MANHATTAN DISTANCE FOR MODEL ORDER CHANGE-POINT ESTIMATION IN GARCH MODELS

Irene Waithira Irungu

Doctor of Philosophy

Mathematics (Financial Option)

A Thesis submitted to Pan African University Institute of Science, Technology and innovation in partial fulfillment of the requirement for the award of the degree of Doctor of Philosophy in Mathematics (Financial Option)

2018

DECLARATION

This Thesis is my original work and has not been presented for a degree award in any other University

Signature   Date

———————————————   ———————————————

Registration Number   MF400-0001/2015

This Thesis has been submitted for examination with our approval as University Supervisors

Signature   Date

———————————————   ———————————————

Prof. Peter Mwita

Machakos University

Signature   Date

———————————————   ———————————————

Prof. Antony Waititu

Jomo Kenyatta University of Agriculture and Technology

DEDICATION

Dedicated to the memory of my dad, Irungu Kariuki, who worked so hard to ensure I got the best education.

ACKNOWLEDGMENT

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

List of Algorithms

ACRONYMS

CUSUM Cumulative Sum

GARCH Generalized AutoRegressive Conditional Heteroscedastic

K and L test Kokoszka and Leipus test

MGARCH Multivariate GARCH

QMLE Quasi Maximum Likelihood Estimator

SDE   Stochastic Differential Equations

TV-ARCH Time varying ARCH

ABSTRACT

Volatility is considered as a measure of risk of financial assets which is vital for prudent financial decision of different stakeholders. GARCH models have been commonly used to capture volatility dynamics of financial time series, particularly in modeling volatility of stock returns and pricing of futures and options. Despite the generalization of $GARCH\,(p,q)$ model to assume different model orders, $GARCH\,(1,1)$ model continues to be widely used by practitioners when modeling volatility of financial assets returns. A key assumption of the GARCH models utilized is that the processes are stationary. This assumption allows for model identifiability. Financial asset returns, however, often exhibit the volatility clustering property implying that assuming one GARCH model is a poor fit. The IGARCH model may be perceived as a solution to this problem as the assumption of stationarity is relaxed and thus the model is able to model persistent changes in volatility. However, the IGARCH model is prone to a shortcoming where the behavior of the process depends on the intercept. In this work, change-point estimation is proposed as a solution to deal with this problem where observed non-stationary series is assumed to be composed of a series of stationary series. The main objective of this work is therefore to propose an estimator for the change-point which is considered as the point in time at which the series departs from one stationary GARCH model with order $(1,1)$ to another stationary GARCH model with order (p,q). Given that plausible values for the model orders p and q can be arrived at through inspection of sample autocorrelations and sample partial autocorrelations of a squared returns series, a change-point estimator based on the Manhattan distance

of sample autocorrelation of squared series is proposed. The estimator is given as the first point in time at which the Manhattan distance is maximum. To facilitate the detection of multiple change-points, binary segmentation technique is applied to extend the single change-point detection algorithm.

The asymptotic consistency of the estimator is proven theoretically based on some properties specific to sequence of stationary random variables with finite second and fourth moments. The limit theory of the process generating the estimator is also established. The general theory of the sample autocovariance and sample autocorrelation functions of a stationary GARCH process forms the basis. Specifically the point processes theory is utilized to obtain their weak convergence limit at different lags. This is further extended to the change-point process. The limits are found to be generally random as a result of the infinite variance.

Monte Carlo simulations is used to examine the performance of the estimator when the sample size, size and position of change vary using the Adjusted Rand Indices. It is established that ARI increases and tends to one as the size of change increases irrespective of the sample size and of the source of change. Histograms are utilized to assess the sampling distribution of the change-point estimator.

The research culminates with the application of the change-point estimator in pricing American options. Comparison is made between the performance of the fitted GARCH models and Black-Scholes model by examining plots of the option prices against moneyness. The fitted piecewise GARCH model, following change-point detection, gives higher prices compared to the Black-Scholes when the option is out-of-the-money indicating that the volatility dymanics affect the prices of options. It is therefore important for a investor trading in American options to consider change-points within the volatility structure of a financial returns series when choosing an early exercise date.

Chapter 1

INTRODUCTION

1.1 Background of the Study

Modelling volatility of financial asset returns is particularly an important area in Finance. This is because volatility is considered to be a measure of risk when pricing financial instruments. GARCH models have been commonly used to capture volatility dynamics in financial time series particularly in modelling of stock returns, interest rates and pricing futures and options. Despite the generalization of $GARCH\,(p,q)$ model to assume different model orders, $GARCH\,(1,1)$ model continues to be widely used by practitioners when modeling volatility of financial assets returns$(X_t)_{t \in \mathbb{Z}}$. The model is chosen prior without implementing any statistical identification procedure. Following an investigation by Starica (2003) on how truthful the simple volatility dynamics imposed by a GARCH(1,1) process to the evolution of returns of main financial indexes, it was established that GARCH(1,1) process is not the data generating process for the series of returns on the indexes under study. However, it was established that the GARCH(1,1) dynamics provides a poor fit for longer horizon as the model provided a good fit for the data as a local, stationary approximation. Starica asserts this can be attributed to the fact that the parameters of the GARCH(1,1) model change significantly through time. Francq and Zakoian (2011) carried out a similar study and utilized 11

different stock return series and 13 exchange rate return series. It was established that the GARCH(1,2) and GARCH(1,3) models provided the overall best fit for some for stock returns data whereas GARCH(2,1) provided the best fit for some exchange rate return data. Out of the 24 series examined, the GARCH(1,1) model was rejected for 16 series. It can therefore be concluded that studies carried out to compare the suitability of GARCH(1,1) over other models are prone to the limitation of the return series used. According to Francq and Zakoian (2011) the practice of using $GARCH(1,1)$ is motivated by the belief that the model and its simplest asymmetric extensions is sufficient to capture the properties of financial series and that higher-order models may be unnecessarily complicated. This practice is not always statistically justified for a large number of series and conclude that $GARCH(1,1)$ model is certainly overrepresented in empirical studies.

A key assumption when determining the model parameters is that the series $(X_t)_{t\in\mathbb{Z}}$ is stationary as this ensures model identifiability as this ensures that there is a one-to-one correspondence between the distribution of data and the values of model parameters. However, financial time series is particularly is characterized by the property of volatility clustering where large changes in volatility are followed by large ones and small changes are followed by small ones. Volatility clustering can be attributed to different mechanisms used by investors as their trading strategy. Long term investors focus on the long-term behavior of asset prices whereas their counterparts who are the traders focus on short term price fluctuations. The behavior assumed by these investors in a financial market introduces heterogeneity in time scale which can be perceived to be one of the sources of volatility clustering. In a study carried out by LeBaron (2000) on an artificial market to examine the effects of diversity on price dynamics, it was established that the presence of heterogeneity in horizons may lead to an increase in return variability, as well as volatility-volume relationships similar to those of actual markets. Various trading strategies models co-exist but evolve with time where their relative profitability is viewed as a means of natural selection mechanism. These mod-

els can be perceived to be a set of decision rules mapping investors information set, say on price history, trading volume and other economic indicators, to a set of actions to buy, sell or not to trade. These evolutionary models may represent the evolution by a deterministic dynamical system or through the complex price dynamics it generates which have been found to mimic some statistical properties of the returns process, including volatility clustering Hommes et al. (2003). Volatility clustering can also be attributed to behavioral change. Lux-Marchesi model attributes volatility clustering, long memory and heavy tails in asset returns to behavioral switching Lux and Marchesi (2000). With this regard it may be considered that the market reflects behavior switch from that of a fundamentalist to that of a chartist where the former evaluates their performance from anticipated gain and the latter evaluates according to realized gains. This decision-making process is driven by an exogenous factors creating imbalances between demand and supply which may result to transient phases of destabilization. An outbreak of volatility occurs if the fraction of agents using chartist techniques surpasses a certain threshold value though such phases are brought to an end by stabilizing tendencies. Thus the series can be considered to display a stationary behaviour for some time then suddenly the variability changes, it stays constant for some time at this new value until another change occurs. This therefore afffirms that the series is non-stationary and thus fitting one stationary GARCH model is not appropriate.

The IGARCH(p,q) model is proposed to be the solution to this problem. In this model, the stationarity assumption is relaxed and is thus able to model persistent changes in volatility. However, the IGARCH(p,q) is prone to some shortcomings. According to Nelson (1991) the behavior of an IGARCH process depends on the intercept, such that, if the intercept is positive then the unconditional variance of the process grows linearly with time. In practice this means that the amplitude of the clusters of volatility to be parametrized by the model on the average increases over time. The rate of increase need not, however, be particularly rapid. If the intercept is zero in the IGARCH model, the realizations from the process collapse to zero almost surely. However, a potentially

disturbing fact is that the model assumes that the unconditional variance of the process to be modeled does not exist in that the variance may be infinite as shown by Terasvirta (2009) and Polzehl and Spokoiny (2006).

In contrast change-point detection can be considered as an alternative of dealing with the volatility clustering phenomenon. According to Mikosch and Starica (2004) the assumption of parameter constancy in GARCH models may not be appropriate especially when the series to be modeled are long. To overcome this problem of modeling financial time series in the presence of structural changes, the duo suggests that one option is to assume that the parameters change at specific points of time, divide the series into sub-series according to the location of the change-points and fit separate GARCH models to the sub-series. In line with this perspective, the non-stationary series is looked at as a union of several stationary series. This brings about the challenge of determining the location of change-points and their number because they are normally not known in advance.

1.2 Statement of the Problem

Empirical studies reveal the popular use of $GARCH(1,1)$ to model volatility of an entire financial time series as highlighted in the background of the study. Volatility clustering, however, necessitates for the identification of the points in time, change-points, at which the series departs from $GARCH(1,1)$ model to $GARCH(p,q)$ model with different model order specification. Various tests for detection of structural breaks exists. A key assumption made in these tests is that structural breaks are as a result of change in one or more parameters in the GARCH(p,q) model. However, the changes could also be as a result of different orders specifications for the model, say, $GARCH(p,q)$ for $t \in [0,k]$, $GARCH(p^*,q)$ for $t \in [k+1,m]$, $GARCH(p,q^*)$ for $t \in [m+1,n]$ and $GARCH(p^*,q^*)$ for $t \in [n+1,T]$ where $0 < i < m < n < T$ and $p* \neq p$ and $q* \neq q$ and $p \neq q$. The study particularly seeks to investigate for the departure of the series

$(X_t)_{t \in \mathbb{Z}}$ from *GARCH* $(1,1)$ to *GARCH* (p,q) and vice versa. An assessment of the plots in Figures 1.1, 1.2 and 1.3 reinforces this possibility. From the plots, arising from different model-order specification, it is observed that the auto-correlation and the partial auto-correlation functions for different segments of the time series are different.

(a)   (b)
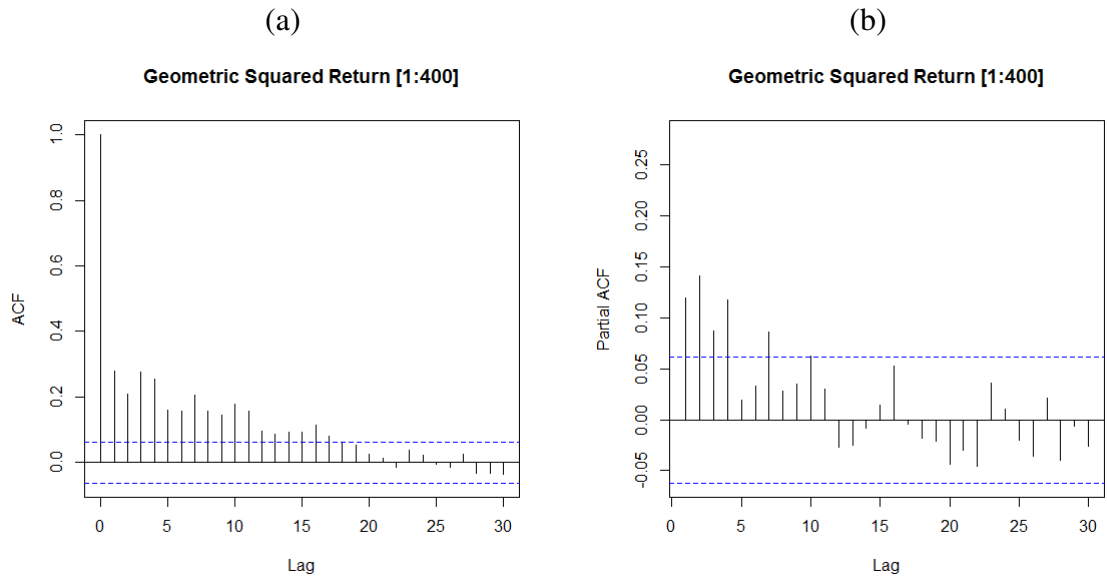
Figure 1.1: ACF and PACF of Geometric Squared Return for Series $t \in [1 : 400]$

(a)   (b)
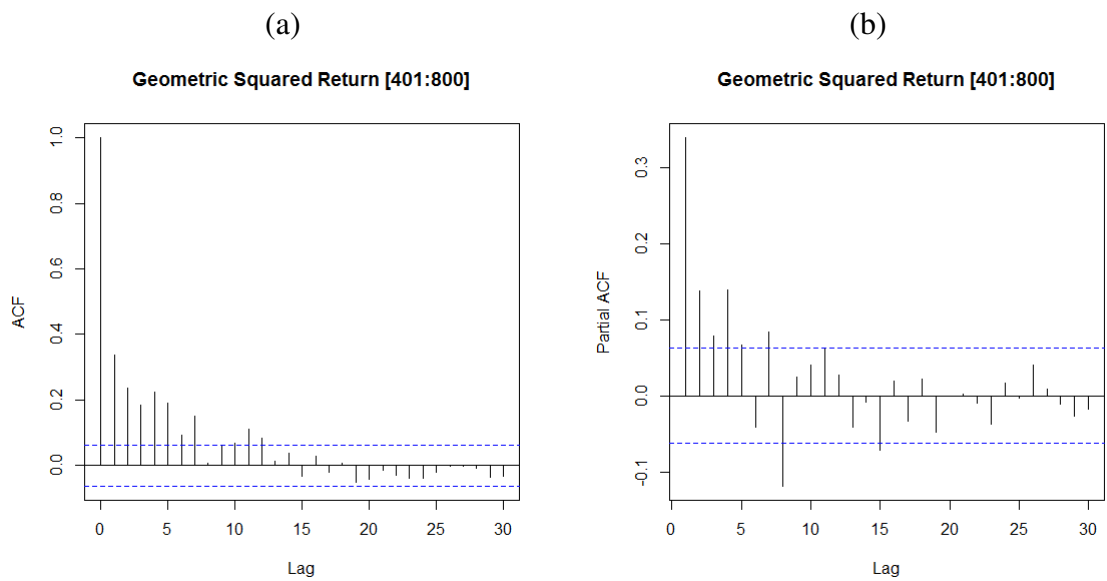
Figure 1.2: ACF and PACF of Geometric Squared Return for Series $t \in [401 : 800]$

(a)	(b)
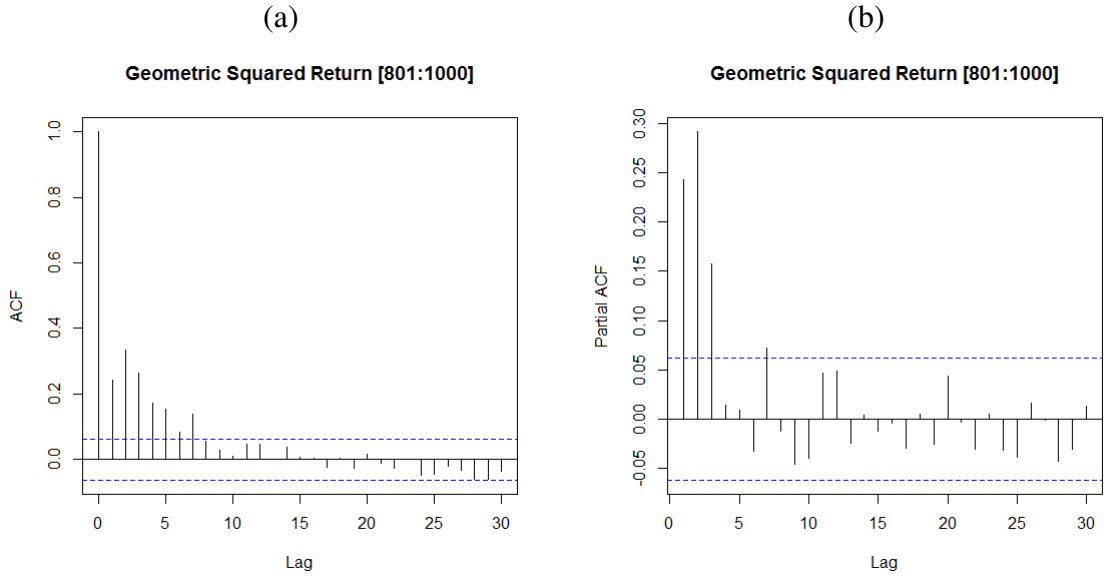
Figure 1.3: ACF and PACF of Geometric Squared Return for Series $t \in [801 : 1000]$

In particular, we can assign different orders (p, q) through the assessment of plots at the by examing where the functions tail off. From Figure 1.1 (a) and (b) it can be seen that the model orders are $p = 1$ and $q = 2$ respectively, which implies that a $GARCH(1, 2)$ is the most suitable model for the geometric return series from time $t = 1$ to time $t = 400$. In Figure 1.2 (a) and (b) it can be seen that the model orders are $p = 1$ and $q = 1$ respectively, which implies that a $GARCH(1, 1)$ is the most suitable model for the geometric return series from time $t = 401$ to time $t = 800$. Similarly, an assessment of Figure 1.3 (a) and (b) it can be seen that the model orders are $p = 2$ and $q = 2$ respectively, which implies that a $GARCH(2, 2)$ is the most suitable model for the geometric return series from time $t = 801$ to time $t = 1000$. This research therefore proposes the use the auto-correlation function as a detector for change-points in the series$(X_t)_{t \in \mathbb{N}}$. The location of change-points may be used to extract further information about the underlying properties of the data and also be used to explain why the break occurred.

1.3 Justification of the Study

Given the changing pace of the underlying economic mechanism and technological progress, modeling economic processes over a long time horizon, it is possible that structural changes may occur. This can cause the time series to deviate from stationarity and result to volatility clustering. The detection of these structural change points is therefore vital to various players in a given economy to ensure timeliness of decisions. A fundamental problem in financial trading is the correct and timely identification of turning points in stock value or exchange-rate series. This detection enables one to make profitable investment decisions, such as buying-at-low and selling-at-high, hence traders require early identification of local troughs and peaks of stock values. In macroeconomics, knowing the beginning of a recession leads to an increase of government expenditure or an expansion of money supply.

1.4 Research Objectives

1.4.1 Main Objective

The general objective of this research is to apply the Manhattan Distance in model order change-point estimation in GARCH models.

1.4.2 Specific Objective

The Specific Objectives are:

1. To develop an estimator for change-point.

2. To determine the consistency of the change-point estimator.

3. To determine the asymptotic distribution of the change-point process.

4. To apply the change-point estimator in pricing American options.

1.5   Scope of the Study

The academic scope of the study is limited to GARCH model order change-point estimation, particularly, assuming departure from $GARCH\left(1,1\right)$ to a general $GARCH\left(p,q\right)$ model with different model order specification.

Chapter 2

LITERATURE REVIEW

This chapter begins with reviewing various change-point estimation methods. This is followed by methods used in examining the consistency of these change-point estimation methods and the corresponding distribution of the change-point estimator. The final section of the chapter describes option pricing models.

2.1   Estimation of Change-point

A change-point is the point in time at there is a change in the underlying distribution of the data. Market and institutional changes are assumed to cause change-points in financial time series. This means that the assumption of homogenous distribution throughout the whole financial time series is not correct. To be able to use all data available for analysis, the change-points need to be located. The following methods outline the theory surrounding change-point estimation with the aims to discover if and where such changes are present in the data.

Zhang (2016) and Bin and Yongmiao (2012) propose a consistent test for smooth structural changes as well as abrupt structural breaks with known or unknown change-points. They estimate smooth time-varying parameters by local smoothing and compare the fitted values of the restricted constant parameter model and the unrestricted

time-varying parameter model. They utilize a smooth transition linear regression function. The transition function allows both the intercept and the slope to change smoothly over time. One advantage of this nonparametric model is that little restriction is imposed on the functional forms of the time-varying intercept and slope, except for the condition that they evolve over time smoothly. The proposed test was also established to have appealing features. First, it is consistent against a large class of smooth time-varying parameter alternatives as well as multiple sudden structural breaks with unknown change-points. Second, no prior information on a structural change alternative is needed. that is, one needs not to know whether the structural changes are smooth or abrupt, and in the cases of abrupt structural breaks nor the dates or the number of breaks. Third, different from many tests for structural breaks in the literature, our test is asymptotically pivotal. The only inputs required are the ordinary least squares and local linear time-varying parameter estimators. The latter is a locally weighted least squares estimator, therefore any standard econometric software can be used to implement the test. Fourth, because only local information is employed in estimating parameters at each time point, the test has symmetric power against structural breaks that occur either in the first or second half of the sample period. The nonparametric methods have also been utilized in change-point estimation in other random variables. However, the validity of the test relies critically on the assumption that the trend function is smooth, namely it does not contain any jump. Although the smoothness assumption provides the basic motivation and technical justification for using nonparametric smoothing methods, it is important to detect whether the trend function has any jump before calling the nonparametric estimation procedure.

Non-parametric estimation is also proposed by Kanamori et al., (2010) who use direct density-ratio estimation for change-point identication. Here, inference is made on the ratio between the densities of the data before and after a change-point. By considering a density ratio, these approaches do not require knowledge about the densities themselves. Gichuhi (2008) examined nonparametric changepoint analysis for Bernoulli

random variables. They used neural network based likelihood ratio test statistic to detect a change-point in a given set of data and derived the limit distribution of the estimator under the assumption that the model is properly specified. The maximum likelihood change-point estimation method was applied to determine the estimator.

A similar approach with respect to time varying parameters was considered under the time varying ARCH (TV-ARCH) model of Dahlhaus and Subba Rao (2006) which was later generalized by Rohan and Ramanathan (2012) to a time varying GARCH (TV-GARCH) (1,1) model, by allowing the parameters of a stationary GARCH model of Bollerslev (1986) to change slowly with time. In their approach, a two-step local polynomial estimator of the parameter functions of the TV-GARCH (p,q) model was initially used. The asymptotic distributions was also investigated where it was established that the asymptotic distribution depends on the parameters of a stationary GARCH process, which is unobservable thus limiting the scope of asymptotic results. The stationary GARCH process is such that it locally approximates the TV-GARCH process at specific time points. It was proved that the distribution of the proposed bootstrap estimator of parameter functions of the TV-GARCH (p,q) model asymptotically coincides with that of the actual local polynomial estimator. Further investigation of the validity of the bootstrapped estimator using a simulation study revealed that the bootstrapped estimator provides a better approximation to normality than the actual local polynomial estimator. However, test based on the least squares and the local polynomial estimators of the parameter functions are not guaranteed to be non-negative. These method results in some of the bootstrapped residual squares to be negative.

Some scholars also argue that volatility clustering can be as a result of investor's inaction. Liu (2000) argues that the notion that a Markovian regime switching mechanism in volatility can lead to volatility clustering is not sufficient to generate long-range dependence in absolute values of returns. According to this research, the time spent in each regime, the duration of regimes, is more important than switching and should have a heavy-tailed distribution. In light of this Luc et al. (2011) proposed the use of

Markov-switching GARCH for modeling GARCH models in the presence of change-points. The motivation for this is bore from the fact that the marginal likelihood os a non-stationary GARCH cannot be established due to the path dependence problem. The path dependence problem occurs because the conditional variance at time t depends on the entire sequence of regimes visited up to time t, due to the recursive nature of the GARCH process. Since the regimes are unobservable, one needs to integrate over all possible regime paths when computing the likelihood function. However, the number of possible paths grows exponentially with t, rendering the likelihood evaluation unfeasible. Thus applying Markov chain Monte Carlo methods as the approach is particularly suitable for conducting inference in non-linear state space models.

A Bayesian approach is considered by Tze Leung and Haipeng (2013). In this approach two time-scales are considered where the "short" time-scale is used to define GARCH dynamics and the "long" time-scale to incorporate parameter jumps. This leads to a Bayesian changepoint ARX-GARCH model, whose unknown parameters can undergo occasional changes at unspecified times and can be estimated by explicit recursive formulas when the hyperparameters of the Bayesian model are specified. Efficient estimators of the hyperparameters of the Bayesian model are developed, yielding empirical Bayes estimates of the piecewise constant parameters in the stochastic changepoint model. The empirical Bayes approach is applied to the frequentist problem of partitioning the time series into segments under sparsity assumptions on the change-points.

Dynamic programming for optimal multiple change-point detection was adopted by R. et al. (2012) and Frick and Sieling (2014) under the assumption of unknown number of change-points. A dynamic programming algorithm retrieves the $k-1$ change-points minimizing the quadratic loss where n is the number of observations. However, at most $O(n2)$ computational complexity of various dynamic programming methods can be achieved only if the objective function can be updated in $O(1)$ time when a new observation is included. This may not be true in general unless the objective function

is chosen as the sum of squares or likelihood function for the Gaussian random variables in change-in-mean and changein-variance cases. In general, all terms in the log-likelihood have to be re-estimated based on the new estimated parameters. Therefore, the updating time may depend on the running sample size since the last change-point. As such, the quadratic complexity in *n* restricts the use of such an algorithm to small or intermediate values of *n*.

CUSUM test for change-points was originally developed for independent processes for estimation a change-point in the mean Page (1954) or the variance Inclan and Tiao (1994). The procedures were extended to β-mixing processes Kokoszka and Leipus (2000). The residual CUSUM test for estimating change-points in the conditional mean examines the point in time when there is a maximum deviation of the the observation from the mean. The squared residual CUSUM test is used for estimating breaks in the conditional variance by examining the point in time when there is a maximum deviation of the the observation from the variance. The K and L CUSUM test as proposed by Kokoszka and Leipus (2000) is utilized to detect for changes in $ARCH(\infty)$ models under the null hypothesis is that there is no change-point and the alternative is that there is one change-point at an unknown time. The model rests on the assumption that the error terms are normally independently distributed. They considered the change-point process to be drawn from a weighted change in the second moment with the change-point being estimated as the point at which there is maximal sample evidence for a change in the squared series. The duo Kokoszka and Leipus (2000) also proved that in the presence of a single break k is a consistent estimator of the unknown change-point where under the null hypothesis which states that there is no break the change-point processes asymptotic distribution was a Brownian bridge. Advantage of the K and L test is its validity under a wide class of processes including long memory, GARCH-type and non-linear time series models. It is also adapted for multiple break points, with the number of breaks determined using the sequential sample segmentation approach. The K and L test has good power only for large and non-monotone

(rather than small and gradual) changes in the GARCH parameters for any of the absolute rather than the squared returns transformations. Similarly it shares good power for detecting changes in the variance of the error term in the GARCH process. As the sample size increases the performance of the test improves even for small change points. However, the CUSUM test have different powers against change-points that have the same magnitudes but occur at different time points.

In order to extend the single change-point estimation method to multiple change-point estimation several methods have been proposed. First is the hierarchical divisive estimation which sequentially identifies change-points via a bisection algorithm. This approach estimates multiple change-points by iteratively applying a procedure for locating a single change-point. A new change point location is estimated at each iteration so that it divides an existing segment. This progresses as a binary tree where the root node corresponds to the case of no change-points, and thus contains the entire time series. The other non-root nodes are either a copy of their parent, or correspond to one of the new segments created by the addition of a change-point to their parent. This is often referred to binary segmentation as the method extends any single change-point method to multiple change-points by iteratively repeating the method on different subsets of the sequence. According to (Matteson and James, 2013) utilize the hierarchical divisive estimation method and shows that this procedure generates strongly consistent change-point estimates. This method requires that an initial segmentation of the data be provided. This initialization helps reduce the computational time of the procedure as well as allowing for the inclusion of a priori knowledge of possible change-point locations. However, in the absence of such assumptions, then each observation is assigned to its own segment.

Secondly is the agglomerative algorithm estimates change-point locations by determining an optimal segmentation. This method requires that an initial segmentation of the data be provided. This initial segmentation can help to reduce the computational time of the procedure. It also allows for the inclusion of a priori knowledge of possible

14

change point locations, however if no such assumptions are made, then each observation can be assigned to its own segment. Neighbouring segments are then sequentially merged to maximize a goodness-of-fit statistic. The estimated change-point locations are determined by the iteration which maximized the penalized goodness-of-fit statistic. When using this procedure it is assumed that there is at least one change-point present within the time series. Within these segmentation approaches, the choices of techniques we can apply are diverse. All these techniques aim at the identification of the prominent points which are used for decisions in the segmentation approaches. The Perceptually Important Points (PIP) method as introduced for time series by (Chung et al., 2001) and later used by (Jiang et al., 2007) and (Fu et al., 2006). Others like (Oliver and Forbes, 1997) pursue a change-point detection approach, (Bao and Yang, 2008) propose turning points sequences applied to financial trading strategies, (Guralnik and Srivastava, 1999) present a special event detection, (Oliver et al., 1998) and (Fitzgibbon et al., 2002) use minimum message length approaches, and (Fancourt and Principe, 1998) tailor PCA to locally stationary time series. (Duncan and Bryant, 1996) suggest the use dynamic programming for time series segmentation which is furthered by (Himberg et al., 2001) through approximating them with Global Iterative Replacement (GIR) algorithm results and they illustrate their segmentation technique with mobile phone applications in context recognition. A piecewise generalized likelihood ratio is used by (Wang and Willett, 2004) who apply it to establish first segmentation and then elaborate the segments further. Recently, (Cho and Fryzlewicz, 2012) segment a piecewise stationary time series with unknown number of breakpoints using a nonparametric locally stationary wavelet model. A combination of recursive and dynamic Principal Component Analysis (PCA) for multivariate time series segmentation is adopted by (Dobos and Abonyi, 2012).

Thirdly are the exact segmentation methods. The Segment Neighbourhood and optimal partitioning method are examples of the exact methods. These methods are based on minimizing the cost function which can be considered as twice the negative log like-

lihood (Chen and Gupta, 2000), quadratic loss and cumulative sums (Rigaill, 2010). The segment neighbourhood begins by setting an upper limit on the size of the segmentation space, the maximum number of changepoint, that is required. This approach searches the entire segmentation space using dynamic programming and continues by computing the cost function for all possible segments. Following (Jackson et al., 2005) the optimal partitioning (OP) method begins by first conditioning on the last point of change. It then relates the optimal value of the cost function to the cost for the optimal partition of the data prior to the last changepoint plus the cost for the segment from the last changepoint to the end of the data. According to (Killick and Eckley, 2012) pruning can be used to increase the computational efficiency of the optimal partitioning method whilst still ensuring that the method finds a global minimum of the cost function. The essence of pruning in this context is to remove those values of k which can never be minima from the minimization performed at each iteration.

2.2 Consistency of the change-point estimation

General theory on examining consistency of estimators is reviewed following (Jacod and Sorensen, 2016). The theory covers consistent estimators obtained by maximising (or minimising), with respect to θ, a function $H_n(\theta; X_1, \cdots, X_n)$ of the data and the parameter. The general set-up is as follows. Consider a measurable space (Ω, F) with a probability measure P, the true measure, and a family $(P^{(\theta)})_{\theta \in \Theta}$ of probability measures indexed by $\Theta \subseteq R^p$, the statistical model. Often it is assumed that there exists a $\theta_0 \in \Theta$, the true parameter value, such that $P = P^{(\theta 0)}$, but this need not be the case. At stage n, we have a set of observations which generates a *P–complete* $\sigma-field$ $\mathscr{F}_n$, that is, $\mathscr{F}_n$ is the *P–complete* of the $\sigma-field$ $\mathscr{F}_n^0$ n generated by those observations. Taking the completion simplifies the mathematical formulation of the results below, and it has no negative consequences from a practical viewpoint since all statements about estimators are always up to a *P − null* set. To be consistent, it is also

suppose that $\mathscr{F}$ is *P–complete*. If a continuous time stochastic process X_t is observed, then $\mathscr{F}_n$ could, for instance, be the complete $\sigma - field$ generated by the variables X_s for $s \in [0, t_n]$ for some increasing sequence t_n, or by the variables $X_{i\Delta n}$ for $i = 0, \cdots, n$ and some $\Delta n > 0$. An estimating function at stage n is a function $(\theta, \omega) \to G_n(\theta, \omega)$ that takes its values in R^p and depends on the statistical parameter $\theta \in \Theta$ and on the observation at this stage, that is $(\theta, \omega) \to G_n(\theta, \omega)$ is measurable with respect to the product of the Borel $\sigma - field$ of Θ with $\mathscr{F}_n$. An estimator is obtained by solving the estimating equation $G_n(\theta) = 0$. For n large enough the estimating equation has a solution that converges to a particular parameter value $\hat{\theta}$. When the statistical model contains the true model, is consistent, the estimating function should preferably be chosen such that $\hat{\theta} = \theta_0$. This approach has been utilized in examining the change-point estimates by Chen and Hong (2012) and Kokoszka and Leipus (2000) reviewed in Section 2.1.

2.3   Distribution of the change-point estimator

The asymptotic distribution of change-point estimators rely on the data generating mechanisms. In particular, we are interested in data generated by the GARCH model. As such the point process theory is reviewed as an appropriate method of examining the limiting structure of partial sums of stationary random variables (Davis and Hsing (1995)). Consider $\{X_t\}$ to be a strictly stationary sequence of random variables with regularly probabilities. Point process methods are considered suitable for determining the weak convergence of the partial sums, $S_n = X_1 + \cdots X_n$, suitably normalized, when $\{X_t\}$ satisfies the mixing condition. This is achieved through characterization of the limit point processes for the sequence of point processes N_n with mass at the points $\{X_t / a_n\}$ for $t = 1, \cdots, n$ and a_n being the $1 - n^{-1}$ quantile of the distribution of $|X_t|$. S_n is asymptotically stable if N_n converges to N for the tail index $\kappa \in (0, 2)$.

2.4 Option Pricing

Consider a financial market in which the total supply of an asset is N. At the beginning of each period t_1, all participants observe the unit asset price $p(t)$. Further assume that all the traders are perfectly competitive; they take the price as given, imagining that they are so small that they cannot affect the price, no matter how much they demand. In particular, we consider the behavior of two types of traders; noise traders and value assets. The noise traders' demand is defined in terms of the cash value they spend on the asset which follows an autoregressive random process. When there are only noise traders the price is set such that demand, D, is equal to the total supply of assets, $D = N$. On the other hand, value investors base their demand D on a mispricing signal on an asset. The presence of the value investors dramatically alters the statistical properties of price returns. Thurner et al. (2012) established that with only noise traders the log returns are nearly normally distributed. When value investors are added without leverage the volatility of prices drops slightly, but the log returns remain approximately normally distributed. However, when leverage is increased, the distribution becomes much more concentrated in the center and the negative returns develop fat tails. In general, when value investors are unleveraged, they will always buy into a falling market since when prices drop there are guaranteed to be buyers thus damping price movements away from the fundamental value. When value investors are leveraged, they sell into a falling market thus amplifying the deviations of price movements away from the fundamental value.

Assume that we begin at a point where the wealth of all value investors is small. In the early stages, all these investors will tend to accumulate wealth, with aggressive funds growing faster than cautious funds. The overall increase in the wealth of funds lowers volatility. As a result, the increase in the wealth of the most aggressive value investors drives up the overall use of leverage. In due course, a substantial downward fluctuation in noise trader demand happens to occur at the same time that one or more aggressive

value investors are fully leveraged. This prompts a large sell-off by the aggressive value investors, which drives prices down, and generates a crash. After the crash the overall wealth of funds is substantially diminished, and as a result volatility goes back up. This leverage cycle results into volatility clustering where volatility patterns change following a leverage point. It is thus important to incorporate this behavior in financial modelling. In particular this study utilizes the change-point estimator to locate the point of change.

More specifically we utilize this in pricing options in a derivative market. There are two basic types of options. A call option (put option) or more simply a call (put) is a derivative giving to the holder the right, but not the obligation, to buy (sell) an agreed quantity of the underlying asset S, from the seller of the option on (or before) the expiration date T , for a specified price K, the strike price or exercise price. The value of an option is a sum of two parts, its intrinsic value and extrinsic value. The intrinsic value is the fundamental value of a financial asset, in this case the payoff of the option. The payoff is the difference between the exercise price and the current stock price. For a call option, the buyer receives the amount $max(S_T - K, 0) = (S_T - K)^+$ since the option will not be exercised unless it is 'in the money'. Similarly, for a put, the payoff at time T is $(K - S_T)^+$. Generally the value of a call option increases as the price of the underlying asset increases and the value of a put option will increase when the value of the underlying asset decreases. Extrinsic value is commonly known as time value. Time value can be measured by subtracting intrinsic value from the price of the option yielding the premium that the investor pays for the current exercise value. Time value will decrease over time as maturity nears since the lesser the time to maturity the lower the probability of a better exercise value. Volatility is also very important in option pricing because after all the changes in price are the reason why options exist. The more volatile the price of the underlying asset, the better the odds are, that the underlying price will end up further away from the current price by the time of expiration. Therefore rise in volatility of the underlying asset's movements

drives up both call and put option prices Hull (2011).

In the financial context, volatility is used to express the uncertainty of the financial returns. If volatility is high the probability that a stock will perform exceptionally well or exceptionally poor is high. Therefore the probability that an option is in-the-money at the time of expiration is higher when the volatility is high. As it can be seen there is a distinct connection between volatility of the underlying asset and the returns of the option. This also leads to the fact that the volatility of the underlying asset has a major impact on the option price. Historical or realized volatility can be calculated from historical data but future volatility is more difficult to estimate. Calculation using the realized volatility is the simplest way but the problem is how long the sample should be. Longer sample might make the estimate more accurate but since volatility is not constant over time the past returns might not be a good estimate of the future volatility. There are numerous ways to estimate volatility statistically in addition to realized volatility. Moving average models smooth the random price fluctuation and form a lagged trend for volatility. Two well-known moving averages are SMA (Simple moving average) and EWMA (exponentially weighted moving average) which gives more weight to the latest observations. Autoregressive models have become popular in academic research due to their good future volatility forecasting ability. Some common autoregressive models are ARMA (autoregressive moving average), ARCH (autoregressevie conditional heteroscedasticity) and GARCH (generalized autoregressive conditional heteroscedasticity). The motivation behind ARCH and GARCH models is that, there is a relationship between todays squared returns and past squared returns. ARCH and GARCH models utilize this relationship to estimate volatility. This way they can forecast future volatility very accurately by using the past returns Abdalla and Winker (2012).

A research on financial speculation in energy and agriculture futures markets was carried out by Matteo et al. (2012). A GARCH model was used to estimate energy and agricultural commodities' returns where returns were explained by macroeconomic

20

variables and a measure of speculation. MGARCH models were also used to investigate the presence of spillovers across commodities. It was established that speculation in one market does not seem to significantly affect returns in other markets. An option pricing model in the context of GARCH asset return process was developed by utilizing the locally risk-neutral valuation relationship Duan. (1995). This was furthered by Heston and Nandi (2000) who developed a closed-form option pricing formula for an asset whose variance follows a GARCH process and allows for correlation between returns of the asset and variance as well as admits multiple lags in the dynamics of the GARCH process. These models have been utilized by others researchers over time with Hsieh and Ritchkeny (2005) particularly comparing the performance of these two models. It was established that the Heston and Nandi Model model can explain a significant portion of the volatility smile whereas the NGARCH model is superior in removing biases from pricing residuals for all moneyness and maturity categories.

The Black-Scholes model is the most widely used model in pricing options. It was originally designed to value options that can be exercised only at maturity and on underlying assets that do not pay dividends. In practice, assets do pay dividends, options sometimes get exercised early and exercising an option can affect the value of the underlying asset. Adjustments to provide partial corrections to the Black- Scholes model exist. This research primarily focusses on the American options which can be exercised any time until expiration. To be able to use the Black-Scholes model it is thus necessary to adjust the value of the option for the possibility of early exercise Damodaran (2012). There are two approaches for doing so. One uses the Black-Scholes to value the option to each potential exercise date. With options on stocks, this basically requires that we value options to each ex-dividend day and choose the maximum of the estimated call values. The second approach is to use a modified version of the binomial model to consider the possibility of early exercise. In this version, the up and down movements for asset prices in each period can be estimated from the variance and the length of each period. In particular, we propose the use of the change-point estimator

to locate the change-points of the asset prices which are assumed to be the early excercise dates. Following the identified change-points, suitable GARCH(p,q) models are fitted to model the period specific variance. Black-Scholes model is then applied.

2.5 Summary and Critique of Existing Literature

Reviewed literature reveal that the use of one model may not be appropriate to model a non-stationary series and as such various change-point estimation methods have proposed. These methods include the use of nonparametric estimation, transition function, markov switching, dynamic programming and CUSUM test. These methods have their pros as discussed in section 2.1. However, they are limited in different ways and their suitability depend on the underlying assumptions. Some methods are based on the use of local smoothed time-varying parameters where a smooth transition regression function is utilized. Where the transition function is linear, both the intercept and the slope to change smoothly over time. However, this does take care of abrupt changes. In contrast the transition function is considered by some to be a polynomial. In such cases the local polynomial estimators of the parameter functions are not guaranteed to be non-negative. When considering the Markovian switching estimation, the regimes are unobservable thus one needs to integrate over all possible regime paths when computing the likelihood function. However, the number of possible paths grows exponentially with t, rendering the likelihood evaluation unfeasible. Bayesian approaches to change-point detection are also prone to the shortconing that require the specication of priors on the number and position of change points and the parameter vectors θ_i between any two change points. The latter can be specied to allow for a specic type of change, that is, in only one element of θ_i. The CUSUM test are focus on univariate processes with changes in the mean only or variance only. The assumption of changes occurring in the mean only can yield severe distortions, even in from a detection point of view. In a similar manner as with the change in mean-only case, if the assumption of changes

occurring only in the variance is inappropriate the consequences can be substantial. In addition, the CUSUM tests have different powers against change-points that have the same magnitudes but occur at different time points. In conclusion, it is arguably that most change-point detection tests focus on changes in parameter values while others are not particular on the data generating mechanisms. This research particularly seeks to estimate change-points assuming that GARCH model is data generating mechanism and that change occurs particularly in the model order specification.

Chapter 3

METHODOLOGY

This chapter discusses concepts required for the achievement of results.

3.1 GARCH Model

3.1.1 Model Definition

In this section we start with a definition of GARCH processes based on the first two conditional moments, the $ARCH(\infty)$, the $ARMA(max(p,q),q)$ and the Stochastic Differential Equations (SDE). The various model definitions are utilized in the proof of results in the methodology.

Definition (GARCH(p,q) Process) A process $(X_t)_{t\in\mathbb{N}}$ is called a $GARCH(p,q)$ process if its first two conditional moments exist and satisfy:

  (i) $E(X_t \mid X_u, u < t) = 0$ for $t \in \mathbb{Z}$

  (ii) There exist constants $\alpha_0 > 0, \alpha_i \geq 0, i = 1, \cdots, p \, and \, \beta_j \geq 0, j = 1, \cdots, q$ such that

$$
\begin{aligned}
\sigma_t^2 &= Var\left(X_t \mid X_u, u < t\right) \\
&= \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2
\end{aligned}
$$

(iii) Let ε_t be an independent and identically distributed sequence with zero mean and unit variance $\varepsilon_t \sim iid\,(0,1)$, then

$$
\begin{aligned}
X_t &= \sigma_t \varepsilon_t \qquad for\, t \in \mathbb{Z} \\
\sigma_t^2 &= \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2
\end{aligned}
\tag{3.1}
$$

Definition (GARCH(p,q) as ARMA(max(p,q),q)) Bollerslev (1986) showed that the GARCH(p,q) model (3.1) can also be represented as an $ARMA(max(p,q),q)$ written

$$
X_t^2 - \sum_{i=1}^{p} \alpha_i X_{t-i}^2 - \sum_{j=1}^{q} \beta_j X_{t-j}^2 = \alpha_0 + u_t - \sum_{j=1}^{q} \beta_j u_{t-j} \qquad for\, t \in \mathbb{Z}
\tag{3.2}
$$

where $u_t = X_t^2 - \sigma_t^2$ *and* $(u_t)_{t \in \mathbb{Z}}$ *is white noise*

Definition (GARCH(p,q) as SDE) By iterating the defining difference equation (3.1) for σ_t^2 the GARCH model can be further expressed as a stochastic differential equation as follows:

$$
\text{Let } Y_t = \begin{pmatrix} \sigma^2_{t+1} \\ \vdots \\ \sigma^2_{t-q+2} \\ X^2_t \\ \\ X^2_{t-p+1} \end{pmatrix}, A_t = \begin{pmatrix} \alpha_1\varepsilon^2_t+\beta_1 & \beta_2 & \cdots & \beta_{q-1} & \beta_q & \alpha_2 & \alpha_3 & \cdots & \alpha_p \\ 1 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 & 0 & \cdots & 0 \\ \varepsilon^2_t & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix},
$$

$B_t = \begin{pmatrix} \alpha_0 & 0 & \cdots & 0 \end{pmatrix}'$ then (Y_t) satisfies the following stochastic differential equation

$$Y_t = A_t Y_{t-1} + B_t \quad for \quad t \in \mathbb{N} \tag{3.3}$$

Theorem 1. *(Strictly Stationary)*

A necessary and sufficient condition for the existence of a strictly stationary solution to the GARCH(p,q) model (3.1) is the negativity of the top Lyapunov exponent γ given by

$$\gamma = \genfrac{}{}{0pt}{}{inf}{n \geq 1} \; n^{-1} E \, log \, ||A_n \cdots A_1|| < 0 \tag{3.4}$$

where γ is the top Lyapunov exponent of the sequence $A_t, t \in \mathbb{Z}$ defined by (3.3).

Proof. The proof of this theorem can be found as Theorem 2.4 in Francq and Zakoian (2011).

$\square$

When the strictly stationary solution exists, it is unique, nonanticipative and ergodic. According to (Bougerol and Picard, 1992) γ cannot be calculated explicitly but a sufficient condition for $\gamma < 0$ is given by

$$\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1 \tag{3.5}$$

Following the works of (Bougerol and Picard, 1992), when $p = q = 1$ in models (3.1) the Lyapunov exponent is obtained as

$$
\begin{aligned}
A_t &= \left(\varepsilon_t^2, 1\right)(\alpha_1, \beta_1) \\
A_t A_{t-1} \cdots A_1 &= \prod_{i=1}^{t-1}\left(\alpha_1 \varepsilon_{t-i}^2 + \beta_1\right) A_t \\
log \|A_t A_{t-1} \cdots A_1\| &= \sum_{i=1}^{t-1} log\left(\alpha_1 \varepsilon_{t-i}^2 + \beta_1\right) + log \|A_t\| \\
\gamma &= E\, log\left(\alpha_1 \varepsilon_{t-i}^2 + \beta_1\right) \tag{3.6}
\end{aligned}
$$

Thus the necessary and sufficient condition for the strict stationarity in this case is $\gamma = E\, log\left(\alpha_1 \varepsilon_t^2 + \beta_1\right) < 0$

Theorem 2. *(Second-Order Stationarity)*

If there exists a GARCH(p,q) process, in the sense of Definition (3.1), which is second-order stationary and nonanticipative, and if $\alpha_0 > 0$, then

$$\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1 \tag{3.7}$$

Conversely, if (3.7) holds, the unique strictly stationary solution of model (3.1) is a weak white noise and thus is second-order stationary.

Proof. The proof of this theorem can be found as Theorem 2.5 in Francq and Zakoian (2011).

$\square$

Since the conditions in Theorems 1 and 2 are necessary and sufficient, we necessarily have

$$\left[\sum_{i=1}^{p}\alpha_i + \sum_{j=1}^{q}\beta_j < 1\right] \Rightarrow \gamma < 0 \tag{3.8}$$

This is in line with (Bougerol and Picard, 1992), who indicate that the sufficient condition for $\gamma < 0$ is given by (3.7) since the Lyapunov exponent γ cannot be calculated explicitly.

3.1.2   Model Order Identification

Order determination is based on inspection of the sample autocorrelations and partial autocorrelations functions. To identify the orders of a $GARCH(p,q)$ model (3.1), the series $(X_t^2)_{t\in\mathbb{Z}}$ is first represented as an $ARMA(p,q)$ as given in definition 3.2. Beguin et al. (1980)propose the Corner Method as a suitable method for studying the sample autocorrelation function and sample partial autocorrelation function patterns. Let $D(i,j)$ denote a $j \times j$ matrix whose determinant is denoted by $\Delta(i,j)$ where

$$D(i,j) = \begin{pmatrix} \rho_i & \rho_{i-1} & \cdots & \rho_{i-j+1} \\ \rho_{i+1} & \rho_i & \cdots & \rho_{i-j} \\ \vdots & \vdots & \vdots & \vdots \\ \rho_{i+j-1} & \cdots & \rho_{i+1} & \rho_i \end{pmatrix} \tag{3.9}$$

Since $\rho_h = 0$ for all $h > Q$, then p and q are minimal orders if and only if

$$\begin{aligned} \Delta(i,j) = 0 & \quad \forall\, i > q \quad \forall\, j > p \\ \Delta(i,p) \neq 0 & \quad \forall\, i \geq q \\ \Delta(q,j) \neq 0 & \quad \forall\, j \geq p \end{aligned} \tag{3.10}$$

The minimal orders p and q are thus characterized by the following table with $\times$ denoting non-zero entries:

AR Order i	MA Order j							
	1	2	$\cdots$	q	q+1	$\cdots$	$\cdots$	$\cdots$	$\cdots$
1	ρ_1	ρ_2		ρ_q	ρ_{q+1}	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$\vdots$									
p				$\times$	$\times$	$\times$	$\times$	$\times$	$\times$
p+1				$\times$	0	0	0	0	0
$\vdots$				$\times$	0	0	0	0	0
$\vdots$				$\times$	0	0	0	0	0
$\vdots$				$\times$	0	0	0	0	0

Table 3.1: Corner Method Table

The orders p and q are thus characterized by a corner of zeros in Table 3.1 with the entries obtained using the recursion on j given by

$$\Delta(i,j)^2 = \Delta(i+1,j)\Delta(i-1,j)+\Delta(i,j+1)\Delta(i,j-1) \qquad (3.11)$$

$$with \quad \Delta(i,0)=1 \; \Delta(i,1)=\rho_i$$

In practice however, we only have a finite number of onservations, say n. This results into a finite number of sample autocorrelations $\hat{\rho}_1,\cdots,\hat{\rho}_n$ which allows for $\hat{\Delta}(j,i)$ to be computed for $i \geq 1$, $j \geq 1$ and $i+j \leq n+1$. Since $\hat{\Delta}(j,i)$ consistently estimates $\Delta(j,i)$, the orders p and q are characterized by a corner of small values in Table 3.1. It is thus prefable to consider Studentized statistics defined for $i = -n,\cdots,n$ and $j = 0,\cdots n-i+1$, by

$$t(i,j) = \sqrt{n}\frac{\hat{\Delta}(j,i)}{\hat{\sigma}_{\hat{\Delta}(j,i)}} \qquad (3.12)$$

$$\hat{\sigma}^2_{\hat{\Delta}(j,i)} = \frac{\partial \hat{\Delta}(j,i)}{\partial \rho'_n}\hat{\Sigma}_{\hat{\rho}_n}\frac{\partial \hat{\Delta}(j,i)}{\partial \rho'_n} \qquad (3.13)$$

where $\hat{\Sigma}_{\hat{\rho}_n}$ is a consistent estimator of the asymptotic covariance matrix of the first n

SACRs. When $\Delta(i,j) = 0$ the statistic $t(i,j)$ asymptotically follows a Standard Normal distribution. In contrast, however, when $\Delta(i,j) \neq 0$ the $\sqrt{n}|t(i,j)| \to \infty$ almost surely as $n \to \infty$. The hypothesis $\Delta(i,j) = 0$ is thus rejected at α if $|t(i,j)| > \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$. The detection of a corner of small values in the Table 3.1 if no entry in this corner is greater than $\left(1 - \frac{\alpha}{2}\right)$ quantile in absolute value.

3.1.3 Model Parameters Estimation

Assume that the observations $X_1, \cdots, X_n$ constitute a realization of length n of a $GARCH(p,q)$ process (3.1). Assume that the orders p and q are known. The vector of parameters $\theta = \left(\theta_1, \cdots, \theta_{p+q+1}\right)' = \left(\alpha_0, \alpha_1, \cdots, \alpha_q, \beta_1, \cdots, \beta_p\right)'$ belongs to a parameter space of the form $\Theta \subset (0, \infty) \times (0, \infty)^{p+q}$,

The true value of the parameter is unknown, and is denoted by $\theta_0 = \left(\alpha_{00}, \alpha_{01}, \cdots, \alpha_{0q}, \beta_{01}, \cdots, \beta_{0p}\right)'$.

To enable us write a likelihood function for this model we make the following assumption

Assumption 1. *(Independence)*

  (i) *$\varepsilon_t's$ are independent and under H_0 identically distributed with zero mean and a variance of one*

 (ii) *$X_t's$ are independent of the $\varepsilon_t's$ for $1 \leq t \leq n$*

Then it follows that :

$$L_n(\theta) = L_n(\theta; X_1, \cdots, X_n) = \prod_{t=1}^{n} \frac{1}{\sigma_t \sqrt{2\pi}} exp\left(-\frac{X_t^2}{2\sigma_t^2}\right) \qquad (3.14)$$

A Quasi Maximum Likelihood Estimator (QMLE) of θ is the measurable solution θ_n of $\tilde{\theta}_n = \underset{\theta \in \Theta}{arg\,max}\, L_n(\theta)$

The log-likelihood function of the parameter vector $\theta \in \Theta$ is given by

$$l_t(\theta) = \left(ln\, \sigma_t^2 - \frac{X_t^2}{\sigma_t^2} \right) \tag{3.15}$$

Maximizing the likelihood is equivalent to minimizing, with respect to θ

$$I_n(\theta) = n^{-1} \prod_{t=1}^{n} l_t \tag{3.16}$$

Thus the QMLE of θ is the measurable solution $\hat{\theta} = \underset{\theta \in \Theta}{argmin}\, I_n(\theta)$

Assumption 2.   *(i) $\theta_0 \in \Theta$ and Θ is compact.*

(ii) $\gamma(A_0) < 0$ and for all$\theta \in \Theta$, and Θ is compact $\sum_{j=1}^{q} \beta_j < 1$.

(iii) ε_t^2 has a nondegenerate distribution and $E\left(\varepsilon_t^2\right) = 1$

(iv) $\kappa_n = E\left(\varepsilon_t^4\right) < \infty$

Theorem 3. *Under assumptions 1 and 2, the QMLE is aymptotically normal*

$$\sqrt{n}\left(\hat{\theta}_n - \theta_0\right) \quad \rightarrow \quad N\left(0, (\kappa_n - 1)J^{-1}\right) \tag{3.17}$$

where J is a positive definite matrix specified as

$$J = E_{\theta_0}\left(\frac{\partial^2 l_t(\theta_0)}{\partial \theta \partial \theta'} \right) = E_{\theta_0}\left(\frac{1}{\sigma_t^4(\theta_0)} \frac{\partial^2 \sigma_t^2(\theta_0)}{\partial \theta} \frac{\partial^2 \sigma_t^2(\theta_0)}{\partial \theta'} \right) \tag{3.18}$$

Proof. The proof of this theorem can be found as Theorem 7.2 in Francq and Zakoian (2011).

$\square$

3.1.4   Testing for Second-Order Stationarity

Testing for second-order stationarity of the $GARCH(p,q)$ model involves testing

$$H_0 : \sum_{i=1}^{p} \alpha_{0i} + \sum_{j=1}^{q} \beta_{0j} < 1 \; against \; H_1 : \sum_{i=1}^{p} \alpha_{0i} + \sum_{j=1}^{q} \beta_{0j} \geq 1 \qquad (3.19)$$

Suppose that $c = (0,1,\cdots,1)^{'} \in \mathbb{R}^{p+q+1}$ then the hypothesis (3.19) to be tested can be represented as

$$H_0 : c^{'}\theta_0 < 1 \; against \; H_1 : c^{'}\theta_0 \geq 1 \qquad (3.20)$$

Following Theorem 3, we have that

$$\sqrt{n}\left(c^{'}\hat{\theta}_n - c^{'}\theta_0\right) \;\; \to \;\; N\left(0,(\kappa_n - 1)c^{'}J^{-1}c\right) \qquad (3.21)$$

An appropriate stationarity test for the hypothesis is the Wald Statistic T_n

$$T_n = \frac{\sqrt{n}\left(\alpha_{0i} + \Sigma_{j=1}^{q}\beta_{0j} - 1\right)}{\sqrt{\left((\kappa_n - 1)c^{'}J^{-1}c\right)}} \qquad (3.22)$$

which converges to a Standard Normal distribution when $c^{'}\theta_0 = 1$. Assuming an α level of significance, the corresponding rejection region is $T_n > \Phi^{-1}(1 - \alpha)$.

3.1.5 Testing for Significance of GARCH Coefficients

Assumption 3. $E(X)^6 < \infty$

Theorem 4. *Under Assumptions 1 and 3, the asymptotic distribution of $\sqrt{n}\left(\hat{\theta}_n - \theta_0\right)$ is λ^Λ where $\lambda^\Lambda = \arg\inf_{\lambda \in \Lambda}(\lambda - Z)^{'}J(\lambda - Z)^{'}$ with $Z \sim N\left(0,(\kappa_n - 1)J^{-1}\right)$.*

Proof. The proof of this theorem can be found as Theorem 8.1 in Francq and Zakoian (2011).

□

Assume that $\theta_0^{(1)} > 0$ and consider the hypothesis testing problem

$$H_0 : \theta_0^{(2)} = 0 \; against \; H_1 : \theta_0^{(2)} \neq 0 \tag{3.23}$$

Following Theorem 4 and under H_0 we have that

$$\sqrt{n}\hat{\theta}_n^{(2)} \xrightarrow{d} K\lambda^\Lambda, \; K = (0_{d_1 \times d_2}, I_{d_2}) \tag{3.24}$$

The Wald test is used to examine the hypothesis (3.23) where the Wald Statistic W_n is defined as

$$W_n = n\hat{\theta}_n^{(2)'} \left(K\hat{\Sigma}K'\right)^{-1} \hat{\theta}_n^{(2)} \tag{3.25}$$

where $\hat{\Sigma}$ is a consistent estimator of $\Sigma = (\kappa_n - 1)J^{-1}$.

At the asymptotic level of significance α, the rejection region is $W_n > \chi_{d_2}^2 (1 - \alpha)$, where $\chi_{d_2}^2 (1 - \alpha)$ is the $(1 - \alpha)$ quantile of the χ^2 distribution with d_2 degrees of freedom.

3.2 Change-Point Detection

The detection of structural changes in macroeconomic and financial time series is crucial for prudent statistical inference, forecasts, and sensible policy implications to be drawn from any underlying model. In particular, ignoring these structural changes may lead to spurious persistence in the conditional volatility dynamics. Here, methods of structural changes in volatility detection of GARCH models are considered.

Definition Let $\mathbf{X} = X_1, \cdots, X_n$ be a sequence of independently distributed random variables. A change-point as the point $k \in (1, n)$ at which the data generating mechanism of $\mathbf{X}$ changes.

In the presence of a single change-point k, $\mathbf{X}$ can be segmented into two parts with different $GARCH(p, q)$ model specifications. Various scholars have dedicated their research work to examining and creating methods of detecting these changes. Frequently,

change in model specification has been attributed to change in parameter specification. This has bore tests that examine change-points for abrupt structural breaks while others check for smooth structural changes in parameters. However, the structural changes could also be attributed to change in model order specification. Change-point detection entails testing the null hypothesis H_0 against the alternative hypothesis H_1. Particularly we consider that H_0 postulates that the model order specification does not change throughout the whole time period. H_1 postulates that the model order remains unchanged as in H_0 up to a certain unknown time point k when the model order changes. The objective of change-point detection is to create an estimator for the possible change-point k given set of random variables.

Assume that the set of random variables given by the data $\{X_t\}_{t=1}^{n}$ describes a financial returns time series modeled using $GARCH(p,q)$ process. A single change-point testing problem is first considered where it is assumed that a change-point can happen only at time k where $1 < k < n - 1$. The hypotheses to be investigated are assumed to follow the following definition:

$$H_0 : X_t \sim GARCH(1,1) \qquad for\, t = 1, \cdots, n$$

$$against$$

$$H_{1:}X_t \sim \begin{cases} GARCH(1,1) & for\, t = 1, \cdots, k \\ GARCH(p,q) & for\, t = k+1, \cdots, n \end{cases} \qquad where\, p,q \in \mathbb{N} \setminus \{0\} \quad (3.26)$$

Location of change-points can be investigated using a segmentation approach which aims at discretization to accurately approximate a time series. The general approach begins by segmenting a time series into subsequences and to then choosing a primitive shape patterns which represent the original time series best. There exist two segmentation procedures for a possible change-point namely hierarchical divisive and agglomerative algorithms.

3.3   Adjusted Rand Index

The performance of the proposed test is evaluated by measuring the degree of similarity between the correct classification and the resultant segmented clusters following identified change-point positions. This is achieved by the use of the Adjusted Rand Index (ARI) lies between 0 and 1. When the two partitions agree perfectly, the ARI is 1. Let $\mathbf{X}$ be a set representing a time series with finite cardinality $|\mathbf{X}| = n$. A clustering $\mathbf{C}$ is a set $\{C_1, C_2, \cdots, C_q\}$ of non-empty disjoint subsets of $\mathbf{X}$ such that their union equals $\mathbf{X}$. Let $\mathscr{P}(\mathbf{X})$ denote the set of all clusterings of $\mathbf{X}$. For a clustering $\mathbf{C}$ we assume $|C_i| > 0$ for all $i = 1, \cdots, l$. Let $\mathbf{C}' = \left\{C'_1, C'_2, \cdots, C'_r\right\} \in \mathscr{P}(\mathbf{X})$ denote the second clustering of $\mathbf{X}$. $\mathbf{C}'$ is a refinement of $\mathbf{C}$ implying that $\mathbf{C}$ is coarser than $\mathbf{C}'$ such that the following is true: $\forall C'_j \in \mathbf{C}' \exists C_i \in \mathbf{C}$ implying $C'_i \subseteq C_i$. Let $M = (m_{ij})$ be a $q \times r$ confusion matrix of the pair $\mathbf{C}, \mathbf{C}'$ with the $i.j - th$ element being equal to the number of elements in the intersection of the clusters C_i and C'_j such that:

$$m_{ij} = |C_i \cap C'_j|,\ 1 \leq i \leq q,\ 1 \leq j \leq r \tag{3.27}$$

To compare clusterings, we utilize the approach of counting pairs that are classified in the same way in both clusterings, that is, pairs of elements of X that are in the same cluster (in different clusters, respectively) under both clusterings. The set of all pairs of elements of X is the disjoint union of the following sets: S_{11}={pairs that are in the same clusters under $\mathbf{C}$ and $\mathbf{C}'$}, S_{00}={pairs that are in the different clusters under $\mathbf{C}$ and $\mathbf{C}'$}, S_{10}={pairs that are in the same cluster under $\mathbf{C}$ but in different ones under $\mathbf{C}'$} and S_{01}={pairs that are in different clusters under $\mathbf{C}$ but in the same under $\mathbf{C}'$}. Let $n_{ab} := |S_{ab}|$ for $a, b \in \{0, 1\}$ denote the repective sizes; then $n_{11} + n_{00} + n_{10} + n_{01} = \binom{n}{2}$ Specifically, the performance of the proposed test is compared to the reviewed existing methods using the Adjusted Rand index (ARI) as specified by Hubert and Arabie Hubert and Arabie (1985) and Fowlkes and Mallows Fowlkes and Mallows (1983).

Fowlkes and Mallows introduced their index as a measure for comparing hierarchical clusterings where the degree of similarity of two clusterings corresponds to the deviation from the expected value under the null hypothesis of independant clusterings with fixed cluster sizes. The index is given by:

$$R_{FM}\left(C,C'\right) = \frac{\sum\limits_{i=1}^{q}\sum\limits_{j=1}^{r} m_{ij}^2 - n}{\sqrt{\left(\sum\limits_{i=1}^{q}|C_i|^2 - n\right)\left(\sum\limits_{j=1}^{r}|C_i'|^2 - n\right)}} \tag{3.28}$$

Under this index, the null hypothesis, is that clusterings are independent. This is however violated in reality as it would be against the intuition to compare two clusterings when assuming that there is no relationship between them since we compare clusterings because we suppose a certain relationship and want to know how strong it is. Hubert and Arabie (ARI) relaxes this assertion and assumes a generalized hypergeometric distribution as null hypothesis where the two clusterings are drawn randomly with a fixed number of clusters and a fixed number of elements in each cluster though the number of clusters in the two clusterings need not be the same. It is considered to be the normalized difference of the Rand Index and its expected value under the null hypothesis and defined as follows:

$$R_{HA}\left(C,C'\right) = \frac{\sum\limits_{i=1}^{q}\sum\limits_{j=1}^{r}\binom{m_{ij}}{2} - \frac{2}{n(n-1)}\left(\sum\limits_{i=1}^{q}\binom{|C_i|}{2}\sum\limits_{j=1}^{r}\binom{|C_j'|}{2}\right)}{\frac{1}{2}\left(\sum\limits_{i=1}^{q}\binom{|C_i|}{2} + \sum\limits_{j=1}^{r}\binom{|C_j'|}{2}\right) - \frac{2}{n(n-1)}\left(\sum\limits_{i=1}^{q}\binom{|C_i|}{2} + \sum\limits_{j=1}^{r}\binom{|C_j'|}{2}\right)} \tag{3.29}$$

3.4 L^p **Space**

When time is modelled as a sequence of discrete dates, then financial time series are sequences of real numbers. An important family of sequences that is utilized to characterize financial time series in this work is the family of L^p spaces. To facilitate the construction of the proposed change-point estimator we make the following definitions.

Definition (L_w^p Space) Let I be a finite or countably index set. Let $w : I \to [0, \infty)$. Given a sequence of scalars $X = (X_i)_{i \in I}$, set

$$
||X||_{p,w} = \begin{cases} \left(\sum\limits_{i \in I} |X_i|^p w_i^p \right)^{\frac{1}{p}} & 0 < p < \infty \\ \\ sup \, |X_i| \, w_i & p = \infty \end{cases}
\tag{3.30}
$$

Then we set weighted L_w^p space as

$$
L_w^p(I) = \left\{ X = (X_i)_{i \in I} : ||X||_{p,w} < \infty \right\}
\tag{3.31}
$$

The L_w^p space is characterized by several useful propoerties as outlined in the subsequent theorems. These properties are used in the proof of theoretical results.

Theorem 5. *For $0 < p < \infty$, the space L^p is a vector space.*

Proof. If f and g are each in L^p and let c be a constant, then

   (i) L^p is closed under scalar multiplication

$$
|cf|^p \le c^p \, |f|^p
\tag{3.32}
$$

   (ii) L^p is closed under addition

$$
|f + g|^p \le [2 \, (|f| \vee |g|)]^p \le 2^p \, (|f|^p + |g|^p)
\tag{3.33}
$$

Thus $f + g$ is also in L^p Faris (2004).

$\square$

Theorem 6. *For $0 < p < \infty$, the space L^p is complete.*

Proof. Suppose that $\sum\limits_{j=1}^{\infty} f_j$ is absolutely convergent in L^p, then

$$\sum_{j=1}^{\infty} \left\| f_j \right\|_p \;=\; B < \infty \tag{3.34}$$

$$\left\| \sum_{j=1}^{\infty} |f_j| \right\|_p \;\leq\; \sum_{j=1}^{\infty} \left\| f_j \right\|_p \leq B \tag{3.35}$$

By the monotone convergence theorem $h = \sum\limits_{j=1}^{\infty} \left| f_j \right|$ is in L^p with L^p semi-norm bounded by B. In particular, it is convergent almost everywhere. It follows that the series $\sum\limits_{j=1}^{\infty} f_j$ converges almost everywhere to some limit g. The sequence $\sum\limits_{j=1}^{k} f_j$ is dominated by h in L^p and converges pointwise to $\sum\limits_{j=1}^{\infty} f_j$. Therefore, by the dominated convergence theorem, it converges to the same limit g in L^p semi-norm Faris (2004).

$\square$

Theorem 7. *(Schwarz inequality)*

Let u, v be vectors in the vector space in L^p, then it follows that

$$|\langle u, v \rangle| \leq \|u\| \, \|v\| \tag{3.36}$$

Proof. The proof of this theorem can be found as Theorem 2.1 in Faris (2004).

$\square$

Theorem 8. *(Triangle inequality)*

Let u, v be vectors in the vector space in L^p, then it follows that

$$\|u + v\| \leq \|u\| + \|v\| \tag{3.37}$$

38

Proof. The proof of this theorem can be found as Theorem 2.2 in Faris (2004).

□

Theorem 9. *(Reverse Triangle inequality)*

Let u, v be vectors in the vector space in L^p, then it follows that

$$\|u\| - \|v\| \leq \|u - v\| \tag{3.38}$$

Proof. Let $u = (u - v) + v$. Taking norms and applying the Triangle inequality (3.37) yields

$$
\begin{aligned}
\|u\| &= \|(u - v) + v\| \\
&\leq \|u - v\| + \|v\| \\
\|u\| - \|v\| &\leq \|u - v\| \tag{3.39}
\end{aligned}
$$

□

Theorem 10. *(Holder's Inequality) Let I be a finite or countable index set. Given $1 \leq p \leq \infty$, if $X = (X_k)_{k \in I} \in L^p(I)$ and $Y = (Y_k)_{k \in I} \in L^{p'}(I)$, where $\frac{1}{p} + \frac{1}{p'} = 1$ then $XY = (X_k Y_k)_{k \in I} \in L^1(I)$ and*

$$\|XY\|_1 \leq \|(X_k)_{k \in I}\|_p \|(Y_k)_{k \in I}\|_{p'} = \left(\sum_{k \in I} |X_k|^p \right)^{\frac{1}{p}} \left(\sum_{k \in I} |Y_k|^{p'} \right)^{\frac{1}{p'}} < \infty \tag{3.40}$$

Proof. The proof of this theorem can be found as Theorem 1.2.9 in Okikiolu (1971).

□

Definition (Absolutely Convergent Series) Let X be a normed space and let $\{f_n\}_{n \in \mathbb{N}}$ be a sequence of elements of X. If

$$\sum_{n=1}^{\infty} \|f_n\| < \infty \tag{3.41}$$

then we say that the series $\sum_n f_n$ is absolutely convergent in X.

3.5 Point Process Theory

Point process techniques are utilized in obtaining the structure of limit variables and limit processes which occur in the theory of summation in time series analysis. By utilizing the point process theory as developed by (Kallenberg, 1983), consider the state space of the point processes $S = \bar{\mathbb{R}}^n$ where $\bar{\mathbb{R}}^n = \mathbb{R} \cup \{\infty\} \cup \{-\infty\}$. Let $\mathscr{B}$ be the collection of Borel sets bounded away from the origin in $\mathbb{R}^n \setminus \{0\}$. Let the number of points of X_i belonging to $B \in \mathscr{B}$ be denoted by

$$|\{X_i \in S : i \in I\} \cap B| = \sum_{i \in I} \varepsilon_{X_i}(B) \tag{3.42}$$

where $\varepsilon_X(B) = \begin{cases} 1 & for \quad X_i \in B \\ 0 & for \quad X_i \notin B \end{cases}$ is a Dirac measure. A point measure μ representing the observed points X_i, $i \in I$ on the $\sigma - field \ \mathscr{B}$ is given by

$$\mu = \sum_{i \in I} \varepsilon_{X_i} \tag{3.43}$$

Write $\mathscr{M}$ for the space of point measures on $\mathbb{R}^n \setminus \{0\}$ such that $\mathscr{M} \equiv \mathscr{M}(S, B)$. Consider subspaces indicated by certain indices so that we have $\mathscr{M}_y \subset \mathscr{M}$ being a collection of measures μ such that $\mu(\{X : |X| > y\}) > 0$ so that $\mathscr{M}_0 = \mathscr{M} \setminus \{0\}$. Define $\mathscr{M}_f = \cup_{i \in I\mathbb{N}_0} \mathscr{M}_i$ as finite point measures which include the zero measure. Let $I\mathbb{N}_0 := I\mathbb{N} \cup \{0\}$ and $I\bar{\mathbb{N}}_0 := I\mathbb{N}_0 \cup \{\infty\}$. Since μ may be identified with $(\mu(B))_{B \in \mathscr{B}}$ then π_B such that $\pi_B : \mathscr{M} \to I\mathbb{N}_0$ or $\pi_B : \mu \to \mu(B)$ is a projection with index B. The $\sigma - field \ \mathscr{M} \equiv \mathscr{M}(S, B)$ introduced by $\mathscr{M}$ is the smallest one such that the projec-

tions π_B are measurable. Thus $\mathscr{M}$ is a $\sigma - field$ generated by sets of the form $\pi_B^{-1}(C)$, that is,

$$\mathscr{M} := \sigma\left(\{\pi_B : B \in \mathscr{B}\}\right) \ = \sigma\left(\left\{\pi_B^{-1}(c) : B \in \mathscr{B}, C \subset I\bar{\mathbb{N}}_0\right\}\right) \tag{3.44}$$

Let $(\Omega, \mathscr{A}, P)$ be a probability space on which the random variables are defined, the mapping

$$N : \Omega \to \mathscr{M}(S, B) \tag{3.45}$$

is called a Point process on (S, B) if it is measurable with respect to $\mathscr{A}$ and $\mathscr{M}(S, B)$ meaning that N is a random point measure on (S, B). $X_1, \cdots X_n$ can be represented by point process

$$N_n = \sum_{i=1}^{n} \varepsilon_{X_i} \tag{3.46}$$

Consider $g : (S, B) \to (T, C)$ to be a measurable mapping. Observing $g(X_i)$ instead of X_i can be rephrased in terms of point processes by having point measure defined as $\sum_{i \in I} \varepsilon_{g(X_i)}$ such that

$$\sum_{i \in I} \varepsilon_{g(X_i)} = \sum_{i \in I} 1_c(g(X_i)) = \sum_{i \in I} \varepsilon_{X_i}\left(g^{-1}(c)\right) = (gu)(c) \tag{3.47}$$

thus the mapping $g^* : \mathscr{M}(S, B) \to \mathscr{M}(T, C)$ can be defined as $g^*(\mu) := g\mu$ such that for a point process N on (S, B) we obtain another point process $g^*(N)$ on (T, C) if g^* is a measurable mapping.

To facilitate the characterization of point processes, define the intensity measure $v(B)$ also known as the mean measure which denotes the expected number of points in the set B as $v(B) = E[N(B)]$

The realisations of a point process N are point measures. Therefore the distribution of N is defined on subsets of point measures. Consequently, the distribution of N is uniquely determined by the family of the distributions of the finite dimensional random

vectors $(N(A_1), \cdots, N(A_m))$ for any choice of $A_1, \cdots, A_m$.

Theorem 11. *(Uniqueness theorem) Let $\mathscr{S}$ be a system generating $\mathscr{B}$ and for $i = 0, 1$*

let

$$N_i : (\Omega, \mathscr{A}, P) \to (M(S, \mathscr{B}), \mathscr{M}(S, \mathscr{B})) \tag{3.48}$$

be point processes, where N_0 is $\sigma - finite$ over $\mathscr{S}$. Then the following assertions are

equivalent:

(i) $N_0 \overset{d}{=} N_1$

(ii) *for every $m \in I\!N$ and $A_1, \cdots, A_m \in \mathscr{S}$*

$$(N_0(A_1), \cdots, N_0(A_m)) \overset{d}{=} (N_1(A_1), \cdots, N_1(A_m)) \tag{3.49}$$

Proof. The proof of this theorem can be found as Theorem 1.1.1 in Reiss (1993).

$\square$

The distribution of the of a random vector is described analytically by the use of the

Laplace transforms.

Definition (Laplace transform) The Laplace transform $\Psi_X : [0, \infty)^n \to [0, 1]$ of a real

valued $[0, \infty)^n$ random vector $\mathbf{X} = (X_1, \cdots, X_n)$ is defined by

$$\Psi_{\mathbf{X}}(\lambda_1, \cdots, \lambda_n) \;=\; \int exp\left(-\sum_{i=1}^{n} \lambda_i X_i\right)$$

Definition (Laplace functional) The Laplace functional

$$\Psi_N : \left\{ g : (S, \mathscr{B}) \to \left([0, \infty), IB\,|_{[0,\infty)}\right) \right\} \to [0, 1] \tag{3.50}$$

of the point process $N : (\Omega, \mathscr{A}, P) \to (IM(S, \mathscr{B}), \mathscr{M}(S, \mathscr{B}))$ is defined for non-negative measurable functions g on the state space S as

$$
\begin{aligned}
\Psi_{N(g)} &= E \exp\left\{ -\int g \, dN \right\} \\
&= \int \exp\left\{ -\int_S g(x) \, d\mu(x) \right\} dP_{N(\mu)}
\end{aligned}
\tag{3.51}
$$

The Laplace functional Ψ_N determines the distribution of a point process uniquely. It is also useful for studying the weak convergence of point processes. Suppose that the real valued function g has compact support, that is, there exists a compact set $K \subset S$ such that $g(x)$ on K^c the complement of K. Then we define

$C_K^+ = \{g : g \text{ is a continuous non} - \text{negative function on } S \text{ with compact support } g\}$.

Theorem 12. *(Criterion for weak convergence of point processes via convergence of Laplace functionals)*

The point processes N_n converge weakly to the point process N if and only if the corresponding Laplace functionals converge for every $g \in C_K^+$ as $n \to \infty$, that is

$$
\Psi_{N_n}(g) = E \exp\left\{ \int_S g \, dN_n \right\} \to \Psi_N(g) = E \exp\left\{ \int_S g \, dN \right\}
\tag{3.52}
$$

Proof. The proof of this theorem can be found as Theorem 8.1.2 in Reiss (1993).

$\square$

Consider a strictly stationary sequence $(X_t)_{t \in \mathbb{N}}$ of random row vectors with values in $\mathbb{R}^n$, that is, $X = (X_1, \cdots, X_n)$. The characterization of the asymptotic behavior of the tails of the random variable X are examined through the regular variation condition.

Theorem 13. *(Regular Variation Condition)*

In light of (Kesten, 1973) assume ε has a density with unbounded support, $\alpha_0 > 0$, $E\left[\ln\left(\alpha_1 \varepsilon^2 + \beta_1\right)\right] < 0$, $E\left|\alpha_1 \varepsilon^2 + \beta_1\right|^{\frac{p}{2}} \geq 1$ and $E|\varepsilon|^p \ln|\varepsilon| < \infty$ for some $p > 0$ holds, then:

43

(i) there exist a number $\kappa \in (0, p]$ which is a unique solution of the equation

$$E\left[\left(\beta_1 + \alpha_1 \varepsilon_n^2\right)^{\kappa/2}\right] = 1 \tag{3.53}$$

and there exist a positive constant $c_0 = c_0(\alpha_0, \alpha_1, \beta_1)$ such that

$$P(\sigma > x) \sim c_0 x^{-\kappa} \text{ as } x \to \infty \tag{3.54}$$

(ii) If $E|\varepsilon|^{\kappa+\xi} < \infty$ for some $\xi > 0$, then

$$P(|X| > x) \sim E|\varepsilon|^{\kappa} P(\sigma > x) \tag{3.55}$$

and the vector (X, σ) is jointly regularly varying such that

$$\frac{P(|X, \sigma| > xt, (X, \sigma) / |X, \sigma| \in B)}{P(|X, \sigma| > t)} \overset{v}{\longrightarrow} x^{-\kappa} P(\Theta \in B)$$

where $\overset{v}{\longrightarrow}$ denotes vague convergence on the Borel $\sigma - field$ of the unit sphere S^1 of $\mathbb{R}^2$, relative to the norm $|\cdot|$ with

$$P(\Theta \in \cdot) = \frac{E|(\varepsilon, 1)|^{\kappa} I_{\{(\varepsilon,1)/|(\varepsilon,1)|\in\cdot\}}}{E|(\varepsilon, 1)|^{\kappa}} \tag{3.56}$$

Proof. Following the works of (Kesten, 1973) and (Breiman, 1965), assume ξ and η are independent non-negative random variables such that $P(\xi > x) \sim L(x)x^{-\kappa}$ for some slowly varying function L and $E\eta^{\kappa+\varepsilon} < \infty$ for some $\varepsilon > 0$, then $P(n\xi > x) \sim E\eta^{\kappa} P(\xi > x)$ as $x \to \infty$.

Applying Theorem 13 yields

$$
\begin{aligned}
P\left(|X,\sigma| > xt, (X,\sigma)/|X,\sigma| \in B\right) &= P\left(\sigma|\varepsilon,1| > xt, (\varepsilon,1)/|\varepsilon,1| \in B\right) \\
&= P\left(\sigma|\varepsilon,1|I_{\{(\varepsilon,1)/|\varepsilon,1|\in B\}} > xt\right) \\
&\sim E|\varepsilon,1|^{\kappa}I_{\{(\varepsilon,1)/|\varepsilon,1|\in B\}}P\left(\sigma > xt\right) \\
&\sim E|\varepsilon,1|^{\kappa}I_{\{(\varepsilon,1)/|\varepsilon,1|\in B\}}x^{-\kappa}P\left(\sigma > t\right) \quad (3.57)
\end{aligned}
$$

also

$$
\begin{aligned}
P\left(|X,\sigma| > t\right) &= P\left(\sigma|\varepsilon,1| > t\right) \\
&\sim E|\varepsilon,1|^{\kappa}P\left(\sigma > t\right) \quad\quad\quad\quad (3.58)
\end{aligned}
$$

which completes proof.

$\square$

Theorem 14. *(Strongly Mixing Condition)*

Let (a_n) be a sequence of positive numbers such that

$$
nP\left(|X| > a_n\right) \to 1 \quad\quad\quad\quad (3.59)
$$

The sequence (a_n) can be chosen as the $\left(1 - n^{-1}\right) - quantile$ of $|X|$. Since $|X|$ is regularly varying, $a_n = n^{\frac{1}{\alpha}}L(n)$ for slowly varying function $L(x)$. The condition (3.59) holds for (X_t) if there exists a sequence of positive integers (r_n) such that $r_n \longrightarrow \infty$, $k_n = [n/r_n] \longrightarrow \infty$ as $n \longrightarrow \infty$ and

$$
E\left[\exp\left\{-\sum_{t=1}^{n} f\left(X_t/a_n\right)\right\}\right] - \left(E\left[\exp\left\{-\sum_{t=1}^{r_n} f\left(X_t/a_n\right)\right\}\right]\right)^{k_n} \to 0 \quad\quad as\ n \longrightarrow \infty, \forall f \in F_s
$$

$$
(3.60)
$$

Proof. For proof of this theorem see Davis and Mikosch (1998).

$\square$

The condition (3.60) implies by the strong mixing condition of the stationary sequence (X_t)Davis and Mikosch (1998). This condition ensures convergence of the Laplace functional form outlined in Theorem 3.5.

Assume that the joint regular variation in Theorem 13 and strongly mixing conditions in Theorem 14 are satisfied for a stationary sequence (X_t), then, the statement can be made for the weak convergence of the sequence of point processes

$$N_n = \sum_{t=1}^{n} \varepsilon_{X_t/a_n}, \quad n = 1, 2, \cdots \tag{3.61}$$

Theorem 15.

Assume that (X_t) is a stationary sequence of random vectors for which all finite-dimensional distributions are jointly regularly varying index $\kappa > 0$. To be specific, let $\theta_{-m}, \cdots, \theta_m$ be the $(2m+1)n-$dimensional random row vector with values in the unit sphere $\left(S^{(2m+1)n-1}\right)$, $m \geq 0$. Assume that the strongly mixing condition for (X_t) and that

$$\lim_{m\to\infty} \lim_{n\to\infty} \sup P\left(\bigvee_{t=1}^{r_n} |X_t| > a_n y \mid |X_0| > a_n y \right) = 0, \qquad y > 0, \tag{3.62}$$

Then the limit

$$\gamma = \lim_{m\to\infty} \frac{E\left(|\theta_0^{(m)}|^{\kappa} - \bigvee_{j=1}^{m} |\theta_j^{(m)}|^{\kappa} \right)_+}{E|\theta_0^{(m)}|^{\kappa}}$$

exists and is the extremal index of the sequence $(|X_t|)$.

  (i) *If $\gamma = 0$, then $N_n \xrightarrow{d} o$*

  (ii) *If $\gamma > 0$, then $N_n \xrightarrow{d} N \overset{d}{=} \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \varepsilon_{P_i Q_{ij}}$*

46

where $\sum_{i=1}^{\infty} \varepsilon_{P_i}$ is a Poisson process on $\mathbb{R}_+$ with P_i describing the radial part of the points and $\sum_{j=1}^{\infty} \varepsilon_{Q_{ij}}$ is a sequence of independent and identically distributed point processes with Q_{ij} decribing the spherical part and a joint distribution Q on $\left(\tilde{M}, B\left(\tilde{M}\right)\right)$, where Q is the weak limit of

$$Q = \lim_{m \to \infty} \frac{E\left(\left|\theta_0^{(m)}\right|^{\kappa} - \bigvee_{j=1}^{m}\left|\theta_j^{(m)}\right|^{\kappa}\right)_+ I\left(\sum_{|t| \leq m} \varepsilon_{\theta_t^{(k)}}\right)}{E\left(|\theta_0^{(m)}|^{\kappa} - \bigvee_{j=1}^{m}|\theta_j^{(m)}|^{\kappa}\right)_+}$$

Proof. For proof of this theorem see Davis and Hsing (1995).

$\square$

Define

$$\tilde{N}_n = \sum_{i=1}^{m_n} \tilde{N}_{r_n,i}, \quad i = 1, 2, \cdots, m_n \tag{3.63}$$

where $\tilde{N}_{r_n,i}$ are independent and identically distributed as $\tilde{N}_{r_n,0} = \sum_{t=1}^{r_n} \varepsilon_{X_t/a_n}$. It therefore follows from Theorems 3.5, 14 and 15 that (N_n) converges weakly if and only if $\tilde{N}_n$ does and they have the same limit N. N is identical in law to the point process $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \varepsilon_{P_iQ_{ij}}$ where $\sum_{i=1}^{\infty} \varepsilon_{P_i}$ is a Poisson process $\mathbb{R}_+$ with P_i describing the radial part of the points and $\sum_{j=1}^{\infty} \varepsilon_{Q_{ij}}$ is a sequence of independent and identically distributed point processes with Q_{ij} describing the spherical part and a joint distribution Q on $\left(\tilde{M}, B\left(\tilde{M}\right)\right)$.

We now consider the convergence of point processes which are products of random variables, which forms the basis of the results on the weak convergence of sample autocovariance and autocorrelation for stationary processes.

Theorem 16. *Let (X_t) be a strictly stationary sequence such that $(X_t) = ((X_t, \cdots, X_{t+m}))$ satisfying the jointly regularly varying condition for some $m \geq 0$ and further assume that Theorem 15 hold, then:*

$$\hat{N}_n = \left(\hat{N}_{n,h}\right)_{h=0,\cdots,m} = \left(\sum_{t=1}^{n} \varepsilon_{a_n^{-1} X_t X_{t+h}}\right)_{h=0,\cdots,m} \xrightarrow{d} \hat{N} = \left(\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \varepsilon_{P_i^2 Q_{ij}^{(0)} Q_{ij}^{(h)}}\right)_{h=0,\cdots,m}$$
(3.64)

where the points $Q_{ij} = \left(Q_{ij}^{(0)}, \cdots, Q_{ij}^{(m)}\right)$ *and* P_i *are as previously defined,* $\hat{N}_n$ *and* $\hat{N}$ *are point processes on* $\bar{\mathbb{R}} \setminus \{0\}$ *meaning that points are not included in the point processes if* $X_t X_{t+h} = 0$ *or* $Q_{ij}^{(0)} Q_{ij}^{(h)} = 0$

Proof. For proof of this theorem we follow the works of Davis and Mikosch (1998) who begin by showing the marginal convergence of $\hat{N}_{n,h}$ to $\hat{N}_h$. For $h = 0, \cdots, m$ define the mapping $\hat{T}_h : (X_0, \cdots, X_m) \in \bar{\mathbb{R}}^{m+1} \setminus \{0\} \to X_0 X_h$. Then $\hat{N}_{n,h} = N_n \circ \hat{T}_h^{-1}$ and $\hat{N}_h = N \circ \hat{T}_h^{-1}$ with all the null points excluded from $\hat{N}_{n,h}$ and $\hat{N}_h$. Let $A_1, \cdots, A_k$ be bounded interval in $\bar{\mathbb{R}}^{m+1} \setminus \{0\}$. By application of Theorems 3.5 and 3.5 then convergence of the finite-dimensional distributions is established as,

$$\left(\hat{N}_{n,h}(A_1), \cdots, \hat{N}_{n,h}(A_k)\right) \xrightarrow{d} \left(\hat{N}_h(A_1), \cdots, \hat{N}_h(A_m)\right)$$
(3.65)

$\square$

We study the weak limit behaviour of the sample autocovariance and sample autocorrelation of a stationary sequence (X_t). Construct from this process the strictly stationary n-dimensional processes $(X_t) = ((X_t, \cdots, X_{t+n}))$, $n \geq 0$. Define the sample autocovariance function

$$\gamma_{n,X}(h) = n^{-1} \sum_{t=1}^{n-h} X_t X_{t+h}, \qquad h \geq 0$$
(3.66)

and the corresponding sample autocorrelation function

$$\rho_{n,X}(h) = \frac{\gamma_{n,X}(h)}{\gamma_{n,X}(0)}, \qquad h \geq 1$$
(3.67)

Define the deterministic counterparts of the autocovariance and autocorrelation functions as follows

$$\gamma_X(h) = EX_0 X_h, \qquad h \geq 0 \qquad (3.68)$$

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)}, \qquad h \geq 1 \qquad (3.69)$$

Theorem 17. *Assume that (X_t) is a strictly stationary sequence of random variables and that for a fixed $m \geq 0$, (X_t) satisfies the regular variation condition and $N_n = \sum_{t=1}^{n} \varepsilon_{X_t/a_n} \xrightarrow{d} N = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \varepsilon_{P_i Q_{ij}}$ where the points $Q_{ij} = \left(Q_{ij}^{(0)}, \cdots, Q_{ij}^{(m)} \right)$ and P_i are as previously defined.*

(i) If $\kappa \in (0,2)$, then

$$\left(n a_n^{-2} \gamma_{n,X}(h) \right)_{h=0,\cdots,m} \xrightarrow{d} (V_h)_{h=0,\cdots,m} \qquad (3.70)$$

$$\left(\rho_{n,X}(h) \right)_{h=1,\cdots,m} \xrightarrow{d} \left(\frac{V_h}{V_0} \right)_{h=1,\cdots,m} \qquad (3.71)$$

where

$$V_h = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} P_i^2 Q_{ij}^{(0)} Q_{ij}^{(h)}, \qquad h = 0, 1, \cdots, m \qquad (3.72)$$

The vector $(V_0, \cdots, V_m)$ is jointly $\kappa/2$ stable in $\mathbb{R}^{m+1}$.

(ii) If $\kappa \in (2,4)$ and for $h = 0, \cdots, m$

$$\lim_{\varepsilon \to 0} \limsup_{n \to \infty} Var \left(a_n^{-2} \sum_{t=1}^{n-h} X_t X_{t+h} I_{\left\{ |X_t X_{t+h}| \leq a_n^2 \varepsilon \right\}} \right) = 0, \qquad (3.73)$$

then

$$\left(na_n^{-2}\left(\gamma_{n,X}\left(h\right)-\gamma_X\left(h\right)\right)\right)_{h=0,\cdots,m} \xrightarrow{d} \left(V_h\right)_{h=0,\cdots,m} \tag{3.74}$$

which implies that

$$\left(na_n^{-2}\left(\rho_{n,X}\left(h\right)-\rho_X\left(h\right)\right)\right)_{h=1,\cdots,m} \xrightarrow{d} \gamma_X^{-1}\left(0\right)\left(V_h-\rho_X\left(h\right)V_0\right)_{h=1,\cdots,m} \tag{3.75}$$

Proof. Part (i) of Theorem 17 follows immediately from Theorem 16 and continous mapping theorem 24. The convergence of $\gamma_{n,X}$ in part (ii) is also a direct consequence of Theorem 16. To prove (3.75) we express (3.74) as

$$\gamma_{n,X}\left(h\right)=\gamma_X\left(h\right)+O_p\left(n^{-1}a_n^2\right) \tag{3.76}$$

Since $n^{-1}a_n^2 \to 0$ as $n \to \infty$ and applying a Taylor series expansion to the function $f\left(x,y\right)=\frac{x}{y}$, then $\rho_{n,X}\left(h\right)$ can be expressed as

$$\rho_{n,X}\left(h\right)=\rho_X\left(h\right)+\frac{\gamma_{n,X}\left(h\right)-\gamma_X\left(h\right)}{\gamma_X\left(0\right)}-\frac{\left(\gamma_{n,X}\left(h\right)-\gamma_X\left(h\right)\right)\gamma_X\left(h\right)}{\gamma_X^2\left(0\right)}+O_p\left(n^{-1}a_n^2\right) \tag{3.77}$$

The result in (3.75) follows from immediate application of Continous Mapping and Slutsky's theorems on equation (3.77) resulting to (3.75).

$\square$

Theorem 17 outlines the limit distributions of the sample autocovariance $\gamma_{n,X}\left(h\right)$ and sample autocorrelation functions $\rho_{n,X}\left(h\right)$. When $\kappa \in \left(0,2\right)$ the sample autocorrelations have a constant limit only if $V_h=\rho_h V_0$ for some constant ρ_h, otherwise the limit is

random. For $\kappa \in (2,4)$, the sample autocorrelations has a random limit and $V_h \neq \rho_h V_0$. The limit distributions are however $\frac{\kappa}{2} - stable$ for $\kappa \in (0,2)$ and regularly varying with a rate of $n^{1-\frac{2}{\kappa}} L(n)$ for $\kappa \in (2,4)$ with some slowly varying function L.

3.6 Option Pricing

Here we outline the formulas for the option pricing.

3.6.1 Black-Scholes Option Pricing Model

Black-Scholes pricing model for an option on a non-dividend paying stock is as follows:

$$C = S_t N(d_1) - Ke^{-r(T-t)} N(d_2) \qquad (3.78)$$

$$P = Ke^{-r(T-t)} N(-d_2) - S_t N(-d_1) \qquad (3.79)$$

$$where \quad d_1 = \frac{log\left(\frac{S_t}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{(T-t)}}$$

$$d_2 = d_1 - \sigma\sqrt{(T-t)}$$

Where C is the price of a call option, P is the price of a put option, S_t is the spot price of the underlying at time t, $N(d_1)$ is a standard normal random variable with the value d_1, K is the strike price, r is the risk-free rate, $(T-t)$ is the time to maturity and σ is the annualized volatility. The volatility for the Black-Scholes model is estimated to be equal to annualized volatility of the underlying asset, that is, $\sigma = sd_{R_t}\sqrt{252}$, where sd_{R_t} is the standard deviation of the log-return series and 252 is the average number of trading days in a year.

3.6.2 Garman Kohlhagen Option Pricing Model

The Garman Kohlhagen model is an extension of Black–Scholes model and generally used on a dividend paying stock with the prices for the call and put given in (3.78) and (3.79) modified as:

$$C = S_t e^{-q(T-t)} N(d_1) - K e^{-r(T-t)} N(d_2) \qquad (3.80)$$

$$P = K e^{-r(T-t)} N(-d_2) - S_t e^{-q(T-t)} N(-d_1) \qquad (3.81)$$

$$\textit{where} \quad d_1 = \frac{\log\left(\frac{S_t}{K}\right) + \left(r - q + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{(T-t)}}$$

$$d_2 = d_1 - \sigma\sqrt{(T-t)}$$

The Garman Kohlhagen model is also extended to cater for two interest rates as in the presence of currency options where (3.80) and (3.81) are modified as:

$$C = S_t e^{-r^f(T-t)} N(d_1) - K e^{-r^d(T-t)} N(d_2) \qquad (3.82)$$

$$P = K e^{-r^d(T-t)} N(-d_2) - S_t e^{-r^f(T-t)} N(-d_1) \qquad (3.83)$$

$$\textit{where} \quad d_1 = \frac{\log\left(\frac{S_t}{K}\right) + \left(r^d - r^f + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{(T-t)}}$$

$$d_2 = d_1 - \sigma\sqrt{(T-t)}$$

Where C is the price of a call option, P is the price of a put option, S_t is the spot price of the underlying at time t, q is the dividend rate, $N(d_1)$ is a standard normal random variable with the value d_1, K is the strike price, r is the risk-free rate, r^d is the is domestic risk free rate, r^d is the is foreign risk free rate, $(T-t)$ is the time to maturity and σ is the annualized volatility. The volatility is estimated to be equal to annualized volatility of the underlying asset, that is, $\sigma = sd_{R_t}\sqrt{252}$, where sd_{R_t} is the standard

deviation of the log-return series and 252 is the average number of trading days in a year.

3.6.3 Heston and Nandi GARCH Option Pricing Model (HN-GARCH)

The Heston and Nandi GARCH option pricing models ($HN-GARCH(p,q)$) as developed by (Heston and Nandi, 2000) model is based on the assumption that the underlying spot price (S_t) follows a GARCH process with the log-returns stochastic volatility model defined as

$$R_t = ln\left(\frac{S_t}{S_{t-1}}\right) = r - \frac{1}{2}\sigma_t^2 + \sigma_t z_t \tag{3.84}$$

$$where \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i (z_{t-i} - \lambda_i \sigma_{t-i})^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$$

$$z_t|\mathscr{F}_{t-1} \overset{\mathbb{P}}{\sim} N(0,1)$$

They arrive at the risk-neutral GARCH (p,q) model and derive an almost closed-form option pricing formula. They derive the following option pricing formula for the European call option with strike price K that expires at time T:

$$C = \frac{1}{2}S_t + \frac{e^{-r(T-t)}}{\pi} Re\left[\frac{K^{-i\phi} f^*(i\phi+1)}{i\phi}\right] d\phi - Ke^{-r(T-t)}\left(\frac{1}{2} + \frac{1}{\pi} Re\left[\frac{K^{-i\phi} f^*(i\phi+1)}{i\phi}\right] d\phi\right) \tag{3.85}$$

where $Re[]$ denotes the real part of a complex number. $f^*(i\phi)$ is the conditional characteristic function of the log asset price using the risk neutral probabilities. i is the imaginary number, $\sqrt{-1}$. They also show that the conditional generating function of the asset price under the physical measure takes the form for the GARCH(1,1) model with $\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i (z_{t-i} - \lambda_i \sigma_{t-i})^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$

Chapter 4

CHANGE-POINT ESTIMATOR

4.1 Single Change-Point Estimator

Consider a financial time series modelled at discrete dates as a sequence of real numbers belonging to the L_p spaces. To facilitate the construction of the proposed change-point estimator the properties inherent to the L_p spaces as outlined in the literature review are utilized. Following **Defintion** (L_w^p Space) in Section 3.4, let I be a finite or countably infinite index sequence. For $p = 2$, then $L_2(I)$ consists of all sequences of scalars $X = (X_k)_{k \in I}$ such that

$$||X||_2 = ||(X_k)_{k \in I}||_2 = \left(\sum_{k \in I} |X_k|^2 \right)^{\frac{1}{2}} < \infty \tag{4.1}$$

Definition (Inner Product)

If X and Y are random variables in L_2 then we define the inner product of X and Y by $\langle X, Y \rangle = E(XY)$ satisfying the following properties

(i) $\langle X, Y \rangle = \langle Y, X \rangle$

(ii) $\langle X, X \rangle \geq 0$ and $\langle X, X \rangle = 0$ if and only if $P(X = 0) = 1$ implying that $X = 0$

(iii) $\langle aX, Y \rangle = a \langle X, Y \rangle$ for any constant a

(iv) $\langle X+Y,Z \rangle = \langle X,Z \rangle + \langle X,Z \rangle$

Definition (Expectation operator)

Let X and Y be random variables, then;

(i) by linearity of the expectation operator we have

$$E(X+Y) = E(X) + E(Y) \tag{4.2}$$

(ii) by absolute value of inequalities of the expectation operator we have

$$|E(X)| \leq E(|X|) \tag{4.3}$$

Let $I = \mathbb{N}$ be a finite index sequence and $(X_t)_{t \in \mathbb{N}}$ be a stationary time series. Let $X = (X_1, X_2, \cdots X_k)$ and $Y = (X_2, X_3, \cdots, X_{n-1})$ be a $(n-1)$ dimensional vectors satisfying Theorem 6 in Section 3.4. Consequently the subsequent Theorems 5, 3.4, 3.4 and 3.4 are satisfied. These Theorems are now applied to construct the proposed change-point estimator.

By application of Theorem 3.4 on Holders inequality in we have that, for a finite and countable index I if $X = (X_k)_{k \in I} \in L_2(I)$ and $Y = (Y_k)_{k \in I} \in L_2(I)$, where $\frac{1}{2} + \frac{1}{2} = 1$ then $XY = (X_k Y_k)_{k \in I} \in L_1(I)$ and

$$||XY||_1 \leq ||(X_k)_{k \in I}||_2 \, ||(Y_k)_{k \in I}||_2 = \left(\sum_{k \in I} |X_k|^2 \right)^{\frac{1}{2}} \left(\sum_{k \in I} |Y_k|^2 \right)^{\frac{1}{2}} < \infty \tag{4.4}$$

On application of the Schwartz inequality in Theorem 3.4 we can express the autocovariance and autocorrelation functions in terms of the inner product defined in Section 4.1 as

$$acovar \, \langle X,Y \rangle = \langle X - E(X), Y - E(Y) \rangle \tag{4.5}$$

55

$$acorr \langle X, Y \rangle = \left\langle \frac{X - E(X)}{sd(X)}, \frac{Y - E(Y)}{sd(Y)} \right\rangle \qquad (4.6)$$

where $sd(X)$ and $sd(Y)$ represents the standard deviation of X and Y respectively which represents an L_2 distance from the mean.

By the expectation operator previously defined in Section 4.1 we have that

$$E(|X||Y|) \leq \sqrt{E(X^2)}\sqrt{E(Y^2)} \qquad (4.7)$$

thus, applying the result in (4.7) to (4.5) and (4.6) yields

$$|acovar(X,Y)| \leq sd(X)\,sd(Y) \qquad \in \quad L_1 space \qquad (4.8)$$

$$|acorr(X,Y)| \leq 1 \qquad \in \quad L_1 space \qquad (4.9)$$

By the assumption that the series $(X_t)_{t\in\mathbb{N}}$ is ergodic, then it is implied that the sample moments converge in probability to the population moments. It therefore follows that the sample autocovariance and autocorrelation converge in probability to the population autocovariance and autocorrelation respectively.

Following (4.9) we can define a sequence of autocorrelation functions $\rho_{i+1,j}$ where for fixed $i = 0$, $1 \leq j \leq n-1$ and for fixed $j = n$, $1 \leq i \leq n-1$ to be such that we have two subsequences $\rho_{1j} = (\rho_{1,1}, \rho_{1,2}, \cdots, \rho_{1,k}, \cdots \rho_{1,n-1})$ and $\rho_{in} = (\rho_{2,n}, \rho_{3,n}, \cdots, \rho_{k+1,n}, \cdots, \rho_{nn})$ where $\rho_{1,k}$ and $\rho_{k+1,n}$ denote the autocorrelation of the sequence $\{X_t^2\}_{t=1}^{k}$ and $\{X_t^2\}_{t=k+1}^{n}$ for $1 \leq k \leq n$.

Assumption 4. *In this work, it will be assumed that the autocorrelation function is bounded such that* $0.05 \leq |\rho_{i,j}| \leq 1$.

We propose a change-point process D_n^k quantifying the deviation between $\rho_{1,k}$ and $\rho_{k+1,n}$ using a divergence measure motivated by the weighted L_p distance, with k de-

noting the change-point. For $p > 0$ we define

$$L_p\left(\rho_{1,k} - \rho_{k+1,n}\right) \quad = \quad \left(\sum_{k=1}^{n} w_k |\phi_k - \phi_{k+1}|^p\right)^{\frac{1}{p}} \tag{4.10}$$

$$where$$

$$\phi_k \quad = \quad \frac{\sum_{t=1}^{k-h} X_t^2 X_{t+h}^2}{\sum_{t=1}^{k} X_t^4} \quad for \quad \begin{array}{l} 0 < k < n \\ \\ 0 < h < n \end{array}$$

Specifically, we assume the case when $p = 1$ in (4.10) hence the Manhattan distance. Applying the Reverse Triangle inequality in Theorem 3.4 and by linearity and absolute value of inequalities of the expectation operator we have

$$\begin{aligned} L_1\left(\rho_{1,k} - \rho_{k+1,n}\right) &= \left(\sum_{k=1}^{n} w_k |\phi_k - \phi_{k+1}|\right) \\ &= E\left(w_k |\phi_k - \phi_{k+1}|\right) \\ &\geq w_k |E\left(\phi_k\right) - E\left(\phi_{k+1}\right)| \end{aligned} \tag{4.11}$$

To facilitate the construction of the proposed estimator the lower bound of the divergence measure $L_1\left(\rho_{1,k} - \rho_{k+1,n}\right)$ is assumed. Further assume that the autocorrelation function is calculated at lag $h : 0 < h < n$. The proposed change-point estimator is thus developed from the process generated by this measure as follows:

$$w_k |E\left(\phi_k\right) - E\left(\phi_{k+1}\right)| = w_k \left| \frac{1}{k} \sum_{i=1}^{k} \phi_i - \frac{1}{n-k} \sum_{i=k+1}^{n} \phi_i \right| \tag{4.12}$$

From (4.12) it can be seen that the proposed change-point process is a weighted difference between the sample autocorrelation functions $\phi_{1,k}$ and $\phi_{k+1,n}$ with w_k denoting the weight. The weight w_k is a measurable function that depends on the sample size n and change-point k. It is arbitrarily chosen such that it satisfies the condition that

$$\sum_{i=1}^{k} \phi_i \quad = \quad \frac{k}{n} \sum_{i=1}^{n} \phi_i$$

$$\Rightarrow \quad \frac{1}{n} \left(\sum_{i=1}^{k} \phi_i - \frac{k}{n} \sum_{i=1}^{n} \phi_i \right) = 0 \qquad (4.13)$$

Equating (4.12) and (4.13) determines the weight w_k as follows:

$$
\begin{aligned}
w_k \left(\frac{1}{k} \sum_{i=1}^{k} \phi_i - \frac{1}{n-k} \sum_{i=k+1}^{n} \phi_i \right) &= \frac{1}{n} \left(\sum_{i=1}^{k} \phi_i - \frac{k}{n} \sum_{i=1}^{n} \phi_i \right) \\
&= \frac{1}{n} \left(\sum_{i=1}^{k} \phi_i - \frac{k}{n} \sum_{i=1}^{k} \phi_i - \frac{k}{n} \sum_{i=k+1}^{n} \phi_i \right) \\
&= \frac{1}{n} \left(\left[1 - \frac{k}{n} \right] \sum_{i=1}^{k} \phi_i - \frac{k}{n} \sum_{i=k+1}^{n} \phi_i \right) \left(\frac{k}{k} \right) \left(\frac{n-k}{n-k} \right) \\
&= \frac{k}{n} \left(1 - \frac{k}{n} \right) \left(\frac{1}{k} \sum_{i=1}^{k} \phi_i - \frac{1}{n-k} \sum_{i=k+1}^{n} \phi_i \right) \\
\Rightarrow \quad w_k &= \frac{k}{n} \left(1 - \frac{k}{n} \right) \qquad (4.14)
\end{aligned}
$$

The resultant change-point process is obtained from (4.12) and (4.14) and defined as

$$D_n^k = \frac{k}{n} \left(1 - \frac{k}{n} \right) \left| \frac{1}{k} \sum_{i=1}^{k} \phi_i - \frac{1}{n-k} \sum_{i=k+1}^{n} \phi_i \right| \qquad (4.15)$$

Since the autocorrelation functions $\rho_{1,k}$ and $\rho_{k+1,n}$ are each in L_1 as shown in (4.9) then by application of Theorem 5 where the L_p is closed under scalar multiplication and addition, it is implied that the divergence measure $L_1 \left(\rho_{1,k} - \rho_{k+1,n} \right)$ specified in equation (4.10) is also in L_1 and thus the change-point process D_n^k also belongs to the L_1 space. On application of Theorem 6 L_1 is complete and under Assumption 4 we can conclude that D_n^k has a supremum.

The change-point estimator $\hat{k}$ of a change-point k^* is the first point at which there is maximal sample evidence for a break in the sample autocorrelation function of the

squared returns process. It is therefore estimated as the least value of k that maximizes the value of D_n^k where $1 < k < n$ is chosen as:

$$\hat{k} = min\left\{ k : D_n^k = \max_{1 < k < n} \left| D_n^k \right| \right\} \qquad (4.16)$$

4.2 Multiple Change-Point Estimation

The proposed change-point estimator $\hat{k}$, see equation (4.16), is designed for a single change-point detection. This can be extended to allow for multiple change-points through binary segmentation procedure. Let $I_i = \{1, k_1, \cdots, k_n, n\}$ denote the partition of the interval $[1, n]$ into sub-intervals $[1, k_1], \cdots, [k_{T-1}, k_T], [k_n, n]$ where $1 < k_1 < \cdots < k_T < n$. The technique is performed as follows:

1. The initial partition is taken to be the whole sample, that is, $I_0 = \{1, n\}$.

2. Each of the sub-intervals given by I_i is tested for change-points k_j using (4.16).

3. The m identified significant change-points k_j are added to the partition so that $I_{i+1} = \{1, k_1, \cdots, k_{T_i}, n\}$ for $T_i = T_{i-1} + m$.

4. Step (2) above is repeated until no more change-points are found.

Chapter 5

CONSISTENCY OF THE

CHANGE-POINT ESTIMATOR

Assume that the data X_t, for $t \in \mathbb{Z}$, are independent and sampled at equispaced points. $(X_t)_{t \in \mathbb{Z}}$ describes a financial returns time series modeled using $GARCH(p,q)$ model specified as model (3.1). Assume that this $GARCH(p,q)$ model has a finite fourth moment and let $E\left(\varepsilon_t^j\right) = v_j, j = 2, 4$. In addition let $\alpha_0 > 0$, $\alpha_i \geq 0$ and $\beta_j \geq 0$ for $i = 1, 2, \cdots, p$ and $j = 1, 2, \cdots, q$ in model (3.1). Let $p = q$ and $c_{i,t-i} = \beta_i + \alpha_i \varepsilon_{t-i}^2$ for $i = 1, 2, \cdots, p$, where $\{c_{i,t}\}$ is a sequence of independent and identically distributed random variables such that $c_{i,t}$ is independent of σ_t. This allows us to rewrite (3.1) as:

$$
\begin{aligned}
X_t &= \sigma_t \varepsilon_t \qquad for\, t \in \mathbb{Z} \\
\sigma_t^2 &= \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \\
&= \alpha_0 + \sum_{i=1}^{p} \left(\beta_i + \alpha_i \varepsilon_{t-i}^2\right) \sigma_{t-i}^2 \\
&= \alpha_0 + \sum_{i=1}^{p} c_{i,t-i} \sigma_{t-i}^2 \qquad\qquad (5.1)
\end{aligned}
$$

Proposition 1. *Consider a sample $X_1^2, X_2^2, \cdots, X_n^2$ satisfying alternative change-point*

hypothesis (3.26) and the change-point estimator $\hat{k}$ given by (4.16). If the sequences $\left\{X_{1,k}^2\right\}$ and $\left\{X_{2,k}^2\right\}$ satisfy

$$\delta = \frac{v_2\gamma_{11}^{h-1}\left[\bar{\gamma}_{11}\left(1-\gamma_{11}^2\right)-v_2\gamma_{11}\left(1-\gamma_{12}\right)\right]}{v_4\left(1-\gamma_{11}^2\right)-v_2^2\left(1-\gamma_{12}\right)} - \frac{v_2\gamma_{S_1}\left(1-\gamma_1\right)M_2\left(h\right)-v_2^2\gamma_{S_2}\left[1-\left(1-\gamma_1\right)M_1\left(h\right)\right]}{v_4\gamma_{S_1}\left(1-\gamma_1\right)-v_2^2\gamma_{S_2}} \neq 0$$

(5.2)

then for $\hat{\tau} = \frac{\hat{k}}{n}$,

$$P\{|\hat{\tau}-\tau^*|>\varepsilon\} \leq \frac{C}{\varepsilon^2\delta^2 n^{\frac{1}{2}}}$$

(5.3)

where C is a positive constant.

Proof. Suppose that $\left\{X_{1,k}^2, k \in \mathbb{Z}\right\}$ and $\left\{X_{2,k}^2, k \in \mathbb{Z}\right\}$ are two *GARCH*(p,q) sequences as defined in model (5.1). Further suppose that we observe a sample $X_1^2, X_2^2, \cdots, X_n^2$ from the model such that

$$X_k^2 = \begin{cases} X_{1,k}^2 & if \quad 1 \leq k \leq k^* \\ X_{2,k}^2 & if \quad k^* < k \leq n \end{cases}$$

(5.4)

where k^* is the unknown change-point. More specifically assume that the two sequences have different model order specification such that

$$X_k^2 = \begin{cases} GARCH\left(p_1,q_1\right) & if \quad 1 \leq k \leq k^* \\ GARCH\left(p_2,q_2\right) & if \quad k^* < k \leq n \end{cases}$$

(5.5)

where $p_1 \neq p_2$ and $q_1 \neq q_2$ but $p_1 = q_1 = 1$ and $p_2 = q_2$. Let $k^* = \tau^* n$ and assume that $0 < \tau < 1$, then in the presence of the change-point, the sequence $\left\{X_k^2\right\}$ is no longer stationary.

The proposed change-point estimator $\hat{k}$ of a change-point k^* is the point at which there is maximal sample evidence for a break in the sample autocorrelation function of the squared returns process. It is therefore estimated as the least value of k that maximizes the value of D_n^k as specified in (4.16) and (4.15) where $1 < k < n$ is chosen as:

$$\hat{k} = min\left\{ k : D_n^k = \max_{1<k<n} \left| D_n^k \right| \right\} \tag{5.6}$$

where

$$D_n^k = \frac{k}{n}\left(1 - \frac{k}{n}\right)\left| \frac{1}{k}\sum_{i=1}^{k} \phi_i - \frac{1}{n-k}\sum_{i=k+1}^{n} \phi_i \right| \tag{5.7}$$

$$\phi_k = \frac{\sum_{t=1}^{k-h} X_t^2 X_{t+h}^2}{\sum_{t=1}^{k} X_t^4} \quad for \quad \begin{array}{c} 0 < k < n \\ 0 < h < n \end{array} \tag{5.8}$$

Model (5.1) is utilised in the proof of consistency of the proposed change-point estimator. The foundation of this proof is based on the second and fourth moments of $\{X_t^2\}$ which will first be derived. Following the assumption that the second moment of $\{X_t\}$ exist it implies that $E(c_{i,t-i}) = \beta_i + \alpha_i v_2 < 1$. Let $\gamma_{i1} = E(c_{i,t-i})$, $\gamma_{i2} = E\left(c_{i,t-i}^2\right)$ and $\gamma_i = \sum_{i=1}^{p} \gamma_{ij}$ for $i = 1, 2, \cdots, p$ and $j = 1, 2$.

$$
\begin{aligned}
E\left(\sigma_t^2\right) &= E\left(\alpha_0 + \sum_{i=1}^{p} c_{i,t-i}\sigma_{t-i}^2\right) \\
&= \alpha_0 + \sum_{i=1}^{p} \gamma_{i1} E\left(\sigma_{t-i}^2\right) \\
&= \alpha_0 + \gamma_1 E\left(\sigma_{t-i}^2\right) \\
E\left(\sigma_t^2\right) &= \frac{\alpha_0}{1 - \gamma_1}
\end{aligned} \tag{5.9}
$$

Equation (5.9) shows that $E\left(\sigma_t^2\right) < \infty$ exists for $\gamma_1 < 1$.

$$
\begin{aligned}
E\left(\sigma_t^4\right) &= E\left(\alpha_0^2 + 2\alpha_0 \sum_{i=1}^{p} c_{i,t-i}\sigma_{t-i}^2 + \sum_{i=1}^{p} c_{i,t-i}^2 \sigma_{t-i}^4 + 2\sum_{l<m}^{p} c_{l,t-l}c_{m,t-m}\sigma_{t-l}^2\sigma_{t-m}^2\right) \\
&= \alpha_0^2 + 2\alpha_0\gamma_1 E\left(\sigma_t^2\right) + \gamma_2 E\left(\sigma_t^4\right) + 2\sum_{l<m} E\left(c_{l,t-l}c_{m,t-m}\sigma_{t-l}^2\sigma_{t-m}^2\right) \quad (5.10)
\end{aligned}
$$

To establish the $E\left(c_{l,t-l}c_{m,t-m}\sigma_{t-l}^2\sigma_{t-m}^2\right)$ we make use of the following Theorem as proved by (Changli and Timo, 1999).

Theorem 18. *Assume that $\lambda\left(\Gamma\right) < 1$. Under this condition,*

$$
E\left(c_{l,t-l}c_{m,t-m}\sigma_{t-l}^2\sigma_{t-m}^2\right) = \alpha_0\gamma_{l1}\gamma_{m1}M_1\left(l,m\right)E\left(\sigma_t^2\right) + \gamma_{l1}M_2\left(l,m\right)E\left(\sigma_t^4\right) \quad (5.11)
$$

where for $m-l > 1$

$$
\begin{aligned}
M_1\left(l,m\right) &= 1 + \gamma_{P\setminus\{m-l\}}' \left[\sum_{i=1}^{m-l-1}\left(\prod_{j=1}^{i}\Gamma_j\right)e_1 + \prod_{i=1}^{m-l}\Gamma_i\left(j_{p-1} + \Gamma_{m-l+1}\left(I_{p^*}-\Gamma\right)^{-1}e_{p-1}\right)\right] \\
j_{p-1} &= \left(1,1,\cdots,1\right)' \ is \ a \ \left(p-1\right)\times 1 \ vector \\
e_{p-1} &= \left(1,\cdots,1,0,\cdots,0\right)' \ is \ a \ p^* \times 1 \ vector \ with \ the \ first \ p-1 \ elements \ equal \ to \ 1 \\
\Gamma_k &= E\left(C_k\right) \ is \ a \ matrix \ of \ order \ \left(p-1\right)\times\left(p-1\right) \\
&\qquad with \ \Gamma_{m-l+1} \ a \ matrix \ of \ order \ \left(p-1\right)\times p^* \\
&\qquad and \ \Gamma \ a \ matrix \ of \ order \ p^* \times p^* \\
\lambda\left(\Gamma\right) &= max\left\{|\lambda_i|\right\} \ is \ maximum \ absolute \ eigenvalue \ of \ the \ matrix \ \Gamma
\end{aligned}
$$

In particular

$$
\begin{aligned}
M_1(m-1,m) &= 1 + \gamma'_{P\backslash\{1\}}\left[j_{p-1} + \Gamma_2(I_{p^*} - \Gamma)^{-1}e_{p-1}\right] \\
M_2(l,m) &= M_{21}(l,m) + \gamma_{m1}\sum_{i=2}^{4}M_{2i}(l,m) \\
M_{21}(l,m) &= \tilde{\gamma}_{m-l,m} + \gamma'_{P\backslash\{m-l\}}\left[\sum_{i=1}^{m-l-1}\left(\prod_{j=1}^{i}\Gamma_j\right)e_1\tilde{\gamma}_{m-l-i,m}\right] \\
M_{22}(l,m) &= \sum_{i=1}^{m-l-1}\gamma_{i1}M_{22}(m-l-i) + \sum_{j=m-l+1}^{p}\tilde{\gamma}_{j-m+l,j} \\
M_{23}(l,m) &= \sum_{i=m-l+1}^{m-l+p-1}\gamma(c(m-l),2,i-1) \\
M_{24}(l,m) &= \gamma'_{P\backslash\{m-l\}}\left[\left(\prod_{j=1}^{m-l+1}\Gamma_j\right)(I_{p^*}-\Gamma)^{-1}\right]\gamma(m-l+p+1,m-l+2p-1)
\end{aligned}
$$

Proof. For proof of Theorem 18 see Appendix 5 of (Changli and Timo, 1999).

$\square$

Substituting (5.9) and (5.11) in (5.10) yields

$$
\begin{aligned}
E\left(\sigma_t^4\right) &= \alpha_0^2 + 2\alpha_0\gamma_1 E\left(\sigma_t^2\right) + \gamma_2 E\left(\sigma_t^4\right) + 2\sum_{l<m}^{p}E\left(c_{l,t-l}c_{m,t-m}\sigma_{t-l}^2\sigma_{t-m}^2\right) \\
&= \alpha_0^2 + 2\alpha_0^2\gamma_1\frac{1}{1-\gamma_1} + \gamma_2 E\left(\sigma_t^4\right) + 2\sum_{l<m}^{p}\alpha_0^2\frac{1}{1-\gamma_1}\gamma_{l1}\gamma_{m1}M_1(l,m) \quad (5.12) \\
&\quad + 2\sum_{l<m}^{p}\gamma_{l1}M_2(l,m)E\left(\sigma_t^4\right)
\end{aligned}
$$

$$E\left(\sigma_t^4\right)\left[1-\gamma_2-2\sum_{l<m}^{p}\gamma_{l1}M_2\left(l,m\right)\right] = \alpha_0^2+2\alpha_0^2\gamma_1\frac{1}{1-\gamma_1}+2\sum_{l<m}^{p}\alpha_0^2\frac{1}{1-\gamma_1}\gamma_{l1}\gamma_{m1}M_1\left(l,m\right)$$

$$= \frac{1}{1-\gamma_1}\left[\alpha_0^2-\alpha_0^2\gamma_1+2\alpha_0^2\gamma_1+2\sum_{l<m}^{p}\alpha_0^2\gamma_{l1}\gamma_{m1}M_1\left(l,m\right)\right]$$

$$= \frac{\alpha_0^2}{1-\gamma_1}\left[1+\gamma_1+2\sum_{l<m}^{p}\gamma_{l1}\gamma_{m1}M_1\left(l,m\right)\right]$$

$$E\left(\sigma_t^4\right) = \frac{\alpha_0^2\left[1+\gamma_1+2\sum_{l<m}^{p}\gamma_{l1}\gamma_{m1}M_1\left(l,m\right)\right]}{\left(1-\gamma_1\right)\left[1-\gamma_2-2\sum_{l<m}^{p}\gamma_{l1}M_2\left(l,m\right)\right]} \tag{5.13}$$

From (5.13) it can be deduced that $E\left(\sigma_t^4\right)<\infty$ for $\gamma_1<1$ and $\gamma_2+2\sum_{l<m}^{p}\gamma_{l1}M_2\left(l,m\right)<1$.

Now we can evaluate the fourth moment of $\{X_t\}$ as

$$E\left(X_t^4\right) = E\left(\sigma_t^4\right)E\left(\varepsilon^4\right)$$

$$= \frac{\alpha_0^2\nu_4\left[1+\gamma_1+2\sum_{l<m}^{p}\gamma_{l1}\gamma_{m1}M_1\left(l,m\right)\right]}{\left(1-\gamma_1\right)\left[1-\gamma_2-2\sum_{l<m}^{p}\gamma_{l1}M_2\left(l,m\right)\right]} \tag{5.14}$$

Equation (5.14) implies that fourth moment of $\{X_t\}$ exist if $\gamma_1<1$ and $\gamma_2+2\sum_{l<m}^{p}\gamma_{l1}M_2(l,m)<1$.

Theorem 19.

The mixed moment $E\left(X_t^2X_{t+h}^2\right)$ has the form

$$E\left(X_t^2X_{t+h}^2\right) = \alpha_0\nu_2^2M_1\left(h\right)E\left(\sigma_t^2\right)+\nu_2M_2\left(h\right)E\left(\sigma_t^4\right) \tag{5.15}$$

where for $h\geq 1$,

$$M_1(h) = \gamma'_{p+1} \Gamma_*^{h-1} \left[e_{\alpha_0} + \Gamma_{h+1}^* (I_{p^*} - \Gamma)^{-1} e_{p-1} \right] \tag{5.16}$$

$$M_2(h) = M_{21}(h) + v_2 \sum_{i=2}^{4} M_{2i}(h) \tag{5.17}$$

$$M_{21}(h) = \bar{\gamma}_{h1} + \gamma'_{P \setminus \{h\}} \left[\sum_{i=1}^{h-1} \left(\prod_{j=1}^{i} \Gamma_j \right) e_1 \bar{\gamma}_{h-i,1} \right] \tag{5.18}$$

$$M_{22}(h) = \sum_{i=1}^{h-1} \gamma_{i1} M_{22}(n-i) + \sum_{j=h+1}^{p} \tilde{\gamma}_{j-h,j} \tag{5.19}$$

$$M_{23}(h) = \sum_{i=h+1}^{h+p-1} \gamma(c(h), 2, i-1) \tag{5.20}$$

$$M_{24}(h) = \gamma'_{P \setminus \{h\}} \left[\left(\prod_{j=1}^{h+1} \Gamma_j \right) (I_{p^*} - \Gamma)^{-1} \right] \gamma(h+p+1, h+2p-1) \tag{5.21}$$

$$e_{\alpha_0} = \left(\alpha_0^{-1}, 0, 1, \cdots, 1 \right)' \text{ is a } (p+1) \times 1 \text{ vector} \tag{5.22}$$

Proof. For proof of Theorem 18 see Appendix 9 of (Changli and Timo, 1999).

$\square$

The expected value of the sample autocorrelation function, $E(\phi_k)$, is first evaluated using (5.14) and (5.15).

$$\begin{aligned} E(\phi_h) &= \frac{E\left(X_t^2 X_{t+h}^2\right)}{E\left(X_t^4\right)} \quad for \quad \begin{array}{l} 0 < k < n \\ 0 < h < n \end{array} \\ &= \frac{v_2 \gamma_{S_1} (1 - \gamma_1) M_2(h) - v_2^2 \gamma_{S_2} \left[1 - (1 - \gamma_1) M_1(h) \right]}{v_4 \gamma_{S_1} (1 - \gamma_1) - v_2^2 \gamma_{S_2}} \end{aligned} \tag{5.23}$$

Further assuming that (5.4) and (5.5) are satisfied for $p_1 = q_1 = 1$, we evaluate (5.23) as follows:

$$
\begin{aligned}
M_1(h) &= \begin{pmatrix} \alpha_0 & \gamma_{11} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \alpha_0 & \gamma_{11} \end{pmatrix}^{h-1} \begin{pmatrix} \alpha_0^{-1} \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} \alpha_0 & \gamma_{11} \end{pmatrix} \begin{pmatrix} \alpha_0^{-1} \\ \sum_{i=0}^{h-2} \left(\prod_{j=1}^{i} \gamma_{11}^{j} \right) \end{pmatrix} \\
&= 1 + \gamma_{11} + \gamma_{11}^2 + \cdots + \gamma_{11}^{h-1} \\
&= \frac{1 - \gamma_{11}^h}{1 - \gamma_{11}}
\end{aligned}
\tag{5.24}
$$

$$
\begin{aligned}
M_2(h) &= \begin{pmatrix} \alpha_0 & \gamma_{11} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \alpha_0 & \gamma_{11} \end{pmatrix}^{h-3} \begin{pmatrix} 0 \\ \gamma_{11} \bar{\gamma}_{11} \end{pmatrix} \\
&= \begin{pmatrix} \alpha_0 & \gamma_{11} \end{pmatrix} \begin{pmatrix} 0 \\ \gamma_{11}^{h-2} \bar{\gamma}_{11} \end{pmatrix} \\
&= \gamma_{11}^{h-2} \bar{\gamma}_{11}
\end{aligned}
\tag{5.25}
$$

For *GARCH* $(1,1)$ model, $\gamma_{21} = \tilde{\gamma}_{12} = \bar{\gamma}_{21} = 0$. Substituting (5.24) and (5.25) in (5.23) results to

$$
E(\phi_h) = \begin{cases} \dfrac{v_2 \gamma_{11}^{h-1} \left[\bar{\gamma}_{11} \left(1 - \gamma_{11}^2\right) - v_2 \gamma_{11} (1 - \gamma_{12}) \right]}{v_4 \left(1 - \gamma_{11}^2\right) - v_2^2 (1 - \gamma_{12})} & for \quad 1 \leq k \leq k^* \\[2ex] \dfrac{v_2 \gamma_{s_1} (1 - \gamma_1) M_2(h) - v_2^2 \gamma_{s_2} \left[1 - (1 - \gamma_1) M_1(h) \right]}{v_4 \gamma_{s_1} (1 - \gamma_1) - v_2^2 \gamma_{s_2}} & for \quad k^* < k \leq n \end{cases}
\tag{5.26}
$$

Equation (5.26) shows that, in the presence of a change-point, the expected value of the sample autocorrelation function before and after the true change-point k^* is not equal. We consider a special case of change from *GARCH* $(1,1)$ to *GARCH* $(2,2)$ where we evaluate (5.23) for $p_2 = q_2 = 2$ as follows:

$$
\begin{aligned}
M_1(h) &= \begin{pmatrix} \alpha_0 & \gamma_{11} & \gamma_{21} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ \alpha_0 & \gamma_{11} & \gamma_{21} \\ 0 & 1 & 0 \end{pmatrix}^{n-1} \begin{pmatrix} \alpha_0^{-1} \\ 0 \\ (1-\gamma_{21})^{-1} \end{pmatrix} \\
&= \begin{pmatrix} \alpha_0 & \gamma_{11} & \gamma_{21} \end{pmatrix} \begin{pmatrix} \alpha_0^{-1} \\ (1-\gamma_{21})^{-1}\sum_{i=0}^{h-2}\left(\prod_{j=1}^{i}\gamma_{11}^{j}\right) \\ (1-\gamma_{21})^{-1}\sum_{i=0}^{h-2}\left(\prod_{j=1}^{i}-\gamma_{21}^{j}\right) \end{pmatrix} \\
&= (1-\gamma_{21})^{-1}\left[1+\frac{1-\gamma_{11}^{h}}{1-\gamma_{11}}-\frac{1-\gamma_{21}^{h}}{1-\gamma_{21}}\right] & \text{(5.27)} \\
M_1(1) &= (1-\gamma_{21})^{-1} & \text{(5.28)}
\end{aligned}
$$

Applying (5.18), (5.19), (5.20), and (5.21) and letting $h=1$ yields $M_2(1)$

$$
\begin{aligned}
M_{21}(1) &= \bar{\gamma}_{11} \\
M_{22}(1) &= \tilde{\gamma}_{12} \\
M_{23}(1) &= \gamma(c(1),2) = E(c_{2,t-3}c_{2,t-4},c_{1,t-4}) = \gamma_{21}\tilde{\gamma}_{12} \\
M_{24}(1) &= \gamma_{21}\Gamma_2(1-\gamma_{21})^{-1}\tilde{\gamma}_{12} = \gamma_{21}\tilde{\gamma}_{12}(1-\gamma_{21})^{-1} \\
M_2(1) &= \frac{\bar{\gamma}_{11}(1-\gamma_{21})+v_2\tilde{\gamma}_{12}}{(1-\gamma_{21})} & \text{(5.29)}
\end{aligned}
$$

The expected value, $E(\phi_k)$, for (5.5) for model order specification $p_1 = q_1 = 1$ and $p_2 = q_2$ for lag 1 results to

$$
E(\phi_1) = \begin{cases} \dfrac{v_2\left[\bar{\gamma}_{11}(1-\gamma_{11}^2)-v_2\gamma_{11}(1-\gamma_{12})\right]}{v_4(1-\gamma_{11}^2)-v_2^2(1-\gamma_{12})} & \text{for } 1\leq k\leq k^* \\[4mm] \dfrac{v_2\gamma_{s_1}\left[\bar{\gamma}_{11}(1-\gamma_{21})+v_2\tilde{\gamma}_{12}\right](1-\gamma_1)-v_2^2\gamma_{s_2}\gamma_{11}}{(1-\gamma_{21})\left[v_4\gamma_{s_1}(1-\gamma_1)-v_2^2\gamma_{s_2}\right]} & \text{for } k^* < k\leq n \end{cases} \quad \text{(5.30)}
$$

From (5.26) and (5.30), it can be seen that the expected values of the sample autocorrelation for *GARCH* $(1,1)$ and *GARCH* $(2,2)$ are not equal and we consequently assert

that

$$\delta = \frac{v_2\gamma_{11}^{h-1}\left[\bar{\gamma}_{11}\left(1-\gamma_{11}^2\right) - v_2\gamma_{11}\left(1-\gamma_{12}\right)\right]}{v_4\left(1-\gamma_{11}^2\right) - v_2^2\left(1-\gamma_{12}\right)} - \frac{v_2\gamma_{S_1}\left(1-\gamma_1\right)M_2\left(h\right) - v_2^2\gamma_{S_2}\left[1-\left(1-\gamma_1\right)M_1\left(h\right)\right]}{v_4\gamma_{S_1}\left(1-\gamma_1\right) - v_2^2\gamma_{S_2}} \neq 0$$

$$(5.31)$$

Thus the $\left|E\left(D_n^k\right)\right|$ is evaluated noting that it reaches its maximum at the point k^* resulting

to

$$E\left(D_n^k\right) = \begin{cases} \delta\tau\left(1-\tau^*\right) & if \quad k \leq k^* \\[2mm] \delta\tau^*\left(1-\tau\right) & if \quad k > k^* \end{cases}$$

$$(5.32)$$

Thus

$$E\left(D_n^{k^*}\right) = \delta\tau^*\left(1-\tau^*\right)$$

$$(5.33)$$

From (5.32) and (5.33) it follows that

$$\left|E\left(D_n^{k^*}\right)\right| - \left|E\left(D_n^k\right)\right| = \begin{cases} |\delta|\left(\tau^*-\tau\right)\left(1-\tau^*\right) & if \quad k \leq k^* \\[2mm] |\delta|\left(\tau-\tau^*\right)\tau^* & if \quad k > k^* \end{cases}$$

implying

$$\left|E\left(D_n^{k^*}\right)\right| - \left|E\left(D_n^k\right)\right| \geq |\delta||\tau^*-\tau|\left(\tau^* \wedge \left(1-\tau^*\right)\right)$$

$$(5.34)$$

We also have that

$$D_n^k - D_n^{k^*} = \left[D_n^k - E\left(D_n^k\right) + E\left(D_n^k\right) \right] - \left[D_n^{k^*} - E\left(D_n^{k^*}\right) + E\left(D_n^{k^*}\right) \right]$$

$$\left| D_n^k \right| - \left| D_n^{k^*} \right| \leq \left| D_n^k - E\left(D_n^k\right) \right| + \left| E\left(D_n^k\right) \right| + \left| D_n^{k^*} - E\left(D_n^{k^*}\right) \right| - \left| E\left(D_n^{k^*}\right) \right|$$

$$\leq 2 \max_{1 \leq k \leq n} \left| D_n^k - E\left(D_n^k\right) \right| + \left| E\left(D_n^k\right) \right| - \left| E\left(D_n^{k^*}\right) \right|$$

implying

$$\left| E\left(D_n^{k^*}\right) \right| - \left| E\left(D_n^k\right) \right| \leq 2 \max_{1 \leq k \leq n} \left| D_n^k - E\left(D_n^k\right) \right| + \left| D_n^{k^*} \right| - \left| D_n^k \right|$$

$$\leq 2 \max_{1 \leq k \leq n} \left| D_n^k - E\left(D_n^k\right) \right| \qquad since \quad \left| D_n^{k^*} \right| \geq \left| D_n^k \right| \qquad (5.35)$$

Thus from (5.34) and (5.35) as well as replacing τ with $\hat{\tau}$ in (5.34) we have that

$$|\delta| |\tau^* - \tau| (\tau^* \wedge (1 - \tau^*)) \leq \left| E\left(D_n^{k^*}\right) \right| - \left| E\left(D_n^k\right) \right|$$

$$\leq 2 \max_{1 \leq k \leq n} \left| D_n^k - E\left(D_n^k\right) \right| \qquad (5.36)$$

Consider D_n^k as given in (5.7), the estimate $\max_{1 \leq k \leq n} \left| D_n^k - E\left(D_n^k\right) \right|$ is now established as follows

$$\left| D_n^k - E\left(D_n^k\right) \right| = \frac{1}{n^2} \left| (n-k) \sum_{i=1}^{k} [\phi_i - E(\phi_i)] - k \sum_{i=k+1}^{n} [\phi_i - E(\phi_i)] \right|$$

$$= \frac{1}{n^2} \left| n \sum_{i=1}^{k} [\phi_i - E(\phi_i)] - k \sum_{i=1}^{k} [\phi_i - E(\phi_i)] - k \sum_{i=k+1}^{n} [\phi_i - E(\phi_i)] \right|$$

$$= \frac{1}{n^2} \left| n \sum_{i=1}^{k} [\phi_i - E(\phi_i)] - k \sum_{i=1}^{n} [\phi_i - E(\phi_i)] \right|$$

$$\leq \frac{1}{n} \left| \sum_{i=1}^{k} [\phi_i - E(\phi_i)] \right| + \frac{1}{n} \frac{k}{n} \left| \sum_{i=1}^{n} [\phi_i - E(\phi_i)] \right|$$

$$\leq \frac{1}{n} \left| \sum_{i=1}^{k} [\phi_i - E(\phi_i)] \right| + \frac{1}{n} \left| \sum_{i=1}^{k} [\phi_i - E(\phi_i)] \right|$$

$$\leq \frac{2}{n} \left| \sum_{i=1}^{k} [\phi_i - E(\phi_i)] \right| \tag{5.37}$$

implying

$$\max_{1 \leq k \leq n} \left| D_n^k - E\left(D_n^k\right) \right| \leq 2 \max_{1 \leq k \leq n} \frac{1}{n} \left| \sum_{i=1}^{k} [\phi_i - E(\phi_i)] \right| \tag{5.38}$$

Theorem 20. *Let $Y_1, Y_2, \cdots, Y_n$ be any random variables with finite second moments and $c_1, c_2, \cdots, c_n$ be any non-negative constants. Then*

$$\varepsilon^2 P\left\{ \max_{m \leq k \leq n} c_k \left| \sum_{i=1}^{k} Y_i \right| > \varepsilon \right\} \leq c_m^2 \sum_{i,j=1}^{m} E(Y_i Y_j) + \sum_{k=m}^{n-1} |c_{k+1}^2 - c_k^2| \sum_{i,j=1}^{k} E(Y_i Y_j)$$

$$+ 2 \sum_{k=m}^{n-1} c_{k+1}^2 E\left(|Y_{k+1}| \left| \sum_{j=1}^{k} Y_j \right| \right) + \sum_{k=m}^{n-1} c_{k+1}^2 E(Y_{k+1}^2) \tag{5.39}$$

Proof. For proof of Theorem 20 see Theorem 4.1 of (Piotr and Remigijus, 2000).

$\square$

Applying Theorem (20) with $m = 1$, $c_1 = c_2 = \cdots = c_n = \frac{1}{n}$, $Y_k = \phi_i - E(\phi_i)$ and Assumption 4 yields

$$\varepsilon^2 P\left\{\max_{1\le k\le n}\frac{1}{n}\left|\sum_{i=1}^{k}\left[\phi_i-E\left(\phi_i\right)\right]\right|>\varepsilon\right\} \le \frac{1}{n^2}\sum_{i,j=1}^{1}E\left(\left(\phi_i-E\left(\phi_i\right)\right)\left(\phi_j-E\left(\phi_j\right)\right)\right)$$

$$+\sum_{k=1}^{n-1}\left|\frac{1}{n^2}-\frac{1}{n^2}\right|\sum_{i,j=1}^{k}E\left(\left(\phi_i-E\left(\phi_i\right)\right)\left(\phi_j-E\left(\phi_j\right)\right)\right)$$

$$+2\sum_{k=1}^{n-1}\frac{1}{n^2}E\left(\left|\phi_{k+1}-E\left(\phi_{k+1}\right)\right|\left|\sum_{j=1}^{k}\left(\phi_j-E\left(\phi_j\right)\right)\right|\right)$$

$$+\sum_{k=1}^{n-1}\frac{1}{n^2}E\left(\phi_{k+1}-E\left(\phi_{k+1}\right)\right)^2$$

$$\le \frac{2}{n^2}\sum_{k=1}^{n-1}E\left(\left|\phi_{k+1}-E\left(\phi_{k+1}\right)\right|\left|\sum_{j=1}^{k}\left(\phi_j-E\left(\phi_j\right)\right)\right|\right)$$

$$+\frac{1}{n^2}\sum_{k=1}^{n-1}E\left(\phi_{k+1}-E\left(\phi_{k+1}\right)\right)^2$$

$$\le \frac{2}{n^2}\sum_{k=1}^{n-1}\sqrt{var\left(\phi_{k+1}\right)}\sqrt{\sum_{j,j'=1}^{k}cov\left(\phi_j^2,\phi_{j'}^2\right)}$$

$$+\frac{1}{n^2}\sum_{k=1}^{n-1}var\left(\phi_{k+1}\right)$$

$$= \frac{C}{n^2}\le\frac{C}{\sqrt{n}} \tag{5.40}$$

$$implying$$

$$P\left\{\max_{1\le k\le n}\frac{1}{n}\left|\sum_{i=1}^{k}\left[\phi_i-E\left(\phi_i\right)\right]\right|>\varepsilon\right\} \le \frac{C}{\sqrt{n}\varepsilon^2}$$

Substituting the result in (5.40) to (5.36) we have that

$$P\left\{\left|\delta\right|\left|\tau^*-\tau\right|\left(\tau^*\wedge\left(1-\tau^*\right)\right)>\varepsilon\right\} \le P\left\{\max_{1\le k\le n}\frac{1}{n}\left|\sum_{i=1}^{k}\left[\phi_i-E\left(\phi_i\right)\right]\right|>\varepsilon\right\}$$

$$\le \frac{C}{\sqrt{n}\varepsilon^2} \tag{5.41}$$

$$P\left\{\left|\tau^*-\tau\right|>\varepsilon\right\} \le \frac{C}{\varepsilon^2\delta^2 n^{\frac{1}{2}}} \tag{5.42}$$

As $n\to\infty$ from (5.41) we can see that $\frac{C}{\varepsilon^2\delta^2 n^{\frac{1}{2}}}\to 0$, which completes proof. Thus the

estimator τ is a consistent estimator of τ^* implying that k

is a consistent estimator of k^*.

Chapter 6

LIMITING DISTRIBUTION OF THE CHANGE-POINT ESTIMATOR

The following proposition is our main result on weak convergence for our proposed change-point process $D_n^k(h)$ for GARCH processes based on the point process theory.

Proposition 2. *Let $(X_t)_{t\in\mathbb{N}}$ be a strictly stationary sequence of random variables irrespective of the distribution of initial value X_0. Specifically, let $(X_t)_{t\in\mathbb{N}}$ be a GARCH $(1,1)$ process defined in the form of a stochastic differential equation (3.3). For fixed $h \geq 0$, set $X_t = (X_t, \cdots, X_{t+h})$. Assume that the regular variation conditions provided in Theorem 13 and the strongly mixing property in Theorem 14 hold. It therefore follows that the convergence of point processes N_n is established and thus Theorem 16 is satisfied. By extension of Theorem 17 then the convergence of the change-point process D_n^k as specified in (4.15) is established such that the following statements hold:*

(i) *If $\kappa \in (0,2)$, then*

$$\left(D_n^k(h)\right)_{h=1,\cdots,m} \xrightarrow{d} C_h\left(\frac{V_h}{V_0}\right)_{h=1,\cdots,n} \tag{6.1}$$

(ii) *If $\kappa \in (2,4)$ and for $h = 0, \cdots, m$*

$$\lim_{\delta \to 0} \lim_{n \to \infty} \operatorname{Sup} \operatorname{Var}\left(a_n^{-4} \sum_{t=1}^{n-h} X_t^2 X_{t+h}^2 I_{\left\{ |X_t^2 X_{t+h}^2| \le a_n^4 \varepsilon \right\}} \right) = 0, \qquad (6.2)$$

then

$$\left(na_n^{-4} \left(D_n^k(h) - D^{k^*}(h) \right) \right)_{h=1,\cdots,m} = na_n^{-4} \left(C_n \rho_{n,X^2}(h) - \rho_{X^2}(h) \right)_{h=1,\cdots,n}$$
$$\xrightarrow{d} \gamma_X^{-1}(0) \left(C_h V_h - \rho_{X^2}(h) C_0 V_0 \right)_{h=1,\cdots,m} \quad (6.3)$$

where

$$C_h = \frac{V_0}{V_h} \left(\frac{V_0^k V_h - V_h^k V_0}{V_0^k \left(V_0 - V_0^k \right)} \right) \qquad (6.4)$$

$$V_h = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} P_i^2 Q_{ij}^{(0)} Q_{ij}^{(h)}, \qquad h = 0, 1, \cdots, n \qquad (6.5)$$

$$V_h^k = \sum_{i=k+1}^{\infty} \sum_{j=k+1}^{\infty} P_i^2 Q_{ij}^{(0)} Q_{ij}^{(h)}, \qquad h = 0, 1, \cdots, n \qquad (6.6)$$

Proof. Consider the GARCH(1,1) model in the context of a stochastic differential equation (3.3) defined as $\sigma_t^2 = \alpha_0 + \left(\alpha_1 \varepsilon_{t-1}^2 + \beta_1 \right) \sigma_{t-1}^2$, then the necessary and sufficient conditions for stationarity are $\alpha_0 > 0$ and $E\left[\log\left(\beta_1 + \alpha_1 \varepsilon_n^2 \right) \right] < 0$ where the latter implies that $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$.

If we assume that the sample vector $X_1, \cdots, X_n$ comes from a stationary model, then the initial values X_0 also have a stationary distribution. This means that the distribution of X_t is stationary whatever the distribution of X_0, given the latter is independent of $(\varepsilon_t)_{t=1,2,\cdots}$ and stationarity conditions. To show this consider two sequences $X_t(X_0)_{t=0,1,2,\cdots}$ and $X_t(Z)_{t=0,1,2,\cdots}$ given the same stochastic differential equation recursion (3.3) but with initial conditions X_0 and Z where both vectors are independent of the future values $(A_t, B_t)_{t=1,2,\cdots}$. Futher assume that X_0 has stationary distribution.

By iteration of stochastic differential equation (3.3) we have

$$Y_t = B_t + \sum_{i=1}^{\infty} A_t \cdots A_{t-i+1} B_{t-i}, \qquad t = 1, 2, \cdots$$

Thus for any initial values Z we have the following recurssion

$$X_t(Z) = A_t \cdots A_1 Z + \sum_{j=1}^{t} A_t \cdots A_{j+1} B_j, \qquad t = 1, 2, \cdots$$

Then for any $\varepsilon > 0$ and for GARCH (1,1) model (3.3) the top Lyapunov exponent $\widetilde{\gamma}$ given by $A_n \cdots A_1 = A_n \prod_{t=1}^{n-1} \left(\beta_1 + \alpha_1 \varepsilon_n^2 \right)$

$$
\begin{aligned}
E \quad |X_t(X_0) - X_t(Z)|^{\varepsilon} \quad &\leq E |A_t \cdots A_1 (X_0 - Z)|^{\varepsilon} \\
&= \quad E |A_1 (X_0 - Z)|^{\varepsilon} \left(E \left| \beta_1 + \alpha_1 \varepsilon_n^2 \right|^{\varepsilon} \right)^{t-1} \\
&\leq \quad E \|A_1\| E |X_0 - Z|^{\varepsilon} \left(E \left| \beta_1 + \alpha_1 \varepsilon_n^2 \right|^{\varepsilon} \right)^{t-1} \quad (6.7)
\end{aligned}
$$

If $E |\varepsilon|^{2\varepsilon} < \infty$, $E |X_0|^{\varepsilon} < \infty$ and $E |Z|^{\varepsilon} < \infty$, then the right hand side is also finite. In addition given the stationary conditions previously stated then $E \mid \beta_1 + \alpha_1 \varepsilon_n^2 \mid^{\varepsilon} < 1$ for some sufficiently small ε. Thus the left hand side of (6.7) decays to zero as $t \to \infty$. Thus we conclude that $(X_t)_{t \in \mathbb{N}}$ is stationary irrespective of the distribution of the initial values X_0.

Consider the change-point process D_n^k to be a point process. Theorems outlined in Section 3.5 are used to determine the limiting behaviour of D_n^k. Begin by considering the sample autocorrelation function as defined in (3.67), then it follows that the following statements hold,

$$\sum_{t=1}^{n-h} X_t^2 X_{t+h}^2 = \sum_{t=1}^{k} X_t^2 X_{t+h}^2 + \sum_{t=k+1}^{n-h} X_t^2 X_{t+h}^2 \tag{6.8}$$

$$\sum_{t=1}^{n-h} X_t^4 = \sum_{t=1}^{k} X_t^4 + \sum_{t=k+1}^{n-h} X_t^4 \tag{6.9}$$

From (6.8) and (6.9) it can be asserted that there exists constants $c_{k,X^2}(h)$ and $c_{n-k,X^2}(h)$ such that the autocorrelation functions $\rho_{k,X^2}(h)$ and $\rho_{n-k,X^2}(h)$ can be expressed in terms of the autocorrelation function $\rho_{n,X^2}(h)$ as follows:

$$\rho_{k,X^2}(h) = \frac{\sum_{t=1}^{k-h} X_t^2 X_{t+h}^2}{\sum_{t=1}^{k} X_t^4} \propto \rho_{n,X}(h) = \frac{\sum_{t=1}^{n-h} X_t^2 X_{t+h}^2}{\sum_{t=1}^{n} X_t^4}$$

$$\rho_{k,X^2}(h) = c_{k,X}(h)\rho_{n,X^2}(h) \tag{6.10}$$

and

$$\rho_{n-k,X^2}(h) = \frac{\sum_{t=k+1}^{n-k-h} X_t^2 X_{t+h}^2}{\sum_{t=k+1}^{n-k} X_t^4} \propto \rho_{n,X}(h) = \frac{\sum_{t=1}^{n-h} X_t^2 X_{t+h}^2}{\sum_{t=1}^{n} X_t^4}$$

$$\rho_{n-k,X^2}(h) = c_{n-k,X^2}(h)\rho_{n,X^2}(h) \tag{6.11}$$

The change-point process (4.15) as in terms of (6.10) and (6.11) as

$$\begin{aligned} D_n^k(h) &= \rho_{k,X^2}(h) - \rho_{n-k,X^2}(h) \\ &= c_{k,X^2}(h)\rho_{n,X^2}(h) - c_{n-k,X^2}(h)\rho_{n,X^2}(h) \\ &= \left(c_{k,X^2}(h) - c_{n-k,X^2}(h)\right)\left(\rho_{n,X^2}(h)\right) \end{aligned} \tag{6.12}$$

The weak limits of the process $D_n^k(h)$ is characterized in terms of the limiting point

processes for the sample autocovariance and autocorrelation functions through the application of the Continuous Mapping Theorem 24. To complete the prove we independently prove the convergence of $c_{k,X^2}(h) - c_{n-k,X^2}(h)$ and $\rho_{n,X^2}(h)$ and apply Theorem 24.

Let $\delta > 0$, $X_t = \left(x_{t,X}^{(0)}, x_{t,\sigma}^{(0)}, \cdots, x_{t,X}^{(n)}, x_{x,\sigma}^{(n)} \right) \in \bar{\mathbb{R}}^{n+1} \setminus \{0\}$. In order to proof the results, we define several mappings

$$T_{h,\delta,X} : M \to \bar{R} \tag{6.13}$$

as follows

$$T_{0,\delta,X}(N_n) = \sum_{t=1}^{\infty} n_t \left(x_{t,X}^{(0)} \right) I_{\left\{ \left| x_{t,X}^{(0)} \right| > \delta \right\}} \tag{6.14}$$

$$T_{1,\delta,X}(N_n) = \sum_{t=1}^{\infty} n_t \left(x_{t,X}^{(1)} \right) I_{\left\{ \left| x_{t,X}^{(0)} \right| > \delta \right\}} \tag{6.15}$$

$$T_{h,\delta,X}(N_n) = \sum_{t=1}^{\infty} n_t \left(x_{t,X}^{(0)} \right) \left(x_{t,X}^{(h-1)} \right) I_{\left\{ \left| x_{t,X}^{(0)} \right| > \delta \right\}} \quad h \in [2,n] \tag{6.16}$$

The set $\left\{ X_t \in \bar{\mathbb{R}}^{n+1} \setminus \{0\} \,\middle|\, \left| x^{(h)} \right| > \delta \right\}$ is bounded for any $h \geq 0$ and thus the mappings are continos with respect to the limit point processes N. Consequently by Continuous Mapping Theorem 24 we have that

$$T(N_n) \xrightarrow{d} T(N) \tag{6.17}$$

where

$$T(N) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} P_i^2 Q_{ij}^{(0)} Q_{ij}^{(h)} I_{\left\{ \left| P_i Q_{ij}^{(0)} \right| > \delta \right\}}$$

The prove of the convergence of of $\rho_{n,X^2}(h)$ is examined for $\kappa \in (0,2)$and $\kappa \in (2,4)$. For the case of $\kappa \in (0,2)$, the point process results of Theorem 17 holds and a direct application of Theorem 17 yields:

78

$$\left(na_n^{-4}\gamma_{n,X^2}(h)\right)_{h=0,\cdots,m} \xrightarrow{d} (V_h)_{h=0,\cdots,m} \qquad (6.18)$$

$$\left(\rho_{n,X^2}(h)\right)_{h=1,\cdots,m} \xrightarrow{d} \left(\frac{V_h}{V_0}\right)_{h=1,\cdots,m} \qquad (6.19)$$

Lemma 1. *Assuming that the regular variation condition holds, $E\left(\varepsilon_t^4\right) < \infty$ and that the convergence of point processes as given in (6.17) is satisfied, then,*

$$Var\left[\alpha_1^2 a_n^{-4}\sum_{t=1}^{n}\left[\sigma_t^4\left(\varepsilon_{t+1}^2-1\right)^2\right]I_{\{\sigma_t\leq a_n\delta\}}\right] \leq a_n^{-8}\sum_{t=1}^{n}E\left(\left(\sigma_t^4\right)^2 I_{\{\sigma_t\leq a_n\delta\}}\right)E\left(\left(\varepsilon_{t+1}^2-1\right)^2\right)$$

$$\sim \quad const\ \delta^{8-\kappa} \qquad as\ n\to\infty \qquad (6.20)$$

$$\to \quad 0 \qquad\qquad\quad as\ \delta\to 0$$

For $\kappa\in(2,4)$ we commence with the $\left\{\sigma_t^2\right\}$ sequence and establish the convergence of $\gamma_{n,\sigma^2}(0)$. We rewrite $\gamma_{n,\sigma^2}(0)$ using the recurrence structure of the SDE (3.3) so that $\varepsilon_t^2 = \alpha_1^{-1}\left((\alpha_1\varepsilon_t^2+\beta_1)-\beta_1\right) = \alpha_1^{-1}\left(A_{t+1}-\beta_1\right)$ and $\sigma_t^2 = \alpha_0 + A_t\sigma_{t-1}^2 \approx A_t\sigma_{t-1}^2 = \left(\alpha_1\left(\varepsilon_t^2-1\right)+(\alpha_1+\beta_1)\right)\sigma_{t-1}^2$. Now using this representation yields:

$$na_n^{-4}\left(\gamma_{n,\sigma^2}(0)-\gamma_{\sigma^2}(0)\right)$$

$$= a_n^{-4}\sum_{t=1}^{n}\left[\sigma_t^4-E\left(\sigma^4\right)\right]$$

$$= a_n^{-4}\sum_{t=1}^{n}\left[\left\{\left(\alpha_1\left(\varepsilon_t^2-1\right)+\left(\alpha_1+\beta_1\right)\right)\sigma_{t-1}^2\right\}^2-E\left(\sigma^4\right)\right]$$

$$= a_n^{-4}\sum_{t=1}^{n}\sigma_{t-1}^4\left(\varepsilon_t^2-1\right)^2\alpha_1^2+2\alpha_1\left(\alpha_1+\beta_1\right)\left(\varepsilon_t^2-1\right)\sigma_{t-1}^2+$$

$$+\left(\alpha_1+\beta_1\right)^2\sigma_{t-1}^4-\left(\alpha_1+\beta_1\right)^2E\left(\sigma^4\right)+\left(\alpha_1+\beta_1\right)^2E\left(\sigma^4\right)-E\left(\sigma^4\right) \qquad (6.21)$$

$$= \alpha_1^2a_n^{-4}\sum_{t=1}^{n}\left[\sigma_{t-1}^4\left(\varepsilon_t^2-1\right)^2\right]+\left(\alpha_1+\beta_1\right)^2a_n^{-4}\sum_{t=1}^{n}\left[\sigma_{t-1}^4-E\left(\sigma^4\right)\right]+nE\left(\sigma^4\right)\left[\left(\alpha_1+\beta_1\right)^2-1\right]$$

$$= \alpha_1^2a_n^{-4}\sum_{t=1}^{n}\left[\sigma_{t-1}^4\left(\varepsilon_t^2-1\right)^2\right]+\left(\alpha_1+\beta_1\right)^2a_n^{-4}\sum_{t=1}^{n}\left[\sigma_{t-1}^4-E\left(\sigma^4\right)\right]+Op(1)$$

Grouping like terms in 6.21 yields

$$= \left[1-\left(\alpha_1+\beta_1\right)^2\right]na_n^{-4}\left(\gamma_{n,\sigma^2}(0)-\gamma_{\sigma^2}(0)\right)$$

$$= \alpha_1^2a_n^{-4}\sum_{t=1}^{n}\left[\sigma_{t-1}^4\left(\varepsilon_t^2-1\right)^2\right]+Op(1)$$

$$= \alpha_1^2a_n^{-4}\sum_{t=1}^{n}\left[\sigma_t^4\left(\varepsilon_t^2-1\right)^2\right]I_{\{\sigma_t>a_n\delta\}}+\alpha_1a_n^{-4}\sum_{t=1}^{n}\left[\sigma_t^4\left(\varepsilon_t^2-1\right)^2\right]I_{\{\sigma_t\leq a_n\delta\}}+Op(1)$$

$$(6.22)$$

By application of Lemma 1 we have

$$na_n^{-4}\left(\gamma_{n,\sigma^2}(0)-\gamma_{\sigma^2}(0)\right) = a_n^{-4}\sum_{t=1}^{n}\sigma_{t+1}^4I_{\{\sigma_t>a_n\delta\}}-\left(\alpha_1+\beta_1\right)^2a_n^{-4}\sum_{t=1}^{n}\sigma_t^4I_{\{\sigma_t>a_n\delta\}}+Op(1) \quad(6.23)$$

Lemma 2. *Assuming that the regular variation condition holds, $E\left(\varepsilon_t^4\right)<\infty$ and that the convergence of point processes as given in (6.17) is satisfied, then convergence of (6.23) is established such that:*

$$
\begin{aligned}
na_n^{-4}\left(\gamma_{n,\sigma^2}(0)-\gamma_{\sigma^2}(0)\right) &= a_n^{-4}\sum_{t=1}^{n}\sigma_{t+1}^4 I_{\{\sigma_t>a_n\delta\}}-(\alpha_1+\beta_1)^2 a_n^{-4}\sum_{t=1}^{n}\sigma_t^4 I_{\{\sigma_t>a_n\delta\}}+O_p(1) \\
&= T_{1,\delta,\sigma}\left(N_n^{(2)}\right)-(\alpha_1+\beta_1)^2 T_{0,\delta,\sigma}\left(N_n^{(2)}\right) \\
&\xrightarrow{d} T_{1,\delta,\sigma}\left(N^{(2)}\right)-(\alpha_1+\beta_1)^2 T_{0,\delta,\sigma}\left(N^{(2)}\right) \qquad (6.24)\\
&\simeq S(\delta,\infty)\xrightarrow{d} V_0^*
\end{aligned}
$$

We utilize the argument given in Theorem 27 where $S(\delta,\infty)\xrightarrow{d} V_0^*$ as $\delta\to 0$. Therefore, we finally obtain that:

$$
na_n^{-4}\left(\gamma_{n,\sigma^2}(0)-\gamma_{\sigma^2}(0)\right)\xrightarrow{d}\frac{1}{1-((\alpha_1+\beta_1))}V_0^*\simeq V_0 \qquad (6.25)
$$

Analogously following the same steps, it can be shown that the convergence of $\gamma_{n-k,\sigma}(0)$ is given by

$$
(n-k)a_{n-k}^{-2}\left(\gamma_{n-k,\sigma}(0)-\gamma_{\sigma}(0)\right)\xrightarrow{d}\frac{1}{1-((\alpha_1+\beta_1))}V_0^{k*}\simeq V_0^k \qquad (6.26)
$$

In the presence of a change-point k as hypothesized (3.20) it is evident that $E(A_t)\neq\alpha_1+\beta_1$ for all t but rather

$$
E(A_t)=\begin{cases} \alpha_1+\beta_1 & for \quad 1<t\leq k \\[2mm] E(A) & for \quad k<t<n \end{cases} \qquad (6.27)
$$

Thus the convergence of $\gamma_{k,\sigma^2}(0)$ and $\gamma_{n-k,\sigma^2}(0)$ are respectively given by

$$
ka_k^{-4}\left(\gamma_{k,\sigma^2}(0)-\gamma_{\sigma^2}(0)\right)\xrightarrow{d}\frac{1}{1-((\alpha_1+\beta_1))}V_0^{k*}\simeq V_0^k \qquad (6.28)
$$

$$
(n-k)a_{n-k}^{-4}\left(\gamma_{n-k,\sigma^2}(0)-\gamma_{\sigma^2}(0)\right)\xrightarrow{d}\frac{1}{1-E(A)}V_0^{(n-k)*}\simeq V_0^{n-k} \qquad (6.29)
$$

Following (6.27), (6.28) and (6.29) it is concluded that $V_0^k \neq V_0^{n-k}$.

Convergence of $\gamma_{n,\sigma^2}(1)$ is determined in a similar manner where by application of Lemma 1 we have that:

$$
\begin{aligned}
&= n a_n^{-4} \left(\gamma_{n,\sigma^2}(1) - \gamma_{\sigma^2}(1) \right) \\
&= a_n^{-4} \sum_{t=1}^{n} \left[\sigma_t^2 \sigma_{t+1}^2 - E\left(\sigma^4\right) \right] \\
&= a_n^{-4} \sum_{t=1}^{n} \left[\sigma_t^2 \sigma_{t+1}^2 - \sigma_t^4 EA \right] I_{\{\sigma_t > a_n \delta\}} \quad (6.30)
\end{aligned}
$$

Lemma 3. *Assuming that the regular variation condition holds, $E\left(\varepsilon_t^4\right) < \infty$ and that the convergence of point processes as given in (6.17) is satisfied, then convergence of (6.30) is established such that:*

$$
\begin{aligned}
n a_n^{-4} \left(\gamma_{n,\sigma^2}(1) - \gamma_{\sigma^2}(1) \right) &= T_{2,\delta,\sigma}\left(N_n^{(2)}\right) - EA\, T_{1,\delta,\sigma}\left(N_n^{(2)}\right) \\
&\xrightarrow{d} V_1
\end{aligned}
$$

Consequently for arbitrary lags we have

$$
n a_n^{-4} \left(\gamma_{n,\sigma^2}(h) - \gamma_{\sigma^2}(h) \right) \xrightarrow{d} V_h \quad (6.31)
$$

$$
(n-k) a_{n-k}^{-4} \left(\gamma_{n-k,\sigma^2}(h) - \gamma_{\sigma^2}(h) \right) \xrightarrow{d} V_h^k \quad (6.32)
$$

In the presence of a change-point k the convergence of $\gamma_{k,\sigma^2}(1)$ and $\gamma_{n-k,\sigma^2}(1)$ are respectively given by

$$
k a_k^{-4} \left(\gamma_{k,\sigma^2}(1) - \gamma_{\sigma^2}(1) \right) \xrightarrow{d} V_1^k \quad (6.33)
$$

$$(n-k)a_{n-k}^{-4}\left(\gamma_{n-k,\sigma^2}(1)-\gamma_{\sigma^2}(1)\right)\xrightarrow{d}V_1^{n-k}\tag{6.34}$$

Now we consider the $\{X_t\}$ sequence and establish the the convergence of $\gamma_{n,X}(0)$ by direct application of Lemma 1 and Lemma 2 as follows:

$$\begin{aligned}na_n^{-4}\left(\gamma_{n,X^2}(0)-\gamma_{X^2}(0)\right)&=a_n^{-4}\sum_{t=1}^n\left[X_t^4-E\left(X_0^4\right)\right]\\&=2a_n^{-4}\sum_{t=1}^n\left[\sigma_t^4\left(\varepsilon_t^2-1\right)^2\right]I_{\{\sigma_t>a_n\delta\}}\\&\xrightarrow{d}T_{1,\delta,\sigma}\left(N^{(2)}\right)-(\alpha_1+\beta_1)^2T_{0,\delta,\sigma}\left(N^{(2)}\right)\\&\xrightarrow{d}V_0\end{aligned}\tag{6.35}$$

Thus we have that

$$na_n^{-4}\left(\gamma_{n,X^2}(0)-\gamma_{X^2}(0)\right)\xrightarrow{d}V_0\tag{6.36}$$

Similarly it can be shown that the convergence of $\gamma_{k,X^2}(0)$ and $\gamma_{n-k,X^2}(0)$ are respectively given by

$$ka_k^{-4}\left(\gamma_{k,X^2}(0)-\gamma_{X^2}(0)\right)\xrightarrow{d}V_0^k\tag{6.37}$$

$$(n-k)a_{n-k}^{-4}\left(\gamma_{n-k,X^2}(0)-\gamma_{X^2}(0)\right)\xrightarrow{d}V_0^{n-k}\tag{6.38}$$

Next we consider the $\{X_t^2\}$ sequence and establish the the convergence of $\gamma_{n,X^2}(1)$ by direct application of Lemmas 1 and 3 as follows:

$$na_n^{-4}\left(\gamma_{n,X^2}(1) - \gamma_{X^2}(1)\right) = a_n^{-4}\sum_{t=1}^{n}\left[X_t^2 X_{t+1}^2 - E\left(X_0^2 X_1^2\right)\right]$$

$$= a_n^{-4}\sum_{t=1}^{n}\left[X_t^2 \sigma_{t+1}^2\left(\varepsilon_{t+1}^2 - E(\varepsilon)\right)\right]$$

$$= a_n^{-4}\sum_{t=1}^{n}\left[X_t^2 \sigma_t^2\left(A_{t+1} - EA\right)\right]I_{\{\sigma_t > a_n\delta\}}$$

$$= a_n^{-4}\sum_{t=1}^{n}\left[\sigma_t^2 \sigma_{t+1}^2 I_{\{\sigma_t > a_n\delta\}}\right] - EA\, a_n^{-4}\sum_{t=1}^{n}\left[\sigma_t^4 I_{\{\sigma_t > a_n\delta\}}\right]$$

$$= T_{2,\delta,\sigma}\left(N_n^{(2)}\right) - EA\, T_{1,\delta,\sigma}\left(N_n^{(2)}\right)$$

$$\xrightarrow{d} T_{2,\delta,\sigma}\left(N^{(2)}\right) - EA\, T_{1,\delta,\sigma}\left(N^{(2)}\right)$$

$$\xrightarrow{d} V_1 \tag{6.39}$$

Thus we have that

$$na_n^{-4}\left(\gamma_{n,X^2}(1) - \gamma_{X^2}(1)\right) \xrightarrow{d} V_1 \tag{6.40}$$

By extending to arbitrary lags $h = 0, \cdots n$ the convergence of $\gamma_{n,X^2}(h)$ is given by

$$na_n^{-4}\left(\gamma_{n,X^2}(h) - \gamma_{X^2}(h)\right) \xrightarrow{d} V_h \tag{6.41}$$

Consequently the convergence of $\rho_{n,X^2}(h)$ is given by

$$na_n^{-4}\left(\rho_{n,X^2}(h) - \rho_{X^2}(h)\right) \xrightarrow{d} \gamma_{X^2}^{-1}(0)\left(V_h - \rho_{X^2}(h)V_0\right) \tag{6.42}$$

We have been able to examine the limiting behaviour of $\rho_{n,X^2}(h)$ for two cases. In the first case, when $\kappa \in (0,2)$, the variance of X_n is infinite and thus $\rho_{n,X^2}(h)$ has a random limit without any normalization. When $\kappa \in (2,4)$, the process has a finite variance but infinite fourth moment and $na_n^{-4}\left(\rho_{n,X^2}(h)\right)$ converges to an $\frac{\kappa}{2} - stable$ distribution. By Theorem 26 convergence of $\rho_{n,X^2}(h)$ implies that the sequence is bounded with

$|\rho_{n,X^2}(h)| \leq 1.$

We now examine the convergence of $c_{k,X^2}(h) - c_{n-k,X^2}(h)$. Consider equation (6.12), we can express $c_{k,X^2}(h) - c_{n-k,X^2}(h)$ as follows:

$$c_{k,X^2}(h) - c_{n-k,X^2}(h) = \frac{\rho_{k,X^2}(h) - \rho_{n-k,X^2}(h)}{\rho_{n,X^2}(h)} \qquad (6.43)$$

By the Bolzano-Weierstrass Theorem 21, a bounded sequence has always a convergent subsequence. This is further confirmed through the invariance property of subsequences in Theorem 22 which states that if $\rho_{n,X^2}(h)$ converges, then every subsequence say, $\rho_{k,X^2}(h)$ and $\rho_{n-k,X^2}(h)$ converges. By linearity rule of sequences as prescribed in Theorem 23, $\rho_{k,X^2}(h) - \rho_{n-k,X^2}(h)$ converges. This further implies that $\rho_{k,X^2}(h)$ and $\rho_{n-k,X^2}(h)$ are bounded since every convergent sequence is bounded. The subsequences $\rho_{k,X^2}(h)$ and $\rho_{n-k,X^2}(h)$ are also bounded with $|\rho_{k,X^2}(h)| \leq 1$ and $|\rho_{n-k,X^2}(h)| \leq 1$, thus their absolute difference is also bounded as $|\rho_{k,X^2}(h) - \rho_{n-k,X^2}(h)| \leq 2$. Further assume that we are considering only significant sample autocorrelation coefficients where $|\rho_{n,X^2}(h)| \geq 0.05$, then $c_{k,X^2}(h) - c_{n-k,X^2}(h)$ is also bounded. Applying the quotient property of subsequences in Theorem 23, then $c_{k,X^2}(h) - c_{n-k,X^2}(h)$ is also convergent.

Consider the change-point process $D_n^k(h)$ as defined in (4.15), then we can derive the limit of C_n as follows:

$$\tag{6.44}$$

$$= \left(D_n^k(h) \right)_{h=1,\cdots,m}$$

$$= \rho_{k,X^2}(h) - \rho_{n-k,X^2}(h)$$

$$= \frac{\gamma_{k,X^2}(h)}{\gamma_{k,X^2}(0)} - \frac{\gamma_{n-k,X^2}(h)}{\gamma_{n-k,X^2}(0)}$$

$$= \frac{\sum\limits_{t=1}^{k} X_t^2 X_{t+h}^2}{\sum\limits_{t=1}^{k} X_t^4} - \frac{\sum\limits_{t=k+1}^{n} X_t^2 X_{t+h}^2}{\sum\limits_{t=k+1}^{n} X_t^4}$$

$$= \left(\frac{\sum\limits_{t=1}^{n} X_t^2 X_{t+h}^2 - \sum\limits_{t=k+1}^{n} X_t^2 X_{t+h}^2}{\sum\limits_{t=1}^{n} X_t^4 - \sum\limits_{t=k+1}^{n} X_t^4} \right) - \frac{\sum\limits_{t=k+1}^{n} X_t^2 X_{t+h}^2}{\sum\limits_{t=k+1}^{n} X_t^4}$$

$$= \frac{\sum\limits_{t=k+1}^{n} X_t^4 \left(\sum\limits_{t=1}^{n} X_t^2 X_{t+h}^2 - \sum\limits_{t=k+1}^{n} X_t^2 X_{t+h}^2 \right) - \sum\limits_{t=k+1}^{n} X_t^2 X_{t+h}^2 \left(\sum\limits_{t=1}^{n} X_t^4 - \sum\limits_{t=k+1}^{n} X_t^4 \right)}{\sum\limits_{t=k+1}^{n} X_t^4 \left(\sum\limits_{t=1}^{n} X_t^4 - \sum\limits_{t=k+1}^{n} X_t^4 \right)}$$

$$= \frac{\sum\limits_{t=k+1}^{n} X_t^4 \sum\limits_{t=1}^{n} X_t^2 X_{t+h}^2 - \sum\limits_{t=k+1}^{n} X_t^2 X_{t+h}^2 \sum\limits_{t=1}^{n} X_t^4}{\sum\limits_{t=k+1}^{n} X_t^4 \left(\sum\limits_{t=1}^{n} X_t^4 - \sum\limits_{t=k+1}^{n} X_t^4 \right)} \tag{6.45}$$

Thus applying Theorem 17 and Theorem 25 we have

$$\left(D_n^k(h) \right)_{h=1,\cdots,m} \xrightarrow{d} \frac{V_0^k V_h - V_h^k V_0}{V_0^k \left(V_0 - V_0^k \right)}$$

$$= \frac{V_0}{V_h} \left(\frac{V_0^k V_h - V_h^k V_0}{V_0^k \left(V_0 - V_0^k \right)} \right) \frac{V_h}{V_0} \tag{6.46}$$

From (6.46) above, then the sequence C_n converges in distribution to C_h as follows

$$C_n \overset{d}{\to} \frac{V_0}{V_h}\left(\frac{V_0^k V_h - V_h^k V_0}{V_0^k\left(V_0 - V_0^k\right)}\right) = C_h \tag{6.47}$$

By application of Continous Mapping Theorem 24, we have the limiting behaviour of the change-point process $D_n^k(h)$ for the two cases $\kappa \in (0,2)$ and $\kappa \in (2,4)$ as follows. for $\kappa \in (0,2)$ and by application of Theorem 16 (i) :

$$\begin{aligned}\left(D_n^k(h)\right)_{h=1,\cdots,m} &= \left(C_n(h)\,\rho_{n,X^2}(h)\right)_{h=1,\cdots,m}\\ &\overset{d}{\to} C_h\left(\frac{V_h}{V_0}\right)_{h=1.\cdots,m}\end{aligned} \tag{6.48}$$

for $\kappa \in (2,4)$ and by application of Theorem 16 (ii):

$$\begin{aligned}\left(na_n^{-4}\left(D_n^k(h) - D^{k^*}(h)\right)\right)_{h=1,\cdots,m} &= na_n^{-4}\left(C_n\rho_{n,X^2}(h) - \rho_{X^2}(h)\right)_{h=1,\cdots,n}\\ &\overset{d}{\to} \gamma_X^{-1}(0)\left(C_h V_h - \rho_{X^2}(h)C_0 V_0\right)_{h=1,\cdots,m}\end{aligned} \tag{6.49}$$

which completes proof.

Chapter 7

NUMERICAL RESULTS

7.1 Comparison of the Autocorrelation Structure

In this chapter simulations studies are carried out to confirm theoretical results obtained in chapters 4 to 6 with models generated according to Appendix B. We begin by examining the autocorrelation structure of the different simulated series.
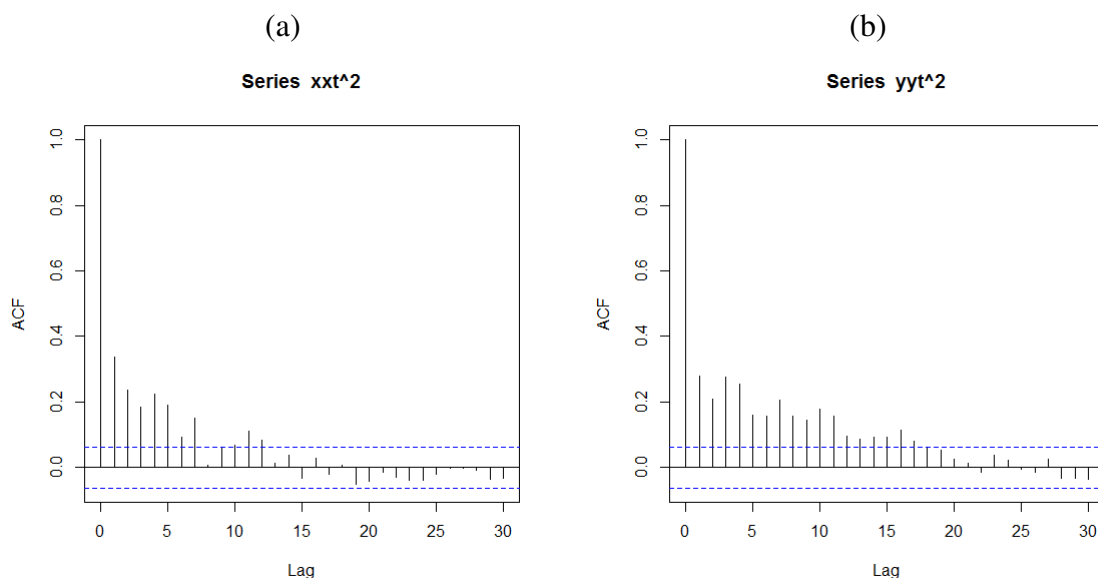
(a)   (b)

Figure 7.1: Comparison of ACF for simulated $GARCH\,(1,1)$ and $GARCH\,(1,2)$

In Figure 7.1 we examine graphically the effect in change of the autocorrelation struc-

ture when the change occurs in model order q. Specifically, Figure 7.1 (a) and (b) compare the autocorrelation of simulated series for $GARCH(1,1)$ and $GARCH(1,2)$ respectively. From the plot, it is clear that the autocorrelation structure of $GARCH(1,1)$ and that of $GARCH(1,2)$ are not similar.

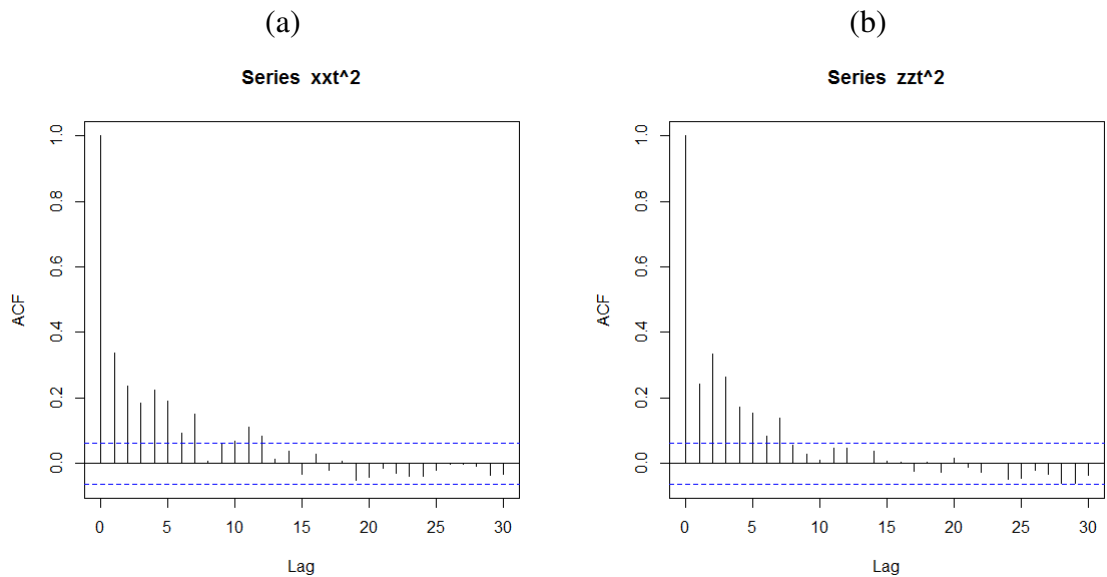(a)   (b)

Figure 7.2: Comparison of ACF for simulated $GARCH(1,1)$ and $GARCH(2,1)$

In Figure 7.2 we examine graphically the effect in change of the autocorrelation structure when the change occurs in model order p. Specifically, Figure 7.2 (a) and (b) compare the autocorrelation of simulated series for $GARCH(1,1)$ and $GARCH(2,1)$ respectively. From the plot, it is clear that the autocorrelation structure of $GARCH(1,1)$ and that of $GARCH(2,1)$ are not similar.
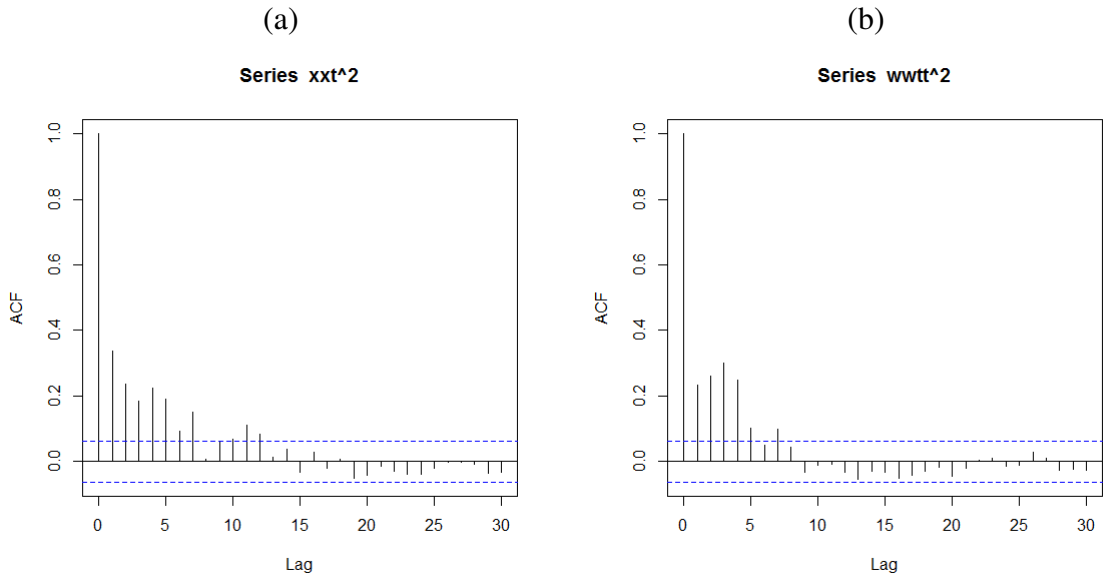
Figure 7.3: Comparison of ACF for simulated $GARCH(1,1)$ and $GARCH(3,3)$

In Figure 7.3 we examine graphically the effect in change of the autocorrelation structure the change occurs in both model order p and q. Specifically, Figure 7.3 (a) and (b) compare the autocorrelation of simulated series for $GARCH(1,1)$ and $GARCH(3,3)$ respectively. From the plot, it is clear that the autocorrelation structure of $GARCH(1,1)$ and that of $GARCH(3,3)$ are not similar.
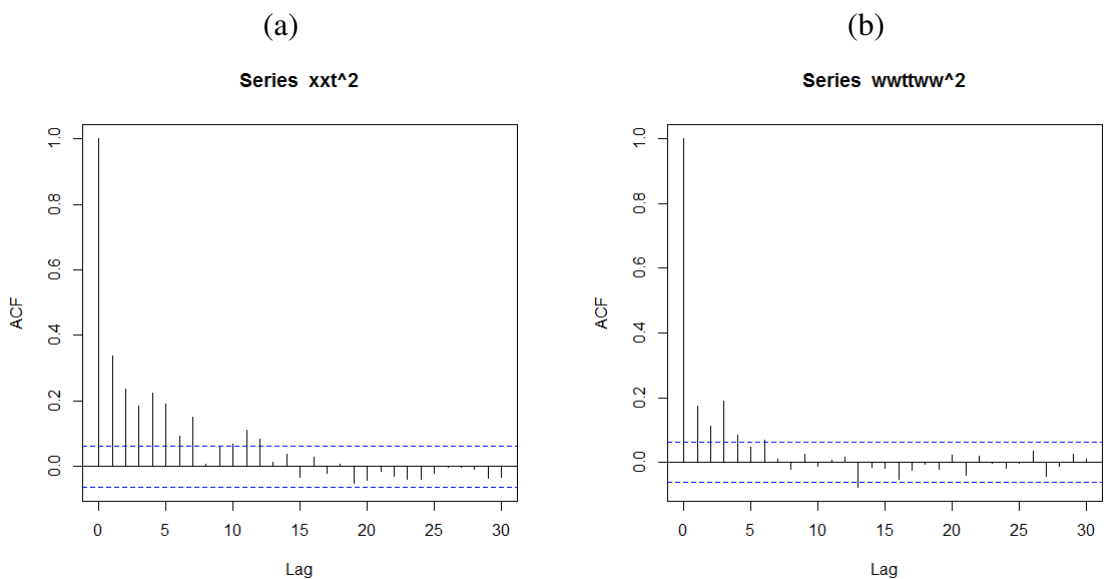

Figure 7.4: Comparison of ACF for simulated $GARCH(1,1)$ and $GARCH(4,4)$

In Figure 7.4 we examine graphically the effect in change of the autocorrelation structure when the change occurs in both model order in p and q and the size of change is increased to three. Specifically, Figure 7.4 (a) and (b) compare the autocorrelation of simulated series for $GARCH(1,1)$ and $GARCH(4,4)$ respectively. From the plot, it is clear that the autocorrelation structure of $GARCH(1,1)$ and that of $GARCH(4,4)$ are not similar even when the size of change is increased.

The Tables 7.1, 7.2 and 7.3 below reinforce the results in Figures 7.1-7.4 by comparing the autocorrelation coefficients across different lags.

Table 7.1: Comparison of ACF across different lags for change in (p)

Lag	$GARCH(1,1)$	$GARCH(2,1)$	$GARCH(4,1)$	$GARCH(1,1)-(2,1)$	$GARCH(1,1)-(4,1)$
0	1.00	1.00	1.00	0.00	0.00
1	0.34	0.24	0.16	0.10	0.18
2	0.24	0.33	0.06	0.09	0.18
3	0.18	0.26	0.04	0.08	0.14
4	0.23	0.17	0.19	0.06	0.04
5	0.19	0.15	-0.04	0.04	0.23

In Table 7.1 the autocorrelation coefficients for $GARCH(2,1)$ and $GARCH(4,1)$ are compared to the autocorrelation for $GARCH(1,1)$ for the first lag. It should be noted that the autocorrelation coefficients for $GARCH(1,1)$, $GARCH(2,1)$ and $GARCH(4,1)$ are largest at lags 1, 2 and 4 respectively. This is as expected since the model order p of a $GARCH(p,q)$ model is determined from the autocorrelation function at the point at which the function cuts off. This means that comparing the autocorrelation coefficients for the different models at specific lags is likely, in most cases, to result into a non-zero absolute difference as shown in column 5 and 6 of Table 7.1.

Table 7.2: Comparison of ACF across different lags for chage in (q)

Lag	$GARCH(1,1)$	$GARCH(1,2)$	$GARCH(1,4)$	$GARCH(1,1)-(1,2)$	$GARCH(1,1)-(1,4)$
0	1.00	1.00	1.00	0.00	0.00
1	0.34	0.28	0.20	0.06	0.14
2	0.24	0.21	0.13	0.03	0.11
3	0.18	0.28	0.08	0.10	0.10
4	0.23	0.25	0.16	0.02	0.07
5	0.19	0.16	0.14	0.03	0.05

In Table 7.2 the autocorrelation coefficients for $GARCH(1,2)$ and $GARCH(1,4)$ are compared to the autocorrelation for $GARCH(1,1)$ for the first lag. It should be noted that the autocorrelation coefficients for $GARCH(1,1)$, $GARCH(1,2)$ and $GARCH(1,4)$ are all largest at lags 1. This is as expected since the model order p of a $GARCH(p,q)$ model is determined from the autocorrelation function at the point at which the function cuts off, whereas the model order q of a $GARCH(p,q)$ model is determined from the partial autocorrelation function at the point at which the function cuts off. However, from the results in Table 7.2 it is observed that the autocorrelation coefficients of the various models at specific lags is different. This implies that comparing the autocorrelation coefficients for the different models at specific lags is likely to result into a non-zero absolute difference as shown in column 5 and 6 of Table 7.2.

Table 7.3: Comparison of ACF across different lags for chage in (p,q)

Lag	$GARCH(1,1)$	$GARCH(3,3)$	$GARCH(4,4)$	$GARCH(1,1)-(3,3)$	$GARCH(1,1)-(4,4)$
0	1.00	1.00	1.00	0.00	0.00
1	0.34	0.23	0.18	0.11	0.16
2	0.24	0.26	0.11	0.02	0.13
3	0.18	0.30	0.19	0.12	0.01
4	0.23	0.25	0.08	0.02	0.15
5	0.19	0.10	0.05	0.09	0.14

In Table 7.3 the autocorrelation coefficients for $GARCH(3,3)$ and $GARCH(4,4)$ are compared to the autocorrelation for $GARCH(1,1)$ for the first lag. It should be noted that the autocorrelation coefficients for $GARCH(1,1)$, $GARCH(3,3)$ and $GARCH(4,4)$ are largest at lags 1, 3 and 4 respectively. This is as expected since the model order p of a $GARCH(p,q)$ model is determined from the autocorrelation function at the point at which the function cuts off. This means that comparing the autocorrelation coefficients for the different models at specific lags is likely, in most cases, to result into a non-zero absolute difference as shown in column 5 and 6 of Table 7.3.

Following the results in Figures 7.1-7.4 and Tables 7.1-7.3 the change-point estimator $\hat{k}$ as defined in (4.16) can be applied to the simulated series such that the true

change-point k^* is fixed at $\frac{n}{2}$ for $n = 500$, $n = 1000$ and $n = 2000$. Assume that $\{X_t\}$ is a stationary $GARCH(p,q)$ process where $p,q \in \mathbb{N} \setminus \{0\}$ such that we are examining change-point hypothesis (3.26) in Section 3.2. The performance of the proposed change-point estimator $\hat{k}$ as given in (4.16) is examined by considering the effects of the location of change points and the size of change in the model order specification.

7.2 Performance of the Change-Point Estimator

The Adjusted Rand Indices are utilized to compare the segmentation created by the proposed change-point estimator $\hat{k}$ under H_1 against the null hypothesis of no change-point under H_0. The effect of the sample size and size of change to the estimator is examined by specifying different sample sizes and different model orders (p,q) in the alternative hypothesis. The performance of the proposed test is compared to that of reviewed method, E-Divisive. The results of the simulations are provided in Tables 7.4, 7.5 and 7.6. Tables 7.4 and 7.5 provide the results for simulations with changes in order p and q respectively while Table 7.6 provides the results for the changes in both order p and q. In these tables, average Adjusted Rand Index is reported for 1000 simulations.

Table 7.4: Fowlkes and Mallows ARI given changes in order p

		Under H_0		Under H_1	
n	$\triangle q$	Proposed Estimator $\hat{k}$	E-Divisive	Proposed Estimator $\hat{k}$	E-Divisive
500	1	1.00	1.00	0.9895613	0.9887145
	2	1.00	1.00	0.9944756	0.9966374
	3	1.00	1.00	0.9944756	0.9993042
1000	1	1.00	1.00	0.9914832	0.9979477
	2	1.00	1.00	0.9885897	0.9987957
	3	1.00	1.00	0.9115919	0.9989981
2000	1	1.00	1.00	0.9971880	0.9758699
	2	1.00	1.00	0.9957313	0.9767283
	3	1.00	1.00	0.9869408	0.9798594

The results of the perfomance analysis in Table 7.4 show that the Fowlkes and Mallows

ARI yields a value of 1 under the null hypothesis H_0 of no change-point as specified in (3.26) for all sample sizes and size of change, implying that there is a perfect similarity. On the other hand, under the alternative hypothesis H_1 as specified in (3.26) with change occuring in model order p, it is observed that as the sample size increases, the similarity index given by Fowlkes and Mallows ARI generally increases for the proposed change-point estimator. However, the index based on the E-Divisive test seems to yield higher ARI values when the sample is 1000 but this value reduces when the sample is increased to 2000. This is indicates that testing for change in order p of the GARCH (p,q) model using the proposed estimator is most suitable for long horizons whereas the edivisive method is suitable for shorter horizons . In addition, the ARI increases as the size of change in order p increases across the different sample sizes. Thus we can conclude that the proposed test performs better when the size of change is greater.

Table 7.5: Fowlkes and Mallows ARI given changes in order q

		Under H_0		Under H_1	
n	$\triangle p$	Proposed Estimator $\hat{k}$	E-Divisive	Proposed Estimator $\hat{k}$	E-Divisive
500	1	1.00	1.00	0.9944461	0.9974525
	2	1.00	1.00	0.9944461	0.9988279
	3	1.00	1.00	0.9945336	0.999690
1000	1	1.00	1.00	0.9882955	0.9918479
	2	1.00	1.00	0.9884021	0.9929368
	3	1.00	1.00	0.9937106	0.9953233
2000	1	1.00	1.00	0.9885593	0.970108
	2	1.00	1.00	0.9938778	0.9747159
	3	1.00	1.00	0.9963049	0.9765529

The results of the perfomance analysis in Table 7.5 show that the Fowlkes and Mallows ARI yields a value of 1 under the null hypothesis H_0 of no change-point as specified in (3.26) for all sample sizes and size of change, implying that there is a perfect similarity. On the other hand, under the alternative hypothesis H_1 as specified in (3.26) with change occuring in model order q, it is observed that as the sample size increases, the similarity index given by Fowlkes and Mallows ARI generally decreases for both

the proposed change-point estimatorand the E-Divisive method. This is indicates that testing for change in order q of the GARCH (p,q) model is most suitable for shorter horizons. In addition, the ARI increases as the size of change in order q increases across the different sample sizes. Thus the proposed test performs better when the size of change is greater.

Table 7.6: Fowlkes and Mallows ARI given changes in order p, q

n	$\triangle(p,q)$	Under H_0		Under H_1	
		Proposed Estimator $\hat{k}$	E-Divisive	Proposed Estimator $\hat{k}$	E-Divisive
500	1	1.00	1.00	0.9944163	0.972161
	2	1.00	1.00	0.9731032	0.9944163
	3	1.00	1.00	0.9945047	0.9983192
1000	1	1.00	1.00	0.9562616	0.9940223
	2	1.00	1.00	0.9937779	0.9941439
	3	1.00	1.00	0.9941203	0.9950048
2000	1	1.00	1.00	0.9942376	0.9607149
	2	1.00	1.00	0.9913397	0.9726061
	3	1.00	1.00	0.9996214	0.9785945

The results of the perfomance analysis in Table 7.6 show that the Fowlkes and Mallows ARI yields a value of 1 under the null hypothesis H_0 of no change-point as specified in (3.26) for all sample sizes and size of change, implying that there is a perfect similarity. On the other hand, under the alternative hypothesis H_1 as specified in (3.26) with change occuring in both model orders p and q, it is observed that as the sample size increases, the similarity index given by Fowlkes and Mallows ARI generally increases for the proposed change-point estimator. However, the index based on the E-Divisive test seems to yield higher ARI values when the sample is 1000 but this value reduces when the sample is increased to 2000. This is indicates that testing for change in order p and q of the GARCH (p,q) model using the proposed estimator is most suitable for long horizons whereas the edivisive method is suitable for shorter horizons . In addition, the ARI increases as the size of change in order p increases across the different sample sizes. Thus we can conclude that the proposed test performs better when the size of change is greater.

It is therefore concluded that the proposed change-point estimator is able to perform nonparametric change point analysis of a financial returns series. The test is able to determine the location of the change-point without prior specification of the number of change-points. The only necessary assumption made is that financial returns data follow stationary GARCH (p,q) series with different orders.

7.3   Single Change-Point Estimation

The Figure below displays the plots for the simulated series with a superimposed location of change-point estimator $\hat{k}$ and the corresponding change-point process D_n^k as defined in (4.15) and change-point location for various sample sizes.

(a)   (b)
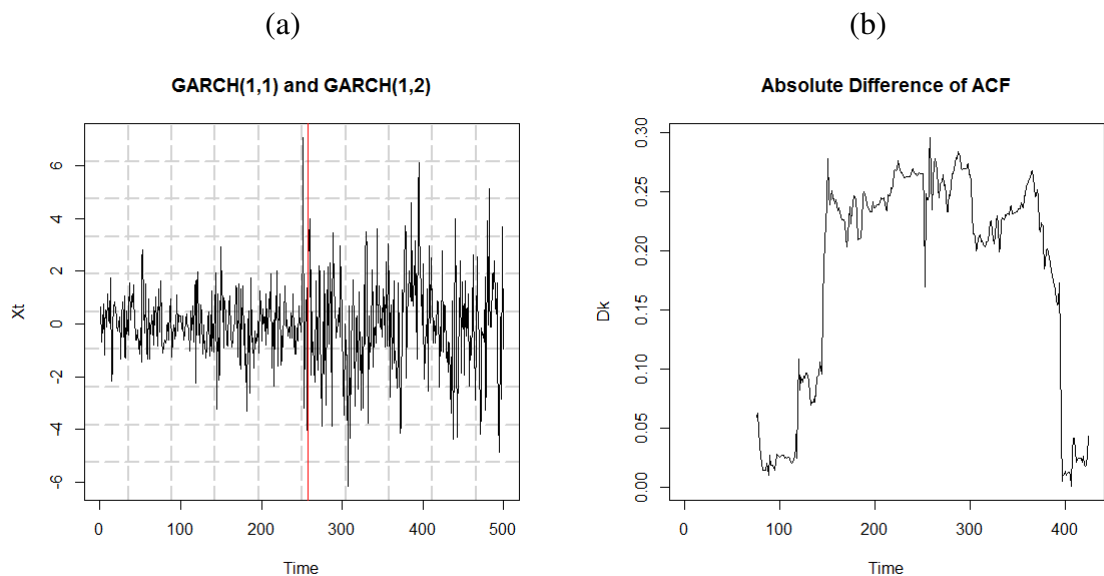
Figure 7.5: Single Change-Point for Stationary Series GARCH X_t

Figure 7.5 displays results of simulated series comprised of $GARCH\,(1,1)$ for time $t \in [1, 250]$ and $GARCH\,(1,2)$ for time $t \in [251, 500]$ with a sample size $n = 500$ where the true change-point k^* is positioned at time $t = 250$. The estimator $\hat{k}$ is able to locate the change-point at $k = 256$ as shown in Figure 7.5(a) with 99.44% similarity compared to a threshold of 100% similarity index as displayed in the ARI Table 7.5. The resultant

change-point process D_n^k is as plotted in Figure 7.5(b) which seems to be random in nature.
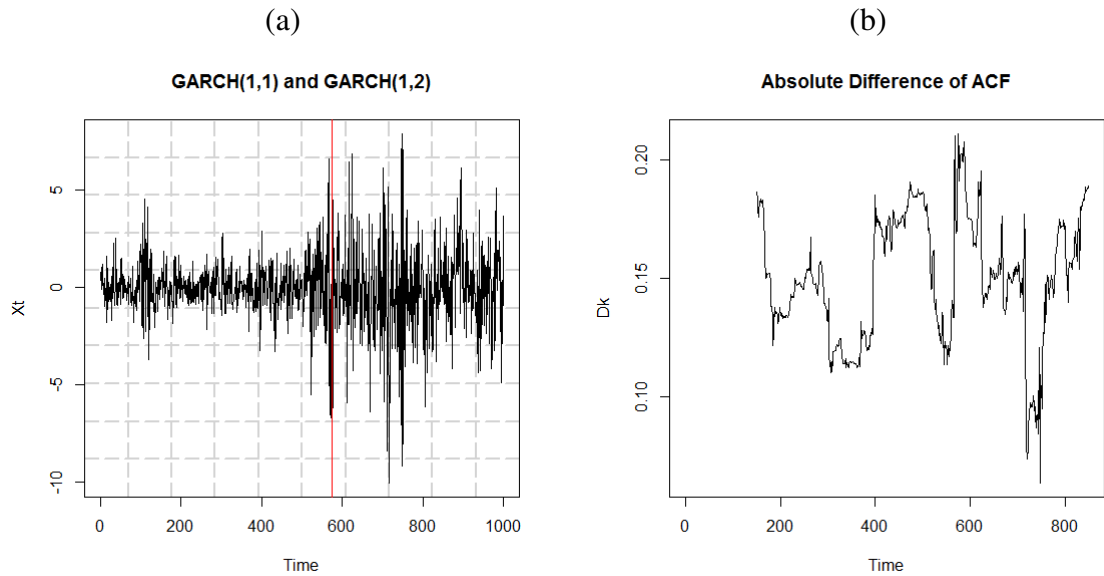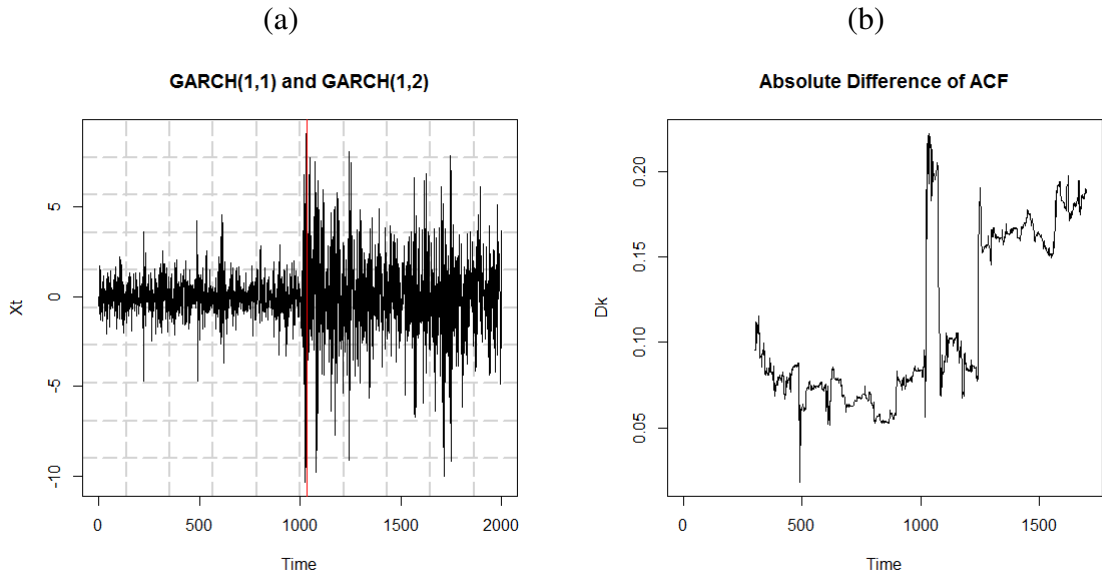

Figure 7.6: Single Change-Point for Stationary Series GARCH X_t

Figure 7.6 displays results of simulated series comprised of $GARCH\,(1,1)$ for time $t \in [1,500]$ and $GARCH\,(1,2)$ for time $t \in [501,1000]$ with a sample size $n = 1000$ where the true change-point k^* is positioned at time $t = 500$. The estimator $\hat{k}$ is able to locate the change-point at $k = 580$ as shown in Figure 7.6(a) with 98.83% similarity compared to a threshold of 100% similarity index as displayed in the ARI Table 7.5. The resultant change-point process D_n^k is as plotted in Figure 7.5(b) which seems to be random in nature.

Figure 7.7: Single Change-Point for Stationary Series GARCH X_t

Figure 7.7 displays results of simulated series comprised of $GARCH\,(1,1)$ for time $t \in [1,1000]$ and $GARCH\,(1,2)$ for time $t \in [1001,2000]$ with a sample size $n = 2000$ where the true change-point k^* is positioned at time $t = 1000$. The estimator $\hat{k}$ is able to locate the change-point at $k = 1016$ as shown in Figure 7.7(a) with 98.86% similarity compared to a threshold of 100% similarity index as displayed in the ARI Table 7.5. The resultant change-point process D_n^k is as plotted in Figure 7.5(b) which seems to be random in nature.

We now examine the change-point estimator $\hat{k}$ as defined in (4.16) when the true change-point k^* is no longer positioned at the middle of the sample size but is rather fixed at $\frac{n}{4}$ and $\frac{3n}{4}$ for $n = 500$, $n = 1000$ and $n = 2000$.
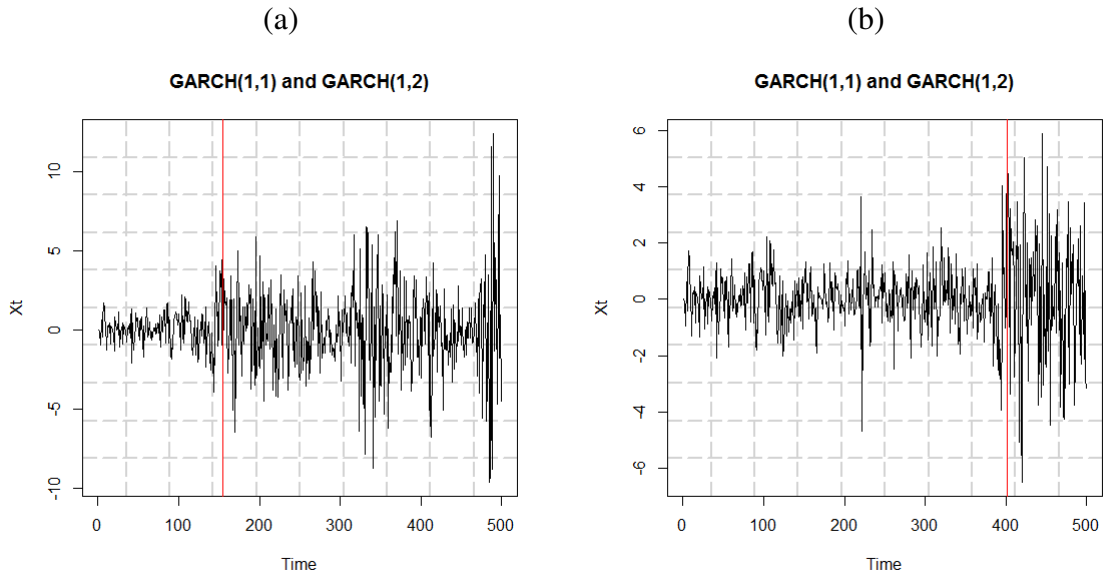
<table>
<tr><td>(a)</td><td>(b)</td></tr>
<tr><td>GARCH(1,1) and GARCH(1,2)</td><td>GARCH(1,1) and GARCH(1,2)</td></tr>
</table>

Figure 7.8: Single Change-Point for Stationary Series GARCH X_t

Figure 7.8 (a) displays results of simulated series comprised of *GARCH* $(1,1)$ for time $t \in [1,125]$ and *GARCH* $(1,2)$ for time $t \in [126,500]$ whereas Figure 7.8 (b) displays results of simulated series comprised of *GARCH* $(1,1)$ for time $t \in [1,375]$ and *GARCH* $(1,2)$ for time $t \in [376,500]$ each with a sample size $n = 500$. The true change-point k^* is positioned at time $t = 125$ and $t = 125$ for Figure 7.8 (a) and Figure 7.8 (b) respectively. The estimator $\hat{k}$ is able to locate the change-point at $k = 155$ as shown in Figure 7.8(a) and at $k = 402$ as shown in Figure 7.8(b) with 92.99% and 98.98% similarity index respectively compared to a threshold of 100% similarity index. For the sample size $n = 500$, the estimator $\hat{k}$ seems not to accurately estimate the true change-point due to the large deviation from k^*.
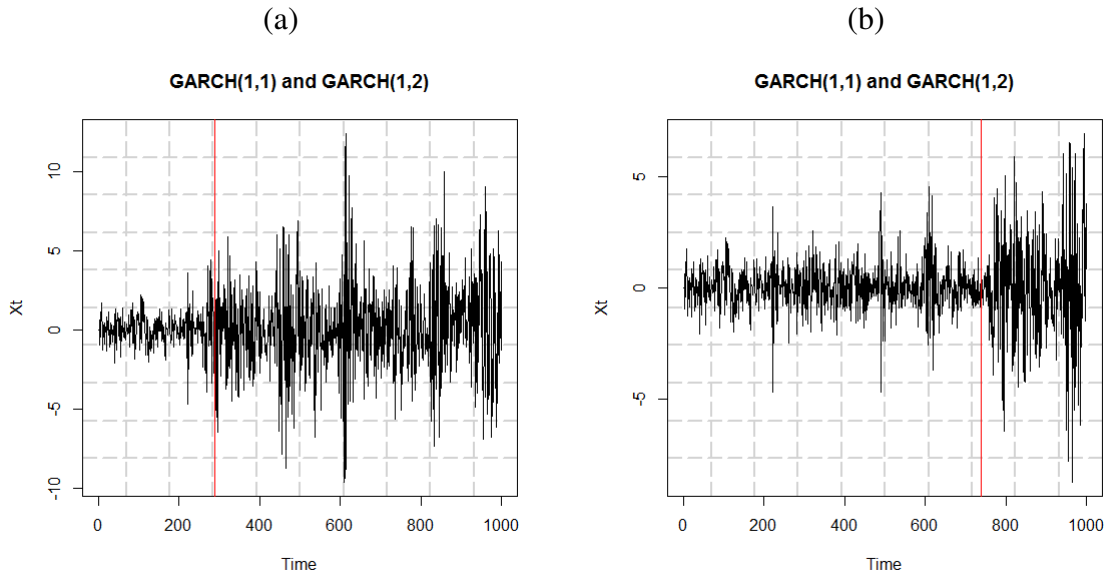
Figure 7.9: Single Change-Point for Stationary Series GARCH X_t

Figure 7.9 (a) displays results of simulated series comprised of $GARCH\,(1,1)$ for time $t \in [1,250]$ and $GARCH\,(1,2)$ for time $t \in [251,1000]$ whereas Figure 7.9 (b) displays results of simulated series comprised of $GARCH\,(1,1)$ for time $t \in [1,750]$ and $GARCH\,(1,2)$ for time $t \in [751,1000]$ each with a sample size $n = 1000$. The true change-point k^* is positioned at time $t = 250$ and $t = 125$ for Figure 7.9 (a) and Figure 7.9 (b) respectively. The estimator $\hat{k}$ is able to locate the change-point at $k = 288$ as shown in Figure 7.9(a) and at $k = 739$ as shown in Figure 7.9(b) with 93.57% and 99.16% similarity index respectively compared to a threshold of 100% similarity index. For the sample size $n = 1000$, the estimator $\hat{k}$ performance seems to have slightly improved compared to the results obtained when $n = 500$ from the smaller deviation from k^*.
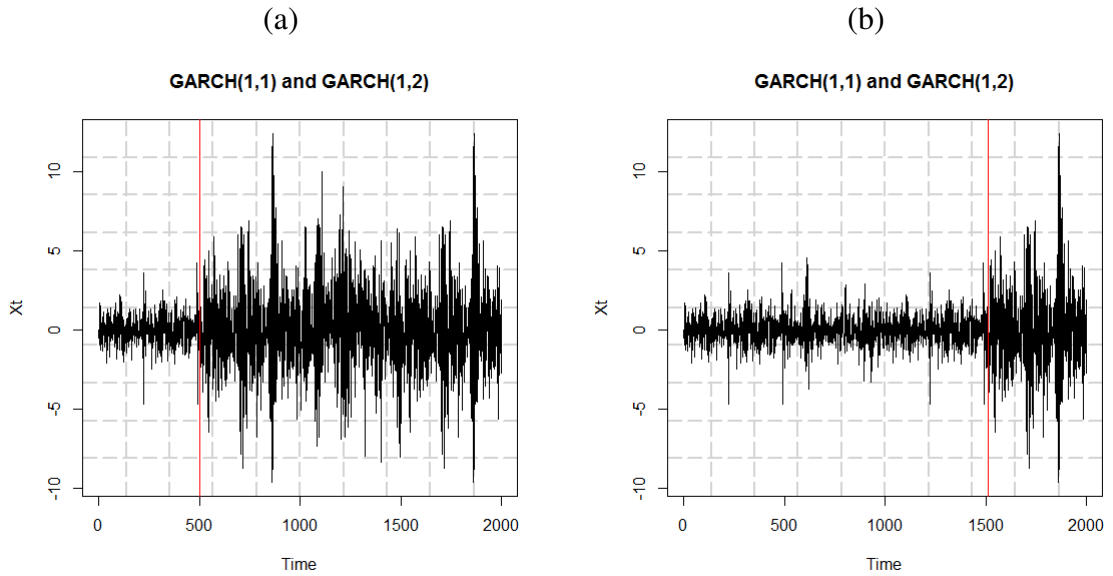
Figure 7.10: Single Change-Point for Stationary Series GARCH X_t

Figure 7.10 (a) displays results of simulated series comprised of $GARCH\,(1,1)$ for time $t \in [1,500]$ and $GARCH\,(1,2)$ for time $t \in [501,2000]$ whereas Figure 7.10 (b) displays results of simulated series comprised of $GARCH\,(1,1)$ for time $t \in [1,1500]$ and $GARCH\,(1,2)$ for time $t \in [1501,2000]$ each with a sample size $n = 2000$. The true change-point k^* is positioned at time $t = 500$ and $t = 1500$ for Figure 7.10 (a) and Figure 7.10 (b) respectively. The estimator $\hat{k}$ is able to locate the change-point at $k = 504$ as shown in Figure 7.10(a) and at $k = 1507$ as shown in Figure 7.10(b) with 99.20% and 99.29% similarity index respectively compared to a threshold of 100% similarity index. For the sample size $n = 2000$, the estimator $\hat{k}$ performance seems to have slightly improved compared to the results obtained when $n = 500$ and $n = 1000$ from the smaller deviation from k^*. This could be attributed to adequate data available in the resultant segments created.

7.4 Multiple Change-Point Estimation

We now consider the multiple change-points testing problem where the change-points k_1 and k_2 are fixed at intervals of $\frac{n}{3}$ for $n = 1500$ and $n = 3000$ and $(p,q) \neq (p^*, q^*)$. The figure below displays the plots for change-point location as estimated by the change-point estimator $\hat{k}$ as defined in (4.16) with application of the hierarchical/binary segmentation procedure for the various sample sizes as outlined of the literature review.

$$H_0 : X_t \sim GARCH(1,1) \qquad for \, t = 1, \cdots, n$$
$$against$$
$$H_1{:}X_t \sim GARCH(p,q) \qquad for \, t = 1, \cdots, k_1 \qquad (7.1)$$
$$X_t \sim GARCH(1,1) \qquad for \, t = k_1 + 1, \cdots, k_2$$
$$X_t \sim GARCH(p^*, q^*) \qquad for \, t = k_2 + 1, \cdots, n$$
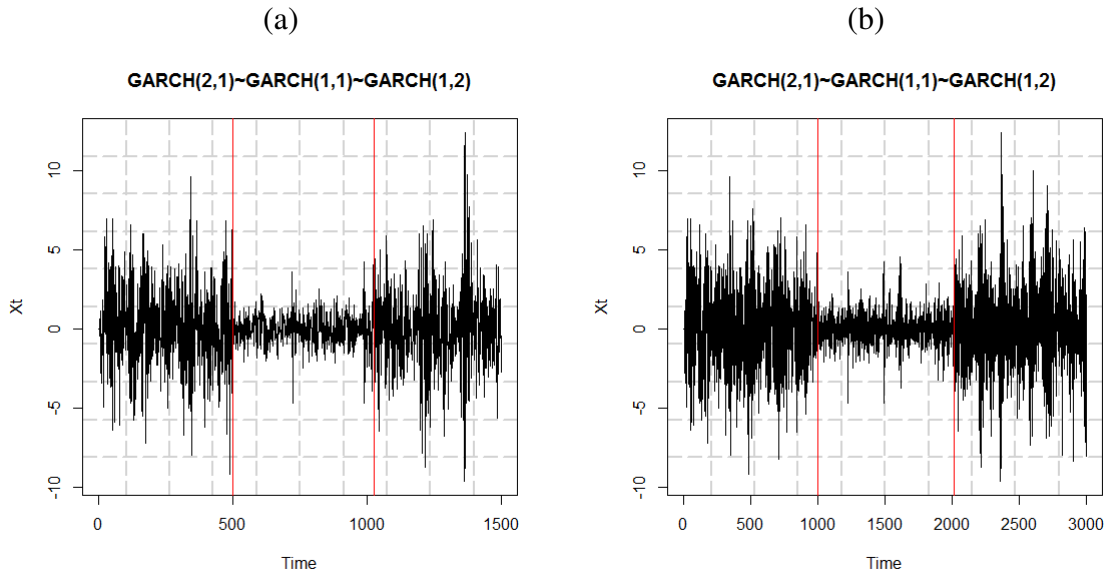
(a)   (b)

Figure 7.11: Multiple Changepoints for Stationary GARCH series X_t

Figure 7.11 (a) displays results of simulated series comprised of $GARCH(2,1)$ for time $t \in [1, 500]$, $GARCH(1,1)$ for time $t \in [501, 1000]$ and $GARCH(1,2)$ for time $t \in [1001, 1500]$ whereas Figure 7.11 (b) displays results of simulated series comprised

of $GARCH\,(2,1)$ for time $t \in [1,1000]$, $GARCH\,(1,1)$ for time $t \in [1001,2000]$ and $GARCH\,(1,2)$ for time $t \in [2001,3000]$ with a sample size of $n = 1500$ and $n = 3000$ respectively. The true change-points $k*$ are positioned at time $t = (500,100)$ and $t = (1000,2000)$ for Figure 7.11 (a) and Figure 7.11 (b) respectively. The estimator $\hat{k}$ is able to locate the change-point at $k = (501,1024)$ as shown in Figure 7.11(a) and at $k = (1002,2013)$ as shown in Figure 7.11(b). It can therefore be asserted that the estimator $\hat{k}$ works well for both identification of single change-points and multiple change-points as long we are considering change of a series from $GARCH\,(1,1)$ to any other $GARCH\,(p,q)$ model.

7.5 Distribution of Change-Point Estimator

We discuss the sampling distribution of the change-point estimator $\hat{k}$. Histograms are used to illustrate the inherent properties. Each histogram in Figures 7.12-7.14 describes the sample distribution of 1000 replications of $\hat{k}$ for several $GARCH(p,q)$ models with change-points as given in (3.26). Each of the estimated series has a sample size of $n = 1000$.

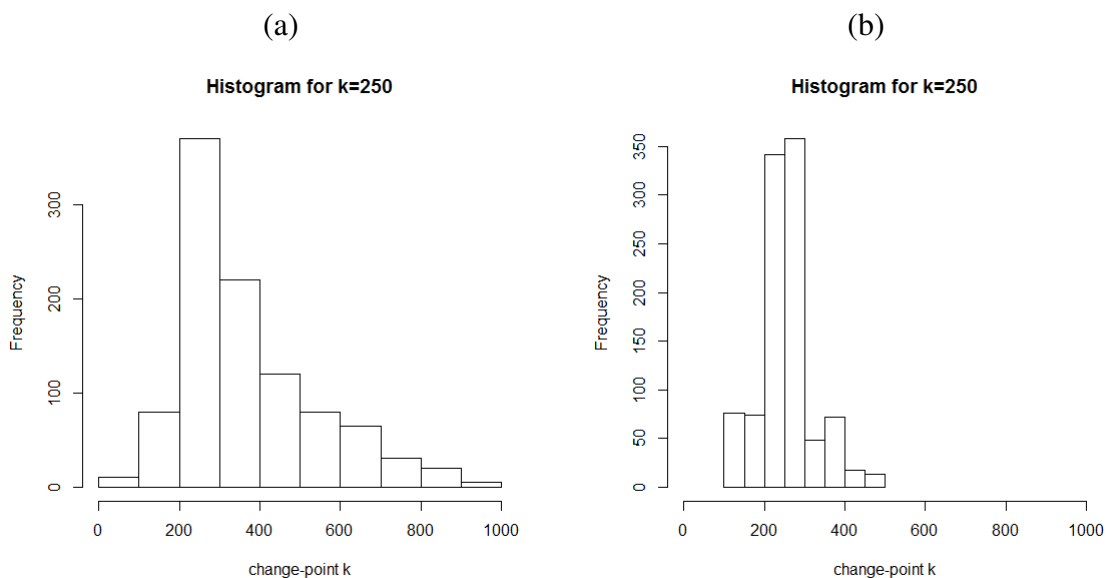(a)   (b)

Figure 7.12: Sampling Distribution of $\hat{k}$ for $k^* = 250$

Figure 7.12 displays results for the sampling distribution of the change-point estimator $\hat{k}$ when the true change-point is positioned at $k^* = 250$. Figure 7.12 (a) depends on the simulated series with a single change-point where it follows a $GARCH(1,1)$ for $t \in [1:250]$ and $GARCH(2,2)$ for $t \in [251:1000]$ whereas Figure 7.12 (b) depends on the simulated series with a single change-point where it follows a $GARCH(1,1)$ for $t \in [1:250]$ and $GARCH(4,4)$ for $t \in [251:1000]$. This represents a size of change of 1 and 3 im model orders (p,q) respectively. Figure 7.12 shows that the distribution is positively skewed as depicted by the longer tail to the right. However, comparing Figure 7.12 (a) and (b), it seems that the degree of skewness reduces as the size of model order change increases.
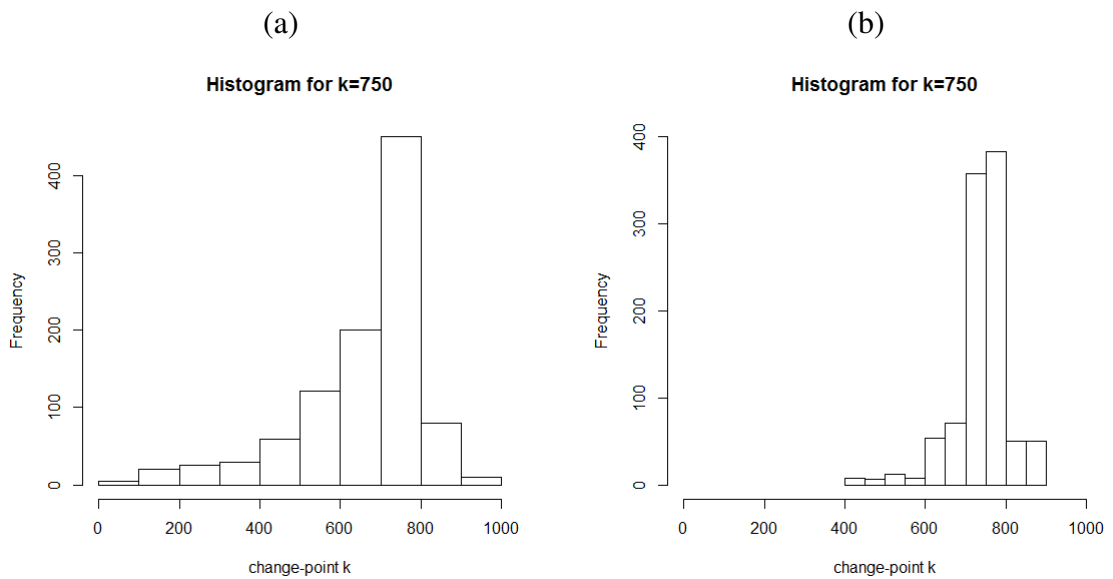
(a)   (b)

Figure 7.13: Sampling Distribution of $\hat{k}$ for $k^* = 750$

Figure 7.13 displays results for the sampling distribution of the change-point estimator $\hat{k}$ when the true change-point is positioned at $k^* = 750$. Figure 7.12 (a) depends on the simulated series with a single change-point where it follows a $GARCH(1,1)$ for $t \in [1:750]$ and $GARCH(2,2)$ for $t \in [751:1000]$ whereas Figure 7.13 (b) depends on the simulated series with a single change-point where it follows a $GARCH(1,1)$ for $t \in [1:750]$ and $GARCH(4,4)$ for $t \in [751:1000]$. This represents a size of change

of 1 and 3 im model order (p,q) respectively. Figure 7.12 shows that the distribution is negatively skewed as depicted by the longer tail to the left. However, comparing Figure 7.13 (a) and (b), it seems that the degree of skewness reduces as the size of model order change increases.
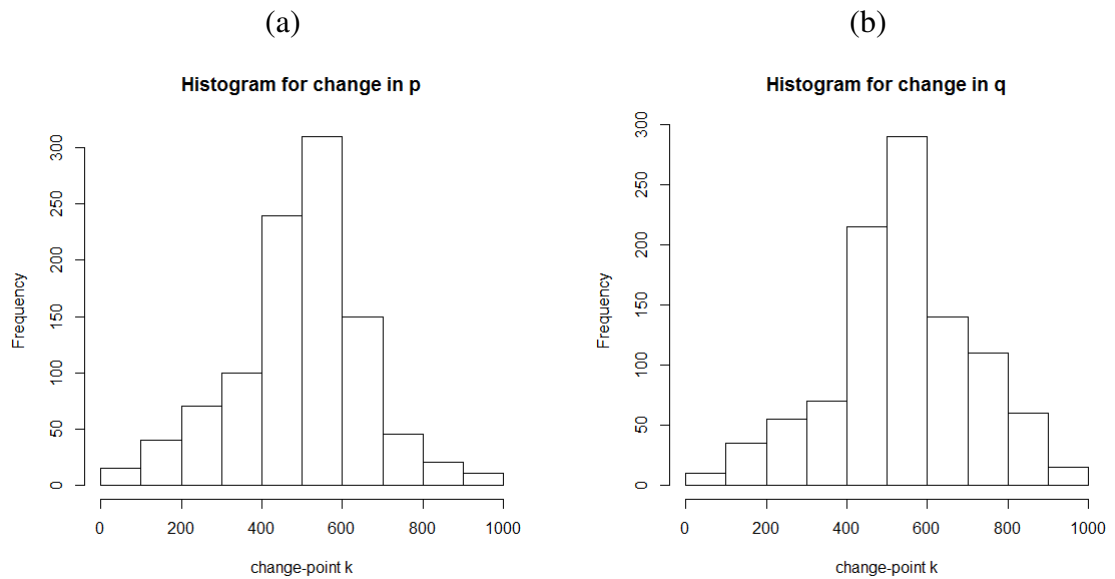
(a)   (b)

Figure 7.14: Sampling Distribution of $\hat{k}$ for $k^* = 500$ (a) $\triangle p = 1$ (b) $\triangle q = 1$

Figure 7.14 displays results for the sampling distribution of the change-point estimator $\hat{k}$ when the true change-point is positioned at $k^* = 500$. Figure 7.14 (a) depends on the simulated series with a single change-point where it follows a $GARCH(1,1)$ for $t \in [1:500]$ and $GARCH(2,1)$ for $t \in [501:1000]$ whereas Figure 7.14 (b) depends on the simulated series with a single change-point where it follows a $GARCH(1,1)$ for $t \in [1:500]$ and $GARCH(1,2)$ for $t \in [501:1000]$. Figure 7.14 shows that the distribution is approximately Normal as depicted by the equal distribution of tails to the left and to the right.

Chapter 8

APPLICATION OF THE CHANGE-POINT ESTIMATOR

A key principle in option pricing is the absence of arbitrage opportunity. To ensure that this holds we utilize the locally risk-neutral valuation relationship to determine the risk-neutral pricing measure Q. Duan. (1995) proposed the following conditional, lognormally distributed stock price process, with stochastic volatility, under the P measure.

$$S_t = S_{t-1} exp \left(r \times (T-t) - \frac{1}{2}\sigma_t^2 + \lambda \sigma_t + \sigma_t \varepsilon_t \right) \tag{8.1}$$

where $\varepsilon_t | \mathscr{F}_{t-1} \overset{P}{\sim} N(0,1)$ is the conditional error process λ is the unit risk premium, $\mathscr{F}_{t-1}$ is the $\sigma - algebra$ of information up to time t, $r = r^d - r^f$ is the yearly risk-free rate of return with r^d and r^f being the domestic and foreign risk-free rate respectively and σ_t^2 is the conditional variance GARCH(p,q) process defined as

$$\sigma_t^2 \quad = \quad \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{8.2}$$

A probability measure $\mathbb{Q}$ is said to be a local risk-neutral probability measure if

1. $\mathbb{Q}$ is equivalent to measure P

2. $E^{\mathbb{Q}}\left[ln\left(\frac{S_t}{S_{t-1}}\right)|\mathscr{F}_{t-1}\right] = r$

3. $Var^{\mathbb{Q}}\left[ln\left(\frac{S_t}{S_{t-1}}\right)|\mathscr{F}_{t-1}\right] = Var^{\mathbb{P}}\left[ln\left(\frac{S_t}{S_{t-1}}\right)|\mathscr{F}_{t-1}\right]$

Following this definition, the probability measure $\mathbb{Q}$ is determined from equation (8.1) as follows:

$$
\begin{aligned}
\frac{S_t}{S_{t-1}} &= e^{v_t + \sigma_t \xi_t} \\
E^{\mathbb{Q}}\left[\left(\frac{S_t}{S_{t-1}}\right)|\mathscr{F}_{t-1}\right] &= E^{\mathbb{Q}}\left[e^{v_t + \sigma_t \xi_t}|\mathscr{F}_{t-1}\right] \\
&= e^{v_t} E^{\mathbb{Q}}\left[e^{\sigma_t \xi_t}|\mathscr{F}_{t-1}\right] \\
&= e^{v_t + \frac{1}{2}Var^{\mathbb{Q}}\left[ln\left(\frac{S_t}{S_{t-1}}\right)|\mathscr{F}_{t-1}\right]} E^{\mathbb{Q}}\left[1|\mathscr{F}_{t-1}\right] \\
&\quad since\, Var^{\mathbb{Q}}\left[ln\left(\frac{S_t}{S_{t-1}}\right)|\mathscr{F}_{t-1}\right] = Var^{\mathbb{P}}\left[ln\left(\frac{S_t}{S_{t-1}}\right)|\mathscr{F}_{t-1}\right] \\
&= e^{v_t + \frac{1}{2}\sigma_t} \\
&= e^{r} \\
where \quad v_t &+ \frac{1}{2}\sigma_t = r \\
v_t &= r - \frac{1}{2}\sigma_t
\end{aligned}
\tag{8.3}
$$

The log-returns process with $GARCH(p,q)$ volatility under the P measure is given by

$$
ln\left(\frac{S_t}{S_{t-1}}\right) = r + \lambda \sigma_t - \frac{1}{2}\sigma_t^2 + \sigma_t \varepsilon_t
\tag{8.4}
$$

and the process implied by derivation from equation (8.3) above under measure $\mathbb{Q}$ is

$$
ln\left(\frac{S_t}{S_{t-1}}\right) = r - \frac{1}{2}\sigma_t^2 + \sigma_t \xi_t
\tag{8.5}
$$

using $Var^{\mathbb{Q}}\left[ln\left(\frac{S_t}{S_{t-1}}\right)|\mathscr{F}_{t-1}\right] = Var^{\mathbb{P}}\left[ln\left(\frac{S_t}{S_{t-1}}\right)|\mathscr{F}_{t-1}\right]$ it can be seen that

$$r + \lambda \sigma_t - \frac{1}{2}\sigma_t^2 + \sigma_t \varepsilon_t \quad = \quad r - \frac{1}{2}\sigma_t^2 + \sigma_t \xi_t \tag{8.6}$$

$$\Rightarrow \quad \varepsilon_t = \xi_t - \lambda \tag{8.7}$$

Substituting the result in equation (8.7) into equation (8.4) yields

$$ln\left(\frac{S_t}{S_{t-1}}\right) = r - \frac{1}{2}\sigma_t^2 + \sigma_t \xi_t \tag{8.8}$$

Substituting the result in equation (8.7) into the GARCH(p,q) process deined in equation (8.2) yields

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \sigma_{t-i}^2 (\xi_{t-i} - \lambda)^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{8.9}$$

Thus, under the $\mathbb{Q} - measure$, implied by the Locally Risk Neutral Valuation Relationship, the dynamics the log-returns processes is modelled by

$$R_t = ln\left(\frac{S_t}{S_{t-1}}\right) = r - \frac{1}{2}\sigma_t^2 + \sigma_t \xi_t \tag{8.10}$$

where $\xi_t | \mathscr{F}_{t-1} \overset{\mathbb{Q}}{\sim} N(0,1)$ and $\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \sigma_{t-i}^2 (\xi_{t-i} - \lambda)^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$ which is a $NGARCH(p,q)$ process.

The price of a call stock option can thus be determined using the formula

$$C_t \quad = \quad E^{\mathbb{Q}}\left[exp\left(-\sum_{i=t}^{T} r_t\right)(S_T - K)^+ | \mathscr{F}_{t-1}\right]$$

$$= \quad E^{\mathbb{Q}}\left[exp\left(-r(T-t)\right) max\left(S_T - K, 0\right) | \mathscr{F}_{t-1}\right] \tag{8.11}$$

Simulations are used to recursively apply the dynamics in (8.11) in order to generate values for S_T since there is no closed form solution. The option price formula obtained

by simulations is given by

$$
\begin{aligned}
C_t &= exp\left(-r\left(T-t\right)\right)\frac{1}{N}\sum_{i=t}^{N}\left[max\left(S_{T,i}-K,0\right)\right] \\
&\quad where \\
S_{T,i} &= S_0 exp\left(\sum_{j=1}^{T-t}R_{i,j}\right) \\
for &\quad i=1,2,\cdots,100,000\ simulations
\end{aligned}
\tag{8.12}
$$

More specifically, following the identified change-point k using the proposed estimator the log returns process R_t is modelled as a piecewise function as

$$
R_t = \begin{cases}
r-\frac{1}{2}(\sigma_t^{11})^2+\sigma_t^{11}\xi_t & for\ t\in[1,k] \\
r-\frac{1}{2}(\sigma_t^{12})^2+\sigma_t^{12}\xi_t & for\ t\in[k+1,T-t]
\end{cases}
$$

$$
\begin{aligned}
&\quad where \\
\left(\sigma_t^{11}\right)^2 &= \alpha_0+\alpha_1\left(\sigma_{t-1}^{11}\right)^2\left(\xi_{t-1}-\lambda\right)^2+\beta_1\left(\sigma_{t-1}^{11}\right)^2 \\
\left(\sigma_t^{12}\right)^2 &= \alpha_0^*+\alpha_1^*\left(\sigma_{t-1}^{12}\right)^2\left(\xi_{t-1}^*-\lambda^*\right)^2+\beta_1^*\left(\sigma_{t-1}^{12}\right)^2+\beta_2^*\left(\sigma_{t-2}^{12}\right)^2
\end{aligned}
\tag{8.13}
$$

The change-point estimator is applied to geometric returns of KES/USD Exchange rate data. The series is found to have multiple change-points where two change-points are estimated at $\hat{k}=400$ and $\hat{k}=754$ through binary segmentation as shown in the Figure 8.1 below.

Figure 8.1: Change-Point Detection on Geometric Returns Data
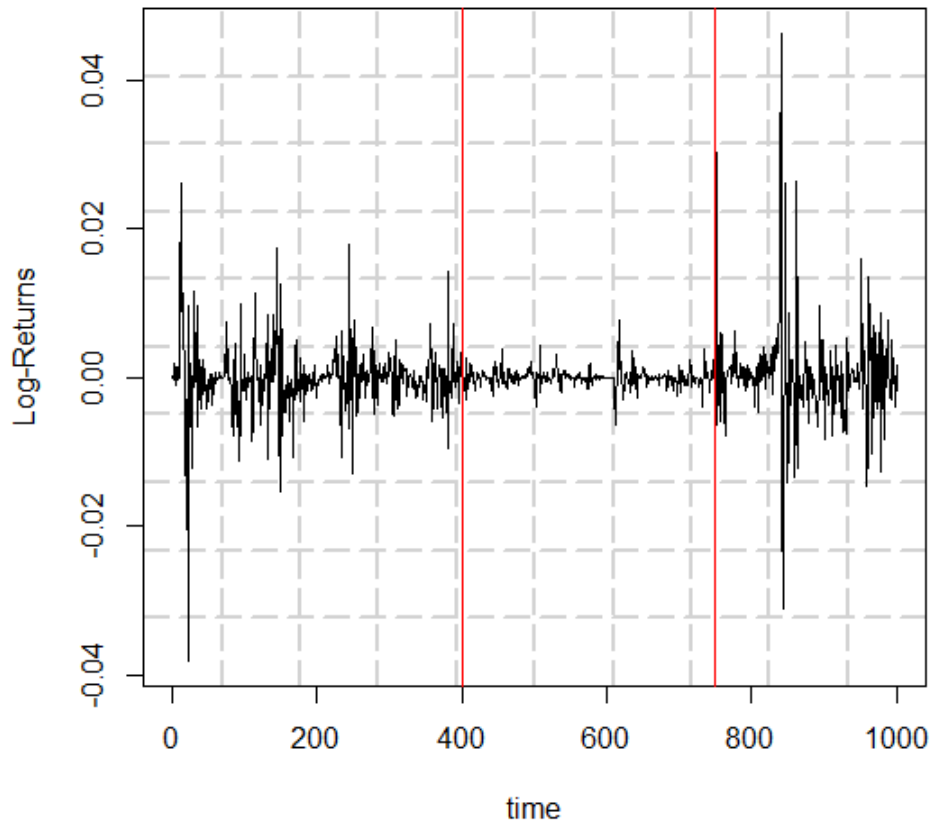
To facilitate the simulation of option prices according to model (8.12) various GARCH models are fitted to the geometric returns data displayed in Figure 8.1 with the incorporation of estimated change-points so as to determine the best model fits for (8.13).

Figure 8.2: Plot of Superimposed Geometric Returns Data

Plot of Superimposed Geometric Returns
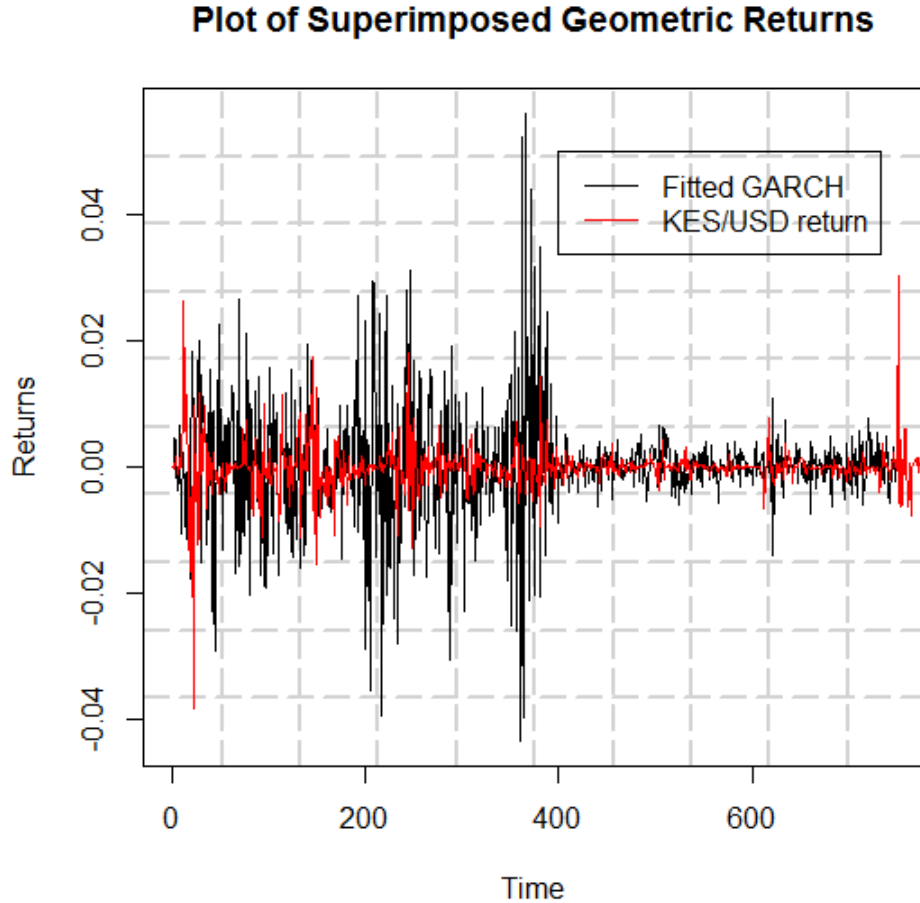
Following the identified change-points, $GARCH(1,2)$ and $GARCH(1,1)$ provided for the best fit for $t = [1, 400]$ and $t = [401, 754]$ respectively. Figure 8.2 displays a plot of the geometric returns of KES/USD Exchange rate data similar to 8.1 with a superimposed plot of the fitted $GARCH(1,2)$ and $GARCH(1,1)$. The ARI established a similarity index of 96.46% and 98.71% for $t = [1, 400]$ and $t = [401, 754]$ respectively when the fitted GARCH model and the KES/USD exchange rate return data were compared.

In line with the risk neutral probabilities, NGARCH models with these model specifications were fitted. The table below shows the values of estimated parameters after fitting $NGARCH(1,2)$ model for $t \in [1, 400]$ and $NGARCH(1,1)$ model for for

$t \in [401, 754]$.

Parameters	$NGARCH(1,2)$	Std Error	p-value	$NGARCH(1,1)$	Std Error	p-value
α_0	2.337×10^{-9}	3.655×10^{-7}	4.579×10^{-2}	1.106×10^{-7}	4.786×10^{-7}	2.08×10^{-2}
α_1	4.938×10^{-6}	1.396×10^{-1}	6.930×10^{-4}	5.636×10^{-6}	7.635×10^{-2}	6.59×10^{-6}
β_1	4.127×10^{-1}	7.393×10^{-2}	9.590×10^{-6}	7.844×10^{-1}	1.401×10^{-1}	2.16×10^{-7}
β_2	3.077×10^{-1}	5.978×10^{-2}	2.69×10^{-7}	-		
λ	-4.190×10^{1}	1.051×10^{-4}	4.38753×10^{-2}	-1.544×10^{-1}	1.017×10^{-4}	3.092×10^{-1}

Table 8.1: Parameter Estimates for $NGARCH(1,2) \sim NGARCH(1,1)$

The table below shows the values of estimated parameters after fitting $HN-GARCH(1,1)$ model for $t \in [1, 754]$.

Parameters	Heston-Nandi
α_0	7.296×10^{-13}
α_1	3.553×10^{-6}
β_1	7.694×10^{-1}
λ	-2.705×10^{1}
Log-Likelihood	183.57

Table 8.2: Parameter Estimates for Heston and Nandi model

The option prices obtained using equations (8.12) and (8.13) are compared with those obtained from Black-Scholes, Heston and Nandi GARCH and the NGARCH(1,1) models. Here we are considering an American option with an expiration date at $T = 754$ but with a possible early excersise considered to be $T = 400$. Under Black-Scholes, the pricing follows model described in model (3.78) with constant volatility for $t \in [1, 754]$ as outlined in section 3.6. The Heston and Nandi GARCH pricing models follow model (3.85) where volatility follows equation (3.84) for $t \in [1, 754]$ as outlined in section 3.6. The $NGARCH(1,2)$ $NGARCH(1,1)$ model incorporates the estimated change-points and utilizes the pricing model (8.12) and (8.13) so that the volatility is modelled using $NGARCH(1,2)$ for $t \in [1, 400]$ and $NGARCH(1,1)$ for $t \in [401, 754]$.

S_0	Black-Scholes	Heston-Nandi	NGARCH(1,2)~NGARCH(1,1)
90	0.32580	0.348465	0.53555
95	1.022495	1.50228	1.88978
100	4.14853	4.20374	4.5999
105	8.334915	8.96734	9.0012
110	13.03975	13.2435	13.4325

Table 8.3: Comparison of Call Option Prices for $T = 400$

From the Table 8.3 it is evident that time variation in volatility is important, and that GARCH models out- perform the Black-Scholes model since the latter seems to under-price. This is because the Black-Scholes model is homoscedastic which does not reflect the true dynamics of the financial time series. On the other hand, though the Heston-Nandi *GARCH* and *NGARCH*$(1,1)$ model are able to capture heteroscedaticity these models may not exhibit sufficient volatility dynamics as it is unlikely that one variance function is sufficient to explain the variation in option prices across time. Thus we propose the use of a piecewise variance function following identified change-points which in this case resulted to *NGARCH*$(1,2) \sim NGARCH(1,1)$ under the locally risk-neutral probability measure $\mathbb{Q}$.
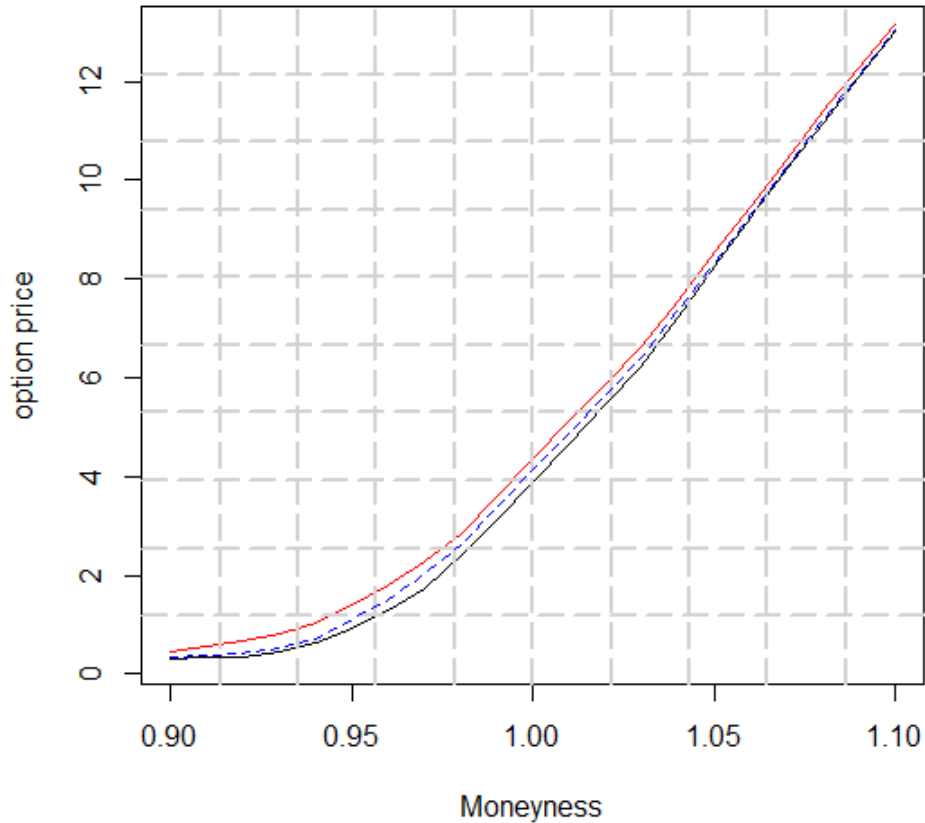
Figure 8.3: Plot of Option Prices against Moneyness

Figure 8.3 shows a plot of the the option prices against moneyness for a early excerise date at time $t = 400$. The black solid line, blue dotted line and red solid line represent the option prices following the Black-Scholes model, Heston-Nandi GARCH model and the $NGARCH(1,2) \sim NGARCH(1,1)$ model repectively. This follows parameter estimates given in Tables 8.1 and 8.2 which yielded the prices in Table 8.3. From the plot we see that the pricing models seem to converge when the option is in-the-money. However, when the option is out-of-the-money, the $NGARCH(1,2) \sim NGARCH(1,1)$ model gives higher prices compared to the Black-Scholes and Heston-Nandi GARCH models. This is attributed to the high volatility experienced in during the period $t \in [1 : 400]$ where under the $NGARCH(1,2) \sim NGARCH(1,1)$ pricing model, the

NGARCH(1,2) model is used to model the returns process. In contrast the volatility is assumed to be constant under the Black-Scholes and assuming one heteroscedastic GARCH model under Heston-Nandi GARCH model fot $t \in [1, 754]$ yet the volatility is high during $t \in [1 : 400]$ and low during $t \in [401 : 754]$.

S_0	Black-Scholes	NGARCH(1,2)~NGARCH(1,1)
90	0.32034	0.43284
95	0.900067	1.38978
100	3.99578	4.00192
105	8.26304	8.55054
110	13.00850	13.1535

Table 8.4: Comparison of Call Option Prices for $T = 754$

Figure 8.3 shows a plot of the the option prices against moneyness for an early excerise date at time $t = 400$. The black solid line, blue dotted line and red solid line represent the option prices following the Black-Scholes model, Heston-Nandi GARCH model and the $NGARCH(1,2) \sim NGARCH(1,1)$ model repectively. This follows parameter estimates given in Tables 8.1 and 8.2 which yielded the prices in Table 8.3. From the plot we see that the pricing models seem to converge when the option is in-the-money. However, when the option is out-of-the-money, the $NGARCH(1,2) \sim NGARCH(1,1)$ model gives higher prices compared to the Black-Scholes and Heston-Nandi GARCH models. This is attributed to the high volatility experienced in during the period $t \in [1 : 400]$ where under the $NGARCH(1,2) \sim NGARCH(1,1)$ pricing model, the $NGARCH(1,2)$ model is used to model the returns process. In contrast the volatility is assumed to be constant under the Black-Scholes and assuming one heteroscedastic GARCH model under Heston-Nandi GARCH model fot $t \in [1, 754]$ yet the volatility is high during $t \in [1 : 400]$ and low during $t \in [401 : 754]$.
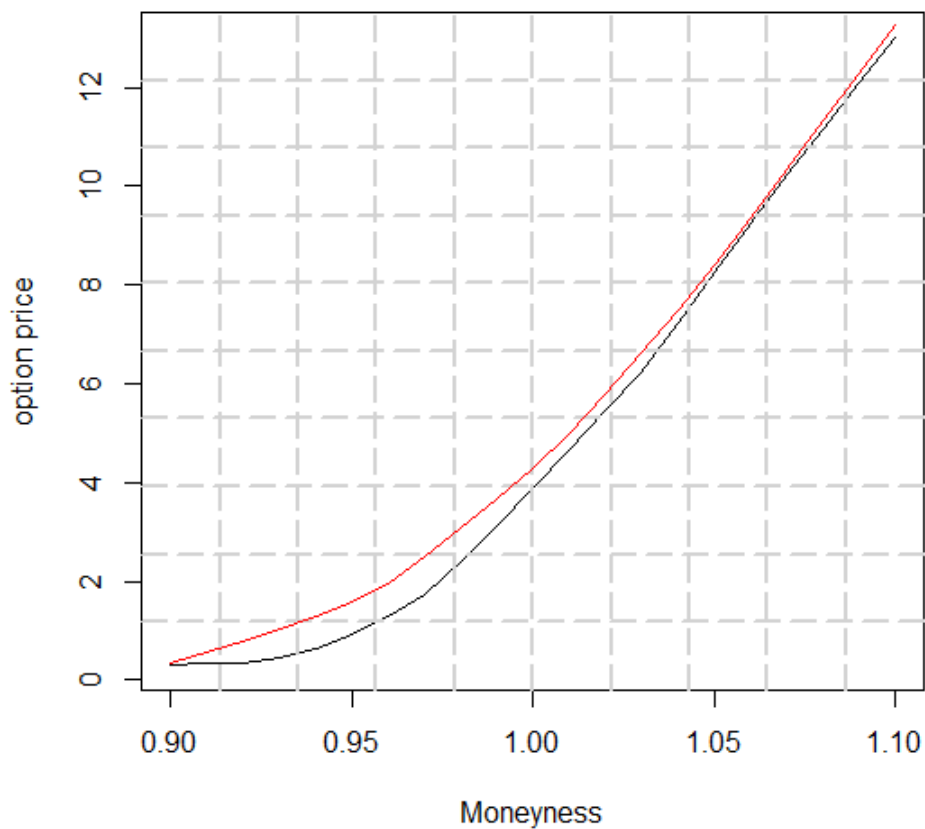
Plot of Option Prices against Moneyness

Figure 8.4: Plot of Option Prices against Moneyness for $T = 400$

Figure 8.4 shows a plot of the the option prices against moneyness for an early excerise date at time $t = 400$. The black solid line and red solid line represent the option prices following the Black-Scholes model and the $NGARCH(1,2) \sim NGARCH(1,1)$ model repectively.
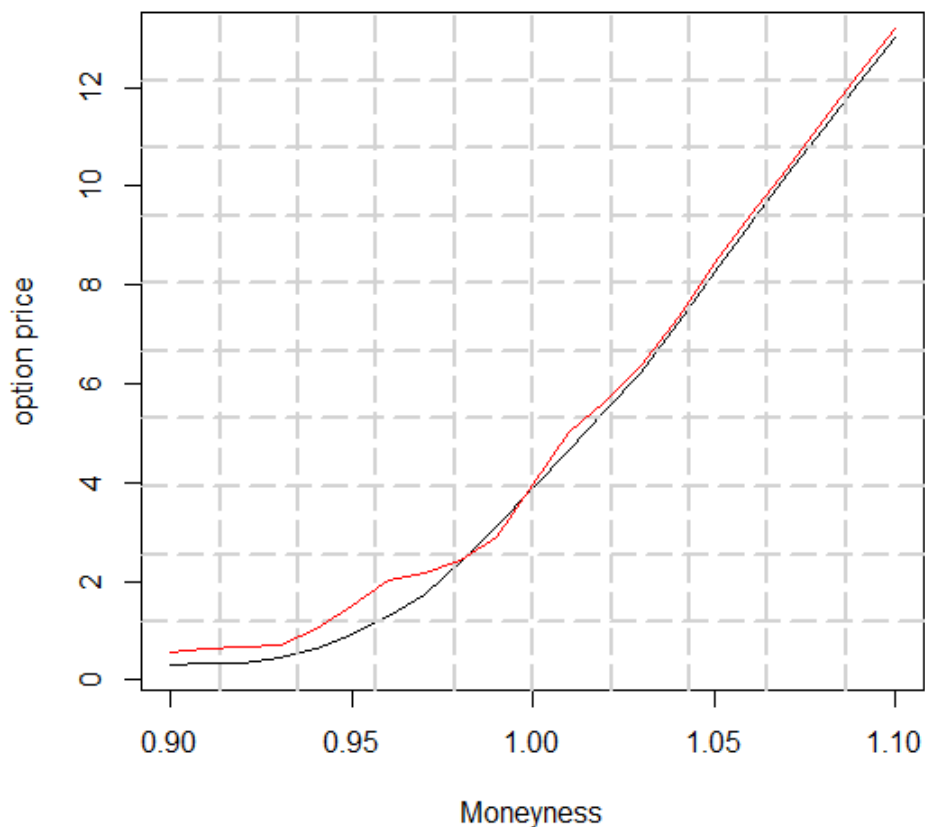
Plot of Option Prices against Moneyness

Figure 8.5: Plot of Option Prices against Moneyness for $T = 754$

Figure 8.5 shows a plot of the the option prices against moneyness for an option ex-cerised at the expiration date at time $t = 754$. The black solid line and red solid line rep-resent the option prices following the Black-Scholes model and the $NGARCH(1,2) \sim NGARCH(1,1)$ model repectively. From the plot we see that the pricing models are equal at two points as the option approaches the at-the-money position. This is in con-trast to the results obtained in Figure 8.4 where the $NGARCH(1,2) \sim NGARCH(1,1)$ model gives higher prices compared to the Black-Scholes when the option is out-of-the-money and at-the-money. This is attributed to the lower volatility experienced during the period $t \in [401, 754]$ where under the former model, volatility is modeled using $NGARCH(1,1)$ during that period. This in effect reduces the option pricess.

However, when the option is out-of-the-money, the $NGARCH(1,2) \sim NGARCH(1,1)$ model gives higher prices compared to the Black-Scholes as was the case in Figure 8.4.

Suppose an investor instructs a broker to buy an KES/USD currency American call option contract with a strike of Kes. 100 and an expiration date at $T = 754$. The broker will release these instructions to a trader at the Securities Exchange. This trader will find another trader who wants to sell a call option contract on the KES/USD currency option with a strike price of Kes. 100 and a price will be agreed. For purposes of this illustration, we ignore the bid-offer spread and assume that the prices are at 4.14853 or 4.59999 assuming the Black-Scholes and $NGARCH(1,2) \sim NGARCH(1,1)$ models respectively as given in Table 8.3. These are the prices for an option to buy one unit. In most markets however, an option contract is a contact to buy or sell 100 shares. Therefore the investor who assumes a long position must arrange for 414.853 or 459.999 to be remmited to the exchange through the broker. The exchange will then arrange for this amount to be passed on to the party on the other side of the transaction who assumes the short position. As a result, the investor has obtained at a cost of 414.853 or 459.999 the right to but 100 units of the KES/USD currency American call option for Kes. 100 each. The party on the other side of the transaction has received 414.853 or 459.999 and has agreed to sell 100 units of the KES/USD currency American call option for Kes. 100 each if the investor chooses to exercise the option.

The investor wishes to exercise the option at time $t = 400$. If the KES/USD exchange rate has not risen above Kes. 100, the option is not exercised and the investor ends up losing 414.853 or 459.999 under the Black-Scholes and $NGARCH(1,2) \sim NGARCH(1,1)$ pricing models repectively. However, if the KES/USD exchange rate has risen above Kes. 100 to say Kes. 105, the investor is able to buy 100 units at Kes. 100 each when they are worth Kes. 105. This results into a gain of 85.147 or 40.001 under the Black-Scholes and $NGARCH(1,2) \sim NGARCH(1,1)$ pricing models repectively. This indicates that the investor is likely to make more profit under

the Black-Scholes pricing model as opposed to the $NGARCH(1,2) \sim NGARCH(1,1)$ model. However, the Black-Scholes model is homoscedastic and does not reflect the true dynamics of volatility as displayed in Figure 8.1. In reality, the heteroscedastic model $NGARCH(1,2) \sim NGARCH(1,1)$ model is more adequate to capture the volatility dynamics inherent.

Suppose the same investor goes for a long position at a price of 414.853 in contract A and a short position at a price 459.999 in contract B. Therefore, they will pay 414.853 for contract A and receive 459.999 for contract B. The investor wishes to exercise the option at time $t = 400$. If the KES/USD exchange rate has not risen above Kes. 100, the investor and the counter party will not excersise the options for contract A and B respectively. This woul mean that the investor would lose 414.853 from contact A but will retain 459.999 from contact B. As a result, the investor will make a profit of 45.146. On the other hand, if the KES/USD exchange rate has risen above Kes. 100 to say Kes. 105, the investor and the counter party would both exercise their options. The investor would gain 85.147 from excersing contact A. However, they would lose 40.001 from contract B. Consequently, the investor would make a profit of 45.146. Therefore by taking a long position on a call option assuming lower volatilty and a short position assuming higher volatility results into the same profit whether the option is in-the-money or out-of-the-money as demonstrated by the illustration above.

This is in contrast to assuming a long and a short position at the same price, say 459.999. If the KES/USD exchange rate does not rise above Kes. 100, the investor will lose 459.999 assuming the long position and receive 459.999 assuming the short position. This would result into zero profit. If the KES/USD exchange rate rises to Kes. 105, the investor would make a profit of 40.001 assuming a long position and a loss of 40.001 assuming a short position. This would result into zero profit. Thus, the volatility dymanics will affect the profits realized. It is therefore important for a investor trading in American options to consider change-points within the volatility structure of a financial returns series when choosing an early exercise date.

Chapter 9

CONCLUSION AND

RECOMMENDATIONS

9.1   Conclusion

In this thesis, a change-point estimator for estimating change attributed to change in GARCH model order specification has been proposed. Given that plausible values for the model orders p and q can be arrived at through inspection of sample autocorrelations and sample partial autocorrelations of a squared returns series it is hypothesized that a change-point in a series occurs when the autocorrelation structure changes. By the Holders Inequality it is shown that the autocorrelation function belongs to the L_1 space and thus the Manhattan distance of sample autocorrelation is proposed as a suitable divergence measure. Procedures are based on the sample autocorrelation function of squared series. Two sequences of autocorrelations with the same dimension are obtained by sequentially dividing the series. A change-point process is generated by assuming a lower bound of a weighted Manhattan distance of the two sequences. The change-point estimator is the first point at which there is maximal sample evidence for a break in the sample autocorrelation function given by the change-point process. To facilitate the detection of multiple change-points, binary segmentation technique is

applied to extend the single change-point detection algorithm.

Data is generated through Monte Carlo simulations under different GARCH modes specifications. Plots of the autocorrelation functions for the simulated series are examined. A comparison of the autocorrelation function for $GARCH(1,1)$ and $GARCH(2,1)$ reveals that the autocorrelation structure is different whether there is a change in model order p. On comparing autocorrelation function for $GARCH(1,1)$ and $GARCH(1,2)$ it is also revealed that the autocorrelation structure is different whether there is a change in model order q. On further comparison of the autocorrelation function for $GARCH(1,1)$ and $GARCH(3,3)$ it is revealed that the autocorrelation structure is different whether there is a change in both model orders p and q. These graphical results are further reinforced by comparing the autocorrelation coefficients for different models across different lags.

The proposed estimator is also observed to correctly identify the change-points when the true change-point is located in the middle of series and when the sample is large. To affirm the performance of the estimator, the similarity index given by Adjusted Rand Indices are utilized to compare the segmentation created by the proposed change-point estimator and the true change-point where an index of one implies a perfect match between segments. It is established that ARI increases and tends to one as the size of change increases irrespective of the sample size and of the source of change. It is observed that as the sample size increases, the ARI generally reduces for the proposed estimator which indicates that testing for change in order p is most suitable for short-term time horizon. It is however not clear of the trend in ARI as the sample size increases when the change occurs in model order q, though the similarity index are close to one as in the former case. When the change is attributed to both model order p and q, it is seen that the estimator is more suitable for short-term time horizons and for greater sizes of change.

It is however established that the change-point process generating the estimator is random in nature does not conform to a known distribution. Histograms are utilized to

121

assess the sampling distribution of the change-point estimator. It is observed that the distribution is approximately Normal when the change-point is positioned at the middle due to the equal distribution of tails to the left and to the right of the true change-point irrespective of the source of change. However, when the change-point is positioned quarter-way in the data, the distribution seems to be positively skewed as depicted by the longer tail to the right of the true change-point. Conversely, the distribution is negatively skewed when the changepoint is positioned three-quarters through the data as shown by the longer tail to the left of the true change-point. In general, however, it seen that the degree of skewness reduces as the size of model order change increases. The asymptotic consistency of the estimator is prooven theoretically based on some properties specific to sequence of stationary random variables with finite second and fourth moments. A condition is established under the probability of observing values outside this true change-point will tend to zero as the sample size increases. By the consistency of the proposed estimator, it means that the estimates converge to the true change-point parameter.

The limit theory of the process generating the estimator is also established. The general theory of the sample autocovariance and sample autocorrelation functions of a stationary GARCH process forms the basis. Specifically the point processes theory is utilized to obtain their weak convergence limit at different lags. This is further extended to the change-point process. The limits are found to be generally random as a result of the infinite variance.

The research culminates with the application of the change-point estimator in pricing American options. The estimator is applied to geometric returns of KES/USD exchange rate data where the series is found to have multiple change-points with two change-points being estimated. Suitable GARCH models are fitted following the estimated change-points and consequently used to model volatility when pricing the American options with the identified change-points as the possible early exercise dates. Comparison is made between the performance of the fitted GARCH models and

Black-Scholes model by examining plots of the option prices against moneyness. It is observed that the two pricing models are equal at as the option is at at-the-money position. The fitted piecewise GARCH model gives higher prices compared to the Black-Scholes when the option is out-of-the-money. The prices are very close when the option is in-the-money which could be attributed to volatility change in the series. Thus, the volatility dymanics affect the prices of options. It is therefore important for a investor trading in American options to consider changepoints within the volatility structure of a financial returns series when choosing an early exercise date.

In conclusion the estimator is deemed to correctly estimate the change-point when the source of change is in the model order p. This could be linked to the fact that the sample autocorrelation function is used in determining the model order p and not q. In addition the estimator works best when the change-point is located at the middle of a sample probably because there is ample data for estimation before and after the change-point. Although the estimator yield ARI close to one irrespective of the sample size, it is seen that the estimator is more suitable for short-term time horizons.

9.2   Recommendation

It should be noted that the proposed estimator only looks at estimation of the change-point when the series is departing from $GARCH(1,1)$ to any other $GARCH(p,q)$ models. Future research should look into generalizing this work so as to examine departure from other model order specification other that $GARCH(1,1)$ such as from $GARCH(p,q)$ to $GARCH(p^*,q^*)$ where $(p,q) \neq (1,1)$ and $(p^*,q^*) \neq (1,1)$.

The estimator has been applied to historical data only. We recommend the application of the estimator to forecasted non-stationary series to enable estimation of future change-points. The foreseen challenge here in would be that with the relaxation of the stationarity assumption as specified in (3.7) would result into generation of different non-stationary series. It would be therefore of interest to establish if the the estimator

converges to some point given the non-stationary series with increased simulations.

It was also established that the estimator works best when the source of change is in model order p. This is attibuted to the fact that the estimator is based on the Manhattan distance of the sample autocorrelation function yet the sample autocorrelation is used in determining the model order p. Given that the model order q is drawn from the sample partial autocorrelation function, we recommend that proposed estimator should be extended to also incorporate the partial autocorrelation function.

From the assessment of the histograms it was established that the sampling distribution of the change-point estimator is approximately normal when the change-point is located at the middle of the series but skewed as the change-point moves away from the middle. Further studies can be carried out to look into the sampling distribution of the change-point estimator k^*.

The asymptotic behavior of the change-point process is established on the basis of examining the asymptotic behavior of the sample autocovariance and sample autocorrelation functions. The limits of the sample autocovariance and sample autocorrelation functions are expressed in terms of the limiting point processes. The limit distributions are the difference of ratios of the infinite variance stable vectors or functions of such vectors. As a result, determination of the quantiles from the limit distributions is difficult. The limits are also generally random as a result of the infinite variance. Future work will be aimed at identifying the limit distributions so as to make the results directly applicable for hypothesis testing purposes.

REFERENCES

Abdalla, S. Z. S. and Winker, P. (2012). Modelling stock market volatility using univariate garch models: Evidence from sudan and egypt. *International Journal of Economics and Finance*, 4(8).

Bao, D. and Yang, Z. (2008). Intelligent stock trading system by turning point confirming and probabilistic reasoning. *Expert Systems with Applications*, 34(1):620–627.

Beguin, J., Gourieroux, C., and Monfort (1980). Identification of a mixed autoregressive-moving average process: the corner method. *Time Series*, pages 423–435.

Bin, C. and Yongmiao, H. (2012). Testing for smooth structural changes in time series models via nonparametric regression. *Econometrica*, 80(3):1157â1183.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, 31:307–327.

Bougerol, P. and Picard, N. (1992). Strict stationarity of generalized autoregressive processes. *Annals of Probability*, 20:1714–1730.

Breiman, L. (1965). On some limit theorems similar to arc-sin law. *Theory Probability Applications*, 10:323–331.

Changli, H. and Timo, T. (1999). Fourth moment structure of the garch(p,q) process. *Econometric Theory*, 15(6):824–846.

Chen, J. and Gupta, A. K. (2000). Parametric statistical change point analysis. *Birkhauser*.

Cho, H. and Fryzlewicz, P. (2012). Multiscale and multilevel technique for consistent segmentation of nonstationary time series. *Statistica Sinica*, 22(1):207–229.

Chung, F., Fu, T., Luk, R., and Ng, V. (2001). Flexible time series pattern matching based on perceptually important points. *In International joint conference on artificial intelligence workshop on learning from temporal and spatial data*, pages 1–7.

Dahlhaus, R. and Subba Rao, S. (2006). Statistical inference for time-varying arch processes. *Annals of Statistics*, 34:1075–1114.

Damodaran, A. (2012). *Investment Valuation:Tools and Techniques for Determining the Value of Any Asset,Third Edition*. John Wiley Sons.

Davis, R. A. and Hsing, T. (1995). Point process and partial sum convergence for weakly dependent random variables with infinite variance. *Annals Probability*, 23:879–917.

Davis, R. A. and Mikosch, T. (1998). The sample autocorrelations of heavy-tailed processes with applications to arch. *Ann. Statist.*, 26(5):2049–2080.

Dobos, L. and Abonyi, J. (2012). On-line detection of homogeneous operation ranges by dynamic principal component analysis based time-series segmentation. *Chemical Engineering Science*, 75:96–105.

Duan., J.-C. (1995). The garch option pricing model. *Mathematical Finance*, 5:13–32.

Duncan, S. R. and Bryant, G. F. (1996). A new algorithm for segmenting data from time series. *Proceedings of the 35th IEEE Conference on Decision and Control*, 3:3123–3128.

Fancourt, C. L. and Principe, J. C. (1998). Competitive principal component analysis for locally stationary time series. *Signal Processing, IEEE*, 46(11):3068–3081.

Faris, W. G. (2004). Real analysis: Part ii. *Arizona Math*.

Fitzgibbon, L. J., Dowe, D. L., and Allison, L. (2002). Change-point estimation using new minimum message length approximations. *PRICAI 2002: Trends in Artificial Intelligence. Springer*, pages 244–254.

Fowlkes, E. B. and Mallows, C. L. (1983). A method for comparing two hierarchical clusterings. *Journal of the American Statistical Association*, 78(383):553–569.

Francq, C. and Zakoian, J. (2011). *GARCH Models: Structure, Statistical Inference and Financial Applications*. Wiley.

Frick, K., M. A. and Sieling, H. (2014). Multiscale change point inference. *Journal of the Royal Statistical Society, Series B (with discussion)*, 76:495–580.

Fu, T. C., F.I., C., and M., N. C. (2006). Financial time series segmentation based on specialized binary tree representatios. *In DMIN*, pages 3–9.

Gichuhi, A. W. (2008). *Nonparametric changepoint analysis for bernoulli random variables based on neural networks*. doctoralthesis, Technische Universit Kaiserslautern.

Guralnik, V. and Srivastava, J. (1999). Event detection from time series data. *In Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining, ACM*, pages 33–42.

Heston, S. L. and Nandi, S. (2000). A closed-form garch option valuation model. *The Review of Financial Studies*, 13(3):585–625. Available online at http://www.jstor.org/stable/2645997.

Himberg, J., Korpiaho, K., Mannila, H., Tikanmaki, J., and Toivonen, H. T. (2001). Time series segmentation for context recognition in mobile devices. *ICDM Proceedings IEEE International Conference on Data Mining*, pages 203–210.

Hommes, C., Gaunersdorfer, A., and Wagener, F. O. (2003). Bifurcation routes to volatility clustering under evolutionary learning. *working paper, CenDEF*.

Hsieh, K. C. and Ritchkeny, P. (2005). An empirical comparison of garch option pricing models. *Review of Derivative Research*, 8(3):129–150. Available online at https://doi.org/10.1007/s11147-006-9001-3.

Hubert, L. and Arabie, P. (1985). Comparing partitions. *Journal of Classification*, 2(1).

Hull, J. C. (2011). *Options, futures and other derivatives*. Prentice Hall.

Inclan, C. and Tiao, G. (1994). Use of cumulative sums of squares for retrospective detection of changes of variance. *Journal of the American Statistical Association*, 89:913–923.

Jackson, B., Sargle, J. D., Barnes, D., Arabhi, S., A. A., Gioumousis, P., Gwin, E., Sangtrakulcharoen, P., Tan, L., and Tsai, T. T. (2005). An algorithm for optimal partitioning of data on an interval. *IEEE, Signal Processing Letters*, 12(2):105–108.

Jiang, J., Zhang, Z., and Wang, H. (2007). A new segmentation algorithm to stock time series based on pip approach. *In WiCom, International Conference on Wireless Communications, Networking and Mobile Computing, IEEE*, pages 5609–5612.

Kallenberg, O. (1983). Random measures. *Akademie-Verlag, Berlin*, 3.

Kesten, H. (1973). Random difference equations and renewal theory for products of random matrices. *Acta Mathematica*, 131:207–248.

Killick, R., a. F. P. and Eckley, I. (2012). Continuous inspection schemes. *ArXiv e-prints*.

Kokoszka, P. and Leipus, R. (2000). Change-point estimation in arch models. *Bernoulli*, 6:513–539.

LeBaron, B. (2000). Agent-based computational finance:suggested readings and early research. *Journal of Economic Dynamics and Control*, pages 679–702.

Liu, M. (2000). Modeling long memory in stock market volatility. *Journal Econometrics*, (99):139–171.

Luc, B., Arnaud, D., and Jeroen, V. R. (2011). Marginal likelihood for markov-switching and change-point garch models.

Lux, T. and Marchesi, M. (2000). Volatility clustering in financial markets:a micro simulation of interacting agents. *International Journal of Theoretical and Applied Finance*, (3):675–702.

Matteo, M., Marcella, N., and Ilaria, V. (2012). Financial speculation in energy and agriculture futures markets: A multivariate garch approach.

Matteson, D. and James, N. (2013). A nonparametric approach for multiple change point analysis of multivariate data. *Journal of the American Statistical Associationl*.

Mikosch, T. and Starica, C. (2004). Nonstationarities in financial time series, the long-range dependence and the igarch effects. *Review of Economics and Statistics*, 86:378–390.

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

Okikiolu, O. G. (1971). *Aspects of the theory of bounded integral operators in Lp-spaces*. London, New York, Academic Press.

Oliver, J. J., Baxter, R. A., and Wallace, C. S. (1998). Minimum message length segmentation. *Research and Development in Knowledge Discovery and Data Mining. Springer*, pages 222–223.

Oliver, J. J. and Forbes, C. S. (1997). Bayesian approaches to segmenting a simple time series. *Citeseer*.

Oman, G. (2017). A short proof of the bolzano-weierstrass theorem. *The College Mathematics Journal*.

Page, E. (1954). Continuous inspection schemes. *Biometrika*, 14:100–115.

Piotr, K. and Remigijus, L. (2000). Change-point estimation in arch models. *Bernoulli*, 6(3):513–539.

Polzehl, J. and Spokoiny, V. (2006). Varying coefficient garch versus local constant volatility modeling. comparison of the predictive power. *The name of the Journal*.

Ponnusamy, S. (2012). *Foundations of Mathematical Analysis*. Springer.

R., K., P., F., and I.A., E. (2012). Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association*, 107(500):1590–1598.

Reiss, R. D. (1993). *A Course on Point Processes*. Springer-Verlag.

Rigaill, G. (2010). Pruned dynamic programming for optimal multiple change-point detection. *ArXiv e-prints*.

Rohan, N. and Ramanathan, T. V. (2012). Non-parametric estimation of a time varying garch model. *Journal of Nonparametric Statistics*.

Shinazi, R. (2012). *From Calculus to Analysis*. Springer.

Terasvirta, T. (2009). An introduction to univariate garch models. *Springer-Verlag Berlin Heidelberg*.

Thurner, S. J., Farmer, D., and Geanakoplos, J. (2012). Leverage causes fat tails and clustered volatility. *Quantitative Finance*, 12(5):695–707.

Tze Leung, L. and Haipeng, X. (2013). Stochastic change-point arx-garch model and their applications to econometric time series. *Statistica Sinica*, 23:1573–1594.

Wang, Z. J. and Willett, P. (2004). Joint segmentation and classification of time series using class-specific features. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE*, 34(2):1056–1067.

Zhang, T. (2016). Testing for jumps in the presence of smooth changes in trends of nonstationary time series. *Electron. J. Statist.*, 10(1):706–735.

Appendix A

APPENDIX OF THEOREMS

In addition to the previously stated Theorems 15 to 16 on point processes, the following additional Theorems are utilized in the proof of the Proposition 2.

Theorem 21. *(Bolzano-Weierstrass)*

Let $\{A_n\}_{n\in\mathbb{N}}$ be a sequence of real numbers that is bounded. Then there exists a subsequence $\{A_{n_k}\}_{n_k\in\mathbb{N}}$ that converges.

Proof. For proof os this Theorem, see Oman (2017).

□

Theorem 22. *(Invariance property of subsequences)*

If $\{A_n\}_{n\in\mathbb{N}}$ is a convergent sequence, then every subsequence of that sequence converges to the same limit.

Proof. For proof os this Theorem, see Ponnusamy (2012).

□

Theorem 23. *(Algebra on Sequences)*

If the sequences $\{A_n\}_{n\in\mathbb{N}}$ converges to L and $\{B_n\}_{n\in\mathbb{N}}$ converges to M then the following hold:

(i) $\lim\limits_{n\to\infty}(A_n+B_n)=\lim\limits_{n\to\infty}A_n+\lim\limits_{n\to\infty}B_n=L+M$

(ii) $\lim\limits_{n\to\infty}(A_n \cdot B_n) = \lim\limits_{n\to\infty} A_n \ \lim\limits_{n\to\infty} B_n = L \cdot M$

if $B_n \neq 0, \forall n \in \mathbb{N}$ and $M \neq 0$ then

(iii) $\lim\limits_{n\to\infty} \frac{A_n}{B_n} = \frac{\lim\limits_{n\to\infty} A_n}{\lim\limits_{n\to\infty} B_n} = \frac{L}{M}$

Proof. For proof os this Theorem, see Shinazi (2012).

$\square$

Theorem 24. *(Continuous Mapping)*

Let a function $g \colon \mathbb{R}^k \to \mathbb{R}^m$ be continous in every point of a set C such that $P(X \in C) = 1$. Then if $X_n \to X$ then $g(X_n) \to g(X)$.

Theorem 25. *(Algebra on Series)*

Let $\sum A_n$ and $\sum B_n$ be two absolutely convergent series. Then:

 (i) the sum of the two series is again absolutely convergent. Its limit is the sum of the limit of the two series.

 (ii) the difference of the two series is again absolutely convergent. Its limit is the difference of the limit of the two series.

 product of the two series is again absolutely convergent. Its limit is the product of the limit of the two series.

Proof. For proof os this Theorem, see Ponnusamy (2012).

$\square$

Theorem 26. *(Convergent sequences are bounded)*

Let $\{A_n\}_{n \in \mathbb{N}}$ be a convergent sequence. Then the sequence is bounded and the limit is unique.

Theorem 27. *Let $\{X_t\}_{t \in \mathbb{N}}$ be a strictly stationary sequence. Define the partial sums of the sequence by $S_n = \sum_{t=1}^{n} X_t$.*

(i) if $\kappa \in (0,2)$ then

$$a_n^{-1} S_n \overset{d}{\to} S \tag{A.1}$$

where $S = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} P_i Q_{ij}$ has a stable distribution

$\kappa \in (2,4)$ and for all $\varepsilon > 0$, $\lim_{\varepsilon \to 0} \lim_{n \to \infty} \sup P\left[|S_n(0,\delta] - ES_n(0,\delta]| > \varepsilon\right] = 0$ then

$$a_n^{-1} S_n - ES_n(0,1] \overset{d}{\to} S \tag{A.2}$$

where S is the distributional limit of

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} P_i Q_{ij} I_{\{|P_i Q_{ij}| > a_n \delta\}} - \} \delta < |x| \le 1 x \mu(dx) \tag{A.3}$$

as $\delta \to 0$, μ is the measure in section 2.1 which has a stable distribution.

For every $\delta > 0$, the mapping from M in section 2.1 into $\mathbb{R}$ is defined by

$$T : \sum_{t=1}^{\infty} \varepsilon_{x_t} \to \sum_{t=1}^{\infty} x_t I_{\{|x_t| > \delta\}}$$

and is almost surely continous with respect to the point process N. Thus by continous mapping theorem 24

$$S_n(\delta, \infty) = T(N_n) \overset{d}{\to} T(N) = S(\delta, \infty) \tag{A.4}$$

As $\delta \to 0$, $S(\delta, \infty) \to S(0, \infty) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} P_i Q_{ij}$

Appendix B

APPENDIX OF ALGORITHMS

LIST OF ALGORITHMS B.1 Simulating Stationary GARCH (3,3) Model with no change-point

---

```
#————————-Simulating GARCH(3,3)with no change point ————————-#
RR=1000
#mu=-8.135e-05
omega=6.015e-07
alpha1=2.3945e-01
alpha2=2.097e-01
alpha3=1.589e-01
beta1=1.011e-01
beta2=1.523e-01
beta3=1.034e-01
epsilon=rnorm(RR,mean=0,sd=1)
wwt=rep(0,RR)
sigma=rep(0,RR)
for(i in 4:RR) {
sigma[i]=omega+(alpha1*(wwt[i-1]^2))+(alpha2*(wwt[i-2]^2))+(alpha3*(wwt[i-
3]^2))                   +(beta1*sigma[i-1])+(beta2*sigma[i-2])+(beta3*sigma[i-3])
wwt[i]=(epsilon[i]*sqrt(sigma[i])) }
```

LIST OF ALGORITHMS B.2 Simulating Stationary GARCH(4,4) Models with no change-point

```
#————————-Simulating GARCH(4,4)with no change point ————————-#
RRR=1000
#mu=-8.135e-05
omega=6.015e-07
alpha1=2.689e-01
alpha2=1.078e-01
alpha3=1.536e-01
alpha4=1.240e-01
beta1=5.985e-02
beta2=1.598e-01
beta3=1.013e-01
beta4=5.679e-02
epsilon=rnorm(RRR,mean=0,sd=1)
wwttw=rep(0,RRR)
sigma=rep(0,RRR)
for(i in 5:RRR) {
sigma[i]=omega+(alpha1*(wwttw[i-1]^2))+(alpha2*(wwttw[i-
2]^2))+(alpha3*(wwttw[i-3]^2))
+(alpha4*(wwttw[i-4]^2))+(beta1*sigma[i-1])+(beta2*sigma[i-2])+(beta3*sigma[i-
3]) +(beta4*sigma[i-4]) wwttw[i]=(epsilon[i]*sqrt(sigma[i])) }
```

LIST OF ALGORITHMS B.3 Simulating Stationary GARCH Models with single change-point

#Simulating GARCH(1,1)with one change point with respect to change of order
N=1000
omega=6.004e-07
alpha1=3.414e-01
beta1=7.287e-01
beta2=3.075e-01
epsilon1=rnorm(500)
epsilon2=rnorm(500)
series=rep(0,1000)
xt=rep(0,500)
yt=rep(0,500)
sigma1=rep(0,500)
sigma2=rep(0,500)
for(k in 1:1000){
for(i in 2:(500)) {
sigma1[i]=omega+(alpha1*(xt[i-1]^2))+(beta1*sigma1[i-1])
xt[i]=epsilon1[i]*sqrt(sigma1[i]) }
for(j in 503:1000) {
sigma2[j]=omega+(alpha1*(yt[j-1]^2))+(beta1*sigma2[j-1])+(beta2*sigma2[j-2])
yt[j]=epsilon2[j]*sqrt(sigma2[j]) }
series[k]=c(xt[i],y[j])}

LIST OF ALGORITHMS B.4 Autocorrelation and Partial Autocorrelation for model order change (p,q)

```
#————————-Change in ACF for model order change (p,q)————————-#
rhoo11=acf((xxt^2),lag.max=10)
rhoo22=acf((wwt^2),lag.max=10)
rhoo33=acf((wwtt^2),lag.max=10)
rhoo44=acf((wwttww^2),lag.max=10)
Lagg=round(rhoo11$lag,2)
ACF11=round(rhoo11$acf,2)
ACF22=round(rhoo22$acf,2)
ACF33=round(rhoo33$acf,2)
ACF44=round(rhoo44$acf,2)
diff22=ACF11-ACF22
diff33=ACF11-ACF33
diff44=ACF11-ACF44
data.frame(Lagg,ACF11,ACF22,ACF33,ACF44,diff22,diff33,diff44)
#————————-Change in PACF for model order change (p,q)————————-#
prhoo11=pacf((xxt^2),lag.max=10)
prhoo22=pacf((wwt^2),lag.max=10)
prhoo33=pacf((wwtt^2),lag.max=10)
prhoo44=pacf((wwttww^2),lag.max=10)
Lagg=round(prhoo11$lag,2)
PACF11=round(prhoo11$acf,2)
PACF22=round(prhoo22$acf,2)
PACF33=round(prhoo33$acf,2)
PACF44=round(prhoo44$acf,2)
pdiff22=PACF11-PACF22
pdiff33=PACF11-PACF33
pdiff44=PACF11-PACF44
data.frame(Lagg,PACF11,PACF22,PACF33,PACF44,pdiff22,pdiff33,pdiff44)
```

LIST OF ALGORITHMS B.5 Autocorrelation and Partial Autocorrelation for model order change p and q

```
 #————————-Change in ACF for model order change (p)————————-#
rhoo11=acf((xxt^2),lag.max=10)
rhoo21=acf((zzt^2),lag.max=10)
rhoo31=acf((zztt^2),lag.max=10)
rhoo41=acf((zzttz^2),lag.max=10)
Lagg=round(rhoo11$lag,2)
ACF11=round(rhoo11$acf,2)
ACF21=round(rhoo21$acf,2)
ACF31=round(rhoo31$acf,2)
ACF41=round(rhoo41$acf,2)
diff21=ACF11-ACF21
diff31=ACF11-ACF31
diff41=ACF11-ACF41
data.frame(Lagg,ACF11,ACF21,ACF31,ACF41,diff21,diff31,diff41)
 #————————-Change in ACF for model order change (q)————————-#
## change in ACF for model order change (q)
rhoo11=acf((xxt^2),lag.max=10)
rhoo12=acf((zzt^2),lag.max=10)
rhoo13=acf((zztt^2),lag.max=10)
rhoo14=acf((zzttz^2),lag.max=10)
Lagg=round(rhoo11$lag,2)
ACF11=round(rhoo11$acf,2)
ACF12=round(rhoo12$acf,2)
ACF13=round(rhoo13$acf,2)
ACF14=round(rhoo14$acf,2)
diff12=ACF11-ACF12
diff13=ACF11-ACF13
diff14=ACF11-ACF14
data.frame(Lagg,ACF11,ACF12,ACF13,ACF14,diff12,diff13,diff14)
```

LIST OF ALGORITHMS B.6 Change-Point Detection Dependent functions

```
#————————-Creating an acf function for lag 1——————-#
ACFFF=function(x) {
mu=mean(x)
num<-c(0)
denom<-c(0)
n=length(x)-1
for(i in 1:n) {
num[i]=((x[i+1]-mu)*(x[i]-mu))
denom[i]=((x[i]-mu)^2)
numerator=sum(num)
denominator=sum(denom) }
return(numerator/denominator) }
ACFFF(x)
        #————————-Trimmer for stabilizing process——————-#
trimmed<-function(Dk,tr){
n<-length(Dk)
trim<-c()
for(i in 1:n) {
if(i>(tr*n)&(i<=((1-tr)*n))) {
trim[i]<-Dk[i] }else{"NA"} }
return(trim) }
#trimmed(Dk,tr=0)
            #————————-Change-Point Position——————-#
ChangePoint<-function(DkReal,Dn) {
postn<-c(0)
for(j in 1:length(DkReal)) {
if(DkReal[j]!=Dn)
next print(paste("The change point is",j))
postn<-j }
return(postn) }
```

```
#——————————-Change-Point Detection——————————-#
library(KernSmooth) library(timeDate) library(timeSeries)
change<-function(x) {
n=length(x)-1
Dk<-c(0)
DkReal<-c()
for(i in (1+1):n) {
Dk[i]<-(abs(ACFFF(x[1:i])-ACFFF(x[(i+1):n])))
DkReal<-substituteNA(Dk,type="mean",interp=linear)
DkTrim<-trimmed(Dk,tr=0.10)
}
position<-ChangePoint(DkReal,Dn)
#return(plot(density(trimmed(Dk,tr=0.20),bw=band,kernel="cosine",na.rm      =
TRUE),type="l"))
#return(plot(trimmed(Dk,tr=0.20),type="l",xlab="Time",
#ylab="Dk",main="Absolute Difference of ACF"))
return(Dn, position)
}
x<-xxt^2
change(x)
plot(series,type="l",panel.first    =    grid(10,    lty   =   5,    lwd    =    2),
xlab="Time",ylab="Xt",main="GARCH(1,1) and GARCH(1,2)")
abline(v=c(400,754),col='blue') # change-points detected
```

LIST OF ALGORITHMS B.8 Edivisive Change-point Detection

```
library(Rcpp) library(ecp)
#method:e.divisive
series1<-matrix(c(xxt[1:1000],yyt[1:1000]),ncol=1)
series2<-matrix(c(xxt[1:1000],yytt[1:1000]),ncol=1)
series3<-matrix(c(xxt[1:1000],yytty[1:1000]),ncol=1)
series4<-matrix(c(xxt[1:1000],zzt[1:1000]),ncol=1)
series5<-matrix(c(xxt[1:1000],zztt[1:1000]),ncol=1)
series6<-matrix(c(xxt[1:1000],zzttz[1:1000]),ncol=1)
series7<-matrix(c(xxt[1:1000],wwt[1:1000]),ncol=1)
series8<-matrix(c(xxt[1:1000],wwtt[1:1000]),ncol=1)
series9<-matrix(c(xxt[1:1000],wwttww[1:1000]),ncol=1)
output1 <- e.divisive(series1, R = 100, alpha = 2)
output2 <- e.divisive(series2, R = 100, alpha = 2)
output3 <- e.divisive(series3, R = 100, alpha = 2)
output4 <- e.divisive(series4, R = 100, alpha = 2)
output5 <- e.divisive(series5, R = 100, alpha = 2)
output6 <- e.divisive(series6, R = 100, alpha = 2)
output7 <- e.divisive(series7, R = 100, alpha = 2)
output8 <- e.divisive(series8, R = 100, alpha = 2)
output9 <- e.divisive(series9, R = 100, alpha = 2)
output1$estimates
output2$estimates
output1$k.hat
output1$order.found
output1$considered.last
output1$p.values
ts.plot(series,ylab='Value',main='Change in a Stationary GARCH Sequence')
#alpha:The moment index used for determining the distance #between and within seg-
ments.
# change-points detected for alpha=2 abline(v=c(1016,1754),col='blue')
```

LIST OF ALGORITHMS B.9 Adjusted Rand Index for eDivisive Method

```
#N=2000 ##output given for series
series1<-c(xxt,yyt)##changepoints:1003
series2<-c(xxt,yytt)##changepoints:1008
series3<-c(xxt,yytty)##changepoints:1006
series4<-c(xxt,zzt)##changepoints:1003
series5<-c(xxt,zztt)##changepoints:1005
series6<-c(xxt,zzttz)##changepoints:1005
series7<-c(xxt,wwt)##changepoints:1004
series8<-c(xxt,wwtt)##changepoints:1004
series9<-c(xxt,wwttww)##changepoints:1006
xt<-xxt
U11<-xt[1:1003]
U22<-xt[1:1008]
U33<-xt[1:1006]
U44<-xt[1:1003]
U55<-xt[1:1005]
U66<-xt[1:1005]
U77<-xt[1:1004]
U88<-xt[1:1004]
U99<-xt[1:1006]
V11<-series1[1:1003]
V22<-series2[1:1008]
V33<-series3[1:1006]
V44<-series4[1:1003]
V55<-series5[1:1005]
V66<-series6[1:1005]
V77<-series7[1:1004]
V88<-series8[1:1006]
V99<-series9[1:266]
RAND11 <- adjustedRand(U11,V11)
RAND22 <- adjustedRand(U22,V22)
RAND33 <- adjustedRand(U33,V33)
RAND44 <- adjustedRand(U44,V44)
RAND55 <- adjustedRand(U55,V55)
RAND66 <- adjustedRand(U66,V66)
RAND77 <- adjustedRand(U77,V77)
RAND88 <- adjustedRand(U88,V88)
RAND99 <- adjustedRand(U99,V99)
```

LIST OF ALGORITHMS B.10 Adjusted Rand Index for eDivisive Method

```
#N=2000 ##output given for series
series1<-c(xxt,yyt)##changepoints:1003
series2<-c(xxt,yytt)##changepoints:1008
series3<-c(xxt,yytty)##changepoints:1006
series4<-c(xxt,zzt)##changepoints:1003
series5<-c(xxt,zztt)##changepoints:1005
series6<-c(xxt,zzttz)##changepoints:1005
series7<-c(xxt,wwt)##changepoints:1004
series8<-c(xxt,wwtt)##changepoints:1004
series9<-c(xxt,wwttww)##changepoints:1006
xt<-xxt
U11<-xt[1:1003]
U22<-xt[1:1008]
U33<-xt[1:1006]
U44<-xt[1:1003]
U55<-xt[1:1005]
U66<-xt[1:1005]
U77<-xt[1:1004]
U88<-xt[1:1004]
U99<-xt[1:1006]
V11<-series1[1:1003]
V22<-series2[1:1008]
V33<-series3[1:1006]
V44<-series4[1:1003]
V55<-series5[1:1005]
V66<-series6[1:1005]
V77<-series7[1:1004]
V88<-series8[1:1006]
V99<-series9[1:266]
RAND11 <- adjustedRand(U11,V11)
RAND22 <- adjustedRand(U22,V22)
RAND33 <- adjustedRand(U33,V33)
RAND44 <- adjustedRand(U44,V44)
RAND55 <- adjustedRand(U55,V55)
RAND66 <- adjustedRand(U66,V66)
RAND77 <- adjustedRand(U77,V77)
RAND88 <- adjustedRand(U88,V88)
RAND99 <- adjustedRand(U99,V99)
```

LIST OF ALGORITHMS B.11 Histograms

```
data=read.csv("D:\\PhD_PAU\\HISTOGRAM.csv",header=TRUE)
k500=data$K500 k750=data$k750 k250=data$k250
hist(k250, breaks = 10, plot = TRUE, xlim=c(0,1000), xlab="change-point
k",main="Histogram for k=250")
hist(k500, breaks = 10, plot = TRUE,xlim=c(0,1000), xlab="change-point
k",main="Histogram for k=500")
hist(k750, breaks = 10, plot = TRUE,xlim=c(0,1000), xlab="change-point
k",main="Histogram for k=750")
k250pq=data$k250pq k500pq=data$k500pq k750pq=data$k750pq
hist(k250pq, breaks = 10, plot = TRUE,xlim=c(0,1000), xlab="change-point
k",main="Histogram for k=250")
hist(k500pq, breaks = 10, plot = TRUE,xlim=c(0,1000), xlab="change-point
k",main="Histogram for k=500")
hist(k750pq, breaks = 10, plot = TRUE,ylim=c(0,400),xlim=c(0,1000), xlab="change-
point k",main="Histogram for k=750")
k500p=data$k500p k500q=data$k500q
hist(k500p, breaks = 10, plot = TRUE,xlim=c(0,1000), xlab="change-point
k",main="Histogram for change in p")
hist(k500q, breaks = 10, plot = TRUE,xlim=c(0,1000), xlab="change-point
k",main="Histogram for change in q")
library(tcltk) library(asbio)
skew(k250,method="unbiased")                    skew(k500,method="unbiased")
skew(k750,method="unbiased")
skew(k250pq,method="unbiased")                  skew(k500pq,method="unbiased")
skew(k750pq,method="unbiased")
skew(k500p,method="unbiased") skew(k500q,method="unbiased")
kurt(k250,method="unbiased")                     kurt(k500,method="unbiased")
kurt(k750,method="unbiased")
skew(k500p,method="unbiased")                    skew(k500q,method="unbiased")
skew(k500pq,method="unbiased")
```

LIST OF ALGORITHMS B.12 Exchange-Rate Data

```
data=read.csv("D:\\PhD_PAU\\ExchangeRate.csv",header=TRUE)
xt=log(data[2:2000]/data[1:1999])
plot(xt,type="l",panel.first = grid(10, lty = 5, lwd = 2))
    #——-Comparing the acf and pacf of (xt*xt) to determine model orders——-#
par(mfrow=c(2,2)) par(new=TRUE)
acf(xt*xt)
pacf(xt*xt)
rhoo=acf(xt^2,lag.max=10)
Lagg=round(rhoo$lag,2)
ACF=round(rhoo$acf,2)
data.frame(Lagg,ACF)
prhoo=pacf(xt^2,lag.max=10)
Lagg2=round(rhoo2$lag,2)
PACF=round(rhoo2$acf,2)
data.frame(Lagg2,PACF)
        #——-Fitting a GARCH Model to Exchage Rate Time Series——-#
#load: fGarch, fBasics, timeDate, timeSeries
library(timeDate)
library(timeSeries)
library(fBasics)
library(fGarch)
output1<-garchFit(formula=~garch(1,1),data=xt,include.mean          =
TRUE,cond.dist="QMLE")
output2<-garchFit(formula=~garch(1,2),data=xt,include.mean          =
FALSE,cond.dist="QMLE")
output3<-garchFit(formula=~garch(2,1),data=xt,include.mean          =
FALSE,cond.dist="QMLE")
output4<-garchFit(formula=~garch(2,2),data=xt,include.mean          =
FALSE,cond.dist="QMLE")
output5<-garchFit(formula=~garch(4,1),data=xt,include.mean          =
FALSE,cond.dist="QMLE")
        #————————-Change-Point Detection——————-#
x<-xt^2
change(x)
```

LIST OF ALGORITHMS B.13 Black-Scholes and Heston-Nandi GARCH Option Pricing

```
modell = list(lambda = 4, omega = 8e-5, alpha = 6e-5, beta = 0.7, gamma = 0, rf = 0)
# HN-GARCH log likelihood Parameter Estimation:
mle = hngarchFit(model = modell, x = rtt, symmetric = TRUE)
summary.hngarch(mle)
## summary.hngarch - # HN-GARCH Diagnostic Analysis:
par(mfrow = c(3, 1), cex = 0.75)
summary(mle)
## hngarchStats - # HN-GARCH Moments:
hngarchStats(mle$model)
modelfit = list(lambda = -2.705e+01, omega = 7.296e-13, alpha = 3.553e-06, beta =
7.694e-01, gamma = 0.000e+00)
S = 90 X = 100 Time.inDays = 252 r.daily = 0.03/Time.inDays
sigma.daily = sqrt((modelfit$omega + modelfit$alpha) / (1 - modelfit$beta - mod-
elfit$alpha * modelfit$gamma^2))
data.frame(S, X, r.daily, sigma.daily)
library(timeDate) library(timeSeries) library(fBasics) library(fOptions) library(zoo) li-
brary(xts)
## HNGOption - # Compute HNG Call-Put and compare with GBS Call-Put:
HNG = GBS = Diff = NULL for (TypeFlag in c("c",))
{ HNG = c(HNG, HNGOption(TypeFlag, model = modelfit, S = S, X = X, Time.inDays
= Time.inDays, r.daily = r.daily)$price )
GBS = c(GBS, GBSOption(TypeFlag, S = S, X = X, Time = Time.inDays, r = r.daily,
b = r.daily, sigma = sigma.daily)$price) }
Options = cbind(HNG, GBS, Diff = round(100*(HNG-GBS)/GBS, digits=2))
row.names(Options) <- c("Call")
data.frame(Options)
   #————————-Plots for Option Prices against Moneyness———————-#
option=read.csv("D:\\PhD_PAU\\OptionPrices.csv",header=TRUE)
plot(lowess(option$BS~option$moneyness,f=0.2),type="l",panel.first = grid(10, lty =
5, lwd = 2))
lines(lowess(option$NS~option$moneyness, f=0.2),type="l",col="red",lty=1,panel.first
= grid(10, lty = 5, lwd = 2))
lines(lowess(option$HN~option$moneyness,f=0.2),type="l",col="blue",lty=8,panel.first
= grid(10, lty = 5, lwd = 2))
```

LIST OF ALGORITHMS B.14 Simulated GARCH Option Pricing

```
X0=100
K=100
T=252
r=0.05/T
Time.to.maturity=252
sims<-function() {
sim11<-c() sim12<-c() sum<-c()
for(i in 1:10000){
spec11 = garchSpec(model = list(omega = 0.000000067352, alpha = 0.1, beta
=0.58934))
sim11[i]<-as.numeric(garchSim(spec11, n = 10000))
spec12= garchSpec(model = list(omega=7.311741e-07,alpha = 0.1, beta =
c(3.282616e-01, 3.077644e-01)))
sim12[i]<-as.numeric(garchSim(spec12,        n      =      10000))        sum[i]<-
sum(c(sim12[i],sim11[i])) } return(X0*exp(sum)) }
sims()
Call_Price<-function(X0,K,r,T) {
X_T=sims()
for(i in 1:10000) {
payoff[i]<-max(X_T[i]-K,0)
} mu<-mean(payoff)
call.price<-(exp(-r*T))*mu return(call.price) }
Call_Price(X0,K,r,T)
```