DOKUZ EYLÜL UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

NEW BOOTSTRAP METHODS FOR EXCHANGE RATE PREDICTION UNDER GARCH(P,Q) PROCESS

by
Beste Hamiye BEYAZTAŞ

February, 2018
İZMİR

NEW BOOTSTRAP METHODS FOR EXCHANGE RATE PREDICTION UNDER GARCH(P,Q) PROCESS

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

by

Beste Hamiye BEYAZTAŞ

February, 2018
İZMİR

Ph.D. THESIS EXAMINATION RESULT FORM

We have read the thesis entitled **"NEW BOOTSTRAP METHODS FOR EXCHANGE RATE PREDICTION UNDER GARCH(P,Q) PROCESS"** completed by **BESTE HAMİYE BEYAZTAŞ** under supervision of **PROF. DR. ESİN FİRUZAN** and we certify that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy.

.....

Prof. Dr. Esin FİRUZAN

Supervisor

.....

Assoc. Prof. Dr. Burcu ÜÇER

Thesis Committee Member

.....

Prof. Dr. Hakan KAHYAOĞLU

Thesis Committee Member

Prof. Dr. Hüseyin TATLIDİL .....

Examining Committee Member

Prof. Dr. Erol EĞRİOĞLU .....

Examining Committee Member

Prof. Dr. Kadriye ERTEKİN

Director

Graduate School of Natural and Applied Sciences

ACKNOWLEDGMENT

I would like to express my sincere gratitude to my supervisor Prof. Dr. Esin FİRUZAN for her guidance, invaluable assistance, encouraging my research and generous support of all kinds during whole doctoral studies. This dissertation would not have completed without her professional experience, knowledge and help.

Besides my supervisor, I would like to thank Assoc. Prof. Dr. Soutir BANDYOPADHYAY and Prof. Dr. Wei-Min HUANG for their precious support, encouragement, constructive suggestions and contributions on my PhD research studies during my visit to Lehigh University in USA. I also would like to thank my committee members, Assoc. Prof. Dr. Burcu ÜÇER and Prof. Dr. Hakan KAHYAOĞLU for serving as my committee members with their precious time, guidance and brilliant suggestions.

Special thanks go to my beloved husband Ufuk BEYAZTAŞ, who spent sleepless nights working together, for his love and being truly supportive. Without his valuable help on R coding in simulation studies, this journey would have been more difficult.

At the end I wish to express my deepest thanks to my family. Words can not express how grateful I am to my mother, my father, and my sister for their endless love and all of the sacrifices.

Beste Hamiye BEYAZTAŞ

NEW BOOTSTRAP METHODS FOR EXCHANGE PREDICTION UNDER GARCH(P,Q) PROCESS

ABSTRACT

The price changes forms an important part of the financial time series analysis. Modeling of financial time series such as exchange rates, market indices, interest rates and option prices, succeeding accurate parameter estimations of the model and prediction of future values play a critical role in assessing risk and uncertainty. In foreign exchange market, the generalized autoregressive conditional heteroscedasticity model is one of the most commonly used techniques for modeling volatility and obtaining dynamic prediction intervals for returns as well as volatilities. Technically, construction of such prediction intervals requires some distributional assumptions which are generally unknown in practice. Moreover, the constructed prediction intervals along with the estimated parameter values can be affected due to any departure from the assumptions and may lead us to unreliable results. One of the remedy to construct prediction intervals without considering distributional assumptions is to use the well known resampling methods, e.g., the bootstrap.

In this dissertation, two new block bootstrap based methods are proposed to obtain valid prediction intervals in conditionally heteroscedastic models. Firstly, the proposed block bootstrap methods are discussed in the context of linear time series models. Secondly, these methods are applied to heteroscedastic models to obtain prediction intervals. Performances of the proposed methods have been compared with the existing methods via both real-world examples and simulation studies. The results are reveal that our proposed methods outperform other existing methods for parameter estimation. Also, they produce better prediction intervals compared to existing ones.

Keywords: Block bootstrap, financial time series, generalized autoregressive conditional heteroscedastic models, prediction, sufficient bootstrap

GARCH(P,Q) MODELLERİ ALTINDA DÖVİZ KURU ÖNGÖRÜSÜ İÇİN YENİ BOOTSTRAP YÖNTEMLERİ

ÖZ

Fiyat değişimleri finansal zaman serileri analizinin önemli bir parçasını oluşturur. Döviz kurları, piyasa endeksleri, faiz oranları ve opsiyon fiyatları gibi finansal zaman serilerinin modellenmesi, modelin doğru parametre tahminlerinin elde edilmesi ve gelecek değerlerin öngörülmesi, risk ve belirsizliğin değerlendirilmesinde kritik bir rol oynar. Uluslararası döviz piyasasında, genelleştirilmiş otoregresif koşullu değişen varyans modeli, oynaklığın modellenmesi ve getirilerin yanında oynaklıkların dinamik öngörü aralıklarının elde edilmesi için en yaygın olarak kullanılan yöntemlerden biridir. Teknik olarak, bu öngörü aralıklarının oluşturulması, uygulamada genellikle bilinmeyen bazı dağılımsal varsayımları gerektirir. Dahası, tahminlenen parametre değerleriyle oluşturulan öngörü aralıkları, varsayımlardan herhangi bir sapma nedeniyle etkilenebilir ve güvenilmez sonuçlara yol açabilir. Dağılımsal varsayımları dikkate almadan öngörü aralıkları oluşturmanın bir yolu, bootstrap gibi yeniden örnekleme yöntemlerinin kullanılmasıdır.

Bu tezde, koşullu değişen varyans modellerinde geçerli öngörü aralıklarının elde edilmesi için iki yeni blok bootstrap tabanlı yöntem önerilmiştir. İlk olarak, önerilen blok bootstrap yöntemleri doğrusal zaman serisi modelleri kapsamında tartışılmıştır. İkinci olarak, bu yöntemler öngörü aralıkları elde etmek için heteroskedastik modellere uygulanmıştır. Önerilen yöntemlerin performansları hem gerçek dünya örnekleri hem de simülasyon çalışmaları aracılığıyla varolan yöntemlerle karşılaştırılmıştır. Parametre tahmini için elde ettiğimiz sonuçlar, önerdiğimiz yöntemlerin varolan yöntemlerden daha iyi performansa sahip olduğunu göstermektedir. Ayrıca, önerdiğimiz yöntemler varolan yöntemlere göre daha iyi öngörü aralıkları üretmektedir.

Anahtar kelimeler: Blok bootstrap, finansal zaman serileri, genelleştirilmiş otoregresif koşullu değişen varyans modelleri, öngörü, yeterli bootstrap

CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER ONE
INTRODUCTION

1.1 Introduction

The fundamental concept in financial markets is the price such as bond price, price of stock, currency price etc. In recent years, the intensive research on financial studies has focused on price changes, instead of prices of assets. As mentioned by Campbell, Lo, & MacKinlay (1997), using price changes or returns have gained importance in financial time series analysis since asset return is a scale-free summary of investment performance and it has more attractive statistical properties than price such as stationarity. Most of the financial time series have some typical features. These characteristics are defined as follows:

- Price series of financial assets generally creates non-stationary process while return series related to asset prices are usually stationary.
- The asset returns have uncorrelated structure or weak correlation while squared or absolute returns of assets show stronger autocorrelation than the original one.
- Excess volatility caused by nonsystematic factors can appear in the most of asset return series. Furthermore, observed level of volatilities are locally clustered in groups of low or high variabilities.
- The asset return series is unconditionally heavy-tailed distributed with positive excess kurtosis, generally.
- Leverage effect that means the negative correlation between past returns and future volatilities is observable from some of the financial time series.

One of the major current areas in financial time series analysis and risk management is accurately modelling and predicting exchange rate volatility. The Chicago Board Options Exchange was established for the exchange-traded stock options as first in 1973, and well-known Black-Scholes option pricing model was introduced in the same year. Since then, for many countries, the global perspective of

"fixed exchange rate regime" has changed to the "fluctuating exchange rate". Volatility of the underlying currency pair, a measure of the variability in an underlying exchange rate, has been an essential concept that needs to be emphasized in evaluation of exchange rate risk because of oscillations in exchange rate of currencies. Subsection 1.2 provides some important concepts regarding the volatility of exchange rates and a detailed information about exchange rate modelling.

1.2 Exchange Rate Time Series Modelling

Classical econometric models used in exchange rate analysis make certain assumptions on the errors such as independence, uncorrelated structure and constant variability of error terms. However, many macroeconomic and financial time series vary over wide range around mean, and very large or small prediction errors may occur in practice. Since financial markets are sensitive to political events, speculations, changes in monetary policy etc., this variability in the error terms may occur. As notted by Üstünel (1999) the over-falls and over-rises in the returns of shares due to speculative rate moves force the Normal Distribution assumption. Also, this implies that the variance of the errors may not be constant and it changes over time so the errors can be serially correlated in financial data. Additionally, one of the uncertain and decisive factors in pricing of exchange rate risk is volatility as a measure of dispersion and an indicator of magnitude of fluctuations of the asset price series since volatility is not directly observable. As the volatility of the exchange rate increases, exchange rate risk may climbs for international traders and investors because high volatility implies the high level of uncertainty of prices and high risk of loss or gain. The other determinants of exchange rates are the economic indicators depend on the international trade relations between two countries such as differentials in inflation, interest rate differentials, public debt, political stability, commercial balance etc.. On the other hand, volatility shapes risk perception of foreign exchange markets related to underlying currency pair and indicates the amount of uncertainty, so volatility modelling has a great importance in exchange rate valuation, option pricing and other financial applications. Volatility is expressed by the conditional standard deviation or conditional variance of the financial return

series as a statistical measure and there are two main types of volatility referred to as historical (statistical) and implied volatility. Historical volatility indicates the recent price movements of the financial instrument for a certain period of time while implied volatility is obtained from the exchange rate based on market expectations and uses Black-Scholes model for European options. To derive implied volatility, Black-Scholes model assumes that the underlying asset price series follows geometric Brownian motion. The Black-Scholes formula is expressed as a partial differential equation by using the relations between elements of option price. Chen (2015) has explained the Black-Scholes equation as a deterministic representation of lognormal processes. Also, this model makes constant volatility assumption during the life of the option. Tsay (2005) has stated that the Black-Scholes approach is frequently criticized since some assumptions of the model may not hold in practice and implied volatility obtained from this model can be different from the real volatility. Also, Gereben (2005) has pointed out that implied volatility leads to obtaining a biased estimate, and does not include the information provided by generalized autoregressive conditional heteroscedastic and autoregressive-moving average predictors for volatility computed from historical exchange rate series. On the other hand, in empirical finance studies, volatility has certain features which are seen especially in financial return series. One of these empirical findings in financial asset returns is the clustered volatility which can be seen in groups of low/high values of volatilities for certain periods of time. As has been pointed out by Mandelbrot (1963), large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes. Also, Bollerslev (1990), Brooks & Burke (1998), and Fiser & Horvath (2010) have examined the characteristics of exchange rate volatility and they emphasized the presence of volatility clustering in the exchange rate return series. For the second feature of volatility, as it has been stated in Tsay (2005), volatility evolves over time in a continuous manner, that is, jumps in volatility are not common. The other characteristic of the volatility is that the variability of the volatility is restricted with some fixed range. Statistically, this characteristic often implies stationarity of the volatility. Also, in case of large price changes, volatility and asset returns act in the opposite directions, called to as leverage effect.

In summary, since volatility is the unobservable component of exchange rates, it should be modeled correctly to obtain efficient parameter estimation and improve the accuracy of prediction intervals for assessing uncertainty in risk management. The assumptions of classical econometric models generally can not be achieved for the financial time series in practice and this situation has led to widespread use of the conditionally heteroscedastic models which allow modelling non-constant variance. Besides, all of the aforementioned characteristics of return series have yielded development and diversity of the conditional heteroscedastic time series models for the volatility modelling. Next subsections present the detailed information about linear and conditionally heteroscedastic time series models.

1.2.1 Linear Time Series Models

In the quantitative finance literature, linear and nonlinear time series models depending on the frequency of the data are commonly used for modelling and forecasting of exchange rate returns and volatilities. For example, the autoregressive and autoregressive moving average models have been recently used for the purpose of model estimation and forecasting of BDT/USD exchange rate in the study of Alam (2012). Also, Ghalayini (2013) build an autoregressive moving average model for US Dollar/Euro exchange rate. The most commonly used linear stationary time series models for the analysis of return series of various financial assets are autoregressive (AR), moving average (MA) and mixed autoregressive moving average models (ARMA).

1.2.1.1 Autoregressive Models

A realization at a certain time of a *p*th order autoregressive time series (AR(*p*)) is explained by a weighted linear sum of *p* previous observations of the series and a noise term. Mathematically, an autoregressive model of order *p* is defined as

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + Z_t, \qquad t = 1, \cdots, T \tag{1.1}$$

where $\phi_0, \phi_1, \cdots, \phi_p$ are unknown parameters and $\{Z_t\}$ is independent and identically distributed (i.i.d.) white noise sequence, that is, $E(Z_t)=0$, $Var(Z_t)=\sigma^2$ and $E(Z_t Z_s)=0$ for all $t \neq s$ $(\{Z_t\} \sim WN(0,\sigma^2))$. An AR($p$) model has the same form as a multiple linear regression model in which Y_t is the dependent variable and the lagged values of dependent variable, $\{Y_{t-1}, \cdots, Y_{t-p}\}$, are the explanatory variables. The AR(p) model can also be expressed using backshift operator B as

$$\begin{aligned} Y_t &= \phi_0 + \sum_{i=1}^{p} \phi_i Y_{t-i} + Z_t \\ &= \phi_0 + \sum_{i=1}^{p} \phi_i B^i Y_t + Z_t \end{aligned} \tag{1.2}$$

so that $\Phi(B)Y_t = \phi_0 + Z_t$ where $BY_t = Y_{t-1}$, $\Phi(B) = (1 - \phi_1 B - \cdots - \phi_p B^p)$ and $\phi_0 = E(Y_t)\Phi(B) = \mu\Phi(B)$. The process Y_t is said to be stationary which means some statistical properties of the process such as mean, variance etc. do not change over time if and only if all roots of the characteristic polynomial, $\Phi(B) = (1 - \phi_1 B - \cdots - \phi_p B^p)$, lie outside of the unit circle, that is, $\phi(B) \neq 0$ for all $|B| \leq 1$. Also, the AR(p) process is always invertible, since $\sum_{i=1}^{p} |\phi_i| < \infty$.

The mean of a stationary AR(p) process is given as follows

$$E(Y_t) = \mu = \frac{\phi_0}{1 - \phi_1 - \cdots - \phi_p} \tag{1.3}$$

The pth order autoregression can also be written as in mean adjusted form by using Equation (1.3) as follows

$$\begin{aligned} &(Y_t - \mu) = \phi_1 (Y_{t-1} - \mu) + \cdots + \phi_p (Y_{t-p} - \mu) + Z_t, \qquad t = 1, \cdots, T \\ &\dot{Y}_t = \phi_1 \dot{Y}_{t-1} + \cdots + \phi_p \dot{Y}_{t-p} + Z_t \end{aligned} \tag{1.4}$$

By using the mean adjusted form of the process, autocovariance function (ACVF) at lag k and variance of the process are calculated as in Equations (1.5) and (1.6) given below

$$\gamma_Y(k) = Cov(Y_t, Y_{t-k}) = \sum_{i=1}^{p} \phi_i \gamma_Y(k-i) + \begin{cases} 0 & \text{for } k > 0 \\ \sigma^2 & \text{for } k = 0 \end{cases} \tag{1.5}$$

$$Var(Y_t) = \gamma_Y(0) = \sum_{i=1}^{p} \phi_i \gamma(i) + \sigma^2 \tag{1.6}$$

Also, by dividing $\gamma_Y(k)$ to $\gamma_Y(0)$, the autocorrelation function (ACF) of the process which produces the Yule-Walker equations is obtained as $\rho_Y(k) = \sum_{i=1}^{p} \phi_i \rho_Y(k-i)$ for $k>0$.

Several methods such as method of moments (MM), conditional and unconditional maximum likelihood (CML-ML), conditional and unconditional least-square methods (CLS-LS) are available to estimate the parameters of autoregressive models. We restrict our attention in this dissertation on ML methods for the estimation of linear time series models with Gaussian (Normal) white noise sequence, $\{Z_t\} \sim$ i.i.d. $N(0, \sigma^2)$.

For the stationary AR(p) model, the joint probability density of $\mathbf{Z} = (z_1, \cdots, z_T)'$ is written as

$$P(\mathbf{Z} | \boldsymbol{\Phi}, \sigma^2) = L_*(\boldsymbol{\Phi}, \sigma^2) = (2\pi\sigma^2)^{-T/2} \exp\left[-\frac{1}{2\sigma^2} \sum_{t=1}^{T} Z_t^2 \right] \tag{1.7}$$

where $\boldsymbol{\Phi} \equiv (\phi_0, \phi_1, \cdots, \phi_p)$ and $L_*(\boldsymbol{\Phi}, \sigma^2)$ represent the parameter vector and the conditional likelihood function, respectively. Let $\mathbf{Y} = (y_1, \cdots, y_T)'$ be the observed sequence of time series. Under the normality assumption, the conditional log-likelihood function is given as follows

$$\ell_*(\boldsymbol{\Phi}, \sigma^2) = \log L_*(\boldsymbol{\Phi}, \sigma^2) = -\frac{T}{2} \log(2\pi\sigma^2) - \frac{S_*(\boldsymbol{\Phi})}{2\sigma^2} \tag{1.8}$$

where $S_*(\boldsymbol{\Phi}) = \sum_{t=1}^{T} Z_t^2(\boldsymbol{\Phi} | \mathbf{Y}_*, \mathbf{Y})$ and $Z_t(\boldsymbol{\Phi} | \mathbf{Y}_*, \mathbf{Y})$ respectively, denote the conditional sum of squares (CSS) function and the error sequence, where $\mathbf{Y}_* = (y_{1-p}, \cdots, y_0)'$ is the vector of initial values. Since the initial conditions are unobservable in practice, by taking $\mathbf{Y}_* = (y_1, \cdots, y_p)'$ as initial values, the log-likelihood and sum of squares

functions conditional on the first p observations, respectively, can be rewritten as follows

$$\ell_*\left(\mathbf{\Phi},\sigma^2\right)=\log f_{Y_{P+1},\cdots,Y_T|Y_1,\cdots,Y_P}\left(\mathbf{Y}|\mathbf{Y}_*;\mathbf{\Phi}\right)=-\frac{(T-p)}{2}\log\left(2\pi\sigma^2\right)-\frac{S_*\left(\mathbf{\Phi}\right)}{2\sigma^2} \tag{1.9}$$

$$S_*\left(\mathbf{\Phi}\right)=\sum_{\mathrm{t}=p+1}^{\mathrm{T}}Z_t^2\left(\mathbf{\Phi}|\mathbf{Y}_*,\mathbf{Y}\right) \tag{1.10}$$

where $\mathbf{Y}=\left(y_{p+1},\cdots,y_T\right)'$ is the observed values of the process and $\log f_{Y_{P+1},\cdots,Y_T|Y_1,\cdots,Y_P}\left(\mathbf{Y}|\mathbf{Y}_*;\mathbf{\Phi}\right)$ shows the joint probability density of $\mathbf{Y}$ conditional on $\mathbf{Y}_*$. Then, the conditional ML estimates, $\hat{\mathbf{\Phi}}$ and $\hat{\sigma}^2$, of the unknown parameter vector and σ^2 are obtained by maximizing the conditional log-likelihood function or equivalently minimizing the CSS given as in Equations(1.11) and (1.12)

$$\hat{\mathbf{\Phi}}=\arg\max_{\Phi}\ell_*\left(\mathbf{\Phi},\sigma^2\right)=\arg\min_{\Phi}S_*\left(\mathbf{\Phi}\right) \tag{1.11}$$

$$\hat{\sigma}^2=\frac{S_*\left(\mathbf{\Phi}\right)}{T-p} \tag{1.12}$$

1.2.1.2 Moving Average Models

One of the simple and useful approaches for modeling return series in finance is the moving-average model. A moving average process of order q abbreviated as MA(q) is expressed as a linear combination of q past periods and present value of the innovations and it is given as

$$Y_t=\phi_0+Z_t-\theta_1Z_{t-1}+\cdots-\theta_qZ_{t-q},\qquad t=1,\cdots,T \tag{1.13}$$

where $\phi_0,\theta_1,\cdots,\theta_q$ are the coefficients of the model and $\{Z_t\}$ is a white noise process with zero mean and a constant variance σ^2. Also, qth order moving average model in terms of backshift operator can be equivalently written in following form

$$\begin{aligned}Y_t&=\phi_0+Z_t\left(1-\theta_1B-\cdots-\theta_qB^q\right)\\&=\phi_0+Z_t\left(1-\sum_{j=1}^{q}\theta_jB^j\right)\\&=\phi_0+\Theta(B)Z_t\end{aligned} \tag{1.14}$$

where $\Theta(B)=\left(1-\sum_{i=1}^{q}\theta_j B^j\right)$. An important property of MA(q) model with finite weights $\left(1+\theta_1^2+\cdots+\theta_q^2<\infty\right)$ is always weakly stationary, implying that the first two moments of the process are independent of time. Furthermore, the MA process is said to be invertible, that is a finite order moving average process can be written in the form of infinite order autoregression (AR(∞)), if and only if all roots of the equation $\Theta(B)$ lie outside of the unit circle provided that $\Theta(B)\neq 0$ for all $|B|\leq 1$.

By taking expectation and variance on both sides of the Equation (1.13), the mean and variance of MA(q) process is obtained as in Equations (1.15) and (1.16)

$$E(Y_t)=\mu=\phi_0 \tag{1.15}$$

$$\mathrm{Var}(Y_t)=\gamma_Y(0)=\sigma^2\left(1+\theta_1^2+\cdots+\theta_q^2\right) \tag{1.16}$$

Moreover, by multiplying Equation (1.14) by Y_{t+k} for $k\geq 0$ and taking expectations, the ACVF at lag k of the MA(q) model is found as

$$\begin{aligned}\gamma_Y(k)&=Cov(Y_t,Y_{t+k})\\&=\begin{cases}\sigma^2\left(-\theta_k+\theta_1\theta_{k+1}+\cdots+\theta_{q-k}\theta_q\right), & k=1,2,\cdots,q\\0, & k>q\end{cases}\end{aligned} \tag{1.17}$$

The ACF of this process by the calculation of $\frac{\gamma_Y(k)}{\gamma_Y(0)}$ is expressed as

$$\rho_Y(k)=\begin{cases}\dfrac{-\theta_k+\theta_1\theta_{k+1}+\cdots+\theta_{q-k}\theta_q}{1+\theta_1^2+\cdots+\theta_q^2}, & k=1,2,\cdots,q\\0, & k>q\end{cases} \tag{1.18}$$

For an invertible MA(q) model, the joint probability density, $f_{Y_1,\cdots,Y_T|\mathbf{Z}_*=\mathbf{0}}(\mathbf{Y}|\mathbf{Z}_*=\mathbf{0};\Theta)$, of the sample, $\mathbf{Y}=(y_1,\cdots,y_T)'$, conditioning on $\mathbf{Z}_*=(z_0,\cdots,z_{1-q})'=\mathbf{0}$ is defined as

$$f_{Y_1,\cdots,Y_T|\mathbf{Z}_*=\mathbf{0}}(\mathbf{Y}|\mathbf{Z}_*=\mathbf{0};\Theta)=L_*(\boldsymbol{\Theta},\sigma^2)=(2\pi\sigma^2)^{-T/2}\exp\left[-\frac{1}{2\sigma^2}\sum_{t=1}^{T}Z_t^2\right] \tag{1.19}$$

where $\boldsymbol{\Theta} \equiv (\phi_0, \theta_1, \cdots, \theta_q)$ and $L_*(\boldsymbol{\Theta}, \sigma^2)$, represent the parameter vector and conditional likelihood function, respectively.

Let $S_*(\boldsymbol{\Theta}) = \sum_{t=1}^{T} Z_t^2(\boldsymbol{\Theta}|\mathbf{Z}_*, \mathbf{Y})$ denotes the CSS. Then, the conditional log-likelihood function and the conditional ML estimates, $\hat{\boldsymbol{\Theta}}$ and $\hat{\sigma}^2$, are obtained as in Equations (1.20) - (1.22)

$$\begin{aligned} \ell_*(\boldsymbol{\Theta}, \sigma^2) &= \log L_*(\boldsymbol{\Theta}, \sigma^2) = \log f_{Y_1,\cdots,Y_T|\mathbf{Z}_*=\mathbf{0}}(\mathbf{Y}|\mathbf{Z}_* = \mathbf{0}; \boldsymbol{\Theta}) \\ &= -\frac{T}{2}\log(2\pi\sigma^2) - \frac{S_*(\boldsymbol{\Theta})}{2\sigma^2} \end{aligned} \tag{1.20}$$

$$\hat{\boldsymbol{\Theta}} = \arg\max_{\boldsymbol{\Theta}} \ell_*(\boldsymbol{\Theta}, \sigma^2) = \arg\min_{\boldsymbol{\Theta}} S_*(\boldsymbol{\Theta}) \tag{1.21}$$

$$\hat{\sigma}^2 = \frac{S_*(\boldsymbol{\Theta})}{T} \tag{1.22}$$

1.2.1.3 Autoregressive Moving Average Models

On some occasions, a long AR or long MA model with many parameters is often needed to capture the dynamic structure of the time series and to approximate any pattern of the model weights. However, a large number of parameters causes to a loss of efficiency in parameter estimation. Thus, instead of increasing the order of the AR or MA model, a model with few parameters is usually desirable in model building and fitting. To achieve parameter parsimony, ARMA model is proposed by Box, Jenkins, & Reinsel (1994).

An ARMA model of orders *p* and *q* consists of a linear combination of AR(*p*) and MA(*q*) processes and it is given as in Equation (1.23).

$$\begin{aligned} Y_t &= \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + Z_t - \theta_1 Z_{t-1} - \theta_2 Z_{t-2} - \cdots - \theta_q Z_{t-q} \\ &= \phi_0 + \sum_{i=1}^{p} \phi_i Y_{t-i} + Z_t - \sum_{j=1}^{q} \theta_j Z_{t-j} \end{aligned} \tag{1.23}$$

where $\phi_1, \cdots, \phi_p$ and $\theta_1, \cdots, \theta_q$ denote, respectively, the parameters of autoregressive and moving average parts, and ϕ_0 is the constant term of the model. Also, Z_t is an i.i.d. white noise sequence, $\{Z_t\} \sim WN(0, \sigma^2)$. An ARMA model of orders *p* and *q*

denoted by ARMA(p,q) is regarded as a dynamic regression model in which the p previous values of the process, and q lagged values of the innovation sequence are explanatory variables. Additionally, for $p=0$, the process reduces to MA(q) model, and it is said to be an AR(p) process for $q=0$. Using backshift operators, the ARMA(p,q) process can be expressed shortly as in Equation (1.24).

$$\Phi(B)Y_t = \phi_0 + \Theta(B)Z_t \tag{1.24}$$

where $\Phi(B)=\left(1-\phi_1 B-\cdots-\phi_p B^p\right)$ and $\Theta(B)=\left(1-\theta_1 B-\cdots-\theta_q B^q\right)$ indicate pth order autoregressive and qth order moving average polynomials, respectively. The ARMA(p,q) process is said to be stationary and invertible provided that all roots of the characteristic polynomials lie outside of the unit circle.

Taking expectation of both sides of Equation (1.23), the mean of the process is found as

$$E(Y_t)=\mu=\frac{\phi_0}{1-\phi_1-\cdots-\phi_p} \tag{1.25}$$

Also, the ARMA(p,q) model can be written in terms of deviations from the mean given as follows

$$\begin{aligned}(Y_t-\mu)&=\phi_1(Y_{t-1}-\mu)+\cdots+\phi_p(Y_{t-p}-\mu)+Z_t-\theta_1 Z_{t-1}-\theta_2 Z_{t-2}-\cdots-\theta_q Z_{t-q}\\ \dot{Y}_t&=\phi_1\dot{Y}_{t-1}+\cdots+\phi_p\dot{Y}_{t-p}+Z_t-\theta_1 Z_{t-1}-\theta_2 Z_{t-2}-\cdots-\theta_q Z_{t-q}\end{aligned} \tag{1.26}$$

Multiplying both sides of mean-adjusted form by $\dot{Y}_{t-k}$ and taking expectations, the ACVF at lag k of the model is obtained as

$$\gamma_Y(k)=\phi_1\gamma_Y(k-1)+\cdots+\phi_p\gamma_Y(k-p) \quad \text{for} \quad k\geq q+1 \tag{1.27}$$

Thus, the ACF at lag k is expressed as in Equation (1.28).

$$\rho_Y(k)=\phi_1\rho_Y(k-1)+\cdots+\phi_p\rho_Y(k-p) \quad \text{for} \quad k\geq q+1 \tag{1.28}$$

As it is understood from the property of the ACF in general class of ARMA models, one of the characteristic features of these processes is short term memory.

For a general ARMA(p,q) model with Gaussian white noise sequence, $\{Z_t\}\sim WN(0,\sigma^2)$,the joint probability density function of $\mathbf{Z}=(z_1,\cdots,z_T)'$ is

$$P\left(\mathbf{Z}\middle|\mathbf{\Phi},\mathbf{\Theta},\sigma^2\right)=L_*\left(\mathbf{\Phi},\mathbf{\Theta},\sigma^2\right)=\left(2\pi\sigma^2\right)^{-T/2}\exp\left[-\frac{1}{2\sigma^2}\sum_{t=1}^{T}Z_t^2\right] \tag{1.29}$$

where $\mathbf{\Phi}\equiv\left(\phi_0,\phi_1,\cdots,\phi_p\right)$, $\mathbf{\Theta}\equiv\left(\phi_0,\theta_1,\cdots,\theta_q\right)$, and $L_*\left(\mathbf{\Phi},\mathbf{\Theta},\sigma^2\right)$represent the parameter vectors and the conditional likelihood function, respectively. Let $\mathbf{Y}=\left(y_1,\cdots,y_T\right)'$ be the sample. Then, the conditional log-likelihood function is obtained as

$$\ell_*\left(\mathbf{\Phi},\mathbf{\Theta},\sigma^2\right)=\log L_*\left(\mathbf{\Phi},\mathbf{\Theta},\sigma^2\right)=-\frac{T}{2}\log\left(2\pi\sigma^2\right)-\frac{S_*\left(\mathbf{\Phi},\mathbf{\Theta}\right)}{2\sigma^2} \tag{1.30}$$

where $S_*\left(\mathbf{\Phi},\mathbf{\Theta}\right)=\sum_{\mathrm{t}=1}^{\mathrm{T}}Z_t^2\left(\mathbf{\Phi},\mathbf{\Theta}\middle|\mathbf{Y}_*,\mathbf{Z}_*,\mathbf{Y}\right)$ and $Z_t^2\left(\mathbf{\Phi},\mathbf{\Theta}\middle|\mathbf{Y}_*,\mathbf{Z}_*,\mathbf{Y}\right)$, respectively, denote the CSS and the residuals with the initial conditions $\mathbf{Y}_*=\left(y_{1-p},\cdots,y_0\right)'$ and $\mathbf{Z}_*=\left(z_0,\cdots,z_{1-q}\right)'$. In order to determine the initial conditions, some procedures have been proposed. Box & Jenkins (1976) suggested that the initial values are set as $\mathbf{Y}_*=\left(y_1,\cdots,y_p\right)'$ and $\mathbf{Z}_*=\left(z_p,\cdots,z_{p+1-q}\right)'=\mathbf{0}$. Hence, by setting $\mathbf{Z}_*=\mathbf{0}$, the log-likelihood and sum of squares functions conditional on $\mathbf{Y}_*$ is given as

$$\begin{aligned}\ell_*\left(\mathbf{\Phi},\mathbf{\Theta},\sigma^2\right)&=\log f_{Y_{P+1},\cdots,Y_T|Y_1,\cdots,Y_P,\mathbf{Z}_*}\left(\mathbf{Y}\middle|\mathbf{Y}_*,\mathbf{Z}_*;\mathbf{\Phi},\mathbf{\Theta}\right)\\&=-\frac{(T-p)}{2}\log\left(2\pi\sigma^2\right)-\frac{S_*\left(\mathbf{\Phi},\mathbf{\Theta}\right)}{2\sigma^2}\end{aligned} \tag{1.31}$$

$$S_*\left(\mathbf{\Phi},\mathbf{\Theta}\right)=\sum_{\mathrm{t}=p+1}^{\mathrm{T}}Z_t^2\left(\mathbf{\Phi},\mathbf{\Theta}\middle|\mathbf{Y}_*,\mathbf{Z}_*,\mathbf{Y}\right) \tag{1.32}$$

where $\mathbf{Y}=\left(y_{p+1},\cdots,y_T\right)'$ and $\log f_{Y_{P+1},\cdots,Y_T|Y_1,\cdots,Y_P,\mathbf{Z}_*}\left(\mathbf{Y}\middle|\mathbf{Y}_*,\mathbf{Z}_*;\mathbf{\Phi},\mathbf{\Theta}\right)$ are the realized values of the time series and the conditional joint probability density function, respectively. So, the conditional ML estimates, $\hat{\mathbf{\Phi}}$, $\hat{\mathbf{\Theta}}$, and $\hat{\sigma}^2$ are calculated as in Equations(1.33) and (1.34)

$$\hat{\mathbf{\Phi}}=\underset{(\mathbf{\Phi},\mathbf{\Theta})}{\arg\max}\,\ell_*\left(\mathbf{\Phi},\mathbf{\Theta},\sigma^2\right)=\underset{(\mathbf{\Phi},\mathbf{\Theta})}{\arg\min}\,S_*\left(\mathbf{\Phi},\mathbf{\Theta}\right) \tag{1.33}$$

$$\hat{\sigma}^2=\frac{S_*\left(\mathbf{\Phi},\mathbf{\Theta}\right)}{T-\left(2p+q+1\right)} \tag{1.34}$$

1.2.2 Conditionally Heteroscedastic Time Series Models

The traditional econometric techniques used to explain economic events have some assumptions which are i.i.d. errors and homoscedasticity. As previously stated, in the economic and financial time series, the assumption of homogeneity of error variances is generally violated, indicating the presence of autocorrelation in the errors. Hence, the Autoregressive Conditionally Heteroscedastic Model (ARCH) and its extended version, Generalized Autoregressive Conditionally Heteroscedastic Model (GARCH) have been introduced respectively by Engle (1982) and Bollerslev (1986) to capture the stylized facts of financial asset returns such as leptokurtic (heavy-tailed) distribution of the returns, time-varying volatility, volatility clustering etc. Besides these, the different volatility models have been suggested depending on the empirical findings of the return series. For example, the exponential GARCH (EGARCH) of Nelson (1991), the quadratic ARCH (QARCH) of Sentana (1995) that allow to model asymmetry in the heteroscedastic processes, nonlinear GARCH (NGARCH) of Higgins & Bera (1992) which takes into account leverage effect etc. are available in the literature.

Throughout this Chapter, Y_t represents the logarithmic return series of the financial asset and it is calculated as follows

$$Y_t = \log\left(\frac{P_t}{P_{t-1}}\right) \tag{1.35}$$

where P_t is price of the underlying asset at time *t*. Also, the conditional variance (volatility) of Y_t is defined as

$$Var\left(Y_t \middle| F_{t-1}\right) = E\left(Y_{t-1}^2 \middle| F_{t-1}\right) = \sigma_t^2 \tag{1.36}$$

where $F_{t-1} = \sigma\left(Y_{t-1}, Y_{t-2}, \ldots\right)$ indicates the set of all information available up to time $t-1$.

1.2.2.1 Autoregressive Conditionally Heteroscedastic Models

The ARCH models take into account the empirical findings in financial asset returns and also have many application fields such as asset pricing, option pricing modeling risk premiums and exchange rates. The most important concept for ARCH models is the conditional variance given the past (volatility). In *p*th order autoregressive conditional heteroscedastic, ARCH(*p*), model, the unconditional variance of the model is constant but the variance conditional on the *p* past values of the process changes over time. The volatility is expressed as a quadratic function of the *p* lagged values of the return series. An ARCH(*p*) process, Y_t, and the volatility of the process, σ_t^2, are described as follows

$$Y_t = \sigma_t \varepsilon_t \tag{1.37}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i Y_{t-i}^2 \tag{1.38}$$

where ε_t is assumed to be a sequence of i.i.d. random variables with zero mean and variance 1, ω and α_i for $i = 1, \ldots, p$ are the unknown parameters. The sufficient condition to ensure a non-negative variance ($\sigma_t^2 \geq 0$) is $\omega \geq 0$ and $\alpha_i \geq 0$ for $i = 1, \ldots, p$. The unconditional mean of the process is obtained as in Equation (1.39).

$$E(Y_t) = E\left[E\left(Y_t | F_{t-1}\right)\right] = E\left[\sigma_t E(\varepsilon_t)\right] = 0 \tag{1.39}$$

If $\sum_{i=1}^{p} \alpha_i < 1$, then ARCH(*p*) process is said to be weakly stationary and its unconditional variance is given by

$$\begin{aligned} Var(Y_t) &= E\left(Y_t^2\right) = E\left[E\left(Y_t^2 | F_{t-1}\right)\right] \\ &= E\left(\omega + \sum_{i=1}^{p} \alpha_i Y_{t-i}^2\right) \\ &= \omega + \sum_{i=1}^{p} \alpha_i E\left(Y_{t-i}^2\right) \\ &= \frac{\omega}{1 - \sum_{i=1}^{p} \alpha_i} \end{aligned} \tag{1.40}$$

An ARCH(*p*) model can be expressed as in the form of AR(*p*) by defining a new noise sequence as $\nu_t = Y_t^2 - \sigma_t^2$ with zero mean and it is obtained as

$$Y_t^2 = \omega + \sum_{i=1}^{p} \alpha_i Y_{t-i}^2 + \nu_t \tag{1.41}$$

1.2.2.2 Generalized Autoregressive Conditionally Heteroscedastic Models

In many empirical applications, the ARCH models frequently require a large number of lags of squared return series to adequately model the volatility process. This causes a non-parsimonious volatility model, and the non-negativity constraint on the model parameters may not be ensured in a long ARCH model. To overcome this difficulty, the GARCH(*p*,*q*) model has been proposed as an improved form of the ARCH(*p*) model by linearly adding the *q* lagged values of the conditional variance in expressing current volatility. Hence, the GARCH model is more preferable, useful and a more parsimonious way than a high-order ARCH model in modeling volatility dynamics. Also, the GARCH(*p*,*q*) model reduces to an ARCH(*p*) process if *q* is equal to zero. The GARCH(*p*,*q*) process has the following representation

$$Y_t = \sigma_t \varepsilon_t \tag{1.42}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i Y_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2, \quad t = 1, \cdots, T \tag{1.43}$$

where $\{\varepsilon_t\}$ is a sequence of i.i.d. random variables with zero mean and unit variance and $E(\varepsilon_t^4) < \infty$, ω, α_i and β_j are unknown parameters satisfying $\omega \geq 0$, $\alpha_i \geq 0$ and $\beta_j \geq 0$ for $i = 1, \ldots, p$ and $j = 1, \ldots, q$. The stochastic process σ_t is assumed to be independent of ε_t.

If $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$ holds, then the GARCH(*p*,*q*) model with the positive constant, $\omega > 0$, is second-order stationary. The unconditional variance of GARCH(*p*,*q*) process is obtained as in Equation (1.44) given below.

$$
\begin{aligned}
Var(Y_t) &= E(Y_t^2) = E\left[E(Y_t^2 | F_{t-1})\right] \\
&= E\left(\omega + \sum_{i=1}^{p} \alpha_i Y_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \right) \\
&= \omega + \sum_{i=1}^{p} \alpha_i E(Y_{t-i}^2) + \sum_{j=1}^{q} \beta_j E(\sigma_{t-j}^2) \\
&= \frac{\omega}{1 - \sum_{i=1}^{p} \alpha_i - \sum_{j=1}^{q} \beta_j}
\end{aligned}
\tag{1.44}
$$

The conditional variance in the GARCH models has a very similar mathematical expression with the conditional expectation in the ARMA models. In this dissertation, we assume that the process $\{Y_t\}$ is strictly stationary, i.e., $\sum_{i=1}^{r}(\alpha_i + \beta_i) < 1$ where $r = \max(p, q)$ and the strict stationary conditions of Y_t as in Bollerslev (1986), Bougerol & Picard (1992a), Bougerol & Picard (1992b) hold. A GARCH(*p*,*q*) process $\{Y_t\}$ is represented in the form of ARMA(*r*,*q*) as follows

$$
Y_t^2 = \omega + \sum_{i=1}^{r}(\alpha_i + \beta_i) Y_{t-i}^2 + \nu_t - \sum_{j=1}^{q} \beta_j \nu_{t-j} \tag{1.45}
$$

where $\alpha_i = 0$ for $i > p$ and $\beta_i = 0$ for $i > q$, and the innovation $\nu_t = Y_t^2 - \sigma_t^2$ is a white noise (not i.i.d. in general) and identically distributed under the strict stationary assumption of Y_t.

1.3 Resampling Methods

The non-parametric bootstrap which is a computer-intensive method was introduced by Efron(1979). This technique provides good solutions to almost all statistical inference problems without making any distributional assumption, and it has become a general framework to approximate the sampling distributions of the statistics and wide variety of statistical applications. Also, this method provides more accurate approximations than first order asymptotic theory. The main idea behind the bootstrap is to rebuild the relation between the "population" and the "sample". The sample plays the role of the population, and resamples are repeatedly generated from

this sample to form an analogue of the given sample. So, the resample together with the sample at hand can be used to picturize the relation between population and sample if the resample scheme is implemented properly. In this way, the statistician uses the sample and the resamples instead of dealing with unknown population distribution and its characteristics.

Suppose that $X_1, X_2 \ldots$ is a sequence of i.i.d. random variables with common unknown distribution *F*. Let $\chi_n = \{X_1, X_2, \ldots, X_n\}$ denotes the random sample of size *n*. Also, let $T_n = t_n(\chi_n; F)\ n \geq 1$ be a random variable of interest with some functional t_n, where T_n depends on only the data and underlying distribution function *F*. The most commonly used statistics, T_n, are the normalized sample mean $T_n \equiv \sqrt{n}(\bar{X}_n - \mu)/\sigma$ and the studentized sample mean $T_n \equiv \sqrt{n}(\bar{X}_n - \mu)/s_n$ where $\bar{X}_n = n^{-1}\sum_{i=1}^{n} X_i$, $s_n^2 = n^{-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$, $\mu = E(X_1)$ and $\sigma^2 = Var(X_1)$. The aim is to approximate the sampling distribution of T_n, G_n, or to approximate some important population characteristics of T_n, e.g., the standard error, the confidence interval, or the critical value of T_n.

A simple random sample of size *n* $\chi_n^* = \{X_1^*, X_2^*, \ldots, X_n^*\}$ is drawn with replacement from a given sample χ_n. Here, the bootstrapped values $\{X_1^*, X_2^*, \ldots, X_n^*\}$ are i.i.d. random variables conditional on χ_n, and the conditional probability is given by

$$P_*(X_1^* = X_i) = \frac{1}{n}, \quad 1 \leq i \leq n \tag{1.46}$$

where P_* denotes the conditional probability given χ_n. So, the common distribution of X_i's is the empirical distribution function given by

$$F_n = \frac{1}{n}\sum_{i=1}^{n} I(X_i \leq x) \text{ where } I(X_i \leq x) = \begin{cases} 1 & \text{if} \quad X_i \leq x \\ 0 & \text{if} \quad X_i > x \end{cases} \tag{1.47}$$

where $I(.)$ denotes the indicator function.

The bootstrap analogue of T_n, T_n^*, is obtained by using plug-in principle, that is by replacing χ_n with χ_n^* and F with F_n as

$$T_n^* = t_n\left(\chi_n^*; F_n\right) \tag{1.48}$$

Conditional on χ_n, the bootstrap estimator of unknown sampling distribution of T_n, G_n^*, is obtained by using plugging-in principle so that the bootstrap estimator of $\varphi(G_n)$ is given by $\varphi(G_n^*)$ with some functional $\varphi(.)$. For instance, if $\varphi(G_n) = Var(T_n) = \int x^2 dG_n(x) - \left(\int x dG_n(x)\right)^2$, then the bootstrap estimator is given by $\varphi(G_n^*) = Var(T_n^* | \chi_n) = \int x^2 dG_n^*(x) - \left(\int x dG_n^*(x)\right)^2$. In addition, the bootstrapped versions of the standardized and studentized sample means are respectively given by $T_n^* = \sqrt{n}\left(\overline{X}_n^* - \overline{X}_n\right)/s_n$ and $T_n^* = \sqrt{n}\left(\overline{X}_n^* - \overline{X}_n\right)/\sigma^*$ where $\sigma^{*2} = n^{-1}\sum_{i=1}^{n}\left(X_i^* - \overline{X}_n^*\right)^2$. The bootstrap provides consistent estimators for the sampling distribution of T_n, and it is well-known that the bootstrap sampling distribution is a strong consistent estimator under Kolmogorov metric.

Under some moment conditions, the sampling distribution of i.i.d. bootstrap is consistent with the distribution of statistics of interest. However, this method does not always provide satisfactory answers for all statistical inference problems or all types of random processes. Singh (1981) gave a simple example of the inadequacy of the i.i.d. bootstrap for dependent data. Let $\{X_n\}_{n\geq 1}$ be a *m*-dependent sequence of random variables with $EX_1 = \mu$ and $EX_1^2 < \infty$ for some positive integer $m \geq 0$ as long as $\{X_1, \ldots, X_t\}$ and $\{X_{t+m+1}, \ldots\}$ are independent for all $t \geq 1$. Let $\sigma_m^2 = Var(X_1) + 2\sum_{i=1}^{m-1} Cov(X_1, X_{1+i}) < \infty$. The Central Limit Theorem for *m*-dependent random variables indicates that $\sqrt{n}\left(\overline{X}_n - \mu\right) \xrightarrow{d} N\left(0, \sigma_m^2\right)$. In this case, if we approximate the sampling distribution of the random variable $T_n = \sqrt{n}\left(\overline{X}_n - \mu\right)$ by the i.i.d. bootstrap sampling distribution of random variable $T_n^* = \sqrt{n}\left(\overline{X}_n^* - \overline{X}_n\right)$, then

the conditional distribution of T_n^* will converge to a normal distribution with wrong variance. So,

$$\lim_{n\to\infty}\left[P_*\left(T_n^* \le x\right) - P\left(T_n \le x\right)\right] = \left[\Phi\left(x/\sigma\right) - \Phi\left(x/\sigma_\infty\right)\right] \neq 0 \tag{1.49}$$

where $\sigma_\infty^2 = Var\left(X_1\right) + 2\sum_{i=1}^{m} Cov\left(X_1, X_{1+i}\right) < \infty$ and $\sum_{i=1}^{m} Cov\left(X_1, X_{1+i}\right) \neq 0$ and $Var\left(X_1\right) \neq 0$. This means that the i.i.d. bootstrap estimator of $P\left(T_n \le x\right)$ will have non-zero mean squared error in the limit since it ignores the lag covariance terms in the asymptotic variance. Thus, it fails to provide a consistent estimator for the distribution of the random variable T_n under the dependency case. Several methods have been suggested to extend this method to the serially correlated data. Next subsection describes these methods in detail.

1.3.1 Resampling Methods for Dependent Data

In the literature, several resampling techniques have been proposed to motivate the bootstrap methods to the dependent data. These techniques can be divided into the two main groups; residual based and block based bootstrap methods. Lahiri (2003) provides an excellent overview of bootstrap methods for dependent data. The subsection 1.3.1.1 presents the residual based bootstrap while subsection 1.3.1.2 illustrates some of the block bootstrap methods which are studied in this dissertation.

1.3.1.1 Residual Based Bootstrap

Let us suppose that the sequence of random variables $\{X_n\}_{n\ge1}$ satisfies the equality given by

$$X_n = g\left(X_{n-1}, \ldots, X_{n-p}; \alpha\right) + \varepsilon_n \tag{1.50}$$

where $n > p$, α is a $q \times 1$ vector of unknown parameters, $g : R^{p+q} \to R$ is a known Borel measurable function, and $\{\varepsilon_n\}_{n>p}$, which is independent of $X_i\ i = 1, \ldots, p$ and $E\left(\varepsilon_1\right) = 0$, is a sequence of i.i.d. random variables with unknown distribution *F*. This is an example of autoregressive model of order *p*. The i.i.d. bootstrap technique can

easily be extended to this process since the ε_i's are i.i.d. Let $\chi_n = \{X_1, X_2, \ldots, X_n\}$ denotes the random sample of size *n*, and $\hat{\alpha}_n$ be an estimator of α. Let us consider the estimation of the sampling distribution of the random variable $T_n = t_n(\chi_n; F, \alpha)$ with some functional t_n. Then, the algorithm of residual bootstrap can be defined as follows

- First, obtain the residuals by using some regression models such as least square $\hat{\varepsilon}_i = X_i - g(X_{i-1}, \ldots, X_{i-p}; \hat{\alpha}_n), \quad p < i \le n$.
- Define the centered residuals $\tilde{\varepsilon}_i = \hat{\varepsilon}_i - \bar{\varepsilon}_n, \quad p < i \le n$ where $\bar{\varepsilon}_n \equiv (n-p)^{-1} \sum_{i=1}^{n-p} \hat{\varepsilon}_{i+p} \neq 0$.
- Draw a simple random sample of size $(m-p)$ from $\{\tilde{\varepsilon}_i, p < i \le n\}$ with replacement and obtain $\varepsilon^*_{p+1}, \ldots, \varepsilon^*_m$ to generate bootstrap pseudo-observations by the model

$$X^*_i = X_i \quad \text{for} \quad i = 1, \ldots p$$
$$X^*_i = g(X^*_{i-1}, \ldots, X^*_{i-p}; \hat{\alpha}_n) + \varepsilon^*_i, \quad p < i \le m$$

- Obtain the bootstrap estimator of $T_n = t_n(\chi_n; F, \alpha)$ by $T^*_n = t_n(\chi^*_n; F_n, \hat{\alpha}_n)$ where $\chi^*_n = (X^*_1, \ldots, X^*_n)$ and F_n represents the empirical distribution function of the centered residuals $\tilde{\varepsilon}_i$.

1.3.1.2 Block Bootstrap Methods

In block based bootstrap methods, the series of size *n* is divided into blocks consisting of consecutive observations. These blocks are then resampled with equal probability, and then pasted end-to-end to form the bootstrap series, in which the dependency structure of the original data can be preserved, at least within the adjacent observations in each block. The blocks may be comprised of nonoverlapping or overlapping subsets from the original series. Non-overlapping block bootstrap (NBB) approach is proposed by Carlstein (1986), and overlapping

blocks known as moving block bootstrap (MBB) are independently proposed by Kunsch (1989) and Liu & Singh (1992). In addition to these methods, the circular block bootstrap (CBB) method suggested by Politis & Romano (1992a) and Shao & Yu (1993). Also, Politis & Romano (1992b) suggested blocks of blocks bootstrap procedure in order to obtain valid inference for the parameters of the infinite dimensional joint distribution process. Moreover, Politis & Romano (1994) proposed stationary bootstrap method with random block lengths where they have a geometric distribution.

Suppose $\{X_n\}_{n\geq 1}$ be a sequence of stationary random variables, and let $\chi_n = \{X_1, X_2, \ldots, X_n\}$ denotes the sample. Let ℓ denotes the block length where $1 \leq \ell \leq n$, $\ell \in Z^+$, satisfying $l \to \infty$ and $n^{-1}\ell \to 0$ as $n \to \infty$. Let $B_i = (X_i, \ldots, X_{i+\ell-1})$ denote the blocks with size ℓ where $1 \leq i \leq N$ and $N = n - \ell + 1$. It should be noted that each block has the same size, ℓ. A suitable number of blocks are drawn from the collection of blocks $\{B_1, \ldots, B_N\}$ to obtain the bootstrap sample by using MBB technique. In other words, the bootstrap analogue of the original data $B_1^*, \ldots, B_k^*$ are drawn with replacement from the set $\{B_1, \ldots, B_N\}$. Then, these block are pasted end-to-end to obtain bootstrap sample, $X_1^*, \ldots, X_m^*$. Here, $B_i^* = (X_{(i-1)\ell+1}, \ldots, X_{i\ell})$, and $m \equiv k\ell$ denotes the bootstrap sample size. Suppose we want to estimate $\hat{\theta}_n = T(F_n)$ where F_n and $T(.)$ denote the empirical distribution function of $X_1, \ldots, X_n$ and a functional of F_n, respectively. Then, the corresponding MBB estimator of $\hat{\theta}_n$ is described as $\theta_m^* = T(F_m^*)$ where F_m^* represents the empirical distribution function of bootstrapped resample. Note that the bootstrap conditional probability for overlapping blocks is $P_*(B_j^* = B_i) = N^{-1}$ for $1 \leq i \leq N$ and $1 \leq j \leq k$.

Suppose we are interested in estimating the sample mean of $X_1, X_2, \ldots, X_n$, $\hat{\theta}_n = n^{-1}\sum_{j=1}^{n} X_j$. Then, the MBB estimator of $\hat{\theta}_n$ is $\hat{\theta}_m^* = m^{-1}\sum_{j=1}^{m} X_j^*$. So we get the expression for the MBB sample mean

$$\begin{aligned} E_*\left(\hat{\theta}_m^*\right) &= E_*\left(\ell^{-1}\sum_{i=1}^{\ell} X_i^*\right) \\ &= N^{-1}\sum_{j=1}^{N}\left(\ell^{-1}\sum_{i=1}^{\ell} X_{j+i-1}\right) \\ &= N^{-1}\left\{n\overline{X}_n - \ell^{-1}\sum_{j=1}^{\ell-1}(\ell-j)\left(X_j + X_{n-j+1}\right)\right\} \end{aligned} \tag{1.51}$$

In the NBB method, the data of size *n* is divided into the *b* consecutive blocks which do not contain the same observations. As in the MBB method, the dependency structure of the original data is preserved with the adjacent observations into the blocks. However, under the NBB method, the number of blocks which consist of nonoverlapping segments of the data is less than the collection of blocks obtained from the MBB method. Again, let $\ell \in [1, n]$ be an integer and denotes the block length, also let $b \geq 1$ is an integer such that $b = \lfloor n/\ell \rfloor$ where $\lfloor x \rfloor$ denotes the largest integer not exceeding x. Then, the blocks of size ℓ obtained by nonoverlapping observations are defined as $B_i = \left(X_{(i-1)\ell+1}, \ldots, X_{i\ell}\right)$, $i = 1, \ldots, b$. To obtain NBB analogue of χ_n, χ_m^* where $m \equiv b\ell$, $B_1^*, \ldots, B_b^*$ are drawn with replacement from the collection of $\{B_1, \ldots, B_b\}$. In this technique, the bootstrap conditional probability for the non-overlapping block is $P_*\left(B_j^* = B_i\right) = b^{-1}$ for $1 \leq i, j \leq b$. Let F_m^* denotes the empirical distribution of the NBB sample $\left\{\left(X_1^*, \ldots, X_\ell^*\right), \left(X_{(\ell+1)}^*, \ldots, X_{2\ell}^*\right), \ldots, \left(X_{(b-1)\ell+1}^*, \ldots, X_m^*\right)\right\}$. Suppose that we are interested in estimating some functional of the sample distribution $\hat{\theta}_n = T\left(F_n\right)$, then the corresponding NBB estimator is obtained by $\hat{\theta}_m^* = T\left(F_m^*\right)$ by using plug-in principle.

Let us consider the aforementioned example where $\hat{\theta}_n = n^{-1}\sum_{j=1}^{n} X_j$. The NBB version of $\hat{\theta}_n$ is $\hat{\theta}_m^* = m^{-1}\sum_{j=1}^{m} X_j^*$. Then, we get the expression for NBB estimator as

$$\begin{aligned} E_*\left(\hat{\theta}_m^*\right) &= E_*\left(\ell^{-1}\sum_{i=1}^{\ell} X_i^*\right) \\ &= b^{-1}\sum_{j=1}^{b}\left(\ell^{-1}\sum_{i=1}^{\ell} X_{(j-1)\ell+1}\right) \\ &= (b\ell)^{-1}\left\{n\bar{X}_n - \sum_{i=b\ell+1}^{n} X_i\right\} \end{aligned} \tag{1.52}$$

Note that if ℓ is a multiple integer of *n*, then the NBB method produces an unbiased estimator for the sample mean. Otherwise, it produces a biased estimator for $\bar{X}_n$.

Both MBB and NBB methods have some disadvantages, for example, in NBB, if the sample size *n* is not exactly divisible by ℓ, some of the observations will be left out the blocks. When MBB is used as a method, first and last $\ell - 1$ observations appear less than the rest. To overcome this problem, Politis & Romano (1992a) and Shao & Yu (1993) suggested CBB method. In this technique, the data is wrapped around a circle so that each of the observation in the original data set has an equal probability. To do this, first suppose that $Y_{n,i}$, $i \geq 1$ represents the wrapped data set. Let $i = j_i$ represents the modulo *n* such that $i = k_i n + j_i$ where $j \in [1, n]$ and $k_i \geq 0$. Then, we can define the variables $Y_{n,i}$ as $Y_{n,i} = X_{j_i}$. In other words, writing the series $X_1, X_2, \ldots, X_n$ continuously end-to-end on a line denote them as $Y_{n,i}, i \geq 1$. Let B denotes the vector of overlapping blocks of size ℓ so that $\left\{\left(X_1, \ldots, X_\ell\right), \left(X_2, \ldots, X_{\ell+1}\right), \ldots, \left(X_n, \ldots, X_{\ell-1}\right)\right\}'$. Thus, this methods has *n* blocks with size ℓ meaning each observation has an equal probability n^{-1} of being selected to CBB resample. So, the edge effects caused by the MBB and NBB methods disappear under the CBB. Accordingly, $k \geq 1$ blocks are selected from B with probability n^{-1} to construct the CBB resample $X_1^*, \ldots, X_m^*$ where $m \equiv k\ell$.

If we want to estimate the functional of empirical distribution of $X_1, X_2, \ldots, X_n$, $\hat{\theta}_n = T(F_n)$, then the CBB version of $\hat{\theta}_n$ is $\hat{\theta}_m^* = T(F_m^*)$ where F_m^* is the empirical distribution of $X_1^*, \ldots, X_m^*$ obtained by the CBB. As in the sample mean example, the CBB estimator is $\hat{\theta}_m^* = m^{-1}\sum_{j=1}^{m} X_j^*$, and its expression is obtained as follows

$$\begin{aligned} E_*\left(\hat{\theta}_m^*\right) &= E_*\left[m^{-1}\sum_{i=1}^{m} X_i^* \right] \\ &= \ell^{-1} E_*\left[\sum_{i=1}^{\ell} X_i^* \right] \\ &= \ell^{-1}\left[n^{-1}\sum_{j=1}^{n}\left\{ \sum_{i=1}^{\ell} Y_{n,(j+i-1)} \right\} \right] \\ &= \ell^{-1}\left[\ell \bar{X}_n \right] \\ &= \bar{X}_n \end{aligned} \tag{1.53}$$

If $\ell = 1$, then the constructed resample $X_1^*, \ldots, X_m^*$ by the block bootstrap methods reduces to one obtained by i.i.d. bootstrap. Otherwise, the ℓ-dimensional joint distribution of the process $\{X_n\}_{n\geq 1}$ is preserved by the adjacent observations in the resampled blocks. Finally, it should be noted that, the accuracy of the block bootstrap estimators strictly depend on the size of block length ℓ. If ℓ is too small, the bootstrap samples do not reflect the correlation structure of the original data, and if ℓ is too large, the replicates do not exhibit enough randomness for bootstrapping. Generally, ℓ is obtained in the form of $\ell = Cn^{\delta}$ where *C* is a positive constant $C > 0$ and $\delta \in (0, 1/2)$. The block length should increase as sample size *n* increase to obtain consistent estimators from the bootstrap distribution function and asymptotically correct coverage probabilities for confidence intervals. For the block length selection, Hall, Horowitz, & Jing (1995) showed that for fixed ℓ, the asymptotically optimal block length is proportional to Cn^{δ}, where *C* is a positive constant, $C > 0$, and δ is proposed to as $\delta = 1/3$ to estimate the bias and variance. Additionally, $\delta = 1/4$ and $\delta = 1/5$ are proposed to estimate the one-sided distribution function and symmetrical distribution function, respectively. Also, it is required that $\ell^{-1} + n^{-1}\ell \to 0$ as $n \to \infty$ to ensure the consistency of the bootstrap estimator.

The remaining part of this dissertation is organized as follows. In Chapter 2 we propose two new block bootstrap methods to estimate the precision and accuracy of the parameters in linear time series models. Chapter 3 extends the proposed methods to the GARCH models to obtain valid prediction intervals for future returns and volatilities. Conclusions and remarks are given in Chapter 4.

CHAPTER TWO
NEW BLOCK BOOTSTRAP METHODS: SUFFICIENT AND/OR ORDERED

2.1 Introduction

In this chapter, we propose two new ideas to improve the performance of block-based-bootstrap. Our first concept is based on using sufficient time series bootstrap methods in order to obtain better results than conventional non-overlapping block bootstrap, but with less computing time and lower standard errors of estimation. The second premise is to use ordered bootstrapped blocks, to preserve the dependency structure of the original data better. Additionally, we propose the combination of these two ideas which we call "sufficient ordered block bootstrap".

2.2 Proposed Methods

Basu (1958) and Raj & Khamis (1958) worked on unique observations in a sample with different sampling schemes, and they emphasized that using only distinct units in a sample obtained by sampling with replacement produces more efficient results than overall sample mean for estimating population mean. Singh & Sedory (2011) proposed sufficient bootstrap by implementing the findings of Basu (1958) and Raj & Khamis (1958) to the traditional bootstrapping procedure. A sufficient bootstrap sample is obtained by removing the repeated observations from the bootstrapped data. In general, let $\chi_n = \{X_1, X_2, ..., X_n\}$ be a sample of i.i.d. random variables from an unknown distribution *F*, and let $\chi_\nu^* = \{X_1^*, X_2^*, ..., X_\nu^*\}$ be a sufficient bootstrap sample of size ν, which consists of only distinct observations. Each observation in a sample with size *n* has probability $1-(1-1/n)^n$ to appear sufficient bootstrap samples. Hence, the expected sample size of the sufficient bootstrap method is $\nu = n \times [1-(1-1/n)^n]$. Since $\nu < n$, the computing time of the sufficient bootstrap will be less than conventional bootstrap. For instance, Beyaztas & Alin (2014) combined sufficient and jackknife-after-bootstrap methods, and they

saved about 30% (approximately 25 hours) of computing time in their simulation studies.

In the time series concept, sufficient bootstrap can only be used with the NBB method. Because, for MBB or CBB, although the selected blocks are distinct, the observations in blocks may be the same since the consecutive blocks have $\ell-1$ the same observations. Let our time series data be $Y=(y_1,y_2,...,y_{12})$ and let $\ell=3$, then the data are represented with blocks as $B_1=(y_1,y_2,y_3)$, $B_2=(y_4,y_5,y_6)$, $B_3=(y_7,y_8,y_9)$, and $B_4=(y_{10},y_{11},y_{12})$. Suppose the bootstrapped blocks are $B_1^*=(y_1,y_2,y_3)$, $B_2^*=(y_1,y_2,y_3)$, $B_3^*=(y_{10},y_{11},y_{12})$ and $B_4^*=(y_7,y_8,y_9)$. Then, the new data are obtained by NBB as $Y_1^*=(y_1,y_2,y_3,y_1,y_2,y_3,y_{10},y_{11},y_{12},y_7,y_8,y_9)$. However, it is obtained as $Y_{\nu_1}^*=(y_1,y_2,y_3,y_{10},y_{11},y_{12},y_7,y_8,y_9)$ when sufficient NBB (SNBB) is used. Since the first and second blocks are the same, sufficient block bootstrap discards one of them. Hence, this method uses fewer blocks compared to conventional block bootstrap. Consequently, this means that calculations are performed with fewer observations and less computing time. For sufficient NBB, each block has an equal probability $1-(1-1/b)^b$ to appear in a resample so that the expected size of a new series obtained by this method is $\nu_b=[b\ell][1-(1-1/b)^b]$.

Let ℓ be $\ell_n[1,n]$, and suppose that the number of blocks is $b=\lfloor n/\ell \rfloor$. In conventional NBB method, the original time series data are divided by *b* blocks such that $B_i=(Y_{(i-1)\ell+1},...,Y_{i\ell})$ for $i=1,...,b$. Then, *b* blocks are selected randomly from $B=\{B_1,...,B_b\}$ with equal probability, b^{-1}. On the other hand, for sufficient NBB, $b_\nu=\lfloor b[1-(1-1/b)^b] \rfloor$, which only has distinct blocks, is selected randomly from $B=\{B_1,...,B_b\}$. Let P_{ν^*} denotes the sufficient bootstrap probability operator. Following Lahiri (2003) notations, since $P_{\nu^*}[\ (Y_{(i-1)\ell+1}^*,...,Y_{i\ell}^*)=(Y_{(j-1)\ell+1},...,Y_{j\ell})\ ]=b^{-1}$ where $i,j=1,...,b$, the sufficient NBB

sample mean is $\hat{\theta}_{\nu}^{*}=\frac{1}{\nu}\sum_{j=1}^{\nu}Y_{j}^{*}=\overline{Y}_{\nu}^{*}$. Let also $U_i=\left(Y_{(i-1)\ell+1}+\ldots+Y_{i\ell}\right)/\ell$ be the average value of each block in NBB, and $U_i^{*}=\left(Y_{(i-1)\ell+1}^{*}+\ldots+Y_{i\ell}^{*}\right)/\ell$ be the bootstrap version of U_i. Then, we have the following lemma for sufficient NBB sample mean $E_{\nu^{*}}\left(\overline{Y}_{\nu}^{*}\right)=E_{\nu^{*}}\left(\frac{1}{b_{\nu}}\sum_{i=1}^{b_{\nu}}U_i^{*}\right)$, where $E_{\nu^{*}}(\cdot)$ denotes the sufficient bootstrap expected value operator.

Lemma 2.1: The sufficient NBB sample mean is an unbiased estimator for the original sample mean $\overline{Y}_n$.

Proof: The proof of *Lemma* directly follows Singh & Sedory's (2011) *Theorem 4.1*. because $U_1^{*},\ldots,U_{b_{\nu}}^{*}$ have the identical distribution,

$$E_{\nu^{*}}\left(\overline{Y}_{\nu}^{*}\right)=E_1E_2\left[\frac{1}{b_{\nu}}\sum_{i=1}^{b_{\nu}}U_i^{*}\right]=E_1\left(\frac{1}{b}\sum_{i=1}^{b}U_i\right)=(b\ell)^{-1}\left(n\overline{Y}_n\right) \tag{2.1}$$

where E_1 and E_2 denote the expected value over all NBB samples and expected value for the given number of distinct blocks b_{ν}, respectively. If ℓ is the multiple integer of *n*, then $E_{\nu^{*}}\left(\overline{Y}_{\nu}^{*}\right)$ is an unbiased estimate of $\overline{Y}_n$. In addition, the variance of $\overline{Y}_{\nu}^{*}$, which follows the *Theorem 4.2* of Singh & Sedory (2011), is obtained as

$$Var_{\nu^{*}}\left(\overline{Y}_{\nu}^{*}\right)=\left(\frac{1}{\nu}-\frac{1}{n}\right)s^2 \tag{2.2}$$

where s^2 and $Var_{\nu^{*}}(\cdot)$ denote the sample variance and sufficient bootstrap variance operator, respectively. Sufficient NBB produces less standard error for the sample mean compared to conventional NBB method, so it yields more efficient results for μ.

For the process $\{Y_i\}_{i\in Z}$ that has a strong $\alpha(\cdot)$ mixing condition (for more information about $\alpha(\cdot)$ mixing see Billingsley, (1994)), Athreya & Lahiri (2006)

showed that under mild moment ($E|Y_1|^{2+\delta} < \infty$ and $\sum_{n=1}^{\infty} \alpha(n)^{\delta/2+\delta} < \infty$ for $\delta > 0$) and $\ell^{-1} + n^{-1}\ell = o(1)$ as $n \to \infty$ conditions, the NBB variance estimator of the statistic $T_n = \sqrt{n}(\bar{Y}_n - \mu), T_n^* = \sqrt{n}(\bar{Y}_n^* - E_*(\bar{Y}_n^*))$, is $Var(T_n^*) \to \sigma_\infty^2$ as $n \to \infty$ where $\sigma_\infty^2 = \sum_{i=-\infty}^{\infty} EZ_1 Z_{1+i}$ and $Z_i = Y_i - \mu$. Also, *Theorem 17.4.3* in Athreya & Lahiri (2006) showed that $\sup_{y \in \Re^d} |P_*(T_n^* \le y) - P(T_n \le y)| \to 0$ as $n \to \infty$. This means the sampling distribution G_n and its NBB estimator $\hat{G}_n$ are consistent when ℓ goes to infinity at a slower rate compared to sample size n ($l^{-1} + n^{-1}l = o(1)$ as $n \to \infty$). Since ν is a function of n so that $\nu_b = \lfloor b\ell \rfloor \left[1 - (1 - 1/b)^b\right] \to \infty$ as $n \to \infty$, the same results are obtained for distribution function of sufficient NBB method $\hat{G}_\nu$ as $\ell^{-1} + \nu^{-1}\ell = o(1)$ as $\nu \to \infty$ and $n \to \infty$. Thus, it produces consistent estimators as well.

As a second approach, we propose the ordered NBB and its sufficient version to capture more dependence structure compared to the conventional NBB method. To do this, first we give the labels to each original block. For the same example given in the first approach, the labels are determined as $B_1 = 1$, $B_2 = 2$, $B_3 = 3$, and $B_4 = 4$, and then each bootstrapped block is sorted according to these labels $B_1^* = 1$, $B_2^* = 1$, $B_4^* = 3$, and $B_3^* = 4$, and pasted end-to-end to obtain new data. Again, to obtain sufficient ordered NBB data, the repeated blocks are discarded. Hence, the new dataset is obtained by the ordered NBB as $Y_{O,n}^* = (y_1, y_2, y_{3,} y_1, y_2, y_3, y_7, y_8, y_9, y_{10}, y_{11}, y_{12})$. However, it is obtained as $Y_{O,\nu}^* = (y_1, y_2, y_3, y_7, y_8, y_9, y_{10}, y_{11}, y_{12})$ when the sufficient ordered NBB method is used. Thus, more representative datasets are formed by these methods yielding more reliable results. The ordered versions of the NBB methods are also consistent, because sorting the bootstrapped blocks does not change the estimated values such as μ, σ^2, median, etc. (Politis, 2014 personal contact).

2.3 Numerical Results

The performances of NBB, MBB, CBB, sufficient NBB and ordered versions of NBB methods were assessed using five real-world data sets *Rio Negro*, *Liver Operation Time*, *Number of Earthquakes*, *Chemical Concentration* and *Unemployment Rate* and a simulation study with different sample sizes and block lengths. The results of our study reveal that the sufficient NBB yields almost the same results but has less computing time compared to conventional NBB, MBB, and CBB methods. They also indicate that the ordered versions of conventional and sufficient NBB results are the best, especially for sufficient ordered NBB since it provides close estimates to the true parameters under consideration.

2.3.1 Simulation Results

We performed a simulation study to assess the performances of the aforementioned block bootstrap methods, and we considered two cases where the time series datasets come from MA(2) and AR(2) processes as follows:

$$Y_t = Z_t + 0.5Z_{t-1} + 0.5Z_{t-2} \tag{2.3}$$

$$Y_t = 0.9Y_{t-1} - 0.5Y_{t-2} + Z_t \tag{2.4}$$

where $Z_t \sim N(0,1)$. The sample sizes were chosen as $[\, n_1,\ n_2 \,] = [\, 100,\ 150 \,]$. Also, we considered two different block lengths for $n_1 = 100$ as $[\, \ell_1,\ \ell_2 \,] = [\, 5,\ 10 \,]$, and three different block lengths for $n_2 = 150$ as $[\, \ell_1,\ \ell_2,\ \ell_3 \,] = [\, 5,\ 10,\ 15 \,]$. The comparisons were done in terms of the estimation of μ, σ^2, θ's, ϕ's, and the coverage probabilities of true parameters for nominal size $\alpha = 0.05$. For each statistic, S = 1000 simulations were performed, and for each simulation, B = 1000 resamples of size $n = b\ell$ were created. It should be noted that the values given in parentheses are the standard errors of the corresponding estimation. The results are presented in Tables 2.1-2.4.

For both sample sizes, the sufficient NBB yields almost the same but more efficient results for μ compared to the conventional NBB. In addition, the estimated values of the parameters for MA(2) and AR(2) processes obtained by both methods are similar.

Table 2.1 Simulation results when n_1 = 100, ℓ = 5, and ℓ = 10 for the MA (2) process with parameters $\theta_1 = \theta_2 = 0.5$ and $\mu = 0, \sigma^2 = 1$

Block length	Model	Conventional NBB	Sufficient NBB	Ordered NBB	Sufficient Ordered NBB	MBB	CBB
ℓ =5	**MA(2)**						
	μ	0.0023	0.0023	0.0023	0.0023	0.0030	0.0016
		(0.1743)	(0.1358)	(0.1743)	(0.1358)	(0.1742)	(0.1757)
	σ^2	1.4470	1.4679	1.4470	1.4679	1.4440	1.4456
	θ_1	0.3815	0.3892	0.4571	0.4526	0.3774	0.3726
		(0.0987)	(0.0986)	(0.0961)	(0.0957)	(0.0992)	(0.0999)
	Cov. Prob.	0.799	0.832	0.964	0.929	0.817	0.806
	θ_2	0.2486	0.2550	0.4660	0.4010	0.2365	0.2300
		(0.1123)	(0.1235)	(0.1588)	(0.1274)	(0.1075)	(0.1072)
	Cov. Prob.	0.367	0.542	0.991	0.956	0.133	0.102
ℓ =10	**MA(2)**						
	μ	-0.0066	-0.0067	-0.0066	-0.0067	-0.0084	-0.0073
		(0.1756)	(0.1402)	(0.1756)	(0.1402)	(0.1763)	(0.1773)
	σ^2	1.4382	1.4587	1.4382	1.4587	1.4366	1.4335
	θ_1	0.4396	0.4482	0.4576	0.4793	0.4310	0.4268
		(0.0950)	(0.0915)	(0.0963)	(0.0883)	(0.0949)	(0.0957)
	Cov. Prob.	0.904	0.905	0.912	0.917	0.907	0.900
	θ_2	0.3661	0.3838	0.4054	0.4600	0.3495	0.3409
		(0.1037)	(0.1179)	(0.0999)	(0.1138)	(0.1057)	(0.1060)
	Cov. Prob.	0.799	0.910	0.875	0.943	0.823	0.805

On the other hand, ordered NBB and its sufficient counterpart enable more reliable results. The sufficient ordered NBB method produces close estimations to the true parameter values especially to the second parameter with the increasing ℓ. Also, it has better coverage probability values compared to other methods. The results are not surprising since the ordered counterparts of the NBB method preserve the dependency structure of the original data more than when obtained from the blocks. The MBB and CBB results are not satisfactory, compared to the ordered methods. According to our results, using sufficient bootstrap reduces the computing time as well. The calculations were carried out using R 3.0.2 on an Intel Core i7-3630QM

2.4 GHz PC, and the computing time (in hours) was obtained for the case where n = 150 and ℓ = 15 as in Table 2.5. As seen from Table 2.5, using sufficient NBB saves about 28% computing time compared to the conventional NBB, and this recovery of time will be highly important for practitioners whose studies require more computing time.

We plotted the sampling distributions of all block bootstrap methods used in simulation study for θ_1, θ_2, ϕ_1, and ϕ_2 based on 1000 bootstrap simulation results where n = 150 and ℓ = 10, and these plots are presented in Figure 2.1. According to these plots, the true values of the parameters are covered in an optimum manner by the sufficient ordered NBB method and this clearly explains the results of coverage probabilities.

Table 2.2 Simulation results when n_1 = 100, ℓ = 5, and ℓ = 10 for the AR (2) process with parameters $\phi_1 = 0.9,\ \phi_2 = -0.5$ and $\mu = 0,\ \sigma^2 = 1$

Block length	Model	Conventional NBB	Sufficient NBB	Ordered NBB	Sufficient Ordered NBB	MBB	CBB
	AR(2)						
	μ	0.0026	0.0024	0.0026	0.0024	0.0021	0.0019
		(0.1818)	(0.1419)	(0.1818)	(0.1419)	(0.1828)	(0.1842)
	σ^2	2.0343	2.0599	2.0343	2.0599	2.0334	2.0331
ℓ =5	ϕ_1	0.6046	0.6177	0.6745	0.7660	0.5974	0.5838
		(0.0982)	(0.1045)	(0.0957)	(0.1087)	(0.1008)	(0.1023)
	Cov. Prob.	0.080	0.152	0.241	0.764	0.007	0.005
	ϕ_2	-0.2958	-0.2962	-0.3691	-0.4037	-0.2840	-0.2743
		(0.0884)	(0.0957)	(0.0976)	(0.096)	(0.0890)	(0.0900)
	Cov. Prob.	0.255	0.367	0.737	0.838	0.118	0.083
	AR(2)						
	μ	0.0002	0.0002	0.0002	0.0002	0.0025	-0.0001
		(0.1611)	(0.1287)	(0.1611)	(0.1287)	(0.1635)	(0.1642)
	σ^2	2.0721	2.0941	2.0721	2.0941	2.0688	2.0684
ℓ =10	ϕ_1	0.7319	0.7452	0.7763	0.8209	0.7238	0.7121
		(0.0953)	(0.0951)	(0.0951)	(0.0882)	(0.0989)	(0.0999)
	Cov. Prob.	0.568	0.652	0.695	0.788	0.570	0.517
	ϕ_2	-0.3955	-0.4027	-0.4283	-0.4627	-0.3860	-0.3767
		(0.0852)	(0.0868)	(0.0815)	(0.0809)	(0.0887)	(0.0890)
	Cov. Prob.	0.783	0.821	0.842	0.888	0.794	0.769

Table 2.3 Simulation results when n_2 = 150, ℓ = 5, ℓ = 10, and ℓ = 15 for the MA (2) process with parameters $\theta_1 = \theta_2 = 0.5$ and $\mu = 0, \sigma^2 = 1$

Block length	Model	Conventional NBB	Sufficient NBB	Ordered NBB	Sufficient Ordered NBB	MBB	CBB
	MA(2)						
	μ	0.0060	0.0060	0.0060	0.0060	0.0052	0.0055
		(0.1444)	(0.1117)	(0.1444)	(0.1117)	(0.1446)	(0.1450)
	σ^2	1.4697	1.4838	1.4697	1.4838	1.4695	1.4685
ℓ =5	θ_1	0.3832	0.3880	0.4609	0.4507	0.3792	0.3750
		(0.0790)	(0.0756)	(0.0742)	(0.0732)	(0.0791)	(0.0797)
	Cov. Prob.	0.698	0.697	0.959	0.900	0.692	0.679
	θ_2	0.2474	0.2509	0.4766	0.3952	0.2390	0.2339
		(0.0859)	(0.0916)	(0.1171)	(0.0938)	(0.0841)	(0.0840)
	Cov. Prob.	0.101	0.182	0.994	0.890	0.002	0.002
	MA(2)						
	μ	-0.0034	-0.0031	-0.0034	-0.0031	-0.0032	-0.0026
		(0.1485)	(0.1167)	(0.1485)	(0.1167)	(0.1482)	(0.1488)
ℓ =10	σ^2	1.4614	1.4758	1.4614	1.4758	1.4612	1.4600
	θ_1	0.4379	0.4428	0.4579	0.4765	0.4340	0.4310
		(0.0761)	(0.0703)	(0.0773)	(0.0680)	(0.0759)	(0.0765)
	Cov. Prob.	0.895	0.880	0.919	0.905	0.889	0.879
	θ_2	0.3592	0.3694	0.3997	0.4497	0.3503	0.3433
		(0.0807)	(0.0860)	(0.0782)	(0.0854)	(0.0828)	(0.0834)
	Cov. Prob.	0.613	0.760	0.787	0.934	0.625	0.570
	MA(2)						
	μ	-0.0016	-0.0015	-0.0016	-0.0015	-0.0005	-0.0018
		(0.1483)	(0.1185)	(0.1483)	(0.1185)	(0.1485)	(0.1490)
	σ^2	1.4567	1.4706	1.4567	1.4706	1.4557	1.4570
ℓ =15	θ_1	0.4529	0.4578	0.4663	0.4796	0.4489	0.4447
		(0.0738)	(0.0679)	(0.0732)	(0.0651)	(0.0742)	(0.0748)
	Cov. Prob.	0.881	0.873	0.911	0.895	0.902	0.896
	θ_2	0.4114	0.4216	0.4445	0.4743	0.3996	0.3921
		(0.0812)	(0.0841)	(0.0826)	(0.0794)	(0.0818)	(0.0823)
	Cov. Prob.	0.852	0.895	0.906	0.930	0.874	0.841

Table 2.4 Simulation results when n_2 = 150, ℓ = 5, ℓ = 10 and ℓ = 15 for the AR (2) process with parameters $\phi_1 = 0.9,\ \phi_2 = -0.5$ and $\mu = 0,\ \sigma^2 = 1$

Block length	**Model**	**Conventional NBB**	**Sufficient NBB**	**Ordered NBB**	**Sufficient Ordered NBB**	**MBB**	**CBB**
	AR(2)						
	μ	-0.0025	-0.0024	-0.0025	-0.0024	-0.0029	-0.0024
		(0.1506)	(0.1165)	(0.1506)	(0.1165)	(0.1509)	(0.1518)
	σ^2	2.0664	2.0836	2.0664	2.0836	2.0677	2.0675
ℓ =5	ϕ_1	0.6079	0.6165	0.6815	0.7698	0.6010	0.5918
		(0.0800)	(0.0838)	(0.0778)	(0.0880)	(0.0819)	(0.0827)
	Cov. Prob.	0.013	0.021	0.092	0.648	0.000	0.000
	ϕ_2	-0.2893	-0.2891	-0.3672	-0.4011	-0.2797	-0.2730
		(0.0723)	(0.0776)	(0.0809)	(0.0791)	(0.0730)	(0.0735)
	Cov. Prob.	0.058	0.088	0.603	0.755	0.001	0.002
	AR(2)						
	μ	0.0013	0.0012	0.0013	0.0012	0.0006	0.0001
		(0.1349)	(0.1060)	(0.1349)	(0.1060)	(0.1369)	(0.1376)
ℓ =10	σ^2	2.0496	2.0649	2.0496	2.0649	2.0491	2.0465
	ϕ_1	0.7383	0.7474	0.7852	0.8282	0.7310	0.7223
		(0.0797)	(0.0784)	(0.0799)	(0.0733)	(0.0820)	(0.0827)
	Cov. Prob.	0.431	0.491	0.661	0.771	0.405	0.357
	ϕ_2	-0.3861	-0.3906	-0.4213	-0.4552	-0.3780	-0.3712
		(0.0708)	(0.0708)	(0.0681)	(0.0668)	(0.0734)	(0.0739)
	Cov. Prob.	0.641	0.684	0.784	0.868	0.622	0.579
	AR(2)						
	μ	0.0035	0.0035	0.0070	0.0035	0.0039	0.0030
		(0.1295)	(0.1034)	(0.1303)	(0.1034)	(0.1302)	(0.1312)
	σ^2	2.0630	2.0776	2.0757	2.0776	2.0650	2.0633
ℓ =15	ϕ_1	0.7912	0.8006	0.8217	0.8516	0.7823	0.7726
		(0.0759)	(0.0726)	(0.0752)	(0.0668)	(0.0787)	(0.0796)
	Cov. Prob.	0.674	0.697	0.777	0.818	0.677	0.641
	ϕ_2	-0.4282	-0.4331	-0.4530	-0.4748	-0.4203	-0.4124
		(0.0689)	(0.0661)	(0.0680)	(0.0615)	(0.0715)	(0.0722)
	Cov. Prob.	0.817	0.821	0.861	0.856	0.823	0.790

Table 2.5 Computing time for all bootstrap methods (in hours), n = 150, ℓ = 15

Method	Conventional NBB	Sufficient NBB	MBB	CBB
Time	6.1	4.4	7.0	7.0

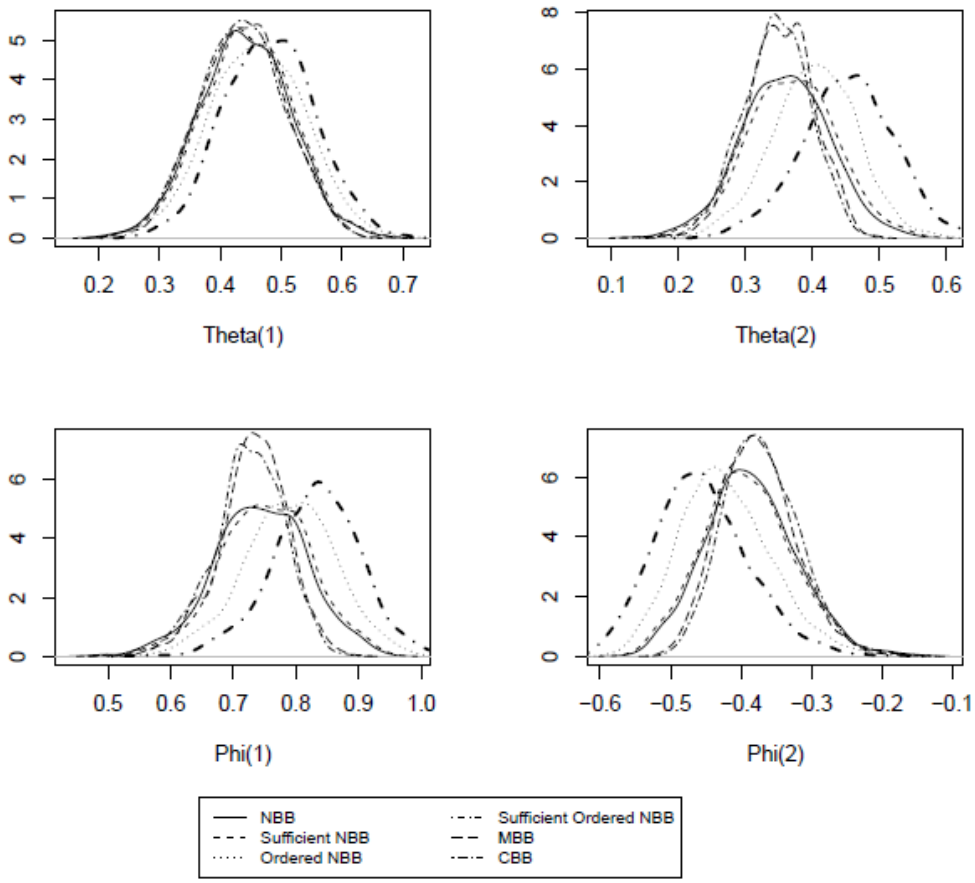

Figure 2.1 Density plots of block bootstrap methods for θ_1, θ_2, ϕ_1 and ϕ_2, respectively, when n = 150 and ℓ = 10

2.3.2 Real-world Examples

2.3.2.1 Rio-Negro Data

The real-world data called "Rio Negro" includes monthly averages of the daily stages (height) of the RioNegro at Manaus from January 1903 to December 1992, and it is available in R "*boot*" package. In the previous studies, this dataset is most likely modeled as AR(2) process. After removing the seasonal component, the parameters are calculated as $\phi_1 = 1.1042$ and $\phi_2 = -0.2952$ by using the maximum likelihood method. To compute the bootstrap estimates of the parameters, we considered two different block lengths $\ell = 12$ and $\ell = 24$, and the number of bootstrap B was chosen as 1000. The results are presented in Table 2.6, and it is clear

that when compared with the others, sufficient ordered NBB and the ordered NBB produce the best results.

Table 2.6 Numerical results for Rio Negro data ($\phi_1 = 1.1042$, $\phi_2 = -0.2952$)

Block length	Model	Conventional NBB	Sufficient NBB	Ordered NBB	Sufficient Ordered NBB	MBB	CBB
ℓ =12	**AR(2)**						
	μ	-0.0046 (0.1144)	-0.0035 (0.0882)	-0.0046 (0.1144)	-0.0035 (0.0882)	0.0152 (0.1129)	0.0003 (0.1122)
	σ^2	2.3447	2.3557	2.3447	2.3557	2.3554	2.3707
	ϕ_1	0.9588 (0.0398)	0.9604 (0.0379)	1.0049 (0.0372)	1.0398 (0.0343)	0.9282 (0.0350)	0.9213 (0.0334)
	ϕ_2	-0.1981 (0.0334)	-0.1981 (0.0330)	-0.2186 (0.0320)	-0.2462 (0.0291)	-0.1892 (0.0285)	-0.1840 (0.0271)
ℓ =24	**AR(2)**						
	μ	-0.0015 (0.1250)	0.0001 (0.0974)	-0.0015 (0.1250)	0.0001 (0.0974)	0.0214 (0.1262)	-0.0030 (0.1319)
	σ^2	2.3391	2.3502	2.3391	2.3502	2.3380	2.3480
	ϕ_1	1.0149 (0.0335)	1.0167 (0.0304)	1.0614 (0.0335)	1.0754 (0.0241)	1.0003 (0.0348)	1.0029 (0.0347)
	ϕ_2	-0.2336 (0.0276)	-0.2342 (0.0262)	-0.2662 (0.0283)	-0.2765 (0.0194)	-0.2280 (0.0285)	-0.2301 (0.0285)

The descriptive statistics of the remaining four real-world data are given in Table 2.7 Also, the descriptive statistics and Anderson-Darling normality test results for the residuals of the fitted models are given in Table 2.8.

Table 2.7 The descriptive statistics of the real-world data

	Operation time	Number of earthquake	Chemical concentration	Unemployment rate
Sample size	70	71	180	241
Mean	488.360	20.141	17.064	4.388
Std. dev.	158.750	7.374	0.395	2.776
Skewness	0.147	0.568	0.229	0.911
Kurtosis	-0.413	0.446	-0.128	-0.440
Block lengths	2, 5, 7, 10	2, 5, 7, 10	3, 5, 9, 12, 15	4, 6, 8, 10, 12, 16

Table 2.8 The descriptive statistics and normality test results for the residuals of the fitted models

Real-world data	Fitted model	Mean	Std. deviation	Skewness	Kurtosis	Anderson-Darling normality test
Liver operation time	AR(1)	-0.336	151.556	0.266	-0.214	0.231 (p-value=0.797)
Number of earthquakes	IMA(1,1)	-0.075	6.088	0.382	0.630	0.557 (p-value=0.144)
Chemical concentration	ARMA(1,1)	-0.003	0.309	0.356	0.895	0.531 (p-value=0.172)
Unemployment rate	ARI(1,1)	0.000	0.136	0.600	4.070	2.934 (p-value=0.000)

2.3.2.2 Liver transplantation operation time

The operation time of liver transplantation patients who had surgery in an university hospital between January 2003 and December 2013 is examined in detail. The liver transplantation operation time is modeled as AR(1) process with ML estimate as in Equation (2.5).

$$Y_t = 344.06 + 0.298Y_{t-1} + Z_t, \quad t \in \mathrm{Z} \tag{2.5}$$

where Z_t follows a $N(0,1)$ distribution.

As reported in Table 2.9, the calculated *p*-value < 0.05 of *t*-tests indicated that both constant term δ and ϕ_1 are statistically significant at %5 level of significance. Beside, Ljung-Box Q-statistics (please see Table 2.10) to test for autocorrelations in residuals shows that the residuals are not auto-correlated. Also, the *p-value*> 0.05 of Anderson-Darling test (see Table 2.8) indicates that the residuals are normally distributed. All of exploratory analysis indicates that AR(1) model is a suitable choice to model operation time data.

The parameter estimates and their MSEs obtained by the block bootstrap methods are presented in Figure 2.2. As it is seen from this figure, the true parameter value is captured by the SONBB method in a better way. Although the results of the SNBB are not as good as the results of the SONBB for parameter estimation, it has almost the same MSE values with the SONBB since SNBB provides smallest variances for $\hat{\phi}_1$ compared to others. It means that the variance of parameter which is obtained by

the SNBB is smaller than the one obtained by the SONBB, but more biased for parameter estimation in comparison with the SONBB. The results show that using SONBB method gets less biasedness and less error for AR(1) parameter.

The worst estimation performance belongs to the MBB, additionally CBB and SB have the similar performances which are not satisfactory. When the block length is 7, all the non-overlapping block bootstrap methods' (i.e. NBB, SNBB, ONBB, SONBB) performances are in the same manner.

Table 2.9 AR(1) model estimates for liver operation time data

Parameter	Standard error (s.e.)	t-stat.	p-value
δ	17.981	19.231	0.000
ϕ_1	0.115	2.589	0.009

Table 2.10 Ljung-Box Q-Statistics for the residuals of the estimated AR(1) model for liver operation time data

Considered lag	Ljung-Box test	Degrees of freedom	Significance level
12	5.285	10	0.871
24	10.324	22	0.983
36	20.515	34	0.967

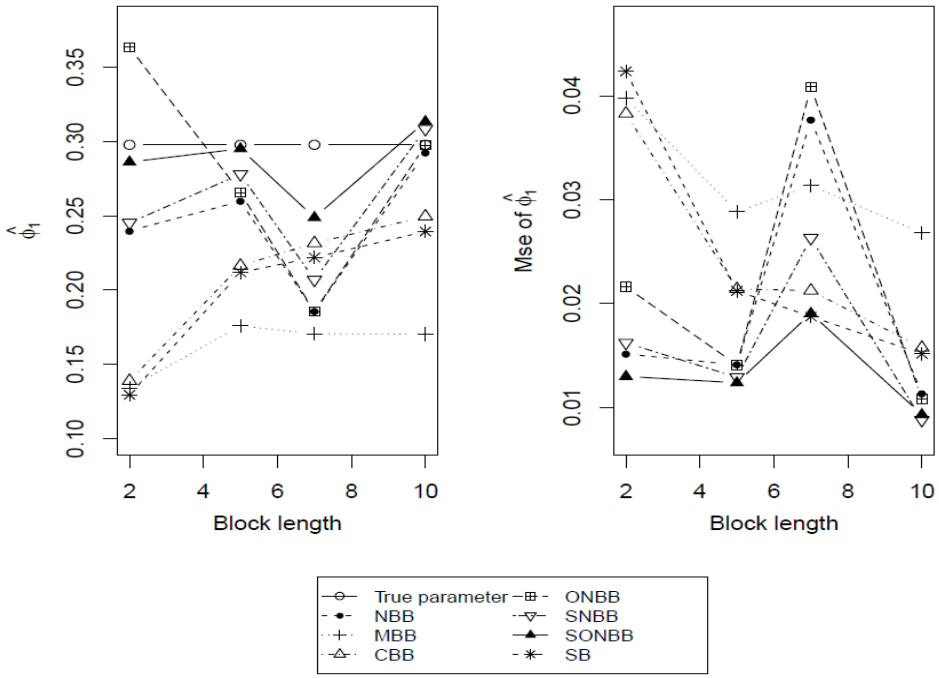

Figure 2.2 Plots of block bootstrap methods for $\hat{\phi}_1$ and its MSE for liver operation time data

2.3.2.3 Number of earthquakes

The number of earthquakes per year which has occurred greater than magnitude 7.0 over the 1928-1998 period is analyzed. The data is obtained from http://datamarket.com/data/list. For this data set, by using Box-Jenkins' methodology the integrated moving average model of order 1 (IMA(1,1)) is fitted as in Equation (2.6) and the ML estimate of θ_1 is statistically significant at the 0.05 level (see Table 2.11).

$$\Delta Y_t = Z_t - 0.623 Z_{t-1},\ t \in \mathrm{Z} \tag{2.6}$$

where $\{Z_t\}$ is a sequence of normally distributed random variables, and $\Delta^1 Y_t = (1-B)^1 Y_t = Y_t - Y_{t-1}$ show the first-order difference of $\{Y_t\}$ where B and Δ denote the backshift and difference operators, respectively.

The results of Ljung-Box autocorrelation test for this data set is given in Table 2.12, and it shows that the residuals do not have statistically significant autocorrelation. Moreover, descriptive statistics and normality test results of the residuals given in Table 2.8 show that the residuals follow the normal distribution. In Figure 2.3, we plot the estimated parameters and MSE values of the block bootstrap estimators. As in the first example, SONBB provides better results among the others even in small block lengths which is the point of our study. The other methods tend to have the same trend for both estimation of $\hat{\theta}_1$ and its MSE. The results of all the bootstrap methods become more consistent with increasing block length.

Table 2.11 IMA(1,1) model estimates for number of earthquakes data

Parameter	Standard error (s.e.)	t-stat.	p-value
θ_1	0.088	7.046	0.000

Table 2.12 Ljung-Box Q-Statistics for the residuals of the estimated IMA(1,1) model for number of earthquakes data

Considered lag	Ljung-Box test	Degrees of freedom	Significance level
12	10.404	11	0.494
24	25.024	23	0.349
36	41.329	35	0.214

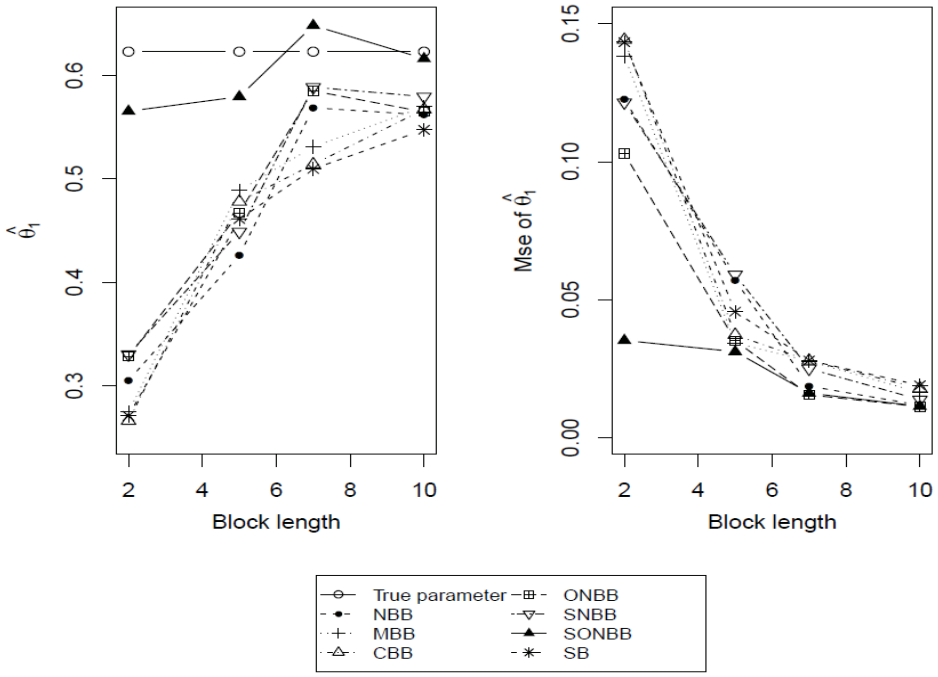

Figure 2.3 Plots of block bootstrap methods for $\hat{\theta}_1$ and its MSE for number of earthquakes data

2.3.2.4 Chemical concentration

Chemical concentration data is chosen because of being an example from ARMA(1,1) model. The original data is given as Series A in Box-Jenkins (1976). The analysis is implemented for between 14th - 193th observations. The estimated ARMA(1,1) model based on Box-Jenkins methodology is represented as in Equation (2.7). The results of residual analysis given in Table 2.14 and Table 2.8 indicate that the residuals of the fitted model are normally distributed and not serially correlated. Therefore, the *p*-value < 0.05 of *t*-tests presented in Table 2.13 shows that ARMA(1,1) model appears adequate at the significance level under 0.05.

$$Y_t - 0.927Y_{t-1} = 1.247 + Z_t - 0.612Z_{t-1},\ t \in Z \tag{2.7}$$

where $\{Z_t\}$ is a sequence of normally distributed random variables.

Figures 2.4 and 2.5, respectively, show the results of the block bootstrap estimators of ϕ_1 and θ_1. It is clear that the ONBB and SONBB produce more preferable results for both parameter estimations and their MSEs while the other methods behave in the same manner and their estimated values get close to the

results of the ONBB and SONBB as ℓ increases. The point of interest in this example is that the results of the ONBB and SONBB grow away from the true parameter value as block length increases. The interpretation of this could be that the ONBB and SONBB may be preferred for small block lengths unlike the other methods.

Table 2.13 ARMA(1,1) model estimates for chemical concentration data

Parameter	Standard error (s.e.)	t-stat.	p-value
δ	0.048	19.154	0.000
ϕ_1	0.008	143.681	0.000
θ_1	0.113	5.406	0.000

Table 2.14 Ljung-Box Q-Statistics for the residuals of the estimated ARMA(1,1) model for chemical concentration data

Considered lag	Ljung-Box test	Degrees of freedom	Significance level
12	13.951	9	0.124
24	23.397	21	0.323
36	45.228	33	0.076

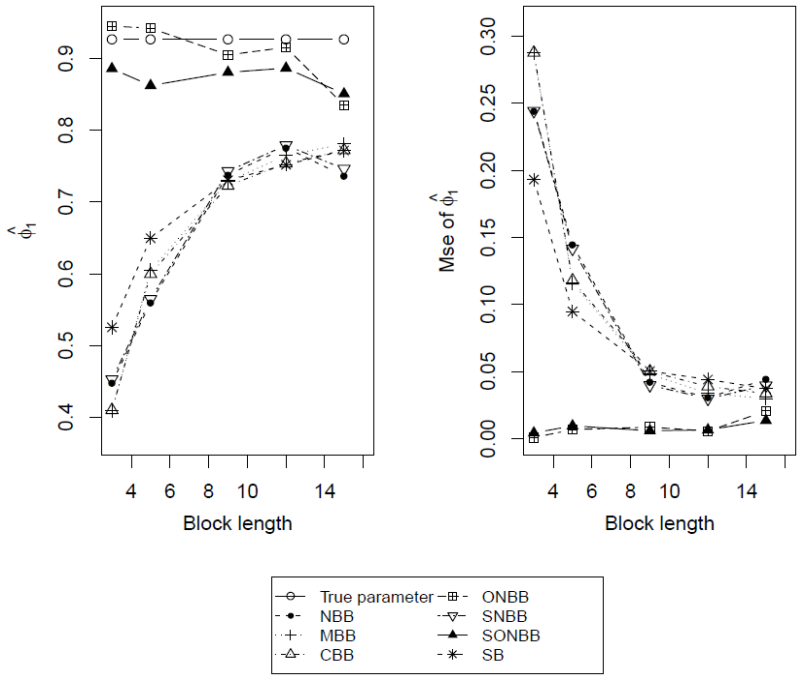

Figure 2.4 Plots of block bootstrap methods for $\hat{\phi}_1$ and its MSE for chemical concentration data

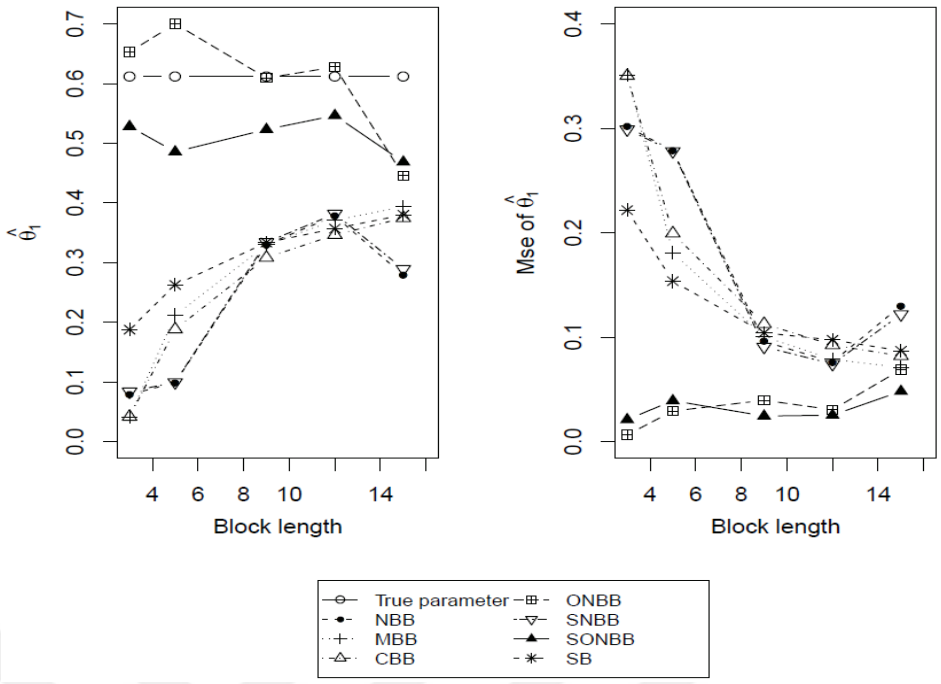

Figure 2.5 Plots of block bootstrap methods for $\hat{\theta}_1$ and its MSE for chemical concentration data

2.3.2.5 Unemployment rate

The seasonally adjusted registered unemployment rate of United Kingdom is discussed as an example of labour market statistics. The quarterly data (01/01/1955 – 01/01/2015 time period) were obtained from https://research.stlouisfed.org/fred2/series.rate. The autoregressive integrated model (ARI(1,1)) is fitted as in Equation (2.8), and Table 2.15 shows that the estimated coefficient is statistically significant at %5 level of significance. Also, Ljung-Box Q-statistics reported in Table 2.16 indicate that there is no evidence of autocorrelation in the residuals.

$$\Delta Y_t = 0.842\Delta Y_{t-1} + Z_t, \quad t \in Z \tag{2.8}$$

where $\{Z_t\}$ is a sequence of random variables which has normal distribution, and $\Delta^1 Y_t = Y_t - Y_{t-1}$ is the first difference of $\{Y_t\}$ process.

Table 2.8 indicates that the residuals are not normally distributed. All of empirical analysis lead us to consider an ARI(1,1) model as a suitable choice in the context of linear time series models although the residuals do not follow normal distribution. In this example, the results of the SONBB are better than the others especially in small ℓ values as shown in Figure 2.6. But the difference between the performances is not statistically significant for large block lengths.

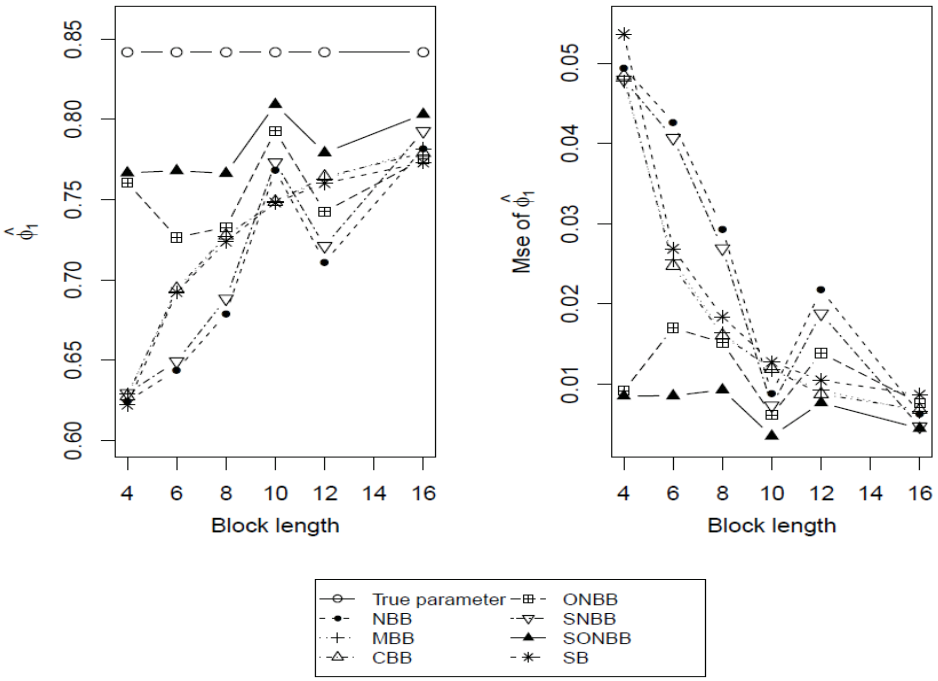

Figure 2.6 Plots of block bootstrap methods for $\hat{\phi}_1$ and its MSE for unemployment rate data

Table 2.15 ARI(1,1) model estimates for unemployment rate data

Parameter	Standard error (s.e.)	t-stat.	p-value
ϕ_1	0.034	24.548	0.000

Table 2.16 Ljung-Box Q-Statistics for the residuals of the estimated ARI(1,1) model for unemployment rate data

Considered lag	Ljung-Box test	Degrees of freedom	Significance level
12	18.684	11	0.067
24	34.537	23	0.058
36	47.971	35	0.071

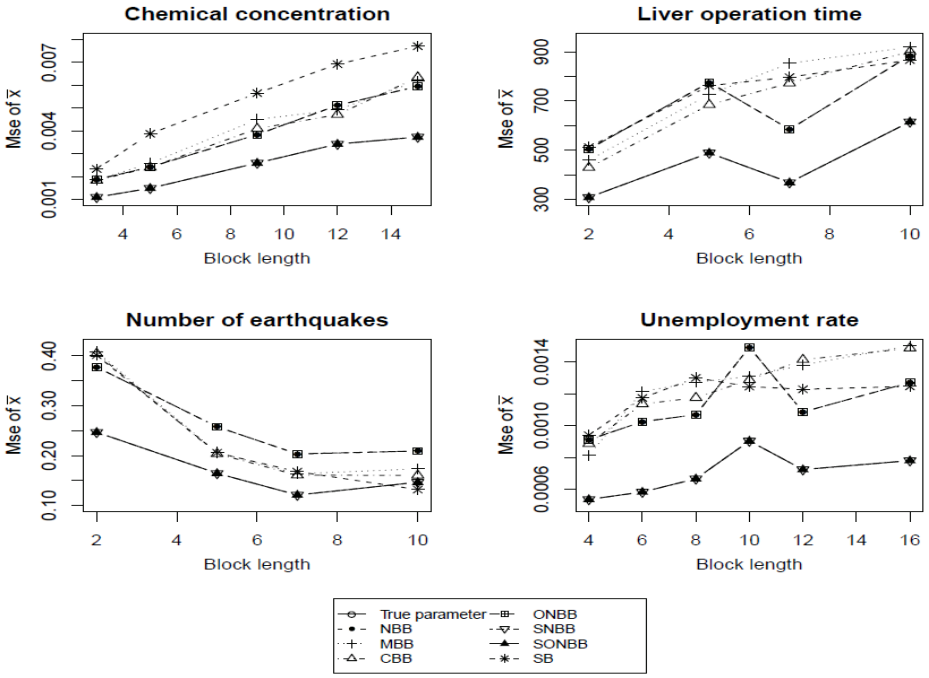

Figure 2.7 Plots of block bootstrap methods for MSE of the sample means

As it is seen from the Figure 2.7, sufficient versions of the NBB method have the smallest MSE values for the sample mean compared with the conventional block bootstrap methods. The reason of this is based on using distinct observations in the resamples. For more information see Pathak (1962).

CHAPTER THREE
NEW AND FAST BLOCK BOOTSTRAP BASED PREDICTION INTERVALS FOR GARCH(1,1) PROCESS WITH APPLICATION TO EXCHANGE RATES

3.1 Introduction

In this chapter, we propose a new bootstrap algorithm to obtain prediction intervals for GARCH(1,1) process which can be applied to construct prediction intervals for future volatilities. Measuring volatility and construction of valid predictions for future returns and volatilities have an important role in assessing risk and uncertainty in the financial market. To this end, the generalized autoregressive conditionally heteroscedastic (GARCH) model proposed by Bollerslev (1986) is one of the most commonly used techniques for modeling volatility and obtaining dynamic prediction intervals for returns as well as volatilities. Andersen & Bollerslev (1998), Andersen, Bollerslev, Diebold, & Labys (2001), Baillie & Bollerslev (1992), and Engle & Patton (2001) provide an excellent overview of research on prediction intervals for future returns in financial time series analysis. However, those works only consider point forecasts of volatility even though prediction intervals provide better inferences taking into account uncertainty of unobservable sequence of volatilities. Technically, construction of such prediction intervals requires some distributional assumptions which are generally unknown in practice. Moreover, the constructed prediction intervals along with the estimated parameter values can be affected due to any departure from the assumptions and may lead us to unreliable results. One of the remedy to construct prediction intervals without considering distributional assumptions is to use the well known resampling methods, e.g., the bootstrap.

It is well-known that the original nonparametric bootstrap proposed by Efron (1979) fails to provide satisfactory answers to general statistical inference problems for dependent data, since the assumption of independently and identically distributed (i.i.d.) data is violated (See, Lahiri (2003); Hall (1992) for more details). As a way to

deal with dependent data several types of resampling techniques were proposed. Among those, one of the most general tools to approximate the properties of estimators for serially correlated data is the method of block bootstrap. The main idea behind this method is to construct a resample of the data of size *n* by dividing the data into several blocks and choosing among them till the bootstrap sample is obtained. The commonly used procedures to implement block bootstrap called "non-overlapping" and "overlapping" blocking are proposed by Hall (1985) in the context of spatial data. In the univariate time series context, the non-overlapping block bootstrap (NBB) approach is proposed by Carlstein (1986), and overlapping blocks known as moving block bootstrap (MBB) is proposed by Kunsch (1989). In addition to these methods, a circular block bootstrap (CBB) method is suggested by Politis & Romano (1992a), where the data is wrapped around a circle so that each observation in the original data set has an equal probability to appear in a bootstrap sample. Also, the stationary bootstrap (SB) method which deals with random block lengths which have a geometric distribution is proposed by Politis & Romano (1994).

In all of the above blocking techniques the idea is to specify a sufficiently large block length ℓ so that the data points which are ℓ units apart are practically independent. Then, the dependence structure of the original data is attempted to be captured by these ℓ consecutive observations in each block drawn independently. However while doing so, the correlation structure is broken while moving from one block to another block. Conceptually, obtaining better estimates could be achieved by creating resamples 'similar' to the original data, as these could help us in preserving a dependency structure close to the original leading us in obtaining more precise estimates of the actual parameters. Ordered non-overlapping block bootstrap (ONBB), proposed by Beyaztas, Firuzan, & Beyaztas (2017), improve the performance of the block bootstrap technique by taking into account the correlations between the blocks. The authors empirically proved that the ONBB method often exhibits improved performance over the conventional block bootstrap methods in terms of parameter and coverage probability estimations for univariate linear time series models. They performed a simulation study based on autoregressive (AR) of order 2 and moving average (MA) of order 2 models with different sample sizes and

block lengths. Their results show that the ONBB method produces close estimations to the true values of the statistics especially to the second parameter with the increasing ℓ, and this result yields a confidence interval having better coverage probabilities. On the other hand, they failed to provide any information about the correlation structure between the blocks. In this study, (*i*) we show that the Spearman rank correlation between the ONBB resample and original data is always positive (and ≥ 0.5) and stronger than those for conventional block bootstrap methods which gives a justification for the superiority of the ONBB method. The similarities of the resamples obtained by the block bootstrap to the original data is shown using the dynamic time warping measure. (*ii*) Also, we extend the ONBB method to GARCH(1,1) process to obtain prediction intervals for future returns and volatilities. In summary, our extension works as follows: First, we use the squares of the GARCH process, which have the autoregressive-moving average (ARMA) representation, to make the the parameter estimation process linear. The ordinary least squares estimators of the ARMA model are calculated by a high order autoregressive model of order *m*, and the residuals are computed. Then the ONBB method is applied to the data to obtain the bootstrap sample of the returns which are used to calculate the ONBB estimators of the ARMA coefficients and the bootstrap sample of the volatilities. Finally, the future values of the returns and volatilities of the GARCH process are obtained by means of bootstrap replicates and quantiles of the Monte Carlo estimates of the ONBB distribution.

The rest of this chapter is organized as follows. We describe the data in Chapter 3.2. In Chapter 3.3 we provide detailed information on the ONBB method and the correlation structures between the resampled and original blocks. In Chapter 3.4, we propose a new, computationally efficient bootstrap algorithm based on the ONBB method to obtain prediction intervals for future returns and volatilities of GARCH(1,1) process. An extensive Monte Carlo simulation is conducted to examine the finite sample performance of the proposed method and the results are presented in Chapter 3.5. Finally, the AUD/USD daily exchange rate data is analyzed using the new method and the results are presented in Chapter 3.6.

3.2 Data

The exchange rate is regarded as the value of a specific country's currency in terms of another currency and obtaining valid prediction intervals of exchange rates is often essential to evaluate foreign denominated cash flows related to international transactions. For instance, exchange brokers, central banks, international traders and investors require prediction intervals of future returns and volatilities for many reasons, e.g., option pricing, to determine the next target zone of the exchange rate of interest, and for international portfolio diversification. In recent years, Australian Dollar (AUD) has become one of the most traded currencies in the world and it has an important position in Asian import market. Also, AUD is an attractive currency for the investors and it is often used in carrying out trades with other currencies due to the strength of Australian economy. On the other hand, from global perspective, the bilateral exchange rate with U.S. dollar (USD) has a great effect on the trading volume of the AUD in the foreign exchange market because USD is the dominant currency against almost all currencies in the world. Therefore, multi-step ahead prediction intervals of levels and volatilities of the bilateral AUD/USD exchange rate are crucial for the international firms and investors who import and export with Australia and use AUD as an investment.

The AUD/USD daily exchange rate data were obtained starting from 29th July, 2011 and ending on 3rd November, 2015 (available at https://www.stlouisfed.org/). After excluding observations on weekends and inactive days, our final data consisted a total of 1070 observations. From the original data the daily logarithmic returns were obtained as $y_t = 100 * \log\left(P_t / P_{t-1}\right)$, where P_t was the closing price on *t*-th day. The time series plots of the exchange rates and returns are presented in Figure 3.1. We checked the stationary status of the return series by applying the Ljung-Box (LB) and Augmented Dickey-Fuller (ADF) t-statistic tests and small *p*-values (*p*-value = 0.017 for the LB test and 0.010 for the ADF test) suggest that the return series is a mean-zero stationary process. Table 3.1 reports the sample statistics of y_t series, and it shows that the estimated kurtosis is higher than 3 which indicates that the

distribution of the returns was leptokurtic. Next, we checked for the Gaussianity of the return series and the *p*-value < 0.001 of Jarque-Bera test indicated that y_t was not Gaussian. Further, we performed the LB test to test for auto-correlations in the absolute and squared returns and smaller *p*-values indicated that the absolute and squared returns are highly auto-correlated. The auto-correlations of returns, absolute and squared returns are presented in Table 3.2. All of our preliminary exploratory analyses suggested the presence of conditional heteroscedasticity in the series. To find the optimal lag for the GARCH model to model the return series we defined many possible subsets of the GARCH(*p*,*q*) models with different *p* and *q* values. To choose the best model we used Akaike information (AIC) criterion (since it is proposed to determine the best model for forecasting) and the results show that GARCH(1,1) model is optimal according to AIC. Also, Hansen & Lunde (2005) compares a large number of volatility models to describe the conditional variance in an extensive empirical study based on exchange rate data. Their results show that the GARCH(1,1) model provides significantly better forecasts for exchange rates than the other models. In light of these results, we consider a GARCH(1,1) model as a suitable choice to model the return series.

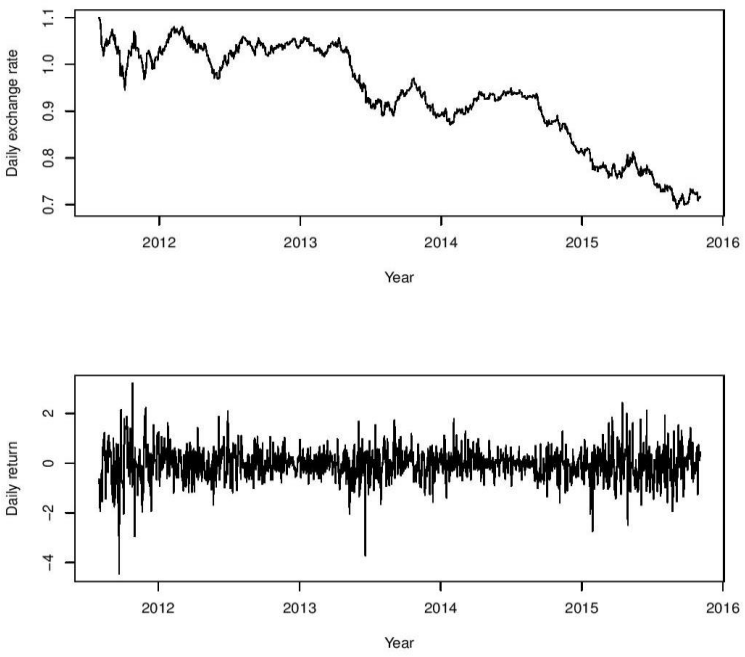

Figure 3.1 Time series plots of AUD/USD daily exchange rates and returns from 29th July, 2011 to 3rd November, 2015

Table 3.1 Sample statistics for y_t

T	Mean	Median	SD	Skewness	Kurtosis	Min.	Max.
1069	-0.04	-0.04	0.7	-0.27	6.15	-4.46	3.21

Table 3.2 Autocorrelations of y_t at lag *k*

Autocorrelations	r(1)	r(2)	r(3)	r(8)	r(9)	r(10)	r(19)	r(20)
y_t	-0.021	0.015	-0.027	0.042	0.030	-0.097	0.072	-0.019
$\lvert y_t \rvert$	0.058	0.101	0.114	0.120	0.167	0.064	0.134	0.089
y_t^2	0.016	0.040	0.166	0.102	0.146	0.046	0.096	0.061

3.3 Methodology

Let $\chi_n = \{X_1, \ldots, X_n\}$ be a sequence of stationary dependent random variables of size *n* having an unknown common distribution function *F*, whose parameter θ_0 is of our interest. We further assume that the distribution has a finite mean μ and a finite variance σ^2, both unknown. Let $\hat{\theta}_n$ be the estimator of θ_0 based on χ_n. Suppose $B_1, \ldots, B_b$ be the non-overlapping blocks where $B_i = \left(X_{(i-1)\ell+1}, \ldots, X_{i\ell}\right)$ for $i = 1, \ldots, b$. In conventional NBB, *b* blocks are drawn independently from $B_1, \ldots, B_b$ and pasted end-to-end to form a bootstrap sample. ONBB is proposed as ordering the bootstrapped blocks according to given labels to each original block for capturing more dependence structure compared to the conventional NBB method. In more detail, suppose the data is divided into the four independent blocks which are nonoverlapping. In this case, the labels are determined as $B_1 = 1$, $B_2 = 2$, $B_3 = 3$ and $B_4 = 4$, and let the bootstrapped blocks are $B_1^* = B_4$, $B_2^* = B_2$, $B_3^* = B_3$ and $B_4^* = B_3$. As a consequence, the new data is obtained using the NBB method as $\chi_{NBB}^* = \{B_4 \vdots B_2 \vdots B_3 \vdots B_3\}$ whereas it is obtained as $\chi_{ONBB}^* = \{B_2 \vdots B_3 \vdots B_3 \vdots B_4\}$ when ONBB is used. From this example it can be seen immediately that more 'representative' data sets can be formed by the ONBB method.

To provide a statistical explanation on the superiority of ONBB we use Spearman's rank correlation between the given labels of the original and bootstrapped blocks. Let j and $k_{(j)}$ be the given labels of the original and NBB blocks in the jth order, respectively, where j, $k_{(j)} = 1, \cdots, b$. Also let $R_{k_{(j)}}$ and m_k denote the rank of $k_{(j)}$ and frequency of the block k respectively, where $R_k = \sum_{i=1}^{k-1} m_{k-i} + (m_k + 1)/2$, $0 \le R_k \le b$, $\sum_{k=1}^{b} m_k = b$, and $m_k = 0,1,\ldots,b$. Then, the Spearman's rank correlation coefficient between the original and NBB blocks labels is obtained as $\rho_{Original,NBB} = 1 - \left(6\sum_{j=1}^{b} d_j^2\right) \Big/ \left(b^3 - b\right)$ where $\sum_{j=1}^{b} d_j^2 = \sum_{j=1}^{b} \left(j - R_{k_{(j)}}\right)^2 = \left(1 - R_{k_{(1)}}\right)^2 + \cdots + \left(b - R_{k_{(b)}}\right)^2$.

Based on the above notations, the following theorem provides an explanation for why ONBB based sample should be better representative of the original sample than that obtained using the NBB method.

Theorem 3.1: It can be shown that,

$$\rho_{Original,NBB} \le \rho_{Original,ONBB}.$$

***Proof*:** We examine *Theorem 3.1* under three cases given below, but it can be generalized to all other cases using a similar logic.

Case 1: Suppose that all the bootstrapped blocks are untied and in a dual order of j. In this case, for $j = 2,\ldots,b-1$, $k_{(j)} < k_{(j-1)}$ and $R_{k_{(j)}} < R_{k_{(j-1)}}$ for the NBB while it is $k_{(j)} > k_{(j-1)}$ and $R_{k_{(j)}} > R_{k_{(j-1)}}$ for the ONBB. Note that $m_k = 1$ for $k = 1,\ldots,b$. Thus,

$$\rho_{Original,NBB} = -1$$

since $\sum_{j=1}^{b} d_j^2 = \left(b^3 - b\right)/3$, and

$$\rho_{Original,ONBB} = 1$$

holds since $\sum_{j=1}^{b} d_j^2 = 0$ and $j = R_{k(j)}$ for all $j = 1,\ldots,b$ and $k_{(j)} = 1,\ldots,b$.

Case 2: Suppose that all the bootstrapped blocks are untied but in the same order of *j*. In this case, $k_{(j)} > k_{(j-1)}$, $R_{k_{(j)}} > R_{k_{(j-1)}}$ and $\sum_{j=1}^{b} d_j^2 = 0$ for both methods. So,

$$\rho_{Original,NBB} = \rho_{Original,ONBB} = 1$$

On the other hand, suppose we change the positions of two blocks, and let *t* and *z* represent the blocks whose positions are changed. Let $s_{t,z} = |z - t|$ denotes the distance between two positions where $t, z = 1,\ldots,b$. It is clear that $\sum_{j=1}^{b} d_j^2 = 2\sum_{t \neq z} s_{t,z}^2$. So,

$$\rho_{Original,ONBB} - \rho_{Original,NBB} = \left(12\sum_{t \neq z} s_{t,z}^2\right) \Big/ \left(b^3 - b\right) > 0$$

Case 3: Suppose that all the bootstrapped blocks are tied so that $k_{(j)} = k_{(j-1)}$ and $R_{k_{(j)}} = (b+1)/2$. Note that $m_k = b$. In this case,

$$\sum_{j=1}^{b} d_j^2 = (1-(b+1)/2)^2 + \ldots + (b-(b+1)/2)^2 = (b^3 - b)/12.$$

So,

$$\rho_{Original,NBB} = \rho_{Original,ONBB} = 0.5$$

Let us consider a more general case given below where only the positions of label groups 1 and 2 are different for NBB and ONBB, while the other block labels have the same positions and frequencies.

$$NBB\ block\ labels = 2,\ldots,2,1,\ldots,1,\ldots,3,\ldots,b$$
$$ONBB\ block\ labels = 1,\ldots 1,2,\ldots,2,\ldots,3,\ldots,b$$

In this case, $\rho_{Original,ONBB} - \rho_{Original,NBB}$ depends on $\sum_{j=1}^{b} d_j^2$. Let $\sum_{j=1}^{b} d_{j_{NBB}}^2$ and $\sum_{j=1}^{b} d_{j_{ONBB}}^2$ be the $\sum_{j=1}^{b} d_j^2$ values obtained by NBB and ONBB, respectively. Then we have,

$$\sum_{j=1}^{b} d_{j_{NBB}}^2 - \sum_{j=1}^{b} d_{j_{ONBB}}^2 = \left(\left(1 - R_{2_{(1)}}\right)^2 - \left(1 - R_{1_{(1)}}\right)^2\right) + \ldots$$
$$+ \left(\left(m_2 - R_{2_{(m_2)}}\right)^2 - \left(m_1 - R_{1_{(m_1)}}\right)^2\right) + \ldots$$
$$+ \left(\left((m_1 + m_2) - R_{1_{(m_1+m_2)}}\right)^2 - \left((m_1 + m_2) - R_{2_{(m_1+m_2)}}\right)^2\right) = 2m^3$$

Thus, $\rho_{Original,ONBB} - \rho_{Original,NBB} = \left(12m^3\right)/\left(b^3 - b\right) > 0$ when $m_1 = m_2 = m$. For the case where $m_1 \neq m_2$, we have,

$$\sum_{j=1}^{b} d_{j_{NBB}}^2 - \sum_{j=1}^{b} d_{j_{ONBB}}^2 = (m_1 m_2)(m_1 + m_2).$$

Therefore, $\rho_{Original,ONBB} - \rho_{Original,NBB} = \left(6(m_1 m_2)(m_1 + m_2)\right)/\left(b^3 - b\right) > 0$.

Theorem 3.1 follows directly combining the above three cases, and it shows that the bootstrap data obtained by ONBB is more representative of the original data than the one obtained by the NBB method. Thus, ONBB allows us to obtain better estimates and more reliable bootstrap quantities such as confidence interval of $\hat{\theta}_n$.

Corollary 3.1: It can further be shown that, the individual Spearman's correlation coefficients have the following ranges.

$$-1 \leq \rho_{Original,NBB} \leq 1, \text{ and, } 0.5 \leq \rho_{Original,ONBB} \leq 1$$

Proof: The proof follows as a direct consequence of the above derivation.

Remark 3.1 For the process $\{X_t\}_{t \in Z}$ which has a strong $\alpha(\cdot)$ mixing condition (see Bilingsley (1994)), Athreya & Lahiri (2006) show that under mild moment conditions, the NBB variance estimator of the statistic $T_n = \sqrt{n}\left(\overline{X}_n - \mu\right)$, $T_n^* = \sqrt{n}\left(\overline{X}_n^* - E^*\left(\overline{X}_n\right)\right)$, is $Var\left(T_n^*\right) \to \sigma_\infty^2$ as $n \to \infty$ where $Z_i = X_i - \mu$ and $\sigma_\infty^2 = \sum_{i=-\infty}^{\infty} EZ_1 Z_{i+1}$. Also, *Theorem 17.4.3* in Athreya & Lahiri (2006) show that $\sup_{x \in R}\left|P^*\left(T_n^* \leq x\right) - P\left(T_n \leq x\right)\right| \to 0$ as $n \to \infty$. This means the sampling distribution G_n and its NBB estimator $\hat{G}_n$ are consistent when ℓ goes to infinity at a slower rate compared to sample size $n\left(\ell^{-1} + n^{-1}\ell = o(1) \text{ as } n \to \infty\right)$. It is clear to say that the

ONBB method provides consistent estimators since sorting the bootstrapped blocks does not change the estimated values such as μ, σ^2, median etc. As mentioned in Chapter3.1, the ONBB method improves the performance of the block bootstrap technique by taking into account the correlations between the blocks, which leads us to have closer results to the true values of the statistics of interest. Also, the better estimates obtained by changing the correlation structure of the resampled time series affects the size of the prediction interval and provides more reliable results (please see the numerical results given in Chapter3.5). It should be noted that the superiority of the ONBB is not a consequence of either Edgeworth or Cornish-Fisher expansion. The main reason is based on having more representative bootstrap data sets by this method as shown in *Theorem 3.1*.

Remark 3.2 The construction of the ONBB bears the question of whether the new bootstrap method produces bootstrap samples that generate enough "new" information on the time series or if the suggested ordering might limit bootstrap samples look too "similar". As it is known each block has an equal probability $1/b$ to appear in a resample so that there are b^b number of distinct NBB samples. On the other hand, for the ONBB method, there are $\binom{2b-1}{b}$ number of distinct bootstrap samples since the order of the bootstrapped blocks is not important (it automatically puts in order the bootstrapped blocks). Let #NBB and #ONBB denote the number of distinct resamples generated by the NBB and ONBB methods, respectively. To answer this question we carry out a simulation with B = 10000 bootstrap resamples, and calculate the proportion of ($\#ONBB/\#NBB$) along with the number of blocks b. The results are shown in Figure 3.2. Note that by Figure 3.2 we can say that the ONBB produces bootstrap samples that generate enough new information on the time series as $n \to \infty$ and $\ell^{-1} + n^{-1}\ell = o(1)$.

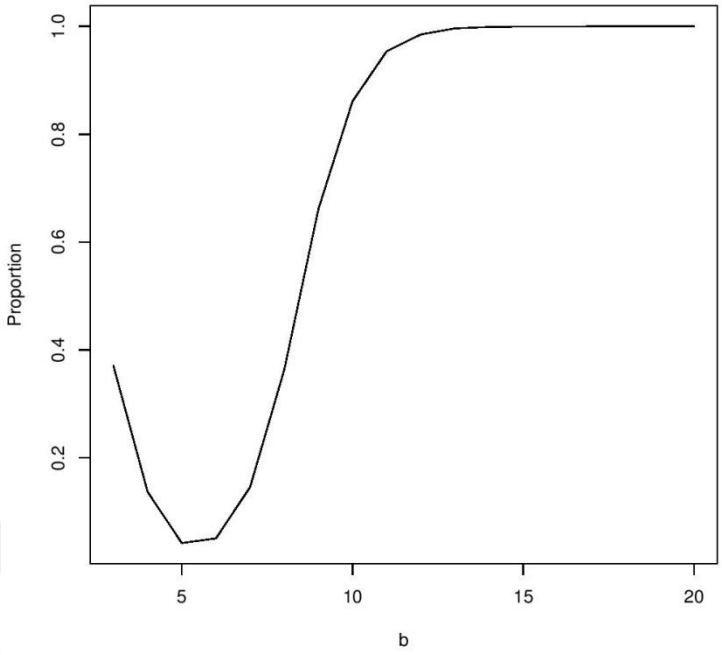

Figure 3.2 Proportion of the number of distinct resamples generated by the NBB and ONBB methods

To show the superiority of the ONBB we use the dynamic time warping (DTW) dissimilarity measure. DTW considers time axis offsets of two time series, say $X=\{x_1,\ldots,x_n\}$ and $Y=\{y_1,\ldots,y_m\}$. The computed value is based on a distance matrix $D_{n\times m}$ whose elements $d_{i,j}$ represents the distance $d(x_i,y_j)=(x_i-y_j)^2$, that is the (i,j) element measures the strength of alignment of points x_i and y_j. This measure is used to find the best synchronization of these two time series by minimizing warping cost $DTW(X,Y)=\min\left(\sqrt{\sum_{k=1}^{K}\omega_k}\right)$, $\max(m,n)\le K\le m+n-1$, where ω_k is the element of warping path *W*, a function of the distance matrix *D*. That is, this measure can be used to find out which block bootstrap method has the best matched resample with the original time series. A detailed discussion on the DTW measure can be found in Giorgino (2009) and Ratanamahatana & Keogh (2004). DTW measure has also been implemented in TSclust and dtw packages in R software.

Generally, DTW is an algorithm for comparing and aligning two sequences of data. The aim of the algorithm is to find an optimal match between two sequences by warping the time axis. To compare the block bootstrap methods in terms of their representativeness to the original dataset, we performed thorough simulation studies for AR(1) and ARMA(1,1) models with various parameters and sample sizes with block length $\ell = n^{1/3}$. Since the results and our conclusions do not vary significantly with different choices of parameter values, therefore to save space we present only the results obtained for the choices of auto-regressive parameter $\alpha = 0.2$ and, moving average parameter $\beta = 0.4$. The number of bootstrap replications *B* and Monte Carlo simulations *MC* are set at *B* = *MC* = 1000. For each simulation, we record DTW distances for all block bootstrap methods. The results are presented in Figure 3.3. Since the calculated values are distances, the method which has the smallest DTW values can be considered as the best method compared to others. As it is shown in Figure 3.3, the DTW distance values obtained by the NBB, MBB, CBB and SB are very close to each other and their lines are overlapping. On the other hand, it is clear that the ONBB has considerable small values compared to other block bootstrap methods. Also, in Figure 3.4, we plot the DTW densities of the block bootstrap methods for a simulated AR(1) sequence with the sample size *n* = 64. In this figure, *Reference index* and *Query index* stand for the time index of original time series and resampled series, respectively obtained by the block bootstrap method. The blue trace represents the warping path of the corresponding block bootstrap method. In this figure, the best alignment between two sequences is equivalent as finding the shortest path to go from the bottom-left to the top-right of the plot. Thus, the block bootstrap method which has a working path close to the diagonal produces more representative resamples of the original time series data. Clearly, considering the DTW analysis results in Figure 3.3 and 3.4, the most representative resamples are produced by the ONBB method compared to other block bootstrap methods, and these results further support *Theorem 3.1*.

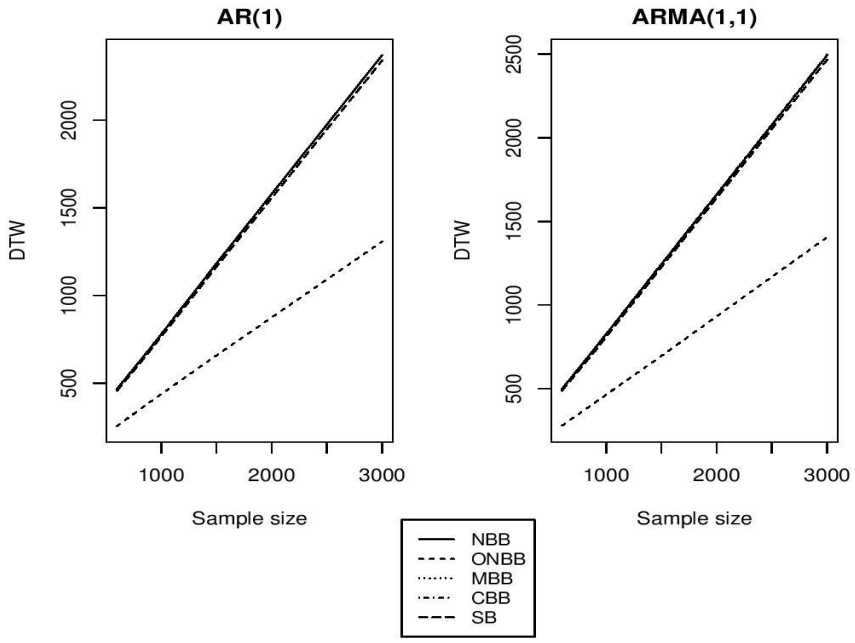

Figure 3.3 DTW distance values of block bootstrap methods

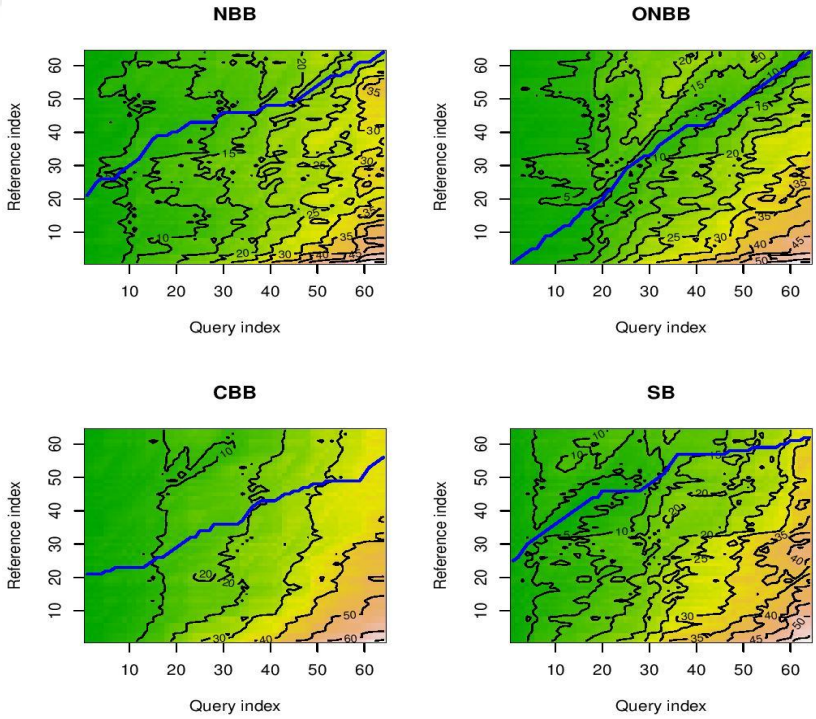

Figure 3.4 DTW density plots of block bootstrap methods

3.4 ONBB Prediction Intervals for GARCH(1,1) Model

As noted earlier, construction of prediction intervals for future returns and volatilities is an important problem in financial markets. However, the estimation of parameters and the construction of prediction intervals may be affected by a great amount due to any departure from these assumptions and may lead us to unreliable results. Resampling based prediction interval is one of the possible solutions to overcome this vexing issue since it does not require the full knowledge of the underlying data and distributional assumptions. In this context, bootstrap-based prediction intervals of autoregressive conditionally heteroscedastic (ARCH) model for future returns and volatilities by resampling residuals are proposed by Miguel & Olave (1999) and Reeves (2005). Pascual, Romo, & Ruiz (2006) further extends the previous works to construct bootstrap-based prediction intervals for returns and volatilities for GARCH(1,1) models. Later, Chen, Gel, Balakrishna, & Abraham (2011) suggests a computationally efficient bootstrap prediction intervals for ARCH and GARCH processes in the context for financial time series. Also, Hwang & Shin (2013) develops a stationary bootstrap prediction interval for GARCH models, and provides a mathematical justification for this method. Generally, block bootstrap is not suitable for construction of prediction intervals in conditionally heteroscedastic time series models because of its poor finite sample performance. However, our ONBB method overcomes this shortcoming and can be used to obtain reliable prediction intervals for future returns and volatilities.

To start with, we use ARMA representation of a GARCH(1,1) model and its least squares (LS) estimators in order to employ ONBB method for constructing prediction intervals. The GARCH(1,1) process has the following representation.

$$y_t = \sigma_t \varepsilon_t \tag{3.1}$$

$$\sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2, \quad t = 1, \cdots, T \tag{3.2}$$

where $\{\varepsilon_t\}$ is a sequence of i.i.d. random variables with zero mean, unit variance and $E(\varepsilon^4) < \infty$, ω, α, and β are unknown parameters satisfying $\omega \geq 0$, $\alpha \geq 0$ and $\beta \geq 0$.The stochastic process σ_t is assumed to be independent of ε_t. Throughout

this chapter, we assume that the process $\{y_t\}$ is strictly stationary, i.e., $E[\log(\beta + \alpha\varepsilon_t^2)] < 1$ and the strict stationary conditions of y_t as in Nelson (1990) hold. A GARCH (1,1) process $\{y_t\}$ is represented in the form of ARMA(1,1) as follows.

$$y_t^2 = \omega + (\alpha + \beta) y_{t-1}^2 + \nu_t - \beta \nu_{t-1} \tag{3.3}$$

where the innovation $\nu_t = y_t^2 - \sigma_t^2$ is a white noise (not i.i.d. in general) and identically distributed under the strict stationary assumption of y_t. According to Hannan & Rissanen (1982), the LS estimators for an ARMA(1,1) model are obtained as follows: (a) Fit a high order autoregressive model of order *m*, AR(*m*), with *m*> 1, to the data by Yule-Walker method to obtain $\hat{\nu}_t$. (b) A linear regression of y_t^2 onto y_{t-1}^2 is fitted to estimate the parameter vector $\phi = (\omega, (\alpha + \beta), -\beta)'$. In more detail, let $y_t^2 - \alpha y_{t-1}^2 = \nu_t + \beta \nu_{t-1}$ be the ARMA(1,1) representation of the underlying GARCH(1,1) process, where $\{\nu_t\} \sim wn(0, \sigma^2)$. Then, in step (a), an AR(*m*) model is fitted to the data to obtain $\hat{\nu}_t$ such that $\hat{\nu}_t = y_t^2 - \hat{\alpha}_{m1} y_{t-1}^2 - \ldots - \hat{\alpha}_{mm} y_{t-m}^2$ for $t = m+1, \ldots, n$. In step (b), a linear regression $y_t^2 = \omega + (\alpha + \beta) y_{t-1}^2 - \beta \hat{\nu}_{t-1} + (\nu_t - \beta(\nu_{t-1} - \hat{\nu}_{t-1}))$ is fitted to obtain the LS estimator of ϕ, where the term given in bracket, $\xi = (\nu_t - \beta(\nu_{t-1} - \hat{\nu}_{t-1}))$ is the error term. In matrix notations, let $\mathbf{Z}_n$ and $\mathbf{X}$ are as follows.

$$\mathbf{Z}_n = \begin{bmatrix} y_{m+1}^2 \\ \vdots \\ y_n^2 \end{bmatrix} \text{ and } \mathbf{X} = \begin{bmatrix} 1 & y_m^2 & \hat{\nu}_m \\ \vdots & \vdots & \vdots \\ 1 & y_{n-1}^2 & \hat{\nu}_{n-1} \end{bmatrix}$$

Then, the LS estimator $\hat{\phi} = (\hat{\omega}, \widehat{(\alpha + \beta)}, -\hat{\beta})'$ is obtained as

$$\hat{\phi} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z_n} \tag{3.4}$$

Given $\mathbf{X}'\mathbf{X}$ is non-singular. The corresponding $\hat{\alpha}$ is calculated as $\hat{\alpha} = \widehat{(\alpha + \beta)} - \hat{\beta}$.

Based on the above, the complete algorithm of the ONBB prediction intervals for future returns and volatilities is as follows.

Step 1: For a realization of GARCH(1,1) process $\{y_0, y_1, \cdots, y_T\}$, calculate the LS estimates of ARMA coefficients as in Equation (3.4).

Step 2: For $t = 1, \cdots, T$, calculate the residuals $\hat{\varepsilon}_t = y_t / \hat{\sigma}_t$ where $\hat{\sigma}_t^2 = \hat{\omega} + \hat{\alpha} y_{t-1}^2 + \hat{\beta}\hat{\sigma}_{t-1}^2$ and $\hat{\sigma}_0^2 = \hat{\omega}/\left(1 - \left(\hat{\alpha} + \hat{\beta}\right)\right)$. Let $\hat{F}_\varepsilon$ be the empirical distribution function of the centered and rescaled residuals.

Step 3: Obtain ONBB observations from the ARMA representation of GARCH process.

Step 4: Compute ONBB estimators of ARMA coefficients as

$$\hat{\phi}^*_{\square} = \left(\mathbf{X}^{*'}\mathbf{X}\right)^{-1}\mathbf{X}^{*'}\mathbf{Z}^*_{\mathbf{n}} = \left(\hat{\omega}^*, \widehat{(\alpha + \beta)}^*, -\hat{\beta}^*\right)'$$

and calculate the corresponding $\hat{\alpha}^*$ as $\hat{\alpha}^*_{\square} = \widehat{(\alpha + \beta)}^* - \hat{\beta}^*$.

Step 5: Obtain ONBB volatilities as $\hat{\sigma}_t^{2*} = \hat{\omega}^* + \hat{\alpha}^* y_{t-1}^{2*} + \hat{\beta}^* \hat{\sigma}_{t-1}^{2*}$ with $\hat{\sigma}_0^{2*} = \hat{\omega}/\left(1 - \left(\hat{\alpha} + \hat{\beta}\right)\right)$.

Step 6: Calculate $h = 1, 2, \ldots$ steps ahead ONBB future returns and volatilities with the following recursion:

$$\hat{\sigma}_{T+h}^{2*} = \hat{\omega}^* + \hat{\alpha}^* y_{T+h-1}^{2*} + \hat{\beta}^* \hat{\sigma}_{T+h-1}^{2*}$$
$$y_{T+h}^* = \hat{\sigma}_{T+h}^{2*} \hat{\varepsilon}_{T+h}^*$$

where $y_{T+h}^* = y_{T+h}$ for $h \le 0$ and $\hat{\varepsilon}_{T+h}^*$ is randomly drawn from $\hat{F}_\varepsilon$.

Step 7: Repeat steps 3-6 *B* times to obtain bootstrap replicates of returns and volatilities $\left\{y_{T+h}^{*,1}, \ldots, y_{T+h}^{*,B}\right\}$ and $\left\{\hat{\sigma}_{T+h}^{2*,1}, \ldots, \hat{\sigma}_{T+h}^{2*,B}\right\}$ for each *h*.

As noted in Pascual et al. (2006), the one-step conditional variance is perfectly predictable if the model parameters are known, and the only uncertainty which is caused by the parameter estimation, is associated with the prediction of σ_{t+1}^2. On the other hand, there are further uncertainties about future errors when predicting two or more step ahead variances. Thus, it is more interesting issue to have prediction intervals for future volatilities. Now, let $G_y^*(h) = P\left(y_{T+h}^* \le h\right)$ and

$G^*_{\sigma^2}(h) = P\left(\hat{\sigma}^{*2}_{T+h} \le h\right)$ be the ONBB distribution functions of unknown distribution functions of y_{T+h} and σ^2_{T+h}, respectively, for $h = 1,2,\ldots$. Also let $G^*_{y,B}(h) = \#\left(y^{*,b}_{T+h} \le h\right)/B$ and $G^*_{\sigma^2,B}(h) = \#\left(\hat{\sigma}^{*2}_{T+h} \le h\right)/B$, for $b = 1,\ldots,B$, be the corresponding Monte Carlo estimates. Then, a $100(1-\gamma)\%$ bootstrap prediction interval for y_{T+h} and σ^2_{T+h}, respectively, are given by

$$\left[L^*_{y,B}(y), U^*_{y,B}(y)\right] = \left[Q^*_{y,B}(\gamma/2), Q^*_{y,B}(1-\gamma/2)\right],$$

$$\left[L^*_{\sigma^2,B}(y), U^*_{\sigma^2,B}(y)\right] = \left[Q^*_{\sigma^2,B}(\gamma/2), Q^*_{\sigma^2,B}(1-\gamma/2)\right]$$

where $Q^*_{y,B} = G^{*-1}_y(h)$ and $Q^*_{\sigma^2,B} = G^{*-1}_{\sigma^2}(h)$.

We have the following proposition which shows the large sample validity of the ONBB prediction intervals.

Proposition**3.**1:

(i) $\sup_x \left| P^*\left(\sqrt{n}\left[\hat{\phi}^* - \hat{\phi}\right] \le x\right) - P\left(\sqrt{n}\left[\hat{\phi} - \phi\right] \le x\right)\right| \xrightarrow{p} 0$.

Hence $\hat{\phi}^* \xrightarrow{p^*} \phi$

(ii) $y^*_{T+h} \xrightarrow{d^*} y_{T+h}$ and $\hat{\sigma}^{*2}_{T+h} \xrightarrow{d^*} \sigma^2_{T+h}$ as $n \to \infty$.

(iii) $\lim_{n\to\infty} \lim_{B\to\infty} P\left[L^*_{y,B} \le Y_{T+h} \le U^*_{y,B}\right] = 1-\gamma$

$\lim_{n\to\infty} \lim_{B\to\infty} P\left[L^*_{\sigma^2,B} \le \sigma^2_{T+h} \le U^*_{\sigma^2,B}\right] = 1-\gamma$

where, for the random variables X_n and X, $X_n \xrightarrow{p} X$, $X_n \xrightarrow{p^*} X$ and $X_n \xrightarrow{d^*} X$ represent the convergence in probability, (conditional) convergence in probability and (conditional) convergence in distribution conditional on a given sample $y = \{y_0,\ldots,y_T\}$, respectively.

Proof: The LS estimator of an ARMA model $\hat{\phi}$ satisfies

$$\sqrt{n}\left[\hat{\phi} - \phi\right] = \left(\frac{\mathbf{X}'\mathbf{X}}{n}\right)^{-1} \frac{1}{\sqrt{n}} \mathbf{X}'\boldsymbol{\xi} \xrightarrow{d} N\left(\mathbf{0}, \mathbf{V}_\phi\right)$$

where the covariance matrix $\mathbf{V}_\phi$ is given by $\mathbf{V}_\varphi = \mathbf{D}^{-1} I \mathbf{D}^{-1}$ such that

$$\left(\frac{\mathbf{X}'\mathbf{X}}{n}\right)^{-1} \xrightarrow{p} \mathbf{D}^{-1}$$

$$\frac{1}{\sqrt{n}}\mathbf{X}'\boldsymbol{\xi} \xrightarrow{d} N(0,\Gamma)$$

where the non-diagonal matrix Γ is related to the covariance matrix of the moving-average model. The bootstrap estimate $\hat{\phi}^*$ can be written as

$$\sqrt{n}\left[\hat{\phi}^* - \hat{\phi}\right] = \left(\frac{\mathbf{X}^{*'}\mathbf{X}^*}{n}\right)^{-1} \frac{1}{\sqrt{n}}\mathbf{X}^{*'}\hat{\boldsymbol{\xi}}^*$$

To prove part (*i*) it is suffice to show that $\mathbf{X}^{*'}\mathbf{X}^*/n$ converges in probability to $\mathbf{X}'\mathbf{X}/n$, and $\mathbf{X}^{*'}\hat{\boldsymbol{\xi}}^*/\sqrt{n}$ convergence in distribution to $\mathbf{X}'\boldsymbol{\xi}/\sqrt{n}$. We write $\mathbf{X}'\mathbf{X} = \sum_{i=1}^{n}\mathbf{X}_i'\mathbf{X}_i$ and $\mathbf{X}^{*'}\mathbf{X}^* = \sum_{i=1}^{n}\mathbf{X}_i^{*'}\mathbf{X}_i^*$. For simplicity, we assume that $n = b\ell$. Let $\mathbf{U}_{k,\ell}$ be the mean of $\mathbf{X}_{nj}'\mathbf{X}_{nj}$ in block B_k such that $B_k = \{\mathbf{X}_{nj} : (k-1)\ell \le j \le k\ell\}$, for $k = 1,\ldots,b$. That is

$$\mathbf{U}_{k,\ell} = \sum_{j=(k-1)\ell+1}^{k\ell}\mathbf{X}_{nj}'\mathbf{X}_{nj}/\ell, \qquad k = 1,\ldots,b$$

Let

$$\mathbf{U}_{k,\ell}^* = \sum_{j=(k-1)\ell+1}^{k\ell}\mathbf{X}_{nj}^{*'}\mathbf{X}_{nj}^*/\ell, \qquad k = 1,\ldots,b$$

be the NBB version of $\mathbf{U}_{k,\ell}$. Then we have

$$\begin{aligned}&\left|\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i^{*'}\mathbf{X}_i^* - \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i'\mathbf{X}_i\right| \\ &\le \left|\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i^{*'}\mathbf{X}_i^* - E^*\left(\mathbf{U}_{k,\ell}^*\right)\right| + \left|E^*\left(\mathbf{U}_{k,\ell}^*\right) - \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i'\mathbf{X}_i\right|\end{aligned} \tag{3.5}$$

where E^* denotes the conditional expectation under the bootstrap distribution. The first term of the right hand side of Equation (3.5) tends to 0 by the weak law of large numbers. Since $P^*\left((X_1^*,\ldots,X_\ell^*) = (X_{(j-1)\ell+1},\ldots,X_{b\ell})\right) = 1/b$, for $j = 1,\ldots,b$, for the second term, we have

$$E^*\left(\mathbf{U}^*_{k,\ell}\right)=\frac{1}{b}\sum_{k=1}^{b}\left(\frac{1}{\ell}\sum_{i=1}^{\ell}\mathbf{X}'_{(k-1)\ell+i}\mathbf{X}_{(k-1)\ell+i}\right)=\frac{1}{b\ell}\sum_{i=1}^{n}\mathbf{X}'_i\mathbf{X}_i$$

Hence, the second term tends to 0 as $n\to\infty$, which proves that $\mathbf{X}^{*'}\mathbf{X}^*/n\xrightarrow{p^*}\mathbf{X}'\mathbf{X}/n$.

To show $\frac{1}{\sqrt{n}}\mathbf{X}^{*'}\hat{\boldsymbol{\xi}}^*\xrightarrow{d^*}N(0,\Gamma)$, write $\mathbf{X}'\boldsymbol{\xi}=\sum_{i=1}^{n}\mathbf{X}'_i\boldsymbol{\xi}_i$ and $\mathbf{X}^{*'}\hat{\boldsymbol{\xi}}^*=\sum_{i=1}^{n}\mathbf{X}^{*'}_i\hat{\boldsymbol{\xi}}^*_i$. Let $\mathbf{V}_{k,\ell}=\sum_{j=(k-1)\ell+1}^{k\ell}\mathbf{X}'_{nj}\hat{\boldsymbol{\xi}}_{nj}\Big/\ell$ and let $\mathbf{V}^*_{k,\ell}$ be the bootstrap version of $\mathbf{V}_{k,\ell}$. Then we have

$$\begin{aligned}&\frac{1}{\sqrt{n}}\mathbf{X}^{*'}\hat{\boldsymbol{\xi}}^*\\&=\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{X}^{*'}_i\hat{\boldsymbol{\xi}}^*_i-\sqrt{n}E^*\left(\mathbf{V}^*_{k,\ell}\right)\right)+\sqrt{n}E^*\left(\mathbf{V}^*_{k,\ell}\right)\\&=\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}^{*'}_i\hat{\boldsymbol{\xi}}^*_i-E^*\left(\mathbf{V}^*_{k,\ell}\right)\right)+\sqrt{n}E^*\left(\mathbf{V}^*_{k,\ell}\right)\end{aligned}\qquad(3.6)$$

For the first term of Equation (3.6) we have

$$\begin{aligned}&\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}^{*'}_i\hat{\boldsymbol{\xi}}^*_i-E^*\left(\mathbf{V}^*_{k,\ell}\right)\right)\\&\le\sqrt{n}\left|\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}^{*'}_i\hat{\boldsymbol{\xi}}^*_i-\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}'_i\hat{\boldsymbol{\xi}}_i\right|\\&\le\sqrt{n}\left[\left|\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}^{*'}_i\hat{\boldsymbol{\xi}}^*_i-E^*\left(\mathbf{V}^*_{k,\ell}\right)\right|+\left|E^*\left(\mathbf{V}^*_{k,\ell}\right)-\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}'_i\hat{\boldsymbol{\xi}}_i\right|\right]\xrightarrow{p^*}0\end{aligned}\qquad(3.7)$$

by weak law of large numbers and since $E^*\left(\mathbf{V}^*_{k,\ell}\right)-\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}'_i\hat{\boldsymbol{\xi}}_i=o_p(1)$. For the second term,

$$\sqrt{n}E^*\left(\mathbf{V}^*_{k,\ell}\right)=\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}'_i\hat{\boldsymbol{\xi}}_i\right)=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{X}'_i\hat{\boldsymbol{\xi}}_i\xrightarrow{d^*}N(0,\Gamma)$$

holds by the Slutsky's theorem. Hence, $\hat{\phi}^*\xrightarrow{p^*}\phi$. By using the proof of part (*i*), the proofs of (*ii*) and (*iii*) directly follow from the proof of *Theorem 3.1* of Thombs & Schucany (1990).

3.5 Numerical Results

To investigate the performance of our proposed ONBB prediction intervals we conduct a simulation study under GARCH (1,1) model given in Equation (3.8) below, and we compare our results with the method proposed by Pascual et al. (2006) (abbreviated as 'PRR') by means of coverage probabilities and length of prediction intervals. It is worth to mention that we also compared the performance of our proposed method with other existing block bootstrap methods mentioned in Chapter 3.1, e.g., NBB, MBB, CBB and SB. Our method performed considerably better compared to them, and therefore to save space we only report the comparative study with ONBB and PRR. Roughly, we observed the coverage probabilities of other block bootstrap methods range in between 90%-94% for future returns while those range only in between 25%-60% for future volatilities.

To discuss the numerical study we present here, let us start with the following GARCH(1,1) model.

$$\begin{aligned} y_t &= \sigma_t \varepsilon_t \\ \sigma_t^2 &= 0.05 + 0.1 y_{t-1}^2 + 0.85 \sigma_{t-1}^2 \end{aligned} \tag{3.8}$$

where, ε_t follows a $N(0,1)$ distribution. The significance level γ is set to 0.05 to obtain 95% prediction intervals for future returns and volatilities. Since the block bootstrap methods are sensitive to the choice of the block length ℓ, we choose three different block lengths in our simulation study: $n^{1/3}$, $n^{1/4}$, $n^{1/5}$ as proposed by Hall et al. (1995). Let $h = 1,2,\ldots,s$, $s \geq 1$, be defined as the lead time. We obtain the prediction intervals for next s = 20 observations. The experimental design is similar to those of Pascual et al. (2006) which is as follows:

Step 1: Simulate a GARCH(1,1) series with the parameters given in Equation (3.8) and generate R = 1000 future values y_{T+h} and σ_{T+h}^2 for $h = 1,2,\ldots,s$.

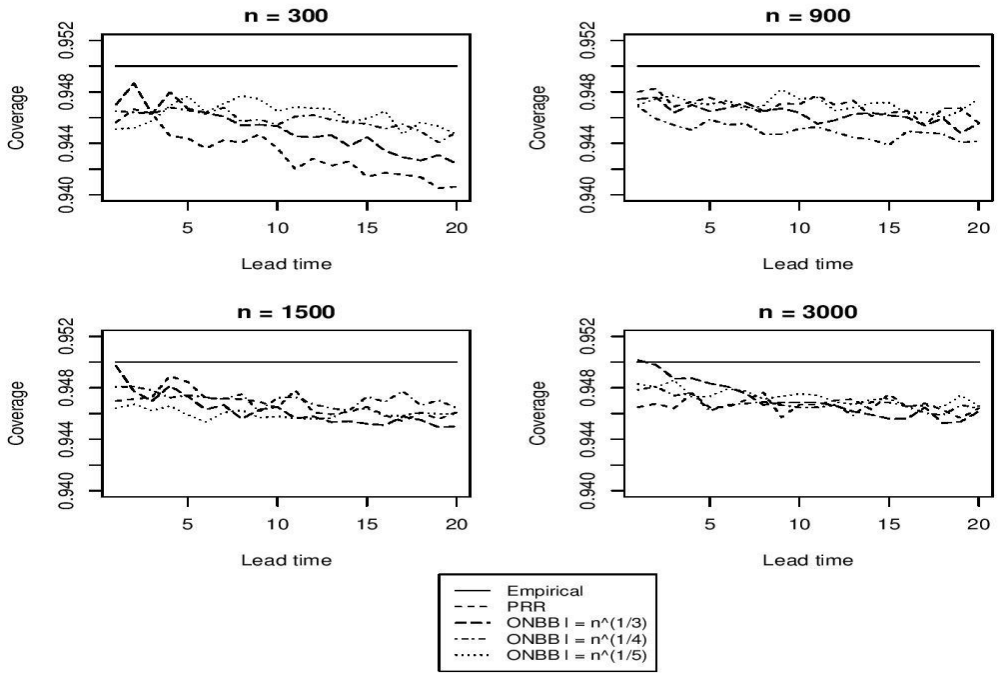

Figure 3.5 Estimated coverage probabilities of returns using PRR and ONBB

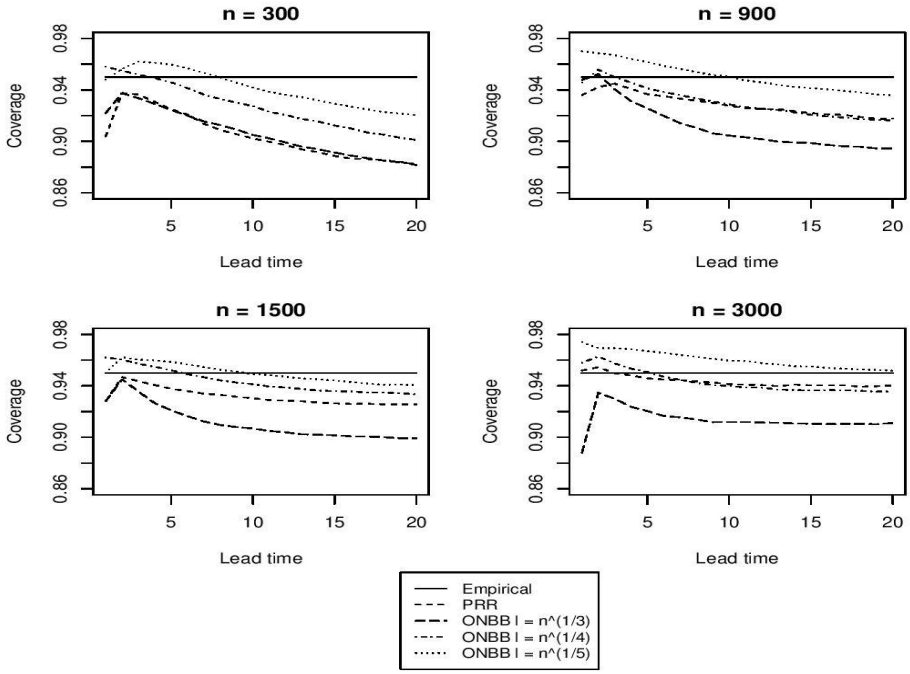

Figure 3.6 Estimated coverage probabilities of volatilities using PRR and ONBB

Step 2: Calculate bootstrap future values $y_{T+h}^{*,b}$ and $\sigma_{T+h}^{*2,b}$ for $h=1,2,\ldots,s$ and $b=1,2,\ldots,B$. Then estimate the coverage probabilities $\left(C^{*}\right)$ of bootstrap prediction intervals for y_{T+h} and σ_{T+h}^{2} as

$$C_{y_{T+h}}^{*,i}=\frac{1}{R}\sum_{r=1}^{R}\mathbf{1}\left\{Q_{y_{T+h}}^{*,i}\left(\gamma/2\right)\le y_{T+h}^{r}\le Q_{y_{T+h}}^{*,i}\left(1-\gamma/2\right)\right\}$$
$$C_{\sigma_{T+h}^{2}}^{*,i}=\frac{1}{R}\sum_{r=1}^{R}\mathbf{1}\left\{Q_{\sigma_{T+h}^{2}}^{*,i}\left(\gamma/2\right)\le \sigma_{T+h}^{2,r}\le Q_{\sigma_{T+h}^{2}}^{*,i}\left(1-\gamma/2\right)\right\}$$

where **1** represents the indicator function. The corresponding interval lengths $\left(L^{*}\right)$ are calculated by

$$L_{y_{T+h}}^{*,i}=Q_{y_{T+h}}^{*,i}\left(1-\gamma/2\right)-Q_{y_{T+h}}^{*,i}\left(\gamma/2\right)$$
$$L_{\sigma_{T+h}^{2}}^{*,i}=Q_{\sigma_{T+h}^{2}}^{*,i}\left(1-\gamma/2\right)-Q_{\sigma_{T+h}^{2}}^{*,i}\left(\gamma/2\right)$$

Step 3: Repeat Steps 1-2, *MC* = 1000 times to calculate the average values of $C_{y_{T+h}}^{*}$, $C_{\sigma_{T+h}^{2}}^{*}$, $L_{y_{T+h}}^{*}$ and $L_{\sigma_{T+h}^{2}}^{*}$ as

$$ave\left(C_{y_{T+h}}^{*}\right)=\sum_{i=1}^{MC}\frac{C_{y_{T+h}}^{*,i}}{MC}\quad ave\left(C_{\sigma_{T+h}^{2}}^{*}\right)=\sum_{i=1}^{MC}\frac{C_{\sigma_{T+h}^{2}}^{*,i}}{MC}$$
$$ave\left(L_{y_{T+h}}^{*}\right)=\sum_{i=1}^{MC}\frac{L_{y_{T+h}}^{*,i}}{MC}\quad ave\left(L_{\sigma_{T+h}^{2}}^{*}\right)=\sum_{i=1}^{MC}\frac{L_{\sigma_{T+h}^{2}}^{*,i}}{MC}$$

Also, calculate the standard errors of the estimated coverage probabilities and interval lengths by

$$s.e\left(C_{y_{T+h}}^{*}\right)=\left\{\sum_{i=1}^{MC}\left[C_{y_{T+h}}^{*,i}-ave\left(C_{y_{T+h}}^{*}\right)\right]^{2}/MC\right\}^{1/2}$$
$$s.e\left(C_{\sigma_{T+h}^{2}}^{*}\right)=\left\{\sum_{i=1}^{MC}\left[C_{\sigma_{T+h}^{2}}^{*,i}-ave\left(C_{\sigma_{T+h}^{2}}^{*}\right)\right]^{2}/MC\right\}^{1/2}$$
$$s.e\left(L_{y_{T+h}}^{*}\right)=\left\{\sum_{i=1}^{MC}\left[L_{y_{T+h}}^{*,i}-ave\left(L_{y_{T+h}}^{*}\right)\right]^{2}/MC\right\}^{1/2}$$
$$s.e\left(L_{\sigma_{T+h}^{2}}^{*}\right)=\left\{\sum_{i=1}^{MC}\left[L_{\sigma_{T+h}^{2}}^{*,i}-ave\left(L_{\sigma_{T+h}^{2}}^{*}\right)\right]^{2}/MC\right\}^{1/2}$$

A short summary of the simulation results is given in Table 3.3. More detailed results are presented in Figures 3.5 - 3.8. Our findings show that ONBB outperforms PRR in general. For the prediction intervals of future returns (see Figure 3.5) the

performances of both methods are almost same. Also, ONBB provides competitive interval lengths for returns (see Figure 3.7). The accuracy of the prediction intervals for volatilities obtained by ONBB is sensitive to the choice of block length parameter ℓ, and the higher coverage probabilities are obtained when $\ell = n^{1/4}$ and $n^{1/5}$ are used.

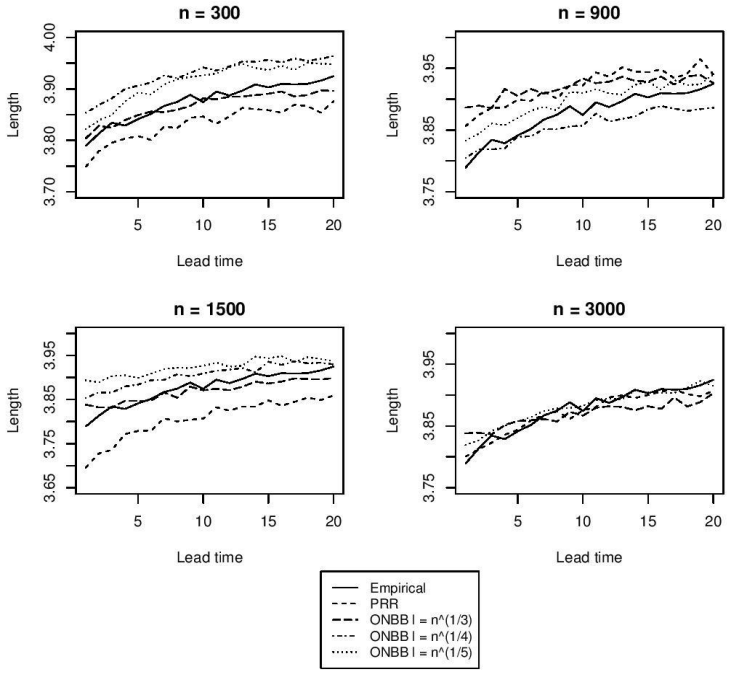

Figure 3.7 Estimated lengths of prediction intervals of returns using PRR and ONBB

The performance of our proposed method is always better than PRR in small sample sizes, and it outperforms PRR also in large samples especially for long-term forecasts as it is shown in Table 3.3 and Figure 3.6. Furthermore, in general, ONBB has less standard errors for coverage probabilities compared to PRR. Based on our findings, the proposed method achieves superior performance with $\ell = n^{1/5}$ for the prediction intervals of future returns. For the prediction intervals of volatilities, by taking into account the coverage probabilities and length of intervals, $\ell = n^{1/5}$ and $n^{1/4}$ seems to be the optimal choices for short-term and long-term forecasts, respectively. We also compare the ONBB and PRR in terms of their computing times and Figure 3.9 represents the approximate computing times for various sample sizes based on B = 1000 bootstrap replications and only one Monte Carlo simulation. As

presented in Figure 3.9, ONBB has considerably less computational time as PRR requires about 4-6 times more computing time than ONBB.

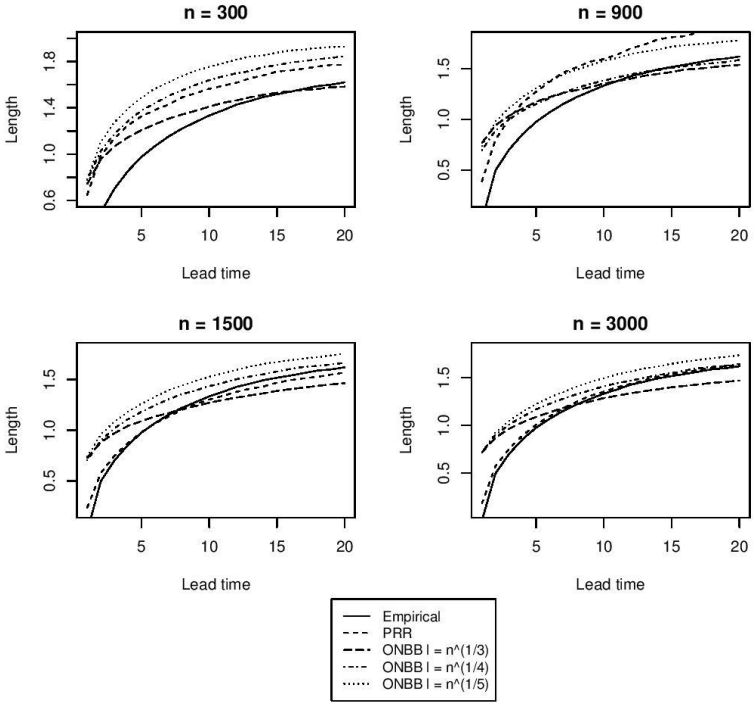

Figure 3.8 Estimated lengths of prediction intervals of volatilities using PRR and ONBB

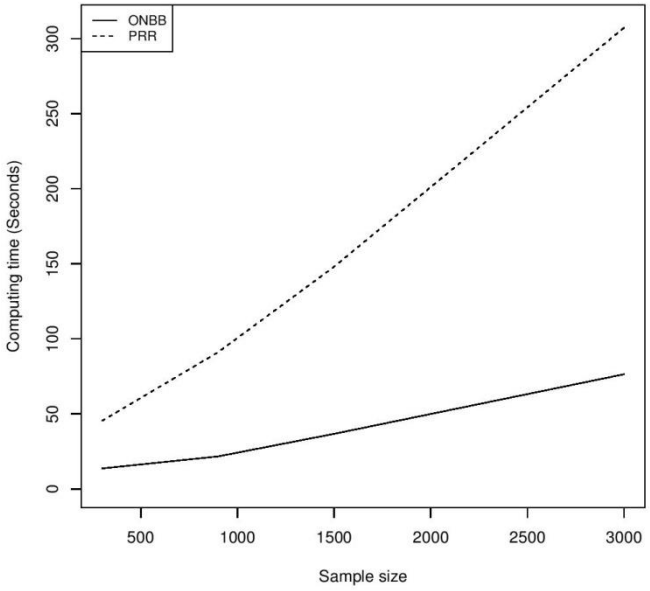

Figure 3.9 Estimated computing times for PRR and ONBB

Moreover, we also compared our method with the one proposed by Chen et al. (2011) [hereafter referred to as the CGBA method] through a simulation study. Like PRR, the CGBA method is also based on residuals. The main difference is that PRR uses quasi-maximum likelihood method to estimate the parameters, and then uses residual based resampling to construct intervals, whereas, the CGBA method utilizes the ARMA representation of a GARCH process to first estimate the parameters of the original data, and then uses the sieve bootstrap to obtain prediction intervals. The CGBA method requires approximately 2.5 more computing time than ours and the coverage performance of our proposed method is significantly better than CGBA method. To save space, we omit the numerical details in this chapter.

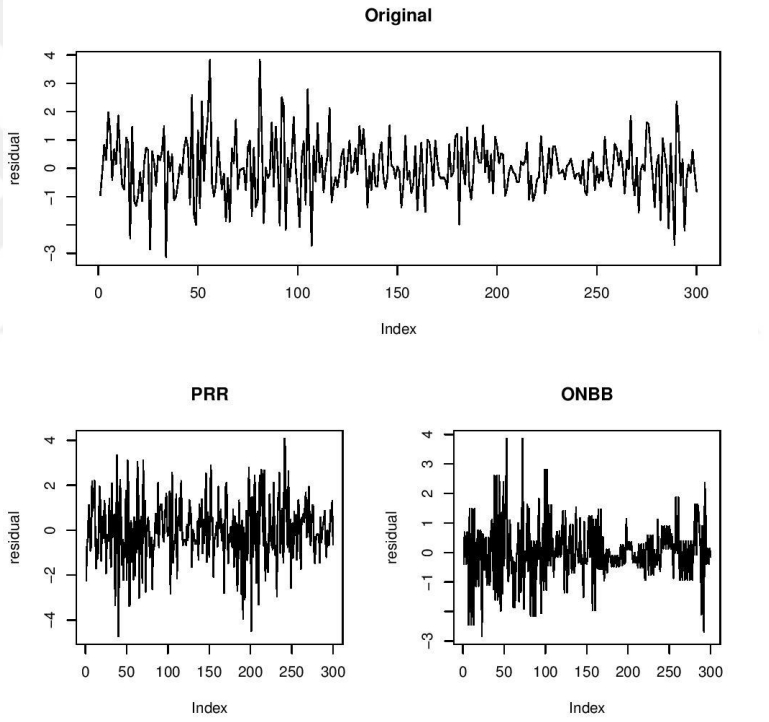

Figure 3.10 Unstandardised residual plots of PRR and ONBB

Finally, to compare PRR and ONBB methods in terms of having more representative resamples, we present the plot of fitted unstandardised residuals obtained by both methods (see Figure 3.10) when the sample size n = 300. It is clear that the residuals obtained by ONBB have more similar fluctuations with the original residuals compared to PRR's residuals. We also conduct a simulation study when$B$ = 1000 and MC = 1000, and for each simulation we recorded the sum of squared differ-

Table 3.3 Prediction intervals for returns and volatilities of GARCH(1, 1) model

Lead time	Sample size	Method	Coverage for return (SE)	Average length for return (SE)	Coverage for volatility (SE)	Average length for volatility (SE)
1	T 300	Empirical	0.95	3.814	0.95	-
		PRR	0.945(0.021)	3.748(0.874)	0.904(0.295)	0.649(0.520)
		$\ell = n^{1/3}$	0.947(0.022)	3.804(0.819)	0.922(0.268)	0.743(0.565)
	ONBB	$\ell = n^{1/4}$	0.946(0.020)	3.853(0.954)	0.958(0.200)	0.779(0.752)
		$\ell = n^{1/5}$	0.945(0.022)	3.821(0.961)	0.948(0.222)	0.781(0.755)
	1500	PRR	0.946(0.013)	3.695(0.748)	0.928(0.259)	0.236(0.181)
		$\ell = n^{1/3}$	0.949(0.018)	3.838(0.804)	0.928(0.258)	0.724(0.842)
	ONBB	$\ell = n^{1/4}$	0.948(0.016)	3.853(0.838)	0.962(0.191)	0.703(0.590)
		$\ell = n^{1/5}$	0.946(0.016)	3.893(0.886)	0.950(0.218)	0.738(0.641)
	3000	PRR	0.946(0.011)	3.800(0.863)	0.952(0.214)	0.181(0.194)
		$\ell = n^{1/3}$	0.950(0.018)	3.838(0.718)	0.888(0.315)	0.718(0.698)
	ONBB	$\ell = n^{1/4}$	0.947(0.016)	3.865(0.845)	0.958(0.200)	0.719(0.590)
		$\ell = n^{1/5}$	0.948(0.015)	3.819(0.879)	0.974(0.159)	0.713(0.633)
10	T 300	Empirical	0.95	3.946	0.95	1.389
		PRR	0.943(0.026)	3.846(0.712)	0.902(0.117)	1.564(1.387)
		$\ell = n^{1/3}$	0.945(0.021)	3.881(0.659)	0.905(0.094)	1.410(0.851)
	ONBB	$\ell = n^{1/4}$	0.945(0.020)	3.941(0.789)	0.927(0.078)	1.638(1.281)
		$\ell = n^{1/5}$	0.946(0.020)	3.926(0.697)	0.941(0.077)	1.753(1.381)
	1500	PRR	0.946(0.014)	3.806(0.527)	0.930(0.056)	1.302(0.689)
		$\ell = n^{1/3}$	0.946(0.016)	3.870(0.617)	0.906(0.061)	1.271(0.871)
	ONBB	$\ell = n^{1/4}$	0.947(0.015)	3.908(0.627)	0.941(0.043)	1.434(0.782)
		$\ell = n^{1/5}$	0.945(0.014)	3.926(0.647)	0.948(0.041)	1.526(0.943)
	3000	PRR	0.946(0.012)	3.875(0.604)	0.941(0.036)	1.354(0.653)
		$\ell = n^{1/3}$	0.946(0.015)	3.866(0.563)	0.911(0.056)	1.287(0.833)
	ONBB	$\ell = n^{1/4}$	0.946(0.014)	3.909(0.645)	0.940(0.040)	1.411(0.786)
		$\ell = n^{1/5}$	0.947(0.012)	3.882(0.644)	0.959(0.027)	1.495(0.864)
20	T 300	Empirical	0.95	3.948	0.95	1.661
		PRR	0.940(0.026)	3.876(0.647)	0.881(0.122)	1.771(1.515)
		$\ell = n^{1/3}$	0.942(0.023)	3.896(0.593)	0.882(0.097)	1.582(0.925)
	ONBB	$\ell = n^{1/4}$	0.944(0.022)	3.964(0.706)	0.901(0.090)	1.847(1.438)
		$\ell = n^{1/5}$	0.944(0.021)	3.947(0.603)	0.920(0.086)	1.928(1.278)
	1500	PRR	0.946(0.015)	3.859(0.399)	0.925(0.059)	1.569(0.705)
		$\ell = n^{1/3}$	0.944(0.015)	3.898(0.519)	0.900(0.060)	1.464(0.879)
	ONBB	$\ell = n^{1/4}$	0.946(0.014)	3.929(0.465)	0.933(0.043)	1.666(0.798)
		$\ell = n^{1/5}$	0.945(0.015)	3.937(0.503)	0.940(0.040)	1.757(0.981)
	3000	PRR	0.946(0.012)	3.907(0.444)	0.940(0.033)	1.634(0.627)
		$\ell = n^{1/3}$	0.946(0.015)	3.903(0.470)	0.911(0.049)	1.471(0.796)
	ONBB	$\ell = n^{1/4}$	0.946(0.014)	3.934(0.505)	0.935(0.036)	1.645(0.747)
		$\ell = n^{1/5}$	0.946(0.013)	3.914(0.504)	0.951(0.029)	1.736(0.855)

ence between original and bootstrap standardized residuals for both methods.The result is 1.822 for the PRR while it is 0.886, 0.895 and 0.888 for the ONBB when the block size is $\ell = n^{1/3}$, $n^{1/4}$ and $n^{1/5}$, respectively. The results show that the ONBB has more similar resamples than PRR. Moreover, we performed a small simulation study to compare the robustness of the PRR and ONBB. To this end, we generated time series data as $(1-\alpha)$GARCH(1,1) + α AR(1) with AR parameter -0.9 and sample size n = 1000, where $0 < \alpha < 1/2$. This generated observations are "nearly GARCH" and we computed sum of squared difference between true residuals (unstandardised) and the residuals obtained by the bootstrap methods. The method having smaller difference values can be considered more robust than the other method. The results are presented in Table 3.4 for different α values. Consequently, we can say that the ONBB is more robust than the PRR.

Table 3.4 Sum of squared residuals of PRR and ONBB for different α values

Method	α				
	0.1	0.2	0.3	0.4	0.5
PRR	65086.006	1919.690	2321.617	11947.420	13409.390
ONBB	1432.617	1398.811	1555.436	1922.087	2522.636

3.6 Case Study

To obtain out-of sample prediction intervals for the real data described in Chapter 2, we divide the full data into the following two parts: The model is constructed based on the observations from 29th July, 2011 to 21st September, 2015 (1040 observations in total) to calculate 30 steps ahead predictions from 22nd September to 3rd November, 2015 and compare with the actual values. The fitted models for PRR and ONBB are obtained as in Equations (3.9) and (3.10), respectively.

$$\hat{\sigma}_t^2 = 0.0028 + 0.0549 y_{t-1}^2 + 0.9412 \hat{\sigma}_{t-1}^2 \tag{3.9}$$

$$y_t^2 = 0.0078 + 0.9845 y_{t-1}^2 + \nu_t - 0.9408 \nu_{t-1} \tag{3.10}$$

where $\hat{\omega} = 0.0078$, $\hat{\alpha}_1 = 0.0436$ and $\hat{\beta}_1 = 0.9408$ for the model estimated by Equation (3.4). The 30 step ahead prediction intervals of the PRR and ONBB for returns y_{T+h} based on the models given in Equations (3.9) and (3.10), together with the true returns are presented in Figure 3.11. The intervals obtained using both

methods are similar and they include all of the true values of returns. Note that, the ONBB prediction interval for $\ell = n^{1/5}$ is slightly narrower than the others.

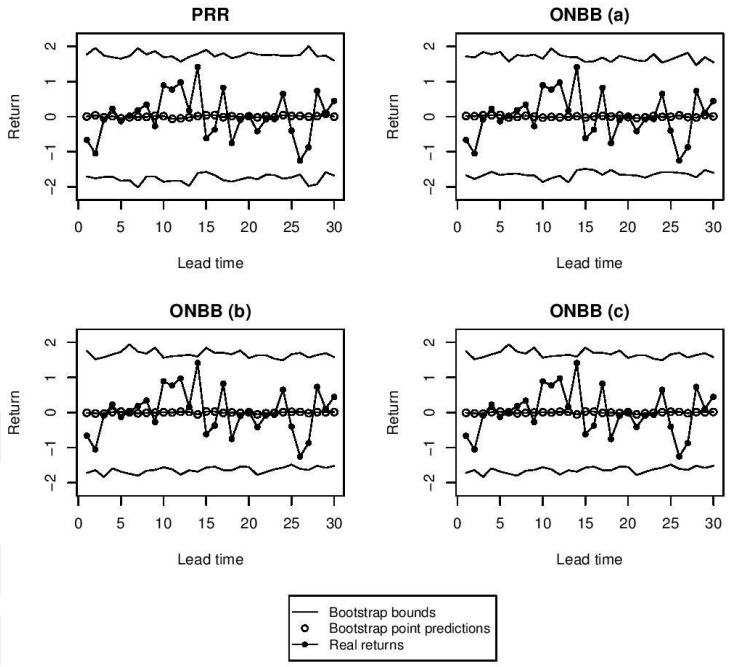

Figure 3.11 95% prediction intervals of returns from 22nd September, 2015 to 3rd November, 2015, where (a), (b) and (c) denote results obtained by choosing block lengths $\ell = n^{1/3}$, $n^{1/4}$ and $n^{1/5}$, respectively

Figure 3.12 shows the predicted intervals for 30 step ahead volatilities σ^2_{T+h}. The true values of the volatilities can not be observed directly.We calculate the realized volatility by summing squared returns at day *t*, $\sigma_t^2 = y_{t,1}^2 + \ldots + y_{t,n}^2$, where *n* is the number of observations recorded during day *t* as proposed by Andersen & Bollerslev (1998). Since our data is from 24 hour open trading market, the realized volatilities are computed by using one-minute returns based on tick-by-tick prices such that *n* = 1440 approximately. Figure 3.12 indicates that the ONBB prediction intervals are significantly narrower than the PRR's for all block lengths. Moreover, by looking at this figure carefully it can be seen that the point forecasts of the volatilities obtained by ONBB are closer to the realized values than the results based on the PRR method. This result clearly explains the supremacy of the ONBB based prediction intervals over the existing ones. Additionally, we perform an extra simulation study from a

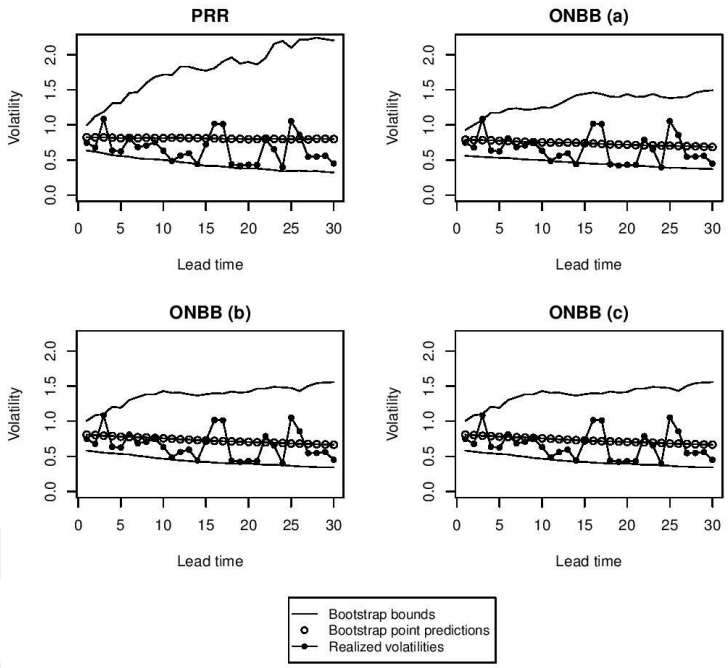

Figure 3.12 95% prediction intervals of volatilities from 22nd September, 2015 to 3rd November, 2015, where (a), (b) and (c) denote the results obtained choosing block lengths $\ell = n^{1/3}$, $n^{1/4}$ and $n^{1/5}$, respectively

GARCH(1,1) model with the parameters as in Equation (3.9) and sample size n = 1040 to compare the performances of the PRR and ONBB, and the results are presented in Figure 3.13. The results for the coverage probabilities of the returns are consistent with the simulation results given in Chapter 3.5. For coverage probabilities of the volatilities, the ONBB outperforms PRR when $\ell = n^{1/3}$ while the other block lengths ONBB over estimates the coverage probability for all lead times. PRR provides narrower intervals than the ONBB but the difference is not too significant especially when $\ell = n^{1/3}$.

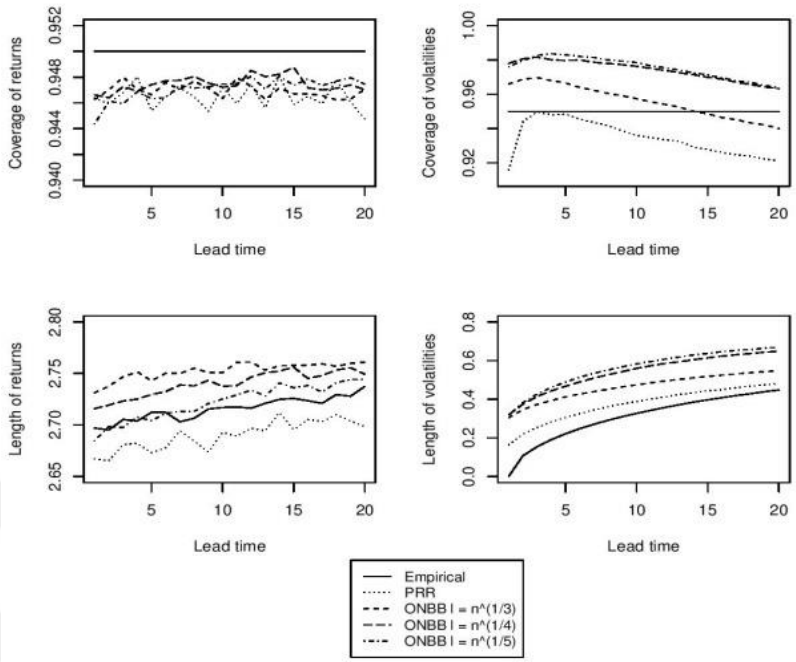

Figure 3.13 Simulation results from the GARCH(1,1) model fitted to the exchange rate data

CHAPTER FOUR
CONCLUSION

Our main aim in this dissertation is to develop new block bootstrap based methods for general class of financial time series models especially used in foreign exchange market analysis. Throughout this dissertation, we first propose sufficient, ordered, and sufficient ordered NBB methods to approximate the sampling distributions of the statistics under consideration. We evaluated the performances of our proposed methods by a simulation study for MA(2) and AR(2) processes and five real-world data sets, and compared their performances with the existing block bootstrap methods. The results show that the sufficient NBB yields the same results compared to the conventional NBB with less standard error for μ and it also reduces computing time. Also, although the difference is not statistically significant, sufficient NBB has better results for coverage probabilities. On the other hand, ordering the blocks causes the best performance for estimating the statistics especially for the sufficient bootstrap method. Since it produces close results to the true values of the statistics, it has appropriate coverage probabilities for those statistics. It should be noted tha while the sufficient bootstrap can only be used with NBB method, ordered bootstrapped blocks can be extended for other block bootstrap methods.

Secondly, we examine our proposed ONBB method in detail and we show its superiority over the traditional block bootstrap methods by Spearman's rank correlation coefficient. Our DTW simulation results show that the ONBB resamples are more similar to the original time series on time axis compared to other block bootstrap methods. Therefore, the ONBB method comprises of more dependency structure of the original series and produces more reliable results among the others.We also propose a novel, computationally efficient resampling algorithm to obtain better prediction intervals for returns and volatilities under GARCH models by using ONBB method, and we compare the performance of our method with the existing PRR method by both simulations and a case study. The important result

produced by our proposed algorithm is that the short-term and long-term forecasting can be done with considerably narrower intervals especially for future volatilities.

REFERENCES

Alam, Z. (2012). Forecasting the BDT/USD exchange rate using autoregressive model. *Global Journal of Management and Business Research, 12,* 85-96.

Andersen, T. G., & Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review, 39,* 885-905.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association, 96,* 42-55.

Athreya, K. B., & Lahiri, S. N. (2006). *Measure theory and probability theory*. NewYork: Springer.

Baillie, R. T., & Bollerslev, T. (1992). Prediction in dynamic models with time-dependent conditional variances. *Journal of Econometrics, 52,* 91-113.

Basu, D. (1958). On sampling with and without replacement. *Sankhya Series A, 20,* 287-294.

Beyaztas, U., & Alin, A. (2014). Sufficient jackknife-after-bootstrap method for detection of influential observations in linear regression models. *Statistical Papers, 55,* 1001-1018.

Beyaztas, B., Firuzan, E., & Beyaztas, U. (2017). New block bootstrap methods: Sufficient and/or ordered. *Communications in Statistics-Simulation and Computation, 46,* 3942-3951.

Billingsley, P. (1994). *Probability and Measure*. NewYork: Wiley.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *The Journal of Political Economy, 81*, 637-654.

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

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch model. *Review of Economics and Statistics, 72*, 498-505.

Bougerol, P., & Picard, N. (1992a). Strict stationarity of generalized autoregressive processes. *The Annals of Probability, 20*, 1714-1730.

Bougerol, P., & Picard, N. (1992b). Stationarity of garch processes and of some nonnegative time series. *Journal of Econometrics, 52*, 115-127.

Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis: Forecasting and control*, San Francisco:Holden-Day.

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (1994). *Time series analysis: Forecasting and control*. New Jersey: PrenticeHall.

Brooks, C., & Burke, S. P. (1998). Forecasting exchange rate volatility using conditional variance models selected by information criteria. *Economics Letters, 61*, 273-278.

Campbell, J. Y., Lo, A. W., & MacKinlay, A. C. (1997). *The econometrics of financial markets*. New Jersey: Princeton University Press.

Carlstein, E. (1986). The use of subseries values for estimating the variance of general statistics from a stationary sequence. *The Annals of Statistics, 14*, 1171-1194.

Chen, B., Gel, Y. R., Balakrishna, N., & Abraham, B. (2011). Computationally efficient bootstrap prediction intervals for returns and volatilities in arch and garch processes. *Journal of Forecasting*, *30*, 51-71.

Chen, J. (2015). *The unity of science and economics: A new foundation of economic theory*. New York: Springer.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, *7*, 1-26.

Efron, B., & Tibshirani, R. J. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science*, *1*, 54-77.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation.*Econometrica*, *50*, 987-1008.

Engle, R. F., & Patton, A. J. (2001). What good is a volatility model. *Quantitative Finance*, *1*, 237-245.

Fiser, R., & Horvath, R. (2010). Central bank communication and exchange rate volatility: a GARCH analysis. *Macroeconomics and Finance in Emerging Market Economies*, *3*, 25-31.

Freedman, D. A. (1981). Bootstrapping regression models. *The Annals of Statistics*, *9*, 1218-1228.

Gereben, A., & Pinter, K. (2005). Implied volatility of foreign exchange options: is it worth tracking?. *MNB Occasional Papers*, *39*, 1-45.

Ghalayini, L. (2014). Modeling and forecasting the US Dollar/Euro exchange rate. *International Journal of Economics and Finance*, *6*, 194-207.

Giorgino, T. (2009). Computing and visualizing dynamic time warping alignments in r: the dtw package. *Journal of Statistical Software*, *31*, 1-24.

Hall, P. (1985). Resampling a coverage pattern. *Stochastic Processes and their Applications*, *20*, 231-246.

Hall, P. (1992). *The bootstrap and Edgeworth expansion*. New York: Springer-Verlag.

Hall, P., Horowitz, J. L., & Jing, B. (1995). On blocking rules for the bootstrap with dependent data. *Biometrika*, *82*, 561-574.

Hannan, E. J., & Rissanen, J. (1982). Recursive estimation of mixed autoregressive-moving average order. *Biometrika*, *69*, 81-94.

Hansen, P. R., & Lunde, A. (2005). A forecast comparison of volatility models: does anything beat a garch(1,1)? *Journal of Applied Econometrics*, *20*, 873-889.

Hardle,W., & Bowman, A. (1988). Bootstrapping in nonparametric regression: Local adaptive smoothing and confidence bands. *Journal of the American Statistical Association*, *83*, 100-110.

Higgins, M. L., & Bera, A. K. (1992).A class of nonlinear arch models.*International Economic Review*, *33*, 137-158.

Hwang, E., & Shin, D. W. (2013). Stationary bootstrap prediction intervals for garch(p, q). *Communications for Statistical Applications and Methods*, *20*, 41-52.

Kunsch, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics*, *17*, 1217-1241.

Lahiri, S. N. (2003). *Resampling methods for dependent data*. NewYork: Springer.

Liu, R. Y. (1988). Bootstrap procedures under some non i.i.d. models. *The Annals of Statistics*, *16*, 1697-1708.

Liu, R. Y., & Singh, K. (1992). *Moving blocks jackknife and bootstrap capture weak dependence, Exploring the Limits of Bootstrap*. NewYork: Wiley.

Mandelbrot, B. (1963). The variation of certain speculative prices.*The Journal of Business*, *36*, 394-419.

Miguel, J. A., & Olave, P. (1999). Bootstrapping forecast intervals in arch models. *Test*, *8*, 345-364.

Nelson, D. B. (1990). Stationarity and persistence in the garch(1,1) model. *Econometric Theory*, *6*, 318-334.

Nelson, D. B. (1991). Conditional heteroscedasticity in asset returns: A new approach. *Econometrica*, *59*, 347-370.

Pascual, L., Romo, J., & Ruiz, E. (2006). Bootstrap prediction for returns and volatilities in garch models. *Computational Statistics and Data Analysis*, *50*, 2293-2312.

Pathak, P. K. (1962). On simple random sampling with replacement. *The Indian Journal of Statistics, Series A*, *24*, 287-302.

Politis, D. N., & Romano, J. P. (1992a). *A circular block resampling procedure for stationary data, Exploring the Limits of Bootstrap*. NewYork: Wiley.

Politis, D. N., & Romano, J. P. (1992b). A general resampling scheme for triangular arrays of α-mixing random variables with application to the problem of spectral density estimation. *The Annals ofStatistics*, *20*, 1985-2007.

Politis, D. N., & Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical Association*, *89*, 1303-1313.

R Core Team. (2014). *R:* A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Available at: http://www.R-project.org/

Raj, D., & Khamis, S. H. (1958). Some remarks on sampling with replacement. *The Annals of Mathematical Statistics*, *29*, 550-557.

Ratanamahatana, C. A., & Keogh, E. (2004). Everything you know about dynamic time warping is wrong. *Proceedings of International Conference on Knowledge Discovery and Data Mining*.

Reeves, J. J. (2005). Bootstrap prediction intervals for arch models. *International Journal of Forecasting*, *21*, 237-248.

Sentana, E. (1995). Quadratic arch models.*The Review of Economic Studies*, *62*, 639-661.

Shao, J., & Yu, H. (1993). Bootstrapping the sample means for stationary mixing sequences. *Stochastic Processes and Their Applications*, *48*, 175-190.

Singh, K. (1981). On the asymptotic accuracy of Efron's Bootstrap. *The Annals of Statistics*, *9*, 1187-1195.

Singh, S., & Sedory, S. A. (2011). Sufficient bootstrapping. *Computational Statistics and Data Analysis*, *55*, 1629-1637.

Thombs, L. A., & Schucany, W. R. (1990). Bootstrap prediction intervals for autoregression. *Journal of the American Statistical Association*, *85*, 486-492.

Tsay, R. S. (2005). *Analysis of financial time series*. New Jersey: Wiley.

Üstünel, İ. E. (1999). *Durağan portföy analizi ve İMKB verilerine uygulaması*. Phd Thesis, Hacettepe University, Ankara.