	90.022	106.003	122.477	140.287	158.342
BL	6.046	12.673	22.110	34.837	47.738	61.329	75.975	91.717	107.445	124.114	141.595	159.279
ESTAR	<u>6.003</u>	<u>12.581</u>	<u>21.427</u>	<u>32.381</u>	<u>44.027</u>	<u>55.921</u>	<u>64.903</u>	<u>76.760</u>	<u>91.612</u>	<u>105.890</u>	<u>122.237</u>	<u>140.005</u>
Britain												
RW	6.442	17.490	28.630	40.485	52.582	64.432	76.385	88.179	100.526	113.447	126.614	140.357
RWD	6.425	17.417	28.459	40.164	52.046	63.640	75.279	86.744	98.739	111.261	123.988	137.301
AR	5.537	12.067	<u>23.582</u>	37.396	49.802	61.417	73.161	84.567	96.609	108.888	<u>121.464</u>	<u>134.830</u>
BL	5.027	<u>11.781</u>	<u>24.511</u>	38.435	51.702	64.623	79.559	93.208	106.594	122.451	141.028	157.757
ESTAR	5.314	11.974	24.224	<u>36.977</u>	<u>48.857</u>	<u>60.529</u>	<u>71.619</u>	<u>82.510</u>	<u>94.568</u>	<u>108.833</u>	126.451	142.986
Japan												
RW	8.231	22.327	37.246	53.521	70.123	86.008	100.332	114.592	130.903	148.419	166.955	187.982
RWD	8.155	22.015	36.553	52.276	68.136	83.079	96.222	109.058	123.744	139.382	155.845	174.518
AR	7.025	15.006	25.416	39.226	54.831	72.022	88.921	105.530	122.665	140.557	151.250	142.082
BL	<u>6.456</u>	<u>14.068</u>	<u>24.957</u>	<u>38.750</u>	<u>54.164</u>	<u>69.619</u>	<u>84.951</u>	<u>99.673</u>	<u>114.049</u>	<u>130.178</u>	<u>141.892</u>	<u>134.690</u>
ESTAR	6.625	14.988	29.506	45.439	62.045	80.161	98.309	115.680	133.353	152.225	163.723	152.493

Table 4.11: MSPEs for out-of-sample forecasts

Model\Step	1	2	3	4	5	6	7	8	9	10	11	12
Canada												
RW	1.399	2.923	4.140	5.459	7.164	8.851	10.048	11.574	13.137	14.803	16.526	18.164
RWD	<u>1.358</u>	<u>2.787</u>	<u>3.864</u>	<u>4.973</u>	<u>6.357</u>	<u>7.678</u>	<u>8.507</u>	<u>9.637</u>	<u>10.790</u>	<u>12.062</u>	<u>13.247</u>	<u>14.450</u>
AR	<u>1.501</u>	<u>2.618</u>	<u>4.064</u>	<u>5.650</u>	<u>8.093</u>	<u>10.531</u>	<u>12.443</u>	<u>14.939</u>	<u>13.633</u>	<u>15.159</u>	<u>11.680</u>	<u>12.074</u>
BL	1.482	<u>2.527</u>	3.871	5.329	7.581	9.823	11.552	13.801	12.716	14.027	<u>11.009</u>	11.390
ESTAR	1.453	2.650	4.092	5.577	7.896	10.293	12.179	14.621	13.773	14.848	11.699	<u>11.179</u>
France												
RW	5.758	14.967	23.453	31.202	37.726	46.053	56.086	66.194	76.044	84.153	94.133	105.761
RWD	5.645	14.523	22.473	29.398	34.740	41.611	49.998	58.275	66.081	72.016	79.480	88.413
AR	5.175	10.893	17.105	23.717	29.228	36.426	44.470	52.364	60.236	66.379	74.280	82.989
BL	<u>5.171</u>	<u>10.866</u>	<u>17.049</u>	<u>23.637</u>	<u>29.123</u>	<u>36.304</u>	<u>44.381</u>	<u>52.326</u>	60.183	66.276	74.195	82.929
ESTAR	<u>5.563</u>	<u>12.550</u>	<u>18.561</u>	<u>20.010</u>	<u>22.258</u>	<u>27.683</u>	<u>34.934</u>	<u>42.238</u>	<u>48.639</u>	<u>53.740</u>	<u>62.069</u>	<u>70.605</u>
Britain												
RW	3.106	6.201	8.217	10.239	12.447	15.666	19.582	22.465	24.768	26.365	27.829	29.018
RWD	<u>3.100</u>	<u>6.177</u>	<u>8.193</u>	<u>10.190</u>	<u>12.266</u>	<u>15.331</u>	<u>19.053</u>	<u>21.830</u>	<u>23.997</u>	<u>25.452</u>	<u>26.784</u>	<u>28.042</u>
AR	<u>3.634</u>	3.952	6.730	9.875	12.051	14.950	18.214	20.779	23.549	25.282	26.749	27.871
BL	3.686	<u>3.911</u>	<u>6.238</u>	<u>9.124</u>	<u>11.300</u>	<u>14.193</u>	<u>17.421</u>	<u>19.896</u>	<u>22.585</u>	<u>24.317</u>	<u>25.847</u>	<u>27.075</u>
ESTAR	3.594	3.931	6.457	9.164	11.428	14.350	17.601	20.373	23.122	24.783	26.983	28.340
Japan												
RW	9.008	23.296	38.659	55.806	68.185	76.026	80.176	86.357	<u>100.792</u>	<u>119.274</u>	<u>140.936</u>	167.280
RWD	9.204	24.002	40.205	58.503	72.227	81.703	87.690	96.468	114.558	137.324	163.978	195.793
AR	8.198	16.407	28.116	44.662	57.047	67.338	75.125	<u>85.360</u>	103.734	125.156	150.135	180.439
BL	7.994	<u>14.636</u>	<u>23.594</u>	<u>37.468</u>	<u>48.186</u>	<u>59.109</u>	<u>70.965</u>	85.611	105.948	128.158	145.900	<u>150.254</u>
ESTAR	<u>7.258</u>	15.226	29.625	45.802	57.787	70.258	82.951	99.064	118.620	140.928	156.208	150.501

$$\frac{S_3 - \frac{T(T+1)}{4}}{\sqrt{\frac{T(T+1)(2T+1)}{24}}}$$

where T is the forecast horizon, $S_3 = \sum_{t=1}^T I_+(d_t) \times \text{rank}(|d_t|)$ with $I_+(d_t) = \begin{cases} 1 & \text{if } d_t > 0 \\ 0 & \text{if } d_t \leq 0 \end{cases}$. The D-M test statistic is asymptotically distributed standard normal, so at significance level 5%, the critical value of the D-M test is ± 1.96 .

Tables 4.12 and 4.13 include the D-M test statistics when forecast A is forecast from the RWD, AR, BL, or ESTAR model and the forecast B is from the RW model. Tables 4.14 and 4.15 list the D-M statistics testing the median of the squared forecast errors difference between forecasts from the AR, BL, or ESTAR model and forecasts from the RWD model. Statistics with absolute values that are smaller than 1.96 are marked with underlines. From these two tables, we can see that many forecast differences found in Tables 4.11 and 4.10 are insignificant. For example, the RWD model is not more accurate than the linear or nonlinear models for ex-ante forecasting of the Canada dollar; and for the British pound, the superiority of the ESTAR model in in-sample forecasting and the BL model in out-of-sample forecasting is not that solid after the forecasts by these two nonlinear models are tested whether or not different from the forecasts from the random walk type models.

Here we establish a counter example for Meese and Rogoff⁴. For the out-of-sample forecasting accuracy measured by the traditional MSPE, the random walk is not always the best. And the random walk model without drift is not always better than the random walk model with drift. However, although the foreign exchange rate data for the four countries are tested and estimated to have nonlinearities, there is no unanimous conclusion to be drawn in favor of the nonlinear model either when it comes to the forecasting. Which will win the battle in out-of sample forecasting? The ESTAR model? The BL model? The random walk model? Or the AR model? There is simply no conclusive conclusion to make. It depends on the country, on the forecast horizon, on the data length, on the measurement, etc. Maybe that is why the topic is still attractive after so many years.

⁴Maybe not exactly for the example Meese and Rogoff give in their 1983 paper, since we use monthly average data and they use monthly point-sample data.

Table 4.12: D-M statistics for in-sample forecasts with RW as the base model

Model\Step	1	2	3	4	5	6	7	8	9	10	11	12
Canada												
RWD_RW	<u>-0.969</u>	<u>-1.250</u>	<u>-1.513</u>	<u>-1.829</u>	<u>-2.247</u>	<u>-2.591</u>	<u>-2.860</u>	<u>-3.069</u>	<u>-3.161</u>	<u>-3.295</u>	<u>-3.324</u>	<u>-3.277</u>
AR_RW	<u>-2.392</u>	<u>-4.627</u>	<u>-3.920</u>	<u>-2.331</u>	<u>-2.142</u>	<u>-1.733</u>	<u>-1.209</u>	<u>-1.533</u>	<u>-3.890</u>	<u>-4.932</u>	<u>-7.351</u>	<u>-7.570</u>
BL_RW	<u>-3.025</u>	<u>-3.499</u>	<u>-2.580</u>	<u>-0.579</u>	<u>-0.704</u>	<u>0.144</u>	<u>0.191</u>	<u>-0.161</u>	<u>-2.771</u>	<u>-3.715</u>	<u>-6.158</u>	<u>-6.454</u>
ESTAR_RW	<u>-3.335</u>	<u>-5.085</u>	<u>-4.193</u>	<u>-2.995</u>	<u>-2.795</u>	<u>-2.231</u>	<u>-1.802</u>	<u>-2.013</u>	<u>-3.899</u>	<u>-5.196</u>	<u>-7.612</u>	<u>-8.673</u>
France												
RWD_RW	<u>-0.504</u>	<u>-0.370</u>	<u>-0.372</u>	<u>-0.481</u>	<u>-0.620</u>	<u>-0.799</u>	<u>-0.990</u>	<u>-1.023</u>	<u>-1.103</u>	<u>-1.227</u>	<u>-1.352</u>	<u>-1.456</u>
AR_RW	<u>-3.077</u>	<u>-10.929</u>	<u>-12.149</u>	<u>-10.355</u>	<u>-8.881</u>	<u>-7.641</u>	<u>-6.884</u>	<u>-6.546</u>	<u>-5.962</u>	<u>-5.249</u>	<u>-4.726</u>	<u>-4.448</u>
BL_RW	<u>-3.657</u>	<u>-9.911</u>	<u>-9.226</u>	<u>-6.538</u>	<u>-5.482</u>	<u>-5.303</u>	<u>-5.213</u>	<u>-5.101</u>	<u>-4.934</u>	<u>-4.599</u>	<u>-4.370</u>	<u>-4.233</u>
ESTAR_RW	<u>-3.313</u>	<u>-9.924</u>	<u>-9.594</u>	<u>-8.315</u>	<u>-8.204</u>	<u>-8.388</u>	<u>-9.932</u>	<u>-10.628</u>	<u>-10.246</u>	<u>-9.676</u>	<u>-9.433</u>	<u>-8.936</u>
Britain												
RWD_RW	<u>-0.189</u>	<u>-0.394</u>	<u>-0.372</u>	<u>-0.255</u>	<u>-0.197</u>	<u>-0.255</u>	<u>-0.444</u>	<u>-0.591</u>	<u>-0.703</u>	<u>-0.835</u>	<u>-0.904</u>	<u>-1.007</u>
AR_RW	<u>-3.275</u>	<u>-11.100</u>	<u>-6.759</u>	<u>-2.804</u>	<u>-1.482</u>	<u>-0.799</u>	<u>-1.026</u>	<u>-1.109</u>	<u>-0.974</u>	<u>-1.014</u>	<u>-0.926</u>	<u>-0.987</u>
BL_RW	<u>-2.828</u>	<u>-6.515</u>	<u>-3.566</u>	<u>-1.232</u>	<u>0.022</u>	<u>0.454</u>	<u>0.363</u>	<u>0.314</u>	<u>0.644</u>	<u>0.958</u>	<u>1.190</u>	<u>1.262</u>
ESTAR_RW	<u>-3.301</u>	<u>-9.809</u>	<u>-5.475</u>	<u>-2.534</u>	<u>-1.698</u>	<u>-1.367</u>	<u>-1.499</u>	<u>-1.274</u>	<u>-1.067</u>	<u>-0.796</u>	<u>0.100</u>	<u>0.556</u>
Japan												
RWD_RW	<u>0.071</u>	<u>-0.252</u>	<u>-0.534</u>	<u>-0.709</u>	<u>-0.894</u>	<u>-1.137</u>	<u>-1.419</u>	<u>-1.732</u>	<u>-1.929</u>	<u>-2.123</u>	<u>-2.230</u>	<u>-2.378</u>
AR_RW	<u>-2.611</u>	<u>-8.595</u>	<u>-8.400</u>	<u>-7.231</u>	<u>-5.484</u>	<u>-3.905</u>	<u>-2.955</u>	<u>-2.307</u>	<u>-1.981</u>	<u>-2.149</u>	<u>-2.554</u>	<u>-4.760</u>
BL_RW	<u>-3.578</u>	<u>-7.894</u>	<u>-7.558</u>	<u>-6.524</u>	<u>-5.411</u>	<u>-4.243</u>	<u>-3.203</u>	<u>-2.975</u>	<u>-2.995</u>	<u>-3.109</u>	<u>-3.420</u>	<u>-5.321</u>
ESTAR_RW	<u>-2.727</u>	<u>-8.375</u>	<u>-6.746</u>	<u>-5.096</u>	<u>-3.542</u>	<u>-2.101</u>	<u>-1.281</u>	<u>-0.831</u>	<u>-0.578</u>	<u>-0.877</u>	<u>-1.346</u>	<u>-3.739</u>

Table 4.13: D-M statistics for out-of-sample forecasts with RW as the base model

Model\Step	1	2	3	4	5	6	7	8	9	10	11	12
Canada												
RWD_RW	<u>-0.906</u>	<u>-1.141</u>	<u>-1.001</u>	<u>-1.325</u>	<u>-1.642</u>	<u>-1.944</u>	<u>-2.039</u>	<u>-2.025</u>	<u>-1.958</u>	<u>-2.025</u>	<u>-2.047</u>	<u>-2.054</u>
AR_RW	<u>0.501</u>	<u>-1.701</u>	<u>-1.237</u>	<u>-0.685</u>	<u>-0.486</u>	<u>-0.295</u>	<u>-0.059</u>	<u>-0.037</u>	<u>-0.994</u>	<u>-0.758</u>	<u>-2.172</u>	<u>-1.958</u>
BL_RW	<u>0.398</u>	<u>-1.988</u>	<u>-1.450</u>	<u>-0.979</u>	<u>-0.832</u>	<u>-0.530</u>	<u>-0.302</u>	<u>-0.376</u>	<u>-1.296</u>	<u>-0.920</u>	<u>-2.429</u>	<u>-2.319</u>
ESTAR_RW	<u>0.442</u>	<u>-1.796</u>	<u>-1.568</u>	<u>-1.163</u>	<u>-0.736</u>	<u>-0.302</u>	<u>-0.037</u>	<u>0.155</u>	<u>-0.766</u>	<u>-0.766</u>	<u>-2.032</u>	<u>-2.312</u>
France												
RWD_RW	<u>-2.142</u>	<u>-2.459</u>	<u>-2.716</u>	<u>-3.033</u>	<u>-3.673</u>	<u>-4.240</u>	<u>-4.763</u>	<u>-4.829</u>	<u>-5.013</u>	<u>-5.249</u>	<u>-5.396</u>	<u>-5.543</u>
AR_RW	<u>-1.487</u>	<u>-4.925</u>	<u>-5.293</u>	<u>-5.050</u>	<u>-4.866</u>	<u>-4.594</u>	<u>-4.756</u>	<u>-5.242</u>	<u>-5.588</u>	<u>-5.595</u>	<u>-5.448</u>	<u>-5.455</u>
BL_RW	<u>-1.472</u>	<u>-4.932</u>	<u>-5.323</u>	<u>-5.087</u>	<u>-4.851</u>	<u>-4.572</u>	<u>-4.719</u>	<u>-5.242</u>	<u>-5.580</u>	<u>-5.588</u>	<u>-5.462</u>	<u>-5.477</u>
ESTAR_RW	<u>-0.832</u>	<u>-2.002</u>	<u>-2.555</u>	<u>-4.837</u>	<u>-5.661</u>	<u>-5.477</u>	<u>-5.381</u>	<u>-5.433</u>	<u>-5.669</u>	<u>-5.956</u>	<u>-5.801</u>	<u>-5.985</u>
Britain												
RWD_RW	<u>0.015</u>	<u>-0.368</u>	<u>-0.295</u>	<u>-0.169</u>	<u>-0.110</u>	<u>-0.169</u>	<u>-0.191</u>	<u>-0.103</u>	<u>-0.199</u>	<u>-0.221</u>	<u>-0.265</u>	<u>-0.272</u>
AR_RW	<u>0.979</u>	<u>-3.696</u>	<u>-2.260</u>	<u>-0.641</u>	<u>-0.169</u>	<u>-0.236</u>	<u>-0.950</u>	<u>-0.633</u>	<u>-0.368</u>	<u>-0.611</u>	<u>-0.353</u>	<u>-0.376</u>
BL_RW	<u>1.112</u>	<u>-3.674</u>	<u>-3.423</u>	<u>-1.590</u>	<u>-0.825</u>	<u>-0.633</u>	<u>-1.546</u>	<u>-1.134</u>	<u>-0.722</u>	<u>-0.825</u>	<u>-0.641</u>	<u>-0.626</u>
ESTAR_RW	<u>0.891</u>	<u>-3.755</u>	<u>-2.753</u>	<u>-1.031</u>	<u>-0.236</u>	<u>-0.118</u>	<u>-0.707</u>	<u>-0.324</u>	<u>-0.184</u>	<u>-0.339</u>	<u>-0.066</u>	<u>-0.052</u>
Japan												
RWD_RW	<u>1.605</u>	<u>1.980</u>	<u>2.113</u>	<u>2.128</u>	<u>2.047</u>	<u>2.113</u>	<u>2.069</u>	<u>2.017</u>	<u>1.973</u>	<u>2.054</u>	<u>2.260</u>	<u>2.665</u>
AR_RW	<u>0.221</u>	<u>-3.777</u>	<u>-2.452</u>	<u>-1.318</u>	<u>-0.243</u>	<u>0.957</u>	<u>1.414</u>	<u>1.141</u>	<u>0.920</u>	<u>0.714</u>	<u>0.883</u>	<u>1.509</u>
BL_RW	<u>-0.133</u>	<u>-3.651</u>	<u>-4.152</u>	<u>-3.593</u>	<u>-2.768</u>	<u>-1.480</u>	<u>-0.552</u>	<u>0.052</u>	<u>-0.020</u>	<u>0.110</u>	<u>-0.206</u>	<u>-0.898</u>
ESTAR_RW	<u>-0.501</u>	<u>-3.401</u>	<u>-3.269</u>	<u>-2.371</u>	<u>-1.605</u>	<u>-0.442</u>	<u>0.250</u>	<u>0.722</u>	<u>0.876</u>	<u>1.001</u>	<u>0.523</u>	<u>-0.758</u>

Table 4.14: D-M statistics for in-sample forecasts with RWD as the base model

Model\Step	1	2	3	4	5	6	7	8	9	10	11	12
Canada												
AR.RWD	-2.642	-4.897	-4.381	-2.392	<u>-1.525</u>	<u>-0.844</u>	<u>-0.009</u>	<u>-0.280</u>	-2.844	-3.965	-6.907	-7.362
BL.RWD	-2.595	-3.844	-2.744	<u>-0.238</u>	<u>-0.332</u>	<u>0.252</u>	<u>0.758</u>	<u>0.473</u>	-2.313	-3.359	-6.076	-6.512
ESTAR.RWD	-3.102	-5.016	-4.564	-2.518	<u>-1.513</u>	<u>-1.679</u>	<u>0.106</u>	<u>-0.140</u>	-2.458	-3.898	-6.721	-7.702
France												
AR.RWD	-3.198	-11.303	-13.203	-12.275	-11.497	-10.325	-9.738	-9.460	-9.070	-8.571	-8.166	-7.959
BL.RWD	-3.820	-10.046	-9.763	-7.259	-6.333	-5.786	-5.773	-5.595	-5.551	-5.382	-5.299	-5.172
ESTAR.RWD	-3.355	-9.932	-9.919	-8.637	-8.273	-8.496	-10.358	-11.409	-11.268	-11.001	-10.885	-10.599
Britain												
AR.RWD	-2.991	-11.281	-7.384	-3.454	-2.263	-2.017	-2.428	-2.397	-2.228	-2.403	-2.336	-2.244
BL.RWD	-2.706	-6.330	-3.175	<u>-1.403</u>	<u>-0.015</u>	<u>0.380</u>	<u>0.258</u>	<u>0.316</u>	<u>0.869</u>	<u>1.417</u>	<u>1.593</u>	<u>1.783</u>
ESTAR.RWD	-3.292	-10.118	-5.991	-3.450	-2.799	-2.522	-2.485	-2.294	-2.268	-2.166	<u>-0.667</u>	<u>0.326</u>
Japan												
AR.RWD	-2.841	-9.378	-9.436	-7.853	-6.039	-3.883	-2.715	<u>-1.897</u>	<u>-1.191</u>	<u>-1.133</u>	<u>-1.575</u>	-4.217
BL.RWD	-3.865	-8.144	-7.666	-6.479	-5.112	-3.544	-2.735	-2.187	-1.995	-2.076	-2.483	-4.859
ESTAR.RWD	-3.122	-9.031	-7.007	-5.179	-3.682	<u>-1.779</u>	<u>-0.822</u>	<u>-0.241</u>	<u>0.303</u>	<u>0.206</u>	<u>-0.257</u>	-2.816

Table 4.15: D-M statistics for out-of-sample forecasts with RWD as the base model

Model\Step	1	2	3	4	5	6	7	8	9	10	11	12
Canada												
AR.RWD	<u>1.045</u>	<u>-1.296</u>	<u>-1.068</u>	<u>-0.199</u>	<u>0.287</u>	<u>0.979</u>	<u>1.097</u>	<u>1.274</u>	<u>0.169</u>	<u>0.376</u>	<u>-1.671</u>	<u>-1.487</u>
BL.RWD	<u>0.913</u>	<u>-1.553</u>	<u>-1.288</u>	<u>-0.604</u>	<u>-0.110</u>	<u>0.825</u>	<u>0.942</u>	<u>1.068</u>	<u>-0.110</u>	<u>-0.015</u>	<u>-1.760</u>	<u>-1.782</u>
ESTAR.RWD	<u>0.802</u>	<u>-1.443</u>	<u>-1.119</u>	<u>-0.390</u>	<u>0.162</u>	<u>1.148</u>	<u>0.957</u>	<u>1.082</u>	<u>0.250</u>	<u>0.368</u>	<u>-1.465</u>	<u>-1.789</u>
France												
AR.RWD	<u>-1.414</u>	<u>-4.874</u>	<u>-5.404</u>	<u>-5.293</u>	<u>-4.837</u>	<u>-3.769</u>	<u>-3.210</u>	<u>-3.585</u>	<u>-3.710</u>	<u>-3.431</u>	<u>-3.224</u>	<u>-2.908</u>
BL.RWD	<u>-1.406</u>	<u>-4.918</u>	<u>-5.462</u>	<u>-5.300</u>	<u>-4.866</u>	<u>-3.813</u>	<u>-3.173</u>	<u>-3.578</u>	<u>-3.666</u>	<u>-3.409</u>	<u>-3.188</u>	<u>-2.915</u>
ESTAR.RWD	<u>-0.795</u>	<u>-2.142</u>	<u>-2.547</u>	<u>-4.815</u>	<u>-5.669</u>	<u>-5.330</u>	<u>-4.667</u>	<u>-4.528</u>	<u>-4.859</u>	<u>-5.131</u>	<u>-4.866</u>	<u>-4.719</u>
Britain												
AR.RWD	<u>0.979</u>	<u>-3.718</u>	<u>-2.172</u>	<u>-0.420</u>	<u>-0.110</u>	<u>-0.103</u>	<u>-0.832</u>	<u>-0.405</u>	<u>0.081</u>	<u>-0.515</u>	<u>-0.250</u>	<u>0.000</u>
BL.RWD	<u>1.229</u>	<u>-3.659</u>	<u>-2.790</u>	<u>-1.112</u>	<u>-0.567</u>	<u>-0.398</u>	<u>-1.325</u>	<u>-0.906</u>	<u>-0.471</u>	<u>-0.876</u>	<u>-0.773</u>	<u>-0.545</u>
ESTAR.RWD	<u>0.891</u>	<u>-3.975</u>	<u>-2.621</u>	<u>-1.252</u>	<u>-0.501</u>	<u>-0.471</u>	<u>-0.729</u>	<u>-0.206</u>	<u>0.169</u>	<u>-0.317</u>	<u>0.147</u>	<u>0.611</u>
Japan												
AR.RWD	<u>-0.618</u>	<u>-5.190</u>	<u>-5.750</u>	<u>-5.926</u>	<u>-6.022</u>	<u>-5.205</u>	<u>-4.469</u>	<u>-4.262</u>	<u>-3.939</u>	<u>-3.968</u>	<u>-4.255</u>	<u>-4.189</u>
BL.RWD	<u>-0.663</u>	<u>-4.550</u>	<u>-5.050</u>	<u>-5.080</u>	<u>-4.741</u>	<u>-3.732</u>	<u>-2.584</u>	<u>-1.855</u>	<u>-1.634</u>	<u>-1.274</u>	<u>-1.737</u>	<u>-2.746</u>
ESTAR.RWD	<u>-1.053</u>	<u>-4.285</u>	<u>-3.747</u>	<u>-3.563</u>	<u>-3.033</u>	<u>-1.944</u>	<u>-0.972</u>	<u>-0.611</u>	<u>-0.442</u>	<u>-0.346</u>	<u>-0.825</u>	<u>-2.186</u>

BIBLIOGRAPHY

- [1] Aitchison, J., and Silvey, S. (1958). "Maximum-likelihood Estimation of Parameters Subject to Restraints". *Annals of Mathematical Statistics*, 29: 813-828.
- [2] Aitchison, J., and Silvey, S. (1960). "Maximum-likelihood Estimation Procedures and Associated Tests of Significance". *Journal of the Royal Statistical Society, Series B*, 22: 154-171.
- [3] Ashley, R., Patterson, D., and Hinich, M. (1986). "A Diagnostic Test for Non-linear Serial Dependence in Time Series Fitting Errors". *Journal of Time Series Analysis*, 7: 165-178.
- [4] Bacon, D., and Watts, D. (1971). "Estimating the Transition Between Two Intersecting Straight Lines". *Biometrika*, 58: 525-534.
- [5] Baillie, R., and Bollerslev, T. (1989). "The Message in Daily Exchange Rates". *Journal of Economics and Business Statistics*, 7: 297-305.
- [6] Balke, N., and Fomby, T. (1997). "Threshold Cointegration". *International Economic Review*, 38: 627-643.
- [7] Barnett, W., Medio, M., and Serletis, A. (1997). "Nonlinear and Complex Dynamics in Economics". Washington University in St. Louis, *Economics Working Paper Archive*, no. WAB-97-13 (24 September).
- [8] Baum, C., Barkoulas, J., and Caglayan, M. (2001). "Nonlinear Adjustment to Purchasing Power Parity in the Post-Bretton Woods Era". *Journal of International Money and Finance*, 20: 379-399.
- [9] Beaudry, P., and Koop, G. (1993). "Do Recessions Permanently Change Output?". *Journal of Monetary Economics*, 31: 149-163.
- [10] Bollerslev, T. (1986). "Generalized Autoregressive Conditional Heteroskedasticity". *Journal of Econometrics*, 31: 307-327.

- [11] Bollerslev, T., Chou, R., and Kroner, K. (1992). "ARCH Modeling in Finance". *Journal of Econometrics*, 52: 5-59.
- [12] Bollerslev, T., Engle, R., and Nelson, D. (1994). "ARCH Models". *Handbook of Econometrics*, Vol. IV, 2959-3038, Engle, R., and McFadden, D. (Eds.). North-holland, Amsterdam.
- [13] Breusch, T. (1980). "Useful Invariance Results for Generalized Regression Models". *Journal of Econometrics*, 13: 327-340.
- [14] Breusch, T., and Pagan, A. (1979). "A Simple Test for Heteroskedasticity and Random Coefficient Variation". *Econometrica*, 47: 1287-1294.
- [15] Brock, W., Dechert, W., and Scheinkman, J. (1987). "A Test for Independence Based on the Correlation Dimension". University of Wisconsin at Madison, *Social Systems Research Institute Working Paper*, no. 8702; reprinted in *Dynamic Econometric Modelling: Proceedings of the Third International Symposium on Economic Theory and Econometrics, 1988*, Barnett, W., Berndt, E., and White, H. (Eds). Cambridge University Press, Cambridge.
- [16] Brock, W., Dechert, W., Scheinkman, J., and LeBaron, B. (1996). "A Test for Independence Based on the Correlation Dimension". *Econometric Review*, 15: 197-235.
- [17] Brock, W., and Hommes, C. (1998). "Heterogeneous Beliefs and Routes to Chaos in a Simple Asset Pricing Model". *Journal of Economic Dynamics and Control*, 22: 1235-1274.
- [18] Brock, W., Hsieh, D., and LeBaron, B. (1991). *Nonlinear Dynamics, Chaos, and Instabilities: Statistical Theory and Economic Evidence*. M.I.T. Press, Reading, Massachusetts.
- [19] Brock, W., and Sayers, C. (1988). "Is the Business Cycle Characterized by Deterministic Chaos?". *Journal of Monetary Economics*, 22: 71-90.
- [20] Brockett, R. (1976). "Volterra Series and Geometric Control Theory". *Automatica*, 12: 61-72.
- [21] Burgess, S. (1992). "Nonlinear Dynamics in a Structural Model of Employment". *Journal of Applied Econometrics*, 7: 101-118.
- [22] Byers, J., and Peel, D. (1995). "Bilinear Quadratic ARCH and Volatility Spillovers in Inter-war Exchange Rates". *Applied Economics Letters*, 2: 215-219.

- [23] Cao, C., and Tsay, R. (1992). "Nonlinear Time-Series Analysis of Stock Volatilities". *Journal of Applied Econometrics*, 7: S165-S185.
- [24] Clements, M., and Krolzig, H. (1998). "A Comparison of the Forecast Performance of Markov-switching and Threshold Autoregressive Models of US GNP". *Econometrics Journal*, 1: C47-C75.
- [25] Clements, M., and Smith, J. (1997). "The Performance of Alternative Forecasting Methods for SETAR Models". *International Journal of Forecasting*, 13: 463-475.
- [26] Clements, M., and Smith, J. (1999). "A Monte Carlo Study of the Forecasting Performance of Empirical SETAR Models". *Journal of Applied Econometrics*, 14: 123-141.
- [27] Clements, M., and Smith, J. (2001). "Evaluating Forecasts from SETAR Models of Exchange Rates". *Journal of International Money and Finance*, 20: 133-148.
- [28] Crowder, M. (1976). "Maximum Likelihood Estimation for Dependent Observation". *Journal of the Royal Statistical Society, Series B*, 24: 406-424.
- [29] De Grauwe, P., Dewachter, H., and Embrechts, M. (1993). *Exchange Rates Theory: Chaotic Models of Foreign Exchange Markets*. Blackwell.
- [30] Dickey, D., and Fuller, W. (1979). "Distribution of the Estimates for Autoregressive Time Series with a Unit Root". *Journal of the American Statistical Association*, 74: 427-431.
- [31] Dickey, D., and Fuller, W. (1981). "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root". *Econometrica*, 49: 1057-1072.
- [32] Dickey, D., and Pantula S. (1987). "Determining the Order of Differencing in Autoregressive Processes". *Journal of Business and Economics Statistics*, 15: 455-461.
- [33] Diebold, F., and Mariano, R. (1995). "Comparing Predictive Accuracy". *Journal of Business and Economic Statistics*, 13: 253-263.
- [34] Diebold, F., and Nason, J. (1990). "Non-parametric Exchange Rate Prediction?". *Journal of International Economics*, 28: 315-322.
- [35] Diebold, F., and Pauly, P. (1986). "Evidence on the Small Sample Properties of Some Asymptoti-

cally Equivalent Tests for Autoregressive Conditional Heteroskedasticity". Paper presented to the Econometric Society European Meeting, Budapest.

- [36] Dumas, B. (1992). "Dynamic Equilibrium and the Real Exchange Rate in a Spatially Separated World". *Review of Financial Studies*, 5(2): 153-180.
- [37] Edison, H. (1987). "Purchasing Power Parity in the Long Run, A Test of the Dollar/Pound Exchange Rate (1890-1978)". *Journal of Money, Credit, and Banking*, 19: 376-387.
- [38] Efron, B. (1979). "Bootstrap Methods: Another Look at the Jackknife". *The Annals of Statistics*, 7: 1-26.
- [39] Efron, B., and Tibshirani, R. (1993). *An Introduction to the Bootstrap*. Chapman and Hall, New York.
- [40] Eitrheim, Ø., and Teräsvirta, T. (1996). "Testing the Adequacy of Smooth Transition Autoregressive Models". *Journal of Econometrics*, 74: 59-75.
- [41] Enders, W., and Granger, C. (1998). "Unit-root Tests and Asymmetric Adjustment with an Example Using the Term Structure of Interest Rates". *Journal of Business and Economic Statistics*, 16: 304-311.
- [42] Engel, C. (1994). "Can the Markov Switching Model Forecast Exchange Rates". *Journal of International Economics*, 36: 151-165.
- [43] Engel, C., and Hamilton, J. (1990). "Long Swings in the Dollar: Are They in the Data and Do Markets Know it?". *American Economic Review*, 80: 689-713.
- [44] Engle, R. (1982). "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation". *Econometrica*, 50: 987-1008.
- [45] Engle, R. (1995). *ARCH: Selected Readings*. Oxford University Press, Oxford, UK.
- [46] Engle, R., and Ng, V. (1993). "Measuring and Testing the Impact of News and Volatility". *Journal of Finance*, 5: 1749-1778.
- [47] Flood, R., Rose, A., and Mathieson, D. (1990). "Is the EMS the Perfect Fix? An Empirical

Exploration of Exchange Rate Target Zones". Board of Governors of the Federal Reserve System. International Finance Discussion Paper No. 388.

- [48] Frankel, J. (1986). "International Capital Mobility and Crowding-out in the US Economy: Imperfect Integration of Financial Markets or Goods Markets". *How Open is the US Economy?*, 33-67, Hafer, R. (Ed.). Lexington Books.
- [49] French, M., and Sichel, D. (1993). "Cyclical Patterns in the Variance of Economics Activity". *Journal of Business and Economic Statistics*, 11: 113-119.
- [50] Ghysels, E., Harvey, A., and Renault, E. (1996). "Stochastic Volatility". *Handbook of Statistics*, Vol. 14, 119-191, Maddala, G. (Ed.). North Holland, Amsterdam.
- [51] Granger, C., and Anderson, A. (1978). *Introduction to Bilinear Time Series Models*. Vandenhoeck and Ruprecht, Göttingen.
- [52] Granger, C., and Anderson, A. (1978). *Nonlinear Time Series Modelling in Applied Time Series Analysis*, Findley, F. (Ed). New York, Academic Press.
- [53] Granger, C., and Teräsvirta, T. (1993). *Modelling Nonlinear Economic Relationships*. Oxford, U.K.: Oxford University Press.
- [54] Guegan, D., and Pham, T. (1992). "Power of the Score Test against Bilinear Time Series Models". *Statistica Sinica*, 2: 157-169.
- [55] Guillaume, D., Dacorogna, M., Dave, R., Muller, U., Olsen, R., and Pictet, O. (1997). "From the Bird's Eye to the Microscope: A Survey of New Stylized Facts of the Intra-daily Foreign Exchange Markets". *Finance and Stochastics*, 1: 95-129.
- [56] Haggan, V., and Oyetunji, O. (1980). "On the Selection of Subset Autoregressive Time Series Models". UMIST Tech. Report 124, Department of Mathematics, UMIST.
- [57] Haggan, V., and Ozaki, T. (1981). "Modelling Nonlinear Random Vibrations Using an Amplitude-dependent Autoregressive Time-series Model". *Biometrika*, 68: 189-196.
- [58] Hamilton, J. (1989). "A New Approach to the Economic Analysis of Nonstationary time Series and the Business Cycle". *Econometrica*, 57: 357-384.

- [59] Hansen, B. (1996). "Inference when A Nuisance Parameter is not Identified under the Null Hypothesis". *Econometrica*, 64: 413-430.
- [60] Hansen, B. (1997). "Inference in TAR models". *Studies in Nonlinear Dynamics and Econometrics*, 1(2): 1-14.
- [61] Hansen, B. (1999). "Testing for Linearity". *Journal of Economic Surveys*, 13: 551-576.
- [62] Harvey, A. (1981). *The Econometric Analysis of Time Series*. Philip Allan, Deddington.
- [63] Hinich, M. (1982). "Testing for Gaussianity and Linearity of a Stationary Time Series". *Journal of Time Series Analysis*, 3: 169-176.
- [64] Hjorth, J. (1994). *Computer Intensive Statistical Methods Validation Model Selection and Bootstrap*. Chapman and Hall, London.
- [65] Hsieh, D. (1989). "Testing for Nonlinear Dependence in Daily Foreign Exchange Rates". *Journal of Business*, 62: 339-368.
- [66] Hsieh, D. (1992). "A Nonlinear Stochastic Rational Expectations Model of Exchange Rates". *Journal of International Money and Finance*, 11: 235-250.
- [67] Kanzler, L. (1999). "Very Fast and Correctly Sized Estimation of the BDS Statistic". *Working paper*, Christ Church, Oxford University, England.
- [68] Keenan, D. (1985). "A Turkey Non-additivity-type Test for Time Series Nonlinearity". *Biometrika*, 72: 39-44.
- [69] Keynes, J. (1936). *The General Theory of Employment, Interest and Money*. Macmillan, London.
- [70] Kilian, L., and Taylor, M. (2001). "Why is it so Difficult to Beat the Random Walk Forecast of Exchange Rates?". Revised version of Department of Economics Working Paper No. 01-01. University of Michigan.
- [71] Kräger, H., and Kugler, P. (1993). "Non-linearities in Foreign Exchange Markets: A Different Perspective". *Journal of International Money and Finance*, 12: 195-208.

- [72] Lawrance, A., and Lewis, P. (1977). "An Exponential Moving Average Sequence and Point Process, EMA(1)". *Journal of Applied Probability*, 14: 98-113.
- [73] Lawrance, A., and Lewis, P. (1980). "The Exponential Autoregressive-Moving Average EARMA(p, q) Process". *Journal of the Royal Statistical Society*, B42: 150-161.
- [74] LeBaron, B. (1997). "A Fast Algorithm for the BDS Statistic". *Studies in Nonlinear Dynamics and Econometrics*, vol. 2, no. 2, 53-59.
- [75] LePage, U., and Billard, L. (1992). *Exploring the Limits of Bootstrap*. John Wiley, New York.
- [76] Leybourne, S., and Mizen, P. (1999). "Understanding the Disinflation in Australia, Canada and New Zealand using Evidence from Smooth Transition Analysis". *Journal of International Money and Finance*, 18: 799-816.
- [77] Luukkonen, R., Saikkonen, P., and Teräsvirta, T. (1988). "Testing Linearity in Univariate Time Series Models". *Scandinavian Journal of Statistics*, 15: 161-175.
- [78] Luukkonen, R., Saikkonen, P., and Teräsvirta, T. (1988). "Testing Linearity against Smooth Transition Autoregressive Models". *Biometrika*, 75: 491-499.
- [79] Maravall, A. (1983). "An Application of Nonlinear Time Series Forecasting". *Journal of Business and Economic Statistics*, 1: 66-74.
- [80] McLeod, A., and Li, W. (1983). "Diagnostic Checking ARMA Time Series Models Using Squared-residual Autocorrelations". *Journal of Time Series Analysis*, 4: 269-273.
- [81] Meese, R., and Rogoff, K. (1983). "Empirical Exchange Rate Models of the Seventies: Do They Fit Out-of-sample?". *Journal of International Economics*, 14: 3-24.
- [82] Meese, R., and Rose, A. (1991). "An Empirical Assessment of Non-Linearities in Models of Exchange Rate Determination". *Review of Economic Studies*, 58: 603-619.
- [83] Michael, P., Nobay, A., and Peel, D. (1997). "Transactions Costs and Nonlinear Adjustment in Real Exchange Rates: an Empirical Investigation". *Journal of Political Economics*, 105: 862-879.

- [84] Mitchell, W. (1927). *Business Cycles. The Problem and its Setting*. National Bureau of Economic Research, New York.
- [85] Mohler, R. (1973). *Bilinear Control Processes: with Applications to Engineering, Ecology, and Medicine*. New York, Academic Press.
- [86] Neftci, S. (1984). "Are Economic Time Series Asymmetric over the Business Cycle?". *Journal of Political Economy*, 92: 11-48.
- [87] Nelson, D. (1991). "Conditional Heteroskedasticity in Asset Returns: A New Approach". *Econometrica*, 59: 347-370.
- [88] Neter, J., Kutner, M., Nachtsheim, C., and Wasserman, W. (1996). *Applied Linear Statistical Models*. Irwin.
- [89] Nychka, D., Ellner, S., Gallant, R., and McCaffrey, D. (1992). "Finding Chaos in Noisy Systems". *Journal of the Royal Statistical Society*, B54(2): 399-426.
- [90] Öcal, N., and Osborn, D. (2000). "Business Cycle Non-linearities in UK Consumption and Production". *Journal of Applied Econometrics*, 15: 27-43.
- [91] Ozaki, T. (1980). "Non-linear Time-series Models for Non-linear Random Vibrations". *Journal of Applied Probability*, 17: 84-93.
- [92] Ozaki, T., and Oda, H. (1978). "Non-linear Time Series Model Identification by Akaike's Information Criterion". Proc. IFAC Workshop on Information and Systems, Compeign, France. October, 1977.
- [93] Pagan, A. (1978). "Some Simple Tests for Nonlinear Time Series Models". CORE Discussion Paper no. 7812.
- [94] Peel, D., and Speight, A. (1998). "The Nonlinear Time Series Properties of Unemployment Rates: Some Further Evidence". *Applied Economics*, 30: 287-294.
- [95] Peel, D., and Speight, A. (1998). "Modelling Business Cycle Nonlinearity in Conditional Mean and Conditional Variance: Some International and Sectoral Evidence". *Economica*, 65: 211-229.

- [96] Pesaran, M., and Potter, S. (1997). "A Floor and Ceiling Model of US Output". *Journal of Economic Dynamics and Control*, 21: 661-695.
- [97] Peters, E. (1994). *Fractal Market Analysis: Applying Chaos Theory to Investment and Economics*. John Wiley and Sons.
- [98] Petrucci, J., and Davies, N. (1986). "A Portmanteau Test for Self-Exciting Threshold Autoregressive-Type Nonlinearity in Time Series". *Biometrika*, 73: 687-694.
- [99] Phillips, P. (1987). "Time Series Regression with Unit Roots". *Econometrica*, 55: 277-302.
- [100] Phillips, P., and Perron, P. (1988). "Testing for a Unit Roots in Time series Regression". *Biometrika*, 75: 335-346.
- [101] Pippenger, M., and Goering, G. (1998). "Exchange Rate Forecasting: Results from a Threshold Autoregressive Model". *Open Economies Review*, 9: 157-170.
- [102] Potter, S. (1995). "A Nonlinear Approach to US GNP". *Journal of Applied Econometrics*, 10: 109-125.
- [103] Rose, A., and Svensson, L. (1991). "Expected and Predicted Realignments: the FF/DM Exchange Rate During the EMS". Board of Governors of the Federal Reserve System. International Finance Discussion Paper No. 395.
- [104] Rothman, P. (1991). "Further Evidence on the Asymmetric Behavior of Unemployment Rates over the Business Cycle". *Journal of Macroeconomics*, spring: 291-298.
- [105] Rothman, P. (1998). "Forecasting Asymmetric Unemployment Rates". *The Review of Economics and Statistics*, 80(1): 164-168.
- [106] Saikkonen, P., and Luukkonen, R. (1988). "Lagrange Multiplier Test for Testing Non-linearities in Time Series Models". *Scandinavian Journal of Statistics*, 15: 55-68.
- [107] Sarantis, N. (1999). "Modeling Non-linearities in Real Effective Exchange Rates", *Journal of International Money and Finance*, 18: 27-45.

- [108] Sercu, P., Uppal, R., and Van Huue, C. (1995). "The Exchange Rate in the Presence of Transactions Costs: Implications for Tests of Purchasing Power Parity". *Journal of Finance*, 50(4): 1309-1319.
- [109] Shephard, N. (1996). "Statistical Aspects of ARCH and Stochastic Volatility Models". *Time Series Models in Econometrics, Finance, and Other Fields*, 1-67, Cox, D., Hinkley, D., and Barndorff-Nielsen, O. (Eds.). Chapman and Hall, London, UK.
- [110] Sichel, D. (1993). "Business Cycle Asymmetry: A Deeper Look". *Economic Inquiry*, 31: 224-236.
- [111] Subba Rao, T., and Gabr, M. (1980). "A Test for Linearity of Stationary Time Series". *Journal of Time Series Analysis*, 1: 145-152.
- [112] Subba Rao, T., and Gabr, M. (1984). *An Introduction to Bispectral Analysis and Bilinear Time Series Models*, Springer-Verlag.
- [113] Taylor, M., and Peel, D. (2000). "Nonlinear Adjustment, Long-run Equilibrium and Exchange Rate Fundamentals". *Journal of International Money and Finance*, 19: 33-53.
- [114] Taylor, M., Peel, D., and Sarno, L. (2001). "Nonlinear Adjustment in Real Exchange Rates: Toward a Solution to the Purchasing Power Parity Puzzles". *International Economic Review*, forthcoming.
- [115] Teräsvirta, T. (1994). "Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models". *Journal of the American Statistical Association*, 89: 208-218.
- [116] Teräsvirta, T., and Anderson, H. (1992). "Characterizing Nonlinearities in Business Cycles Using Smooth Transition Autoregressive Models". *Journal of Applied Econometrics*, 7: S119-S136.
- [117] Terdik, G. (1999). *Bilinear Stochastic Models and Related Problems of Nonlinear Time Series Analysis*. Springer.
- [118] Tiao, G., and Tsay, R. (1994). "Some Advances in Nonlinear and Adaptive Modeling in Time Series Analysis". *Journal of Forecasting*, 13: 109-131.
- [119] Tong, H. (1978). *On a Threshold Model in Pattern Recognition and Signal Processing*, Chen, C. (Ed.). Amsterdam: Sijhoff & Noordhoff.

- [120] Tong, H. (1983). *Threshold Models in Nonlinear Time Series Analysis (Lecture Notes in Statistics No. 21)*. New York: Springer-Verlag.
- [121] Tong, H. (1990). *Nonlinear Time-series: A Dynamical Systems Approach*. New York: Oxford University Press.
- [122] Tong, H., and Lim, K. (1980). "Threshold Autoregression, Limit Cycles and Cyclical Data" (with discussion). *Journal of the Royal Statistical Society*, B42: 245-292.
- [123] Tsay, R. (1986). "Non-linearity Tests for Time Series". *Biometrika*, 73: 461-466.
- [124] Tsay, R. (1989). "Testing and Modeling Threshold Autoregressive Processes". *Journal of the American Statistical Association*, 84: 231-240.
- [125] Tukey, J. (1958). "Bias and Confidence in Not-Quite Large Samples" (abstract). *The Annals of Mathematical Statistics*, 29: 614.
- [126] Volterra, V. (1930). *Theory of Functionals and of Integral and Integro-differential Equations*. London, Glasgow, Blackie & Son Limited.
- [127] Weiss, A. (1984). "ARMA Models with ARCH Errors". *Journal of Time Series Analysis*, 5: 129-143.
- [128] Weiss, A. (1986). "ARCH and Bilinear Time Series Models: Comparison and Combination". *Journal of Business and Economic Statistics*, 4: 59-70.
- [129] White, H. (1989). "Some Asymptotic Results for Learning in Single Hidden-layer Feedforward Network Models". *Journal of the American Statistical Association*, 84, no. 408, 1003-1013.
- [130] White, H. (1989). "An Additional Hidden Unit Test for Neglected Nonlinearity in Multilayer Feedforward Networks". *Proceedings of the International Joint Conference On Neural Networks* (IEEE Press, New York) Vol II: 451-455.
- [131] Wold, H. (1938). *A Study in the Analysis of Stationary Time Series*. Almqvist & Wiskell, Uppsala.

ACKNOWLEDGMENTS

Along the journey of writing this dissertation, I am very luck to get guidance, help and love from many people. Without them, I could not be able to finish the dissertation.

First, I would like to thank my father. He suggested me to choose economics as my major when I graduated from high school. This is the first stepping stone for the dissertation.

Second, I would like to thank my major professor, Dr. Barry Falk. He has a very busy schedule of teaching, research and advising. Yet, as just one of his many advisees, I have never been feeling neglected for one moment. From the topic to the wording, from the theory to the programming, he is always there to help me. His academic guidance and working attitude will affect me far beyond this dissertation.

Third, I would like to give many thanks to people who helped me during my difficult time. Dr. Enders, Dr. Carriquiry, Dr. Bunzel, Dr. Breidt, Dr. Herriges, Dr. Mattila, Dr. Yang and Julie, Danette, Diana, my classmates and friends, sorry I can not complete all of them here. These thanks should be sent out long time ago, but the appreciations are so enormous that I did not know how to express them appropriately then. I am blessed that my individual trajectory crosses with theirs. Without their love and encouragement, I could not come so far.

Last but not the least, I would like to thank my husband, Jiwu Sun. Five years ago, we got married and came to US to study. Since then, no matter what comes out of the life, happiness or joy, sadness or sorrow, I know I have a person I can always rely on. Without his support, I can hardly imagine I could finish the dissertation.