Bootstrap-based Test for Volatility Shifts in GARCH Against Long-range Dependence

by

Yu Wang

(Under the direction of Cheolwoo Park)

Abstract

Measuring volatility has been one of the most critical steps in financial modeling. One characteristic of volatility is persistence. Numerous methods have been proposed to test whether such persistency is due to volatility shifts in the market or a natural fluctuation explained by stationary long-range dependence. Lee et al. (2014) propose a residual-based cumulative sum test statistic to test volatility shifts in generalized autoregressive conditional heteroscedasticity model against long-range dependence. This thesis continues the study of Lee et al. (2014). It compares the asymptotic size and power proposed by them and the bootstrap size and power. Baek and Pipiras (2012) revise the test statistics based on the local Whittle estimation of the self-similar parameter. This thesis also compares the bootstrap size and power of the local Whittle statistics to the test statistics for no volatility shift by Lee et al. (2014).

Index words: Bootstrap, Change point estimation, GARCH models, Hypothesis test, Long-range dependence, Size and power, Volatility shifts,

Bootstrap-based Test for Volatility Shifts in GARCH Against Long-range Dependence

by

Yu Wang

B.B.A., Yantai University, 2011

M.S.B.A., Colorado State University, 2013

A Dissertation Submitted to the Graduate Faculty
of The University of Georgia in Partial Fulfillment
of the
Requirements for the Degree

Master of Science

Athens, Georgia

2015

Bootstrap-based Test for Volatility Shifts in GARCH Against Long-range Dependence

by

Yu Wang

Approved:

Major Professor: Cheolwoo Park

Committee: Ping Ma
Sue Ryung Chang

Electronic Version Approved:

Julie Coffield
Interim Dean of the Graduate School
The University of Georgia
May 2015

Acknowledgments

I take this opportunity to express gratitude to my advisor Dr. Cheolwoo Park and the visiting professor Dr. Taewook Lee from HanKuk University of Foreign Studies in Souel, Korea for their valuable guidance on this thesis. I also thank Dr. Ping Ma and Dr. Sue Ryung Chang for their time serving as the committee. Special thanks to my families and friends for their support through out my Master's study.

Contents

List of Tables

Chapter 1

Introduction

1.1 Motivation

In finance, volatility is a variation measure for return of a financial instrument over time. Measuring volatility has been one of the most critical steps in financial modeling. A volatility model can be used to forecast the absolute magnitude of returns or the density of the volatility in risk management, asset pricing, and portfolio selection etc.

Generally, there are two classes of volatility models (Engle and Patton, 2001), latent volatility or (misleading) stochastic volatility models that shocks are not all functions of observable and volatility models that are functions of observable, such as autoregressive conditional heteroscedasticity (ARCH) and generalized autoregressive conditional heteroscedasticity (GARCH). For the former one, the models are usually difficult to estimate and forecast because they are arbitrarily elaborated with jumps and fat-tailed shocks, multiple factors, general types of nonlinearities and so on. They are also hard to simulate because not all the functions and distributions are known. Therefore, when referencing to volatility models, most research forces on the class that can be expressed as an ARCH or GARCH model, though the former class might be more realistic.

Engle and Patton (2001) document some common features of asset price volatility processes. They include:

1. Persistence. This behavior means that high volatility period will be followed by high volatility period and low volatility period by low volatility period.

2. Mean reverting. High volatility will eventually die out and the series will go back to normal volatility level, and low volatility period back to normal volatility level.

3. The conditional volatility of the underlying asset may be affected differently by the positive and negative shocks.

4. Exogenous variables may influence the volatility.

5. Unconditional distribution of assets returns might have heavy tails.

This thesis, as well as some of the research in the area of financial time series, focuses mainly on the persistence character of the volatility. Numerous methods have been proposed to test whether such persistency is due to volatility shifts in the market or a natural fluctuation explained by stationary long-range dependence.

1.2 Background

GARCH Model

Traditional econometrics assumes constant variance over the period of study. Researchers also use the strategy called rolling standard deviation or variance in a short rolling window to capture the time-varying conditional standard deviation or conditional variance. Engle (1982) introduces the autoregressive conditional heteroscedastic (ARCH) process. It has constant unconditional variance and non-constant conditional variance. The forecasted current period value depends on the changing conditional variance.

Engle (1982) mentions several attractive characteristics of the ARCH model. It explains the change in variance of financial assets portfolios and other models in the theory of finance that are inappropriate explained by exogenous variables. It also has the ability to model and predict the varies from one period to another.

A more general process, generalized autoregressive conditional heteroscedasticity (GARCH) model, has also been a popular tool to analyze financial time series data since introduced by Bollerslev (1986). It takes a weighted average of lagged squared returns and the lagged conditional variance. An univariate GARCH (p, q) model satisfies:

$$x_t = \sigma_t \epsilon_t, \ \sigma_t^2 = \omega^\circ + \sum_{i=1}^{q} \alpha_i^\circ x_{t-i}^2 + \sum_{j=1}^{p} \beta_j^\circ \sigma_{t-j}^2, \tag{1.1}$$

where the innovation $\{\epsilon_t\}_{t \in \mathbb{Z}}$ is a sequence of standard i.i.d.random variables. It is also assumed that $\omega^\circ > 0$, $\alpha_i^\circ \geq 0$ for all $i = 1, \ldots, q$ and $\beta_j^\circ \geq 0$ for all $j = 1, \ldots, p$. The GARCH models are conditionally heteroskedastic, but have a constant unconditional variance. It is said that the volatility is highly persistent when the sum of the estimated coefficients of the squared lagged returns and the lagged conditional variance terms in a GARCH model is close to 1.

Empirically, both ARCH and GARCH models can successfully model financial time series with clustering of volatilities. In finance, volatility clustering can be understood as news clustering. While asset pricing attempts to reflect the market news on the price as accurate as possible, successfully modeling of the volatility clustering makes ARCH and GARCH models attractive to researchers in this area.

Long-range Dependent

One way of dealing with the high persistence of volatility is to use long-range dependent (LRD) modeling of time series on transformed data, such as squared return. Beran (1994) points out that scientists in various areas have observed correlations between observations that decay to zero at a slow rate. This long memory phenomenon has been documented long before suitable models were developed. Ding et al. (1993) investigate the long-memory property of the stock market return series. They find that other than the absolute returns, the power transformation of the absolute returns also has "quite high autocorrelations for long lags".

A LRD time series $\{x_t\}_{t\in\mathbb{Z}}$ is defined as a second-order stationary time series model with a slowly decaying autocovariance function,

$$\gamma(h) = Cov(x_t, x_{t+h}) \sim Ch^{2d-1} = Ch^{2H-2}, \text{ as } h \to \infty, \tag{1.2}$$

where $C > 0$ is a constant,

$$d \in \left(0, \frac{1}{2}\right)$$

is the long-range dependence parameter, and

$$H = d + \frac{1}{2} \in \left(\frac{1}{2}, 1\right)$$

is the self-similarity (SS) parameter.

Based on Equation (1.2), for LRD series, the autocovariances are not absolute summable. That is, $\sum_{h=-\infty}^{\infty} |\gamma(h)| = +\infty$. If, on the other hand, the autocovariances are absolute summable, $\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$, the series is short-range dependence (SRD).

Bootstrap Method

Another key tool of the thesis is the bootstrap method that first proposed by Efron (1979).

Traditionally, in a hypothesis test, if we know the cumulative distribution function (CDF) of the test statistics under the null hypothesis, we would reject the null hypothesis when the p-value of the realized test statistic is smaller than the given significant level. Or, alternatively, in a upper-tailed test, we would reject the null hypothesis if the calculated critical value under the CDF is greater than the critical value at a given significant level.

If the CDF is unknown, researchers would provide an asymptotic one. Most of the asymptotic distribution results are based on the large-sample properties. Therefore, results can be inaccurate when the sample size is small. Bootstrap has been developed to solve the inaccuracy caused by small sample size in statistical inference.

According to MacKinnon (2006), there are many different types of bootstrap methods. Not all of them are easy to implement or work well in certain cases. For example, Gonçalves and Kilian (2004) propose three easy-to-implement bootstrap methods: fixed-design wild bootstrap, the recursive-design wild bootstrap, and the pairwise bootstrap. They find these methods tend to be more accurate in the simulation study on small samples, and argue that these methods "should routinely replace conventional bootstrap procedures for autoregressions based on the i.i.d. error assumption". In this thesis, however, we use the standard residual-based bootstrap for time series in our method, given the i.i.d. innovation series. The bootstrap procedures will be presented in details in Chapter 3.

1.3 Objective of the Study

We attempt to distinguish between volatility shift and LRD based on a bootstrap approach. The main goal of this thesis is to compare the asymptotic and the bootstrap size and power of the test for volatility-shift versus LRD, using the residual-based cumulative sum (CUSUM) test statistic adapted by Lee et al. (2014). We also compare the bootstrap size and power between the test statistic by Lee et al. (2014) and the one based on the local Whittle estimator of the SS parameter proposed by Baek and Pipiras (2012).

The organization of the thesis is as follows. Chapter 2 introduces more details about the tests on volatility shift versus LRD. Chapter 3 proposes the standard residual-based bootstrap method on time series data and the revised test statistic based on the local Whittle estimator of the SS parameter. Simulation results are reported in Chapter 4. We conclude in Chapter 5.

Chapter 2

Literature Review

2.1 Residual-based CUSUM Test for No Volatility Shifts Against LRD

For numbers of statistical tests to distinguish LRD and volatility shift, the procedure can be divided into two classes according to the null hypotheses: either LRD as the null hypothesis (e.g. Qu, 2011; Kuswanto, 2011), or the volatility-shift model as the null (Jach and Kokoszka, 2008). In this thesis, we use the latter one as the null hypothesis. Specifically, the tests consider:

$$H_0^{(R)} : \text{VS - R model versus } H_1 : \text{LRD model},$$

where $VS-R$ represents that volatility shifts occurred at R *unknown* locations. We fit each of the $(R+1)$ regimes using a GARCH model.

Researchers propose various ways to estimate the unknown change location for the volatility. The most popular one is the cumulative sum (CUSUM) type of test statistic that is first considered by Kokoszka et al. (2000), where they use the squared asset returns for a single volatility shift. According to Andreou and Ghysels (2002), this test suffers from size distortions under finite sample because of the strong dependence on observations. The Bartlett long-run variance estimator can be used to correct the size distortion (see Berkes et al. 2006, and Zhang et al. 2007). However, the Bartlett long-run variance estimator is very sensitive to choice of the kernel bandwidth. To overcome the flaw, Lee et al. (2014) adapts the residual-based CUSUM test. The test uses the standardized residuals from an estimated GARCH model, which imitate the innovation series in GARCH model and therefore reduce the dependence among observations. Moreover, unlike the statistic proposed by Lee et al. (2004), the residual-based CUSUM test statistic does not require tuning parameters for truncation and can be used for various types of volatility models.

It is widely reported that the power transformation of returns, such $\{x_t^2\}_{t\in\mathbb{Z}}$, displays recognizable autocorrelations. Therefore, the series $\{x_t^2\}_{t\in\mathbb{Z}}$ satisfies (1.2) and will be used for the alternative hypothesis. One of the popular models for LRD is the fractionally integrated GARCH (FIGARCH) model of Baillie et al. (1996). We will use this as the alternative model in the power of the test. This will be described in more details in the simulation chapter. To be more specific, the first test we consider is:

$H_0^{(0)}$: The observed data $\{x_t\}_{t\in\mathbb{Z}}$ follow the VS-0 model.

H_1 : The observed data $\{x_t^2\}_{t\in\mathbb{Z}}$ follow the LRD model.

In order to conduct the test, one might consider the CUSUM test statistic basesd on $\{x_t\}_{t\in\mathbb{Z}}$:

$$T_n = \frac{1}{\sqrt{n}s_n} \max_{1\le k\le n} \left| \sum_{t=1}^{k} x_t^2 - \frac{k}{n} \sum_{t=1}^{n} x_t^2 \right|,$$

where n is the sample size, and s_n^2 is the estimator of the long-run variance $\sigma^2 = \sum_{h=-\infty}^{\infty} Cov(x_t^2, x_{t+h}^2)$. Note that the CUSUM test for GARCH model suffers from size distortions and low powers under finite sample size due to the strong correlation between observations. Therefore, Lee et al. (2014) consider the CUSUM test based on the i.i.d.

innovation series $\{\epsilon_t\}$ in equation (1.1), which are uncorrelated observations. That is,

$$\hat{T}_n = \frac{1}{\sqrt{n}\tau} \max_{1\le k\le n} \left| \sum_{t=1}^{k} \epsilon_t^2 - \frac{k}{n}\sum_{t=1}^{n} \epsilon_t^2 \right|.$$

Note that $\{\epsilon_t\}$ is unobservable. One can estimate the innovation series by

$$\tilde{\epsilon}_t := \frac{x_t}{\tilde{\sigma}_t(\hat{\theta})},\ t = 1, \ldots, n, \tag{2.1}$$

where $\hat{\theta} = (\hat{\omega}_0, \hat{\alpha}_1, \ldots, \hat{\alpha}_q, \hat{\beta}_1, \ldots, \hat{\beta}_p)'$ is the estimated parameter vector of the GARCH model, and $\tilde{\sigma}_t^2(\theta)$ is calculated recursively from

$$\tilde{\sigma}_t^2(\theta) = \omega + \sum_{i=1}^{q} \alpha_i x_{t-i}^2 + \sum_{j=1}^{p} \beta_j \tilde{\sigma}_{t-j}^2(\theta),$$

where $\theta = (\omega, \alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_p)'$ is the parameter vector. Note that calculating $\tilde{\sigma}_t^2$ recursively requires the initial values for $x_0^2, \ldots, x_{1-q}^2$and $\tilde{\sigma}_0^2, \ldots, \tilde{\sigma}_{1-p}^2$. The `fGarch` package in R provides the initial values to be the sample average of $x_1^2, , \ldots, x_n^2$. Finally, the residual-based CUSUM statistic is defined as

$$\tilde{T}_n = \frac{1}{\sqrt{n}\hat{\tau}} \max_{1\le k\le n} \left| \sum_{t=1}^{k} \tilde{\epsilon}_t^2 - \frac{k}{n}\sum_{t=1}^{n} \tilde{\epsilon}_t^2 \right|, \tag{2.2}$$

where

$$\hat{\tau}^2 = \frac{1}{n}\sum_{t=1}^{n} \tilde{\epsilon}_t^4 - \Big(\frac{1}{n}\sum_{t=1}^{n} \tilde{\epsilon}_t^2\Big)^2$$

is a method of moment estimator of $\mathrm{Var}(\epsilon_1^2)$.

Lee et al. (2014) also prove that under certain assumptions and conditions, and under $H_0^{(0)}$, as $n \to \infty$,

$$\tilde{T}_n \xrightarrow{d} \sup_{0\le u\le 1} |\mathcal{B}^\circ(u)|,$$

where $\mathcal{B}^{\circ}(u)$ is a standard Brownian bridge. We will use this as the asymptotic distribution of the residual-based CUSUM test statistic $\tilde{T}_n$. According to Resnick (1992), for a standard Brownian motion, the formula for the distribution of $\tilde{T}_n$ can be derived as

$$\Pr[\tilde{T}_n \leq \upsilon] = 1 + 2\sum_{k=1}^{\infty}(-1)^k e^{-2k^2\upsilon^2},\ \upsilon > 0. \tag{2.3}$$

The p-value calculated based on (2.3) will be shown in the simulation chapter.

2.2 Residual-based CUSUM Test for A Single Volatility Shift Against LRD

Once the null hypothesis of no volatility shift is rejected, one will consider the test on volatility shift for *known* R number of times with *unknown* locations against LRD. Therefore, the hypothesis would become:

$H_0^{(R)}$: The observed data $\{x_t\}_{t\in\mathbb{Z}}$ follow the VS-R model.

H_1 : The observed data $\{x_t^2\}_{t\in\mathbb{Z}}$ follow the LRD model.

We follow Lee et al. (2014) and only consider the simple case when R = 1 in this thesis, which is the single-volatility-shift case. The testing procedure can be easily extended to multiple volatility shifts.

Suppose the whole series consists of two GARCH (p, q) models, namely $\{x_{1,t}\}_{t\in\mathbb{Z}}$ and $\{x_{2,t}\}_{t\in\mathbb{Z}}$, with the same standard i.i.d. innovations. That is,

$$x_t = \begin{cases} x_{1,t},\ \text{if } 1 \leq t \leq k^*, \\ x_{2,t},\ \text{if } k^* < t \leq n, \end{cases} \tag{2.4}$$

where k^* is an unknown change point. By Kokoszka et al. (2000), it then can be estimated by

$$\hat{k} = \underset{1\le k\le n}{\operatorname{argmax}} \left| \frac{k(n-k)}{n^2} \left(\frac{1}{k} \sum_{t=1}^{k} x_t^2 - \frac{1}{n-k} \sum_{t=k+1}^{n} x_t^2 \right) \right| \tag{2.5}$$

and

$$\begin{cases} x_{1,t} = \sigma_{1,t}\epsilon_t,\ \sigma_{1,t}^2 = \omega_1^\circ + \sum_{i=1}^q \alpha_{1,i}^\circ x_{1,t-i}^2 + \sum_{j=1}^p \beta_{1,j}^\circ \sigma_{1,t-j}^2, \\ x_{2,t} = \sigma_{2,t}\epsilon_t,\ \sigma_{2,t}^2 = \omega_2^\circ + \sum_{i=1}^q \alpha_{2,i}^\circ x_{2,t-i}^2 + \sum_{j=1}^p \beta_{2,j}^\circ \sigma_{2,t-j}^2. \end{cases} \tag{2.6}$$

Assume again that $\omega_m^\circ > 0$, $\alpha_{m,i}^\circ \ge 0$ for all $i = 1, \dots, q$ and $\beta_{m,j}^\circ \ge 0$ for all $j = 1, \dots, p$, for $m = 1, 2$. The test statistic is given by Lee et al. (2014) as

$$M_n = \max\{\tilde{T}_{n,1}, \tilde{T}_{n,2}\}, \tag{2.7}$$

where

$$\tilde{T}_{n,1} = \frac{1}{\hat{\tau}_1} \hat{k}^{\frac{1}{2}} \max_{1\le k\le \hat{k}} \left| \sum_{t=1}^{k} \tilde{\epsilon}_t^2 - \frac{k}{\hat{k}} \sum_{t=1}^{\hat{k}} \tilde{\epsilon}_t^2 \right|,\ \tilde{T}_{n,2} = \frac{1}{\hat{\tau}_2} \left(n - \hat{k} \right)^{\frac{1}{2}} \max_{\hat{k}<k\le n} \left| \sum_{t=\hat{k}+1}^{k} \tilde{\epsilon}_t^2 - \frac{k-\hat{k}}{n-\hat{k}} \sum_{t=\hat{k}+1}^{n} \tilde{\epsilon}_t^2 \right|,$$

and

$$\hat{\tau}_1^2 = \frac{1}{\hat{k}} \sum_{t=1}^{\hat{k}} \tilde{\epsilon}_t^4 - \left(\frac{1}{\hat{k}} \sum_{t=1}^{\hat{k}} \tilde{\epsilon}_t^2 \right)^2,\ \ \hat{\tau}_2^2 = \frac{1}{n-\hat{k}} \sum_{t=\hat{k}+1}^{n} \tilde{\epsilon}_t^4 - \left(\frac{1}{n-\hat{k}} \sum_{t=\hat{k}+1}^{n} \tilde{\epsilon}_t^2 \right)^2.$$

Similar to the last section, define recursively that

$$\tilde{\sigma}_{1,t}^2(\theta_1) = \omega_1 + \sum_{i=1}^{q} \alpha_{1,i} \tilde{x}_{1,t-i}^2 + \sum_{j=1}^{p} \beta_{1,j} \tilde{\sigma}_{1,t-j}^2(\theta_1),\ \ t = 1, \dots, \hat{k},$$

$$\tilde{\sigma}_{2,t}^2(\theta_2) = \omega_2 + \sum_{i=1}^{q} \alpha_{2,i} \tilde{x}_{2,t-i}^2 + \sum_{j=1}^{p} \beta_{2,j} \tilde{\sigma}_{2,t-j}^2(\theta_2),\ \ t = \hat{k}+1, \dots, n$$

with appropriate initial values. Using these equations, Lee et al. (2014) then calculate $\{\tilde{\epsilon}_t, t = 1, \ldots, n\}$ by

$$\tilde{\epsilon}_t = \begin{cases} \frac{x_t}{\tilde{\sigma}_{1,t}\left(\hat{\theta}_1\right)}, \ t = 1, \ldots, \hat{k}, \\ \frac{x_t}{\tilde{\sigma}_{2,t}\left(\hat{\theta}_2\right)}, \ t = \hat{k}+1, \ldots, n. \end{cases} \tag{2.8}$$

Similarly, they also prove that if certain conditions hold, for the VS-1 model (2.6) under $H_0^{(1)}$, as $n \to \infty$,

$$M_n \xrightarrow{d} \max\left\{ \sup_{0 \le u \le 1} |\mathcal{B}_1^\circ(u)|, \ \sup_{0 \le u \le 1} |\mathcal{B}_2^\circ(u)| \right\},$$

where $\mathcal{B}_1^\circ(u)$ and $\mathcal{B}_2^\circ(u)$ are independent standard Brownian bridges. After calculating M_n, we can then use equation (2.3) to calculate the p-value of the test.

Note that we consider only when the single volatility shift is known *in priori*. If the number of volatility shift is *unknown*, Bai (1997) mentions the so-called binary segmentation (BS) method to estimate the number of change points. To apply it in our case, the procedure will be:

1. Start with the whole sample $\{x_t\}_{t \in \mathbb{Z}}$, and perform the test of $H_0^{(R)} : VS-0$ using the test statistic in (2.2).

2. If the null hypothesis is not rejected at a given level of significance, we conclude there is no volatility shift in the series and stop at this step. If the null is rejected at the significant level, continue to next step.

3. To test $H_0^{(R)} : VS-1$, estimate the first change point $\hat{k}_1$ as in (2.5), and calculate the test statistic given in (2.7).

4. If the null hypothesis is not rejected at a given level of significance, set $\hat{R} = 1$ and stop at this step.

5. If the null is rejected, split the whole sample into two subsamples, $\{r_1, \ldots, r_{\hat{k}_1}\}$ and $\{r_{\hat{k}_1+1}, \ldots, r_n\}$. Then for each of the subsamples, repeat Steps 3 - 4 until the null is not rejected.

Since we only consider the simplest case where the only volatility shift is *known*, the above procedures will not be implemented in this thesis. But they are necessary when using the real data and the number of volatility shifts is *unknown*.

Chapter 3

Proposed Methods

3.1 Bootstrap Method

The use of the asymptotic test statistics for the finite sample might give size distortion and lower power. Efron (1979) introduces the bootstrap method to study the distributions of estimators (sampling distribution) and test statistics by resampling the finite sample data. He calls the method bootstrap to express the situation that a person falls into the water and has to pull his/her own bootstrap in order to get him/herself out of the water.

In this thesis, we compare the size and power of the tests presented in Sections 2.1 and 2.2 using the asymptotic and bootstrap-based methods. We also compare the size and power of the test in Section 2.1 using the bootstrap method and the local Whittle method, which will be described in details in the next section.

If the data $\{x_1, \ldots, x_n\}$ are independent observations, a commonly used bootstrap procedure is to sample m observation, $\{x_1^*, \ldots, x_m^*\}$, from the original data with replacement, and calculate the resample's estimators or test statistics, say t^*. Repeat this procedure B times and get $\{t_1^*, \ldots, t_m^*\}$. Then the distribution of bootstrap replication can approximate the distribution of the estimator or test statistic under the null hypothesis. If one performs bootstrap for the test statistic, then the right-sided bootstrap p-value can be calculated as

$$\hat{p}^* = \frac{1}{B} \sum_{b=1}^{B} I(t_b^* \geq t),$$

where $I(\cdot)$ is the indicator function and t is the test statistic from the original data.

However, time series data are autocorrelated, and therefore are not independent to each other. Simply resampling the original data might yield even worse results than the asymptotic estimation does. One method to deal with such problem in time series is to bootstrap the estimated i.i.d. innovation series $\{\tilde{\epsilon}_t\}$, and then generate recursively a bootstrap sample $\{x_t^*\}$ (Ahlgren and Catani, 2012). Specifically, the algorithm for bootstrap residual-based CUSUM test for no volatility shift against LRD is as follows:

1. Compute the test statistic $\tilde{T}_n$ by equation (2.2) from the original data. Obtain the estimated GARCH model parameters and calculate the innovation series $\{\tilde{\epsilon}_t\}$ according to equation (2.1).

2. Construct the bootstrap innovation series $\{\tilde{\epsilon}_t^*\}$ with size equal to the original sample size n, by drawing independently with replacement from $\{\tilde{\epsilon}_t\}$.

3. Generate recursively a bootstrap sample $\{x_t^*\}$ using the estimated parameters and the bootstrap innovation series from Steps 1 and 2, respectively.

4. Compute the bootstrap test statistics $\tilde{T}_n^*$ from the bootstrap sample.

5. Repeat Steps 2 - 4 B times to obtain $\tilde{T}_{n_1}^*, \ldots, \tilde{T}_{n_B}^*$.

6. Estimate the bootstrap p-value by

$$\hat{p}^* = \frac{1}{B} \sum_{b=1}^{B} I(\tilde{T}_{n_b}^* \geq \tilde{T}_n),$$

   where $I(\cdot)$ is the indicator function.

For the test in Section 2.2 and the test statistic M_n, the bootstrap procedure is given as:

1. Compute the test statistic M_n from the original data and get the estimated change point k^* by equation (2.7) and (2.5), respectively.

2. Obtain the estimated parameters of the two GARCH models as defined by equation (2.4), call them model 1 and model 2, respectively. Calculate two innovation series by equation (2.8), call them $\{\tilde{\epsilon}_{1,t}\}$ and $\{\tilde{\epsilon}_{2,t}\}$ for model 1 and model 2, respectively.

3. Construct the bootstrap innovation series $\{\tilde{\epsilon}^*_{s,t}\}$ with size equal to the original sample size, by drawing independently with replacement from $\{\tilde{\epsilon}_{s,t}\}$, for $s = 1, 2$.

4. Generate recursively a bootstrap sample $\{x^*_{s,t}\}$ using the estimated parameters and the bootstrap innovation series from Steps 2 - 3, for $s = 1, 2$.

5. Obtain the bootstrap test statistics M^*_n by comparing $\tilde{T}^*_{n,s}$ calculated from the bootstrap sample $\{x^*_{s,t}\}$, for $s = 1, 2$.

6. Repeat Steps 2 - 5 B times to obtain $M^*_{n_1}, \ldots, M^*_{n_B}$.

7. Estimate the bootstrap p-value by

$$\hat{p}^* = \frac{1}{B} \sum_{b=1}^{B} I(M^*_{n_b} \geq M_n),$$

   where $I(\cdot)$ is the indicator function.

Refer to the Appendix for the R code of these two bootstrap algorithms.

3.2 Test of No Volatility Shifts Based on the Local Whittle Estimation of SS Parameter

Baek and Pipiras (2012) propose the local Whittle estimator for the test of a single change in *mean* against LRD. They define the local Whittle estimator of the SS parameter for the series $\{x_t\}_{t\in\mathbb{Z}}$ as

$$\hat{H}_{lw} = \underset{H\in\Theta}{\operatorname{argmin}}\, x(H),$$

where $\Theta = [\Delta_1,\ \Delta_2]$ with $0 < \Delta_1 < \Delta_2 < 1$ and

$$x(H) = \log\left(\frac{1}{m}\sum_{l=1}^{m}\omega_l^{2H-1} l_x(\omega_l)\right) - (2H-1)\frac{1}{m}\sum_{l=1}^{m}\log\omega_l, \tag{3.1}$$

with m denoting the number of low frequencies used in estimation (Robinson, 1995). Moreover,

$$l_x(\omega_l) = \frac{1}{2\pi n}\left|\sum_{j=1}^{n} x_j e^{-ij\omega_l}\right|^2 \tag{3.2}$$

denotes the periodogram of the series $\{x_t\}_{t\in\mathbb{Z}}$ at the Fourier frequencies

$$\omega_l = \frac{2\pi l}{t},\ l = 1,\ldots,t.$$

Baek and Pipiras (2012) then define the test statistics as

$$M_{lw} = \sqrt{m}\left(\hat{H}_{lw} - \frac{1}{2}\right). \tag{3.3}$$

If one substitutes the x_t in equations (3.1) and (3.2) with its second order x_t^2, the test statistic can then be used for the test of volatility shift. Baek and Pipiras (2012) show that under the null hypothesis of change in mean at one unknown location (CM-1), equation (3.3) converges to the normal distribution with mean 0 and standard deviation 1/2. However, this is the case under null hypothesis of CM-1. They do not show the asymptotics under the null of VR-0, and it is not clear if there exists one. Bootstrap is a reliable method to compare the size and power of two test statistics under the situation of no asymptotic distribution. Hence, we also compare the bootstrap size and power of the test for no volatility shift against LRD presented in Section 2.1 and the revised test statistic M_{lw} based on the local Whittle estimation method. The bootstrap procedure for the latter one follows the similar steps as that for the test statistic $\tilde{T}_n$ in Section 2.1.

Chapter 4

Simulation Study

4.1 Simulation Settings

The main goal of the thesis is to test the volatility shift in GARCH models. We consider the univariate GARCH(1,1) models with non-negative coefficients satisfying $\alpha + \beta < 1$. That is,

$$x_t = \sigma_t \epsilon_t, \ \sigma_t^2 = \omega + \alpha x_{t-1}^2 + \beta \sigma_{t-1}^2,$$

where $\{\epsilon_t\}_{t \in \mathbb{Z}}$ is the the innovation series following specific distributions.

In the simulations for test for no volatility shift against LRD, we consider the model with two types of innovations: i.i.d. standard normal and t-distribution with a degree of freedom 5, denoted by $t(5)$, following the setting from Lee et al. (2014). We also follow their 6 sets of GARCH(1,1) parameters (ω, α, β) in the simulation, with $\alpha + \beta$ ranging from 0.5 to very close to 1. In the test for one volatility shift against LRD, the innovation series are only considered as i.i.d. standard normal. The GARCH(1,1) parameters shift at the midpoint of the samples from $(\omega, \alpha, \beta) = (0.1, 0.1, 0.8)$ to 4 other sets with $\alpha + \beta$ ranging from 0.5 to very close to 1. Note that the closer $\alpha + \beta$ to 1, the more persistent the series is. A favorable performance of empirical size on such a parameter setting indicates a promising test statistic.

For the empirical power, we consider only the Fractionally Integrated Generalized AutoRegressive Conditionally Heteroskedastic (FIGARCH) (1,d,1), by Baillie et al. (1996), as the alternative. That is,

$$x_t = \sigma_t \epsilon_t,\ (1-\beta L)\sigma_t^2 = \omega + \big((1-\beta L) + (1-\phi L)(1-L)^d\big)x_t^2,$$

where $E[\epsilon_t|I^{t-1}] = 0, Var[\epsilon_t|I^{t-1}] = 1$, given that $I^{t-1} = \{\ldots, \epsilon_{t-2}, \epsilon_{t-1}\}$ is the information set up to time $t-1$. Here, L is the backshift operator. In the model, the power d of the fractional differencing operator, $(1-L)^d$, is allowed to be non-integer, and the term can be expanded as

$$(1-L)^d = \sum_{i=0}^{\infty} \pi_i L^i$$

where

$$\pi_i = \prod_{0 \le k \le i} \frac{k-1-d}{k},$$

and $d \in (-1/2, 1/2)$. For $d > 0$, the process is long memory. The FIGARCH model implies a slow hyperbolic rate of decay for the lagged squared innovations (Baillie et al., 1996), and combines features of the fractionally integrated process for the mean together with the regular GARCH process for the conditional variance. It provides good estimation for the samples typically encountered with finance data.

In the simulation of empirical power, we set FIGARCH$(1, d, 1)$ parameters as $(\omega, \beta, \phi) = (0.6, 0.1, 0.2)$ and $(0.2, 0.2, 0.2)$, each with three different values of d.

All results are based on 1000 replications on sample sizes 1000, 2000, and 5000 with significant level 5%. The bootstrap replication B is 100.

4.2 Simulation Results

Table 4.1 presents the empirical size of the test for no volatility shift against LRD (refer to the hypothesis test in Chapter 2) under different sample size and types of innovations.

It contains the asymptotic (columns under "Asymptotic") and bootstrap (columns under "Bootstrap") size from the test statistics $\tilde{T}_n$ by Lee et al. (2014).

By comparing the result from test $\tilde{T}_n$ under two methods with different sample sizes and innovation types, the bootstrap method yields a better (not too conservative) size within acceptable range (smaller than 0.06) than the asymptotic ones, even $\alpha+\beta$ are very close to 1. This difference is more evident when the innovation part has the heavy-tailed distribution. Moreover, the bootstrap size is closer to the nominal 5% significance level when sample size is small. When the sample size becomes larger, the bootstrap size gets similar to the asymptotic size when the innovation is normal, but still better than the asymptotic one when the innovation is heavy-tailed. Hence, the bootstrap method might be a good choice when data show the evidence of heavy tailed innovation or the sample size is small.

The power comparison is summarized in Table 4.2. the bootstrap method provides more power than the asymptotic one, even if the data are strongly long-range dependent ($d = 0.45$). But as the sample size increases, the power of both methods increases dramatically, and the advantage of the bootstrap diminishes.

When comparing the bootstrap size between $\tilde{T}_n$ and M_{lw} in Table 4.3, it is clear that the former one is more promising when the sample size is small. When $n = 5000$, bootstrap M_{lw} has larger size closer to the nominal 5%. However, it also suffers from size distortion when $\alpha + \beta$ getting close to 1.

For test M_{lw} in Table 4.4, the power also increases as the sample size gets larger as $\tilde{T}_n$ does, under both parameter settings. But M_{lw} loses its power when the series is strongly long-range dependent.

Table 4.5 presents the empirical size of the test for one volatility shift against LRD (refer to the hypotheses in Chapter 2). The data clearly show that, for test $\tilde{M}_n$, the bootstrap size is more assuring under small sample size. The asymptotic size gets close to the nominal significance level as the sample size gets larger, while size distortion is observed for bootstrap size. It is noteworthy that one might observe size distortion when the series after the change point has different weight on lag volatility, holding other parameters the same. Hence, the asymptotic method provides acceptable size under different sample sizes and parameters, and the bootstrap method should be used with more caution due to the size distortion.

Table 4.6 shows the empirical power of test for one volatility shift. Overall, the power increases significantly as the sample size increases, given the same long-range dependence. Power also increases with the dependence of the series. Similar to the size of this test, bootstrap method yields an much better power than the asymptotic method when the sample size is small. This difference is barely observed as the sample size gets larger or the series gets more dependent. In conclusion, bootstrap is a better method for the test for one volatility shift against LRD. Both methods are acceptable when sample size is large, despite the dependence of the sample.

Table 4.1: Empirical Size of Test for No Volatility Shifts Under Test Statistic $\tilde{T}_n$

(ω, α, β)	$N(0,1)$		$t(5)$	
$n = 1000$	Asymptotic	Bootstrap	Asymptotic	Bootstrap
$(0.1, 0.1, 0.8)$	0.034	0.052	0.014	0.036
$(0.1, 0.1, 0.6)$	0.037	0.057	0.029	0.049
$(0.1, 0.1, 0.4)$	0.036	0.052	0.021	0.043
$(0.1, 0.2, 0.6)$	0.043	0.05	0.032	0.051
$(0.3, 0.1, 0.8)$	0.034	0.052	0.029	0.044
$(0.3, 0.1, 0.89)$	0.02	0.049	0.014	0.034
$n = 2000$				
$(0.1, 0.1, 0.8)$	0.04	0.048	0.027	0.041
$(0.1, 0.1, 0.6)$	0.047	0.058	0.028	0.046
$(0.1, 0.1, 0.4)$	0.051	0.054	0.025	0.04
$(0.1, 0.2, 0.6)$	0.052	0.054	0.023	0.034
$(0.3, 0.1, 0.8)$	0.04	0.048	0.035	0.047
$(0.3, 0.1, 0.89)$	0.027	0.044	0.035	0.047
$n = 5000$				
$(0.1, 0.1, 0.8)$	0.034	0.033	0.027	0.033
$(0.1, 0.1, 0.6)$	0.033	0.038	0.03	0.04
$(0.1, 0.1, 0.4)$	0.035	0.038	0.033	0.049
$(0.1, 0.2, 0.6)$	0.036	0.039	0.036	0.049
$(0.3, 0.1, 0.8)$	0.034	0.033	0.025	0.037
$(0.3, 0.1, 0.89)$	0.029	0.04	0.032	0.045

Table 4.2: Empirical Power of Test for No Volatility Shifts Under Test Statistic $\tilde{T}_n$

	$n = 1000$		$n = 2000$		$n = 5000$	
d	Asymptotic	Bootstrap	Asymptotic	Bootstrap	Asymptotic	Bootstrap
			FIGARCH$(1, d, 1)(\omega = 0.6, \beta = 0.1)$			
0.25	0.342	0.379	0.558	0.573	0.806	0.801
0.35	0.446	0.477	0.667	0.685	0.885	0.887
0.45	0.489	0.528	0.713	0.719	0.899	0.899
			FIGARCH$(1, d, 1)(\omega = 0.2, \beta = 0.2)$			
0.25	0.26	0.302	0.444	0.474	0.694	0.717
0.35	0.37	0.401	0.586	0.61	0.837	0.83
0.45	0.433	0.453	0.631	0.651	0.861	0.859

Table 4.3: Bootstrap Empirical Size of Test for No Volatility Shifts Under Test Statistic $\tilde{T}_n$ and Revised M_{lw}

(ω, α, β)	$N(0,1)$		$t(5)$	
$n = 1000$	$\tilde{T}_n$ Bootstrap	Revised M_{lw}	$\tilde{T}_n$ Bootstrap	Revised M_{lw}
$(0.1, 0.1, 0.8)$	0.052	0.023	0.036	0.028
$(0.1, 0.1, 0.6)$	0.057	0.025	0.049	0.024
$(0.1, 0.1, 0.4)$	0.052	0.034	0.043	0.028
$(0.1, 0.2, 0.6)$	0.05	0.037	0.051	0.037
$(0.3, 0.1, 0.8)$	0.052	0.023	0.044	0.028
$(0.3, 0.1, 0.89)$	0.049	0.062	0.034	0.066
$n = 2000$				
$(0.1, 0.1, 0.8)$	0.048	0.034	0.041	0.035
$(0.1, 0.1, 0.6)$	0.058	0.054	0.046	0.05
$(0.1, 0.1, 0.4)$	0.054	0.065	0.04	0.049
$(0.1, 0.2, 0.6)$	0.054	0.052	0.034	0.047
$(0.3, 0.1, 0.8)$	0.048	0.034	0.047	0.035
$(0.3, 0.1, 0.89)$	0.044	0.038	0.047	0.054
$n = 5000$				
$(0.1, 0.1, 0.8)$	0.033	0.045	0.033	0.044
$(0.1, 0.1, 0.6)$	0.038	0.056	0.04	0.045
$(0.1, 0.1, 0.4)$	0.038	0.059	0.049	0.039
$(0.1, 0.2, 0.6)$	0.039	0.062	0.049	0.053
$(0.3, 0.1, 0.8)$	0.033	0.045	0.037	0.044
$(0.3, 0.1, 0.89)$	0.04	0.064	0.045	0.053

Table 4.4: Bootstrap Empirical Power of Test for No Volatility Shifts Under Test Statistic $\tilde{T}_n$ and Revised M_{lw}

	$n = 1000$		$n = 2000$		$n = 5000$	
d	$\tilde{T}_n$	Revised M_{lw}	$\tilde{T}_n$	Revised M_{lw}	$\tilde{T}_n$	Revised M_{lw}
			FIGARCH$(1, d, 1)(\omega = 0.6, \beta = 0.1)$			
0.25	0.379	0.277	0.573	0.565	0.801	0.883
0.35	0.477	0.264	0.685	0.497	0.887	0.742
0.45	0.528	0.222	0.719	0.402	0.899	0.637
			FIGARCH$(1, d, 1)(\omega = 0.2, \beta = 0.2)$			
0.25	0.302	0.184	0.474	0.406	0.717	0.83
0.35	0.401	0.177	0.61	0.327	0.83	0.636
0.45	0.453	0.151	0.651	0.251	0.859	0.49

Table 4.5: Empirical Size of Test for One Volatility Shift

$(0.1, 0.1, 0.8)$	$n = 1000$		$n = 2000$		$n = 5000$	
to $(\omega_2, \alpha_2, \beta_2)$	Asymptotic	Bootstrap	Asymptotic	Bootstrap	Asymptotic	Bootstrap
(0.1, 0.1, 0.6)	0.011	0.058	0.037	0.072	0.057	0.069
(0.1, 0.1, 0.4)	0.015	0.044	0.032	0.061	0.05	0.068
(0.1, 0.2, 0.6)	0.016	0.058	0.032	0.058	0.032	0.042
(0.3, 0.1, 0.8)	0.01	0.043	0.023	0.058	0.043	0.055

Table 4.6: Empirical Power of Test for One Volatility Shift

	$n = 1000$		$n = 2000$		$n = 5000$	
d	Asymptotic	Bootstrap	Asymptotic	Bootstrap	Asymptotic	Bootstrap
			FIGARCH$(1, d, 1)(\omega = 0.6, \beta = 0.1)$			
0.25	0.188	0.27	0.41	0.452	0.686	0.696
0.35	0.245	0.346	0.513	0.564	0.813	0.821
0.45	0.301	0.45	0.55	0.597	0.841	0.847
			FIGARCH$(1, d, 1)(\omega = 0.2, \beta = 0.2)$			
0.25	0.115	0.2	0.272	0.347	0.575	0.616
0.35	0.166	0.264	0.416	0.483	0.736	0.747
0.45	0.218	0.313	0.457	0.539	0.762	0.773

Chapter 5

Conclusion

Lee et al. (2014) proposed test statistics for test for volatility shift against LRD based on the CUSUM statistic from residual series obtained after fitting the GARCH model. In this thesis, we have continued their study and compared the empirical size and power of the test statistics using their proposed asymptotic distribution and the bootstrap procedure from Ahlgren and Catani (2012). We have run tests based on the test statistics proposed by Lee et al. (2014). They are:

$$H_0^{(R)} : \text{VS - 0 model versus } H_1 : \text{LRD model},$$

and

$$H_0^{(R)} : \text{VS - 1 model versus } H_1 : \text{LRD model}.$$

We observed that, generally, the bootstrap method is more promising than the asymptotic one for the test for no volatility shift (size close to the nominal significant level and higher power). It can also yield higher power for the test for one volatility shift. But one might observe size distortion from the bootstrap result. It would be interesting to further study the bootstrap procedure of the test for one volatility shift, in order to generate reasonable size and higher power.

We also have revised the test statistics by Baek and Pipiras (2012) by substituting the x_t with the second power x_t^2 to test hypothesis of no volatility shift against LRD. Since the existence of an asymptotic distribution for the revised test statistic is unknown, we have only used the bootstrap size and power. We have compared the bootstrap result from this test statistic and the one by Lee et al. (2014). In short, the test statistic by Lee et al. (2014) might be a better choice given its higher power and no size distortion.

Bibliography

Niklas Ahlgren and Paul Catani. Wild bootstrap tests for autocorrelation in vector autoregressive models'. Technical report, Working Papers, 2012.

Elena Andreou and Eric Ghysels. Detecting multiple breaks in financial market volatility dynamics. *Journal of Applied Econometrics*, 17(5):579–600, 2002.

Changryong Baek and Vladas Pipiras. Statistical tests for a single change in mean against long-range dependence. *Journal of Time Series Analysis*, 33(1):131–151, 2012.

Jushan Bai. Estimation of a change point in multiple regression models. *Review of Economics and Statistics*, 79(4):551–563, 1997.

Richard T Baillie, Tim Bollerslev, and Hans Ole Mikkelsen. Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 74(1):3–30, 1996.

Jan Beran. *Statistics for long-memory processes*, volume 61. CRC Press, 1994.

István Berkes, Lajos Horváth, Piotr Kokoszka, and Qi-Man Shao. On discriminating between long-range dependence and changes in mean. *The annals of statistics*, pages 1140–1165, 2006.

Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327, 1986.

Zhuanxin Ding, Clive WJ Granger, and Robert F Engle. A long memory property of stock market returns and a new model. *Journal of empirical finance*, 1(1):83–106, 1993.

Bradley Efron. Bootstrap methods: another look at the jackknife. *The annals of Statistics*, pages 1–26, 1979.

Robert F Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007, 1982.

Robert F Engle and Andrew J Patton. What good is a volatility model. *Quantitative finance*, 1(2):237–245, 2001.

Sılvia Gonçalves and Lutz Kilian. Bootstrapping autoregressions with conditional heteroskedasticity of unknown form. *Journal of Econometrics*, 123(1):89–120, 2004.

Agnieszka Jach and Piotr Kokoszka. Wavelet-domain test for long-range dependence in the presence of a trend. *Statistics*, 42(2):101–113, 2008.

Piotr Kokoszka, Remigijus Leipus, et al. Change-point estimation in ARCH models. *Bernoulli*, 6(3):513–539, 2000.

Heri Kuswanto. A new simple test against spurious long memory using temporal aggregation. *Journal of Statistical Computation and Simulation*, 81(10):1297–1311, 2011.

Sangyeol Lee, Yasuyoshi Tokutsu, and Koichi Maekawa. The cusum test for parameter change in regression models with arch errors. *Journal of the Japan Statistical Society*, 34 (2):173–188, 2004.

Taewook Lee, Moosup Kim, and Changryong Baek. Tests for volatility shifts in GARCH against long-range dependence. *Journal of Time Series Analysis*, 2014.

James G MacKinnon. Bootstrap methods in econometrics*. *Economic Record*, 82(s1):S2–S18, 2006.

Zhongjun Qu. A test against spurious long memory. *Journal of Business & Economic Statistics*, 29(3):423–438, 2011.

Sidney I. Resnick. *Adventures in Stochastic Processes*. Birkhäuser Boston, Reading, Massachusetts, 1992.

Peter M Robinson. Gaussian semiparametric estimation of long range dependence. *The Annals of statistics*, pages 1630–1661, 1995.

Aonan Zhang, Robertas Gabrys, and Piotr Kokoszka. Discriminating between long memory and volatility shifts. *Statistics*, 36(4):253–275, 2007.

Appendix

Algorithm for Bootstrap of Test Statistic $\tilde{T}_n$

(Test for No Volatility Shift)

To calculate the empirical size, the code first simulates a series under the null hypothesis distribution, then calculate $\tilde{T}_n$ in the first step of the algrithm. The estimated parameters (ω, α, β) are stored in variable `est.par` with length 3. The innovation series is calculated and store in the variable `innov`, which is also the input of the bootstrap function `method2`. Other variables include:

`B`: bootstrap times, was set to 100

`n`: sample size

`n.s`: "burn-in period" for the sample

The calculation of the empirical power follow the same procedure except that the series generated under the alternative distribution.

```
    method2 = function(innov)
    # Input: bootstrap innovation series
{
  t.stat = rep(NA, times = B)
   # Store the test statistics calculated from the B time bootstrap
```

```
for (i in 1:B)
{
  # Bootstrap Innovation
  e.tilde = sample(innov,size=(n+n.s), replace = T)

  # Recursive Bootstrap Sample
  # Initial Values: s.boot=unconditional variance estimate
  s.boot = rep(NA, times = n+n.s)             # Bootstrap sigma
  x.boot = rep(NA, times = n+n.s)             # Bootstrap x

  if ( (est.par[2] + est.par[3]) < 1 )
    s.boot[1] = est.par[1] / ( 1- est.par[2] - est.par[3] )
  if ( (est.par[2] + est.par[3]) >= 1 )
    s.boot[1] = var(innov)

  x.boot[1] = sqrt(s.boot[1])*e.tilde[1]

  for (j in 2:(n+n.s))
  {
    s.boot[j] = est.par[1] + est.par[2]*(x.boot[j-1])^2
              + est.par[3]*s.boot[j-1]
    x.boot[j] = sqrt(s.boot[j])*e.tilde[j]
  }

  x.boot = x.boot[(n.s+1):(n.s+n)]
  t.stat[i] = T_hat(x.boot)$t.stat
```

```
    #T_hat is the function for the test statistic for no volatility shift
  }

  return(t.stat)
}

   # Bootstrap Innovation Series and Calculate p-value
    method2.out = method2(innov)
    p.method2[i] = mean(method2.out > Tn.hat[i])
    #Tn.hat[i] is the ith T_n out of 1000 repetition
    #in the first step of the algorithm
```

Algorithm for Bootstrap of Test Statistic $\tilde{M}_n$ (Test for One Volatility Shift)

To calculate the empirical size, the code first simulates a series under the null hypothesis distribution, then estimate the position of the change point k^* and calculate M_n in the first step of the algorithm. The estimated parameters $(\omega_1, \alpha_1, \beta_1)$ and $(\omega_2, \alpha_2, \beta_2)$ are stored in variable `est.par1` and `est.par2`, respectively, each with length 3. The innovation series for the two subsamples are calculated and store in the variable `innov1` and `innov2`, respectively. They are also the input of the bootstrap function `method2`. Other variables include:

`B`: bootstrap times, was set to 100

`n`: sample size

`n.s`: "burn-in period" for the sample

The calculation of the empirical power follow the same procedure except that the series generated under the alternative distribution.

```
method2 = function(innov1,innov2)
# Input: bootstrap innovation series
{

  k = 1
  # Store the test statistics calculated from the B time bootstrap
  t.stat1 = rep(NA, times = B)
  t.stat2 = rep(NA, times = B)
  #Get the length of the subsamples
  size1 = floor(1.1*length(innov1))
  size2 = floor(1.1*length(innov2))
```

```
####### Calculate Tn of every resample and compare
  ### For First Part ###
  while (k < B+1)
  {
    # Bootstrap Innovation
    e.tilde = sample(innov1, size=size1, replace = T)
    # Recursive Bootstrap Sample
    s.boot = rep(NA, times = size1)              # Bootstrap sigma
    x.boot = rep(NA, times = size1)              # Bootstrap x

    s.boot[1] = var(innov1)

    x.boot[1] = sqrt(s.boot[1])*e.tilde[1]

    for (j in 2:size1)
    {
      s.boot[j] = est.par1[1] + est.par1[2]*
                  (x.boot[j-1])^2 + est.par1[3]*s.boot[j-1]
      x.boot[j] = sqrt(s.boot[j])*e.tilde[j]
    }

    x.boot = x.boot[(floor(0.1*length(innov1))+1):size1]
    #T_n
    try.t.stat = try(T_hat(x.boot), TRUE)

    if(class(try.t.stat)=="try-error") {k=k-1
```

```
  else t.stat1[k] = try.t.stat$t.stat
  }

  k = k + 1
}
### End First Part ###

### For Second Part ###
k = 1
while (k < B+1)
{

  # Bootstrap Innovation
  e.tilde = sample(innov2, size=size2, replace = T)
  # Recursive Bootstrap Sample
  s.boot = rep(NA, times = size2)             # Bootstrap sigma
  x.boot = rep(NA, times = size2)             # Bootstrap x

  s.boot[1] = var(innov2)

  x.boot[1] = sqrt(s.boot[1])*e.tilde[1]

  for (j in 2:size2)
  {
    s.boot[j] = est.par2[1] + est.par2[2]*(x.boot[j-1])^2
                + est.par2[3]*s.boot[j-1]
    x.boot[j] = sqrt(s.boot[j])*e.tilde[j]
```

```
      }

      x.boot = x.boot[(floor(0.1*length(innov2))+1):size2]
      #T_n
      try.t.stat = try(T_hat(x.boot), TRUE)


      if(class(try.t.stat)=="try-error") {k=k-1
      else t.stat2[k] = try.t.stat$t.stat
      }

      k = k + 1
    }
    ### End Second Part ###

     t.stat = pmax(t.stat1, t.stat2)

  return(t.stat)
}

  ############ Method 2: Recursive Bootstrap ###########
  # Estimate two parts separately
  est1 = garchFit(formula = ~ garch(1,1),data = dat[1:khat],
                  cond.dist = innov.type, include.mean = FALSE, trace = F)
  est.par1 = est1@fit$par[1:3]
  innov1 = dat[1:khat]/est1@sigma.t
```

```
est2 = garchFit(formula = ~ garch(1,1),data = dat[(khat+1):n],
                cond.dist = innov.type, include.mean = FALSE, trace = F)
est.par2 = est2@fit$par[1:3]
innov2 = dat[(khat+1):n]/est2@sigma.t

# Bootstrap Innovation Series and Calculate p.boot1
method2.out = method2(innov1,innov2)
p.method2[i] = mean(method2.out > Tn.hat[i])
#Tn.hat[i] is the ith M_n out of 1000 repetition
#in the first step of the algorithm
```