ESTIMATION, TESTING, AND MONITORING OF
GENERALIZED AUTOREGRESSIVE CONDITIONALLY HETEROSKEDASTIC
TIME SERIES

by

Aonan Zhang

A dissertation submitted in partial fulfillment
of the requirements for the degree
of
DOCTOR OF PHILOSOPHY
in
Mathematical Sciences

UTAH STATE UNIVERSITY
Logan, Utah
2005
Estimation, Testing, and Monitoring of Generalized Autoregressive Conditionally Heteroskedastic Time Series

by

Aonan Zhang, Doctor of Philosophy
Utah State University, 2005

Major Professor: Dr. Piotr Kokoszka
Department: Mathematics and Statistics

We study in this dissertation Generalized Autoregressive Conditionally Heteroskedastic (GARCH) time series. The research focuses on squared GARCH sequences. Our main results are as follows:

1. We compare three methods of constructing confidence intervals for sample autocorrelations of squared returns modeled by models from the GARCH family. We compare the residual bootstrap, block bootstrap and subsampling methods. The residual bootstrap based on the standard GARCH(1,1) model is seen to perform best. Confidence intervals for cross-correlations of a bivariate GARCH model are also studied.

2. We study a test to discriminate between long memory and volatility changes in financial returns data. Finite sample performance of the test is examined and compared using various variance estimators. The Bartlett kernel estimator with truncation lag determined by a calibrated bandwidth selection procedure is seen to perform best. The testing procedure is robust to various GARCH-type models.

3. We propose several methods of on-line detection of a change in unconditional variance in a conditionally heteroskedastic time series. We follow a paradigm in which the first $m$
observations are assumed to follow a stationary process and the monitoring scheme has asymptotically controlled probability of falsely rejecting the null hypothesis of no change. Our theory is applicable to broad classes of GARCH-type time series and relies on a strong invariance principle which holds for the squares of observations generated by such models. Practical implementation of the procedures is proposed and the performance of the methods is investigated by a simulation study.
ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my advisor, Piotr Kokoszka, for his constant support and guidance during my entire graduate study and dissertation preparation. It probably would have never been completed had it not been for his inspiration, encouragement, and advice. It has been a great pleasure and rewarding experience to work under his supervision.

I would also like to thank Richard Cutler, Michael Minnotte and Jürgen Symanzik for severing on my PhD committee, and for their lectures that provided me with deep statistical knowledge. Special thanks are due to Rimma Shiptsova for agreeing to be a committee member on short notice. I am greatly indebted to Gilles Teyssière and Lajos Horváth for their contributions and advice that make the research so fruitful. The research was partially supported by NSF grant INT-0223262.

I have been fortunate to work in the pleasant atmosphere of the Department of Mathematics and Statistics at Utah State University, Logan. I am grateful for the financial assistantship that the department offered me, and would like to thank all the faculty members and staff for their support.

I dedicate this dissertation to my parents, to whom I am indebted the most, for their unconditional love. I pray for their good health. I thank my sister and her family for taking care of my parents for me. Last but not least at all, I want to thank my beloved wife, Fang Qiu, for her love, patience, and strong belief in me.

Aonan Zhang
# CONTENTS

| ABSTRACT | ii |
| ACKNOWLEDGMENTS | iv |
| LIST OF TABLES | vii |
| LIST OF FIGURES | xii |

## 1 INTRODUCTION

1.1 Objectives and Main Results of the Research
1.2 Financial Time Series and GARCH Models
    1.2.1 Stylized facts of financial returns data
    1.2.2 GARCH models
    1.2.3 Quasi maximum likelihood estimation
    1.2.4 Financial applications of GARCH models
1.3 Resampling Methods
    1.3.1 Efron’s bootstrap
    1.3.2 Resampling methods for dependent observations
1.4 Long Memory and Change-Point Models
    1.4.1 Long memory in financial returns data
    1.4.2 Change point models
1.5 Sequential Testing
1.6 Outline of the Dissertation

## 2 CONFIDENCE INTERVALS FOR THE AUTOCORRELATION OF SQUARED GARCH SEQUENCES

2.1 Introduction
2.2 Confidence Intervals for Autocorrelations of Squared Observations
    2.2.1 Residual bootstrap
    2.2.2 Block bootstrap
    2.2.3 Subsampling
2.3 GARCH Models
2.4 Simulation Results
    2.4.1 Residual bootstrap
    2.4.2 Block bootstrap
    2.4.3 Subsampling
    2.4.4 Summary
2.5 Confidence Intervals for Cross-Correlations
2.6 Simulation Results
    2.6.1 Residual bootstrap for bivariate observations
    2.6.2 Subsampling for bivariate observations
2.6.3 Summary ........................................... 48

3 DISCRIMINATING BETWEEN LONG MEMORY AND CHANGE-POINTS IN VOLATILITY ........................................... 49
  3.1 Introduction ........................................ 49
  3.2 Historical and Empirical Background .................. 50
  3.3 The Test ............................................ 55
  3.4 Variance Estimation .................................. 57
  3.5 Patterns of Volatility ................................ 60
  3.6 Simulation Study .................................... 63

4 MONITORING CONSTANCY OF VARIANCE IN CONDITIONALLY HETEROSKEDASTIC TIME SERIES ........................................... 73
  4.1 Introduction ........................................ 73
  4.2 Problem Formulation and Assumptions .................. 74
  4.3 Detection Algorithms ................................ 76
    4.3.1 CUSUM and fluctuation monitoring schemes .......... 77
    4.3.2 Monitoring schemes based on partial sums of residuals and recursive residuals ........................................... 79
  4.4 Proofs of Theorems 4.3.2 and 4.3.3 .................... 81
  4.5 Proofs of Theorems 4.3.4 and 4.3.5 .................... 86
  4.6 Estimation of the Asymptotic Variance .................. 89
  4.7 Empirical Analysis of Index Data ....................... 92
  4.8 Simulation Study .................................... 96

5 CONCLUSION AND FUTURE RESEARCH .................................. 104

REFERENCES ............................................. 106

APPENDIX .............................................. 121
  A DERIVATIONS OF FORMULAS IN CHAPTER 2 .................. 122
  B ADDITIONAL TABLES AND FIGURES OF CHAPTER 2 ............ 128
  C ADDITIONAL TABLES AND FIGURES OF CHAPTER 3 ............ 137
  D ADDITIONAL TABLES AND FIGURES OF CHAPTER 4 ............ 142

VITA ............................................ 147
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Five standard GARCH(1, 1) models with given parameters, theoretical values of $\gamma_{c2}$ and $\rho_{X^2}(1)$</td>
<td>34</td>
</tr>
<tr>
<td>2.2</td>
<td>Five GJR-GARCH(1, 1) models with given parameters, theoretical values of $\gamma_{c2}$ and $\rho_{X^2}(1)$</td>
<td>35</td>
</tr>
<tr>
<td>2.3</td>
<td>Five NLGARCH(1,1,2) models with given parameters, theoretical values of $\gamma_{c2}$ and $\rho_{X^2}(1)$</td>
<td>35</td>
</tr>
<tr>
<td>2.4</td>
<td>Empirical coverage probabilities of equal-tailed confidence intervals constructed using <em>residual bootstrap</em></td>
<td>36</td>
</tr>
<tr>
<td>2.5</td>
<td>Empirical coverage probabilities of symmetric confidence intervals constructed using <em>residual bootstrap</em></td>
<td>37</td>
</tr>
<tr>
<td>2.6</td>
<td>Empirical coverage probabilities of equal-tailed and symmetric confidence intervals constructed using <em>block bootstrap</em> for standard GARCH models</td>
<td>40</td>
</tr>
<tr>
<td>2.7</td>
<td>Empirical coverage probabilities of equal-tailed and symmetric confidence interval constructed using <em>subsampling</em> for standard GARCH models</td>
<td>41</td>
</tr>
<tr>
<td>2.8</td>
<td>Three CCC-GARCH(1,1) models with given parameters, theoretical values of $r_0$, $r_1$, and $\lambda$</td>
<td>45</td>
</tr>
<tr>
<td>2.9</td>
<td>Empirical coverage probabilities of equal-tailed and symmetric confidence interval based on the method of <em>residual bootstrap</em> for $r_0$ and $r_1$ in the CCC-GARCH model</td>
<td>46</td>
</tr>
<tr>
<td>2.10</td>
<td>Empirical coverage probabilities of equal-tailed and symmetric confidence interval based on the <em>subsampling</em> method for $r_0$ and $r_1$ in the CCC-GARCH model, with $n = 500$</td>
<td>47</td>
</tr>
<tr>
<td>3.1</td>
<td>Asymptotic critical values of the test statistic $M_n$</td>
<td>56</td>
</tr>
<tr>
<td>3.2</td>
<td>Fitted GARCH(1,1) models from the returns on DJIA, NASDAQ, General Electric, Wal-Mart, American Express, and Altria Group</td>
<td>62</td>
</tr>
</tbody>
</table>
3.3 Empirical sizes (in percent) of the test using four different variance estimators applied to simulated series of squared GARCH(1,1) observations following the change-point models estimated from the returns on DJIA, NASDAQ, GE, and Wal-Mart in Table 3.2. The number of replications is \( R = 5000 \), with nominal size \( \alpha = 5\% \).

3.4 Empirical sizes (in percent) of the test using the Bartlett kernel estimator with lags \( q^*_c,\text{arma} \) applied to simulated series of squared GARCH(1,1) observations following the change-point models estimated from the returns on American Express and Altria Group in Table 3.2. The number of replications is \( R = 5000 \).

3.5 Empirical sizes (in percent) of test using the Bartlett kernel estimator with lags \( q^*_c,\text{arma} \) applied to simulated series of squared EGARCH and TGARCH observations following the change-point models estimated from the returns on DJIA, Altria Group, GE, and American Express in Table C.3. The number of replications is \( R = 5000 \).

3.6 Empirical power (in percent) of the test using the Bartlett kernel estimator with lags \( q^*_c,\text{arma} \) applied to simulated series of squared LARCH observations with parameters \( d = 0.35, a = 0.85, \) and \( b_0 = 0.35 \). The number of replications is \( R = 1000 \).

3.7 Realized test statistic \( M_n's \) using the Bartlett kernel estimator with truncation lag \( q^*_c,\text{arma} \) for the six sets of real returns data shown in Figure 3.1.

4.1 Fitted models of the returns data.

4.2 Empirical sizes (in percent) for the four monitoring methods applied to simulated series of squared GARCH(1,1) observations following Model 1 and Model 7 in Table 4.1. The number of replications is \( R = 1000 \) with historical size \( m = 500 \) and autocovariances computed sequentially.
4.3 Empirical power and ARL of the four methods applied to the two types of variance changes. The number of replications is $R = 1000$, with controlled size of 10%, and autocovariances computed sequentially. The monitoring horizon is $q = 500$. 

4.4 Comparison of empirical sizes of the four methods considered in this chapter with the method based on likelihood scores. Results are reported for the Bartlett kernel, sequentially computed autocovariances, $m = q = 1000$ and are based on $R = 2000$ replications.

B.1 Empirical coverage probabilities of equal-tailed confidence intervals based on the method of block bootstrap.

B.2 Empirical coverage probabilities of symmetric confidence intervals based on the method of block bootstrap.

B.3 Empirical coverage probabilities of equal-tailed confidence intervals based on the method of subsampling.

B.4 Empirical coverage probabilities of symmetric confidence intervals based on the method of subsampling.

B.5 Empirical coverage probabilities of equal-tailed and symmetric confidence interval based on the subsampling method for $r_0$ and $r_1$ in the CCC-GARCH model, with $n = 1000$.

C.1 Empirical sizes (in percent) of the test using four different variance estimators applied to simulated series of squared GARCH(1,1) observations following the change-point models estimated from the returns on DJIA, NASDAQ, GE, and Wal-Mart in Table 4.1. The number of replications is $R = 5000$, with nominal size $\alpha = 1\%$. 

C.2 Empirical sizes (in percent) of the test using four different variance estimators applied to simulated series of squared GARCH(1,1) observations following the change-point models estimated from the returns on DJIA, NASDAQ, GE, and Wal-Mart in Table 4.1. The number of replications is \( R = 5000 \), with nominal size \( \alpha = 10\% \).......

C.3 Fitted EGARCH(1,1) and TGARCH(1,1) models from the returns on DJIA and Altria Group, and General Electric and American Express, respectively.

D.1 Empirical sizes (in percent) for the four monitoring methods applied to simulated series of squared GARCH(1,1) observations following Model 1 and Model 7 in Table 4.1. The number of replications is \( R = 1000 \) with historical size \( m = 100 \) and autocovariances computed sequentially.

D.2 Empirical sizes (in percent) for the four monitoring methods applied to simulated series of squared GARCH(1,1) observations following Model 1 and Model 7 in Table 4.1. The number of replications is \( R = 1000 \) with historical size \( m = 200 \) and autocovariances computed sequentially.

D.3 Empirical sizes (in percent) for the four monitoring methods applied to simulated series of squared GARCH(1,1) observations following Model 1 and Model 7 in Table 4.1. The number of replications is \( R = 1000 \) with historical size \( m = 300 \) and autocovariances computed sequentially.

D.4 Empirical sizes (in percent) for the four monitoring methods applied to simulated series of squared GARCH(1,1) observations following Model 1 and Model 7 in Table 4.1. The number of replications is \( R = 1000 \) with historical size \( m = 400 \) and autocovariances computed sequentially.

D.5 Empirical sizes (in percent) for the four monitoring methods applied to simulated series of squared GARCH(1,1) observations following Model 1 and Model 7 in Table 4.1. The number of replications is \( R = 1000 \) with historical size \( m = 500 \) and autocovariances computed using the historical \( m \) observations.
D.6 Empirical power and ARL of the four methods applied to the two types of variance changes. The number of replications is \( R = 1000 \), with controlled size of 10\%, and \( \hat{\eta} \) computed using the historical \( m \) observations. The monitoring horizon is \( q = 500 \).
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Daily returns on Dow Jones Industrial Averages Index from 01/01/1990 to 12/31/1999.</td>
</tr>
<tr>
<td>1.2</td>
<td>Sample autocorrelations for log-returns and squared log-returns on DJIA. At every lag ( h ) the two dashed horizontal lines mark an asymptotic 95% acceptance region of the null hypothesis of zero autocorrelation (ACF). We refer to Section 7.2 in [28] for a derivation of this region.</td>
</tr>
<tr>
<td>1.3</td>
<td>Left-hand: Kernel estimate of the density of DJIA log-returns, with fitted Gaussian curve. Right-hand: Normal Quantile-Quantile plot of the DJIA log-returns.</td>
</tr>
<tr>
<td>2.1</td>
<td>Comparison of ECP’s for symmetric residual bootstrap confidence intervals based on standard GARCH and a correct specification. The nominal coverage of 95% is marked by solid horizontal line.</td>
</tr>
<tr>
<td>2.2</td>
<td>Of symmetric confidence intervals, performance comparison among the three methods. The lengths of sequences are ( n = 500 ) and ( n = 1000 ) in the two plots, respectively. The nominal coverage 95% is marked by solid horizontal line.</td>
</tr>
<tr>
<td>3.1</td>
<td>Daily returns on stocks and stock indexes. Dashed lines indicate the borderlines of the subsamples, i.e., the location of the estimated change-point.</td>
</tr>
<tr>
<td>3.2</td>
<td>Comparison of empirical sizes for truncation lags ( q_{\text{arma}}^* ) and ( q_{\text{c arma}}^* ). Results are reported for the change-point models for DJIA and GE. The straight line indicates the nominal size of 5%.</td>
</tr>
<tr>
<td>4.1</td>
<td>Daily returns on four stock indexes. Dotted lines indicate the borderlines of the subsamples, i.e., the locations of the expected change-points.</td>
</tr>
<tr>
<td>4.2</td>
<td>Estimate of break point of variance. Dotted lines indicate the same dates as in Figure 4.1.</td>
</tr>
</tbody>
</table>
4.3 Comparison of empirical sizes. Results are reported for Model 1 with either Bartlett kernel or QS kernel. The monitoring horizon is \( q = 3 \). The straight line marks the target size of 10%. 99

4.4 Comparison of the distributions of the first hitting time. Results are reported for \( m = 300 \) with break at \( n = 331 \), 10% controlled size and the Bartlett kernel. The number of replications is \( R = 1000 \). 101

B.1 Of residual bootstrap method, performance comparison between the equal-tailed and the symmetric confidence interval. The lengths of sequences are \( n = 100, 250, 500, 1000 \) in the four plots, respectively. The nominal coverage 95% is marked by solid horizontal line. 129

B.2 Of block bootstrap method, performance comparison between the smallest and greatest value of block size \( b \) used for \( n = 500, 1000 \). The top two plots are for the equal-tailed confidence interval, the bottom two for the symmetric confidence interval. The nominal coverage of 95% is marked by solid horizontal line. 132

B.3 Of subsampling method, performance comparison between the smallest and greatest value of block size \( b \) used for \( n = 500, 1000 \). The top two plots are for the equal-tailed confidence interval, the bottom two for the symmetric confidence interval. The nominal coverage of 95% is marked by solid horizontal line. 135

B.4 Of symmetric confidence intervals, performance comparison between the methods of residual bootstrap and subsampling. The length of sequence is \( n = 1000 \), and \( b = 10 \) is used for subsampling. The nominal coverage 95% is marked by solid horizontal line. 136

C.1 Comparison of empirical sizes for truncation lags \( q_{\text{arma}}^* \) and \( q_{\text{c arma}}^* \). Results are reported for the change-point models for NASDAQ and Wal-Mart. The straight line indicates the nominal size of 5%. 140
C.2 Simulated LARCH observations with $a = 0.85$, $b_0 = 0.35$, and $d = 0.35$. The lengths of the series in the four panels are $n = 500, 1000, 2000$, and 1500 clockwise.

D.1 Comparison of the distributions of the first hitting times. Results are reported for $m = 100$ with break at $n = 111$, 10% controlled size and the Bartlett kernel. The number of replications is $R = 1000$.

D.2 Comparison of the distributions of the first hitting times. Results are reported for $m = 500$ with break at $n = 551$, 10% controlled size and the Bartlett kernel. The number of replications is $R = 1000$. 
CHAPTER 1
INTRODUCTION

1.1 Objectives and Main Results of the Research

Since the introduction of the Autoregressive Conditional Heteroskedastic (ARCH) models by Engle in 1982, the development and applications of various conditionally heteroskedastic models have become overwhelming. Among various extensions, the GARCH (Generalized ARCH) models form a generic class of models that has gained increasing popularity in academia and industry. GARCH models are applicable to almost all classes of financial asset returns. Our study aims at providing several empirical and theoretical analyses of GARCH time series in different settings.

This dissertation consists of three main chapters. Each of them focuses on a different aspect of the GARCH model. We next give a brief introduction to what we have done and to the objectives of each chapter. All concepts and quantities referred to in the remainder of this section are explained in detail in the following sections of the Introduction and in the relevant chapters of the dissertation.

Our study in Chapter 2 is motivated by the rich dependence structure in squared financial returns data. To illustrate the idea, suppose we have a series of returns $r_t$ and denote by $X_t = r_t^2$ the squared returns. We focus on squared returns because the returns themselves are practically uncorrelated and the $r_t^2$ give information about the volatility (variability) of the returns. This information is essential in many applications, as explained in Section 1.2.4. We construct confidence intervals for the autocorrelations of $X_t$ to estimate the strength of the dependence among the squared returns. To do so, we use various resampling methods for dependent observations. They are residual bootstrap, block bootstrap and subsampling. The first method relies on a model specification for the returns $r_t$, for which we use various members of the GARCH family. The latter two methods are model free. Their performance however depends largely on a block size $b$, and choosing appropriate $b$ is often difficult. We examine and compare how each method works by means of an extensive simulation study. The objective is to come up with an effective method of estimating how strong the squared
financial returns are correlated and to provide practical guidance in selecting the block size $b$ when either block bootstrap or subsampling is used. The residual bootstrap based on the standard GARCH(1,1) model is seen to perform best. The method also turns out to be superior in estimating cross-correlations for a bivariate GARCH model, which has wide applications in portfolio management.

There is a large amount of empirical evidence that financial market volatility exhibits long-range dependence. This concept, which is explained in greater detail in Section 1.4.1 and in Chapter 3, refers roughly to the persistent correlation of squared returns separated by long periods of time. Two competing modeling approaches for this phenomenon coexist in the literature: long memory models versus change-point models, see Section 1.4. The choice between the two positions has important implications in various financial applications, particularly in risk measurement, asset allocation and option pricing. We study in Chapter 3 a test designed to distinguish between long memory and volatility changes in financial returns. The test statistic requires an estimate of the long-run variance of the squared series. The difficulty lies in how to obtain such an estimate. We use various estimators and assess their performance. The objective of the comparative study is to find a testing procedure, which has a controlled small probability of falsely rejecting volatility changes when the volatility does change in the data and, in the meanwhile, can effectively reject volatility changes when the underlying model of the series is indeed a long memory one. We use real returns data covering different time periods to obtain practically relevant change-point GARCH models that we use in simulations. The testing procedure using the variance estimator proposed by us performs best. It also proves to be robust to various GARCH-type models.

In financial markets, volatility is closely related to risk and, in a sense, more volatile assets are riskier. Investors tend to expect more returns from a riskier asset. Volatility is one of the most important quantities in many financial applications. We study in Chapter 4 four monitoring schemes for detecting volatility changes in financial returns. The challenge is to provide theoretical justifications for the procedures. For that, we provide rigorous
proofs. We assume the constancy of volatility within the historical $m$ observations and begin to monitor to decide if there is a change from the $(m+1)$th data point. The objective is to develop an algorithm that falsely signals a change with a controlled small probability and, at the same time, can correctly detect a change with high probability and short delay as well. To simulate observations that have a volatility change, we use returns on four major US stock indexes and estimate GARCH models before and after the estimated breaks. From a practical point of view, we consider various implementations of the monitoring schemes and give comparative discussions and suggestions. Our methods perform better than those existing in the literature.

1.2 Financial Time Series and GARCH Models

This section gives introduction to financial time series and GARCH models. Estimation methods and financial applications of GARCH models are also discussed.

1.2.1 Stylized facts of financial returns data

Financial time series usually consist of a sequence of prices $P_t, t = 0, \ldots, n$, of a certain asset. The financial asset could be, for example, the stock of a company or a stock index, an exchange rate between two different currencies, or a commodity, such as gold or oil. It is assumed that the observations are equally spaced. Time separation can range from 1 minute to 1 year. The prices form a nonstationary and therefore are often transformed to log-returns that appear to be stationary. The log-returns are defined by

$$r_t = \log P_t - \log P_{t-1} = \log \left(1 + \frac{P_t - P_{t-1}}{P_{t-1}}\right).$$

Applying first-order Taylor series approximation to the natural log function, it can be readily seen that the log-returns are almost indistinguishable from relative returns, provided the latter are small, that is,

$$r_t \approx \frac{P_t - P_{t-1}}{P_{t-1}}.$$

Due to the product rule of logarithmic functions, log-returns are advantageous over relative returns for they are additive with respect to different time spans. One of the earliest and
most enduring models of the behavior of financial prices $P_t$ is the Random Walk Hypothesis, an idea that was conceived in the sixteenth century as a model of games of chance. Its first serious application to financial markets can be traced back to Paul Samuelson ([155], 1965), who argued that, in an informationally efficient market, price changes must be unforecastable if they fully incorporate the expectations and information of all market participants. Almost a decade later, Samuelson ([156], 1973) used geometric Brownian motion to model speculative prices in continuous time. The discretization of this model leads to the conclusion that the log-prices ($\log P_t$) evolve according to a random walk, and log-returns $r_t$ thus constitute a series of independent and identically distributed (i.i.d.) Gaussian random variables. The empirical studies of, for example, Mandelbrot [126] and Fama [64], on several US stock log-returns time series however rejected the random walk hypothesis with Gaussian increments as a realistic model for financial returns data. Their conclusion was based on the following so-called “stylized-facts” of financial returns:

1. returns $r_t$ exhibit serial dependence,
2. there is time-varying volatility, or conditional heteroskedasticity in the data,
3. the distribution of the data is heavy-tailed and asymmetric and hence non-Gaussian.
We present in Figure 1.1 the daily log-returns on Dow Jones Industrial Average index (DJIA) during the ten years time span of the 1990s. As one of the oldest and most widely quoted indexes, DJIA is a price-weighted average of thirty US blue chip companies. For easy visualization and following a common practice, the returns in this plot and subsequent figures are $100r_t$, with $r_t$ computed according to (1.2.1). Before using this particular time series to demonstrate the three facts in detail, let us first recall some definitions in classical time series analysis.

**Definition 1.2.1** Let $\{X_t\}$ be a time series with $EX_t^2 < \infty$. The mean function of $X_t$ is

$$\mu_X(t) = E(X_t).$$

The covariance function of $\{X_t\}$ is

$$\gamma_X(r, s) = \text{Cov}(X_r, X_s) = E[(X_r - \mu_X(r))(X_s - \mu_X(s))]$$

for all integers $r$ and $s$.

**Definition 1.2.2** $\{X_t\}$ is (weakly) stationary if

(i) $\mu_X(t)$ is independent of $t$,

and

(ii) $\gamma_X(t + h, t)$ is independent of $t$ for each $h$.

Strict stationarity of a time series $\{X_t, t = 1, \pm 1 \ldots \}$ is defined by the condition that $(X_1, \ldots, X_n)$ and $(X_{1+h}, \ldots, X_{n+h})$ have the same joint distribution for all integers $h$ and $n > 0$. A strictly stationary time series with finite second moment is also weakly stationary, but usually not conversely. The term stationary refers to weakly stationary in this dissertation.

**Definition 1.2.3** Let $\{X_t\}$ be a stationary time series. The autocovariance function (ACVF) of $\{X_t\}$ is

$$\gamma_X(h) = \text{Cov}(X_{t+h}, X_t).$$
The autocorrelation function (ACF) of \( \{X_t\} \) is

\[
\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = \text{Cor}(X_{t+h}, X_t).
\]

That is, the autocovariance \( \gamma_X(h) \) and the autocorrelation \( \rho_X(h) \) at lag \( h \in \mathbb{Z} \) is the covariance and the correlation, respectively, between observations that are \( h \) lags apart.

If the autocovariance function \( \gamma_X(h) \) goes to zero sufficiently quickly as \( h \) becomes large, which is usually quantified by

\[
\sum_{h=0}^{\infty} |\gamma_X(h)| < \infty,
\]

we say the process \( \{X_t\} \) is weakly dependent. The sample counterparts of autocovariance and autocorrelation functions are defined as follows.

**Definition 1.2.4** Let \( X_1, \ldots, X_n \) be observations of a time series. The sample autocovariance function (SACVF) is

\[
\hat{\gamma}_{n,X}(h) = \frac{1}{n} \sum_{t=1}^{n-h} \left( X_t - \frac{1}{n} \sum_{t=1}^{n} X_t \right) \left( X_{t+h} - \frac{1}{n} \sum_{t+1}^{n} X_t \right),
\]

for \( 0 \leq h < n \). The sample autocorrelation function (SACF) is

\[
\hat{\rho}_{n,X}(h) = \frac{\hat{\gamma}_{n,X}(h)}{\hat{\gamma}_{n,X}(0)}, \quad 0 \leq h < n.
\]

For a broad class of stationary processes, known as ergodic processes, process \( \{X_t\} \), the \( \hat{\gamma}_{n,X}(h) \) and \( \hat{\rho}_{n,X}(h) \) are strongly consistent for \( \gamma_{n,X}(h) \) and \( \rho_{n,X}(h) \), respectively, that is,

\[
\hat{\gamma}_{n,X}(h) \xrightarrow{a.s.} \gamma_X(h) \quad \text{and} \quad \hat{\rho}_{n,X}(h) \xrightarrow{a.s.} \rho_X(h), \quad n \to \infty.
\]

Using the time series in Figure 1.1, we now come back to the discussion of the stylized facts of financial returns data. Figure 1.2 presents the sample autocorrelation (SACF) plots for, from left to right, the log-returns and the squared log-returns on DJIA. It shows, as it is well believed, that the SACF of real-life returns data are not significantly different from zero at any lag \( h \neq 0 \), except perhaps at the first lag \( |h| = 1 \). This observation does not however imply the independence of the returns data, because the SACF of the absolute and squared log-returns, denoted by \( \hat{\rho}_{n,|r|}(h) \) and \( \hat{\rho}_{n,r^2}(h) \), respectively, are different from zero
Figure 1.2. Sample autocorrelations for log-returns and squared log-returns on DJIA. At every lag $h$ the two dashed horizontal lines mark an asymptotic 95% acceptance region of the null hypothesis of zero autocorrelation (ACF). We refer to Section 7.2 in [28] for a derivation of this region.

for a large number of lags $h$. In the right-hand panel of the figure, we can clearly see that the $\hat{\rho}_{n,t-2}(h)$ does not vanish even at lags $h > 30$. If the $r_t$ were independent, then so would be the $r_t^2$.

In the past two decades, it has been carefully and convincingly established that stock market volatility is not constant but rather changes over time. See for example [157], [158], and [33]. In the language of statistics, the conditional variance of returns given the past, i.e., $\text{Var}[r_t|r_{t-1}, r_{t-2}, \ldots]$ is not constant over time and the underlying stochastic process $\{r_t\}$ is conditionally heteroskedastic. Econometricians would also say that the volatility of returns

$$\sigma_t = (\text{Var}[r_t|r_{t-1}, r_{t-2}, \ldots])^{1/2}$$

changes over time. Possible explanations of this characteristic of financial returns include political disorders, economic crises, and other catastrophic events that are believed to cause the prices of financial assets to fluctuate enormously. Stock volatility is also related to the time-varying volatility of a variety of economic variables, such as inflation and money
growth, and even to stock trading activity. It is evident from Figure 1.1 that there are stretches of time where the volatility is relatively high and stretches of time where the volatility is relatively low, that is, volatility values tend to cluster together in time, with more or less smooth transitions from higher to lower volatility and conversely.

Many financial log-returns time series seem to indicate that the unconditional distribution of $r_t$ is heavy-tailed, i.e., not all moments of $r_t$ exist. Distributions with elongated upper (lower) tails are called positively (negatively) skewed. Visually, skewed sample distributions have one “heavier” (longer) and one “lighter” (shorter) tail. We present in Figure 1.3 the density distribution and Normal Quantile-Quantile plot of the DJIA returns. Both plots show that the distribution has high and low outliers and thus possesses heavier, especially on the left, tails than Gaussian. Furthermore, the kurtosis coefficient for the returns data is $\kappa = 5.19$. Recall that the empirical kurtosis is conventionally defined as:

$$\kappa = \frac{\frac{1}{n} \sum_{t=1}^{n} (r_t - \bar{r})^4}{\left[ \frac{1}{n} \sum_{t=1}^{n} (r_t - \bar{r})^2 \right]^2},$$

where $\bar{r} = 1/n \sum_{t=1}^{n} r_t$ is the sample mean. Kurtosis is a measure of whether the data are peaked, or flat relative to a Gaussian distribution. It should be close to 3 in the case of
a Gaussian series $r_t$. Density of data sets with kurtosis higher than 3, as the returns on DJIA, tend to have a distinct peak near the mean, decline rather rapidly, and have heavy tails. All these features are clearly seen in Figure 1.3.

There are both empirical evidence and theoretical considerations that suggest that volatility tends to respond asymmetrically to positive and negative log-returns, that is, volatility tends to rise in response to negative shocks ("bad news") and tends to fall in response to positive shocks ("good news"). This negative correlation is also referred to as leverage effect in [18] by Black. Black reasons that when a company's stock price falls, its value of equity also falls. As a result, the company's leverage, or its debt-to-equity ratio, increases. In economics terms, leverage is generally interpreted as an indicator of a company riskiness. When the leverage ratio increases, investments on the company is considered more risky, and a higher degree of risk is associated with higher volatility. See [137] for a readable account on this subject.

1.2.2 GARCH models

Time series models have been initially introduced for descriptive purposes, prediction and seasonal correction and for dynamic control. Time series analysis of the 1970s focused on the so-called autoregressive moving average processes (ARMA). It is a specific class of time series models where the current value of the series of interest is specified as a linear function of its own lagged values and current and past values of some noise process, which can be interpreted as innovations to the system. The main drawback of the approach, which causes the poor performance of ARMA models in financial and monetary applications, is that its essential linear setup automatically restricts the type of dynamics to be approximated. Financial time series however feature various forms of nonlinear dynamics. The crucial one, which motivated Engle ([60]) to introduce ARCH (Autoregressive Conditionally Heteroskedastic) models, is the strong dependence of the instantaneous variability of the process on its own past, i.e., the time varying variance discussed in the previous section. His work in this area was recognized in 2004 by a Nobel Prize in Economics. Among the vari-
ous variations and extensions of the ARCH model, the GARCH\((p,q)\) (Generalized ARCH) model developed by Bollerslev ([20]) is the leading generic model for almost all classes of asset returns.

Before we define (G)ARCH models, we introduce ARMA processes because they form the backbone of classical time series analysis and because squared GARCH processes can be interpreted as ARMA processes, see [91] pp. 665-666 for the details.

A stationary process \(\{X_t\}\) follows an ARMA\((p,q)\) model if for every \(t \in \mathbb{Z}\)
\[
X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q},
\]
where polynomials \((1 - \phi_1 z - \cdots - \phi_p z^p)\) and \((1 + \theta_1 z + \cdots + \theta_q z^q)\) have no common factors and where \(Z_t\) are i.i.d. and are usually called innovations or errors. The process \(\{Z_t\}\) is also often called the noise sequence. The process \(\{X_t\}\) is said to be an ARMA\((p,q)\) with mean \(\mu\) if \(\{X_t - \mu\}\) is an ARMA\((p,q)\) process. Alternatively, we can write (1.2.13) in a more concise form
\[
\phi(B)X_t = \theta(B)Z_t,
\]
where \(B\) denotes the backward shift operator \((B^j X_t = X_{t-j}, B^j Z_t = Z_{t-j}, j = 0, \pm 1, \ldots)\) and \(\phi(\cdot)\) and \(\theta(\cdot)\) are \(p\)th and \(q\)th degree polynomials in (1.2.13).

We now define GARCH models. A time series \(\{X_t\}\) is called a GARCH\((p,q)\) process if it satisfies
\[
X_t = \sigma_t Z_t, \quad t \in \mathbb{Z}
\]
where \(Z_t\) are i.i.d. with mean zero and unit variance and are usually called the innovations (or the noise process), and where \(\sigma_t\) evolves according to
\[
\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2, \quad t \in \mathbb{Z},
\]
with \(\omega > 0, \alpha_i \geq 0 \text{ for } i = 1, 2, \ldots, p, \beta_j \geq 0 \text{ for } j = 1, 2, \ldots, q\). These constraints ensure that \(\sigma_t^2\) is nonnegative for all realizations of \(\{X_t\}\). Moreover, the stationarity condition for the process is given by
\[
\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1.
\]
When \( q = 0 \), (1.2.15) and (1.2.16) define an ARCH(\( p \)) process. By the statistical properties of \( Z_t \), the conditional variance of \( X_t \) given all the past information up to time \( t-1 \), denoted by \( I_{t-1} \), is

\[
\text{Var}(X_t|I_{t-1}) = \mathbb{E}(X_t^2|I_{t-1}) = \sigma_t^2,
\]

which obeys the recurrence equation (1.2.16) and is not constant over time. We thus can see that (G)ARCH models were formulated in a way to recognize and quantify the time-varying variance that is often observed in financial time series. Indeed, that is one of the underlying reasons why conditionally heteroskedastic models gained popularity in financial econometrics and among practitioners.

In financial markets, the risk of an asset can be equated to its variance and its expected return in equilibrium must be dependent on this risk. Thus the conditional variance of an asset interact with its conditional mean. This is the basic structure proposed in [58] as the GARCH-in-the-mean (GARCH-M) model. The conditional mean in this model is an arbitrary function of volatility \( \sigma_t \), as opposed to zero in standard GARCH models. Among GARCH models that allow for leverage effects, EGARCH proposed in [137] and GJR-GARCH developed in [75] are usually preferred over others in practice. For example, in a set of diagnostic tests of [63] to assess the ability of different models to forecast the Nikkie volatility, EGARCH and GJR-GARCH proved the best. To allow for high persistence and long memory in conditional variance of financial data, there are, for example, integrated GARCH (IGARCH) in [62], fractionally integrated GARCH (FIGARCH) in [8], and Linear ARCH (LARCH) studied in [154] and others. In IGARCH(1,1), we have \( \alpha_1 + \beta_1 = 1 \) in (1.2.16), which makes it clear that \( X_t^2 \) is integrated in this model. Among multivariate heteroskedastic models, the Constant Conditional Correlation model of multivariate GARCH (CCC-GARCH) proposed in [21] has proved to be one of the most stable and reliable of the multivariate processes. In this model, although variance, covariance, and correlations may vary over time, the conditional correlations are assumed constant. There are a wide range of other related models with names like PGARCH, PNP-GARCH, STARCH and multivariate FACTOR-GARCH.
1.2.3 Quasi maximum likelihood estimation

An estimate that maximizes a possibly misspecified likelihood function (for example, an MLE calculated under the assumption of a Gaussian process when the true data are non-Gaussian) is known as a quasi maximum likelihood estimate (QMLE). The QMLE is such a commonly used procedure for estimating GARCH parameters that most statistical packages for financial time series analysis contain the QMLE as a built-in function.

To illustrate the idea and focus attention, we first discuss the maximum likelihood estimation for GARCH(1,1) process with innovations that are Gaussian, i.e., \( Z_t \)'s are i.i.d. \( \sim N(0,1) \). For easy reference, a GARCH(1,1) process \( \{X_t\} \) is defined by

\[
(1.2.19) \quad X_t = \sigma_t Z_t, \quad \text{with } \sigma_t^2 = \omega + \alpha X_{t-1}^2 + \beta \sigma_{t-1}^2, \quad t \in \mathbb{Z}.
\]

Even under the assumption that \( Z_t \)'s are i.i.d. standard normal, there is no explicit expression of the probability density of an observed GARCH(1,1) vector \( (X_1, \ldots, X_n)^T \) since the distribution of \( (\sigma_1, \ldots, \sigma_n)^T \) is not known in closed form. We however can consider a conditional log-likelihood instead, because any likelihood can be decomposed into its conditional densities. Given \( X_0 \) and \( \sigma_0^2 \), the random variable \( X_1, \ldots, X_n \) are conditionally Gaussian with mean zero and variance \( \sigma_t^2(\theta) \), where \( \theta = (\omega, \alpha, \beta) \) denote the model parameters to be estimated and

\[
(1.2.20) \quad \sigma_t^2(\theta) = \omega + \alpha X_{t-1}^2 + \beta \sigma_{t-1}^2(\theta), \quad \text{for } 1 \leq t \leq n,
\]

whereas for \( t = 0 \) we let \( \sigma_0^2(\theta) = \sigma_0^2 \). We thus have the conditional Gaussian log-likelihood

\[
(1.2.21) \quad \ell(\theta) = \log f_\theta(X_1, \ldots, X_n|X_0, \sigma_0^2) = -\frac{n}{2} \log(2\pi) - \sum_{t=1}^{n} \log(\sigma_t(\theta)) - \frac{1}{2} \sum_{t=1}^{n} \frac{X_t^2}{\sigma_t^2(\theta)}.
\]

Since \( X_0 \) is not available and \( \sigma_0^2 \) is not observable, the log-likelihood (1.2.21) cannot be numerically evaluated without a certain initialization for \( X_0 \) and \( \sigma_0^2 \). It is shown however in [15] that the initial values are asymptotically irrelevant to the distribution of the estimates \( \hat{\theta} = (\hat{\omega}, \hat{\alpha}, \hat{\beta}) \). Because analytical forms of the MLE's are extremely difficult to derive, the log-likelihood \( \ell(\theta) \) is numerically maximized. With the derivative of (1.2.21) of the
The $t$th observation with respect to the parameter vector $\theta^T$, known as the $t$th score $s_t(\theta)$, the log-likelihood function can be maximized using the method of scoring as in [60], p. 997 or using the Berndt, Hall, Hall, and Hausman algorithm ([16]) as in [20], p. 317. Alternatively, the gradient of the log-likelihood can be calculated analytically from the sum of the scores

$$\nabla \ell(\theta) = \sum_{t=1}^{T} s_t(\theta)$$

or by numerical differentiation of $\ell(\theta)$. In addition, provided that the second derivative of the log-likelihood function $\ell(\theta)$ exists and the function $\ell(\theta)$ is concave, meaning that $-1$ times the matrix of second derivatives is everywhere positive definite, the MLE’s can be numerically obtained by using Newton-Raphson method. Section 5.7 of [91] discusses in detail other numerical maximization methods. Note that for general GARCH$(p, q)$ models the observations $X_0, \ldots, X_{p-1}$ and the squared volatilities $\sigma^2_0, \ldots, \sigma^2_{q-1}$ have to be initialized for the maximization of $\ell(\theta)$, in which case $\theta = (\omega, \alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q)$.

The preceding formulation of the log-likelihood function assumed that $Z_t$ has a Gaussian distribution. However, as we have discussed, the unconditional distribution of many financial time series seem to have heavier tails than allowed by the Gaussian family. Some of this can be explained by the GARCH model; even if $Z_t$ has a Gaussian distribution, the unconditional distribution of $X_t$ is non-Gaussian with fatter tails than a Gaussian distribution, see for example [133] or [20], p. 313. Even so, there is a fair amount of evidence that the distribution of $Z_t$ is non-Gaussian and thus the conditional distribution of $X_t$ is non-Gaussian as well and therefore $\ell(\theta)$ in (1.2.21) is invalid. The parameter estimates that maximize the log-likelihood in this case are QMLE. In [15] it is established that, under weak assumptions on the distribution of $Z_t$ (not necessarily Gaussian) and on model parameters, the consistency and asymptotic normality of QMLE are preserved. The standard errors of the QMLE however need to be adjusted from those obtained under Gaussian $Z_t$. To concentrate on important ideas rather than technicalities, we omit the conditions and detailed derivations of the asymptotic normality of the QMLE. They are available in [15] and [91].
1.2.4 Financial applications of GARCH models

In this section, we present several financial applications of GARCH models proposed in the literature and used by practitioners. The discussion covers financial decisions concerning risk analysis and option pricing and the interpretation of GARCH models as discrete approximations of continuous time models.

As we noted in Section 1.2.2, GARCH models capture the dynamic behavior of market volatility using specific volatility equations of the asset-return series of interest. They can be used to forecast the absolute magnitude of returns, and also to predict quantiles of the unconditional density of the returns, which has widespread applications in financial risk management. Value at Risk (VaR) in one of the most important measures of the market risk that has been widely used by institutions including banks, regulators and portfolio managers. It measures the potential risk of capital loss caused by adverse price moves in financial markets. A risk manager can use VaR to have a sense on the minimum amount that is expected to be lost with a small probability \( \alpha \) over a given holding period, denoted by \( k \), of usually one day or 10 days. For example, an \( \alpha = 5\% \) one-day VaR of $1 million tells us that one out of 20 days, we could expect to realize a loss of at least $1 million on the next business day. Alternatively, we could say that the maximum loss we could expect on 19 out of 20 days is $1 million. Indeed, VaR defines the maximum loss over a given period of time at a give confidence level. Mathematically, a \( k \)-day VaR of a long position (holding the asset) on day \( t \) is defined by

\[
P(P_{t+k} - P_t \leq \text{VaR}(t, k, \alpha)) = 1 - \alpha. \tag{1.2.23}
\]

Recall that \( P_t \) denotes the price of a financial asset on day \( t \). Given the distribution of the returns, VaR can be determined and expressed in terms of a percentile of the return distribution. See [52] and [106]. Specifically, if \( q_\alpha \) is the \( \alpha \)th percentile of the unconditional distribution of the returns \( r(t, k) = \log P_{t+k} - \log P_t \), then VaR can be written as

\[
\text{VaR}(t, k, \alpha) = (e^{q_\alpha} - 1)P_t. \tag{1.2.24}
\]

The above expressions indicates that good VaR estimate can only be produced with accu-
rate forecasts of the percentiles $q_\alpha$, which particularly concerns the tail properties of the return distribution. One of the main merits of GARCH process as an attractive modeling tool is that it can adequately describe the heavy-tailedness of the marginal distribution of financial returns data. Indeed, under very general conditions on the innovation process $Z_t$ the GARCH(1,1) in particular, and GARCH($p,q$) processes in general, have Pareto-like marginal distribution (see [47] and [131]), that is,

$$P(X > x) \sim c_0 x^{-\kappa} \quad \text{as } x \to \infty \quad \text{for some } c_0, \kappa > 0.$$  

Since the 70's financial derivatives, such as options and futures, have become a key component of investors' portfolio. The correct valuation of derivatives is of crucial importance for practitioners in any financial market. In their seminal paper [19], Black and Scholes derived a closed form solution for the price of a European option. The Black-Scholes formula has been celebrated as one of the major successes of modern financial economics. Empirical studies however has found several systematic pricing errors when the standard Black-Scholes model is used. These valuation errors are summarized in the so-called “smile” effect: When the implied volatilities, derived from the observed option prices, are plotted against money-ness and maturity, the resulting surface deviates significantly from a flat surface, which is predicted by the model. The assumptions underlying the model thus have been widely criticized. Apart from assuming continuous trading, a crucial assumption of the Black-Scholes method is the constant variance framework of using the geometric Brownian motion (the lognormality in discrete case) as the stochastic process of the financial asset that options trade on. The lognormal model however fails to explain a number of empirical regularities found in financial returns data, the most important of which are the heavy-tailedness of the unconditional distribution and the volatility clustering phenomenon.

As we have seen, GARCH models can in principle accommodate both excess kurtosis and volatility clustering. Efforts have been made towards deriving option pricing models with GARCH type volatility specifications in an equilibrium framework. Comparative

---

1 There are two kinds of option in terms of the time of exercising the contract. European options can be exercised only on the expiration date itself. American options can be exercised at any time up to the expiration date.
studies, for example, [59], [9], and [135], documented that the errors of the Black-Scholes model can be indeed be reduced if the correct model of the underlying assets are used. It is argued in [53] that for European call options the GARCH option pricing model with Gaussian innovations can explain some of the regularities found in the empirical option pricing literature like the smiles in implied Black-Scholes volatilities and the underpricing of short term options. Options on the Standard and Poor's 500 stock index have been the focus of a large amount of research. In [22] and [23], the authors have successfully used the GARCH option pricing framework together with fractionally integrated GARCH processes to price long-term European options on this particular index. In [119], the performance of three option pricing models are studied. They are the Black-Schole model and the valuation methods using GARCH and stochastic volatility (SV) as the respective volatility specifications. In terms of pricing errors, the GARCH option models was found clearly dominates the SV and the benchmark Black-Scholes model.

The main technical problem when using a GARCH option pricing model is that the distribution of future asset prices cannot be derived in closed form. Thus, generally no analytical formula for the option price exists and instead numerical methods have to be used. Recently, methods using lattice algorithms and Markov chain approximations have been proposed in [153] and [54], respectively.

We finally briefly discuss the interpretation of GARCH models as discrete approximation of continuous time models. Continuous time models have been widely used in theoretical finance, mainly because Ito's calculus usually provides us an elegant and relatively straightforward way of analyzing problems. Empirical difficulties however may arise when the data are available only at discrete times. The distribution of the series should be derived from the stochastic differential equations defining the model. In rare cases a closed form solutions can be derived. As an alternative approach, a recursive equation can be used as an approximation of the differential equations and the likelihood function can be maximized. As the length of the discrete time intervals between observations goes to zero, [136] investigated the convergence of stochastic difference equations, such as GARCH type models,
to stochastic differential equations. For detailed and rather technical exploitations of the subject, the read is referred to Chapter 5 of [76], where other financial applications of (G)ARCH models, such as the modeling aspects and the random walk hypothesis, are also discussed.

1.3 Resampling Methods

We start this section with a brief discussion of Efron's bootstrap for independent data and then extend the discussion to various resampling methods for dependent observations.

1.3.1 Efron's bootstrap

Following Efron's (1979) profound paper [56] introducing the bootstrap, substantial effort has been devoted to the development of bootstrap, and other resampling methods such as block bootstrap and subsampling, among others. The primary goal that these computer-intensive methods have achieved is that they provide alternative statistical tools of solving complex inferential difficulties without imposing unrealistic or unverifiable assumptions about the data-generating mechanism.

Both bootstrap and traditional parametric inference seek to use the limited information in the observed data to estimate the sampling distribution of a statistic. To focus attention and understand how bootstrap works, let us consider a simple example where we are interested in a parameter $\theta$ of a unknown population probability distribution. Let $\hat{\theta}_n$ be an estimator of $\theta$ calculated from a random sample, consisting of observations $X_1, \ldots, X_n$, drawn from the population. With knowledge of the sampling distribution, denoted by $F$, of the estimate $\hat{\theta}_n$, statisticians can next make inference about the parameter $\theta$. The key difference between traditional inferential methods and bootstrap methods lies in how the sampling distribution is obtained. Whereas parametric inference utilizes a priori assumptions about the distribution of the observations, bootstrap is not dependent on any particular class of distributions. It is done by randomly drawing a large number, $B$, of "resamples" of size $n$ from the original sample (of size $n$ too) with replacement, denoted by $X_1(b), \ldots, X_n(b), b = 1, \ldots, B$. So, although each of the $B$ resamples has the same
number of observations as the original sample, it could include some of the original data points more than once, and some not included. From each resample, the statistic \( \hat{\theta}_n(b) \) is calculated. The central assertion of bootstrapping is that the empirical distribution \( F_n^* \) of these \( B \) values \( \hat{\theta}_n(b) \) serves as a good approximation of the sampling distribution \( F_n \) of \( \hat{\theta}_n \). See [11], [87], [57], and [48] for detailed discussions of bootstrap methods and their properties.

With dependence in the data, the construction of the empirical distribution function \( F_n^* \) becomes more complicated and far less obvious than the basic procedure for the independent setup described above. In what follows, we discuss several resampling methods for dependent observations, especially for time series data. We use these methods in Chapter 2 to construct confidence intervals for the autocorrelations of squared GARCH sequences.

1.3.2 Resampling methods for dependent observations

It is well known that inference methods for i.i.d. data or, more generally, independent data need to be modified to be applicable and yield consistent results when the underlying process exhibits certain dependence structure.

Historically, the first extension was made under situations where a general regression model can be relied on. In those cases, the residuals of the regression models are resampled, not the original observations. The idea is to fit a suitable model to the data, to construct residuals from the estimated model, and then to generate new series by incorporating bootstrapped residuals into the fitted model. The residuals are typically recentered to have the same mean as the innovations of the model. Examples where this model-based resampling method is successful include linear regression (see, among others, [66], [67], [168], [122]) and autoregressive time series (for example, [57], [25]). Our interest in this dissertation is the finite sample performance of the residual bootstrap for conditionally heteroskedastic models.

The block bootstrap is probably the best-known method for implementing the bootstrap with time series data. It consists of dividing the data into blocks of consecutive observa-
tions of length $b$ and sampling the blocks $X_{t+1}, \ldots, X_{t+b}$ randomly with replacement: the blocking is used to preserve the original time series structure within a block. It is thus can easily seen that Efron's bootstrap for i.i.d. observations is just a special case of the block bootstrap, namely, if block of size $b = 1$ is used. Proper application of the block bootstrap procedure however involves first an adaptation to the parameter $\theta$ to be estimated. Suppose $\theta$ is a functional of the $m$-dimensional distribution of the time series, we construct the block-resampling on the basis of vectorized observations of length $m$. For example, the lag-1 correlation in a stationary time series corresponds to $m = 2$. The block bootstrap appears in [86], but the breakthrough is [117], in which detailed discussions on how and why the method works are provided. The blocks may be non-overlapping ([86], [36]) or overlapping ([86], [117], [146]). In constructing confidence intervals, regardless of the blocking method that is used, the block length $b$ must increase with increasing sample size $n$ to enable the block bootstrap to achieve asymptotically correct coverage probabilities. The asymptotically optimal block length may be defined as the one that minimizes the deviation of empirical coverage probability (ECP) from the target confidence level. Our investigation focuses on the finite sample performance of the method with different choices of $b$ in constructing confidence intervals.

The last method we examine for conditionally heteroskedastic time series is the subsampling method proposed in [147] and [148]. The basic idea is that the statistic of interest, such as an estimator or test statistic, is evaluated at subsamples of the data, and these subsampled values are used to build an estimated sampling distribution. Similar to the method of block bootstrap, the implementation of subsampling requires an integer $b < n$, such that blocks of consecutive data points of size $b$ will be allowed as proper subsamples, rather than any collection of $b$ observations sampled from the original sequence. Once the value of block size $b$ is determined, we can obtain $n - b + 1$ overlapping subsamples for a series of length $n$. The motivation is that within each subsample the dependence structure of the underlying process is preserved. Theoretically, asymptotic correct inference can ensue if the $b$ is allowed to increase to infinity with the sample size $n$. In the i.i.d.
context, subsampling is closely related to the usual bootstrap, the main difference being resampling the data without replacement versus with replacement. Its main advantage over bootstrapping is the very general setting in which the approximation with subsampling is consistent. See Section 3 of [148] for details. Historically, the idea of using subsamplings in statistical inference goes back to Mahalanobis ([125], 1946) who used it under the name “inter-penetrating network of subsampling” to assess and control non-sampling errors in the study of crop yields. Other early contributors to the methodology based on subsampling include, Quenouille ([152], 1949), Tukey ([163], 1958), and Hartigan ([92], 1969 and [93], 1975). The subsampling method provides inference tools for a wide variety of situations. It leads to valid large-sample statistical inferences in general estimation situations and data structures (including i.i.d. and stationary data), provided that the employed estimator converges to the unknown parameter at a known rate $\tau_n$ and possesses (when normalized by the rate $\tau_n$) a non-degenerate asymptotic distribution. 

In this section we merely give a brief introduction to some resampling methods for dependent data. The reader is referred to Chapter 2 for the detailed step-by-step procedures for constructing confidence intervals for parameters of conditionally heteroskedastic time series.

1.4 Long Memory and Change-Point Models

There are roughly two positions concerning how to model long time series observed in financial markets. One is in favor of long memory models, such as FIGARCH, FIEGARCH, and LARCH models in the GARCH family. The other prefers change-point models. In this section we qualitatively describe these two competing approaches.

1.4.1 Long memory in financial returns data

It has been a hypothesis of many early theories of the trade and business cycles that economic time series can exhibit long-term dependence. Researches in this line of thought have been motivated by the distinct but nonperiodic cyclical patterns that typified plots of many economic aggregates over time. In spectral domain such time series have power
at low frequencies. This particular feature is so common that in [77] it was regarded as the “typical spectral shape of an economic variable.” Long-range dependence has also been well-documented in hydrology, geophysics, and meteorology. As the ultimate sources of the unpredictable components in economics are natural phenomena like rainfall or earthquakes, we might also expect long memory in economic time series. Economic variables that are believed to exhibit persistent behavior over time include nominal and real interest rates, real exchange rates, exchange rate forward premiums, and, more recently, volatility measures of financial assets. A heuristic economic explanation to the cause of long memory is that unexpected shocks to the underlying variable have long lasting effects.

Recall in our previous discussion that although log-returns appear uncorrelated (sample autocorrelations (SACF) are zero at lags $h \neq 0$), the SACF of absolute and squared values are significantly different from zero even for large $h$. This long memory type of behavior of the SACF of absolute and squared log-returns is commonly believed as evidence for long memory in the volatility of financial returns data. The presence of long memory in assets returns has profound implications for many of the paradigms used in modern financial economics. For example, if stock returns exhibit such persistence, optimal ratio of consumption to savings and portfolio management may become extremely sensitive to the investment horizon. The conclusions of tests of “efficient” markets hypotheses or stock markets rationality rely greatly on the presence or absence of long-range dependence. Problems may also arise in the pricing of financial derivatives (such as futures and options) with martingale methods widely used, since the continuous time stochastic models are inconsistent with long-range dependence. Traditional tests of the capital asset pricing model and the arbitrage pricing theory are no longer valid since the usual statistical inferential methods are not directly applicable to long-range dependent time series.

Among the first to have studied the possibility and implications of persistent statistical dependence in financial returns was Mandelbrot ([127]). Since then, numerous empirical studies supported Mandelbrot’s finding. For example, [81], in which the authors claimed to have found long-term dependence in the daily returns of many stocks listed in the New York
Stock Exchange. “Anomalous” behavior in long-horizon stock returns has been uncovered by more recent investigations, which can be alternatively attributed to speculative fads and to time-varying conditional expected returns. These long-run swings may be further evidence of the long-range dependence in financial returns data. For more examples and detailed discussions, see [65], [105], and [150], among others.

1.4.2 Change point models

Consider a time series $X_1, \ldots, X_n$. A “change-point” refers to a unknown time $m < n$, at which the distribution function of the underlying data generating process switches from $F(X|\theta)$ that $X_1, \ldots, X_m$ follow to $F(X|\theta') (\theta' \neq \theta)$ that $X_{m+1}, \ldots, X_n$ come from. The number of change-points however need not be restricted to one. Suppose there are $N > 1$ breaks in a sample, then we have $N$ sub-series such that any consecutive two have different population distributions, or they come from different regimes. In econometric applications, the “change” may be generally interpreted as a change in model parameters, from $\theta$ to $\theta'$.

Structural stability or the constancy of model parameters is essential in statistical modeling and inference. Its importance is reflected in a large body of research on the subject of change-point detection. For surveys concerned with the general statistical methodology for change-point detection, we refer to [10], [29], [30], and [44].

Many recent papers have highlighted the fact that structural instability seems to be present in a wide variety of macroeconomic and financial time series. See for example [7] and [160]. Such dramatic breaks may result from events such as wars, financial panics, or significant changes in government policies. Structural stability is of prime importance in applied time series econometrics. The negative consequences of ignoring this instability include meaningless estimates, severely biased inferences, and inaccurate forecasts. Such pitfalls have been stressed by, among many others, [40], [41], and [112]. This has inspired a wide range of change-point models. There are two main approaches. The first approach is to estimate a model with a small number of change-points, usually not more than two. Recent influential empirical work using single change-point model includes [129], in which
the authors present evidence that the volatility of US economic activity abruptly fell in early 1984. In the other approach, one estimates a time varying parameter models where the parameters are allowed to change at any time, which can be interpreted as having \( n - 1 \) possible breaks in a sample of size \( n \). In this framework, [42] modeled inflation dynamics in the US as a continuous variable evolving over time.

A tractable mathematical model of structural changes and discrete market regimes is the Markov regime switching autoregressive process introduced by [90], and subsequently considered by [1]. Given that financial time series appear interdependent, in terms of both their level and volatility, see, for example, [88], a vector joint regime switching process would seem to be an attractive description of the data. In estimation of regime switching models, Bayesian methods have proved popular since they incorporate flexible relations between parameters in different regime and are computationally tractable. Suppose there are \( M \) different regimes in a model, information about coefficients in the \( j^{\text{th}} \) regime (or the duration of the \( j^{\text{th}} \) regime) can be related to information in other regimes by using hierarchical priors. Such approach is particularly useful for forecasting in the presence of multiple structural breaks since it allows for the possibility of out-of-sample breaks. For a readable account on the subject of modeling time series with changes in regime, see Chapter 22 of [91].

Here we are interested in parameter changes that lead to changes in unconditional variance of conditionally heteroskedastic time series. Chapter 3 of the dissertation focuses on a testing procedure that is developed to distinguish between long memory and volatility changes for financial returns data.

1.5 Sequential Testing

Change-point detection procedures fall into two categories: off-line, retrospective or \textit{a posteriori} tests and on-line, sequential or \textit{a priori} tests. The former, as it is put in [39], is a "one shot" test: given a historical dataset of fixed size, the test attempts to detect a structural break within the dataset. Whereas an on-line test monitors continuously, as new
data arrive, to decide if the current model is still adequate.

Sequential analysis was born in response to demands for more efficient testing of antiaircraft gunnery during World War II, culminating in Wald’s development of the sequential probability ratio test (SPRT) in 1943 (cf. [165]). Wald’s SPRT is closely analogous to the classical Neyman-Pearson fixed sample size test. To focus attention, we restrict the discussion to simple hypothesis test. Let $X_1, X_2, \ldots$ be i.i.d. random variables with a common distribution $P$. The null hypothesis is $H_0 : P = P_0$ and the alternative is $H_a : P = P_1$. The SPRT stops sampling at stage

$$N = \inf\{n \geq 1 : R_n \geq A \text{ or } R_n \leq B\},$$

where $A > 1 > B > 0$ are stopping boundaries and where $R_n = \prod_{i=1}^{n} \left( f_1(X_i) / f_0(X_i) \right)$ is the likelihood ratio, $f_i$ being the density of $P_i$, $i = 0, 1$. When stopping happens, a decision regarding acceptance or rejection of $H_0$ has to be made; when $R_n \geq A$ the null is rejected and when $R_n \leq B$ we fail to reject the null. The two constants $A$ and $B$ are chosen such that $\alpha = P_0\{R_N \geq A\}$ and $\beta = P_1\{R_N \leq B\}$, with $\alpha$ and $\beta$ denoting the probability of type I and type II error, respectively. The optimality of this simple test was proved by [164], in the sense that the SPRT minimizes both $E_0(N)$ and $E_1(N)$ among all tests whose stopping time $N$ has a finite expectation under both the null and alternative, and whose error probabilities satisfy

$$P_0\{\text{Reject } H_0\} \leq \alpha \quad \text{and} \quad P_1\{\text{Reject } H_1\} \leq \beta.$$ 

As in SPRT, the Neyman-Pearson test also involves the likelihood ratios $R_n$ and is a solution to a optimization problem under the constraint $P_0\{\text{Reject } H_0\} \leq \alpha$. There are however two main differences between the two tests. The first one is obvious; SPRT is an online, sequential test, whereas Neyman-Pearson is an off-line, retrospective test. The second difference is that the Neyman-Pearson optimization criterion is to maximize the power $P_1\{\text{Reject } H_0\}$ for a given sample size $n$ and type I error bound $\alpha$, the Wald-Wolfowitz criterion is to minimize both $E_0(N)$ and $E_1(N)$ under the type I and type II error constraints (1.5.27).
We thus can see that in Wald's theory of sequential testing collecting data is often assumed costly, a decision whether the null hypothesis or the alternative is true needs to be made reasonably fast. However, the sequential analysis of economic and financial data is somewhat different from engineering applications. In such applications, data arrive steadily and sampling is costless under the no change null hypothesis and no action is needed if the coming process is still "in control," that is, there is no change in the parameters of the data generating process. Since false alarm is inevitable due to chance, the probability of stopping under the no change null hypothesis however should be less than a given level $0 < \alpha < 1$. On the other hand, it is desirable to stop eventually with probability one if a change uncured. A commonly used performance measure of sequential procedures is average run length (ARL). The ARL is defined as the expected stopping time. We are usually more concerned with the ARL under the alternative that there is a change, and want the ARL to be small.

In Chapter 4 of the dissertation, we are interested in sequential detecting procedures for change-points in volatility of financial returns. Changes in stock market volatility have important implications for investment allocation and equilibrium asset returns. Prominent contributions include [82] on dynamic portfolio management, [68] and [166] on market efficiency, [69] and [167] on equilibrium stock and bond returns, [124] and [88] on business cycle dynamics.

1.6 Outline of the Dissertation

To conclude, we give a short description of the following three chapters:

**Chapter 2.** We propose and examine the finite sample performance of three resampling methods of constructing confidence intervals for autocorrelations of squared GARCH processes.

**Chapter 3.** We consider a testing procedure designed to distinguish between long memory and volatility changes in financial log-returns data. We propose several implementations of this procedure that utilize different long-run variance estimates. We assess the
finite sample performance of these implementations by means of a simulation study.

Chapter 4. We propose four on-line monitoring procedures for the constancy of unconditional variance in conditionally heteroskedastic time series. Asymptotic distribution of the detectors are derived. We assess empirical size and power of the tests and include practical guidelines.

Chapter 5. We conclude and discuss future research.
CHAPTER 2
CONFIDENCE INTERVALS FOR THE AUTOCORRELATION OF
SQUARED GARCH SEQUENCES

2.1 Introduction

This chapter is concerned with assessing finite sample performance of several methods of constructing confidence intervals for autocorrelations of squared returns on speculative assets. Although the returns themselves are essentially uncorrelated, and most econometric and financial models explicitly imply that they are so, their squares exhibit a rich dependence structure. The autocorrelation function measures the strength of linear dependence and is a standard tool for assessing the impact of past information on the present.

We compare the performance of the various methods by means of their empirical coverage probability (ECP). Suppose we have a method of constructing, say, a 95% confidence interval \((\hat{r}_n, \hat{u}_n)\) from an observed realization \(X_1, X_2, \ldots, X_n\). We simulate a large number \(R\) of realizations from a specific GARCH type model from which we construct \(R\) confidence intervals \((\hat{r}_{n}^{(r)}, \hat{u}_{n}^{(r)})\), \(r = 1, 2, \ldots, R\). The percentage of these confidence intervals that contain the population autocorrelation is the ECP, which we want to be as close as possible to the nominal coverage probability of 95%.

Our objective is to provide answers to the following questions: 1) Does any method have better ECP than the others? 2) If not, what is the range of optimal applicability of each method? 3) Is it better to use equal-tailed or symmetric confidence intervals (see Section 2.2.1)? 4) How does the coverage depend on the value of the fourth moment existence parameter \(\gamma_4\)? 5) For a given series length \(n\), how should one choose the block length \(b\) for the block bootstrap and subsampling? 6) For what lengths \(n\) do these methods yield useful confidence intervals?

The ultimate goal is to recommend a practical procedure for finding confidence intervals for squared autocorrelations which assumes minimal prior knowledge of the stochastic mechanism generating the returns.

In Section 2.2, we describe the three methods. Section 2.3 introduces the various univari-
ate GARCH models we use for the comparison. The results of our simulations are presented in Section 2.4 with general conclusions summarized in Section 2.4.4. Section 2.5 considers a bivariate GARCH model, specifically the constant conditional correlation GARCH model (CCC-GARCH), for which we are interested in cross-correlations at lags 0 and 1. We report results of simulation study based on residual bootstrap and subsampling in Section 2.6.

For ease of reference, recall in Definition 1.2.4 that the sample autocovariances of the squared returns are

\[
\hat{\gamma}_{n,X^2}(h) = \frac{1}{n} \sum_{t=1}^{n-h} \left( X_t^2 - \frac{1}{n-h} \sum_{t=1}^{n-h} X_t^2 \right) \left( X_{t+h}^2 - \frac{1}{n-h} \sum_{t=h+1}^{n} X_t^2 \right)
\]

whereas the population autocovariances are

\[
\gamma_{X^2}(h) = E [(X_0^2 - EX_0^2)(X_h^2 - EX_h^2)].
\]

The corresponding autocorrelations are

\[
\hat{\rho}_{n,X^2}(h) = \frac{\hat{\gamma}_{n,X^2}(h)}{\hat{\gamma}_{n,X^2}(0)}, \quad \rho_{X^2}(h) = \frac{\gamma_{X^2}(h)}{\gamma_{X^2}(0)}.
\]

2.2 Confidence Intervals for Autocorrelations of Squared Observations

In this section we describe in detail the three methods of constructing confidence intervals that we wish to compare. For concreteness, we focus on lag 1 autocorrelations, but the methods apply with minimal modifications (needed only for block bootstrap) to any lag.

2.2.1 Residual bootstrap

To illustrate the idea and focus attention, we describe the method of residual bootstrap for the ARCH(1) model given by

\[
X_t = \sigma_t Z_t, \quad \sigma_t^2 = \omega + \alpha X_{t-1}^2.
\]

As we describe in Section 2.3, the method can be readily extended to any parametric model defined by GARCH type equations by computing the residuals \( \hat{Z}_t = [\hat{h}_t]^{-1} X_t \), see Eq. (2.3.8). Since the conditional volatility \( h_t \) is a function of the model parameters, past
observations and past innovations, \( \hat{h}_t \) can be computed recursively once parameter estimates are available.

Assuming then that the observations \( X_1, X_2, \ldots, X_n \) follow the ARCH(1) model (2.2.4), we proceed as follows:

1) Find estimates \( \hat{\omega} \) and \( \hat{\alpha} \) and form the residuals

\[
\hat{Z}_t = [\hat{\omega} + \hat{\alpha} X_{t-1}^2]^{-1/2} X_t
\]

with suitably chosen \( X_0 \), e.g. the average of all \( X_t \).

2) Form \( B \) bootstrap realizations

\[
X_t^2(b) = [\hat{\omega} + \hat{\alpha} X_{t-1}^2(b)] \hat{Z}_t^2(b), \quad t = 1, 2, \ldots, n,
\]

where \( \hat{Z}_1^2(b), \ldots, \hat{Z}_n^2(b), \ b = 1, 2, \ldots, B \) are the \( B \) bootstrap samples selected with replacement from the squared residuals \( \hat{Z}_1^2, \ldots, \hat{Z}_n^2 \).

3) Calculate the bootstrap autocorrelations \( \hat{\rho}_{n,X^2(1)}(b), \ b = 1, 2, \ldots, B \) and use their empirical quantiles to find a confidence interval for \( \hat{\rho}_{n,X^2(1)} \).

In step 1), we use quasi maximum likelihood estimators (QMLE’s) of model parameters which maximize the likelihood function computed under the assumption that the innovations \( Z_t \) are standard normal.

We now expand on step 3). Denote by \( F^*_\rho(1) \) the EDF (empirical distribution function) of the \( \hat{\rho}_{n,X^2(1)}(b), \ b = 1, 2, \ldots, B \). The \((\alpha/2)\)th and \((1-\alpha/2)\)th quantiles of \( F^*_\rho(1) \) will yield an equal-tailed \((1-\alpha)\) level confidence interval. To construct a symmetric confidence interval centered at \( \hat{\rho}_{n,X^2(1)} \), we need the empirical distribution \( F^*_\rho(1,|.|) \) of the \( B \) values \(|\rho_{n,X^2(1)} - \hat{\rho}_{n,X^2(1)}|\). Denote by \( q_{|.|}(1-\alpha) \) the \((1-\alpha)\) quantile of \( F^*_\rho(1,|.|) \). Then the symmetric confidence interval is

\[
(\hat{\rho}_{n,X^2(1)} - q_{|.|}(1-\alpha), \ \hat{\rho}_{n,X^2(1)} + q_{|.|}(1-\alpha)).
\]

A usual criticism of methods based on a parametric model is that misspecification can lead to large biases. In many applications however these biases have only negligible impact on a statistical procedure of interest. In our setting, it may well be the case that the residual
bootstrap confidence intervals based on a misspecified model can produce good coverage probabilities. This point is taken up again after empirical evidence has been presented.

We conclude this section by noting that since the innovations $Z_t$ are assumed to have zero mean and unit variance, the normalized residuals $\hat{N}_t = \left( \hat{Z}_t - \bar{Z} \right) / \sigma_Z$, where $\bar{Z} = \frac{1}{n} \sum_{t=1}^{n} \hat{Z}_t$, $\sigma_Z^2 = \frac{1}{n} \sum_{t=1}^{n} \left( \hat{Z}_t - \bar{Z} \right)^2$, are often used to form bootstrap observations, see step 2) above. In our context however replacing $\hat{Z}_t$ by $\hat{N}_t$ did not lead to any improvement in empirical coverages.

2.2.2 Block bootstrap

In this section we describe how the popular block-bootstrap of [117] can be used to construct confidence intervals for autocorrelations. This method does not require a model specification, but it relies on a choice of the block size $b$ which is often a difficult task. A good account of block bootstrap is given in [31].

Focusing again on lag one sample autocorrelation of the squared observations, we proceed follows:

Having observed the sample $X_1, \ldots, X_n$ form the vectors

$$Y_2 = \left[ X_1^2, X_2^2 \right]', \quad Y_3 = \left[ X_2^2, X_3^2 \right]', \ldots, \quad Y_n = \left[ X_{n-1}^2, X_n^2 \right]' .$$

There are $n - 1$ such vectors. Now choose a block length $b$ and compute the number of blocks $k = \left[ (n - 1)/b \right] + 1$ (if $(n - 1)/b$ is an integer we take $k = (n - 1)/b$). Choose $k$ blocks with replacement to obtain $kb$ vectors. Choosing the $k$ blocks corresponds to generating $k$ observations from the uniform distribution on $\{2, 3, \ldots, n-b+1\}$. Denote these observations $j_1, j_2, \ldots, j_k$. We thus obtain the $kb$ vectors

$$Y_{j_1}, Y_{j_1+1}, \ldots, Y_{j_1+b-1}, \ldots, Y_{j_k}, Y_{j_k+1}, \ldots, Y_{j_k+b-1} .$$

If $(n - 1)/b$ is not an integer, remove the last few vectors to have exactly $n - 1$ vectors. This gives us the bootstrap vector process

$$Y'_2 = \left[ X_1'^2, X_2'^2 \right]', \quad Y'_3 = \left[ X_2'^2, X_3'^2 \right]', \ldots, \quad Y'_n = \left[ X_{n-1}'^2, X_n'^2 \right]' .$$
The bootstrap sample autocovariances are computed according to (2.1.1) with the $X_t$ replaced by the $X_t^*$ defined above. The empirical distribution of $\hat{\rho}_{n,X^2}(1)$ is then an approximation to the distribution of $\hat{\rho}_{n,X^2}(1)$. As described in Section 2.2.1, the quantiles of the empirical distribution of $|\hat{\rho}_{n,X^2}(1) - \hat{\rho}_{n,X^2}(1)|$ can be used to construct symmetric confidence intervals.

Pararoditis and Politis proposed in [139] a new version of block bootstrap which attenuates the effects of joining the blocks together by applying kernel type weights which smoothly approach zero at the end points of the blocks. This method is superior to block bootstrap when the goal is to estimate the variance of the mean of a stationary series, but the simulations in [140] show that it does not posses a clear advantage over the standard block bootstrap when constructing confidence intervals. For this reason we do not investigate this method here.

### 2.2.3 Subsampling

The subsampling methodology is described in detail in [148]. Subsampling confidence intervals for autocorrelations of linear time series models like ARMA were investigated in [149]. In this section we adapt their methodology to the squares of GARCH processes.

To simplify the notation, denote

\begin{equation}
U_t = X_t^2 - \frac{1}{n} \sum_{j=1}^{n} X_j^2
\end{equation}

and suppress the subscript $X^2$ in the following formulas in which use definitions (2.1.1) and (2.1.3). Set

\begin{equation}
s_n^2(h) = \frac{1}{n} \sum_{j=1}^{n-h} (U_{j+h} - \hat{\rho}_n(h)U_j)^2, \quad \hat{\sigma}_n^2(h) = \frac{s_n^2(h)}{\sum_{j=h}^{n} U_j^2}
\end{equation}

and consider the studentized statistic

\begin{equation}
\hat{\xi}_n = \frac{\hat{\rho}_n(h) - \rho_n(h)}{\hat{\sigma}_n(h)}.
\end{equation}

To construct equal-tailed and symmetric confidence intervals, we need to know the sampling distribution of $\hat{\xi}_n$ and $|\hat{\xi}_n|$, respectively. We use subsampling to approximate these distributions: Consider an integer $b < n$ and the $n - b + 1$ blocks of data $X_t, \ldots, X_{t+b-1}$, $t =$
From each of these blocks compute \( \hat{p}_{b,t}(h) \) and \( \hat{\sigma}_{b,t}(h) \) according to (2.1.1), (2.1.3) and (2.2.6), respectively, but replacing the original data \( X_1, \ldots, X_n \) by \( X_t, \ldots, X_{t+b-1} \). Next, compute the subsampling counterpart of the studentized statistic (2.2.7)

\[
\hat{\xi}_{b,t}(h) = \frac{\hat{p}_{b,t}(h) - \hat{p}_n(h)}{\hat{\sigma}_{b,t}(h)}
\]

and construct the empirical distribution functions

\[
L_b(x) = \frac{1}{n-b+1} \sum_{t=1}^{n-b+1} 1\{\hat{\xi}_{b,t}(h) \leq x\},
\]

\[
L_{b,|1|}(x) = \frac{1}{n-b+1} \sum_{t=1}^{n-b+1} 1\{|\hat{\xi}_{b,t}(h)| \leq x\}.
\]

The empirical quantiles of \( L_b \) and \( L_{b,|1|} \) allow us to construct, respectively, equal-tailed and symmetric confidence intervals. For example, denoting by \( q_{b,|1|}(1-\alpha) \) the \((1-\alpha)\)th quantile of \( L_{b,|1|} \), a subsampling symmetric \( 1-\alpha \) level confidence interval for \( \rho_n(h) \) is

\[
(\hat{\rho}_n(h) - \hat{\sigma}_n(h)q_{b,|1|}(1-\alpha), \hat{\rho}_n(h) + \hat{\sigma}_n(h)q_{b,|1|}(1-\alpha)).
\]

### 2.3 GARCH Models

In this section we describe the GARCH models which we used to assess the performance of confidence interval building procedures. Our goal was to consider a variety of dependence structures, focusing, however, on those models which have been extensively used and whose properties have been thoroughly investigated. To present the various models in a uniform framework, it is convenient to concentrate on a general class of GARCH\((1,1)\) type models studied by [94]. The observations \( X_t \) are thus assumed to satisfy

(2.3.8)

\[
X_t = Z_t h_t,
\]

where \( Z_t \) is a sequence of independent identically distributed random variables with zero mean and

(2.3.9)

\[
h_t^k = g(Z_{t-1}) + c(Z_{t-1})h_{t-1}^k
\]

for \( k = 1 \) or \( k = 2 \). We consider only specifications (2.3.9) with \( k = 2 \) in which the function \( g(\cdot) \) is a constant: \( g(\cdot) = \omega \) and the \( Z_t \) are standard normal. Denoting \( c_t = c(Z_t) \)
and \( \gamma_{ci} = Ec^i_c \), [94] proved that under the above assumptions a sufficient and necessary condition for the existence of the \( km \)th unconditional moment of \( X_t \) is \( \gamma_{cm} = Ec^m_c < 1 \).

Since we consider \( k = 2 \), the fourth unconditional moment of \( X_t \) exists if and only if

\[
(2.3.10) \quad \gamma_{c2} = Ec^2_c \in [0, 1).
\]

We considered the following three specific models:

1) The standard GARCH(1, 1) model ([20]; [161], pp. 78-79) for which \( c_{t-1} = \beta + \alpha Z^2_{t-1} \), that is

\[
(2.3.11) \quad h^2_t = \omega + \alpha X^2_{t-1} + \beta h^2_{t-1}.
\]

For this model \( \gamma_{c2} = E (\beta + \alpha Z^2_t)^2 \).

2) The GJR-GARCH(1, 1) model ([75]) with

\[
(2.3.12) \quad h^2_t = \omega + (\alpha + \phi I(Z_{t-1})) X^2_{t-1} + \beta h^2_{t-1},
\]

where \( I(Z_{t-1}) = 1 \) if \( Z_{t-1} < 0 \), and \( I(Z_{t-1}) = 0 \) otherwise. For this model

\[
\gamma_{c2} = E [\beta + (\alpha + \phi I(Z_t))Z^2_t]^2.
\]

3) The nonlinear GARCH(1, 1) model (NLGARCH(1,1,2); see [61]) with

\[
(2.3.13) \quad h^2_t = \omega + \alpha (1 - 2\eta \text{sign}(Z_{t-1}) + \eta^2) X^2_{t-1} + \beta h^2_{t-1},
\]

For this model

\[
\gamma_{c2} = E [\beta + \alpha(1 - 2\eta \text{sign}(Z_t) + \eta^2)Z^2_t]^2.
\]

Under the assumption that the innovations follow the standard normal distribution, such that \( E(Z^2_t) = 1 \), and \( E(Z^4_t) = 3 \), Eq. (21) of [94] is reduced to

\[
(2.3.14) \quad \rho_{X^2}(1) = \frac{\gamma_c (1 - \gamma^2_c) - \gamma_c (1 - \gamma_{c2})}{3 (1 - \gamma^2_c) - (1 - \gamma_{c2})}.
\]
Table 2.1. Five standard GARCH(1, 1) models with given parameters, theoretical values of $\gamma_{c2}$ and $\rho_{X^2(1)}$.

<table>
<thead>
<tr>
<th>Model</th>
<th>$\omega$</th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$\gamma_{c2}$</th>
<th>$\rho_{X^2(1)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0.13</td>
<td>0.13</td>
<td>0.101</td>
<td>0.132</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0.2</td>
<td>0.28</td>
<td>0.311</td>
<td>0.214</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0.28</td>
<td>0.3</td>
<td>0.493</td>
<td>0.312</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>0.33</td>
<td>0.37</td>
<td>0.708</td>
<td>0.395</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>0.38</td>
<td>0.4</td>
<td>0.897</td>
<td>0.488</td>
</tr>
</tbody>
</table>

where $\gamma_c = E(Z_t^2 c_t)$, $\gamma_c = E(c_t)$, and recall in Eq. (2.3.10) that $\gamma_{c2} = E(c_t^2)$. Elementary but lengthy calculations yield the following expressions for $\gamma_{c2}$ and $\rho_{X^2(1)}$ in terms of the parameters of the three models. See Appendix A for the derivations of Eqs. (2.3.15)-(2.3.20).

The values of $\rho_{X^2(1)}$ are needed to assess the performance of confidence intervals. As we shall see in the sequel, this performance depends to a large degree on the value of $\gamma_{c2}$.

1) Standard GARCH:

\[
\gamma_{c2} = \beta^2 + 2\alpha\beta + 3\alpha^2 \tag{2.3.15}
\]

\[
\rho_{X^2(1)} = \frac{\alpha (1 - \alpha\beta - \beta^2)}{1 - 2\alpha\beta - \beta^2} \tag{2.3.16}
\]

2) GJR-GARCH:

\[
\gamma_{c2} = \beta^2 + 3\alpha^2 + \frac{2}{3}\phi^2 + 2\alpha\beta + \beta\phi + 3\alpha\phi \tag{2.3.17}
\]

\[
\rho_{X^2(1)} = \frac{2\alpha\beta(\alpha + \phi) + \frac{5}{4}\beta\phi^2 + 2(\alpha + \frac{1}{2}\phi) (1 - \beta^2 - 2\alpha\beta - \beta\phi + \frac{3}{4}\phi^2)}{2(1 - \beta^2 - 2\alpha\beta - \beta\phi + \frac{3}{4}\phi^2)} \tag{2.3.18}
\]

3) NLGARCH:

\[
\gamma_{c2} = \beta^2 + 3\alpha^2 + 2\alpha\beta + 18\alpha^2\eta^2 + 3\alpha^2\eta^4 + 2\alpha\beta\eta^2 \tag{2.3.19}
\]

\[
\rho_{X^2(1)} = \alpha (1 + \eta^2) + \frac{\alpha^2\beta (\eta^4 + 8\eta^2 + 1)}{1 - \beta^2 - 2\alpha\beta - 2\alpha\beta\eta^2 + 6\alpha^2\eta^2} \tag{2.3.20}
\]
Table 2.2. Five GJR-GARCH(1, 1) models with given parameters, theoretical values of \( \gamma_c^2 \) and \( \rho_{X^2(1)} \).

<table>
<thead>
<tr>
<th>Model</th>
<th>( \omega )</th>
<th>( \alpha )</th>
<th>( \beta )</th>
<th>( \phi )</th>
<th>( \gamma_c^2 )</th>
<th>( \rho_{X^2(1)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.115</td>
<td>0.153</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0.15</td>
<td>0.2</td>
<td>0.15</td>
<td>0.299</td>
<td>0.238</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0.2</td>
<td>0.25</td>
<td>0.2</td>
<td>0.513</td>
<td>0.333</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
<td>0.719</td>
<td>0.428</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>0.29</td>
<td>0.27</td>
<td>0.27</td>
<td>0.899</td>
<td>0.502</td>
</tr>
</tbody>
</table>

Table 2.3. Five NLGARCH(1,1,2) models with given parameters, theoretical values of \( \gamma_c^2 \) and \( \rho_{X^2(1)} \).

<table>
<thead>
<tr>
<th>Model</th>
<th>( \omega )</th>
<th>( \alpha )</th>
<th>( \beta )</th>
<th>( \eta )</th>
<th>( \gamma_c^2 )</th>
<th>( \rho_{X^2(1)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0.15</td>
<td>0.1</td>
<td>0.1</td>
<td>0.112</td>
<td>0.154</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0.2</td>
<td>0.22</td>
<td>0.22</td>
<td>0.296</td>
<td>0.224</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0.25</td>
<td>0.28</td>
<td>0.28</td>
<td>0.506</td>
<td>0.305</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.704</td>
<td>0.388</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>0.3</td>
<td>0.4</td>
<td>0.35</td>
<td>0.902</td>
<td>0.450</td>
</tr>
</tbody>
</table>

For each of the three models, we considered five parameter choices, which we labeled as models 1 through 5 in Tables 2.1, 2.2, and 2.3. The parameters were chosen in such a way that the models with the same index have similar values of \( \gamma_c^2 \) and \( \rho_{X^2(1)} \). The values of \( \gamma_c^2 \) and \( \rho_{X^2(1)} \) increase as the model index increases from 1 to 5. Such an arrangement will be seen to facilitate the comparisons presented in Section 2.4. Note that neither \( \gamma_c^2 \) nor \( \rho_{X^2(1)} \) depends on the parameter \( \omega \) in Eqs. (2.3.15)-(2.3.20), so we fixed its value to be 1.

2.4 Simulation Results

The objective of the simulations presented in this section was to investigate the performance of the three methods described in Section 2.2. This is done by comparing the empirical coverage probabilities (ECP's) for the fifteen data generating processes (DGP's) introduced in Section 2.3. We generated one thousand replications of each DGP and considered realizations of length \( n = 100, 250, 500, 1000 \). We focused on the most commonly used confidence level of 95%. The standard errors in all tables are about 0.5% and are always smaller than 1%.
Table 2.4. Empirical coverage probabilities of equal-tailed confidence intervals constructed using residual bootstrap.

<table>
<thead>
<tr>
<th>n</th>
<th>e.c.p. (%)</th>
<th>e.c.p. (%)</th>
<th>e.c.p. (%)</th>
<th>e.c.p. (%)</th>
<th>e.c.p. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>STD GARCH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>100</td>
<td>90.1</td>
<td>87.4</td>
<td>84.5</td>
<td>79.7</td>
</tr>
<tr>
<td>250</td>
<td>94.0</td>
<td>91.7</td>
<td>93.9</td>
<td>93.7</td>
<td>90.3</td>
</tr>
<tr>
<td>500</td>
<td>94.5</td>
<td>95.2</td>
<td>96.5</td>
<td>96.1</td>
<td>92.8</td>
</tr>
<tr>
<td>1000</td>
<td>95.6</td>
<td>97.0</td>
<td>97.5</td>
<td>98.9</td>
<td>98.2</td>
</tr>
<tr>
<td>GJR GARCH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>100</td>
<td>94.5</td>
<td>94.3</td>
<td>91.5</td>
<td>87.8</td>
</tr>
<tr>
<td>250</td>
<td>97.9</td>
<td>97.6</td>
<td>97.0</td>
<td>97.7</td>
<td>95.0</td>
</tr>
<tr>
<td>500</td>
<td>98.4</td>
<td>99.1</td>
<td>99.5</td>
<td>99.5</td>
<td>98.3</td>
</tr>
<tr>
<td>1000</td>
<td>92.6</td>
<td>99.6</td>
<td>99.8</td>
<td>99.9</td>
<td>99.7</td>
</tr>
<tr>
<td>NL GARCH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>58.3</td>
<td>65.0</td>
<td>54.8</td>
<td>48.8</td>
<td>36.4</td>
</tr>
<tr>
<td>250</td>
<td>77.5</td>
<td>63.5</td>
<td>57.4</td>
<td>48.6</td>
<td>37.4</td>
</tr>
<tr>
<td>500</td>
<td>81.7</td>
<td>67.0</td>
<td>52.7</td>
<td>46.2</td>
<td>36.6</td>
</tr>
<tr>
<td>1000</td>
<td>81.7</td>
<td>59.5</td>
<td>45.9</td>
<td>39.2</td>
<td>29.9</td>
</tr>
</tbody>
</table>

Below are some questions our simulations were intended to answer:

- Does any of the three methods have uniformly (i.e. for all 15 DGP’s and all 4 sample sizes) better coverage probability than the remaining two methods. If not, what is the range of optimal applicability of each method?

- Is it more advisable to use equal-tailed or symmetric confidence intervals?

- How does the coverage depend on the value of $\gamma_{c2}$ (see eq. (2.3.10))?

- For a given series length $n$, how should one choose the block length $b$ for block bootstrap or subsampling?

- For what lengths $n$ do these methods yield useful confidence intervals?

The ultimate goal is to recommend a practical procedure for finding confidence intervals for squared autocorrelations which assumes minimal prior knowledge of the stochastic mechanism generating the observations.

We first examine in Sections 2.4.1–2.4.3 each method separately which allows us to focus on specific issues pertaining to each method. We then combine and summarize our findings in Section 2.4.4.
Table 2.5. Empirical coverage probabilities of symmetric confidence intervals constructed using residual bootstrap.

<table>
<thead>
<tr>
<th>n</th>
<th>e.c.p. (%)</th>
<th>e.c.p. (%)</th>
<th>e.c.p. (%)</th>
<th>e.c.p. (%)</th>
<th>e.c.p. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>99.6</td>
<td>85.3</td>
<td>86.0</td>
<td>80.4</td>
<td>77.4</td>
</tr>
<tr>
<td>250</td>
<td>92.9</td>
<td>91.3</td>
<td>92.1</td>
<td>89.4</td>
<td>84.4</td>
</tr>
<tr>
<td>500</td>
<td>93.4</td>
<td>93.4</td>
<td>94.1</td>
<td>93.7</td>
<td>92.7</td>
</tr>
<tr>
<td>1000</td>
<td>95.1</td>
<td>96.8</td>
<td>97.6</td>
<td>97.6</td>
<td>94.4</td>
</tr>
<tr>
<td>100</td>
<td>97.7</td>
<td>94.8</td>
<td>92.0</td>
<td>89.5</td>
<td>81.5</td>
</tr>
<tr>
<td>250</td>
<td>96.2</td>
<td>96.6</td>
<td>97.0</td>
<td>96.4</td>
<td>92.3</td>
</tr>
<tr>
<td>500</td>
<td>98.3</td>
<td>99.2</td>
<td>98.9</td>
<td>99.1</td>
<td>96.5</td>
</tr>
<tr>
<td>1000</td>
<td>99.0</td>
<td>99.4</td>
<td>99.6</td>
<td>99.8</td>
<td>98.8</td>
</tr>
<tr>
<td>100</td>
<td>95.5</td>
<td>83.8</td>
<td>79.8</td>
<td>74.7</td>
<td>66.0</td>
</tr>
<tr>
<td>250</td>
<td>91.7</td>
<td>87.3</td>
<td>84.3</td>
<td>81.0</td>
<td>73.6</td>
</tr>
<tr>
<td>500</td>
<td>91.7</td>
<td>93.1</td>
<td>88.5</td>
<td>82.1</td>
<td>77.3</td>
</tr>
<tr>
<td>1000</td>
<td>96.4</td>
<td>93.3</td>
<td>92.9</td>
<td>87.0</td>
<td>81.0</td>
</tr>
</tbody>
</table>

2.4.1 Residual bootstrap

Table 2.4 presents the empirical coverage probabilities of the equal-tailed confidence interval for the three GARCH models, whereas Table 2.5 gives the results of the symmetric confidence interval.

Several conclusions can be readily drawn from the examination of Tables 2.4 and 2.5.

1. Equal-tailed and symmetric confidence intervals perform equally well for the standard GARCH and GJR-GARCH. However, for the NLGARCH, the symmetric interval is better than the equal-tailed. It is thus seen that the symmetric confidence interval is preferred over the equal-tailed. See Figure B.1 for a graphical comparison between the two confidence intervals across the three GARCH processes for different values of $n$.

2. The ECP decreases as the value of $\gamma_{c2}$ approaches 1. Recall that $\gamma_{c2} < 1$ is required for the population autocovariances to exist. When $\gamma_{c2} \approx 0.9$, at least 250 observations are needed to ensure reasonable ECP for the standard GARCH and the GJR-GARCH. For the NL GARCH, even series length of 1000, does not produce satisfactory results.

3. For the standard GARCH and the GJR-GARCH increasing the sample size from 500
Figure 2.1. Comparison of ECP's for *symmetric residual bootstrap* confidence intervals based on standard GARCH and a correct specification. The nominal coverage of 95% is marked by solid horizontal line.

to 1000 does not improve the ECP. For the NL GARCH a sample size of 1000 observations is needed, except when $\gamma_{\text{c2}} \leq 0.3$.

The somewhat worse performance of the residual bootstrap method for the GJR-GARCH which becomes markedly worse for the NL GARCH can be attributed to identification problems, which are particularly acute for the NL GARCH. We noticed that for the latter model biases of parameter estimates are very large when $\eta$ in Eq. (2.3.13) is large. As formula (2.3.19) shows, this corresponds to large values of $\gamma_{\text{c2}}$. On the other hand, for the standard GARCH, while they still do exist, the identification problems are much less severe. It is therefore worthwhile to investigate if estimating the standard GARCH model on all three
DGP's might lead to improvements in ECP's. Figure 2.1 shows that this in fact the case for symmetric confidence intervals and series lengths of 500 and 1000. The results for other series lengths look very much the same and are therefore not presented.

In conclusion, the residual bootstrap method works best if symmetric confidence intervals are used and the standard GARCH model is estimated. Thus, in our context, misspecifying a model improves the performance of the procedure.

2.4.2 Block bootstrap

The implementation of this method requires a choice of the block length $b$, which is an integer less than $n$. We therefore have a multitude of cases to explore: 15 models, two types of confidence intervals (equal-tailed and symmetric), two sample sizes, 500 and 1000, and four choices of $b$. We thus obtained 240 ECP's. Table 2.6 reports the ECP's of equal-tailed and symmetric confidence intervals for standard GARCH models, with $n = 500$. See Tables B.1 and B.2 for complete results.

We conclude, see Figure B.2 for a graphical illustration, that

1. The empirical coverage probabilities are generally too low for all choices of $n$ and $b$ and are in the range of 80% to 90% for $\gamma_{c2} \leq 0.3$ and go down to slightly above 50% for $\gamma_{c2} \approx 0.9$.

2. Irrespective of the value of $\gamma_{c2}$, choosing smaller $b$ gives higher coverage. However, extremely small $b$, like 1 or 2, do not work well. We recommend to use $b = 3$ or $b = 5$. The dependence on $b$ is however not substantial, which is very desirable, as in many other applications choosing optimal $b$ is very difficult.

3. There is not much difference of ECPs between equal-tailed and symmetric confidence intervals.

4. The block bootstrap confidence intervals are generally too short and given that the QML estimates underestimate the true value of the autocorrelation, they are shifted too much to the left which causes the undercoverage.
Table 2.6. Empirical coverage probabilities of *equal-tailed* and *symmetric* confidence intervals constructed using *block bootstrap* for *standard GARCH* models.

<table>
<thead>
<tr>
<th>Model</th>
<th>C.I.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>e.c.p. (%)</td>
<td>e.c.p. (%)</td>
<td>e.c.p. (%)</td>
<td>e.c.p. (%)</td>
<td>e.c.p. (%)</td>
</tr>
<tr>
<td>equal-tailed</td>
<td>500</td>
<td>3</td>
<td>89.0</td>
<td>85.4</td>
<td>78.1</td>
<td>72.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>89.9</td>
<td>83.6</td>
<td>78.0</td>
<td>67.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>87.9</td>
<td>82.9</td>
<td>72.8</td>
<td>62.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>83.3</td>
<td>78.0</td>
<td>68.8</td>
<td>57.0</td>
</tr>
<tr>
<td>symmetric</td>
<td>500</td>
<td>3</td>
<td>87.9</td>
<td>84.7</td>
<td>78.7</td>
<td>70.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>88.8</td>
<td>82.8</td>
<td>76.2</td>
<td>68.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>87.7</td>
<td>84.3</td>
<td>74.2</td>
<td>67.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>85.4</td>
<td>82.4</td>
<td>74.1</td>
<td>63.9</td>
</tr>
</tbody>
</table>

### 2.4.3 Subsampling

This method also requires the block length $b$, an integer less than $n$. We generate sequences of length $n = 500$ and 1000 and examined five choices of $b$, in specific, it ranges from 1% to 20% of $n$ for equal-tailed confidence interval and is less than 10% of $n$ for symmetric confidence interval. We use different values of $b$ for the two kinds of confidence interval. Such selections are based on our preliminary investigations, so as to present as good empirical results as each confidence interval can produce and to reveal the pattern of the effects of block size on the coverage probability. Table 2.7 presents the ECP's of the two types of confidence interval for the standard GARCH, with $n = 500$. See Tables B.3 and B.4 for complete results. We observe, see Figure B.3 for a graphical presentation, that

1. It is noted again that when $\gamma_{c2}$ is getting greater and closer to 1 the empirical coverage probabilities descend in monotone.

2. The ECP's are generally too low for equal-tailed confidence intervals. The highest coverage it can produce is less than 85% for models with $\gamma_{c2} < 0.2$.

3. Symmetric confidence intervals have a much better performance than the equal-tailed. By choosing very short $b$'s, such as 3, we can obtain ECP's that are quite close to 95% for models with $\gamma_{c2} < 0.6$ and fair coverages for models with greater values of $\gamma_{c2}$. 
Table 2.7. Empirical coverage probabilities of equal-tailed and symmetric confidence interval constructed using subsampling for standard GARCH models.

<table>
<thead>
<tr>
<th>Model</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>C.I.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>equal-tailed</td>
<td>n=500</td>
<td>b=5</td>
<td>e.c.p. (%)</td>
<td>e.c.p. (%)</td>
<td>e.c.p. (%)</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td></td>
<td>54.0</td>
<td>39.2</td>
<td>24.8</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>77.5</td>
<td>60.6</td>
<td>41.8</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td></td>
<td>80.7</td>
<td>72.6</td>
<td>59.8</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>77.5</td>
<td>70.3</td>
<td>58.5</td>
</tr>
<tr>
<td>symmetric</td>
<td>n=500</td>
<td>b=3</td>
<td>e.c.p. (%)</td>
<td>e.c.p. (%)</td>
<td>e.c.p. (%)</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td></td>
<td>97.8</td>
<td>96.7</td>
<td>92.9</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td></td>
<td>93.6</td>
<td>88.8</td>
<td>83.1</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td></td>
<td>88.8</td>
<td>85.1</td>
<td>79.8</td>
</tr>
</tbody>
</table>

2.4.4 Summary

We summarize as follows:

1. Generally, as value of $\gamma_{c2}$ gets greater and closer to the boundary value 1 empirical coverage probabilities tend to decrease.

2. The more observations we generate the greater empirical coverage probabilities we can obtain.

3. As far as the choice of block size $b$ is concerned, we suggest using very short $b$'s, less than 5, for sequences of any length when both block bootstrap and subsampling are implemented.

4. Overall, symmetric confidence intervals are preferred over equal-tailed confidence intervals. We conclude that the symmetric confidence interval yielded by the method of residual bootstrap works best among the all that are investigated. Its results are shown in Table 2.5.

Focusing on symmetric confidence intervals, we carry out a comparison of empirical coverage probabilities produced by the three methods for two $n$'s, 500 and 1000. In light of the suggestion in Point 3 above, we select the results corresponding to $b=3$ when $n=500$ and $b=5$ when $n=1000$ for block bootstrap, while for subsampling $b$ equals to 3 for both $n$'s. Figure 2.2 presents the comparison; it also gives graphical support to our findings.
Figure 2.2. Of symmetric confidence intervals, performance comparison among the three methods. The lengths of sequences are \( n = 500 \) and \( n = 1000 \) in the two plots, respectively. The nominal coverage 95% is marked by solid horizontal line.

2.5 Confidence Intervals for Cross-Correlations

In this section we consider confidence intervals for cross-correlations in a multivariate GARCH model and assess their relative performance. We focus on a simple, but commonly used, constant conditional correlation model (CCC-GARCH) introduced by [21]. We first recall some facts about the CCC-GARCH model. We then report results of simulation study based on residual bootstrap and subsampling in Section 2.6.

We follow the exposition in Chapter 6 of [76] and in [95]. In the bivariate CCC-GARCH
model the observations \( X_t = [X_t(1), X_t(2)]^T \) satisfy

\[
(2.5.21) \quad E[X_t | \mathcal{F}_{t-1}] = 0,
\]

\[
(2.5.22) \quad \text{Var}[X_t | \mathcal{F}_{t-1}] = \begin{bmatrix}
\sigma_t^2(1) & \rho \sigma_t(1) \sigma_t(2) \\
\rho \sigma_t(1) \sigma_t(2) & \sigma_t^2(2)
\end{bmatrix},
\]

where \( \mathcal{F}_{t-1} \) is the \( \sigma \)-algebra generated by \( X_{t-1}, X_{t-2}, \ldots \) and

\[
(2.5.23) \quad \sigma_t^2(i) = \omega_i + a_i X_{t-1}^2(i) + b_i \sigma_{t-1}^2(i), \quad i = 1, 2.
\]

A time series satisfying (2.5.21)-(2.5.23) can be generated as

\[
(2.5.24) \quad X_t(i) = \sigma_t(i) Z_t(i), \quad i = 1, 2,
\]

where the random vectors \( [Z_t(1), Z_t(2)]^T \) are independent with mean zero and satisfy

\[
(2.5.25) \quad E[Z_t(i)] = 1, \quad i = 1, 2, \quad E[Z_t(1)Z_t(2)] = \rho.
\]

A random vector \( [Z_t(1), Z_t(2)]^T \) satisfying (2.5.25) is a linear combination of two independent standard normal random variables \( \varepsilon_t(1) \) and \( \varepsilon_t(2) \):

\[
(2.5.26) \quad Z_t(1) = a \varepsilon_t(1) + b \varepsilon_t(2), \quad Z_t(2) = b \varepsilon_t(1) + a \varepsilon_t(2),
\]

where

\[
(2.5.27) \quad 2a = \sqrt{1 + \rho} + \sqrt{1 - \rho}, \quad 2b = \sqrt{1 + \rho} - \sqrt{1 - \rho}.
\]

In order to formulate a sufficient condition derived by [95] for the existence of unconditional fourth moments of the \( X_t(i) \), \( i = 1, 2 \), we need to introduce some additional notation. Setting \( \sigma_t^2 = [\sigma_t^2(1), \sigma_t^2(2)]^T \), we see that Eqs. (2.5.21)-(2.5.23) can be written equivalently in the matrix form as

\[
(2.5.28) \quad \sigma_t^2 = \omega + C_{t-1} \sigma_{t-1}^2,
\]

where

\[
(2.5.29) \quad \omega = \begin{bmatrix}
\omega_1 \\
\omega_2
\end{bmatrix}, \quad C_t = \begin{bmatrix}
a_1 Z_t^2(1) + b_1 & 0 \\
0 & a_2 Z_t^2(2) + b_2
\end{bmatrix}.
\]

Define

\[
(2.5.29) \quad \Gamma_C = E(C_t), \quad \Gamma_{C \otimes C} = E(C_t \otimes C_t),
\]
where $\otimes$ denotes the Kronecher product of two matrices, see [95] for the details. The fourth-order moment matrix $E\left[X_t^2X_t^{2T}\right]$ exists if

$$\lambda(\Gamma_{C\otimes C}) < 1. \tag{2.5.30}$$

where $\lambda(\Gamma)$ denotes the modulus of the largest eigenvalue of $\Gamma$. By (2.5.28),

$$\Gamma_{C\otimes C} = \text{diag}\{\gamma_1, \gamma_2, \gamma_3, \gamma_4\},$$

where

$$\gamma_1 = E\left[(a_1Z_t^2(1) + b_1)^2\right] = 3a_1^2 + 2a_1b_1 + b_1^2,$$

$$\gamma_2 = \gamma_3 = E\left[(a_1Z_t^2(1) + b_1)(a_2Z_t^2(2) + b_2)\right] = (1 + 2\rho^2)a_1a_2 + a_1b_2 + a_2b_1 + b_1b_2,$$

$$\gamma_4 = E\left[(a_2Z_t^2(2) + b_2)^2\right] = 3a_2^2 + 2a_2b_2 + b_2^2.$$

In the above computations we used the identities

$$E[Z_i^2(i)] = 1, \quad E[Z_i^4(i)] = 3, \quad E[Z_t^2(1)Z_t^2(2)] = 1 + 2\rho^2, \tag{2.5.31}$$

which hold true for normal random variables satisfying (2.5.25). The last identity in (2.5.31) can be verified using the bivariate moment generating function of $Z_t(1)$ and $Z_t(2)$, which is

$$M(t_1, t_2) = \exp\left\{t_1^2 + t_2^2 + 2\rho t_1t_2\right\}.$$ 

Therefore,

$$E[Z_t^2(1)Z_t^2(2)] = \left.\frac{\partial^4 M(t_1, t_2)}{\partial t_1^2\partial t_2^2}\right|_{t_1=t_2=0} = 1 + 2\rho^2.$$

Since the eigenvalues of the diagonal matrix $\Gamma_{C\otimes C}$ are just the four diagonal entries, condition (2.5.30) becomes

$$\max(\gamma_1, \gamma_2, \gamma_3, \gamma_4) < 1. \tag{2.5.32}$$

In the sequel, we assume that condition (2.5.32) holds.

Let $R_2(n)$ be the $n$th order autocorrelation matrix of $\{X_t^2(i)\}, \ i = 1, 2, \ n \geq 0$. The $i$th diagonal element $r_{ii}(n), i = 1, 2$ of $R_2(n)$ is the $n$th autocorrelation of the squared observations for the $i$th component of $X_t$ while the $r_{ij}(n), i, j = 1, 2, i \neq j$, the off-diagonal elements
Table 2.8. Three CCC-GARCH(1,1) models with given parameters, theoretical values of $r_0$, $r_1$, and $\lambda$.

<table>
<thead>
<tr>
<th>Model</th>
<th>$\omega_1$</th>
<th>$\omega_2$</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$b_1$</th>
<th>$b_2$</th>
<th>$\rho$</th>
<th>$r_0$</th>
<th>$r_1$</th>
<th>$\lambda$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.4</td>
<td>0.157</td>
<td>0.0159</td>
<td>0.06</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0.25</td>
<td>0.2</td>
<td>0.15</td>
<td>0.2</td>
<td>0.7</td>
<td>0.456</td>
<td>0.118</td>
<td>0.29</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0.4</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.8</td>
<td>0.444</td>
<td>0.200</td>
<td>0.81</td>
</tr>
</tbody>
</table>

of $R_2(n)$, are the cross-correlations between $X_{it}^2$ and $X_{jt-n}^2$. We consider $\text{Cor}(X_{i}^2(1), X_{j}^2(2))$ and $\text{Cor}(X_{i}^2(1), X_{j-1}^2(2))$, i.e. cross-correlations at lags 0 and 1. To lighten the notation, we denote these correlations, respectively, by $r_0$ and $r_1$. The expressions for $r_0$ and $r_1$ in terms of the model parameters are rather complicated. We refer the reader to Appendix A for the derivations.

2.6 Simulation Results

Due to the fact that both $r_0$ and $r_1$ are not dependent on $\omega_1$ or $\omega_2$, we fix their values to 1 constantly. To lighten the notation denote by $\lambda$ the modulus of the largest eigenvalue of $\Gamma_{C\&C}$ in Eq. (2.5.30). We have three specifications of CCC-GARCH(1,1) in Table 2.8, together with the model parameters. The theoretical values of $r_0$, $r_1$, and $\lambda$ are also given. We intentionally let $\lambda$ ascend from model 1 to 3.

We consider only residual bootstrap and subsampling confidence intervals as these two methods were superior to block bootstrap in the univariate case. Confidence level of 95% is used and empirical coverage probabilities (ECP’s) are based on one thousand replications.

2.6.1 Residual bootstrap for bivariate observations

The residual bootstrap method consists in finding QMLE’s of the parameters as described in Section 6.3.3. of [76] and then computing the residuals and generating bootstrap realizations based on (2.5.23) and (2.5.24). Table 2.9 presents the ECP’s of equal-tailed and symmetric confidence interval for the two cross-correlations $r_0$ and $r_1$ in the CCC-GARCH model. Realizations of length $n = 100, 250, 500, 1000$ are considered. We observe

1. Equal-tailed and symmetric confidence intervals perform equally well for the two
Table 2.9. Empirical coverage probabilities of equal-tailed and symmetric confidence interval based on the method of residual bootstrap for \( r_0 \) and \( r_1 \) in the CCC-GARCH model.

<table>
<thead>
<tr>
<th>C.I.</th>
<th>( n )</th>
<th>( r_0 )</th>
<th></th>
<th>( r_1 )</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>e.c.p. (%)</td>
<td>e.c.p. (%)</td>
<td>e.c.p. (%)</td>
<td>e.c.p. (%)</td>
<td>e.c.p. (%)</td>
</tr>
<tr>
<td>equal-tailed</td>
<td>100</td>
<td>94.0</td>
<td>95.6</td>
<td>99.7</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>250</td>
<td>93.2</td>
<td>96.8</td>
<td>99.8</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>94.8</td>
<td>97.9</td>
<td>99.8</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>94.3</td>
<td>98.7</td>
<td>99.9</td>
<td>100</td>
</tr>
<tr>
<td>symmetric</td>
<td>100</td>
<td>91.3</td>
<td>95.5</td>
<td>99.2</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>250</td>
<td>91.5</td>
<td>97.5</td>
<td>99.5</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>93.9</td>
<td>98.8</td>
<td>99.4</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>93.9</td>
<td>99.3</td>
<td>99.6</td>
<td>100</td>
</tr>
</tbody>
</table>

2. The conclusion that more observations yield greater coverages in univariate case seems not applicable here. The differences of ECP with respect to length of realizations are really negligible.

3. The value of \( \gamma \) has barely any effects on the performances, that is, the coverage probabilities change little regarding different model specifications. This is especially true for \( r_1 \), in which case all the ECP’s are basically 100%.

2.6.2 Subsampling for bivariate observations

To describe the subsampling method, set \( U_t(i) = X_t^2(i) \) \( i = 1, 2 \), denote by \( \hat{r}_{0,b,t} \) the sample correlation of \( U_t(1), \ldots, U_{t-b+1}(1) \) and \( U_t(2), \ldots, U_{t-b+1}(2) \), \( t = 1, \ldots, n - b + 1 \). Similarly, denote by \( \hat{r}_{0,b,t} \) the sample correlation of \( U_t(1), \ldots, U_{t-b+1}(1) \) and \( U_{t-1}(2), \ldots, U_{t-b}(2) \), \( t = 2, \ldots, n-b+1 \). Thus \( \hat{r}_{0,n} =: \hat{r}_{0,n} \) and \( \hat{r}_{1,n} =: \hat{r}_{1,n} \) are estimators, respectively of \( r_0 \) and \( r_1 \). The equal-tailed subsampling confidence interval is based on the empirical distributions of

\[
\xi_{m,b,t} := \sqrt{b}(\hat{r}_{m,b,t} - \hat{r}_{m,n}), \quad m = 0, 1.
\]
Table 2.10. Empirical coverage probabilities of equal-tailed and symmetric confidence interval based on the subsampling method for $r_0$ and $r_1$ in the CCC-GARCH model, with $n = 500$.

<table>
<thead>
<tr>
<th>Model</th>
<th>$r_0$</th>
<th></th>
<th>$r_1$</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>e.c.p. (%)</td>
<td>e.c.p. (%)</td>
<td>e.c.p. (%)</td>
<td>e.c.p. (%)</td>
</tr>
<tr>
<td>equal-tailed</td>
<td>88.1</td>
<td>77.4</td>
<td>50.8</td>
<td>97.0</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
<td>90.9</td>
<td>82.5</td>
<td>46.1</td>
</tr>
<tr>
<td>25</td>
<td>91.0</td>
<td>87.4</td>
<td>42.6</td>
<td>95.8</td>
</tr>
<tr>
<td>100</td>
<td>82.7</td>
<td>81.9</td>
<td>39.2</td>
<td>81.7</td>
</tr>
<tr>
<td>symmetric</td>
<td>85.9</td>
<td>85.3</td>
<td>43.2</td>
<td>96.1</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
<td>88.6</td>
<td>82.4</td>
<td>37.1</td>
</tr>
<tr>
<td>25</td>
<td>89.9</td>
<td>85.1</td>
<td>39.3</td>
<td>92.2</td>
</tr>
<tr>
<td>100</td>
<td>83.4</td>
<td>81.8</td>
<td>46.5</td>
<td>84.2</td>
</tr>
</tbody>
</table>

More specifically, denoting by $q_0(\beta)$ the $\beta$th quantile of

$$L_{0,b}(x) = \frac{1}{n - b + 1} \sum_{t=1}^{n-b+1} \mathbb{1}\{\xi_{0,b,t}(h) \leq x\},$$

an equal-tailed level $(1 - \alpha)$ confidence interval for $r_0$ is

$$\left( \hat{r}_{0,n} + n^{-1/2}q_0(\alpha/2), \hat{r}_{0,n} + n^{-1/2}q_0(1 - \alpha/2) \right).$$

The confidence interval for $r_1$ is constructed in the same way, but in the definition of of $L_{1,b}(x)$ the range of $t$ must be adjusted. Symmetric confidence intervals can be constructed in a similar way. Table 2.10 gives the ECP's of equal-tailed and symmetric confidence interval for cross-correlations $r_0$ and $r_1$, with length of realizations $n = 500$. The results for $n = 1000$ are presented in Table B.5. We observe

1. Equal-tailed and symmetric confidence intervals perform equally. The ECP substantially depends on the value of $\lambda$; it can be very close to the nominal coverage of 95% for the first model with $\lambda = 0.06$ in Table 2.8, 85% is the best we can get for model 2 with $\lambda = 0.29$, and it is too low for model 3 with $\lambda = 0.81$.

2. As what we suggested in univariate case, the use of short $b$, less than 10% of $n$, is recommended.
2.6.3 Summary

Residual bootstrap and subsampling are implemented to construct confidence intervals for cross-correlations at lags 0 and 1 of CCC-GARCH(1,1) model. The parametric method residual bootstrap is seen to perform much better than the nonparametric method. See Figure B.4 for a graphical comparison. Moreover, it gives equally good coverages for model specifications with different values of $\gamma$, whereas the ECP’s of subsampling is considerably contingent on this fourth-moment existence parameter. As for the choice of equal-tailed or symmetric confidence interval, no preference is evident.
CHAPTER 3
DISCRIMINATING BETWEEN LONG MEMORY AND
CHANGE-POINTS IN VOLATILITY

3.1 Introduction

We consider the test developed in [14] for distinguishing between a long-range dependent
time series and a weakly dependent time series with change-points in the mean. Denote
by \{r_t\}_t returns on a speculative asset. Based on empirical evidence, we assume that it is
a mean zero process. Working with squared returns \(X_t = r_t^2\), the test of [14] can be used
to discriminate between a long memory and a change point model for the volatility of a
conditionally heteroskedastic time series.

Our null hypothesis is that the observations \(X_t\) follow a change point model, while
under the alternative the observations are long range dependent. The testing procedure is
based on a CUSUM statistic for the partial sums, which is probably the most widely used
statistic for detecting and estimating change-points, see [44]. The test statistic involves
estimating the long-run variances of subsamples obtained by dividing the observations into
two parts, before and after a potential change-point. Our objective is to develop a robust
implementation of the test of [14]. We examine and compare the finite sample performance
of several implementations by considering different variance estimators \(s_{n1}^2\) and \(s_{n2}^2\) in (3.3.5)
and (3.3.6), respectively. We study log-returns on two major US stock indexes, DJIA and
NASDAQ, and four component stocks of DJIA during different time spans so as to obtain
practically relevant models with a change point in volatility that will be used in simulations.
As a long memory alternative we use the LARCH model of [74]. We now explain the set up
of the simulation study. We first fit GARCH(1,1) models to each of the sub-series before
and after the estimated breaks to obtain relevant change point models for returns. We next
use four of the six change point models to conduct an extensive simulation study. Empirical
sizes are computed using processes of different lengths with varying locations of the break.
The Bartlett kernel estimator of \(s_{n1}^2\) and \(s_{n2}^2\) using the truncation lag \(q_{c,arma}\), which is a
calibrated version of the bandwidth selection procedure in [5], is found preferable to other
variance estimators. It leads to a test with best size. We then validate the robustness of this variance estimator by applying it to the other two models and obtain satisfactory results. Another issue we address concerns the robustness of the testing procedure to model mis-specification. We therefore consider two popular asymmetric GARCH(1,1) models and apply the procedure based on standard GARCH(1,1) model to processes simulated according to these different models. We find that our method is robust to model mis-specification. After the investigation of the empirical size, we examine the empirical power of the test by using the LARCH model as the long range dependent conditionally heteroskedastic model satisfying $H_A$. Applied to the six sets of real returns data, the test fails to reject the null of a volatility change in favor of a long memory alternative.

The chapter is organized as follows. After reviewing the relevant literature in Section 3.2, we introduce the testing procedure in Section 3.3. Section 3.4 discusses several methods of estimating the variance. Using the models obtained in Section 3.5 from daily returns on stocks and stock indexes, we study the finite sample performance of the tests by means of simulations in Section 3.6.

### 3.2 Historical and Empirical Background

An empirical phenomenon, identified later as long-range dependence, or long memory, was first observed by a hydrologist [102] when he was analyzing the annual water flow data of the Nile river for the planning of the second Aswan dam. Intuitively, long-range dependence means that observations far away from each other are still strongly correlated. In the time domain, a stationary time series $X_t$ is said to possess long-range dependence if the autocorrelation function $\rho(k)$ is not absolutely summable and, more precisely,

$$\lim_{k \to \infty} \frac{\rho(k)}{k^{2d-1}} = c_\rho,$$

where $c_\rho$ is a positive constant and $d \in (0, 0.5)$ denotes the memory parameter. The autocorrelation function $\rho(k)$ in (3.2.1) decays at a hyperbolic rate as the lag $k$ increases, whereas for stationary short memory processes, such as ARMA (corresponding to $d = 0$), $\rho(k)$ decays exponentially. In the frequency domain, an equivalent definition of long-range...
dependence can be given by using the spectral density $f(\lambda)$ of the process $X_t$,

$$
(3.2.2) \quad \lim_{\lambda \to 0^+} \frac{f(\lambda)}{\lambda^{-2d}} = c_f,
$$

where $c_f$ is a positive constant. The long-term behavior is determined by low frequencies in spectral analysis; the spectral density of a long-memory process has a pole at the origin. Equations (3.2.1) and (3.2.2) can also be expressed in terms of the Hurst parameter $H \in (0.5, 1)$, which relates to the memory parameter $d$ via $d := H - 0.5$. We use $d$ in this chapter because it is commonly used to denote the differencing parameter of the fractional ARIMA processes discussed below. For $d \in (-0.5, 0)$ the series is said to exhibit intermediate or negative memory. In the case $d > 0.5$, the process is no longer covariance stationary but still exhibits persistence, and is usually referred to as a non-stationary long-memory process. We focus on stationary long-range dependence, i.e. $d \in (0, 0.5)$, in the present chapter.

In [128] the authors first proposed to use self-similar processes, most notably the fractional Brownian motion, to model long memory series. Over a decade later, [80] and [101] introduced the autoregressive fractional integrated moving average, or ARFIMA$(p,d,q)$, processes which are approximately self-similar and offer a much greater modeling flexibility. The model is a generalization of the ARIMA$(p,d,q)$ model proposed by [26], where the differencing parameter $d$ is an integer. In an ARFIMA model the order of integration is allowed to be a non-integer, and when $d \in (0,0.5)$ the process is stationary and exhibits long-range dependence. Starting from the 1980's researchers have been using long memory processes to model macroeconomic time series, whereas the 1990's witnessed an increased interest in modeling the volatility of returns on speculative assets by such processes, see [96] for an in depth discussion and relevant references. Initiated by the pioneering work of [120] and [142], self-similar processes have also been widely used to model certain aspect of computer networks traffic, an extensive treatment of such applications and the relevant theory are provided in [141]. See [51] for a recent extensive review of the theory and applications of long-range dependent models.

Similar to the debate in the literature on unit roots versus mean shifts, see [143] and [144], several authors have argued that the empirical evidence for long-range dependence
is attributable to the presence of trends or structural breaks in the data. In the statistical literature, it has been pointed out (see for example, [17], [162], [71]), that statistics computed from short memory processes perturbed by trends or shifts in the mean may exhibit the same properties as those of long-range dependent processes. For example, the sample variance and several statistics akin to the modified R/S statistic of [123] may display long memory type behavior when applied to short-memory processes affected by structural changes. Standard tools, such as ACF plots and sample periodograms can not help to distinguish between long memory and structural breaks, because they all behave in a very similar way under the two alternatives. For example, a positive relation between the number of level shifts and the value of the memory parameter $d$ in finite samples was found by [78] by means of simulation. Liu in [121] shows that long memory in volatility can be thought of as discrete shifts that persist for very long time periods. One can find in the literature a number of long memory tests designed to test the null hypothesis of weak dependence against an alternative of long-range dependence and change-point tests developed to test the same null hypothesis but against a change-point alternative. The easy confusion between long memory and change-points is also reflected by the fact that most long memory tests reject in the presence of change-points and many change-point tests reject in the presence of long memory, see, for example, [98], and [113] for detailed examples and discussion. It thus seems that, despite their rather different mathematical formulations, long-range dependent processes and structural change models can both model the same phenomena, and it is often difficult to decide on purely statistical grounds which approach is preferable.

There have, correspondingly, been two opposite positions concerning how to model the volatility of financial returns. One is in favor of the long-range dependence hypothesis and the other prefers regime changes.

Works that advocate long memory models can be found in [46], [50], [79], [8], [22], [2], among others. In these studies, the memory properties of the volatility have been analyzed using daily and intra-daily stock and exchange rate returns over different time spans, and comprehensive evidence of long-term persistence has been provided. The authors
employed a variety of econometric techniques, both parametric (FIGARCH, FI-EGARCH, ARFIMA) and semi-parametric (stationarity test, log periodogram regression, local Whittle estimators). The main conclusion is that the volatility of financial returns can be well described by long-range dependent models, that is, long memory appears to be the defining feature of the data generating processes. Moreover, the theoretical results of [37] indicate that the degree of persistence is invariant with respect to the sampling frequency.

Researchers who are opposed to the long memory modeling argue that the strong persistence observed in the volatility of financial returns may be a spurious feature due to unaccounted for structural breaks. Since the work of [49], and [118], among others, it is well-known that neglecting level shifts may give rise to misleading inference on the persistence in volatility of financial data. Mikosch and Stărică ([130], [132]) contend that the long-range dependence in the conditional variance of returns on S&P 500 index is, in fact, a manifestation of changes in the parameters of the underlying GARCH models. They also argue that the long-term persistence observed in sample ACFs of absolute and squared log-returns can be explained by shifts in the unconditional variance of the underlying model. In [114] it was observed that long memory in squares of German stock returns disappears once shifting means are properly accounted for. Econometric models where structural change can be modeled endogenously have been proposed by [32], and [89] introduced Markov switching ARCH model, while [55] extended the Markov switching model to a GARCH framework. The main conclusion of these studies is that the long lasting persistence in the volatility processes can be successfully explained by a model of switching heteroskedasticity. Economic interpretations of these shifts relate the dynamics of stock volatility to the business cycle, the variability of fundamentals and the fad components of returns.

An intermediate position of combining long memory and level shifts was recently put forward by some authors to analyze economic and financial time series. Bos et estimated in [24] the differencing parameter $d$ in ARFIMA$(p,d,q)$ model for the inflation rates in the G7 countries and found a significant evidence of long-range dependence and that the addition of a set of level shifts does decrease the degree of persistence for some countries.
Using exchange rates data, [134] concluded that superior forecasts can be obtained at longer horizons by modeling both long memory and structural changes.

The choice between long memory and structural breaks is important in a variety of financial applications, particularly in risk measurement, asset allocation and option pricing. In a simulation study of call option prices for the S&P 500 index, [22] showed that taking into account a long memory structure of the volatilities can sometimes even double the price of options, compared with situations when long-memory is neglected. One of the merits of long memory models is that they may provide a parsimonious description of a long, possibly non-stationary, time series, which is particularly useful for the purpose of forecasting. On the other hand, the first stage of using a structural break model is to quickly detect a change in the pattern or in the parameters of the econometric model, which itself is not an easy task, and then to fit an appropriate model to the most recent stretch of data after the last estimated change-point. Formal statistical tests which would help decide whether a long-range dependent process or a weakly dependent process with change-points is a better fitting model for a particular time series are therefore of value. There has however not been much research in this direction. The periodogram based testing procedure in [116] is developed to distinguish between a long-range dependent process and the process \( X_k = Y_k + f(k) \) with a monotonic function \( f \) and Gaussian weakly dependent \( Y_k \). In [97] tests were constructed for detecting long-range dependence which are based on a smoothed periodogram are robust in the presence of small trends. Sibbertsen and Venetis recently developed these ideas and proposed a test statistic, which is the squared difference between the [70] estimator of \( d \) and its version based on the tapered periodogram. These tests however are not directly applicable to the main subject we want to investigate in the present chapter, that is, discriminating between long memory and volatility changes in financial time series.
3.3 The Test

As in [14], the observations $X_t$ follow a change-point model if

$$(3.3.3) \quad X_t = \begin{cases} \mu + Y_t, & 1 \leq t \leq k^* \\ \mu + \Delta + Y_t, & k^* < t \leq n. \end{cases}$$

In (3.3.3), $k^*$ is the unknown time of a change in mean; the means $\mu$ and $\mu + \Delta$ are also unknown. The sequence $\{Y_t\}$ is assumed to have mean zero and to be stationary and weakly dependent.

We wish to test

$$H_0: \text{The observations } X_t \text{ follow model (3.3.3).}$$

versus

$$H_A: \text{The observations } X_t \text{ are long range dependent.}$$

Let $X_t = r_t^2$, where $\{r_t\}_1^n$ denotes returns on a speculative assets with mean zero, the unknown means $\mu$ and $\mu + \Delta$ in (3.3.3) are then variances of the returns. We thus test to discriminate between long-range dependence in variance and changes in variance. Consider a change point estimator computed from an observed realization $\{X_t\}_1^n$,

$$(3.3.4) \quad \hat{k} = \min \left\{ k : \max_{1 \leq t \leq n} \left| \sum_{1 \leq j \leq t} X_j - \frac{t}{n} \sum_{1 \leq j \leq n} X_j \right| = \max_{1 \leq j \leq k} \left| \sum_{1 \leq j \leq k} X_j - \frac{k}{n} \sum_{1 \leq j \leq n} X_j \right| \right\}.$$  

Let $n_1 = \hat{k}$ and $n_2 = n - \hat{k}$, we separate the realization into sub-series, one before the other after the break, denoted, respectively, by $\{X_t\}_{1}^{n_1}$ and $\{X_t\}_{n_1+1}^{n}$. We next define the statistics

$$(3.3.5) \quad T_{n_1} = \frac{1}{s_{n_1}} n_1^{-1/2} \max_{1 \leq k \leq n_1} \left| \sum_{1 \leq i \leq k} X_i - \frac{k}{n_1} \sum_{1 \leq j \leq n_1} X_j \right|$$

based on $\{X_t\}_{1}^{n_1}$ and

$$(3.3.6) \quad T_{n_2} = \frac{1}{s_{n_2}} n_2^{-1/2} \max_{n_1 < k \leq n} \left| \sum_{n_1 < i \leq k} X_i - \frac{k - n_1}{n_2} \sum_{n_1 < j \leq n} X_j \right|$$

based on $\{X_t\}_{n_1+1}^{n}$. In the definitions above, $s_{n_1}^2$ and $s_{n_2}^2$ are the estimates of the long-run variance of $\{X_t\}_{1}^{n_1}$ and $\{X_t\}_{n_1+1}^{n}$, respectively. In this chapter, we consider nonparametric
Table 3.1. Asymptotic critical values of the test statistic $M_n$.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>.10</th>
<th>.05</th>
<th>.01</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c(\alpha)$</td>
<td>1.36</td>
<td>1.48</td>
<td>1.72</td>
</tr>
</tbody>
</table>

kernel estimators with bandwidth parameters $q$ and also a method referred to as prewhitening that is independent of selecting $q$. In [14] Berkes et al used the Bartlett kernel estimator and established its almost sure convergence for time series under both $H_0$ and $H_A$. Specifically, setting

$$X_{n_1} = \frac{1}{n_1} \sum_{1 \leq t \leq n_1} X_t, \quad X_{n_2} = \frac{1}{n_2} \sum_{n_1 < t \leq n} X_t,$$

and the Bartlett kernel

$$\omega_j(q) = 1 - \frac{j}{q + 1},$$

the variance estimator $s_{n_1}^2$ and $s_{n_2}^2$ are defined as

$$s_{n_1}^2 = \frac{1}{n_1} \sum_{1 \leq t \leq n_1} (X_t - \bar{X}_{n_1})^2$$

$$+ 2 \sum_{1 \leq j \leq q(n_1)} \omega_j(q(n_1)) \frac{1}{n_1} \sum_{1 \leq t \leq n_1 - j} (X_t - \bar{X}_{n_1})(X_{t+j} - \bar{X}_{n_1}),$$

$$s_{n_2}^2 = \frac{1}{n_2} \sum_{n_1 < t \leq n} (X_t - \bar{X}_{n_2})^2$$

$$+ 2 \sum_{1 \leq j \leq q(n_2)} \omega_j(q(n_2)) \frac{1}{n_2} \sum_{n_1 < t \leq n} (X_t - \bar{X}_{n_2})(X_{t+j} - \bar{X}_{n_2}),$$

(3.3.7) (3.3.8)

The test statistic is

$$M_n = \max\{T_{n_1}, T_{n_2}\}.$$  

(3.3.9)

Under $H_0$, Corollary 2.1 of [14] gives

$$M_n \overset{d}{\rightarrow} M = \max \left\{ \sup_{0 \leq t \leq 1} |B(1)(t)|, \sup_{0 \leq t \leq 1} |B(2)(t)| \right\},$$

where $B(1)$ and $B(2)$ are independent Brownian bridges. Since the distribution function of $\sup_{0 \leq t \leq 1} |B(t)|$ is known (cf. Section 1.5 of [45]), the limit distribution of $M_n$ can be computed explicitly. Theorem 2.2 of [14] established that $M_n$ tends to infinity in probability
under $H_A$. Table 3.1 gives asymptotic critical values $c(\alpha)$ defined by $P(M > c(\alpha)) = \alpha$. We reject $H_0$ if $M_n > c(\alpha)$.

### 3.4 Variance Estimation

In this section we investigate various methods of obtaining $s_{n_1}^2$ and $s_{n_2}^2$ in (3.3.5) and (3.3.6), respectively. When a nonparametric kernel function is used to estimate the long-run variance of a time series $\{X_t\}$, the basic guidance in choosing the bandwidth parameter (truncation lag) $q$ is that $q \to \infty$ and $q/n \to 0$ as the sample size $n$ approaches infinity. Although different choices of $q$ yield the same asymptotic distribution for a certain test statistic, the finite sample performance of tests using nonparametric estimates depends greatly on the bandwidth parameter. In what follows we consider three methods for selecting the truncation lags $q(n_1)$ and $q(n_2)$ in (3.3.7) and (3.3.8), respectively, and also a variance estimator that does not rely on $q$.

We first consider a method of choosing $q$ that is merely dependent on sample size $n$, and denote it by $q_n$. Specifically, we let $q_n(n_1) = 10\log_{10}(n_1)$ and $q_n(n_2) = 10\log_{10}(n_2)$. In statistical packages, such as S-PLUS and R, $10\log_{10}(n)$ is the default maximum lag up to which the autocovariance function is computed.

Assume GARCH(1,1) is the underlying model for the returns before and after change-points of volatility under $H_0$. Specifically,

\begin{equation}
(3.4.10) \quad r_t = \sigma_t Z_t,
\end{equation}

where $\{Z_t\}$ is an i.i.d. sequence with mean zero and unit variance and $\sigma_t$ evolves according to

\begin{equation}
(3.4.11) \quad \sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2,
\end{equation}

with $\alpha$ and $\beta$ positive constants such that $\alpha + \beta < 1$. It is known that $X_t = r_t^2$ follow an ARMA(1,1) model with autoregressive coefficient $\rho = \alpha + \beta$ and moving average coefficient $\psi = -\beta$, see [91] pp. 665-666 for details. In the sense of minimizing the asymptotic truncated MSE, [5] derived optimal bandwidth parameters $q^*$ for a class of real-valued
kernel functions. The optimal truncation lag for Bartlett kernel is given by

\[ q^* = 1.1447 \left( a(X)n \right)^{1/3}, \]  

where \( a(X) \) is a function of the unknown spectral function \( f(\lambda) \) of the process \( \{X_t\}^n \). For ARMA(1,1) models with autoregressive parameter \( \rho \) and moving average parameter \( \psi \), the estimate of \( a(X) \) is given by Eq. (6.6) in [5], in which an integer \( p \) is involved. By setting \( p = 1 \), we obtain

\[ \hat{a}(X) = \frac{4 \left( 1 + \hat{\rho} \hat{\psi} \right)^2 \left( \hat{\rho} + \hat{\psi} \right)^2}{\left( 1 - \hat{\rho}^2 \right)^2 \left( 1 + \hat{\psi} \right)^4}, \]  

where \( \hat{\rho} \) and \( \hat{\psi} \) are appropriate estimates. Denote by \( q^{*}_{\text{arma}} \) the optimal truncation lag for ARMA(1,1) model. We now explain in detail the procedure for obtaining it. We first estimate the change-point of volatility of the returns \( \{r_t\}^n \) using \( \hat{k} \) in (3.3.4), and split the process into sub-series before the break \( \{r_t\}^{n_1} \) and after the break \( \{r_t\}^{n_1+1}_n \), where \( n_1 = \hat{k} \). Focusing on \( \{r_t\}^{n_1} \), we compute the quasi-maximum likelihood estimates (QMLE) \( \hat{\alpha}_1 \) and \( \hat{\beta}_1 \) of the GARCH(1,1) model. By setting \( \hat{\rho}_1 = \hat{\alpha}_1 + \hat{\beta}_1 \) and \( \hat{\psi}_1 = -\hat{\beta}_1 \) and plugging them in (3.4.13), we get the estimate \( \hat{a}_1(X) \). As the final step, we calculate the optimal bandwidth \( q^{*}_{\text{arma}}(n_1) \) for the squared series before the change point by substituting \( a(X) \) in (3.4.12) by \( \hat{a}_1(X) \) and \( n \) by \( n_1 \). The lag \( q^{*}_{\text{arma}}(n_2) \), \( n_2 = n - n_1 \), is computed in the same way for the series, \( \{X_t\}^{n_1+1}_n \), after the change point.

As will be seen in the simulation study in Section 3.6, the empirical sizes of the test using \( q^{*}_{\text{arma}} \) are generally much lower than the nominal ones, indicating that the \( q^{*}_{\text{arma}} \) is too long to give satisfactory results. Furthermore, it is noticed in Tables 3.3, C.1, and C.2 that the performance of the test depends not only on the location of the change point \( k^* \), but also on sample size \( n \). The observations lead us to shorten the lags by multiplying by a factor \( c \) less than 1 both \( q^{*}_{\text{arma}}(n_1) \) and \( q^{*}_{\text{arma}}(n_2) \). The constant \( c \) should take account of both \( k^* \) and \( n \). Specifically, denote by \( \text{Var}(X_{n_1}) \) and \( \text{Var}(X_{n_2}) \) the sample variances of the squared sub-series, \( \{X_t\}^{n_1}_1 \) and \( \{X_t\}^{n_2}_{n_1+1} \), before and after the estimated change point \( \hat{k} \).
respectively. The factor $c$ is determined as follows.

(3.4.14) \[ \text{If } \text{Var} (X_{n_1}) \leq \text{Var} (X_{n_2}), \quad c = \log_{10} \left( \frac{n}{100} \right) \frac{1}{n/n^2 + 1.5}; \]

(3.4.15) \[ \text{If } \text{Var} (X_{n_1}) > \text{Var} (X_{n_2}), \quad c = \log_{10} \left( \frac{n}{100} \right) \frac{1}{n/n^2 + 1.5}; \]

The rationale will be explained in Section 3.6, where the effects of $n$ and $k^*$ on empirical sizes are discussed in detail. In the sequel, we denote by $q^*_\text{arma}$ the product of $c$ and $q^*_\text{arma}$, a modified version of selecting $q$ based on the optimal bandwidth of Bartlett kernel for ARMA(1,1) model.

Until now, we have been focusing on finding appropriate bandwidth parameter, so that sample autocovariances at lags up to $q$ can be summed up using Bartlett kernel function to estimate the long-run variance of the process $\{X_t\}^n$. The last method we consider in the chapter does not depend on choosing truncation lag. Specifically, we use the VARHAC (vector autoregression heteroskedasticity and autocorrelation consistent) procedure proposed by [83]. We first fit AR($b$) model with autoregressive order $b$ chosen by Akaike's Information Criterion (AIC). This step is referred to as prewhitening. Denote by $\hat{\rho}_i$, $i = 1, 2, \ldots, b$, the $i$th autoregressive coefficient, and by $\Sigma_{\text{PW.resid}}$ the sample variance of the prewhitened residuals, the estimate of the long-run variance is then

(3.4.16) \[ s^2 = \frac{\Sigma_{\text{PW.resid}}}{\left( 1 - \sum_{i=1}^{b} \hat{\rho}_i \right)^2}. \]

In practice, we apply the procedure to the two squared sub-series, $\{X_t\}^n_1$ and $\{X_t\}^n_{n_1+1}$, to obtain $s^2_{n_1}$ and $s^2_{n_2}$, respectively, in (3.3.5) and (3.3.6). Originally proposed by [151], prewhitening has long been used in time series literature to reduce the bias of kernel-based spectral estimation, especially in the presence of strong temporal dependence. The AR model is not meant to be the true model of the underlying process, instead it is used as a tool to "soak up" some of the temporal dependence (cf. [6], p. 954). The VARHAC method differs in two aspects from the prewhitened kernel estimator in [6]. First, they use AR model of fixed order, AR(1), in their simulations, instead of an order chosen by
AIC. Second, as opposed to using the sample variance $\Sigma_{PW,\text{resid}}$ in (3.4.16) to construct the variance estimate $s^2$, they follow a bandwidth selection procedure by fitting an AR(1) model to the prewhitened residuals and then use a kernel function to obtain the variance of the residuals. They do so in the view that temporal dependences still exist in the prewhitened residuals. In [85] the authors highlight the pitfalls of using a fixed order of VAR prewhitening and find relatively little benefit from applying a kernel-based method to the prewhitened residuals. Our simulations confirm their findings. Indeed, the prewhitened residuals are so close to white noise that fitting an AR(1) model yields the autoregressive coefficient not significantly different from zero.

3.5 Patterns of Volatility

In this section we study the daily returns on major US stocks and stock indexes with the objective to find what the typical model changes and the implied variance changes of real financial time series are. In the simulation study in Section 3.6 we will use these practically relevant models to compare the finite sample performance of various variance estimators.

We consider two indexes: the Dow Jones Industrial Average index (DJIA) of the period from 1/1/1992 to 12/31/1999 and the National Association of Securities Dealers Automated Quotations Composite index (NASDAQ) covering the period from 7/1/1994 to 12/31/1998. In addition, we consider four constituent stocks of DJIA: the General Electric, the Wal-Mart Stores, Inc., and the American Express of the same period from 1/1/2000 to 6/30/2004, and the Altria Group, Inc. of the period from 7/1/1997 to 12/31/1999. The coverage periods of the time series range from two and half years to eight years, with number of observations varying between 631 and 2021. We work with log returns $r_t = 100 \log(P_t/P_{t-1})$, where $P_t$ denotes the index value or stock price at time $t$. For each of the time series, we first estimate the break point of volatility using $\hat{k}$ in (3.3.4), and then separate the returns data with respect to the change-point and fit a GARCH(1,1) model to these sub-series so as to estimate the models before and after the break.

Figure 3.1 presents the time series plots of the six data sets, with the same limits on
Figure 3.1. Daily returns on stocks and stock indexes. Dashed lines indicate the borderlines of the subsamples, i.e., the location of the estimated change-point.
Table 3.2. Fitted GARCH(1,1) models from the returns on DJIA, NASDAQ, General Electric, Wal-Mart, American Express, and Altria Group.

<table>
<thead>
<tr>
<th>Speculative Asset</th>
<th>Period</th>
<th>Model</th>
<th>Number of obs.</th>
<th>$\mu$</th>
<th>$\omega$</th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>Sample Var.</th>
<th>Implied Var.</th>
</tr>
</thead>
<tbody>
<tr>
<td>DJIA</td>
<td>1/1/1992 ~</td>
<td>Model 1</td>
<td>1061</td>
<td>0.076</td>
<td>0.025</td>
<td>0.064</td>
<td>0.879</td>
<td>0.424</td>
<td>0.439</td>
</tr>
<tr>
<td></td>
<td>12/31/1999</td>
<td>Model 2</td>
<td>960</td>
<td>0.073</td>
<td>0.095</td>
<td>0.097</td>
<td>0.839</td>
<td>1.465</td>
<td>1.484</td>
</tr>
<tr>
<td>NASDAQ</td>
<td>7/1/1994 ~</td>
<td>Model 3</td>
<td>775</td>
<td>0.129</td>
<td>0.133</td>
<td>0.132</td>
<td>0.75</td>
<td>1.133</td>
<td>1.127</td>
</tr>
<tr>
<td></td>
<td>12/31/1998</td>
<td>Model 4</td>
<td>363</td>
<td>0.321</td>
<td>0.549</td>
<td>0.171</td>
<td>0.677</td>
<td>3.57</td>
<td>3.612</td>
</tr>
<tr>
<td>General Electric</td>
<td>1/1/2001 ~</td>
<td>Model 5</td>
<td>473</td>
<td>-0.13</td>
<td>4.345</td>
<td>0.094</td>
<td>0.281</td>
<td>7.147</td>
<td>6.952</td>
</tr>
<tr>
<td></td>
<td>6/30/2004</td>
<td>Model 6</td>
<td>402</td>
<td>0.031</td>
<td>0.284</td>
<td>0.082</td>
<td>0.775</td>
<td>2.038</td>
<td>1.986</td>
</tr>
<tr>
<td>Wal-Mart</td>
<td>1/1/2001 ~</td>
<td>Model 7</td>
<td>457</td>
<td>0.027</td>
<td>1.407</td>
<td>0.172</td>
<td>0.499</td>
<td>4.355</td>
<td>4.277</td>
</tr>
<tr>
<td></td>
<td>6/30/2004</td>
<td>Model 8</td>
<td>418</td>
<td>-0.015</td>
<td>0.379</td>
<td>0.038</td>
<td>0.753</td>
<td>1.85</td>
<td>1.813</td>
</tr>
<tr>
<td>American Express</td>
<td>1/1/2001 ~</td>
<td>Model 9</td>
<td>475</td>
<td>-0.09</td>
<td>4.717</td>
<td>0.148</td>
<td>0.274</td>
<td>8.288</td>
<td>8.152</td>
</tr>
<tr>
<td></td>
<td>6/30/2004</td>
<td>Model 10</td>
<td>398</td>
<td>0.074</td>
<td>0.201</td>
<td>0.107</td>
<td>0.796</td>
<td>2.068</td>
<td>2.072</td>
</tr>
<tr>
<td>Altria Group</td>
<td>7/1/1997 ~</td>
<td>Model 11</td>
<td>389</td>
<td>0.092</td>
<td>0.933</td>
<td>0.116</td>
<td>0.618</td>
<td>3.558</td>
<td>3.507</td>
</tr>
<tr>
<td></td>
<td>12/31/1999</td>
<td>Model 12</td>
<td>242</td>
<td>-0.224</td>
<td>1.145</td>
<td>0.088</td>
<td>0.773</td>
<td>8.482</td>
<td>8.237</td>
</tr>
</tbody>
</table>

y-axis. The locations of the estimated break points of volatility are marked by a dashed line. DJIA is a price-weighted average of thirty blue chip companies, and NASDAQ, comprised of more than 5000 domestic and foreign companies, is a weighted index based on market value. We thus expect, as clearly shown in the plots, that the volatilities of the indexes should be relatively smaller than those of individual stocks. It is also seen that, during the two and half years starting from 2001, the returns of the three stocks, GE, Wal-Mart, and American express, exhibit somewhat similar patterns. For instance, the estimated breaks of the three time series happen closely, approximately in the fourth quarter of 2002, and, moreover, the volatilities all decrease after the change-points. This may be attributed to the steady increase of stock prices since then. In contrast, during the time periods consider for DJIA, NASDAQ, and Altria Group, the returns are more volatile after the change-points.

A GARCH(1,1) model, specified in (3.4.10) and (3.4.11), with mean value $\mu$ is estimated
to the sub-series of the six time series. Table 3.2 gives the results. The sample variance and implied variance of each stretch of data are also reported. Sample variances are computed in the usual way. Provided that \( \alpha + \beta < 1 \), which is true for all the models in Table 3.2, the unconditional variance of a GARCH(1,1) process is given by \( \omega/(1 - \alpha - \beta) \), the implied variance is thus determined as \( \hat{\omega}/(1 - \hat{\alpha} - \hat{\beta}) \). The closeness of the two variances reflects the goodness of fit of the model to the returns data.

We have two main categories of models. Represented by DJIA, NASDAQ, and Altria Group, the volatilities of the first group increase after the breaks, while, the returns of the second group, represented by GE, Wal-Mart, and American Express, have more substantial swings before the breaks. We notice in the first group that the sum of \( \hat{\alpha} \) and \( \hat{\beta} \) does not change much after the breaks, the values of \( \hat{\omega} \) however increase. For the second group of models, \( \hat{\alpha} + \hat{\beta} \) increases slightly after the change-points, but the values of \( \hat{\omega} \) drop considerably. We thus conclude from the expression \( \hat{\omega}/(1 - \hat{\alpha} - \hat{\beta}) \) that the increase (decrease) of the parameter \( \omega \) is the main cause of the ups (downs) of the volatilities of the returns. As we observed in Figure 3.1, the magnitudes of volatility of indexes are much lower than those of stocks, Table 3.2 gives precise numerical values. The sample variance of DJIA increases from 0.424 to 1.465, and NASDAQ from 1.133 to 3.57, whereas, the smallest volatility change we observe for stocks is 1.85 for Wal-Mart and the largest is 8.482 for Altria Group.

3.6 Simulation Study

We wish to distinguish between a long-range dependent time series and a weakly dependent time series, exhibiting conditional heteroskedasticity, with a change-point in the variance. We compare the finite sample performance of the test with different implementations of the variance estimators discussed in Section 3.4. Since the primary concern is the type I error, the goal is to find a testing procedure that outperforms the others in terms of empirical size. Its empirical power will then be examined.

We use the change-point models obtained from the returns on DJIA, NASDAQ, GE, and Wal-Mart and generate squared GARCH(1,1) observations of length \( n = 500, 1000, \)
Table 3.3. Empirical sizes (in percent) of the test using four different variance estimators applied to simulated series of squared GARCH(1,1) observations following the change-point models estimated from the returns on DJIA, NASDAQ, GE, and Wal-Mart in Table 3.2. The number of replications is $R = 5000$, with nominal size $\alpha = 5\%$.

<table>
<thead>
<tr>
<th>Method</th>
<th>$n$</th>
<th>$k^*=\frac{1}{4}n$</th>
<th>$k^*=\frac{1}{2}n$</th>
<th>$k^*=\frac{3}{4}n$</th>
<th>$k^*=\frac{1}{4}n$</th>
<th>$k^*=\frac{1}{2}n$</th>
<th>$k^*=\frac{3}{4}n$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DJIA : Model 1 → Model 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$q_n$</td>
<td>500</td>
<td>5.24</td>
<td>2.56</td>
<td>1.52</td>
<td>2.86</td>
<td>1.3</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>26.74</td>
<td>8.16</td>
<td>4.08</td>
<td>18.9</td>
<td>3.8</td>
<td>1.94</td>
</tr>
<tr>
<td></td>
<td>1500</td>
<td>35.28</td>
<td>12.12</td>
<td>6.62</td>
<td>26.96</td>
<td>5.54</td>
<td>2.94</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>39.36</td>
<td>12.58</td>
<td>8.98</td>
<td>29.04</td>
<td>5.4</td>
<td>3.74</td>
</tr>
<tr>
<td>$q_{arma}$</td>
<td>500</td>
<td>0.98</td>
<td>0.58</td>
<td>0.4</td>
<td>1.66</td>
<td>0.88</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>1.66</td>
<td>1.06</td>
<td>1.0</td>
<td>3.56</td>
<td>2.18</td>
<td>1.62</td>
</tr>
<tr>
<td></td>
<td>1500</td>
<td>2.3</td>
<td>1.52</td>
<td>1.48</td>
<td>6.28</td>
<td>3.46</td>
<td>2.36</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>2.6</td>
<td>2.46</td>
<td>1.86</td>
<td>8.92</td>
<td>3.52</td>
<td>2.84</td>
</tr>
<tr>
<td>$q_{c,arma}$</td>
<td>500</td>
<td>5.62</td>
<td>4.4</td>
<td>4.2</td>
<td>9.0</td>
<td>6.58</td>
<td>5.38</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>6.32</td>
<td>4.24</td>
<td>3.64</td>
<td>10.56</td>
<td>6.18</td>
<td>5.02</td>
</tr>
<tr>
<td></td>
<td>1500</td>
<td>6.66</td>
<td>4.84</td>
<td>3.82</td>
<td>14.48</td>
<td>7.02</td>
<td>5.14</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>6.4</td>
<td>5.2</td>
<td>4.8</td>
<td>15.24</td>
<td>6.26</td>
<td>5.02</td>
</tr>
<tr>
<td>Prewhitening</td>
<td>500</td>
<td>19.98</td>
<td>14.78</td>
<td>12.3</td>
<td>12.4</td>
<td>7.04</td>
<td>4.56</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>20.64</td>
<td>12.72</td>
<td>11.44</td>
<td>15.96</td>
<td>5.76</td>
<td>4.32</td>
</tr>
<tr>
<td></td>
<td>1500</td>
<td>21.26</td>
<td>12.02</td>
<td>10.3</td>
<td>19.36</td>
<td>5.12</td>
<td>3.66</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>21.42</td>
<td>10.0</td>
<td>8.76</td>
<td>19.54</td>
<td>4.68</td>
<td>3.82</td>
</tr>
<tr>
<td><strong>NASDAQ : Model 3 → Model 4</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$q_n$</td>
<td>500</td>
<td>0.7</td>
<td>0.42</td>
<td>2.26</td>
<td>0.72</td>
<td>0.54</td>
<td>1.52</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>1.84</td>
<td>2.26</td>
<td>10.18</td>
<td>1.54</td>
<td>2.06</td>
<td>11.16</td>
</tr>
<tr>
<td></td>
<td>1500</td>
<td>2.82</td>
<td>2.96</td>
<td>11.86</td>
<td>2.2</td>
<td>3.04</td>
<td>17.84</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>3.26</td>
<td>3.54</td>
<td>13.08</td>
<td>2.02</td>
<td>3.82</td>
<td>20.8</td>
</tr>
<tr>
<td>$q_{arma}$</td>
<td>500</td>
<td>1.16</td>
<td>1.3</td>
<td>1.84</td>
<td>0.92</td>
<td>1.28</td>
<td>2.46</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>2.18</td>
<td>2.48</td>
<td>2.8</td>
<td>2.26</td>
<td>2.62</td>
<td>7.32</td>
</tr>
<tr>
<td></td>
<td>1500</td>
<td>2.98</td>
<td>3.18</td>
<td>3.44</td>
<td>2.62</td>
<td>3.42</td>
<td>10.16</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>3.68</td>
<td>3.7</td>
<td>3.96</td>
<td>2.48</td>
<td>4.22</td>
<td>11.54</td>
</tr>
<tr>
<td>$q_{c,arma}$</td>
<td>500</td>
<td>3.7</td>
<td>3.56</td>
<td>4.28</td>
<td>2.54</td>
<td>4.18</td>
<td>6.42</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>4.48</td>
<td>3.86</td>
<td>5.08</td>
<td>3.52</td>
<td>4.36</td>
<td>11.02</td>
</tr>
<tr>
<td></td>
<td>1500</td>
<td>4.98</td>
<td>4.16</td>
<td>5.66</td>
<td>3.74</td>
<td>4.62</td>
<td>13.04</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>4.84</td>
<td>4.78</td>
<td>5.14</td>
<td>3.9</td>
<td>5.28</td>
<td>13.92</td>
</tr>
<tr>
<td>Prewhitening</td>
<td>500</td>
<td>4.74</td>
<td>5.24</td>
<td>12.14</td>
<td>3.88</td>
<td>4.3</td>
<td>9.52</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>5.12</td>
<td>5.12</td>
<td>11.62</td>
<td>4.78</td>
<td>4.92</td>
<td>17.6</td>
</tr>
<tr>
<td></td>
<td>1500</td>
<td>5.38</td>
<td>5.38</td>
<td>10.78</td>
<td>4.64</td>
<td>5.8</td>
<td>21.38</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>4.6</td>
<td>5.5</td>
<td>9.36</td>
<td>4.4</td>
<td>6.52</td>
<td>23.4</td>
</tr>
</tbody>
</table>
1500, and 2000, with possible breaks at $k^* = n/4$, $n/2$, and $3n/4$. In the simulation study, the independent identically distributed residuals $Z_t$ follow standard normal distribution. In each series, the test statistic $M_n$ is computed and using the critical values in Table 3.1 a decision on whether to reject $H_0$ at $\alpha = 1\%, 5\%$ or $10\%$ is made. We repeat the process for a large number of times, $R = 5000$ in our simulations, the empirical size is then the number of rejections divided by $R$. Notice that the pattern of the variance change for DJIA and NASDAQ is volatilities increase after breaks, whereas the volatilities for GE and Wal-Mart decrease after the change-points. Once we find the procedure that gives best empirical sizes, we validate its performance by applying it to the other two models for American Express and Altria Group. Its robustness to model mis-specification will also be investigated.

The empirical sizes of the test at the nominal significance level $\alpha = 5\%$ are presented in Table 3.3, the results for $1\%$ and $10\%$ are reported in Tables C.1 and C.2, respectively. The relative performance of the four variance estimators are consistent across different target sizes, so in the following we focus on the most often used $5\%$ level. We first observe that the empirical sizes for the bandwidths $q_n$ are very unstable, varying between $0.42\%$ to $39.36\%$. One concludes that the richness of the dependence structure of squared GARCH(1,1) observations is far more than lags that are just proportional to the sample sizes can capture. Lack of stability is also the main drawback of the method of prewhitening. It yields good sizes that are close enough to the target levels for some models with different positions of the break $k^*$. However, the overrejections are at least two times of $5\%$ in Table 3.3 when the transition is from Model 1 to Model 2 that are estimated from the returns on DJIA.

Generally speaking, the empirical sizes corresponding to the lags $q_{arma}$ are always well under the nominal levels, only with some exceptions when the changes are from Model 3 to Model 4 for NASDAQ with $k^* = n/4$ and from Model 7 to Model 8 for Wal-Mart with $k^* = 3n/4$. From how the test statistic $M_n$ is constructed, one can see that under-rejections are caused by over-estimations of the long-run variance, that is, both $s_{s_1}^2$ and $s_{s_2}^2$ in (3.3.7) and (3.3.8), respectively, are too large, which in turn leads us to conclude that the lags $q_{arma}$ are too long to yield good empirical sizes. We thus propose the bandwidth parameter
that is the product of $q_{\text{arma}}$ and a factor $c$ less than 1. A constant, say, $c = 0.5$, for time series of varying lengths with different locations of change-point, does not work well in our simulations. We may expect to get good rejection rates when, for instance, the length is 1000 with $k^* = 500$ using $c = 0.5$, but the empirical size for $n = 2000$ and $k^* = 1500$ may turn out to be two times of the target level. The reason lies in the fact that sample size and location of breaks jointly influence the empirical sizes. The factor $c$ thus need to take into account of both $n$ and $k^*$. Two plots are presented in Figure 3.2 to illustrate the effects of $n$ and $k^*$ on the empirical sizes, and also to show the improvements of using lags $q_{\text{c arma}}$ over $q_{\text{arma}}$. We now take on the rationale in the expressions of $c$ in (3.4.14) and (3.4.15) that we left over in Section 3.4. It is clear that when lags are determined by $q_{\text{arma}}$ sample size $n$ and empirical rejection rate are positively correlated, i.e., the more observations the higher the frequency of rejection. To alleviate this effect, we thus need to use relatively wider bandwidth to taper the test statistic $M_n$ and get less rejections for longer series. The gradually increasing function $\log_{10}(n/100)$ in (3.4.14) and (3.4.15) is used for that purpose. Now focus on the right panel of Figure 3.2 where the model change is from 1 to 2 for DJIA with volatilities increased after breaks. Notice that as $k^*$ increases toward the end of the series, in spite of the value of $n$, the empirical sizes decrease. In contrast, in the left-hand plot where the volatilities decrease after breaks, related to the change-point model for GE, we observe the exactly opposite situation that increasing the value of $k^*$ increases the empirical sizes. Thus, regarding increase or decrease of volatilities after breaks, we should use different $c$. Specifically, for change-point models with volatilities increased after breaks, we use shorter lags to increase the value of $M_n$ and get relatively more rejections for greater $k^*$. That is why, when $\text{Var}(X_{n_1}) \leq \text{Var}(X_{n_2})$, the term $(n_1/n + 1.5)^{-1}$ in the expression of $c$ in (3.4.14) is used, whose value decreases for larger $n_1 = \hat{k}$, the estimate of the change-point $k^*$. Clearly, we should do the opposite for models with volatility decreased after breaks, i.e., when $\text{Var}(X_{n_1}) > \text{Var}(X_{n_2})$ we extend bandwidth for greater $k^*$. In (3.4.15), the term $(n_2/n + 1.5)^{-1}$ is used instead in the determination of $c$, whose value increases, recalling that $n_2 = n - n_1$, when the estimated break gets farther away from the beginning of the
Figure 3.2. Comparison of empirical sizes for truncation lags $q_{\text{arma}}^*$ and $q_{\text{arma}}^*$. Results are reported for the change-point models for DJIA and GE. The straight line indicates the nominal size of 5%.

The advantage of using lags $q_{\text{arma}}^*$ as opposed to $q_{\text{arma}}^*$ is clear in Figure 3.2. The empirical sizes yielded by the calibrated bandwidth are very close to the nominal level of 5%, with positive and negative disparities less than 2%. Unfortunately, there are some serious over-rejections in Table 3.3. When the changes are from Model 3 to Model 4 for NASDAQ with $k^* = n/4$ and from Model 7 to Model 8 for Wal-Mart with $k^* = 3n/4$, we see rejection rates as high as 15.24% for nominal 5% level. See Figure C.1 for a graphical illustration, where similar plots to those in Figure 3.2 are presented for the change-point models for NASDAQ and Wal-Mart. Indeed, we expect over-rejections as severe as this by recalling that, even for longer lags $q_{\text{arma}}^*$ with under-rejections as the dominating pattern, we get unusually high rejection frequencies for the two particular change-point models. Similar over-rejections are also observed in Tables C.1 and C.2 where target sizes are 1% and 10%, respectively. Overall, among the four variance estimators considered in our simulations, we recommend the use of the Bartlett kernel estimator with truncation lags determined by $q_{\text{arma}}^*$. Its performance however is not very reliable, in the sense that it may deliver too
Table 3.4. Empirical sizes (in percent) of the test using the Bartlett kernel estimator with lags $q_{c,\text{arma}}$ applied to simulated series of squared GARCH(1,1) observations following the change-point models estimated from the returns on American Express and Altria Group in Table 3.2. The number of replications is $R = 5000$.

<table>
<thead>
<tr>
<th>Stock</th>
<th>American Express</th>
<th>Altria Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model change</td>
<td>(Model 9 → Model 10)</td>
<td>(Model 11 → Model 12)</td>
</tr>
<tr>
<td>n</td>
<td>873, $k^* = 475$</td>
<td>631, $k^* = 389$</td>
</tr>
<tr>
<td>Var. change</td>
<td>8.288 → 2.068</td>
<td>3.558 → 8.482</td>
</tr>
<tr>
<td>Nominal level</td>
<td>1% 5% 10%</td>
<td>1% 5% 10%</td>
</tr>
<tr>
<td>Empirical size</td>
<td>0.64 4.58 11.06</td>
<td>1.26 4.62 9.82</td>
</tr>
</tbody>
</table>

much over-rejections when the change-point happens in the vicinity of either the start or the end of the series. With this limitation in mind, we conclude that the procedure can yield reasonably good empirical sizes when the location of the break is close to the middle of the time series.

We next validate the kernel estimator with lags $q_{c,\text{arma}}$ by applying it to the models for American Express and Altria Group in Table 3.2. We, as before, generate series of squared GARCH(1,1) observations accordingly, but, instead of considering different sample sizes with varying locations of the breaks, we let $n$ and $k^*$ equal to those obtained from the time series. The empirical sizes, based on $R = 5000$ replications, at significance levels 1%, 5%, and 10% are presented in Table 3.4. The procedure works quite well that for $\alpha = 5\%$ we get under-rejections that are less than 0.5% away from the target. We have both variance increase and decrease in the two models, and the breaks occur approximately halfway of the series. The results in Table 3.4 thus confirms the conclusion we made above about the applicability of the estimator using lags $q_{c,\text{arma}}$.

Since the bandwidth parameter $q_{c,\text{arma}}$ is based on the estimation of a GARCH model, a usual criticism of such a parametric method is that mis-specification of the model can lead to large under-rejections or over-rejections. We address this issue by assessing the empirical sizes of the procedure under different conditional heteroskedastic processes. Specifically, we consider two asymmetric GARCH(1,1) models, the Exponential GARCH (EGARCH)
Table 3.5. Empirical sizes (in percent) of test using the Bartlett kernel estimator with lags $q_{arma}$ applied to simulated series of squared EGARCH and TGARCH observations following the change-point models estimated from the returns on DJIA, Altria Group, GE, and American Express in Table C.3. The number of replications is $R = 5000$.

<table>
<thead>
<tr>
<th>Asymmetric GARCH</th>
<th>Model change</th>
<th>n</th>
<th>$k^*$</th>
<th>Nominal level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1%</td>
</tr>
<tr>
<td>EGARCH</td>
<td>A $\rightarrow$ B (DJIA)</td>
<td>2021</td>
<td>1061</td>
<td></td>
</tr>
<tr>
<td>(1,1)</td>
<td>C $\rightarrow$ D (Altria Group)</td>
<td>631</td>
<td>389</td>
<td></td>
</tr>
<tr>
<td>TGARCH</td>
<td>E $\rightarrow$ F (GE)</td>
<td>875</td>
<td>473</td>
<td></td>
</tr>
<tr>
<td>(1,1)</td>
<td>G $\rightarrow$ H (American Express)</td>
<td>873</td>
<td>475</td>
<td></td>
</tr>
</tbody>
</table>

and the Threshold GARCH (TGARCH), that are widely used in practice. The two models are developed to incorporate the asymmetric news impact, also referred to as the leverage effect, on the volatility of financial time series, that is, volatility tends to rise in response to negative shocks and tends to fall in response to positive shocks.

We introduce the two models by noting that they differ from the standard GARCH models only in the definition of the conditional variance $\sigma_t^2$, as given in Eq. (3.4.11) for GARCH(1,1) model. For EGARCH(1,1) model proposed in [137], the conditional variance is defined as follows,

$$\sigma_t^2 = e^{ht}$$

where

$$ht = \omega + \alpha (|Z_{t-1}| + \gamma Z_{t-1}) + \beta h_{t-1}.$$  

Note that the total effect of the lagged error $Z_{t-1}$ is $(1 + \gamma)|Z_{t-1}|$ when $Z_{t-1}$ is positive or there is "good news", and is $(1 - \gamma)|Z_{t-1}|$ when $Z_{t-1}$ is negative or there is "bad news." Since bad news can have a larger impact on volatility, the value of $\gamma$ would be expected to be negative.

The TGARCH model is also known as the GJR-GARCH model because [75] essentially proposed the same model. The conditional variance for TGARCH(1,1) is given by

$$\sigma_t^2 = \omega + (\alpha + \gamma I(Z_{t-1})) r_{t-1}^2 + \beta \sigma_{t-1}^2,$$
Table 3.6. Empirical power (in percent) of the test using the Bartlett kernel estimator with lags $q_{c,arma}$ applied to simulated series of squared LARCH observations with parameters $d = 0.35$, $a = 0.85$, and $b_0 = 0.35$. The number of replications is $R = 1000$.

<table>
<thead>
<tr>
<th>Sample size</th>
<th>Nominal level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1%</td>
</tr>
<tr>
<td>500</td>
<td>13.0</td>
</tr>
<tr>
<td>1000</td>
<td>18.0</td>
</tr>
<tr>
<td>1500</td>
<td>20.7</td>
</tr>
<tr>
<td>2000</td>
<td>21.4</td>
</tr>
</tbody>
</table>

where $I(Z_{t-1}) = 1$ if $Z_{t-1} < 0$, and $I(Z_{t-1}) = 0$ otherwise. Note that when the lagged residual $Z_{t-1}$ is above the threshold value 0 the effects of $r_{t-1}^2$ is $\alpha r_{t-1}^2$ and when it is below the threshold the effects are given by $(\alpha + \gamma) r_{t-1}^2$. So one would expect $\gamma$ to be positive because bad news tends to increase volatility.

We estimate EGARCH(1,1) model from the returns on DJIA and Altria Group and estimate TGARCH(1,1) model from the returns on GE and American Express. The four data sets, with their corresponding break points, are exactly the same as those used to estimate GARCH(1,1) in Table 3.2. The results are reported in Table C.3. We notice, as expected, that the estimates of the leverage parameter $\gamma$ are negative for the EGARCH models and are positive for the TGARCH models. We next apply the Bartlett kernel estimator with lags $q_{c,arma}$ to the simulated series of squared EGARCH(1,1) and TGARCH(1,1) observations following the models in Table C.3, with $n$ and $k^*$ equal to those obtained from the time series. The empirical sizes of the test based on $R = 5000$ replications are given in Table 3.5. The rejection frequencies are rather close to the nominal levels, the differences are less than 1.5% for $\alpha = 5\%$ and less than 3% for $\alpha = 10\%$. One can thus conclude that the testing procedure is not sensitive to model mis-specification, and its performance is robust to a variety of conditional heteroskedastic processes.

In what follows we focus on the test using the kernel estimator with lags $q_{c,arma}$, as it performs better than the other tests and is robust to model mis-specification. We now examine the empirical power of the testing procedure. We use the LARCH (Linear ARCH)
model as the long range dependent model satisfying $H_A$, which also exhibits conditional heteroskedasticity. The returns $r_t$ are said to follow a LARCH model if

\[(3.6.20) \quad r_t = \sigma_t Z_t, \quad \sigma_t = a + \sum_{j \geq 1} b_j r_{t-j},\]

where $a \neq 0$, the $b_j$ are real coefficients that are not necessarily nonnegative and the $Z_t$ are independent identically distributed with zero mean and finite fourth moment. The long-memory property is obtained by assuming that $b_j \sim c j^{d-1}$, for some $0 < d < 1/2$ and $c \neq 0$, and "~" indicates that the ratio of left and right sides tends to 1 as $j \to \infty$. The LARCH model was studied by [154], Giraitis et al. ([72], [74], [73]) and [13], among others.

We generate series of squared LARCH observations with $d = 0.35$ and the $b_j$ recursively computed according to $b_j = [b_{j-1}(j + d)]/(j + 1)$ with $b_0 = 0.35$ and $a = 0.85$. These parameter values ensure that the process $r_t$ is fourth order stationary and are chosen by experimentation to make the realizations similar to the returns data. Systematic parameter estimation method have not yet been developed for the LARCH model. The empirical power of the procedure with sample sizes $n = 500, 1000, 1500, \text{ and } 2000$ is reported in Table 3.6. The number of replications is $R = 1000$. We notice that the power increases slightly as sample size $n$ increases. The results are seen to be satisfactory, given that the trajectories of the LARCH process with the parameters selected above typically look like the trajectories of a GARCH process with a change-point. Figure C.2 presents four LARCH series with different lengths, from which we can see that they very much resemble the plots in Figure 3.1 of the daily log-returns for which change point models are believed to be the underlying data generating processes.

We conclude this section by presenting in Table 3.7 the results of applying the test procedure to the six set of returns data we used to estimate the volatility change models satisfying $H_0$. Recall in Table 3.1 that the critical value for $\alpha = 10\%$ is 1.36. The realized values of the test statistic $M_n$ are all less than 1.36, meaning the the test fails to reject the $H_0$ that there is a volatility change in the observations at all the three commonly used significance levels. Moreover, we notice that the statistic $M_n$ for the returns on American
Table 3.7. Realized test statistic $M_n$'s using the Bartlett kernel estimator with truncation lag $q^*_{c,arma}$ for the six sets of real returns data shown in Figure 3.1.

<table>
<thead>
<tr>
<th>Speculative Asset</th>
<th>$M_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>DJIA</td>
<td>1.304</td>
</tr>
<tr>
<td>NASDAQ</td>
<td>1.353</td>
</tr>
<tr>
<td>General Electric</td>
<td>1.059</td>
</tr>
<tr>
<td>Wal-Mart</td>
<td>1.013</td>
</tr>
<tr>
<td>American Express</td>
<td>0.882</td>
</tr>
<tr>
<td>Altria Group</td>
<td>1.280</td>
</tr>
</tbody>
</table>

Express is the smallest, indicating a more significant evidence of volatility change. This is expected, because the variance change for that asset is the largest in Table 3.2 in terms of both absolute change and scale of change.

Our investigation indicates that a change point model may be preferable to a long memory model for the squared returns. It must however be kept in mind that the power of the test is relatively small and models like LARCH may be capable of explaining important facets of the behavior of returns.
CHAPTER 4
MONITORING CONSTANCY OF VARIANCE IN CONDITIONALLY HETEROSKEDASTIC TIME SERIES

4.1 Introduction

This chapter is concerned with on-line detection of a change in unconditional variance in a conditionally heteroskedastic time series. Our sequential testing procedures are similar in spirit to the procedures developed by [39] and [99] who considered monitoring for changes in linear regression models. To explain the idea, suppose \( r_t, \ t = 1, 2, \ldots \) are returns on a speculative asset. We assume that the first \( m \) observations \( r_1, \ldots, r_m \) are a realization of a stationary process and, in particular, the unconditional variance does not change up to time \( m \). We then monitor the observations \( r_{m+1}, r_{m+2}, \ldots \) as they arrive, and at each time \( n > m \) we have to make a decision whether \( r_1, \ldots, r_n \) are a realization of the same process as \( r_1, \ldots, r_m \). The detection algorithms are designed in such a way that they will detect changes which lead to a change in the unconditional variance.

The problem of detecting structural changes in conditionally heteroskedastic time series has received increased attention in recent years. Most work has, however, focused on the so called \textit{a posteriori} change-point problem in which a historical sample of fixed size is given and a decision has to be made whether one model is suitable for the whole sample. There are many variants of this problem depending on whether one is interested in the constancy of the parameters of a parametric model or merely in the constancy of specified moments. Contributions in this direction have been made by [103], [38], Kokoszka and Leipus ([108] [109], [110], [111]), [107], [104], Andreou and Ghysels ([3], [4]), [115], among others. By contrast, very few contributions have been made to the problem of sequential testing for a structural change in conditionally heteroskedastic time series. Mikosch and Stărică ([130], [132]) suggested a sequential change-point detection method based on the periodogram whereas [12] developed a method based on likelihood scores which applies to the same testing setting as in [39] but focuses on GARCH\((p, q)\) models rather than linear regression models.
The methods proposed in this chapter are essentially nonparametric, even though their specific implementations may require estimating an approximate parametric model for the data. This is because the detectors have a general form $D(X_1, \ldots, X_n)/\hat{\sigma}$, where $\hat{\sigma}^2$ is an estimator of the variance of the sample mean of dependent stationary observations. While the function $D(\cdot)$ is fully specified and depends only on the observations, some estimators $\hat{\sigma}^2$ may rely on model assumptions. In our simulation study, we use a kernel estimator of $[5]$ which requires a bandwidth selection procedure. The data-driven procedure of [5] relies on postulating and estimating an approximate parametric model for the observations.

Theoretical justification for the procedures proposed here applies to broad classes of conditionally heteroskedastic time series. The main assumption is that the squared returns obey a strong invariance principle, see Equation (4.2.4). Such a very general assumption is possible because recent research has established that practically all heteroskedastic models of importance in econometrics obey the strong invariance principle (4.2.4), see [35] and further references in Section 4.2. Using a strong approximation leads to straightforward proofs which avoid the often very intricate arguments used in [39] who based their theory on the weak convergence of measures in the Skorokhod space.

The chapter is organized as follows. After formulating the monitoring problem and presenting some further background in Section 4.2, we describe the detection schemes and establish their asymptotic properties in Section 4.3. Section 4.4 and 4.5 contain the proofs of Theorems stated in Section 4.3. Section 4.6 briefly describes the variance estimator used in simulations presented in Section 4.8. Section 4.7 presents an empirical analysis of returns on four major US stock indexes which motivates the subsequent simulations.

4.2 Problem Formulation and Assumptions

Assume that the returns $r_t$ have mean zero and denote $X_t = r_t^2$, $t = 1, 2, \ldots$, so that $\omega_t = EX_t$ is the variance of the $t$th return.

We assume that

$$
\omega_1 = \omega_2 = \ldots = \omega_m =: \omega_0.
$$

(4.2.1)
We wish to test

\[(4.2.2) \quad H_0 : \omega_n = \omega_0 \text{ for } n \geq m + 1\]

against

\[(4.2.3) H_A : \exists n^* \geq m + 1 \omega_n = \omega_0 \text{ for } n \leq n^* - 1 \quad \text{and} \quad \omega_n = \omega_* \neq \omega_0 \text{ for } n \geq n^* \]

We assume that under \(H_0\) the \(X_i\) satisfy the following strong invariance principle.

\[(4.2.4) \quad \sum_{1 \leq i \leq n} (X_i - \omega_0) - \sigma W(n) \overset{a.s.}{=} o(n^\alpha)\]

with some \(0 < \alpha < 1/2\).

The following theorem gives sufficient conditions for \((4.2.4)\) to hold.

**Theorem 4.2.1** Let \(\{X_k\}\) be a weakly stationary sequence of random variables with mean zero and uniformly bounded \((2 + \delta)\)-th moments for some \(0 < \delta \leq 2\). Assume that \(\{X_k\}\) satisfies the strong mixing condition

\[(4.2.5) \quad \sup_{A \in \mathcal{F}_m, B \in \mathcal{F}_{m+n}} |P(A \cap B) - P(A)P(B)| \leq n^{-\gamma} \quad \text{for all } m, n \geq 1\]

with \(\gamma \geq 300(1 + 2/\delta)\), where \(\mathcal{F}_k\) denotes the \(\sigma\)-field generated by \(X_k, \ldots, X_\ell\). Then letting \(S_n = X_1 + \ldots + X_n\), the limit

\[\sigma^2 = \lim_{n \to \infty} \frac{1}{n} E S_n^2\]

exists, and if \(\sigma > 0\), then there exists a Wiener process \(\{W(t), 0 \leq t < \infty\}\) such that

\[S_n - \sigma W(n) \overset{a.s.}{=} O(n^{1/2-\varepsilon}),\]

where \(\varepsilon = \delta/600\).

There are several theorems of this type. Theorem 4.2.1 is due to [145], see their Theorem 8.1. on p. 96. The constants in the theorem are far from optimal, however, the theorem of Philipp and Stout covers most applications.

Recently, [35] established easily verifiable sufficient conditions for exponential \(\beta\)-mixing (absolute regularity) and the existence of finite moments in several important models extending the standard GARCH(1,1). Since exponential \(\beta\)-mixing implies exponential strong
mixing, see eg. [27], eq. (1.7), these conditions imply (4.2.5) whenever $X_k = f(r_k)$ and the $r_k$ follow one of the GARCH(1,1)-type models considered by [35], $f(\cdot)$ being any measurable function. Necessary and sufficient conditions for the existence of higher order moments in these models were established by [94]. The results of [35] also imply that condition (4.2.5) holds if $X_k = f(r_k)$ and the $r_k$ follow a GARCH($p,q$) process which satisfies $\sum_i \alpha_i + \sum_j \beta_j < 1$ and the innovations have a density which is continuous and positive on the whole line, see their Proposition 12.

As a corollary of the above discussion, we conclude that the Theorems in Section 4.3 hold, in particular, if $X_t = r_t^2$ and the $r_t$ follow a strictly stationary process which is strongly mixing with exponential rate and has finite $(4 + \delta)$th moment.

4.3 Detection Algorithms

We now introduce four sequential monitoring methods considered in this chapter. These methods are formulated in terms of the observations $X_t$ which can be interpreted as the squares of mean zero returns. With such an interpretation, the methods are designed to detect a change in unconditional variance of returns. Even though this has been our primary motivation, the proposed algorithms have a much wider applicability. In fact, the $X_t$ can be viewed as a sequence of observations satisfying a week dependence condition (invariance principle (4.2.4)) and the monitoring is then intended to detect a change in the mean of the $X_t$.

The monitoring schemes considered in Section 4.3.1 are motivated by the methods proposed by [39] in the context of detecting parameter changes in linear regression models. The two methods described in Section 4.3.2 are based on procedures proposed in [99], also in the context of linear regression models.

In the following, we use the notation

$$\xi_t = X_t - \omega_t,$$

and let \{W(t), 0 \leq t < \infty\} denote the standard Brownian motion.
4.3.1 CUSUM and fluctuation monitoring schemes

CUSUM monitoring is based on the detector

\[(4.3.6) \quad S_n = \sum_{i=2}^{n} \nu_i^{-1/2}(X_i - \bar{X}_{i-1}), \quad n \geq 2,\]

where

\[(4.3.7) \quad \nu_i = \frac{i}{i-1}, \quad i \geq 2\]

and

\[(4.3.8) \quad \bar{X}_{i-1} = \frac{1}{i-1} \sum_{j=1}^{i-1} X_j, \quad i \geq 2.\]

The conditions imposed on the boundary function \(g(\cdot)\) appearing in Theorems 4.3.2 and 4.3.3 below are collected in the following assumption.

**Assumption 4.3.1** The function \(g : [1, \infty) \to \mathbb{R}\) satisfies

\[(4.3.9) \quad g(\cdot) \text{ is continuous on } [1, \infty);\]

\[(4.3.10) \quad g(t) > ct^{1/2}, \quad 1 \leq t < \infty, \quad \text{for some } c > 0\]

and

\[(4.3.11) \quad \lim_{t \to \infty} \frac{|W(t)|}{g(t)} = 0.\]

Sufficient conditions for (4.3.11) to hold can be obtained from the results formulated in Section 4.1 of [43]. For convenience, we state them in the following proposition.

**Proposition 4.3.1** If \(g : [0, \infty) \to \mathbb{R}\) satisfies the following conditions

\[(4.3.12) \quad \inf_{1 \leq t \leq c} g(t) > 0 \quad \text{for all } t > 1,\]

\[(4.3.13) \quad g(t)/t \text{ is non-increasing on } [c, \infty) \text{ for some } c > 0\]

and

\[(4.3.14) \quad \int_{1}^{\infty} \frac{1}{t} \exp(-cg^2(t)/t)dt < \infty \text{ for all } c > 0,\]

then (4.3.11) holds.
The assumptions we impose on the function $g(\cdot)$ are different from those used by [39] who required that $g(\cdot)$ be regular (in the sense defined on p. 1050 of [39]) and $t^{-1/2}g(t)$ be eventually nondecreasing. Assumption (4.3.9) implies regularity whereas, in the class of eventually positive functions, if $t^{-1/2}g(t)$ is eventually nondecreasing, then assumption (4.3.10) holds. The assumptions formulated in Proposition 4.3.1 were not needed by [39]), but are satisfied by most functions of practical interest. In the empirical applications, we will work with the function

$$g_a(t) = [t(a^2 + \ln t)]^{1/2}, \quad t \geq 1,$$

which meets Assumption 4.3.1. Conditions (4.3.9), (4.3.10), (4.3.12) and (4.3.13) are obviously satisfied. Since

$$\int_1^\infty \frac{1}{t} \exp(-cg_a^2(t)/t)dt = \int_1^\infty e^{-\left(a^2 \frac{1}{t^{1+c}}\right)} dt,$$

assumption (4.3.14) is also met.

It is known that, see [39], p. 1052,

$$P\{|W(t)| \geq g_a(t), \text{ for some } t \geq 1\} = 2[1 - \Phi(a) + a\phi(a)],$$

where $\Phi$ and $\phi$ are, respectively, the cdf and pdf of a standard normal random variable. For boundary crossing probabilities (4.3.16) of 5% and 10%, $a^2$ equals to 7.78 and 6.25, respectively.

Theorem 4.3.2, whose proof is given in Section 4.4, justifies a monitoring scheme based on the detector (4.3.6).

**Theorem 4.3.2** Suppose assumption (4.2.4) holds and the function $g(\cdot)$ satisfies Assumption 4.3.1. Then, under $H_0$,

$$\sup_{m \leq n < \infty} \frac{|S_n|}{m^{1/2}g_a(n/m)} \xrightarrow{d} \sup_{1 \leq t < \infty} \frac{\sigma|W(t)|}{g_a(t)}.$$

Using the function $g_a(\cdot)$ in Theorem 4.3.2 and a consistent estimator $\hat{\sigma}$ of $\sigma$, we obtain the following rejection rule: reject $H_0$ if

$$\frac{1}{\hat{\sigma}} \left| \sum_{i=2}^{n} \nu^{-1/2}_i (X_i - \bar{X}_{i-1}) \right| > m^{1/2} \left[ \frac{n}{m} \left( a^2 + \ln \frac{n}{m} \right) \right]^{1/2}, \text{ for some } n > m.$$
By (4.3.17) and (4.3.16), as $m \to \infty$, the probability of falsely rejecting $H_0$ thus tends to a prescribed significance level $\alpha$ which is controlled by the constant $a$.

We now turn to the fluctuation detector defined as

$$Z_n = n\sigma^{-1}(\bar{X}_n - \bar{X}_m), \quad n > m,$$

with $\bar{X}_n$ as defined in (4.3.8).

The following theorem provides a justification for the monitoring scheme based on the fluctuation detector (4.3.19).

**Theorem 4.3.3** Suppose assumption 4.2.4 holds and the function $g(\cdot)$ satisfies Assumption 4.3.1. Then, under $H_0$,

$$\sup_{m < n < \infty} \frac{|Z_n|}{m^{1/2} \left( \frac{n-m}{m} \right) g_\alpha \left( \frac{n}{n-m} \right)} \overset{d}{\to} \sup_{1 < t < \infty} \frac{|W(t)|}{g_\alpha(t)}.$$

Theorem 4.3.3 is proved in Section 4.4.

Similarly as in CUSUM monitoring, Theorem 4.3.3 leads to the following rejection rule:

reject $H_0$ if for some $n > m$

$$\frac{n}{\sigma} |\bar{X}_n - \bar{X}_m| > m^{1/2} \left( \frac{n-m}{m} \right) \left[ \frac{n}{n-m} \left( a^2 + \ln \frac{n}{n-m} \right) \right]^{1/2}.$$

4.3.2 Monitoring schemes based on partial sums of residuals and recursive residuals

The following two kinds of residuals are used to construct detectors:

$$\hat{\xi}_i = X_i - \bar{X}_m, \quad i > m$$

and

$$\tilde{\xi}_i = X_i - \bar{X}_{i-1}, \quad i \geq 2.$$

First, define the detector

$$\hat{S}(m, k) = \sum_{i=m+1}^{m+k} \hat{\xi}_i, \quad k \geq 1,$$
and consider the boundary function

\[(4.3.24)\]  
\[g_\gamma(m,k) = \left(1 + \frac{k}{m}\right) \left(\frac{k}{m+k}\right)^\gamma, \quad 0 \leq \gamma < \frac{1}{2} - \alpha.\]

Define the critical value \(c_\alpha(\gamma)\) by

\[P\left\{ \sup_{0 \leq t \leq 1} \frac{|W(t)|}{t^\gamma} \geq c_\alpha(\gamma) \right\} = \alpha.\]

The critical values \(c_\alpha(\gamma)\) are tabulated in Table 1 of [99].

**Theorem 4.3.4** If \((4.2.4)\) holds, then under \(H_0\)

\[\lim_{m \to \infty} P\left\{ \sup_{1 \leq k < \infty} \frac{\hat{S}(m,k)}{\sqrt{mg_\gamma(m,k)}} \geq c_\alpha(\gamma) \right\} = P\left\{ \sup_{0 \leq t \leq 1} \frac{\sigma|W(t)|}{t^\gamma} \geq c_\alpha(\gamma) \right\}.

Theorem 4.3.4, which is proved in Section 4.5, leads to the following rejection rule: reject \(H_0\) if

\[(4.3.25)\]  
\[\frac{1}{\sigma} \left| \sum_{i=m+1}^{m+k} (X_i - \bar{X}_m) \right| \geq c_\alpha(\gamma) m^{1/2} \left(1 + \frac{k}{m}\right) \left(\frac{k}{m+k}\right)^\gamma, \text{ for some } k \geq 1,
\]

or, by letting \(n = m + k\), reject if

\[(4.3.26)\]  
\[\frac{1}{\sigma} \left| \sum_{i=m+1}^{n} (X_i - \bar{X}_m) \right| \geq c_\alpha(\gamma) m^{1/2} \left(\frac{n}{m}\right)^{1/2} \left(\frac{n-m}{n}\right)^\gamma, \text{ for some } n > m.
\]

Finally, define the detector

\[(4.3.27)\]  
\[\hat{S}(m,k) = \sum_{i=m+1}^{m+k} \hat{\xi}_i, \quad k \geq 1.
\]

We now denote the boundary function by \(h(t), t \geq 0\), and assume that \(h(t) = g(t+1)\) for a function \(g(u), u \geq 1\) which satisfies Assumption 4.3.1. Thus \(h(\cdot)\) satisfies the following assumption.

**Assumption 4.3.2** The function \(h(\cdot)\) satisfies the following conditions:

\[(4.3.28)\]  
\[h(\cdot) \text{ is continuous on } [0, \infty);\]

\[(4.3.29)\]  
\[h(t) > c(t+1)^{1/2}, \quad 0 \leq t < \infty, \quad \text{for some } c > 0\]

and

\[(4.3.30)\]  
\[\lim_{t \to \infty} \frac{|W(t)|}{h(t)} = 0 \text{ a.s.}\]
Theorem 4.3.5  Suppose assumption (4.2.4) holds and the function \( h(\cdot) \) satisfies Assumption 4.3.2. Then, under \( H_0 \),

\[
\sup_{1 \leq k < \infty} \frac{|\hat{S}(m, k)|}{m^{1/2}h(k/m)} \overset{d}{\to} \sup_{0 < t < \infty} \frac{\sigma|W(t)|}{h(t)}.
\]

Theorem 4.3.5 is proved in Section 4.5.

The discussion in Section 4.3.1 shows that the boundary function

\[
h_a(t) = (t + 1)^{1/2}[a^2 + \ln(t + 1)]^{1/2}
\]

(4.3.31)

satisfies Assumption 4.3.2. It is known that, see Eq.(8) of [39],

\[
P\left\{ \sup_{0 < t < \infty} \frac{|W(t)|}{h_a(t)} \geq 1 \right\} = \exp(-a^2/2).
\]

(4.3.32)

For the asymptotic false alarm rate (4.3.32) of 5% and 10%, \( a^2 \) equals to 6.0 and 4.6, respectively.

Theorem 4.3.5 and (4.3.32) lead to the rejection rule: reject \( H_0 \) if for some \( k \geq 1 \)

\[
\left| \frac{1}{s} \sum_{i=m+1}^{m+k} (X_i - \bar{X}_{i-1}) \right| \geq \sqrt{m} \left( 1 + \frac{k}{m} \right)^{1/2} \left( a^2 + \ln \left( 1 + \frac{k}{m} \right) \right)^{1/2},
\]

(4.3.33)

or, by letting \( n = m + k \), reject if

\[
\left| \frac{1}{s} \sum_{i=m+1}^{n} (X_i - \bar{X}_{i-1}) \right| \geq \sqrt{m} \left( \frac{n}{m} \right)^{1/2} \left( a^2 + \ln \left( \frac{n}{m} \right) \right)^{1/2}, \text{ for some } n > m.
\]

(4.3.34)

4.4 Proofs of Theorems 4.3.2 and 4.3.3

The proof of Theorem 4.3.2 relies on several lemmas and Proposition 4.4.4.

Lemma 4.4.1  If (4.2.4) holds, then

\[
\sum_{1 \leq i \leq n} (X_i - \omega_0) = O \left( (n \log \log n)^{1/2} \right) \quad a.s.
\]

Proof: The Lemma follows from assumption (4.2.4) and the law of the iterated logarithm for the Wiener process.
Lemma 4.4.2 If (4.2.4) holds, then

\[ |X_n| = o(n^\alpha) \quad \text{a.s.} \]

**Proof:** Observe that

\[
|X_n| \leq |X_n - \omega_0| + |\omega_0|
\]

\[
\leq \left| \sum_{i=1}^{n}(X_i - \omega_0) - \sigma W(n) \right| + \left| \sum_{i=1}^{n-1}(X_i - \omega_0) - \sigma W(n-1) \right|
\]

\[
+ \sigma |W(n) - W(n-1)| + |\omega_0|
\]

\[ \overset{a.s.}{=} o(n^\alpha) + \sigma|W(n) - W(n-1)|. \]

Since

\[
P \left[ |W(n) - W(n-1)| \geq (10 \log n)^{1/2} \right] = O(n^{-2}),
\]

by the Borel-Cantelli lemma

\[ |W(n) - W(n-1)| \overset{a.s.}{=} O \left( (\log n)^{1/2} \right). \]

This completes the proof of Lemma 4.4.2.

We will use the decomposition

(4.4.35) \[ S_n = \sum_{2 \leq i \leq n} (X_i - \bar{X}_{i-1}) + \sum_{2 \leq i \leq n} \left[ \left( 1 - \frac{1}{i} \right)^{1/2} - 1 \right] (X_i - \bar{X}_{i-1}). \]

Lemma 4.4.3 If (4.2.4) holds, (4.4.35) satisfies

\[
\sum_{2 \leq i \leq n} \left[ \left( 1 - \frac{1}{i} \right)^{1/2} - 1 \right] (X_i - \bar{X}_{i-1}) \overset{a.s.}{=} o(n^\alpha).
\]

**Proof:** By the mean value theorem, we have

\[
\left| \sum_{2 \leq i \leq n} \left[ \left( 1 - \frac{1}{i} \right)^{1/2} - 1 \right] (X_i - \bar{X}_{i-1}) \right|
\]

\[
= \left| \sum_{2 \leq i \leq n} \left[ \left( 1 - \frac{1}{i} \right)^{1/2} - 1 \right] (X_i - \omega_0 - (\bar{X}_{i-1} - \omega_0)) \right|
\]
By Lemma 4.4.2,
\[ \sum_{2 \leq i \leq n} \frac{1}{i} |X_i - \omega_0| = o(n^\alpha). \]

By Lemma 4.4.1,
\[ |\bar{X}_{i-1} - \omega_0| \overset{a.s.}{=} O \left( \frac{\log \log i}{i} \right)^{1/2} \]
and so
\[ \sum_{2 \leq i \leq n} \frac{1}{i} |\bar{X}_{i-1} - \omega_0| \overset{a.s.}{=} O(1) \sum_{2 \leq i \leq n} \left( \frac{\log \log i}{i^3} \right)^{1/2} = O(1). \]
This completes the proof of Lemma 4.4.3.

Introduce the process

\[ W^*(t) = W(t) - \int_1^{t+1} \frac{1}{x} W(x) dx, \quad t \geq 0. \]

Computing the covariances shows that \( W^*(\cdot) \) is a standard Wiener process.

**Proposition 4.4.4** If condition (4.2.4) holds, then
\[ S_n - \sigma W^*(n) \overset{a.s.}{=} o(n^\alpha). \]

**Proof:** Using decomposition (4.4.35) and Lemma 4.4.3, we observe that
\[ S_n \overset{a.s.}{=} \sum_{2 \leq i \leq n} (X_i - \omega_0) - \sum_{2 \leq i \leq n} \frac{1}{i-1} \sum_{1 \leq j \leq i-1} (X_j - \omega_0) + o(n^\alpha). \]

Note that
\[ \sum_{2 \leq i \leq n} \frac{1}{i-1} \sum_{1 \leq j \leq i-1} (X_j - \omega_0) = \int_1^n \frac{1}{[x]} U(x) dx, \]
where \([x]\) is the integer part of \( x \) and
\[ U(x) = \sum_{1 \leq j \leq x} (X_j - \omega_0). \]

It is easy to see that
\[ \left| \frac{1}{x} - \frac{1}{[x]} \right| = O \left( \frac{1}{x^2} \right), \quad x \geq 1. \]
Therefore by Lemma 4.4.1,
\[
\left| \int_1^{n+1} \frac{1}{x} U(x) dx - \int_1^n \frac{1}{x} U(x) dx \right| \\ \leq \left| \int_1^n \left( \frac{1}{x} - \frac{1}{|x|} \right) U(x) dx \right| + \left| \int_n^{n+1} \frac{1}{x} U(x) dx \right| \\
\overset{a.s.}{=} O(1) \int_1^n \frac{(x \log \log x)^{1/2}}{x^2} dx + \frac{1}{n} (n \log \log n)^{1/2} = O(1).
\]

We can thus conclude that
\[
S_n \overset{a.s.}{=} \sigma W(n) - \int_1^{n+1} \frac{1}{x} U(x) dx + o(n^\alpha).
\]

It remains to note that by (4.2.4) and the modulus of continuity of \( W(\cdot) \) (cf. Lemma 1.1.1 in [45])
\[
U(x) - \sigma W(x) \overset{a.s.}{=} o(x^\alpha), \quad \text{as } x \to \infty
\]
and so
\[
\left| \int_1^{n+1} \frac{1}{x} (U(x) - \sigma W(x)) dx \right| \overset{a.s.}{=} o(1) \int_1^{n+1} \frac{x^\alpha}{x^2} = o(n^\alpha).
\]

This completes the proof of Proposition 4.4.4.

**Proof of Theorem 4.3.2:** First we verify that
\[
(4.4.37) \quad \sup_{m \leq n < \infty} \frac{|S_n|}{m^{1/2} g(n/m)} \overset{a.s.}{=} \sup_{m \leq n < \infty} \frac{\sigma |W^*(n)|}{m^{1/2} g(n/m)} + o(1).
\]
Relation (4.4.37) follows from Proposition 4.4.4 and (4.3.10) because
\[
\sup_{m \leq n < \infty} \frac{|S_n - \sigma W^*(n)|}{m^{1/2} g(n/m)} \overset{a.s.}{=} o(1) \sup_{m \leq n < \infty} \frac{n^{\alpha-1/2}}{(n/m)^{-1/2} g(n/m)} = o(1).
\]
By the scale transformation of the Wiener process
\[
(4.4.38) \quad \sup_{m \leq n < \infty} \frac{|W^*(n)|}{m^{1/2} g(n/m)} \overset{d}{=} \sup_{m \leq n < \infty} \frac{|W^*(n/m)|}{g(n/m)},
\]
so it suffices to verify that
\[
(4.4.39) \quad \sup_{m \leq n < \infty} \frac{|W^*(n/m)|}{g(n/m)} \overset{a.s.}{=} \sup_{1 \leq t < \infty} \frac{|W^*(t)|}{g(t)}.
\]
By the triangle inequality, for $T \geq 1$, we have

$$
\left| \sup_{m<n<\infty} \frac{|W^*(n/m)|}{g(n/m)} - \sup_{1\leq t<\infty} \frac{|W^*(t)|}{g(t)} \right|
\leq \sup_{Tm<n<\infty} \frac{|W^*(n/m)|}{g(n/m)} + \sup_{T<t<\infty} \frac{|W^*(t)|}{g(t)} + \sup_{m\leq n \leq mT} \frac{|W^*(n/m)|}{g(n/m)} - \sup_{1\leq t\leq T} \frac{|W^*(t)|}{g(t)}
$$

For any fixed $T$, by the continuity of $W^*(\cdot)/g(\cdot)$, the last term on the right hand side tends to zero as $m \to \infty$. Combined with assumption (4.3.11), this completes the proof of (4.4.39) and of Theorem 4.3.2.

**Proof of Theorem 4.3.3:** By Assumption (4.2.4),

$$
Z_n = \sigma^{-1} \left[ \sum_{1 \leq i \leq n} (X_i - \omega_0) - \frac{n}{m} \sum_{1 \leq i \leq m} (X_i - \omega_0) \right] \overset{a.s.}{=} W(n) - \frac{n}{m} W(m) + o(n^\alpha).
$$

By assumption (4.3.10)

$$
(4.4.40) \quad \frac{n^\alpha}{m^{1/2} \left( \frac{n-m}{m} \right) g \left( \frac{n}{n-m} \right)} = O(1) \frac{n^{\alpha-1/2}}{(n-m)^{1/2}} = O(n^{\alpha-1/2}) = O(1).
$$

Using the self-similarity of the Wiener process and the fact that

$$
\{W(t) - tW(1), t \in (1, \infty)\} \overset{d}{=} \{(t-1)W \left( \frac{t}{t-1} \right), t \in (1, \infty)\}
$$

we conclude that

$$
(4.4.41) \quad \sup_{m<n<\infty} \frac{|W(n) - \frac{n}{m} W(m)|}{m^{1/2} \left( \frac{n-m}{m} \right) g \left( \frac{n}{n-m} \right)} \overset{d}{=} \sup_{m<n<\infty} \frac{|W(n/m) - \frac{n}{m} W(1)|}{\left( \frac{n-m}{m} \right) g \left( \frac{n}{n-m} \right)}
$$

By (4.4.40) and (4.4.41), it remains to show that

$$
(4.4.42) \quad \sup_{m<n<\infty} \left| \frac{W \left( \frac{n/m}{n/m-1} \right)}{g \left( \frac{n/m}{n/m-1} \right)} \right| \overset{a.s.}{=} \sup_{1<t<\infty} \left| \frac{W \left( \frac{t}{t-1} \right)}{g \left( \frac{t}{t-1} \right)} \right|
$$

Note that

$$
\sup_{1<t<\infty} \left| \frac{W \left( \frac{t}{t-1} \right)}{g \left( \frac{t}{t-1} \right)} \right| \overset{d}{=} \sup_{1<t<\infty} \left| \frac{W(t)}{g(t)} \right|.
$$
Introduce the map $u(t) = t/(t - 1)$, $t > 1$. Since $|u'(t)| \leq (c - 1)^{-2}$ for $t \geq c > 1$, for any fixed $c > 1$

\begin{equation}
\sup_{n \geq cm} \|W(u(n/m))\|_{a.s.} \sup_{t \geq c} \frac{|W(u(t))|}{g(u(t))}.
\end{equation}

Since

\begin{equation}
\sup_{1 < t < c} \frac{|W(u(t))|}{g(u(t))} = \sup_{u \in c/(c - 1)} \frac{|W(u)|}{g(u)},
\end{equation}

assumption (4.3.11) implies that

\begin{equation}
\lim_{c \to 1} \sup_{1 < t < c} \frac{|W(u(t))|}{g(u(t))} = 0.
\end{equation}

Relation (4.4.42) follows from (4.4.43) and (4.4.44). This completes the proof of Theorem 4.3.3.

4.5 Proofs of Theorems 4.3.4 and 4.3.5

**Proof of Theorem 4.3.4:** Observe that under $H_0$, by (4.2.4),

\[ \hat{S}(m, k) = \sum_{m \leq i \leq m+k} (X_i - \omega_0) - \frac{k}{m} \sum_{1 \leq i \leq m} (X_i - \omega_0) \]

\[ \overset{a.s.}{=} \sigma [W(m + k) - W(m) - (k/m)W(m)] + o((m + k)\sigma + (k/m)m\sigma). \]

Therefore,

\[ \sup_{1 \leq k < \infty} \frac{|\hat{S}(m, k)|}{m^{1/2}g_\gamma(m, k)} \overset{a.s.}{=} \sup_{1 \leq k < \infty} \frac{|W(m + k) - W(m) - (k/m)W(m)|}{m^{1/2}g_\gamma(m, k)} + o(1) \sup_{1 \leq k < \infty} \frac{(m + k)^\alpha + (k/m)m^\alpha}{m^{1/2}g_\gamma(m, k)}. \]

Elementary verification shows that

\[ \sup_{1 \leq k < \infty} \frac{(m + k)^\alpha + (k/m)m^\alpha}{m^{1/2}g_\gamma(m, k)} = O(m^{\alpha+\gamma-1/2}), \]

so the assumption $\gamma < 1/2 - \alpha$ yields

\[ \sup_{1 \leq k < \infty} \frac{|\hat{S}(m, k)|}{m^{1/2}g_\gamma(m, k)} \overset{a.s.}{=} \sigma \sup_{1 \leq k < \infty} \frac{|W(m + k) - W(m) - (k/m)W(m)|}{m^{1/2}g_\gamma(m, k)} + o(1). \]
It thus remains to verify that

\begin{equation}
(4.5.45) \quad \sup_{1 \leq k < \infty} \frac{|W(m + k) - W(m) - (k/m)W(m)|}{m^{1/2}g_\gamma(m, k)} \overset{d}{=} \sup_{0 \leq t \leq 1} \frac{|W(t)|}{t^{\gamma}}.
\end{equation}

Notice that if \(k/m \to t\), as \(m \to \infty\), then

\begin{equation}
(4.5.46) \quad g_\gamma(m, k) \to h_\gamma(t) := (1 + t) \left( \frac{t}{1 + t} \right)^\gamma, \quad t \in [0, \infty).
\end{equation}

Note also that, see the proof of Theorem 2.1 in [99],

\begin{equation}
(4.5.47) \quad \sup_{0 \leq t < \infty} \frac{|W(1 + t) - W(1) - tW(1)|}{h_\gamma(t)} \overset{d}{=} \sup_{0 \leq t \leq 1} \frac{|W(t)|}{t^{\gamma}}.
\end{equation}

Using the self-similarity of the Wiener process, (4.5.46) and the modulus of continuity of the Wiener process, and finally (4.5.47), we obtain

\begin{align*}
\sup_{1 \leq k < \infty} \frac{|W(m + k) - W(m) - (k/m)W(m)|}{m^{1/2}g_\gamma(m, k)} &\overset{d}{=} \sup_{1 \leq k < \infty} \frac{|W(1 + k/m) - W(1) - (k/m)W(1)|}{h_\gamma(k/m)} \\
&\overset{a.s.}{\to} \sup_{0 \leq t < \infty} \frac{|W(1 + t) - W(1) - tW(1)|}{h_\gamma(t)} \overset{d}{=} \sup_{0 \leq t \leq 1} \frac{|W(t)|}{t^{\gamma}}.
\end{align*}

This completes the verification of (4.5.45) and the proof of Theorem 4.3.4.

**Proof of Theorem 4.3.5:** Observe that under \(H_0\)

\[ \hat{S}(m, k) = \sum_{m < i \leq m+k} (X_i - \omega_0) - \sum_{m < i \leq m+k} \frac{1}{i - 1} \sum_{j=1}^{m} (X_j - \omega_0) - \sum_{m < i \leq m+k} \frac{1}{i - 1} \sum_{j=m+1}^{i-1} (X_j - \omega_0). \]

By (4.2.4),

\[ \sum_{m < i \leq m+k} (X_i - \omega_0) \overset{a.s.}{=} \sigma [W(m + k) - W(m)] + o((m + k)^{\alpha}); \]

\[ \sum_{m < i \leq m+k} \frac{1}{i - 1} \sum_{j=1}^{m} (X_j - \omega_0) \overset{a.s.}{=} \sum_{m < i \leq m+k} \frac{1}{i - 1} [\sigma W(m) + o(m^{\alpha})] \\
= \sigma W(m) \log((m + k)/m) + o(m^{\alpha}) \log((m + k)/m); \]

\[ \sum_{m < i \leq m+k} \frac{1}{i - 1} \sum_{j=m+1}^{i-1} (X_j - \omega_0) \overset{a.s.}{=} \sigma \sum_{m < i \leq m+k} \frac{1}{i - 1} [W(i - 1) - W(m) + o((i - 1)^{\alpha})] \\
= \sigma \int_{m}^{m+k} \frac{1}{x - 1} [W(x - 1) - W(m)] dx + o((m + k)^{\alpha}). \]
Since $m^\alpha \log((m + k)/m) = O((m + k)^\alpha)$, we conclude that

\[(4.5.48) \quad \sigma^{-1} \tilde{S}(m, k) \overset{a.s.}{=} \left[ W(m + k) - W(m) \right] - \log((m + k)/m)W(m) - \int_m^{m+k} \frac{1}{x-1} [W(x) - W(m)] \, dx + o((m + k)^\alpha). \]

By (4.3.29),

\[(4.5.49) \quad \sup_{1 \leq k < \infty} \frac{(m + k)^\alpha}{m^{1/2} h(k/m)} = O(1). \]

Combining (4.5.48) and (4.5.49), we obtain

\[(4.5.50) \quad \frac{1}{\sigma} \sup_{1 \leq k < \infty} \frac{|\tilde{S}(m, k)|}{m^{1/2} h(k/m)} \overset{a.s.}{=} \frac{\sup_{1 \leq k < \infty} \left| W(m + k) - W(m) \right| - \log((m + k)/m)W(m) - \int_m^{m+k} \frac{1}{x-1} [W(x) - W(m)] \, dx}{m^{1/2} h(k/m)} + o(1). \]

By the self-similarity and the modulus of continuity of the Wiener process,

\[\frac{\sup_{1 \leq k < \infty} \left| W(m + k) - W(m) \right| - \log((m + k)/m)W(m) - \int_m^{m+k} \frac{1}{x-1} [W(x) - W(m)] \, dx}{m^{1/2} h(k/m)} \overset{d}{=} \sup_{1 \leq k < \infty} \frac{|W(1 + k/m) - W(1)| - \log(1 + k/m)W(1) - \int_{1-1/m}^{1+k/m} \frac{1}{s} [W(s) - W(1)] \, ds}{h(k/m)} \overset{a.s.}{=} \sup_{1 \leq k < \infty} \frac{\Gamma(k/m)}{h(k/m)} + o(1), \]

where

\[\Gamma(t) = W(1 + t) - W(1) - \log(1 + t)W(1) - \int_1^{1+t} \frac{1}{s} [W(s) - W(1)] \, ds. \]

As observed in Section 6 of [99], computing the covariances of the process $\Gamma(\cdot)$ shows that it is a Wiener process. Therefore, it remains to show that

\[\sup_{1 \leq k < \infty} \frac{|\Gamma(k/m)|}{h(k/m)} \overset{a.s.}{=} \sup_{0 < t < \infty} \frac{\Gamma(t)}{h(t)}. \]

This can be done by repeating the corresponding argument used in the proof of Theorem 4.3.2. Thus proof of Theorem 4.3.5 is complete.
4.6 Estimation of the Asymptotic Variance

The on-line change-point detection procedures discussed in this chapter rely on the consistent estimation of the asymptotic variance \( \sigma^2 \) given by

\[
\sigma^2 = \sum_{l=-\infty}^{+\infty} \gamma_l, \quad \gamma_l = \text{Cov}(X_k, X_{k+l}).
\]

This estimation problem has been extensively studied, see [5], [6], and [84] for the most relevant contributions. Unlike in the setting of point estimation or hypothesis testing, in the case of sequential monitoring, the criterion for the most useful estimator is not only its MSE, but also the speed of calculation and the properties of the detection delay time of a method which utilizes the estimator. We therefore adapt the original estimator proposed by [5].

We focus primarily on estimators of the form

\[
\hat{\sigma}^2_n = \frac{n}{n - 1} \sum_{l=-n+1}^{n-1} k \left( \frac{l}{W_m} \right) \hat{\gamma}_l(n), \quad n > m,
\]

where \( k(\cdot) \) is a real-valued kernel, \( W_m \) is the bandwidth parameter, and \( \hat{\gamma}_l(n) \) is the \( l \)th sample autocovariance of \( \{X_t\}_t^n \).

We also consider estimators of the form

\[
\hat{\sigma}^2_m = \frac{m}{m - 1} \sum_{l=-m+1}^{m-1} k \left( \frac{l}{W_m} \right) \hat{\gamma}_l(m)
\]

which differ from those defined in (4.6.51) in that they use sample autocovariances computed only from the initial \( m \) observations \( \{X_t\}_t^m \) rather than all observations up to the current time \( n \). The estimators (4.6.52) are thus consistent under both the null and the alternative. Estimators (4.6.51) are consistent only under the null. The consistency of \( \hat{\sigma}^2_m \) was established by [5].

We restrict ourselves to the following two kernel functions:

1. Bartlett : 
   \[
   k_{BT}(x) = \begin{cases} 
   1 - |x| & \text{for } |x| \leq 1, \\
   0 & \text{otherwise}, 
   \end{cases}
   \]

2. Quadratic Spectral : 
   \[
   k_{QS}(x) = \frac{25}{12\pi^2x^2} \left( \frac{\sin(6\pi x/5)}{6\pi x/5} - \cos(6\pi x/5) \right).
   \]
The optimal bandwidths $W_m^*$ for the two kernels are

$$(4.6.53) \quad \text{Bartlett : } W_m^* = 1.1447 \, (a(1)m)^{1/3},$$

$$(4.6.54) \quad \text{QS : } W_m^* = 1.3221 \, (a(2)m)^{1/5},$$

where $a(i)$, $i = 1$ or 2, is a function of the unknown spectral density function $f(\lambda)$ of the process $\{X_t\}$. In the applications and simulations discussed in the following, we assume that the returns follow a GARCH(1,1) model before a change has occurred. We now explain how the constants $a(i)$ can be found under this assumption.

Assume then that $X_t = r_t^2$ and

$$(4.6.55) \quad r_t = h_t Z_t,$$

where $\{Z_t\}$ is an i.i.d. sequence with zero mean and unit variance and $h_t$ evolves according to

$$(4.6.56) \quad h_t^2 = \omega + \alpha r_{t-1}^2 + \beta h_{t-1}^2.$$

Following [91] pp. 665-666, we obtain

$$(4.6.57) \quad X_t = \omega + (\alpha + \beta) X_{t-1} + \nu_t - \beta \nu_{t-1},$$

where $\nu_t = X_t^2 - h_t^2$. Notice that $h_t^2$ is the forecast of $X_t$ based on its own lagged values and thus $\nu_t$ in (4.6.57) is the error associated with this forecast. Hence, $\nu_t$ is a white noise process that is fundamental for $X_t$. Expression (4.6.57) will then be recognized as an ARMA(1,1) process for $X_t$, in which the autoregressive coefficient is $\alpha + \beta$ and moving average coefficient is $-\beta$.

For ARMA(1,1) models with autoregressive parameter $\rho$ and moving average parameter $\psi$, estimates of $a(1)$ and $a(2)$ in (4.6.53) and (4.6.54) are given, respectively, in Eqs (6.6) and (6.5) in [5]. These equations involve an integer parameter $p$, which we set equal to 1. We thus obtain

$$(4.6.58) \quad \hat{a}(1) = \frac{4 \left(1 + \rho \psi \right)^2 \left(\hat{\rho} + \hat{\psi}\right)^2}{(1 - \rho^2)^2 \left(1 + \hat{\psi}\right)^4},$$
Figure 4.1. Daily returns on four stock indexes. Dotted lines indicate the borderlines of the subsamples, i.e., the locations of the expected change-points.

and

\[
\hat{\alpha}(2) = \frac{4 \left(1 + \hat{\rho} \hat{\psi}\right)^2 \left(\hat{\rho} + \hat{\psi}\right)^2}{(1 - \hat{\rho})^4 \left(1 + \psi\right)^4},
\]

where \(\hat{\rho}\) and \(\hat{\psi}\) are appropriate estimates. Using the \(m\) historical observations, we compute the quasi-maximum likelihood estimates (QMLE) \(\hat{\alpha}\) and \(\hat{\beta}\) of GARCH(1,1) model and set \(\hat{\rho} = \hat{\alpha} + \hat{\beta}\) and \(\hat{\psi} = -\hat{\beta}\).

We emphasize that the optimal bandwidth is obtained using the fixed initial \(m\) data points, and so it does not change as we proceed to monitor the data, whereas the sample autocovariance function \(\hat{\gamma}_l\) is either sequentially updated in (4.6.51) or computed only once.
4.7 Empirical Analysis of Index Data

In this section we analyze the daily returns on major US stock indexes with the objective to determine a change-point in variance and estimate GARCH models before and after the change. This analysis will yield practically relevant examples of time series with a change-point in variance which will be used in Section 4.8 to compare the finite sample performance of the various monitoring procedures.

We consider four indexes: the Dow Jones Industrial Average index (DJIA), the Standard and Poor's 500 index (S&P 500), the New York Stock Exchange Composite index (NYSE Composite), and National Association of Securities Dealers Automated Quotations Composite index (NASDAQ). We work with log returns \( r_t = 100 \ln(P_t/P_{t-1}) \), where \( P_t \) denotes the value of index at time \( t \). The four datasets cover the period from 1/1/1990 to 12/31/1999 and consist of 2526 observations each. For each of them, we first determine an approximate date when a variance change occurs, and then split the whole dataset into sub-series with respect to the estimated break point and fit a GARCH(1,1) model to these sub-series so as to estimate the models before and after the change. As indicated in the previous paragraph, our objective is to find what the typical model changes and the implied variance changes of the index data are.

We present the time series plots of the data in Figure 4.1. There are marked similarities among the plots of DJIA, S&P 500, and NYSE Composite, in terms of both shape and the scale of volatility. Visual inspection leads us to suspect that there is a variance change in 1996. Indeed, [159] reports the largest 35 daily increases and the 35 daily decreases in the value of the DJIA from 1885 to 1998 and finds that all but three of the largest increases occurred in the two-year period 1996-1997 and twenty-seven of the largest decreases occurred in that period. Although this observation does not pertain to percent change, a more sensible measurement of volatility, we can certainly say that the stock returns were more volatile since 1996 in comparison with the earlier years of the 1990s. We thus separate each
of the three datasets into two parts: 1/1/1992 ~ 12/31/1995 and 1/1/1996 ~ 12/31/1999, as indicated by the dotted lines in Figure 4.1. We removed the first two years of the data which appears to have a higher volatility than the subsequent four years. In contrast, the volatility of NASDAQ has undergone more substantial swings, especially since the third quarter of 1997, and it peaked around the beginning of 1999. The period coincides with the “mania” stage of speculations on high-tech stocks in the US. Therefore, we divide the data of NASDAQ in a different way: the two subsamples are: 7/1/1994 ~ 6/30/1997 and 7/1/1997 ~ 12/31/1998. In summary, we argue that 1/1/1996 is the change point of variance for the returns on DJIA, S&P 500, and NYSE Composite, and 7/1/1997 is that for NASDAQ. We have, however, no intention to argue that the two dates are the exact break points (if such exist), although we believe that they represent well significant “break periods.” In what follows, we use a change-point estimator to support our decisions.

Kokoszka and Leipus in [108] considered an estimator \( \hat{k} \) of a change point \( k^* \) in the mean of dependent observations defined by

\[
\hat{k} = \min\{k : |U_k| = \max_{1 \leq j < n} |U_j|\},
\]

where

\[
U_k = \left( \frac{k(n-k)}{n} \right)^{1-\lambda} \left( \frac{1}{k} \sum_{j=1}^{k} X_j - \frac{1}{n-k} \sum_{j=k+1}^{n} X_j \right)
\]

with some \( 0 \leq \lambda < 1 \). In our context, \( \{X_j\} \) are squares of the daily log returns and \( n = 2526 \). Thus, the estimate \( \hat{k} \) can be regarded as the break point of variance in the series. The plots of the statistics \( U_k \) versus date are displayed in Figure 4.2, \( \lambda = 0.3 \) is used. We obtained similar plots for different values of \( \lambda \), except when \( \lambda \) equals to 0.8 or 0.9. This is because, as observed in [100], for \( \lambda \) close to 1 the estimator \( \hat{k} \) is not likely to detect a change in the middle of the data, which seems to be the case for our series. The dotted lines in Figure 4.2 mark the same dates as in Figure 4.1. We notice that 1/1/1997 is rather close to the break points estimated by \( \hat{k} \) for DJIA, S&P 500, and NYSE Composite. A similar match is easily seen for 7/1/1997 and the NASDAQ series.

We use the function \textit{garch} in the module \textit{S+Finmetrics} of \textit{S-PLUS} to fit the returns
data. Two models are estimated on each of the eight sub-series:

\[(4.7.62) \quad \text{Model I: } r_t = \mu + \epsilon_t,\]

\[(4.7.63) \quad \text{Model II: } r_t = \mu + \psi r_{t-1} + \epsilon_t,\]

where, in both cases, \(\epsilon_t \sim N(0, \sigma_t^2)\) and

\[(4.7.64) \quad \sigma_t^2 = \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2.\]

Model I is the standard GARCH(1,1) with mean value \(\mu\). We added to Model II the AR(1) term \(\psi r_{t-1}\) to the conditional mean formula. The parameter \(\psi\) takes into account the short range dependence in returns caused by differences in closing times, see [34] for details. The two most common information criteria for model selection are the Akaike (AIC) and Schwarz-Bayesian (BIC). Unfortunately, they are not always in agreement on selecting between Model I and II; Model I fits better for some series while Model II for others. Moreover, the differences in terms of both AIC and BIC for the two models are nearly negligible for all eight series. Thus, applying the parsimony principle, we choose Model I for all eight series. Table 4.1 reports the parameter estimates, which are all highly significant (p-values are not shown here). The last two columns of the table give the sample variance and the implied variance. Sample variances are computed in the usual way. If \(\alpha + \beta < 1\), the unconditional variance of a GARCH(1,1) process is \(\omega/1 - \alpha - \beta\), and the implied variance is thus determined as \(\hat{\omega}/1 - \hat{\alpha} - \hat{\beta}\). The closeness of the two variances renders support, from another angle, to how we separate the data and to the choice we made regarding the model selection.

As shown in both Figure 4.1 and Figure 4.2, there are salient resemblances among the returns of DJIA, S&P 500, and NYSE Composite, while the volatility of the NASDAQ index is more considerable than the others. This observation is reflected in Table 4.1.

We now summarize the patterns of the data before and after breaks.

1. The estimates \(\hat{\mu}, \hat{\omega},\) and \(\hat{\alpha}\) increase after break, whereas \(\hat{\beta}\) slightly decreases. This is consistently observed in all the four indexes.
Figure 4.2. Estimate of break point of variance. Dotted lines indicate the same dates as in Figure 4.1.

Table 4.1. Fitted models of the returns data.

<table>
<thead>
<tr>
<th>Stock Index</th>
<th>Model (Fitting Period)</th>
<th>Number of obs.</th>
<th>μ</th>
<th>ω</th>
<th>α</th>
<th>β</th>
<th>Sample Var.</th>
<th>Implied Var.</th>
</tr>
</thead>
<tbody>
<tr>
<td>DJIA</td>
<td>Model 1. (1/1/92 ~ 12/31/95)</td>
<td>1010</td>
<td>0.070</td>
<td>0.028</td>
<td>0.064</td>
<td>0.869</td>
<td>0.409</td>
<td>0.416</td>
</tr>
<tr>
<td></td>
<td>Model 2. (1/1/96 ~ 12/31/99)</td>
<td>1011</td>
<td>0.081</td>
<td>0.082</td>
<td>0.092</td>
<td>0.853</td>
<td>1.427</td>
<td>1.495</td>
</tr>
<tr>
<td>S&amp;P 500</td>
<td>Model 3. (1/1/92 ~ 12/31/95)</td>
<td>1010</td>
<td>0.062</td>
<td>0.012</td>
<td>0.044</td>
<td>0.924</td>
<td>0.367</td>
<td>0.378</td>
</tr>
<tr>
<td></td>
<td>Model 4. (1/1/96 ~ 12/31/99)</td>
<td>1011</td>
<td>0.082</td>
<td>0.110</td>
<td>0.100</td>
<td>0.832</td>
<td>1.551</td>
<td>1.618</td>
</tr>
<tr>
<td>NYSE</td>
<td>Model 5. (1/1/92 ~ 12/31/95)</td>
<td>1010</td>
<td>0.060</td>
<td>0.015</td>
<td>0.051</td>
<td>0.901</td>
<td>0.304</td>
<td>0.308</td>
</tr>
<tr>
<td>Composite</td>
<td>Model 6. (1/1/96 ~ 12/31/99)</td>
<td>1011</td>
<td>0.074</td>
<td>0.064</td>
<td>0.108</td>
<td>0.841</td>
<td>1.148</td>
<td>1.245</td>
</tr>
<tr>
<td>NASDAQ</td>
<td>Model 7. (7/1/94 ~ 6/30/97)</td>
<td>757</td>
<td>0.130</td>
<td>0.131</td>
<td>0.136</td>
<td>0.749</td>
<td>1.140</td>
<td>1.139</td>
</tr>
<tr>
<td></td>
<td>Model 8. (7/1/97 ~ 12/31/98)</td>
<td>382</td>
<td>0.324</td>
<td>0.534</td>
<td>0.186</td>
<td>0.661</td>
<td>3.431</td>
<td>3.492</td>
</tr>
</tbody>
</table>
2. For DJIA, S&P 500, and NYSE Composite, despite the rise of variances, \( \hat{\mu}, \hat{\omega}, \) and \( \hat{\alpha} \) stay small (less than or around 0.1), \( \hat{\beta} \) is relatively large (0.87 on average), and the sum of \( \hat{\alpha} \) and \( \hat{\beta} \) is quite close to the boundary value 1. The estimates \( \hat{\mu}, \hat{\omega}, \) and \( \hat{\alpha} \) for NASDAQ are comparatively greater than those of the other indexes (about 0.13 for all before the break and 0.324, 0.534, and 0.186, respectively, after the break), but \( \hat{\beta} \) is smaller (changes from 0.749 to 0.661 after the break).

3. We notice that, as implied variances increase after the breaks, the sum \( \hat{\alpha} + \hat{\beta} \) does not vary much, but \( \hat{\omega} \) gets 4 to 5 times greater in general. The expression \( \hat{\omega}/1 - \hat{\alpha} - \hat{\beta} \) implies that it is mainly the increase of the constant term \( \omega \) that causes the upswings of variance of the returns data.

4. Due to the similarities among the parameters for DJIA, S&P 500, and NYSE Composite, we may let Model 1 and Model 2 of DJIA (see Table 4.1) represent this group of indexes. Model 7 and Model 8 of NASDAQ can be regarded as another typical change-point model. In both cases, the variances increase about 3, but less than 4 times.

In the following, we concentrate on the two sets of models determined in point 4 above to conduct a simulation study in next section.

4.8 Simulation Study

The objective of this section is to compare the finite sample performance of the four monitoring schemes introduced in Section 4.3. For ease of reference, the following table lists the schemes and the abbreviations that will be used in the sequel.

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Rejection rule</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>CUSUM</td>
<td>(4.3.18) (Theorem 4.3.2)</td>
<td>CS</td>
</tr>
<tr>
<td>Fluctuation</td>
<td>(4.3.20) (Theorem 4.3.3)</td>
<td>FL</td>
</tr>
<tr>
<td>Partial sum of residuals</td>
<td>(4.3.26) (Theorem 4.3.4)</td>
<td>PS</td>
</tr>
<tr>
<td>Partial sum of recursive residuals</td>
<td>(4.3.34) (Theorem 4.3.5)</td>
<td>RR</td>
</tr>
</tbody>
</table>

Recall that the boundary function \( g_a(t) = [t(a^2 + \ln t)]^{1/2}, \ t \geq 1 \) is used for both CS and FL monitoring and, for asymptotic controlled sizes of 5% and 10%, \( a^2 \) equals to 7.78 and 6.25, respectively. The rejection condition for the method PS involves \( c_a(\gamma) \), see (4.3.26). If \( \gamma = 0.25 \), the value used in our simulations, critical value of \( c_a(\gamma) = 2.386 \)

Table 4.2. Empirical sizes (in percent) for the four monitoring methods applied to simulated series of squared GARCH(1,1) observations following Model 1 and Model 7 in Table 4.1. The number of replications is $R = 1000$ with historical size $m = 500$ and autocovariances computed sequentially.

<table>
<thead>
<tr>
<th>Method</th>
<th>$q$</th>
<th>Model 1</th>
<th></th>
<th></th>
<th>Model 7</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Bartlett</td>
<td>QS</td>
<td>Bartlett</td>
<td>QS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5%</td>
<td>10%</td>
<td>5%</td>
<td>10%</td>
<td>5%</td>
<td>10%</td>
</tr>
<tr>
<td>CS</td>
<td>$m$</td>
<td>2.8</td>
<td>5.1</td>
<td>2.4</td>
<td>4.0</td>
<td>2.8</td>
<td>5.7</td>
</tr>
<tr>
<td></td>
<td>$2m$</td>
<td>3.8</td>
<td>7.0</td>
<td>3.5</td>
<td>6.2</td>
<td>3.6</td>
<td>7.2</td>
</tr>
<tr>
<td></td>
<td>$3m$</td>
<td>4.7</td>
<td>8.0</td>
<td>4.1</td>
<td>7.5</td>
<td>4.3</td>
<td>7.9</td>
</tr>
<tr>
<td>FL</td>
<td>$m$</td>
<td>6.9</td>
<td>10.2</td>
<td>5.0</td>
<td>7.8</td>
<td>8.6</td>
<td>10.6</td>
</tr>
<tr>
<td></td>
<td>$2m$</td>
<td>8.2</td>
<td>12.3</td>
<td>6.3</td>
<td>10.4</td>
<td>9.3</td>
<td>11.6</td>
</tr>
<tr>
<td></td>
<td>$3m$</td>
<td>9.1</td>
<td>13.7</td>
<td>6.7</td>
<td>11.6</td>
<td>9.9</td>
<td>12.3</td>
</tr>
<tr>
<td>PS</td>
<td>$m$</td>
<td>3.6</td>
<td>6.5</td>
<td>2.3</td>
<td>5.1</td>
<td>2.9</td>
<td>5.8</td>
</tr>
<tr>
<td></td>
<td>$2m$</td>
<td>6.0</td>
<td>10.5</td>
<td>5.1</td>
<td>9.1</td>
<td>4.5</td>
<td>9.0</td>
</tr>
<tr>
<td></td>
<td>$3m$</td>
<td>7.3</td>
<td>13.0</td>
<td>6.8</td>
<td>11.1</td>
<td>5.6</td>
<td>10.8</td>
</tr>
<tr>
<td>RR</td>
<td>$m$</td>
<td>0.9</td>
<td>1.6</td>
<td>0.4</td>
<td>1.2</td>
<td>0.2</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>$2m$</td>
<td>2.0</td>
<td>4.2</td>
<td>1.4</td>
<td>3.2</td>
<td>1.2</td>
<td>2.4</td>
</tr>
<tr>
<td></td>
<td>$3m$</td>
<td>3.1</td>
<td>6.1</td>
<td>2.6</td>
<td>4.7</td>
<td>1.6</td>
<td>3.9</td>
</tr>
</tbody>
</table>

(2.106) gives 5% (10%) asymptotic false alarm rate. The boundary function is $h_a(t) = (t + 1)^{1/2}[a^2 + \ln(t + 1)]^{1/2}$ in the RR method, setting $a^2$ equal to 6.0 (4.6) gives $\alpha$ of 5% (10%) asymptotically.

We consider the two kernels, Bartlett and quadratic spectral, for each of the four schemes. We use autocovariances computed either from all observations available up to the current time $n$ or from the initial $m$ observations. We focus mainly on the former way of computing the autocovariances because it will be seen to give better results.

The primary concern regarding the performance of the four sequential monitoring procedures is the false alarm rate $\alpha$, namely, the probability of falsely rejecting a true null hypothesis of no change in variance. To evaluate this probability, we generate GARCH(1,1) time series according to Model 1 and Model 7, the two typical pre-change specifications in Table 4.1. With historical sample sizes $m = 100, 200, 300, 400, 500$, we begin to monitor the squared process from the $(m + 1)$th observation. The monitoring horizon $q$ is set to be one, two, and three times $m$. We replicate the monitoring procedures on a given series of length $(q + 1)m$ for a large number of times, $R = 1000$ in our simulations. The empirical
sizes can then be computed by dividing by $R$ the number of times a boundary is crossed. Theoretically, they should become close to the asymptotic size when $m$ and $q$ approach infinity.

Table 4.2 presents the empirical sizes for the four algorithms when $m = 500$, with autocovariances sequentially computed. We first notice that the sizes produced by methods CS and RR are below the target levels. Almost all over-rejections are in the cells of methods FL and PS, and those of FL are more severe: the worst is 9.9% (13.7%) for the controlled size of 5% (10%) and they equal to or surpass the nominal levels even when the monitoring horizon is only one $m$ long, which equals to 500 in Table 4.2. The four methods give nearly equal performance with respect to the two different models. The empirical sizes, however, do depend on model specification to some degree. Our simulations, whose details are not reported here, show that, when the sum of $\alpha$ and $\beta$ is smaller and not as close to 1 as in models 1 and 7, the problem of over-rejection is much less severe even for the FL scheme. This is probably due to the fact that the greater $\alpha + \beta$ is, the less accurate estimation of the variance of squared GARCH(1,1) observations we can get.

Tables D.1-D.4 give the empirical sizes for the other four values of $m$, 100, 200, 300, and 400, respectively. In order to illustrate the overall conclusions that can be drawn from Tables D.1-D.4, we present a representative graphical comparison in Figure 4.3. We focus on the results for Model 1 with controlled size 10% and $q = 3$. Plots of other cases support our overall conclusions. Figure 4.3 supports the observation made earlier that methods CS and RR are more conservative than methods PS and FL. In addition, empirical size basically decreases as we extend the historical sample size $m$. This is especially true for methods FL and RR with the Bartlett window, whose sizes fall monotonically as $m$ increases from 100 to 500. When the Bartlett kernel is used, it seems that $m = 200$ is long enough for methods CS and RR to secure a size close to the nominal level of 10%, whereas for methods PS and FL even $m = 500$ can only yield sizes that are about 14%. As for the QS kernel, the sizes decrease consistently as $m$ increases from 100 to 300 for all four algorithms, and they roughly level off for longer historical samples, meaning that extending $m$ beyond
Figure 4.3. Comparison of empirical sizes. Results are reported for Model 1 with either Bartlett kernel or QS kernel. The monitoring horizon is $q = 3$. The straight line marks the target size of 10%.

$300$ does not appreciably improve the performance. Because we are more concerned about over-rejections, the QS window is recommended: it works better for methods PS or FL and brings their empirical sizes down to less than 12%, as opposed to 16% for the Bartlett window, when $m = 300$. Overall, however, the difference between the two windows is not large.

We next examine the power of the tests. We consider two typical variance changes which are represented by the passages from Model 1 to Model 2 and from Model 7 to Model 8. Focusing again on the five values of $m$ equal to 100, 200, 300, 400, and 500, we generate the data and let the model transitions happen at $t = 1.1m + 1$. The monitoring starts from the $(m + 1)$th observation, but the monitoring horizon $q$ is fixed at 500, instead of varying with $m$. The empirical power of the tests is the percentage of rejections in $R$ replications. We used $R = 1000$ in our simulations. A commonly used criterion of evaluating sequential procedures is the average run length (ARL) defined as the average of detection delays in the presence of a real break. Empirical ARL can be computed by subtracting the time of real break from the average of the alarm times in the $R$ replications. To save space, we
Table 4.3. Empirical power and ARL of the four methods applied to the two types of variance changes. The number of replications is $R = 1000$, with controlled size of 10%, and autocovariances computed sequentially. The monitoring horizon is $q = 500$.

<table>
<thead>
<tr>
<th>Method</th>
<th>$m$ break</th>
<th>Model 1 $\rightarrow$ Model 2</th>
<th>Model 7 $\rightarrow$ Model 8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Bartlett</td>
<td>QS</td>
</tr>
<tr>
<td>CS</td>
<td>100 111</td>
<td>94.3</td>
<td>(89)</td>
</tr>
<tr>
<td></td>
<td>300 331</td>
<td>98.9</td>
<td>(102)</td>
</tr>
<tr>
<td></td>
<td>500 551</td>
<td>98.5</td>
<td>(115)</td>
</tr>
<tr>
<td>FL</td>
<td>100 111</td>
<td>98.1</td>
<td>(38)</td>
</tr>
<tr>
<td></td>
<td>300 331</td>
<td>99.9</td>
<td>(36)</td>
</tr>
<tr>
<td></td>
<td>500 551</td>
<td>100</td>
<td>(37)</td>
</tr>
<tr>
<td>PS</td>
<td>100 111</td>
<td>99.0</td>
<td>(40)</td>
</tr>
<tr>
<td></td>
<td>300 331</td>
<td>100</td>
<td>(46)</td>
</tr>
<tr>
<td></td>
<td>500 551</td>
<td>100</td>
<td>(51)</td>
</tr>
<tr>
<td>RR</td>
<td>100 111</td>
<td>97.8</td>
<td>(76)</td>
</tr>
<tr>
<td></td>
<td>300 331</td>
<td>99.7</td>
<td>(88)</td>
</tr>
<tr>
<td></td>
<td>500 551</td>
<td>99.9</td>
<td>(105)</td>
</tr>
</tbody>
</table>

only report in Table 4.3 the results for $m = 100, 300, 500$ with controlled size 10%. The empirical ARL is shown in parenthesis next to the power.

The general relation between size and power in hypothesis testing is that a test with a smaller size, i.e. a lower probability of type I error, tends to have a higher probability of type II error, i.e. lower power. With this in mind, it is not surprising to see that methods CS and RR have less power than methods FL and PS, because, as shown in Table 4.2 and Figure 4.3, their sizes are smaller and do not suffer from the problem of over-rejection. This difference is most obvious when $m = 100$, and becomes negligible when $m \geq 300$ and the transition is from Model 1 to Model 2, in which case the empirical powers are all more than 98%. The powers corresponding to the transition from Model 7 to Model 8 are generally lower, with longer ARL. This is may be because the variance increases about 3.5 times in the transition from Model 1 to Model 2 and about 3 times in the transition from Model 7 to Model 8. The difference, is, however practically negligible for methods FL and PS, with $m \geq 300$. An appealing property of the tests is that greater power comes with shorter ARL. In particular, the average detection delays of methods FL and PS, the two tests with
greater power, are around 40 and less than 55 for the first type of change, but are more than
double for the other two algorithms. As regards the choice of kernel function, the Bartlett
kernel gives higher power and shorter ARL than the QS kernel. Overall, we may expect to
get higher power by using more observations as a historical sample, but the gains, in terms
of both power and ARL, of extending \( m \) beyond 300 are not significant.

Using box plots, we present the distributions of the first hitting time in Figure 4.4, with
\( m = 300 \). The plots for methods FL and PS are almost identical, except that the former
has a slightly smaller median and interquartile range (the difference between third and first
quartiles). The distributions for methods CS and RR are clearly more spread out. All
distributions have elongated upper tails, i.e. are positively skewed. For the transition from
Model 7 to Model 8, all summary statistics, including first quartile, median, third quartile,
and interquartile range, increase meaning that all methods are less effective in detecting this
transition. Hitting time distributions for \( m = 100 \) and \( m = 500 \) are shown, respectively, in
Figures D.1 and D.2.
To investigate the effect of the location of the break point on the power and the distribution of the detection time, we run simulations with the structural change occurring at $1.2m + 1$, with the length of the monitoring horizon remaining equal to 500. There are barely any differences in terms of the power of the tests. The ARL however increases; it prolongs by around 20 for the transition from Model 1 to Model 2, and by about 30 for the other transition. The increases of the order statistics, such as median and the quartiles, are small.

We now discuss the effect of using covariances computed from the initial $m$ observations rather than from the observations up to the current time $n$. The simulation results presented in Tables D.5 and D.6, for empirical sizes and powers, respectively, are produced from exactly the same set-up as those in Tables 4.2 and 4.3, except that the sample autocovariances $\hat{\gamma}_n$ are computed using the historical $m$ observations. All four monitoring procedures suffer from the problem of over-rejections now, their relative performance, however, does not change. The methods CS and RR are still more conservative, producing smaller false alarm rates. Nearly all sizes are doubled, some are even more than ten times greater than those in Table 4.2. For the monitoring horizon $q = 3m$, with $m = 500$, in Table D.5, method FL gives, as observed before, the worst results: 20.3% (23.9%) for nominal size 5% (10%). The power of the tests exceeds 97% in all cells in Table D.6, with much shorter ARL, basically halved, than those in Table 4.3. As we discussed earlier, it is legitimate to compute autocovariances sequentially under the null hypothesis what leads to a more accurate variance estimate $\hat{\sigma}$. It explains why Table 4.2 reports better sizes. Since methods with covariances computed sequentially also have satisfactory power, we recommend this way of computing covariances.

In conclusion, we can say that if the primary concern is to control the false alarm rate under the nominal level, the methods CS and RR with sequentially computed covariances and the QS kernel are recommended. The choice of the algorithm is especially important for $m = 100$ and $m = 200$. The false alarm rates of methods PS and FL are much higher, but if we use the QS kernel and historical samples longer than 300, the problem becomes less
Table 4.4. Comparison of empirical sizes of the four methods considered in this chapter with the method based on likelihood scores. Results are reported for the Bartlett kernel, sequentially computed autocovariances, $m = q = 1000$ and are based on $R = 2000$ replications.

<table>
<thead>
<tr>
<th>Model</th>
<th>5%</th>
<th>10%</th>
<th>5%</th>
<th>10%</th>
<th>5%</th>
<th>10%</th>
<th>5%</th>
<th>10%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CS</td>
<td>FL</td>
<td>PS</td>
<td>RR</td>
<td>LS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model A</td>
<td>2.7</td>
<td>7.1</td>
<td>1.4</td>
<td>0.05</td>
<td>13.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>9.75</td>
<td>3.9</td>
<td>0.25</td>
<td>18.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model B</td>
<td>3.1</td>
<td>6.75</td>
<td>1.2</td>
<td>0.05</td>
<td>12.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.1</td>
<td>9.4</td>
<td>3.5</td>
<td>0.15</td>
<td>13.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model C</td>
<td>3.1</td>
<td>7.7</td>
<td>1.65</td>
<td>0.15</td>
<td>53.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.2</td>
<td>10.9</td>
<td>4.3</td>
<td>0.3</td>
<td>61.6</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

severe. Generally, the QS kernel gives somewhat more conservative sizes than the Bartlett kernel. Methods PS and FL have only somewhat higher power than the other two methods, but the detection delay of methods CS and RR is basically two times larger than that of methods PS and FL.

We conclude this section by comparing the methods investigated in this chapter with the method proposed by [12] which is based on likelihood scores and for which we use here the abbreviation LS. Following [12], we consider three GARCH(1,1) models:

Model A: \( \omega = 0.05, \alpha_1 = 0.4, \beta_1 = 0.3 \);

Model B: \( \omega = 0.05, \alpha_1 = 0.5, \beta_1 = 0.0 \);

Model C: \( \omega = 1.0, \alpha_1 = 0.3, \beta_1 = 0.2 \).

Table 4.4 reports empirical sizes for $m = 1000$ and $q = 1000$. i.e. the monitoring period has the same length as the initial period. These values and the models have been chosen to make possible the comparison with the entries in Table 2 of [12]. All methods proposed in this chapter have well controlled sizes. Method RR is too conservative, but the other three methods appear to be preferable to method LS.
CHAPTER 5
CONCLUSION AND FUTURE RESEARCH

Many econometric studies (for example, [138], [18], [62] and [69]) have documented that financial time series tend to be highly heteroskedastic. This has important implications for many areas of modern finance, including risk management and dynamic capital-asset pricing theory such as options pricing. In the time series literature several competing models have been developed which can take account of the empirical heteroskedasticity of financial returns series. A large number of these modeling approaches have been within the framework of Autoregressive Conditional Heteroskedasticity (ARCH) model proposed in [60]. Since its introduction, the variations, extensions and applications have become breathtaking. The Generalized ARCH (GARCH) model developed by [20] has proved to be a leading generic model for almost all classes of asset returns. In this dissertation we studied squared GARCH sequences in different contexts.

In Chapter 2 we are concerned with assessing finite sample performance of several resampling methods of constructing confidence intervals for autocorrelations of squared GARCH sequences. Various GARCH model specifications are used in the comparative study. Among the methods of residual bootstrap, block bootstrap and subsampling, the residual bootstrap based on the standard GARCH models is seen to perform the best. The residual bootstrap method also dominates the others in constructing confidence intervals for cross-correlations of squared sequences that follow a popular bivariate GARCH model. The performances of both block bootstrap and subsampling depend largely on the block size $b$, which is often difficult to choose. Based on extensive simulations, we provided practical guidance for selecting $b$.

There have been historically two competing approaches to modeling observed long memory in stock market volatility: the long-range dependent models versus the change-point models. Based on the test developed in [14], the objective of Chapter 3 is to distinguish between long memory and volatility changes in financial returns data. Using various variance estimators, we examined and compared finite sample performance of the test by means of
an extensive simulation study. In terms of empirical size, the Bartlett kernel estimator with bandwidth parameter determined by a modified version of the truncation selection method in [5] is seen to perform best. The testing procedure also gives fairly satisfactory empirical power and proves to be robust to various GARCH models.

In Chapter 4 we proposed several sequential monitoring schemes to detect a change in unconditional variance of GARCH sequences. The methods are based on the paradigm of [39] and essentially nonparametric. We assume the constancy of the variance within the historical $m$ observations and start to monitor from the $(m + 1)$th observations to decide if there is a change. The methods has asymptotically controlled probability of falsely rejecting the null hypothesis of no change. Our theory has a broad applicability to various GARCH-type sequences and relies on a strong invariance principle which holds for the squared observations generated by such models. In a simulation study, we examined and compared, in terms of both empirical size and power, the procedures with different variance estimators. Our methods are seen to have better controlled sizes than the likelihood scores based method proposed in [12].

Our investigations mainly used daily returns on stocks or stock indexes and the practically relevant conditionally heteroskedastic models estimated from them. It will be interesting to see how our methods and tests work on time series of different time units, such as monthly and intra-daily returns. Long memory also exists in exchange rates, interest rates and the prices of different commodities. More work towards discriminating between long memory and change points in the above series will be interesting as well.
REFERENCES


[33] J. CAMPBELL, M. LETTAU, B. MALKIET, AND Y. XU, *Have individual stocks become


APPENDIX A

DERIVATIONS OF FORMULAS IN CHAPTER 2

Recall that $Z_t$ in Eq. (2.3.8) are standard normal, we thus have $E(Z_t^2) = 1$ and $E(Z_t^4) = 3$. The autocorrelation of squared $X_t$ at lag 1 is

\[ \rho_{X^2}(1) = \frac{\bar{\gamma}_c (1 - \gamma_c^2) - \gamma_c (1 - \gamma_{c2})}{3 (1 - \gamma_c^2) - (1 - \gamma_{c2})}, \]

where $\bar{\gamma}_c = E(Z_t^2 c_t)$, $\gamma_c = E(c_t)$, and $\gamma_{c2} = E(c_t^2)$. We denote the numerator and denominator of $\rho_{X^2}(1)$ by $N(\gamma)$ and $D(\gamma)$, respectively.

1. Derivations of $\gamma_{c2}$ in Eq. (2.3.15) and $\rho_{X^2}(1)$ in Eq. (2.3.16).

For standard GARCH(1,1), $c_t = \beta + \alpha Z_t^2$. We have

\[ \gamma_c = E(Z_t^2 \beta + \alpha Z_t^4) = \beta + 3\alpha, \]

\[ \gamma_c = E(\beta + \alpha Z_t^2) = \beta + \alpha, \]

\[ \gamma_{c2} = E(\beta^2 + 2\alpha \beta Z_t^2 + \alpha^2 Z_t^4) = \beta^2 + 2\alpha \beta + 3\alpha^2. \]

Thus, after simplification,

\[ \rho_{X^2}(1) = \frac{\alpha (1 - \alpha \beta - \beta^2)}{1 - 2\alpha \beta - \beta^2}. \]

\[ \square \]

2. Derivations of $\gamma_{c2}$ in Eq.(2.3.17) and $\rho_{X^2}(1)$ in Eq. (2.3.18).

For GJR-GARCH(1,1), $c_t = \beta + (\alpha + \phi I(Z_t))Z_t^2$. Recall that $I(Z_t) = 1$ if $Z_t < 0$, and $I(Z_t) = 0$ otherwise. In the following calculations, we use $E(I(Z_t)Z_t^2) = \frac{1}{2}$ and $E(I(Z_t)Z_t^4) = \frac{3}{2}$. We get

\[ \gamma_c = E(\beta Z_t^2 + \alpha Z_t^4 + \phi I(Z_t)Z_t^4) \]

\[ = \beta + 3\alpha + \frac{3}{2} \phi, \]

\[ \gamma_c = E(\beta + \alpha Z_t^2 + \phi I(Z_t)Z_t^2) \]

\[ = \beta + \alpha + \frac{1}{2} \phi, \]

\[ \gamma_{c2} = E(\beta^2 + \alpha^2 Z_t^4 + \phi^2 I^2(Z_t)Z_t^4 + 2\alpha \beta Z_t^2 + 2\alpha \phi I(Z_t)Z_t^4 + 2\beta \phi I(Z_t)Z_t^2) \]

\[ = \beta^2 + 3\alpha^2 + \frac{3}{2} \phi^2 + 2\alpha \beta + 3\alpha \phi + \beta \phi. \]
Thus, the numerator
\[
(A.0.9) \quad N(\gamma) = \left[ \beta + 3(\alpha + \frac{1}{2} \phi) \right] \cdot \left[ (1 - \beta^2 - 2\alpha \phi - \beta \phi) - (\alpha^2 + \alpha \phi) - \frac{1}{4} \phi^2 \right] - \\
\left[ \beta + (\alpha + \frac{1}{2} \phi) \right] \cdot \left[ (1 - \beta^2 - 2\alpha \phi - \beta \phi) - 3(\alpha^2 + \alpha \phi) - \frac{3}{2} \phi^2 \right] \\
= 2\alpha \beta (\alpha + \phi) + \frac{5}{4} \beta \phi^2 + 2 \left( \alpha + \frac{1}{2} \phi \right) \left( 1 - \beta^2 - 2\alpha \phi - \beta \phi + \frac{3}{8} \phi^2 \right),
\]
and the denominator
\[
(A.0.10) \quad D(\gamma) = 3 \left( 1 - \beta^2 - 2\alpha \phi - \beta \phi - \alpha^2 + \alpha \phi - \frac{1}{4} \phi^2 \right) - \\
\left( 1 - \beta^2 - 2\alpha \phi - \beta \phi - 3\alpha^2 + \alpha \phi - \frac{3}{2} \phi^2 \right) \\
= 2 (1 - \beta^2 - 2\alpha \phi - \beta \phi) + \frac{3}{4} \phi^2.
\]
Finally, we obtain
\[
(A.0.11) \quad \rho_{\omega^2}(1) = \frac{2\alpha \beta (\alpha + \phi) + \frac{5}{4} \beta \phi^2 + 2 \left( \alpha + \frac{1}{2} \phi \right) \left( 1 - \beta^2 - 2\alpha \phi - \beta \phi + \frac{3}{8} \phi^2 \right)}{2 (1 - \beta^2 - 2\alpha \phi - \beta \phi) + \frac{3}{4} \phi^2}.
\]

3. Derivations of $\gamma_{c2}$ in Eq. (2.3.19) and $\rho_{\omega^2}(1)$ in Eq. (2.3.20).

For NLGARCH(1,1,2), $c_t = \beta + \alpha (1 - 2\eta \text{sign}(Z_t) + \eta^2) Z_t^2$. In the following calculations, we use $E(\text{sign}(Z_t)Z_t^2) = 0$, $E(\text{sign}(Z_t)Z_t^4) = 0$, and $E(\text{sign}^2(Z_t)Z_t^4) = 3$. We get
\[
(A.0.12) \quad \tilde{\gamma}_c = E \left( \beta Z_t^2 + \alpha Z_t^4 \left( 1 - 2\eta \text{sign}(Z_t) + \eta^2 \right) \right) \\
= \beta + 3\alpha + 3\alpha \eta^2,
\]
\[
(A.0.13) \quad \gamma_c = E \left( \beta + \alpha Z_t^2 \left( 1 - 2\eta \text{sign}(Z_t) + \eta^2 \right) \right) \\
= \beta + \alpha + \alpha \eta^2,
\]
\[
(A.0.14) \quad \gamma_{c2} = E \left[ \beta^2 + \alpha^2 Z_t^4 \left( 1 + 4\eta^2 \text{sign}^2(Z_t) + \eta^4 - 4\eta \text{sign}(Z_t) \\
+ 2\eta^2 - 4\eta^3 \text{sign}(Z_t) \right) + 2\alpha \beta Z_t^2 - 4\alpha^2 Z_t^4 \eta \text{sign}(Z_t) + 2\alpha \beta \eta^2 Z_t^2 \right] \\
= \beta^2 + 3\alpha^2 + 18\alpha^2 \eta^2 + 3\alpha^2 \eta^4 + 2\alpha \beta + 2\alpha \beta \eta^2.
\]
Thus, the numerator
\[
N(\gamma) = [\beta + 3(\alpha + \alpha \eta^2)] \cdot [(1 - \beta^2 - 2\alpha \beta - 2\alpha \beta \eta^2) - (\alpha^2 + \alpha^2 \eta^4) - 2\alpha^2 \eta^2] - [\beta + (\alpha + \alpha \eta^2)] \cdot [(1 - \beta^2 - 2\alpha \beta - 2\alpha \beta \eta^2) - 3(\alpha^2 + \alpha^2 \eta^4) - 18\alpha^2 \eta^2]
\]
\[
(A.0.15) = 2\alpha^2 \beta (\eta^4 + 8\eta^2 + 1) + 2\alpha(1 + \eta^2) \cdot (1 - \beta^2 - 2\alpha \beta - 2\alpha \beta \eta^2 + 6\alpha^2 \eta^2),
\]
and the denominator
\[
D(\gamma) = 3(1 - \beta^2 - 2\alpha \beta - 2\alpha \beta \eta^2 - \alpha^2 + \alpha^2 \eta^4 - 2\alpha^2 \eta^2)
\]
\[
- (1 - \beta^2 - 2\alpha \beta - 2\alpha \beta \eta^2 - 3\alpha^2 + \alpha^2 \eta^4 - 18\alpha^2 \eta^2)
\]
\[
= 2(1 - \beta^2 - 2\alpha \beta - 2\alpha \beta \eta^2) + 12\alpha^2 \eta^2.
\]
Finally, we get
\[
(A.0.17) \rho_{X_t^2(1)} = \alpha (1 + \eta^2) + \frac{\alpha^2 \beta (\eta^4 + 8\eta^2 + 1)}{1 - \beta^2 - 2\alpha \beta - 2\alpha \beta \eta^2 + 6\alpha^2 \eta^2}.
\]

4. Derivations of Cor($X_t^2(1), X_t^2(2)$) and Cor($X_t^2(1), X_{t-1}^2(2)$) of bivariate CCC-GARCH(1,1) model.

Recall that
\[
(A.0.18) \quad X_t(i) = \sigma_t(i) Z_t(i), \quad i = 1, 2,
\]
where
\[
(A.0.19) \quad \sigma_t^2(i) = \omega_i + a_i X_{t-1}^2(i) + b_i \sigma_{t-1}^2(i), \quad i = 1, 2.
\]
The random vectors $[Z_t(1), Z_t(2)]^T$ are independent with mean zero and satisfy
\[
(A.0.20) \quad EZ_t^2(i) = 1, \quad i = 1, 2, \quad E [Z_t(1)Z_t(2)] = \rho.
\]
We focus on weakly stationary bivariate CCC-GARCH processes, when the condition in
(2.5.30) is met, the fourth moments of $X_t = [X_t(1), X_t(2)]^T$ are not dependent on time.
The formulas that we need are
\[
(A.0.21) \quad r_0 = \text{Cor}(X_t^2(1), X_t^2(2))
\]
\[
= \frac{E [X_t^4(1)X_t^2(2)] - E [X_t^2(1)] \cdot E [X_t^2(2)]}{\sqrt{(E [X_t^4(1)] - E^2 [X_t^2(1)]) \cdot (E [X_t^4(2)] - E^2 [X_t^2(2)])}},
\]
and

\[ r_1 = \text{Cor}(X^2_t(1), X^2_{t-1}(2)) = \frac{E[X^2_t(1)X^2_{t-1}(2)] - E[X^2_t(1)] \cdot E[X^2_{t-1}(2)]}{\sqrt{(E[X^4_t(1)] - E^2[X^2_t(1)]) \cdot (E[X^4_{t-1}(2)] - E^2[X^2_{t-1}(2)])}}. \]

In the sequel, we use the notations \( \mu_{i2} = E[X^2_t(i)] \) and \( \mu_{i4} = E[X^4_t(i)] \), for \( i = 1, 2 \).

Moreover, denote

\[ E[X^2_t(1)X^2_t(2)] = \gamma_0, \quad E[X^2_t(1)X^2_{t-1}(2)] = \gamma_1. \]

We rewrite Eqs. (A.0.21) and (A.0.22) and get

\[ r_0 = \frac{\gamma_0 - \mu_{12}\mu_{22}}{\sqrt{\mu_{14} - \mu_{12}^2} \cdot (\mu_{24} - \mu_{22}^2)}, \]

and

\[ r_1 = \frac{\gamma_1 - \mu_{12}\mu_{22}}{\sqrt{\mu_{14} - \mu_{12}^2} \cdot (\mu_{24} - \mu_{22}^2)}. \]

Clearly, \( E[X^2_t(i)] = E[\sigma^2_t(i)] \), for \( i = 1, 2 \), in Eq. (A.0.18). By taking expectation of Eq. (A.0.19) and solve, we have

\[ \mu_{12} = \frac{\omega_1}{1 - a_1 - b_1}, \]

and

\[ \mu_{22} = \frac{\omega_2}{1 - a_2 - b_2}. \]

As for fourth-order moment, we find

\[ \mu_{14} = E\left[(Z^2_{t+1}(1)\sigma^2_{t+1}(1))^2\right] = E\left[Z^2_{t+1}(1)(\omega_1 + a_1 X^2_t(1) + b_1 \sigma^2_t(1))^2\right] = E[Z^2_{t+1}(1)] E[\omega^2_1 + a^2_1 X^4_t(1) + b^2_1 \sigma^4_t(1) + 2\omega_1 a_1 X^2_t(1)] + 2\omega_1 b_1 \sigma^2_t(1) + 2a_1 b_1 X^2_t(1)\sigma^2_t(1)] = 3\left(\omega^2_1 + a^2_1 \mu_{14} + \frac{1}{3} b^2_1 \mu_{14} + 2\omega_1 a_1 \mu_{12} + 2\omega_1 b_1 \mu_{12} + \frac{2}{3} a_1 b_1 \mu_{14}\right). \]
We use the equalities $E[\sigma_1^2(1)] = \frac{1}{2}\mu_{14}$, and $E[X_1^2(1)\sigma_1^2(1)] = \frac{1}{2}\mu_{14}$ in the calculations above. Solve Eqs. (A.0.27) for $\mu_{14}$, we get

(A.0.28) \[ \mu_{14} = \frac{3\omega_1^2 + 6\omega_1\mu_{12}(a_1 + b_1)}{1 - (3a_1^2 + b_1^2 + 2a_1b_1)}. \]

Similarly,

(A.0.29) \[ \mu_{24} = \frac{3\omega_2^2 + 6\omega_2\mu_{22}(a_2 + b_2)}{1 - (3a_2^2 + b_2^2 + 2a_2b_2)}. \]

Next, note that

(A.0.30) \[ \gamma_0 = E[Z_1^2(1)Z_2^2(2)] E[\sigma_1^2(1)\sigma_1^2(2)] \]
\[ = (1 + 2\rho^2) E[\sigma_1^2(1)\sigma_1^2(2)], \]

due to the independence between $[Z_1^2(1), Z_2^2(2)]^T$ and $[\sigma_1^2(1), \sigma_1^2(2)]^T$. To lighten the notation, we let $\gamma_{01} = E[\sigma_1^2(1)\sigma_1^2(2)]$. By Eq. (A.0.19), we have

\[ \gamma_{01} = E[(\omega_1 + a_1X_1^2(1) + b_1\sigma_1^2(1)) (\omega_2 + a_2X_2^2(2) + b_2\sigma_1^2(2))] \]
\[ = E[(\omega_1 + a_1X_1^2(1)) (\omega_2 + a_2X_2^2(2))] + E[(\omega_1 + a_1X_1^2(1)) b_2\sigma_1^2(2)] + \]
\[ E[(\omega_2 + a_2X_2^2(2)) b_1\sigma_1^2(1)] + E[b_1b_2\sigma_1^2(1)\sigma_1^2(2)] \]
\[ = (\omega_1\omega_2 + \omega_1a_2\mu_{22} + \omega_2a_1\mu_{12} + a_1a_2\gamma_0) + (\omega_1b_2\mu_{22} + a_1b_2\gamma_{01}) + \]
\[ (\omega_2b_1\mu_{12} + a_2b_1\gamma_{01}) + b_1b_2\gamma_{01}. \]

(A.0.31) \[ = \omega_1\omega_2 + \omega_2(a_1 + b_1)\mu_{12} + \omega_1(a_2 + b_2)\mu_{22} + (a_1b_2 + a_2b_1 + b_1b_2)\gamma_{01} + a_1a_2\gamma_0. \]

In the calculations above we use the facts that $E[X_1^2(1)\sigma_1^2(2)] = E[\sigma_1^2(1)X_2^2(2)] = \gamma_{01}$. In Eq. (A.0.30), $\gamma_0 = (1 + 2\rho^2)\gamma_{01}$, by multiplying (A.0.31) by $(1 + 2\rho^2)$, we get

(A.0.32) \[ \gamma_0 = (1 + 2\rho^2) \left(\omega_1\omega_2 + \omega_2(a_1 + b_1)\mu_{12} + \omega_1(a_2 + b_2)\mu_{22}\right) \]
\[ + (a_1b_2 + a_2b_1 + b_1b_2)\gamma_{01} + a_1a_2(1 + 2\rho^2)\gamma_0. \]

Finally, we obtain

(A.0.33) \[ \gamma_0 = \frac{(1 + 2\rho^2) \left(\omega_1\omega_2 + \omega_2(a_1 + b_1)\mu_{12} + \omega_1(a_2 + b_2)\mu_{22}\right)}{1 - [a_1b_2 + a_2b_1 + b_1b_2 + a_1a_2(1 + 2\rho^2)]}. \]
In summary, we find $\mu_{12}$ and $\mu_{22}$ in Eqs. (A.0.25) and (A.0.26), respectively; $\mu_{14}$ and $\mu_{24}$ in Eqs. (A.0.28) and (A.0.29), respectively; and $\gamma_0$ in Eq. (A.0.33). By plugging them in the expression of $r_0$ in (A.0.23), we, after lengthy simplifications, get

\[ r_0 = \frac{2\rho^2(1-a_1b_2-a_2b_1-b_1b_2)}{1-a_1b_2-a_2b_1-b_1b_2-a_1a_2(1+2\rho^2)(1-a_1-b_1)(1-a_2-b_2)} \]

Now, we just need $\gamma_1$ to compute $r_1$ in Eq. (A.0.24). We get

\[ \gamma_1 = E[X^2_{t+1}(1)X^2_t(2)] \]
\[ = E[Z^2_{t+1}(1)\sigma_t^2(1)Z^2_t(2)\sigma_t^2(2)] \]
\[ = \omega_1E[Z^2_{t+1}(1)Z^2_t(2)\sigma_t^2(2) + a_1E[Z^2_{t+1}(1)X^2_t(2)X^2_t(1)] \]
\[ + b_1E[Z^2_{t+1}(1)Z^2_t(2)\sigma_t^2(2)\sigma_t^2(1)] \]
\[ = \omega_1\mu_{22} + a_1\gamma_0 + b_1\frac{\gamma_0}{1 + 2\rho^2}. \]

We use the equality $E[\sigma^2(1)\sigma^2(2)] = \gamma_0(1 + 2\rho^2)^{-1}$ in the calculations above. Then, we obtain

\[ r_1 = \frac{2a_1\rho^2(1-a_1b_2-b_1b_2)}{1-a_1b_2-a_2b_1-b_1b_2-a_1a_2(1+2\rho^2)(1-a_1-b_1)(1-a_2-b_2)} \]

\[ \sqrt{\frac{3+6\frac{a_1+b_1}{a_1-b_1}}{1-a_1b_2-a_2b_1-b_1b_2-a_1a_2(1+2\rho^2)(1-a_1-b_1)(1-a_2-b_2)}} \]

\[ \frac{3+6\frac{a_2+b_2}{a_2-b_2}}{1-a_1b_2-a_2b_1-b_1b_2-a_1a_2(1+2\rho^2)(1-a_1-b_1)(1-a_2-b_2)} \]

\[ \frac{1}{1-a_1b_2-a_2b_1-b_1b_2-a_1a_2(1+2\rho^2)(1-a_1-b_1)(1-a_2-b_2)} \]

\[ \frac{1}{(1-a_1-b_1)^2} \]

\[ \frac{1}{(1-a_2-b_2)^2} \]

\[ \square \]
APPENDIX B

ADDITIONAL TABLES AND FIGURES OF CHAPTER 2
Figure B.1. Of residual bootstrap method, performance comparison between the equal-tailed and the symmetric confidence interval. The lengths of sequences are $n = 100, 250, 500, 1000$ in the four plots, respectively. The nominal coverage $95\%$ is marked by solid horizontal line.
Table B.1. Empirical coverage probabilities of equal-tailed confidence intervals based on the method of block bootstrap.

<table>
<thead>
<tr>
<th>Model</th>
<th>Model</th>
<th>e.c.p. (%)</th>
<th>e.c.p. (%)</th>
<th>e.c.p. (%)</th>
<th>e.c.p. (%)</th>
<th>e.c.p. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>STD GARCH</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>n=500</td>
<td>b=3</td>
<td>89.0</td>
<td>85.4</td>
<td>78.1</td>
<td>72.8</td>
<td>63.2</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>89.9</td>
<td>83.6</td>
<td>78.0</td>
<td>67.7</td>
<td>57.8</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>87.9</td>
<td>82.9</td>
<td>72.8</td>
<td>62.5</td>
<td>50.2</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>83.3</td>
<td>78.0</td>
<td>68.8</td>
<td>57.0</td>
<td>40.9</td>
</tr>
<tr>
<td>n=1000</td>
<td>b=5</td>
<td>90.1</td>
<td>84.9</td>
<td>78.3</td>
<td>68.0</td>
<td>58.5</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>89.9</td>
<td>83.1</td>
<td>76.3</td>
<td>66.6</td>
<td>54.6</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>87.1</td>
<td>84.8</td>
<td>76.8</td>
<td>62.8</td>
<td>53.0</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>87.2</td>
<td>81.6</td>
<td>73.1</td>
<td>58.9</td>
<td>42.1</td>
</tr>
<tr>
<td>GJR GARCH</td>
<td>1</td>
<td>88.3</td>
<td>83.3</td>
<td>74.8</td>
<td>70.3</td>
<td>66.7</td>
</tr>
<tr>
<td>n=500</td>
<td>b=3</td>
<td>85.9</td>
<td>81.5</td>
<td>69.4</td>
<td>63.5</td>
<td>57.6</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>84.8</td>
<td>78.5</td>
<td>70.6</td>
<td>56.6</td>
<td>48.5</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>82.8</td>
<td>74.9</td>
<td>62.6</td>
<td>49.4</td>
<td>39.8</td>
</tr>
<tr>
<td>n=1000</td>
<td>b=5</td>
<td>87.0</td>
<td>82.1</td>
<td>70.5</td>
<td>64.7</td>
<td>60.8</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>87.5</td>
<td>81.3</td>
<td>70.1</td>
<td>58.8</td>
<td>50.6</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>86.4</td>
<td>79.4</td>
<td>69.0</td>
<td>56.9</td>
<td>46.8</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>86.3</td>
<td>79.5</td>
<td>67.9</td>
<td>52.2</td>
<td>39.2</td>
</tr>
<tr>
<td>NL GARCH</td>
<td>1</td>
<td>87.7</td>
<td>82.6</td>
<td>77.5</td>
<td>68.1</td>
<td>55.9</td>
</tr>
<tr>
<td>n=500</td>
<td>b=3</td>
<td>87.3</td>
<td>82.5</td>
<td>73.8</td>
<td>58.7</td>
<td>50.1</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>85.2</td>
<td>80.1</td>
<td>65.0</td>
<td>56.8</td>
<td>44.2</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>84.8</td>
<td>75.0</td>
<td>67.5</td>
<td>47.5</td>
<td>40.4</td>
</tr>
<tr>
<td>n=1000</td>
<td>b=5</td>
<td>88.1</td>
<td>84.7</td>
<td>71.4</td>
<td>65.3</td>
<td>51.6</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>86.9</td>
<td>80.6</td>
<td>70.9</td>
<td>59.9</td>
<td>49.3</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>87.5</td>
<td>80.7</td>
<td>71.3</td>
<td>57.9</td>
<td>45.5</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>87.0</td>
<td>78.9</td>
<td>68.7</td>
<td>54.5</td>
<td>37.2</td>
</tr>
</tbody>
</table>
Table B.2. Empirical coverage probabilities of symmetric confidence intervals based on the method of block bootstrap.

<table>
<thead>
<tr>
<th>Model</th>
<th>e.c.p. (%)</th>
<th>e.c.p. (%)</th>
<th>e.c.p. (%)</th>
<th>e.c.p. (%)</th>
<th>e.c.p. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>STD GARCH</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>n=500</td>
<td>b=3</td>
<td>87.9</td>
<td>84.7</td>
<td>78.7</td>
<td>70.3</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>88.8</td>
<td>82.8</td>
<td>76.2</td>
<td>68.3</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>87.7</td>
<td>84.3</td>
<td>74.2</td>
<td>67.7</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>85.4</td>
<td>82.4</td>
<td>74.1</td>
<td>63.9</td>
</tr>
<tr>
<td>n=1000</td>
<td>b=5</td>
<td>88.7</td>
<td>85.7</td>
<td>80.0</td>
<td>70.3</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>90.1</td>
<td>84.2</td>
<td>74.3</td>
<td>70.0</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>88.8</td>
<td>84.4</td>
<td>77.3</td>
<td>67.0</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>88.9</td>
<td>84.0</td>
<td>75.1</td>
<td>66.3</td>
</tr>
<tr>
<td>GJR GARCH</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>n=500</td>
<td>b=3</td>
<td>87.0</td>
<td>82.0</td>
<td>78.4</td>
<td>65.5</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>89.1</td>
<td>83.8</td>
<td>73.4</td>
<td>63.0</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>87.9</td>
<td>81.8</td>
<td>71.4</td>
<td>60.6</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>85.3</td>
<td>79.3</td>
<td>67.1</td>
<td>58.7</td>
</tr>
<tr>
<td>n=1000</td>
<td>b=5</td>
<td>87.7</td>
<td>84.4</td>
<td>75.2</td>
<td>67.9</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>88.6</td>
<td>85.1</td>
<td>70.8</td>
<td>61.0</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>89.7</td>
<td>83.0</td>
<td>72.7</td>
<td>63.6</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>85.8</td>
<td>82.4</td>
<td>69.9</td>
<td>59.5</td>
</tr>
<tr>
<td>NL GARCH</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>n=500</td>
<td>b=3</td>
<td>85.4</td>
<td>83.0</td>
<td>73.7</td>
<td>64.0</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>85.6</td>
<td>82.9</td>
<td>74.1</td>
<td>64.1</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>86.5</td>
<td>80.2</td>
<td>70.6</td>
<td>63.9</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>85.1</td>
<td>80.7</td>
<td>69.9</td>
<td>60.6</td>
</tr>
<tr>
<td>n=1000</td>
<td>b=5</td>
<td>88.7</td>
<td>82.6</td>
<td>73.3</td>
<td>67.7</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>89.4</td>
<td>82.3</td>
<td>70.4</td>
<td>60.8</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>89.1</td>
<td>82.9</td>
<td>72.3</td>
<td>62.7</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>86.8</td>
<td>82.4</td>
<td>70.6</td>
<td>63.1</td>
</tr>
</tbody>
</table>
Figure B.2. Of block bootstrap method, performance comparison between the smallest and greatest value of block size $b$ used for $n = 500, 1000$. The top two plots are for the equal-tailed confidence interval, the bottom two for the symmetric confidence interval. The nominal coverage of 95% is marked by solid horizontal line.
Table B.3. Empirical coverage probabilities of equal-tailed confidence intervals based on the method of subsampling.

<table>
<thead>
<tr>
<th>Model</th>
<th>e.c.p. (%)</th>
<th>e.c.p. (%)</th>
<th>e.c.p. (%)</th>
<th>e.c.p. (%)</th>
<th>e.c.p. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>STD GARCH</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>n=500</td>
<td>b=5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>54.0</td>
<td>39.2</td>
<td>24.8</td>
<td>15.9</td>
<td>7.2</td>
</tr>
<tr>
<td>50</td>
<td>77.5</td>
<td>60.6</td>
<td>41.8</td>
<td>25.8</td>
<td>13.5</td>
</tr>
<tr>
<td>100</td>
<td>80.7</td>
<td>72.6</td>
<td>59.8</td>
<td>41.3</td>
<td>26.2</td>
</tr>
<tr>
<td>n=1000</td>
<td>b=10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>74.3</td>
<td>59.6</td>
<td>39.5</td>
<td>24.5</td>
<td>12.1</td>
</tr>
<tr>
<td>25</td>
<td>83.0</td>
<td>71.4</td>
<td>52.0</td>
<td>34.7</td>
<td>20.3</td>
</tr>
<tr>
<td>100</td>
<td>80.1</td>
<td>74.0</td>
<td>61.4</td>
<td>48.4</td>
<td>29.6</td>
</tr>
<tr>
<td>200</td>
<td>76.7</td>
<td>70.6</td>
<td>60.0</td>
<td>45.7</td>
<td>32.4</td>
</tr>
<tr>
<td>GJR GARCH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n=500</td>
<td>b=5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>46.0</td>
<td>32.3</td>
<td>20.4</td>
<td>13.5</td>
<td>7.6</td>
</tr>
<tr>
<td>50</td>
<td>73.7</td>
<td>53.0</td>
<td>36.9</td>
<td>23.9</td>
<td>14.6</td>
</tr>
<tr>
<td>100</td>
<td>80.1</td>
<td>68.5</td>
<td>54.1</td>
<td>39.0</td>
<td>24.3</td>
</tr>
<tr>
<td>n=1000</td>
<td>b=10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>69.2</td>
<td>54.5</td>
<td>34.5</td>
<td>20.8</td>
<td>15.1</td>
</tr>
<tr>
<td>100</td>
<td>79.7</td>
<td>74.1</td>
<td>56.1</td>
<td>39.8</td>
<td>25.4</td>
</tr>
<tr>
<td>200</td>
<td>77.9</td>
<td>66.8</td>
<td>54.9</td>
<td>38.3</td>
<td>32.1</td>
</tr>
<tr>
<td>NL GARCH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n=500</td>
<td>b=5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>45.3</td>
<td>33.2</td>
<td>21.8</td>
<td>14.2</td>
<td>8.0</td>
</tr>
<tr>
<td>50</td>
<td>73.7</td>
<td>56.6</td>
<td>37.5</td>
<td>22.3</td>
<td>15.1</td>
</tr>
<tr>
<td>100</td>
<td>80.5</td>
<td>70.5</td>
<td>52.2</td>
<td>39.3</td>
<td>23.2</td>
</tr>
<tr>
<td>n=1000</td>
<td>b=10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>69.8</td>
<td>54.2</td>
<td>35.1</td>
<td>21.4</td>
<td>13.7</td>
</tr>
<tr>
<td>100</td>
<td>81.6</td>
<td>71.2</td>
<td>57.9</td>
<td>44.6</td>
<td>30.1</td>
</tr>
<tr>
<td>200</td>
<td>76.0</td>
<td>68.5</td>
<td>56.5</td>
<td>41.9</td>
<td>29.5</td>
</tr>
</tbody>
</table>
Table B.4. Empirical coverage probabilities of *symmetric* confidence intervals based on the method of *subsampling*.

<table>
<thead>
<tr>
<th>Model</th>
<th>e.c.p. (%)</th>
<th>e.c.p. (%)</th>
<th>e.c.p. (%)</th>
<th>e.c.p. (%)</th>
<th>e.c.p. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td><strong>STD GARCH</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n=500</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b=3</td>
<td>97.8</td>
<td>96.7</td>
<td>92.9</td>
<td>86.5</td>
<td>75.0</td>
</tr>
<tr>
<td>6</td>
<td>93.6</td>
<td>88.8</td>
<td>83.1</td>
<td>76.9</td>
<td>57.8</td>
</tr>
<tr>
<td>8</td>
<td>88.8</td>
<td>85.1</td>
<td>79.8</td>
<td>66.4</td>
<td>55.8</td>
</tr>
<tr>
<td>50</td>
<td>83.2</td>
<td>76.0</td>
<td>68.3</td>
<td>57.6</td>
<td>46.2</td>
</tr>
<tr>
<td>n=1000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b=3</td>
<td>98.2</td>
<td>96.4</td>
<td>92.5</td>
<td>83.7</td>
<td>66.6</td>
</tr>
<tr>
<td>6</td>
<td>93.6</td>
<td>89.6</td>
<td>79.9</td>
<td>72.6</td>
<td>53.6</td>
</tr>
<tr>
<td>10</td>
<td>86.2</td>
<td>83.0</td>
<td>70.4</td>
<td>61.6</td>
<td>47.2</td>
</tr>
<tr>
<td>100</td>
<td>86.4</td>
<td>75.1</td>
<td>67.8</td>
<td>55.8</td>
<td>40.0</td>
</tr>
<tr>
<td><strong>GJR GARCH</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n=500</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b=3</td>
<td>98.1</td>
<td>96.0</td>
<td>92.0</td>
<td>85.1</td>
<td>71.8</td>
</tr>
<tr>
<td>6</td>
<td>91.3</td>
<td>87.2</td>
<td>78.5</td>
<td>67.8</td>
<td>58.9</td>
</tr>
<tr>
<td>8</td>
<td>89.5</td>
<td>84.1</td>
<td>74.9</td>
<td>65.7</td>
<td>55.4</td>
</tr>
<tr>
<td>50</td>
<td>80.6</td>
<td>73.9</td>
<td>64.6</td>
<td>53.9</td>
<td>42.9</td>
</tr>
<tr>
<td>n=1000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b=3</td>
<td>98.0</td>
<td>94.6</td>
<td>89.8</td>
<td>77.4</td>
<td>67.2</td>
</tr>
<tr>
<td>6</td>
<td>93.7</td>
<td>85.3</td>
<td>75.3</td>
<td>63.4</td>
<td>51.6</td>
</tr>
<tr>
<td>10</td>
<td>86.2</td>
<td>76.7</td>
<td>66.4</td>
<td>53.3</td>
<td>41.1</td>
</tr>
<tr>
<td>100</td>
<td>83.4</td>
<td>75.8</td>
<td>61.3</td>
<td>51.4</td>
<td>42.5</td>
</tr>
<tr>
<td><strong>NL GARCH</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n=500</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b=3</td>
<td>97.2</td>
<td>95.3</td>
<td>91.6</td>
<td>82.3</td>
<td>70.4</td>
</tr>
<tr>
<td>6</td>
<td>94.1</td>
<td>95.5</td>
<td>79.9</td>
<td>67.9</td>
<td>51.5</td>
</tr>
<tr>
<td>8</td>
<td>90.1</td>
<td>83.0</td>
<td>75.1</td>
<td>63.3</td>
<td>50.2</td>
</tr>
<tr>
<td>50</td>
<td>80.2</td>
<td>76.1</td>
<td>63.9</td>
<td>54.1</td>
<td>41.2</td>
</tr>
<tr>
<td>n=1000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b=3</td>
<td>98.5</td>
<td>94.7</td>
<td>88.1</td>
<td>75.6</td>
<td>59.6</td>
</tr>
<tr>
<td>6</td>
<td>92.1</td>
<td>87.7</td>
<td>74.8</td>
<td>61.4</td>
<td>47.5</td>
</tr>
<tr>
<td>10</td>
<td>86.3</td>
<td>81.7</td>
<td>66.8</td>
<td>51.7</td>
<td>42.3</td>
</tr>
<tr>
<td>100</td>
<td>83.7</td>
<td>75.5</td>
<td>64.2</td>
<td>51.7</td>
<td>38.9</td>
</tr>
</tbody>
</table>
Figure B.3. Of subsampling method, performance comparison between the smallest and greatest value of block size $b$ used for $n = 500, 1000$. The top two plots are for the equal-tailed confidence interval, the bottom two for the symmetric confidence interval. The nominal coverage of 95% is marked by solid horizontal line.
Table B.5. Empirical coverage probabilities of equal-tailed and symmetric confidence interval based on the subsampling method for \( r_0 \) and \( r_1 \) in the CCC-GARCH model, with \( n = 1000 \).

<table>
<thead>
<tr>
<th>Model</th>
<th>( r_0 )</th>
<th>( r_1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( n )</td>
<td>( b )</td>
</tr>
<tr>
<td>equal-tailed</td>
<td>1000</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>90.6</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>90.8</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>83.0</td>
</tr>
<tr>
<td>symmetric</td>
<td>1000</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>89.0</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>90.5</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>84.7</td>
</tr>
</tbody>
</table>

Figure B.4. Of symmetric confidence intervals, performance comparison between the methods of residual bootstrap and subsampling. The length of sequence is \( n = 1000 \), and \( b = 10 \) is used for subsampling. The nominal coverage 95% is marked by solid horizontal line.
APPENDIX C
ADDITIONAL TABLES AND FIGURES OF CHAPTER 3
Table C.1. Empirical sizes (in percent) of the test using four different variance estimators applied to simulated series of squared GARCH(1,1) observations following the change-point models estimated from the returns on DJIA, NASDAQ, GE, and Wal-Mart in Table 4.1. The number of replications is $R = 5000$, with nominal size $\alpha = 1\%$.

<table>
<thead>
<tr>
<th>Method</th>
<th>$n$</th>
<th>$k^* = \frac{1}{4}n$</th>
<th>$k^* = \frac{1}{2}n$</th>
<th>$k^* = \frac{3}{4}n$</th>
<th>$k^* = \frac{1}{4}n$</th>
<th>$k^* = \frac{1}{2}n$</th>
<th>$k^* = \frac{3}{4}n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$q_n$</td>
<td>500</td>
<td>0.06</td>
<td>0.1</td>
<td>0.04</td>
<td>0.0</td>
<td>0.06</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>7.68</td>
<td>1.36</td>
<td>0.68</td>
<td>4.14</td>
<td>0.48</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>1500</td>
<td>17.74</td>
<td>2.58</td>
<td>1.08</td>
<td>12.78</td>
<td>0.96</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>23.62</td>
<td>3.34</td>
<td>1.38</td>
<td>16.82</td>
<td>0.9</td>
<td>0.32</td>
</tr>
<tr>
<td>$q_{arma}$</td>
<td>500</td>
<td>0.04</td>
<td>0.0</td>
<td>0.06</td>
<td>0.06</td>
<td>0.08</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.1</td>
<td>0.12</td>
<td>0.04</td>
<td>0.32</td>
<td>0.06</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>1500</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
<td>0.74</td>
<td>0.36</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>0.18</td>
<td>0.28</td>
<td>0.16</td>
<td>1.3</td>
<td>0.32</td>
<td>0.18</td>
</tr>
<tr>
<td>$q_{c arma}$</td>
<td>500</td>
<td>0.48</td>
<td>0.28</td>
<td>0.2</td>
<td>0.78</td>
<td>0.52</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.52</td>
<td>0.44</td>
<td>0.4</td>
<td>2.02</td>
<td>1.1</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>1500</td>
<td>0.68</td>
<td>0.24</td>
<td>0.5</td>
<td>3.34</td>
<td>1.2</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>0.8</td>
<td>0.7</td>
<td>0.54</td>
<td>3.72</td>
<td>1.1</td>
<td>0.72</td>
</tr>
<tr>
<td>Prewhitening</td>
<td>500</td>
<td>7.14</td>
<td>4.68</td>
<td>3.82</td>
<td>3.62</td>
<td>1.7</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>7.88</td>
<td>3.76</td>
<td>3.66</td>
<td>6.02</td>
<td>1.6</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>1500</td>
<td>8.66</td>
<td>3.78</td>
<td>3.06</td>
<td>8.68</td>
<td>1.12</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>9.2</td>
<td>3.0</td>
<td>2.46</td>
<td>9.72</td>
<td>1.06</td>
<td>0.78</td>
</tr>
<tr>
<td>$q_n$</td>
<td>500</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.06</td>
<td>0.1</td>
<td>2.34</td>
<td>0.14</td>
<td>0.18</td>
<td>1.48</td>
</tr>
<tr>
<td></td>
<td>1500</td>
<td>0.32</td>
<td>0.22</td>
<td>3.94</td>
<td>0.22</td>
<td>0.26</td>
<td>6.32</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>0.42</td>
<td>0.46</td>
<td>5.12</td>
<td>0.2</td>
<td>0.48</td>
<td>10.42</td>
</tr>
<tr>
<td>$q_{arma}$</td>
<td>500</td>
<td>0.08</td>
<td>0.08</td>
<td>0.16</td>
<td>0.04</td>
<td>0.1</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.28</td>
<td>0.3</td>
<td>0.36</td>
<td>0.18</td>
<td>0.4</td>
<td>1.34</td>
</tr>
<tr>
<td></td>
<td>1500</td>
<td>0.42</td>
<td>0.4</td>
<td>0.3</td>
<td>0.28</td>
<td>0.28</td>
<td>3.58</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>0.44</td>
<td>0.58</td>
<td>0.52</td>
<td>0.24</td>
<td>0.62</td>
<td>4.32</td>
</tr>
<tr>
<td>$q_{c arma}$</td>
<td>500</td>
<td>0.4</td>
<td>0.26</td>
<td>0.4</td>
<td>0.18</td>
<td>0.44</td>
<td>1.14</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.68</td>
<td>0.54</td>
<td>0.66</td>
<td>0.38</td>
<td>0.74</td>
<td>2.9</td>
</tr>
<tr>
<td></td>
<td>1500</td>
<td>0.64</td>
<td>0.66</td>
<td>0.56</td>
<td>0.62</td>
<td>0.6</td>
<td>5.0</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>0.88</td>
<td>0.74</td>
<td>0.8</td>
<td>0.48</td>
<td>1.06</td>
<td>5.72</td>
</tr>
<tr>
<td>Prewhitening</td>
<td>500</td>
<td>1.12</td>
<td>1.12</td>
<td>4.44</td>
<td>0.78</td>
<td>0.82</td>
<td>2.76</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>1.04</td>
<td>1.32</td>
<td>4.42</td>
<td>0.96</td>
<td>1.02</td>
<td>7.36</td>
</tr>
<tr>
<td></td>
<td>1500</td>
<td>1.18</td>
<td>1.06</td>
<td>3.2</td>
<td>0.84</td>
<td>1.32</td>
<td>10.62</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>0.9</td>
<td>1.24</td>
<td>2.88</td>
<td>0.68</td>
<td>1.58</td>
<td>12.62</td>
</tr>
</tbody>
</table>
Table C.2. Empirical sizes (in percent) of the test using four different variance estimators applied to simulated series of squared GARCH(1,1) observations following the change-point models estimated from the returns on DJIA, NASDAQ, GE, and Wal-Mart in Table 4.1. The number of replications is \( R = 5000 \), with nominal size \( \alpha = 10\% \).

<table>
<thead>
<tr>
<th>Method</th>
<th>DJIA : Model 1 → Model 2</th>
<th>NASDAQ : Model 3 → Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( k^* = \frac{1}{4}n )</td>
<td>( k^* = \frac{1}{3}n )</td>
</tr>
<tr>
<td></td>
<td>( k^* = \frac{1}{4}n )</td>
<td>( k^* = \frac{1}{3}n )</td>
</tr>
<tr>
<td>( q_n )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>17.5</td>
<td>9.24</td>
</tr>
<tr>
<td>1000</td>
<td>39.44</td>
<td>17.32</td>
</tr>
<tr>
<td>1500</td>
<td>46.8</td>
<td>21.14</td>
</tr>
<tr>
<td>2000</td>
<td>48.96</td>
<td>21.98</td>
</tr>
<tr>
<td>( q_{arma} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>3.58</td>
<td>2.08</td>
</tr>
<tr>
<td>1000</td>
<td>4.84</td>
<td>3.38</td>
</tr>
<tr>
<td>1500</td>
<td>7.16</td>
<td>5.4</td>
</tr>
<tr>
<td>2000</td>
<td>7.82</td>
<td>6.58</td>
</tr>
<tr>
<td>( q_{c arma} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>16.66</td>
<td>13.8</td>
</tr>
<tr>
<td>1000</td>
<td>14.56</td>
<td>12.12</td>
</tr>
<tr>
<td>2000</td>
<td>15.5</td>
<td>12.06</td>
</tr>
<tr>
<td>Prewhitening</td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>30.08</td>
<td>23.92</td>
</tr>
<tr>
<td>1000</td>
<td>30.32</td>
<td>20.7</td>
</tr>
<tr>
<td>1500</td>
<td>31.82</td>
<td>18.92</td>
</tr>
<tr>
<td>2000</td>
<td>31.6</td>
<td>16.54</td>
</tr>
<tr>
<td>( q_{arma} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>3.74</td>
<td>3.9</td>
</tr>
<tr>
<td>1000</td>
<td>5.98</td>
<td>6.0</td>
</tr>
<tr>
<td>1500</td>
<td>7.64</td>
<td>6.9</td>
</tr>
<tr>
<td>2000</td>
<td>7.62</td>
<td>8.48</td>
</tr>
<tr>
<td>( q_{c arma} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>9.98</td>
<td>9.68</td>
</tr>
<tr>
<td>1500</td>
<td>11.06</td>
<td>9.46</td>
</tr>
<tr>
<td>Prewhitening</td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>9.28</td>
<td>10.66</td>
</tr>
<tr>
<td>1000</td>
<td>10.24</td>
<td>10.8</td>
</tr>
<tr>
<td>1500</td>
<td>9.86</td>
<td>10.22</td>
</tr>
<tr>
<td>2000</td>
<td>9.46</td>
<td>10.94</td>
</tr>
</tbody>
</table>
Figure C.1. Comparison of empirical sizes for truncation lags $q^*_{arma}$ and $q^*_{c.arma}$. Results are reported for the change-point models for NASDAQ and Wal-Mart. The straight line indicates the nominal size of 5%.

Table C.3. Fitted EGARCH(1,1) and TGARCH(1,1) models from the returns on DJIA and Altria Group, and General Electric and American Express, respectively.

<table>
<thead>
<tr>
<th>Asymmetric GARCH</th>
<th>Speculative Asset</th>
<th>Period</th>
<th>No. of obs.</th>
<th>Period</th>
<th>$\mu$</th>
<th>$\omega$</th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$\gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGARCH (1,1)</td>
<td>DJIA</td>
<td>1/1/1992 ~</td>
<td>A 1061</td>
<td>12/31/1999</td>
<td>0.067</td>
<td>-0.193</td>
<td>0.142</td>
<td>0.907</td>
<td>-0.578</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12/31/1999</td>
<td>B 960</td>
<td></td>
<td>0.046</td>
<td>-0.037</td>
<td>0.048</td>
<td>0.986</td>
<td>-0.999</td>
</tr>
<tr>
<td></td>
<td>Altria</td>
<td>7/1/1997 ~</td>
<td>C 389</td>
<td></td>
<td>0.001</td>
<td>0.312</td>
<td>0.188</td>
<td>0.615</td>
<td>-0.999</td>
</tr>
<tr>
<td></td>
<td>Group</td>
<td>12/31/1999</td>
<td>D 242</td>
<td></td>
<td>-0.088</td>
<td>0.035</td>
<td>0.051</td>
<td>0.958</td>
<td>-0.999</td>
</tr>
<tr>
<td>TGARCH (1,1)</td>
<td>General</td>
<td>1/1/2001 ~</td>
<td>E 473</td>
<td></td>
<td>-0.171</td>
<td>1.301</td>
<td>0.032</td>
<td>0.672</td>
<td>0.245</td>
</tr>
<tr>
<td></td>
<td>Electric</td>
<td>6/30/2004</td>
<td>F 402</td>
<td></td>
<td>0.049</td>
<td>0.021</td>
<td>0.029</td>
<td>0.957</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>American</td>
<td>1/1/2001 ~</td>
<td>G 475</td>
<td></td>
<td>-0.121</td>
<td>0.368</td>
<td>-0.025</td>
<td>0.881</td>
<td>0.198</td>
</tr>
<tr>
<td></td>
<td>Express</td>
<td>6/30/2004</td>
<td>H 398</td>
<td></td>
<td>0.046</td>
<td>-0.004</td>
<td>0.006</td>
<td>0.971</td>
<td>0.053</td>
</tr>
</tbody>
</table>
Figure C.2. Simulated LARCH observations with $a = 0.85$, $b_0 = 0.35$, and $d = 0.35$. The lengths of the series in the four panels are $n = 500, 1000, 2000, \text{ and } 1500$ clockwise.
APPENDIX D

ADDITIONAL TABLES AND FIGURES OF CHAPTER 4
Table D.1. Empirical sizes (in percent) for the four monitoring methods applied to simulated series of squared GARCH(1,1) observations following Model 1 and Model 7 in Table 4.1. The number of replications is $R = 1000$ with historical size $m = 100$ and autocovariances computed sequentially.

<table>
<thead>
<tr>
<th>Method</th>
<th>$q$</th>
<th>Model 1</th>
<th></th>
<th></th>
<th>Model 7</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Bartlett</td>
<td>QS</td>
<td>Bartlett</td>
<td>QS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5%</td>
<td>10%</td>
<td>5%</td>
<td>10%</td>
<td>5%</td>
<td>10%</td>
</tr>
<tr>
<td>CS</td>
<td>$m$</td>
<td>2.9</td>
<td>6.2</td>
<td>3.7</td>
<td>6.3</td>
<td>3.0</td>
<td>6.4</td>
</tr>
<tr>
<td></td>
<td>2m</td>
<td>4.9</td>
<td>9.4</td>
<td>5.6</td>
<td>9.5</td>
<td>5.2</td>
<td>9.5</td>
</tr>
<tr>
<td></td>
<td>3m</td>
<td>6.7</td>
<td>12.1</td>
<td>7.4</td>
<td>12.1</td>
<td>6.5</td>
<td>11.2</td>
</tr>
<tr>
<td>FL</td>
<td>$m$</td>
<td>8.2</td>
<td>12.3</td>
<td>10.1</td>
<td>14.3</td>
<td>8.3</td>
<td>11.1</td>
</tr>
<tr>
<td></td>
<td>2m</td>
<td>10.6</td>
<td>16.3</td>
<td>12.3</td>
<td>17.4</td>
<td>10.4</td>
<td>14.1</td>
</tr>
<tr>
<td></td>
<td>3m</td>
<td>12.6</td>
<td>18.4</td>
<td>13.9</td>
<td>19.4</td>
<td>11.8</td>
<td>16.7</td>
</tr>
<tr>
<td>PS</td>
<td>$m$</td>
<td>5.5</td>
<td>9.6</td>
<td>6.8</td>
<td>10.5</td>
<td>4.7</td>
<td>8.0</td>
</tr>
<tr>
<td></td>
<td>2m</td>
<td>9.6</td>
<td>15.4</td>
<td>10.5</td>
<td>16.8</td>
<td>8.2</td>
<td>12.7</td>
</tr>
<tr>
<td></td>
<td>3m</td>
<td>12.3</td>
<td>18.1</td>
<td>12.8</td>
<td>19.2</td>
<td>11.1</td>
<td>16.4</td>
</tr>
<tr>
<td>RR</td>
<td>$m$</td>
<td>1.5</td>
<td>3.0</td>
<td>1.7</td>
<td>3.1</td>
<td>1.3</td>
<td>2.9</td>
</tr>
<tr>
<td></td>
<td>2m</td>
<td>3.7</td>
<td>7.0</td>
<td>3.9</td>
<td>7.7</td>
<td>3.7</td>
<td>7.2</td>
</tr>
<tr>
<td></td>
<td>3m</td>
<td>6.7</td>
<td>12.1</td>
<td>6.6</td>
<td>11.2</td>
<td>5.7</td>
<td>10.4</td>
</tr>
</tbody>
</table>

Table D.2. Empirical sizes (in percent) for the four monitoring methods applied to simulated series of squared GARCH(1,1) observations following Model 1 and Model 7 in Table 4.1. The number of replications is $R = 1000$ with historical size $m = 200$ and autocovariances computed sequentially.

<table>
<thead>
<tr>
<th>Method</th>
<th>$q$</th>
<th>Model 1</th>
<th></th>
<th></th>
<th>Model 7</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Bartlett</td>
<td>QS</td>
<td>Bartlett</td>
<td>QS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5%</td>
<td>10%</td>
<td>5%</td>
<td>10%</td>
<td>5%</td>
<td>10%</td>
</tr>
<tr>
<td>CS</td>
<td>$m$</td>
<td>3.1</td>
<td>5.4</td>
<td>2.5</td>
<td>5.0</td>
<td>3.4</td>
<td>5.8</td>
</tr>
<tr>
<td></td>
<td>2m</td>
<td>4.7</td>
<td>7.8</td>
<td>4.4</td>
<td>8.1</td>
<td>4.3</td>
<td>7.9</td>
</tr>
<tr>
<td></td>
<td>3m</td>
<td>6.0</td>
<td>10.2</td>
<td>5.9</td>
<td>10.2</td>
<td>5.0</td>
<td>8.9</td>
</tr>
<tr>
<td>FL</td>
<td>$m$</td>
<td>10.6</td>
<td>13.8</td>
<td>7.2</td>
<td>9.7</td>
<td>7.7</td>
<td>10.2</td>
</tr>
<tr>
<td></td>
<td>2m</td>
<td>13.1</td>
<td>16.7</td>
<td>8.6</td>
<td>13.3</td>
<td>8.7</td>
<td>12.2</td>
</tr>
<tr>
<td></td>
<td>3m</td>
<td>13.7</td>
<td>17.8</td>
<td>10.3</td>
<td>14.9</td>
<td>9.3</td>
<td>13.6</td>
</tr>
<tr>
<td>PS</td>
<td>$m$</td>
<td>6.7</td>
<td>11.2</td>
<td>4.2</td>
<td>7.5</td>
<td>4.1</td>
<td>7.5</td>
</tr>
<tr>
<td></td>
<td>2m</td>
<td>10.3</td>
<td>15.7</td>
<td>7.4</td>
<td>13.4</td>
<td>6.6</td>
<td>11.1</td>
</tr>
<tr>
<td></td>
<td>3m</td>
<td>11.8</td>
<td>18.9</td>
<td>9.5</td>
<td>15.5</td>
<td>8.1</td>
<td>13.4</td>
</tr>
<tr>
<td>RR</td>
<td>$m$</td>
<td>1.9</td>
<td>3.7</td>
<td>1.4</td>
<td>2.2</td>
<td>1.0</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>2m</td>
<td>3.7</td>
<td>7.8</td>
<td>3.5</td>
<td>6.1</td>
<td>2.2</td>
<td>4.1</td>
</tr>
<tr>
<td></td>
<td>3m</td>
<td>6.0</td>
<td>10.7</td>
<td>5.7</td>
<td>9.6</td>
<td>2.9</td>
<td>5.5</td>
</tr>
</tbody>
</table>
Table D.3. Empirical sizes (in percent) for the four monitoring methods applied to simulated series of squared GARCH(1,1) observations following Model 1 and Model 7 in Table 4.1. The number of replications is \( R = 1000 \) with historical size \( m = 300 \) and autocovariances computed sequentially.

<table>
<thead>
<tr>
<th>Method</th>
<th>Method</th>
<th>Model 1</th>
<th>Model 7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Bartlett</td>
<td>QS</td>
</tr>
<tr>
<td></td>
<td>( q )</td>
<td>5%</td>
<td>10%</td>
</tr>
<tr>
<td>CS</td>
<td>( m )</td>
<td>4.3</td>
<td>7.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.6</td>
<td>9.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7.4</td>
<td>11.2</td>
</tr>
<tr>
<td>FL</td>
<td>( m )</td>
<td>7.5</td>
<td>12.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9.7</td>
<td>14.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10.5</td>
<td>16.1</td>
</tr>
<tr>
<td>PS</td>
<td>( m )</td>
<td>4.8</td>
<td>8.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.2</td>
<td>13.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10.2</td>
<td>15.8</td>
</tr>
<tr>
<td>RR</td>
<td>( m )</td>
<td>1.7</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.1</td>
<td>7.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.8</td>
<td>9.2</td>
</tr>
</tbody>
</table>

Table D.4. Empirical sizes (in percent) for the four monitoring methods applied to simulated series of squared GARCH(1,1) observations following Model 1 and Model 7 in Table 4.1. The number of replications is \( R = 1000 \) with historical size \( m = 400 \) and autocovariances computed sequentially.

<table>
<thead>
<tr>
<th>Method</th>
<th>Method</th>
<th>Model 1</th>
<th>Model 7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Bartlett</td>
<td>QS</td>
</tr>
<tr>
<td></td>
<td>( q )</td>
<td>5%</td>
<td>10%</td>
</tr>
<tr>
<td>CS</td>
<td>( m )</td>
<td>3.1</td>
<td>5.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.8</td>
<td>7.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.6</td>
<td>8.9</td>
</tr>
<tr>
<td>FL</td>
<td>( m )</td>
<td>7.8</td>
<td>11.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9.4</td>
<td>13.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10.1</td>
<td>14.3</td>
</tr>
<tr>
<td>PS</td>
<td>( m )</td>
<td>4.6</td>
<td>7.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7.5</td>
<td>11.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.7</td>
<td>13.3</td>
</tr>
<tr>
<td>RR</td>
<td>( m )</td>
<td>0.9</td>
<td>1.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.9</td>
<td>5.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.2</td>
<td>6.7</td>
</tr>
</tbody>
</table>
Table D.5. Empirical sizes (in percent) for the four monitoring methods applied to simulated series of squared GARCH(1,1) observations following Model 1 and Model 7 in Table 4.1. The number of replications is $R = 1000$ with historical size $m = 500$ and autocovariances computed using the historical $m$ observations.

<table>
<thead>
<tr>
<th>Method</th>
<th>$q$</th>
<th>Model 1</th>
<th></th>
<th></th>
<th>Model 7</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Bartlett</td>
<td>QS</td>
<td></td>
<td>Bartlett</td>
<td>QS</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5%</td>
<td>10%</td>
<td>5%</td>
<td>10%</td>
<td>5%</td>
<td>10%</td>
</tr>
<tr>
<td>CS</td>
<td>$m$</td>
<td>5.6</td>
<td>9.5</td>
<td>6.7</td>
<td>10.4</td>
<td>7.8</td>
<td>12.0</td>
</tr>
<tr>
<td></td>
<td>$2m$</td>
<td>8.8</td>
<td>13.5</td>
<td>10.3</td>
<td>13.2</td>
<td>11.2</td>
<td>16.2</td>
</tr>
<tr>
<td></td>
<td>$3m$</td>
<td>11.0</td>
<td>16.4</td>
<td>12.4</td>
<td>16.8</td>
<td>13.3</td>
<td>18.5</td>
</tr>
<tr>
<td>FL</td>
<td>$m$</td>
<td>14.0</td>
<td>17.5</td>
<td>14.4</td>
<td>18.5</td>
<td>17.9</td>
<td>20.5</td>
</tr>
<tr>
<td></td>
<td>$2m$</td>
<td>16.1</td>
<td>20.3</td>
<td>16.1</td>
<td>22.0</td>
<td>19.9</td>
<td>22.9</td>
</tr>
<tr>
<td></td>
<td>$3m$</td>
<td>17.3</td>
<td>21.5</td>
<td>17.6</td>
<td>23.3</td>
<td>20.3</td>
<td>23.9</td>
</tr>
<tr>
<td>PS</td>
<td>$m$</td>
<td>9.1</td>
<td>13.8</td>
<td>10.5</td>
<td>14.5</td>
<td>12.7</td>
<td>16.6</td>
</tr>
<tr>
<td></td>
<td>$2m$</td>
<td>12.9</td>
<td>18.3</td>
<td>14.1</td>
<td>19.9</td>
<td>15.8</td>
<td>20.6</td>
</tr>
<tr>
<td></td>
<td>$3m$</td>
<td>14.7</td>
<td>21.4</td>
<td>15.6</td>
<td>22.4</td>
<td>17.1</td>
<td>23.6</td>
</tr>
<tr>
<td>RR</td>
<td>$m$</td>
<td>4.4</td>
<td>6.1</td>
<td>5.1</td>
<td>7.9</td>
<td>5.9</td>
<td>9.0</td>
</tr>
<tr>
<td></td>
<td>$2m$</td>
<td>7.8</td>
<td>10.8</td>
<td>7.9</td>
<td>12.8</td>
<td>9.6</td>
<td>13.9</td>
</tr>
<tr>
<td></td>
<td>$3m$</td>
<td>9.7</td>
<td>14.8</td>
<td>11.1</td>
<td>16.5</td>
<td>12.5</td>
<td>17.2</td>
</tr>
</tbody>
</table>

Figure D.1. Comparison of the distributions of the first hitting times. Results are reported for $m = 100$ with break at $n = 111$, 10% controlled size and the Bartlett kernel. The number of replications is $R = 1000$.

Note: The historical sample size $m = 100$ is subtracted from the hitting times for the ease of visual comparison.
Table D.6. Empirical power and ARL of the four methods applied to the two types of variance changes. The number of replications is $R = 1000$, with controlled size of 10%, and $\gamma$ computed using the historical $m$ observations. The monitoring horizon is $q = 500$.

<table>
<thead>
<tr>
<th>Method</th>
<th>$m$</th>
<th>break</th>
<th>Model 1 → Model 2</th>
<th>Model 7 → Model 8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Bartlett</td>
<td>Q5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>power (ARL)</td>
<td>power (ARL)</td>
</tr>
<tr>
<td>CS</td>
<td>100</td>
<td>111</td>
<td>99.3 (31)</td>
<td>97.6 (53)</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>331</td>
<td>99.8 (42)</td>
<td>99.1 (73)</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>551</td>
<td>100 (51)</td>
<td>98.4 (92)</td>
</tr>
<tr>
<td>FL</td>
<td>100</td>
<td>111</td>
<td>99.6 (20)</td>
<td>97.8 (41)</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>331</td>
<td>100 (24)</td>
<td>99.3 (43)</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>551</td>
<td>100 (26)</td>
<td>99.5 (47)</td>
</tr>
<tr>
<td>PS</td>
<td>100</td>
<td>111</td>
<td>99.6 (22)</td>
<td>98.2 (41)</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>331</td>
<td>100 (30)</td>
<td>99.4 (48)</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>551</td>
<td>100 (34)</td>
<td>99.5 (58)</td>
</tr>
<tr>
<td>RR</td>
<td>100</td>
<td>111</td>
<td>99.6 (29)</td>
<td>98.3 (51)</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>331</td>
<td>99.9 (44)</td>
<td>99.3 (69)</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>551</td>
<td>100 (53)</td>
<td>99.4 (86)</td>
</tr>
</tbody>
</table>

Note: The historical sample size $m = 100$ is subtracted from the hitting times for the ease of visual comparison.

Figure D.2. Comparison of the distributions of the first hitting times. Results are reported for $m = 500$ with break at $n = 551$, 10% controlled size and the Bartlett kernel. The number of replications is $R = 1000$. 
CURRICULUM VITAE

Aonan Zhang
May 2005

Professional objective:
A quantitative researcher position which requires skills in GARCH modeling and empirical analysis of financial data.

Prominent attributes:
◊ Quantitative analysis of financial returns data, ARCH and related time series.
◊ Detection of volatility changes.
◊ Proficient programming skills in C++, S-PLUS, R, and SAS.

Education:
◊ Ph.D. in Mathematical Science, May 2005. Utah State University, Logan, Utah.

Publications and working papers:
◊ Confidence intervals for the autocorrelations of the squared of GARCH sequences, with Piotr Kokoszka and Gilles Teyssiére. Published in "Computational Science - ICCS 2004."
◊ Monitoring constancy of variance in conditionally heteroskedastic time series, with Lajos Horváth and Piotr Kokoszka, forthcoming in "Econometric Theory."
◊ Discriminating between long memory and volatility changes, with Piotr Kokoszka, under revision.

Work experience:
◊ Assistant economist: Bank of Communications, QingDao, China, Aug 1997 – May 1999.
◊ Instructor: Utah State University, Spring 2000 – Fall 2004.