International Journal of Data Science and Analysis
2023; 9(1): 1-12
http://www.sciencepublishinggroup.com/j/ijdsa
doi: 10.11648/j.ijdsa.20230901.11
ISSN: 2575-1883 (Print); ISSN: 2575-1891 (Online)

Detection and Estimation of Change Point in Volatility Function of Foreign Exchange Rate Returns

Josephine Njeri Ngure[1], Anthony Gichuhi Waititu[2], Simon Maina Mundia[3]

[1]Department of Mathematics, Pan African University Institute for Basic Sciences, Technology and Innovation (PAUSTI), Nairobi, Kenya
[2]Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology (JKUAT), Nairobi, Kenya
[3]Department of Statistics and Actuarial Sciences, Dedan Kimathi University of Technology (DEKUT), Nyeri, Kenya

Email address:

jngure@kyu.ac.ke (Josephine Njeri Ngure)

To cite this article:

Josephine Njeri Ngure, Anthony Gichuhi Waititu, Simon Maina Mundia. Detection and Estimation of Change Point in Volatility Function of Foreign Exchange Rate Returns. *International Journal of Data Science and Analysis.* Vol. 9, No. 1, 2023, pp. 1-12.
doi: 10.11648/j.ijdsa.20230901.11

Received: January 13, 2023; **Accepted:** February 22, 2023; **Published:** April 6, 2023

Abstract: This work aims at detection and estimation of a change point in conditional variance function of a Nonparametric Auto-Regressive Conditional Heteroscedastic model. The conditional mean and conditional variance functions are not specified a priori but estimated using Nadaraya Watson kernel. This is because inferences based on nonparametric approaches are robust against misspecification of the conditional mean function and the conditional variance function of returns. The squared residuals obtained after estimating the regression function of the returns are used in estimating the conditional variance function. Further, the squared residuals are used in developing a test statistic for unknown abrupt change point in volatility of the exchange rate returns. The test statistic takes into consideration the conditional heteroskedasticity of the disturbances, dependence of the returns, heterogeneity and fourth moment of returns. This does not require prior knowledge of the marginal or the conditional densities of the returns as opposed to maximum likelihood estimation methods. The estimator for change point is considered as the augmented maximum of the test statistic. The consistency of the estimator is stated as a theorem. The asymptotic distribution associated with the test for unknown break points is the Bessel process distribution. The Bessel process distributions have no known simple closed-form expression for the distribution function which makes it difficult to compute exact p-values. Also, the Bessel process distributions depend on two parameters which makes it hard to tabulate the critical values hence one needs to simulate them. After simulating the critical values, hypothesis testing is done in the presence and absence of a change point in volatility of a simulated time series and the test is shown to reject the null hypothesis in the presence of a change point at alpha level of significance. Further, the test fails to reject the null hypothesis in the absence of a change point at alpha level of significance. An application to United States Dollar/Kenya Shilling historical exchange rates returns is made from 1^{st} January 2010 to 27^{th} November 2020 where the sample size $n = 2839$ is done. Through binary segmentation method, three change points are detected, estimated and accounted for. A significant improvement in describing a time series is expected if a point in time for volatility change has been detected and estimated.

Keywords: Nonparametric, Kernel, Volatility

1. Introduction

The point at which the probabilistic structure of the data changes is known as a change point. This means that the data exhibits different means, different variances or both or higher order moments. The observations therefore have two or more distinct segments each with a unique underlying process. Financial returns typically have constant mean but volatility is rarely constant but changes over time. This could be due to financial liberalization of emerging markets and integration of world equity markets, "shocks" induced by institutional changes, such as changes in exchange rate regimes from a fixed exchange rate regime to a floating exchange rate regime [1], introduction of a single currency like the Euro in Europe, disease outbreaks like COVID-19 pandemic in December 2019 in the world among others.

ln financial instruments, large (small) price movements tend to be followed by large (small) price movements on successive hours, days, weeks or other time durations creating extended regimes of high volatility and others of low volatility. This leads to volatility clustering [2]. Volatility clustering means that volatility behaves like a process with abrupt jumps. The jumps lead to structural breaks in volatility otherwise called change points hence this research. Presence of a change point means that application of one Auto-Regressive Conditional Heteroscedastic model to describe volatility is incorrect and leads to invalid results. A significant improvement in describing the structure of financial returns is obtained if change points in volatility are detected and estimated.

Change in variance of financial returns under the assumption of independence and identical distribution of returns has been done by [3]. Financial returns are however neither independent nor identically distributed. [4], [5], [6], [7] studied change in conditional variance function of time series under different assumptions on dependence, heterogeneity among others.

Change point analysis can be classified into off-line (fixed sample or retrospective) and on-line (sequential) where the classification depends on the sample acquisition approach [8] and [9]. In off-line approach, the data is first observed, then the detection and further estimation of the change-point is done by looking back in time to recognize where the change occurred. These include [10] where a change point in variance in a nonparametric time series regression model with a strong mixing error process using cumulative sum of squares approach introduced by [3] is done and [11] in change point detection in unconditional variance of financial time series. In sequential approaches, new data is continually arriving and is analyzed adaptively. These include but not limited to [12] for change point detection in GARCH(p,q) models and [13] for change point detection in mean. Online approaches are used in quality control, financial risk management, allocation of asset or portfolio selection while off-line approaches are used in genome analysis, linguistics, audiology among others.

2. Method

2.1. NonParametric Auto-Regressive Conditional Heteroscedastic (NP-ARCH) model

Let S_t denote the price of some financial instrument observed at equally spaced time interval. The continuously compounded single period return on the financial instrument from time index $t-1$ to t (t^{th} time interval) for $t = 1, 2, ..., n$ is given by

$$X_t = \ln\left(\frac{S_t}{S_{t-1}}\right) = \ln S_t - \ln S_{t-1} \quad (1)$$

The volatility of the instrument is the standard deviation of the returns. One assumes the existence of a nonparametric and nonlinear relationship between X_t and $X_{t-i}, i = 1, 2, ..., d$ which is modeled by a nonlinear autoregressive process of the form

$$X_t = m(X_{t-1}, X_{t-2}, ..., X_{t-d}) + u_t \quad \text{for} \quad t = 1, 2, ..., n \quad (2)$$

where $\{u_t\}$ is a series of innovations (random shock) which is independent of the past $\{X_t\}$, $m(.)$ is the conditional mean function of the returns in period t given past periods $X_{t-1}, X_{t-2}, ...$ In a nonparametric approach, $m(.)$ is allowed to be from flexible class of functions and is approximated such that precision increases with size of sample. The interest is on the future volatility, so, when innovations are represented as $u_t = \sigma(X_{t-1}, X_{t-2}, ..., X_{t-d})z_t$, Equation (2.2) is extended to a nonlinear Nonparametric Auto Regressive Conditional Heteroscedastic (NP-ARCH) model of the form

$$X_t = m(X_{t-1}, ..., X_{t-d}) + \sigma(X_{t-1}, X_{t-2}, ..., X_{t-d})z_t \qquad t = 1, 2, ..., n \quad (3)$$

where the conditional mean function of the returns in period t given the past periods is given by

$$\mathbb{E}(X_t|X_{t-1} = x_1, ..., X_{t-d} = x_d) = m(X_{t-1}, ..., X_{t-d}) \quad (4)$$

and the conditional variance function of the returns in period t given the past periods is given by

$$\text{Var}(X_t|X_{t-1} = x_1, ..., X_{t-d} = x_d) = \mathbb{E}\left(u_t^2|X_{t-1} = x_1, ..., X_{t-d} = x_d\right) = \sigma^2(X_{t-1}, ..., X_{t-d}) \quad (5)$$

Hence, the model allows for conditional heteroscedasticity. $\{z_t\}$ are independent and identically distributed random variables (errors) which are time invariant with unspecified continuous and positive probability density function f_z, $\mathbb{E}(z_t|\mathcal{F}_{t-1} = X_{t-1}, ..., X_{t-d}) = 0$, $\text{Var}(z_t|\mathcal{F}_{t-1} = X_{t-1}, ..., X_{t-d}) = 1$ and independent of $X_{t-1}, ..., X_{t-d}$ while $\mathbb{E}z_t = 0$, $\mathbb{E}z_t^2 = 1, \mathbb{E}\left(z_t\right)^4 < \infty$.
Model (2.3) is a flexible nonparametric time series model because it does not impose any (parametric) particular form on the conditional mean and conditional variance functions hence avoiding any model misspecification. However, in higher dimensions, there is poor performance called curse of dimensionality, which for $d > 2$ the estimation of the functions in Equation (2.3) is complicated unless one has a very large sample. Hence, setting $d = 1$ Equation (2.3) becomes

$$X_t = m(X_{t-1}) + \sigma(X_{t-1})z_t. \quad (6)$$

In the absence of a change point, X_t is assumed to be strictly stationary and strong mixing an assumption satisfied by most financial time series [14]. Hence the theorem below is applied.

Theorem 2.1. *Strong mixing condition*: Suppose the existence of a probability space $(\Omega, \mathcal{F}, P)$. Let the dependence measure between any two σ fields $\mathcal{A}$ and $\mathcal{B} \subset \mathcal{F}$ as introduced by [15] be defined by

$$\alpha(\mathcal{A}, \mathcal{B}) := \sup_{A \in \mathcal{A}, B \in \mathcal{B}} |P(A \cap B) - P(A) P(B)|.$$

Now, suppose $\{X_t, t \in \mathbb{Z}\}$ is a two-sided sequence of variables on a given probability space $(\Omega, \mathcal{A}, P)$. For $-\infty \leq j < l \leq \infty$, let $\mathcal{F}_j^l = \sigma(X_t, j \leq t \leq l)$ denote the $\sigma -$ field of events which has been generated by the family $\{X_t, j \leq t \leq l\, (t \in \mathbb{Z})\}$. For each $n \in \mathbb{N}$, define the "coefficient of dependence (mixing)" $\alpha(n)$ by,

$$\alpha(n) = \sup_{-\infty < j < \infty} \alpha\left(\mathcal{F}_{-\infty}^{j}, \mathcal{F}_{j+n}^{\infty}\right)$$

where $\mathcal{F}_{-\infty}^{j}$ is the $\sigma -$ field of events contained in the past of the sequence $\{X_t\}$ up to time j and $\mathcal{F}_{j+n}^{\infty}$ is the $\sigma -$ field of events contained in the future of the sequence $\{X_t\}$ from time $j+n$ onwards. The sequence of numbers $(\alpha(n), n \in \mathbb{N})$ is non-increasing (decreasing)in n and are non-negative. The sequence $\{X_t, t \in \mathcal{Z}\}$ is therefore said to be "strong mixing" or "α mixing" if

$$\lim_{n \longrightarrow \infty} \alpha(n) = 0$$

and hence the sequence is asymptotically independent between the past and the future.

Let $\{(X_t, X_{t-1}) \in \mathbb{R} \times \mathbb{R} : t = 1, 2, ..., n\}$ be a sample of size $n \in \mathbb{N}$, $K(\cdot) : \mathbb{R} \longrightarrow \mathbb{R}$ be a kernel function which is a bounded continuous function on $\mathbb{R}$ satisfying the assumption of normalization $\int K(u)du = 1$, which ensures that the method of kernel density estimation results in a probability density function, symmetry about zero $K(u) = K(-u) \quad \forall u$ implying that all the odd moments are zero, $K(u) \geq 0 \ \forall\ u$ implying that $K(u)$ is a probability density function, $\int u\, K(u)du = 0$, $u_2 = \int u^2\, K(u)du < \infty$ and b_n a positive real-valued number, called a bandwidth. The kernel density estimator is defined as

$$\hat{f}(x) = \frac{1}{(n-1)b_n} \sum_{t=2}^{n} K\left(\frac{X_{t-1} - x}{b_n}\right) \quad \text{for } x \in \mathbb{R} \tag{7}$$

X_{t-1} is assumed to have a density function $f(x)$ for $x \in [-1, 1]$ which is actually the support of the second order Epanechnikov kernel function

$$K(u) = \begin{cases} \frac{3}{4}(1-u^2) \text{ support } |u| \leq 1 \\ 0 \quad \text{otherwise} \end{cases} \quad \text{where } u = \left(\frac{X_{t-1} - x}{b_n}\right). \tag{8}$$

The Epanechnikov kernel is used in estimating the regression function since it is the most efficient in minimizing the Mean Integrated Squared Error (MISE) and is therefore optimal putting in mind that the choice of the kernel is not as important as the choice of the bandwidth.

The nonparametric estimator of the regression function $m(x) = \mathbb{E}\left[X_t | X_{t-1} = x\right]$ is obtained by

$$\hat{m}(x) = \begin{cases} \frac{\sum_{t=2}^{n} K\left(\frac{X_{t-1}-x}{b_n}\right) X_t}{\sum_{t=2}^{n} K\left(\frac{X_{t-1}-x}{b_n}\right)}, & \text{if } \hat{f}(x) \neq 0 \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

where $K(\cdot)$ is a kernel function and b_n is the bandwidth. $\hat{m}(x)$ is called a kernel estimator or the Nadaraya Watson kernel estimator developed independently by [16] and [17].

The estimator of the conditional variance function $\sigma^2(x) = \text{Var}(X_t | X_{t-1} = x)$ which is obtained by using the residuals $\mathbb{E}\left(\{X_t - m(X_{t-1})\}^2 | X_{t-1} = x\right) = \sigma^2(x)$ is given by

$$\hat{\sigma}^2(x) = \begin{cases} \frac{\sum_{t=2}^{n} K\left(\frac{X_{t-1}-x}{b_n}\right) \{X_t - \hat{m}(X_{t-1})\}^2}{\sum_{t=2}^{n} K\left(\frac{X_{t-1}-x}{b_n}\right)}, & \text{if } \hat{f}(x) \neq 0 \\ 0, & \text{otherwise} \end{cases} \tag{10}$$

The second order Gaussian Kernel function

$$K(u) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}u^2) \ \text{ for } u = \left(\frac{X_{t-1} - x}{b_n}\right), \quad -\infty < u < \infty \tag{11}$$

is employed when estimating the conditional variance function so as to cater for asymmetric behavior of volatility. [18] showed that $\hat{m}(x)$ is a consistent estimator of $m(x)$.

2.2. Single Change Point Model

Under the null hypothesis H_0 of no change in volatility, Equation (6) is written as

$$X_t = m(X_{t-1}) + \sigma_t(X_{t-1}) z_t \quad t = 1, 2, 3, ..., n \tag{12}$$

This in turn implies that under H_0,

$$\mathbb{E}\left\{X_t - m(X_{t-1})\right\}^2 = \sigma_t^2(X_{t-1}) \Longrightarrow \sigma_t^2(X_{t-1}) = \sigma_{(1)}^2(X_{t-1}) \text{ for } t = 1, 2, ..., n \tag{13}$$

The data structure having changed at a certain point in time means that using one regression model to study the data leaves the data poorly explained by the regression model. Hence, in the presence of a change point in volatility, $\tau \in [2, n-1]$, Equation (6) becomes

$$X_t = m(X_{t-1}) + \sigma_t(X_{t-1})z_t \Longrightarrow \mathbb{E}\left\{X_t - m(X_{t-1})\right\}^2 = \sigma_t^2(X_{t-1}) \tag{14}$$

where the alternative hypothesis H_A becomes

$$\sigma_t^2(X_{t-1}) = \begin{cases} \sigma_{(1)}^2(X_{t-1}) & \text{for} \quad t = 1, 2, ...\tau \\ \sigma_{(2)}^2(X_{t-1}) & \text{for} \quad t = \tau + 1, ..., n \end{cases} \tag{15}$$

2.3. Volatility Change Point Test Statistic and Estimator

Define the residuals obtained from non-parametric estimation of conditional mean function and standardized using the conditional variances obtained from the conditional variance function as $\hat{\varepsilon}_t = \frac{X_t - \hat{m}(X_{t-1})}{\hat{\sigma}(X_{t-1})}$ where $\hat{m}(.)$ is the Nadaraya Watson estimator of the unknown regression function $m(.)$.

The sums of squared residuals among the sample segments (partial sums) are defined as

$$\varepsilon_n = \sum_{t=1}^{n} \hat{\varepsilon}_t^2 \quad \varepsilon_\tau = \sum_{t=1}^{\tau} \hat{\varepsilon}_t^2 \quad \varepsilon_{n-\tau} = \sum_{t=\tau+1}^{n} \hat{\varepsilon}_t^2 \tag{16}$$

Define $\bar{\varepsilon}_{1,\tau}$ as the mean of the first τ squared residuals and $\bar{\varepsilon}_{\tau+1,n}$ as the mean of the last $n-\tau$ squared residuals as

$$\bar{\varepsilon}_{1,\tau} = \frac{\varepsilon_\tau}{\tau} = \frac{1}{\tau}\sum_{t=1}^{\tau} \hat{\varepsilon}_t^2 \quad \bar{\varepsilon}_{\tau+1,n} = \frac{\varepsilon_{n-\tau}}{n-\tau} = \frac{1}{n-\tau}\sum_{t=\tau+1}^{n} \hat{\varepsilon}_t^2 \tag{17}$$

The change point test statistic is obtained by considering a weighted l_2 norm of the conditional variance functions and application of the reverse triangle inequality which follows from Minkowski's inequality leading to

$$\begin{aligned} l_2\left(\sigma_{(1)}^2(X_{t-1}) - \sigma_{(2)}^2(X_{t-1})\right) &= \left(\sum_{\tau=1}^{n} w_\tau \left|\varepsilon_\tau - \varepsilon_{n-\tau}\right|^2\right)^{\frac{1}{2}} \leq \left(\sum_{\tau=1}^{n} w_\tau^{\frac{1}{2}} \left|\varepsilon_\tau - \varepsilon_{n-\tau}\right|\right) \\ &= \mathbb{E}\left(w_\tau^{\frac{1}{2}} \left|\varepsilon_\tau - \varepsilon_{n-\tau}\right|\right) \geq w_\tau^{\frac{1}{2}} \left|\mathbb{E}(\varepsilon_\tau) - \mathbb{E}(\varepsilon_{n-\tau})\right| \end{aligned} \tag{18}$$

Therefore, it follows that

$$w_\tau^{\frac{1}{2}} \left|\mathbb{E}(\varepsilon_\tau) - \mathbb{E}(\varepsilon_{n-\tau})\right| = w_\tau^{\frac{1}{2}} \left|\frac{1}{\tau}\sum_{t=1}^{\tau} \hat{\varepsilon}_t^2 - \frac{1}{n-\tau}\sum_{t=\tau+1}^{n} \hat{\varepsilon}_t^2\right| \tag{19}$$

leading to

$$D_t^n = \left(\frac{\tau}{n}\left(1 - \frac{\tau}{n}\right)\right)^{\frac{1}{2}} \left|\frac{1}{\tau}\sum_{t=1}^{\tau} \hat{\varepsilon}_t^2 - \frac{1}{n-\tau}\sum_{t=\tau+1}^{n} \hat{\varepsilon}_t^2\right| \tag{20}$$

The nonparametric statistic Equation (20) so obtained does not utilize any a priori information about the data like marginal densities of the data.

A good choice of the estimator for the change point τ is where the test statistic has a global maximum since the maximum usually occur in the area of the "true " change point. The estimator $\hat{\tau}$ of change point in volatility τ is thus given by

$$\hat{\tau} = \arg\max_{\tau} |D_t^n| \tag{21}$$

Theorem 3 below for the convergence in probability of the change point fraction estimate $\hat{k} = \frac{\hat{\tau}}{n}$ to the "true " change point fraction $\frac{\tau}{n} = k^*$ under H_A is the main result in proving the consistency of the estimated change point fraction.

Theorem 2.2. *Consistency of the change point estimator*: Consider a sample of squared residuals $\hat{\varepsilon}_1^2, \hat{\varepsilon}_2^2, ..., \hat{\varepsilon}_n^2$ satisfying the alternative change point hypothesis and the change point estimator $\hat{\tau}$ given in (21). If the sequences $\left\{\hat{\varepsilon}_{1,t}^2 \; t \in \mathcal{Z}\right\}$ and $\left\{\hat{\varepsilon}_{2,t}^2 \; t \in \mathcal{Z}\right\}$ satisfy

$$\Delta := \sigma_{(1)}^2(X_{t-1}) - \sigma_{(2)}^2(X_{t-1}) \neq 0$$

where $\Delta < \infty$ denotes the finite magnitude of the jump in the conditional variance function , then for $\hat{k} = \frac{\hat{\tau}}{n}$

$$P\left\{|\hat{k} - k^*| > \epsilon\right\} \leqslant \frac{B}{\epsilon^2 \Delta^2} n^{-\frac{1}{2}}$$

where $0 < B < \infty$ (is a positive constant) and $k^* = \frac{\tau}{n}$.

2.4. Limit Distribution of the Change Point Test Statistic

Besides the asymptotic distribution of the test, it was shown to be consistent and more powerful at the tail end of the data than the ordinary cumulative sum of squares test statistics which have more power at the middle of the data.

From Equation (20), the difference $\left|\frac{1}{\tau}\sum_{t=1}^{\tau}\hat{\varepsilon}_t^2 - \frac{1}{n-\tau}\sum_{t=\tau+1}^{n}\hat{\varepsilon}_t^2\right|$ and thus D_t^n is close to 0 under H_0 and different from 0 under H_A. To obtain the critical values, the limit distribution of D_t^n needs to be derived.

Given a test statistic D_t^n, and a critical value $u(a,b)$ which is obtained as the $(1-\alpha)$ quantile of the limit distribution of the test statistic under H_0, the decision rule of the test is to

$$\text{Reject } H_0 \text{ if } \max_{1<\tau<n} |D_t^n| > u(a,b) \tag{22}$$

hence an asymptotic test of level α. The Functional Central Limit Theorem for strong mixing sequences followed by the Continuous Mapping Theorem are applied in obtaining the limit distribution of D_t^n.

Theorem 2.3. *Functional Central Limit Theorem for strong mixing sequences*: Let $\{\xi_t\}\,(1 \le t \le n)$ be a strong mixing sequence of variables which are strictly stationary, centered at expectation, $\mathbb{E}\{\xi_t\} = \mu = 0$ and satisfying the strong mixing condition as described in Theorem (2.1). Suppose further that the process satisfies

$$\mathbb{E}\left|\xi_t\right|^{2+\psi} < \infty \text{ and } \sum_{n=1}^{\infty}\alpha(n)^{\frac{\psi}{2+\psi}} < \infty \text{ for some } \psi \in (0,\infty)$$

Then $\lim_{n\longrightarrow\infty}\frac{\operatorname{var}\left(\sum_{t=1}^{\lfloor nk\rfloor}\xi_t\right)}{n} = \lim_{n\longrightarrow\infty}\frac{\mathbb{E}\left(\sum_{t=1}^{\lfloor nk\rfloor}\xi_t\right)^2}{n} = \sigma_l^2 = \mathbb{E}\left[\xi_1\right]^2 + 2\sum_{l=2}^{\infty}\mathbb{E}\left[\xi_1,\xi_l\right] > 0$.

Define the process

$$W_n(k) = \frac{\sum_{t=1}^{[nk]}\xi_t}{\sigma_l\sqrt{n}} \text{ for } 0 \le k \le 1$$

Then as $n \longrightarrow \infty$, $W_n(k) = \frac{\sum_{t=1}^{[nk]}\xi_t}{\sigma_l\sqrt{n}} \overset{D([0,1])}{\Longrightarrow} W(k)$ for $0 \le k \le 1$ (in the Skorokhod space $D\left([0,1];\mathbb{R}\right)$). $\overset{D([0,1])}{\Longrightarrow}$ signifies weak convergence in the space $D[0,1]$ (since $\mu = 0$) to the Standard Brownian Motion $W(k)$, [19].

The Skorokhod space $\mathcal{D} = D\left([0,1];\mathbb{R}^1\right)$ is the space of real-valued functions (continuous functions) $f : [0,1] \longrightarrow \mathbb{R}^1$ and admits limit $f(k-)$ from the left at every point $k \in (0,1]$ and limit $f(k+)$ from the right at every point $k \in [0,1)$. These functions are everywhere right continous and have left limits everywhere and are called Cadlag funtions.

The logic behind the FCLT usually rely on the convergence of a sequence of standardized partial sums of disturbances to Standard Brownian Motion.

Lemma 2.1. Let $\{\xi_t, t = 1,2,...,n\}$ be a strong mixing sequence and random variables η and υ be measurable with respect to $\sigma-$algebras $\mathcal{F}_{-\infty}^{j}(\xi)$ and $\mathcal{F}_{j+n}^{\infty}(\xi)$ respectively. If the moments $\mathbb{E}|\eta|^p$ and $\mathbb{E}|\upsilon|^q$ exist for $p,q > 1$ where $\frac{1}{p}+\frac{1}{q} < 1$, then

$$|\mathbb{E}\eta\upsilon - \mathbb{E}\eta\mathbb{E}\upsilon| \le C\left[\mathbb{E}|\eta|^p\right]^{\frac{1}{p}}\left[\mathbb{E}|\upsilon|^q\right]^{\frac{1}{q}}\left[\alpha(n)\right]^{1-\frac{1}{p}-\frac{1}{q}}$$

where $p = q = 2+\psi$, then $1-\frac{1}{p}-\frac{1}{q} = 1-\frac{1}{2+\psi}-\frac{1}{2+\psi} = \frac{\psi}{2+\psi}$ for $\psi > 0$, which means that

$$\begin{aligned}|\mathbb{E}\eta\upsilon - \mathbb{E}\eta\mathbb{E}\upsilon| &\le C\left[\mathbb{E}|\eta|^p\right]^{\frac{1}{p}}\left[\mathbb{E}|\upsilon|^q\right]^{\frac{1}{q}}\left[\alpha(n)\right]^{1-\frac{1}{p}-\frac{1}{q}} \\ &= C\left[\mathbb{E}|\eta|^{2+\psi}\right]^{\frac{1}{2+\psi}}\left[\mathbb{E}|\upsilon|^{2+\psi}\right]^{\frac{1}{2+\psi}}\left[\alpha(n)\right]^{\frac{\psi}{2+\psi}}\end{aligned}$$

see, [20]

Lemma (2.1), a consequence of Theorem (2.3), provides a bound on the covariance of ξ and η. It entails that the autocovariance function of an α mixing stationary process (with enough moments) tends to zero.

Theorem 2.4. *Continuous Mapping Theorem*: Let $\{\xi_n\}_{n=0}^{\infty}$ be a sequence of random variables with

$$\xi_n \xrightarrow{d} \xi$$

as $n \longrightarrow \infty$. For every continuous function $g : R \longrightarrow R$ then

$$g(\xi_n) \xrightarrow{d} g(\xi)$$

Also, if

$$\xi_n \xrightarrow{p} \xi, \text{ then } g(\xi_n) \xrightarrow{p} g(\xi)$$

and if

$$\xi_n \xrightarrow{as} \xi, \text{ then } g(\xi_n) \xrightarrow{a.s} g(\xi).$$

A similar result holds for random functions, [21].

The CMT (2.4) ensures invariance of convergence in distribution under continuous transformations and thus emphasizes that the weak convergence property under continuous mappings is stable.

Under H_0, $\mathbb{E}\left(\hat{\varepsilon}_t^2\right) = \sigma^2 = 1$. Similarly, let $\xi_t = \hat{\varepsilon}_t^2 - \sigma^2$ where $\mathbb{E}\left(\hat{\varepsilon}_t^2\right) = \sigma^2 = 1$ so that $\mathbb{E}\xi_t = 0$ under H_0. A general case where the residuals could be dependent and heterogeneously distributed is considered by letting the long-run fourth moment of the residuals be obtained by

$$\begin{aligned}\operatorname{var}(\varepsilon_n) = \operatorname{var}\left[\sum_{t=1}^{n}\hat{\varepsilon}_t^2\right] &= n\left\{\mathbb{E}\left(\hat{\varepsilon}_1^2 - \mathbb{E}\hat{\varepsilon}_1^2\right)^2 + 2\sum_{l=2}^{n}\mathbb{E}\left(\left(\hat{\varepsilon}_1^2 - \mathbb{E}\hat{\varepsilon}_1^2\right)\left(\hat{\varepsilon}_l^2 - \mathbb{E}\hat{\varepsilon}_l^2\right)\right)\right\} \\ &\quad - 2\sum_{l=2}^{n} l\,\mathbb{E}\left(\left(\hat{\varepsilon}_1^2 - \mathbb{E}\hat{\varepsilon}_1^2\right)\left(\hat{\varepsilon}_l^2 - \mathbb{E}\hat{\varepsilon}_l^2\right)\right)\end{aligned} \tag{23}$$

which as the sample size increases, becomes

$$\begin{aligned} \underset{n\longrightarrow\infty}{lim}\frac{\text{var}\,(\varepsilon_n)}{n} &\longrightarrow\ \text{var}\left[\hat{\varepsilon}_1^2\right]+2\sum_{l=2}^{\infty}\text{cov}\left[\hat{\varepsilon}_1^2,\hat{\varepsilon}_l^2\right] \\ &=\mathbb{E}\left(\hat{\varepsilon}_1^2-\mathbb{E}\hat{\varepsilon}_1^2\right)^2+2\sum_{l=2}^{\infty}\mathbb{E}\left(\left(\hat{\varepsilon}_1^2-\mathbb{E}\hat{\varepsilon}_1^2\right)\left(\hat{\varepsilon}_l^2-\mathbb{E}\hat{\varepsilon}_l^2\right)\right) \\ &=\sigma_l^2<\infty \end{aligned} \tag{24}$$

which is the long-run fourth moment of the residuals. σ_l^2 is used in rescaling the test-statistic for convergence results to be achieved. Due to the fact that σ_l^2 is not known, it is estimated using Bartlett kernel. When H_0 is true, then as $n\longrightarrow\infty,\ D_t^n\longrightarrow 0$ at a rate $\frac{1}{\sqrt{n}}$ and thus the statistic needs to be normalized with $\sqrt{n}$ to become $\sqrt{n}D_t^n$ which is now written as

$$\begin{aligned} \frac{\sqrt{n}}{\sigma_l}\left|D_t^n\right| &= \frac{1}{\sigma_l}\left[\frac{n^2}{\tau\,(n-\tau)}\right]^{\frac{1}{2}}\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{\tau}\hat{\varepsilon}_t^2-\frac{\tau}{n}\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\hat{\varepsilon}_t^2\right)\right| \\ &= \frac{\frac{1}{\sigma_l}\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{\tau}\hat{\varepsilon}_t^2-\frac{\tau}{n}\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\hat{\varepsilon}_t^2\right)\right|}{\left[\frac{\tau}{n}\left(1-\frac{\tau}{n}\right)\right]^{\frac{1}{2}}} \end{aligned} \tag{25}$$

By application of Theorem (2.3) putting in mind that $\xi_t=\hat{\varepsilon}_t^2-\sigma^2$ where $\mathbb{E}\left(\hat{\varepsilon}_t^2\right)=\sigma^2=1$ one obtains,

$$\frac{\sqrt{n}}{\sigma_l}\left|D_t^n\right| = \frac{\frac{1}{\sigma_l}\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{\tau}\xi_t-\frac{\tau}{n}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\xi_t\right|}{\left[\frac{\tau}{n}\left(1-\frac{\tau}{n}\right)\right]^{\frac{1}{2}}} \tag{26}$$

The convergence in distribution of the test statistic follows immediately from the weak convergence of the partial sum process of the errors ξ_t to SBM in the unit interval $\mathcal{D}[0,1]$ following Theorem (2.3).

Theorem 2.5. For a weakly consistent estimator of σ_l^2, under H_0, strict stationary $\alpha-$mixing process $\{\xi_t\}$ setting $\tau=\lfloor nk\rfloor$,

$$\lim_{n\longrightarrow\infty}\frac{\sqrt{n}}{\sigma_l}\left|D_t^n\right| = \lim_{n\longrightarrow\infty}\frac{\frac{1}{\sigma_l}\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{\tau}\xi_t-\frac{\tau}{n}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\xi_t\right|}{\left[\frac{\tau}{n}\left(1-\frac{\tau}{n}\right)\right]^{\frac{1}{2}}}\ \overset{d}{\longrightarrow}\ \frac{|W(k)-kW(1)|}{(k(1-k))^{\frac{1}{2}}}=\frac{|B(k)|}{(k(1-k))^{\frac{1}{2}}}$$

where $\{|B(k)|\ 0\leq k\leq 1\}$ is the absolute value of the Standard Brownian Bridge on the unit interval, $(k(1-k))^{\frac{1}{2}}$ is the asymptotic standard deviation of the Standard Brownian Bridge process. $\frac{|B(k)|}{(k(1-k))^{\frac{1}{2}}}$ is a standardized tied down Bessel process restricted to $0\leq k\leq 1$ [22].

Further mathematical derivations, scale transformation and application of Theorem (2.4) yields,

$$\begin{aligned} \lim_{n\longrightarrow\infty}\max_{1<\tau<n}\frac{\sqrt{n}}{\sigma_l}\left|D_t^n\right| &= \lim_{n\longrightarrow\infty}\max_{1<\tau<n}\frac{\left|\frac{1}{\sigma_l\sqrt{n}}\sum_{t=1}^{\tau}\xi_t-\frac{\tau}{n}\left(\frac{1}{\sigma_l\sqrt{n}}\sum_{t=1}^{n}\xi_t\right)\right|}{\left[\frac{\tau}{n}\left(1-\frac{\tau}{n}\right)\right]^{\frac{1}{2}}} \\ &= \lim_{n\longrightarrow\infty}\sup_{0<k<1}\frac{\left|\frac{1}{\sigma_l\sqrt{n}}\sum_{t=1}^{\lfloor n.k\rfloor}\xi_t-\frac{n.k}{n}\left(\frac{1}{\sigma_l\sqrt{n}}\sum_{t=1}^{\lfloor n.1\rfloor}\xi_t\right)\right|}{\left(k\left(1-k\right)\right)^{\frac{1}{2}}} \end{aligned} \tag{27}$$

which finally leads to convergence in distribution to Standard Brownian Bridge process as

$$\lim_{n\longrightarrow\infty}\frac{\sqrt{n}}{\sigma_l}\sup_{\eta n\leq\tau\leq(1-\eta)n}\left|D_t^n\right|\ \overset{d}{\longrightarrow}\ \sup_{\eta\leq k\leq 1-\eta}\frac{|W(k)-kW(1)|}{(k(1-k))^{\frac{1}{2}}}=\sup_{\eta\leq k\leq 1-\eta}\frac{|B(k)|}{(k(1-k))^{\frac{1}{2}}} \tag{28}$$

2.5. Tail Probabilities Approximations and Critical Values

Theorem 2.6. If H_0 holds and $\xi_t, t=1,2,...,n$ is a strong mixing sequence, then as $n\longrightarrow\infty$,

$$\left|\frac{\sup_{\eta n<\tau<(1-\eta)n}\left|\frac{1}{\sigma_l\sqrt{n}}\sum_{t=1}^{\lfloor n.k\rfloor}\xi_t-\frac{n.k}{n}(\frac{1}{\sigma_l\sqrt{n}}\sum_{t=1}^{\lfloor n.1\rfloor}\xi_t)\right|}{\left(k\left(1-k\right)\right)^{\frac{1}{2}}}-\sup_{\eta\leq k\leq 1-\eta}\frac{|B(k)|}{(k(1-k))^{\frac{1}{2}}}\right|=o_P(\exp(-(\log n)^{1-\epsilon}))$$

$\forall\, 0 < \epsilon < 1$ where $\sigma_l^2 = \lim\limits_{n \longrightarrow \infty} n^{-1} \text{Var}\,(\varepsilon_n)$ and $|B(k)|$ is the absolute value of the Standard Brownian Bridge process on the truncated interval $[\eta, 1-\eta] \subset [0,1]$ for some $\eta \in \left(0, \frac{1}{2}\right)$ [23].

The statements below generalize [23] to the mixing case. To construct rejection region, given a fixed level of statistical significance $0 < \alpha < 1$, one considers the quantiles

$$\begin{aligned} q_n =& q_n\,(1-\alpha) \\ &= \sup \left\{ x \geq 0 : P\left(\frac{\sqrt{n}}{\sigma_l} \sup_{\eta \leq k \leq 1-\eta} |D_t^n| < x \right) \leq 1 - \alpha \right\} \end{aligned} \tag{29}$$

and

$$\begin{aligned} u(a,b) &= u(a,b;1-\alpha) \\ &= \sup \left\{ x \geq 0 : P\left(\sup_{a \leq k \leq 1-b} \frac{|B(k)|}{(k(1-k))^{\frac{1}{2}}} < x \right) = 1 - \alpha \right\} \end{aligned} \tag{30}$$

Rejection of the null hypothesis happens if the $1-\alpha$ quantile of the limit distribution of the test statistic under H_0 has been exceeded.

One can show that $u(a,b)$ is an asymptotically correct size α critical value according to Theorem (2.7)

Theorem 2.7. As in [23] corollary 1.3.1, Theorem 1.3.2 and Theorem 1.3.3 with

$$a_n \geq \frac{1}{n}, \;\; b_n \geq \frac{1}{n} \text{ and for } n \longrightarrow \infty$$

and with

$$\lim_{n \longrightarrow \infty} \sup n\,(a_n + b_n) \exp\left(-\,(\log\; n)^{1-\epsilon^*} \right) < \infty$$

for some $0 < \epsilon^* \leq 1$, then as $n \longrightarrow \infty$ one has that

$$\lim_{n \longrightarrow \infty} P \left\{ \frac{\sqrt{n}}{\sigma_l} \sup_{\eta \leq k \leq 1-\eta} |D_t^n| > u\,(a_n, b_n) \right\} = \alpha$$

and

$$|q_n(1-\alpha) - u\,(a_n, b_n)| = o\left((\log\log\; n)^{-\frac{1}{2}} \right) \qquad n \longrightarrow \infty$$

Contrasting with tests for known break points which normally have asymptotic chi-square distribution, the asymptotic distribution associated with tests for unknown break points is the Bessel process distributions. The Bessel process distributions pose some challenges to researchers one being that there there is no known simple closed-form expression for the distribution function which makes it difficult to compute exact p-values. Also, the Bessel process distributions depend on two parameters which makes it hard to tabulate the critical values hence one needs to simulate them.

For practical applications, the distribution function of $\sup\limits_{a \leq k \leq 1-b} \left\{ \frac{|B(k)|}{(k(1-k))^{\frac{1}{2}}} \right\}$, is approximated by its inverse laplace transform where for each $0 < a < b < 1$, under no change point, it is given by

$$P\left\{ \sup_{a \leq k \leq 1-b} \left(\frac{B(k)}{k(1-k)} \right)^{\frac{1}{2}} \geq x \right\} = \frac{x \exp(-\frac{x^2}{2})}{(2\pi)^{\frac{1}{2}}} \left\{ \log \frac{(1-a)\,(1-b)}{ab} - \frac{1}{x^2} \log \frac{(1-a)\,(1-b)}{ab} + \frac{4}{x^2} + O\left(\frac{1}{x^4} \right) \right\} \tag{31}$$

as $x \longrightarrow \infty$ [23]. To obtain the asymptotic critical value of size α, Monte Carlo simulation of the distribution of the limit variable is performed so as to obtain a roughly good approximation for the quantile $q_n = q_n(1-\alpha)$.

2.6. Simulation of Asymptotic Critical Values

Asymptotic critical values corresponding to the 0.99, 0.95 and 0.90 quantiles of the limit distribution of $\max_\tau \frac{\sqrt{n}}{\sigma_l} |D_t^n|$ which allow tests corresponding to significance levels 0.01, 0.05, 0.10 respectively were simulated. The upper boundary can be set at 3.541899 in the $\frac{\sqrt{n}}{\sigma_l} |D_t^n|$ plot when $n = 100$ and $\alpha = 0.01$. If this boundary is exceeded, then there is a significant change in conditional variance function of the returns series.

The simulated asymptotic critical values from Equation (31) are as in Table 1.

Table 1. Simulated critical values.

Sample size	$1-\alpha$	$u(a,b)$
100	0.99	3.541899
	0.95	2.969163
	0.90	2.658421
200	0.99	3.613201
	0.95	3.063084
	0.90	2.774228
500	0.99	3.686959
	0.95	3.156928
	0.90	2.885142
1000	0.99	3.732673
	0.95	3.213796
	0.90	2.950778
2000	0.99	3.772292
	0.95	3.262439
	0.90	3.006221
4000	0.99	3.807243
	0.95	3.304918
	0.90	3.054156

2.7. Discussion on the Simulated Critical Values

From results in Table 1, holding α constant, the critical values increases as the sample size increases. For example, when $\alpha = 0.01$ as $n \longrightarrow \infty$, the critical values increase tending to converge at 3.8. When $\alpha = 0.05$ as $n \longrightarrow \infty$, the critical values seem to converge at 3.3 and when $\alpha = 0.1$ as $n \longrightarrow \infty$, the critical values seem to converge at 3.0. As α increases for a given sample size, the rejection region for the test increases meaning the critical value reduces.

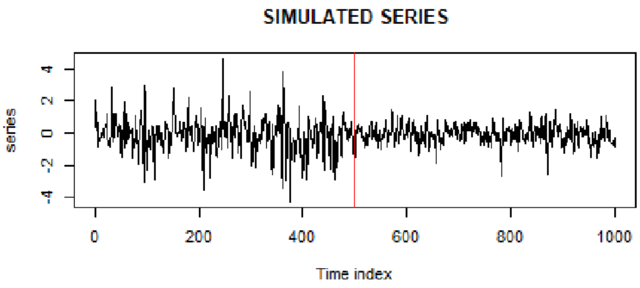
2.8. Simulation Study in the Presence of a Single Change Point

The ability of the test to detect a change point was investigated by simulating an ARMA$(1,1)-$ARCH(1) model.

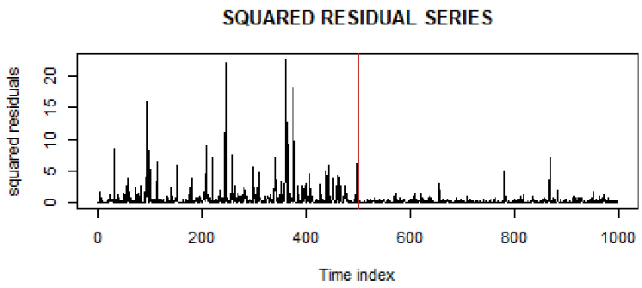
$$\sigma_t^2(X_{t-1}) = \begin{cases} 0.5 + 0.1\epsilon_{t-1}^2 \text{ for } t = 1, 2, ..., \tau \\ 0.1 + 0.1\epsilon_{t-1}^2 \text{ for } t = \tau + 1, ..., n \end{cases} \tag{32}$$

$$X_t = 0.35X_{t-1} + \epsilon_t + 0.4\epsilon_{t-1}, \quad \epsilon_t = \sigma_t(X_{t-1})z_t \quad z_t \sim i.i.d \text{ Normal } (0,1)$$

For the first case, the change point was fixed at $\frac{1}{2}n = 500$.



(a) Simulated series



(b) Squared residuals

Figure 1. *Simulated series and squared residual series.*

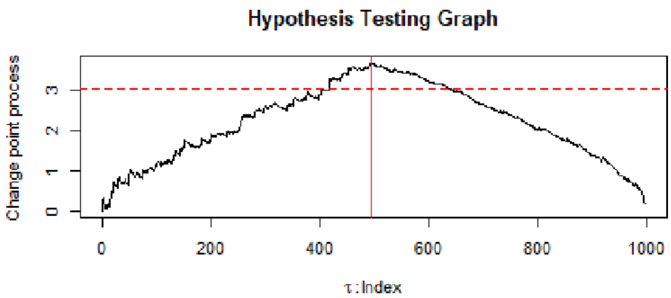


Figure 2. *The sequence* $\frac{\sqrt{n}}{\sigma_l}|D_t^n|$ *for change point located at* $\frac{1}{2}n = 500$ *and at* $\alpha = 0.05$.

The change point testing graph with max$\frac{\sqrt{n}}{\sigma_l}|D_t^n| = 3.828467$ was as shown in Figure 2 where the dashed line indicates the point of asymptotic critical value of 3.213796. Hence H_0 was rejected at 0.05 significant level. H_0 was also rejected at 0.1 and 0.01 significant levels.

2.9. Simulation Study in the Absence of a Change Point

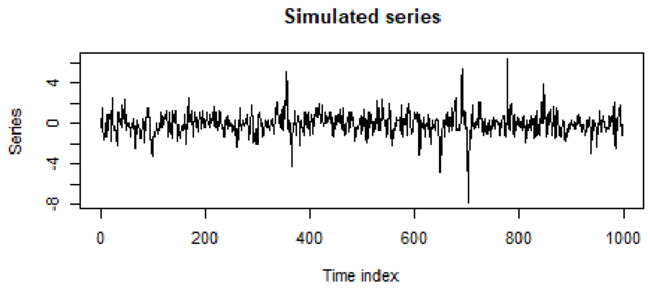
Further simulations were done to test for a change when it was actually not present. An $\text{ARMA}(1,1)-\text{ARCH}(1)$ model with no change point in conditional variance function

$$\sigma_t^2(X_{t-1}) = 0.5 + 0.1\epsilon_{t-1}^2 \text{ for } t = 1, 2, ..., n \tag{33}$$

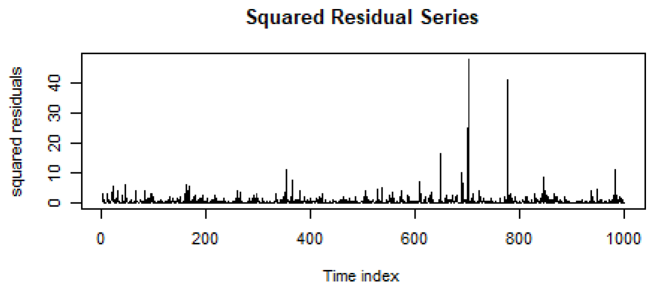
where

$$X_t = 0.35X_{t-1} + \epsilon_t + 0.4\epsilon_{t-1}, \;\; \epsilon_t = \sigma_t(X_{t-1})z_t \quad z_t \sim \text{Normal}(0,1).$$

was considered. A series of length $n = 1000$ was simulated and the ability of the test not to detect a change point when it was actually not there investigated.



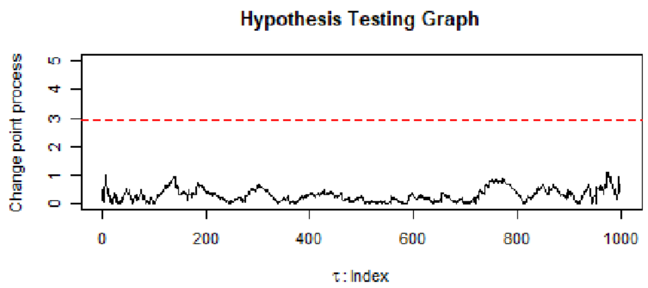
(a) Simulated series



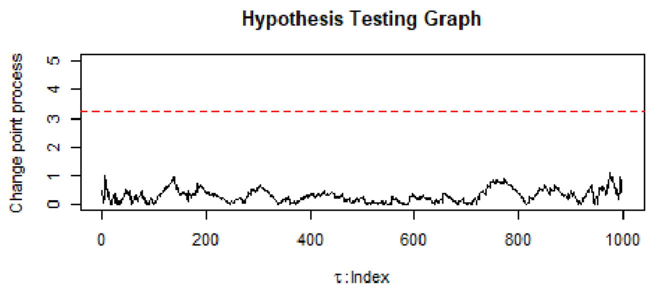
(b) Squared residuals

Figure 3. *Simulated series and squared residual series in the absence of a change point.*

The simulated series for the model under no change point were as shown in Figures 3a and 3b respectively.



(a) The sequence $\frac{\sqrt{n}}{\sigma_l}|D_t^n|$, $\alpha = 0.1$, $n = 1000$



(b) The sequence $\frac{\sqrt{n}}{\sigma_l}|D_t^n|$, $\alpha = 0.05$, $n = 1000$

Figure 4. *Change point testing graphs in the absence of a change point.*

The change point testing graphs when there was no change point were as shown in Figures 4a and 4b for $\alpha = 0.1$ and $\alpha = 0.05$. The dashed line indicates the asymptotic critical value point of 2.950778 at 0.1 significant level and 3.213796 at 0.05 significant level.

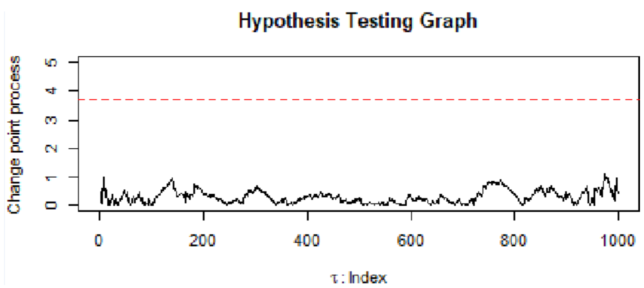


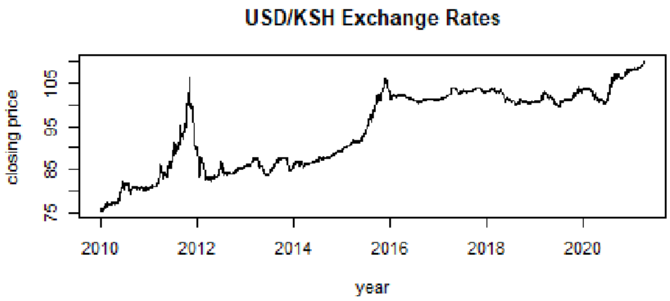
Figure 5. *The sequence* $\frac{\sqrt{n}}{\sigma_l}|D_t^n|$, $\alpha = 0.01$, $n = 1000$.

The change point testing graphs in the absence of a change point at 0.01 significant level was as shown in Figure 5 where the dashed line indicates the point of asymptotic critical value of 3.732673. The test failed to reject H_0 (failed to detect a change point) at 0.1, 0.05 and 0.01 significant levels. Hence, in the absence of a change point in the conditional variance function, the plot of $\frac{\sqrt{n}}{\sigma_l}|D_t^n|$ against τ remains near 0. Hence the test correctly failed to reject the null hypothesis when it was not supposed to be rejected.

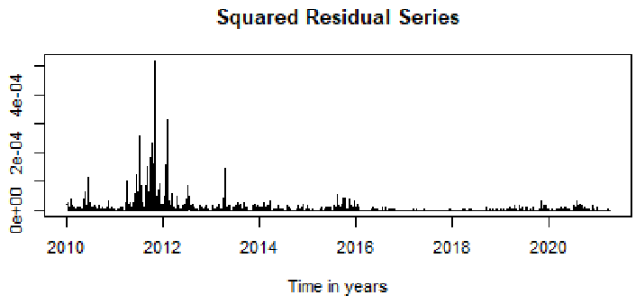
3. Application to Foreign Exchange Rate Data

The existence of a change point in the conditional variance of USD/KSH daily returns from 1^{st} January 2010 to 27^{th}

November 2020 is investigated where $n = 2839$ historical exchange rates. This means obtaining 2838 continuously compounded returns. Lagging the returns by one implies having 2837 continuously compounded returns at lag one.



(a) Daily exchange rate

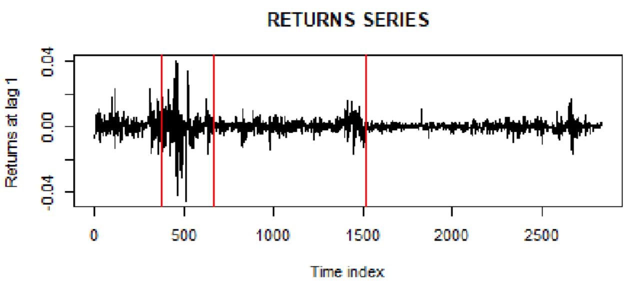


(b) Plot of the squared residuals

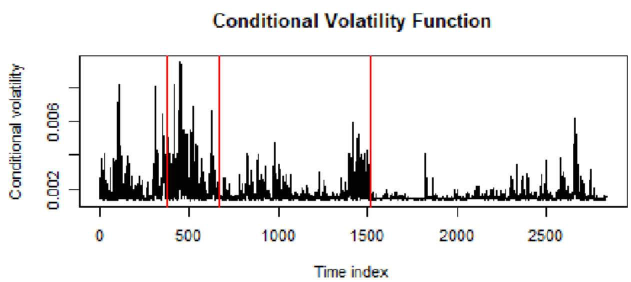
Figure 6. *Exchange rate data and squared residual series.*

The exchange rates are plotted as in Figure 6a above. Since the conditional variances are obtained via the conditional expectation of the squared residuals, the squared residuals plot is as shown in Figure 6b above.

The first change point was detected and estimated as $\hat{\tau}_1 = 668$ corresponding to 26^{th} July 2012. Through binary segmentation, the data is then split into two segments; from $[1 : 668]$ and from $[669 : 2837]$ and change point detection and estimation done on both segments. From the first segment $t \in [1 : 668]$, the second change point in volatility is detected and estimated as $\hat{\tau}_2 = 375$ corresponding to 13^{th} June 2011. From the second segment $t \in [669 : 2837]$, the third change point in volatility of returns is detected and estimated at data point 850 meaning $\hat{\tau}_3 = 1519$ corresponding to 30^{th} October 2015. Further segmentation of the series did not show any other significant change point.



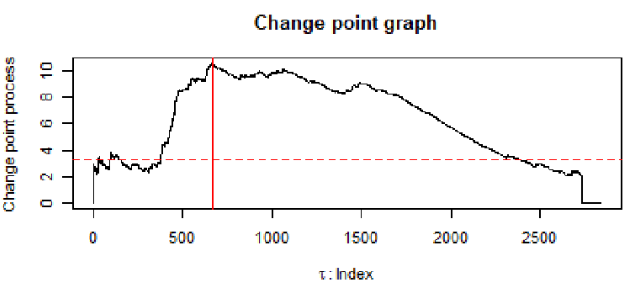
(a) Three estimated change points, $\hat{\tau}_1 = 668$ and $\hat{\tau}_2 = 375$ and $\hat{\tau}_3 = 1519$



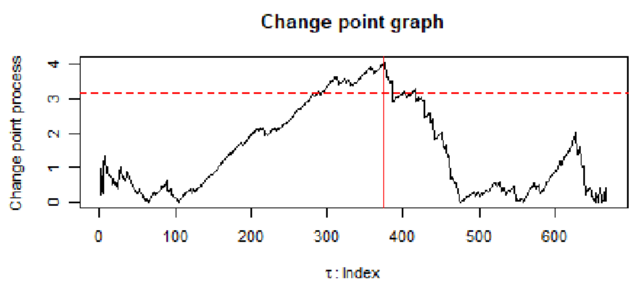
(b) Conditional volatility function, $\hat{\tau}_1 = 668$, $\hat{\tau} = 375$ and $\hat{\tau}_3 = 1519$

Figure 7. *Returns series at lag 1 and the conditional volatility function with three estimated change points.*

The returns series and the conditional volatility function with three estimated change points are shown in Figures 7a and 7b respectively.



(a) The sequence $\frac{\sqrt{n}}{\sigma_l}\,|D_t^n|$ for $t \in [1 : 2837]$ implying $n = 2837$, $\hat{\tau}_1 = 668$



(b) The sequence $\frac{\sqrt{n}}{\sigma_l}\,|D_t^n|$ for $t \in [1 : 668]$ implying $n = 668$, $\hat{\tau}_2 = 375$

Figure 8. *Hypothesis Testing Graphs for $n = 2837$ and $n = 668$ corresponding to (a) and (b) respectively.*

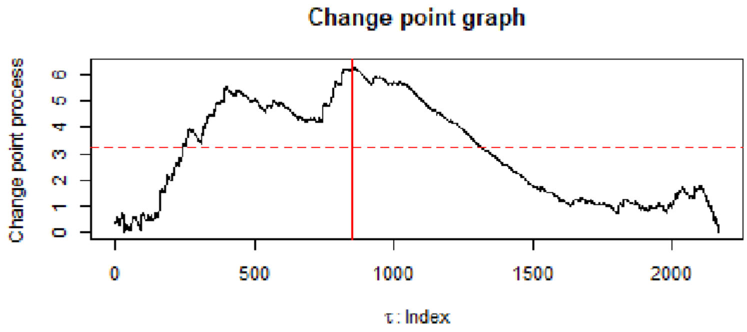


Figure 9. *The sequence* $\frac{\sqrt{n}}{\sigma_l}\,|D_t^n|$ *for* $t \in [669:2837]$ *implying* $n = 2169$, $\hat{\tau}_3 = 850$ *which further represents data point* 1519 *if one considers the original returns series from* $t \in [1:2837]$.

The corresponding plot of the hypothesis testing graph with change points estimates $\hat{\tau}_1 = 668$ and where the red dashed line indicates the boundary corresponding to a critical value of 3.284537 at $\alpha = 0.05$, $\hat{\tau}_2 = 375$ and where the dashed line indicates the boundary corresponding to a critical value of 3.181871 at $\alpha = 0.05$ and $\hat{\tau}_3 = 850$ (where in the original series corresponds to τ index 1519) where the red dashed line indicates the boundary corresponding to a critical value of 3.267697 at $\alpha = 0.05$ are as shown in Figures 8a, 8b and 9 respectively.

4. Results Based on the Real Dataset Application

From the exchange rate plot there is an increasing trend between January 2010 up to October 2011 where the exchange rate prices were at the peak. The USD/KSH exchange rate depreciated from 83.89 to 101.39 between April 2011 to October 2011. After October 2011, there is a decreasing trend after which the exchange rates started to rise again around the year 2014 up to the year 2016. From the year 2016, the USD/KSH seems not to be varying very much. From the year 2020, the USD/KSH exchange rate seems to be taking an upward trend probably due to economic crisis brought about by COVID-19 pandemic like loss of jobs and ban on international travels by governments in most countries of the world.

Accounting for the change point in June 2011, Kenya was importing too much and exporting too little (imbalanced economy) which made it vulnerable to shocks. In 2011, imports soared because of higher costs of fuel and food while exports were stagnant which led to increase in demand for foreign exchange to finance imports. The foreign exchange market witnessed significant volatility between May 2011 and October 2011 as seen in exchange rate plot Figure 6a and the return series 7a reflecting the general volatility in the global financial markets. This resulted in the weakening of the Kenyan Shilling just like other currencies in the region and other global markets whereby the exchange rate for the Kenya shilling against the US dollar depreciated from an average of 84.2 in March 2011 to 101.39 in October 2011 (20.42% depreciation).

Accounting for the change point in volatility of the returns on 26^{th} July 2012, the Kenyan economy experienced slow growth at the beginning of 2012 following high inflation and high interest rates from commercial banks. For the change point on 30^{th} October 2015, the Kenya shilling depreciated against the United States Dollar in the financial year ended June 2015 due to tightening of global financial market conditions and further continued to depreciate in the financial year starting July 2015. The depreciation was further aggregated by increase in food prices due to delayed rains in the same period up to December 2015.

5. Conclusion and Recommendations

In this study, a non-parametric procedure for detecting and estimating a change point in volatility of financial returns is considered. The procedure allows for change point detection and estimation in sequences with conditional heteroscedastic variances and fourth moment. The limit distribution of the test statistic is obtained and critical values simulated. The power of the test and consistency of the test are shown in subsequent papers.

The change point detection and estimation approach can be applied to multidimensional non-parametric models of the form $X_t = m(X_{t-1}, ..., X_{t-d}) + \sigma(X_{t-1}, ..., X_{t-d})$ where the volatility function is changing. In such cases, the functions $m(.)$ and $\sigma(.)$ should be estimated using multivariate kernel methods. One can also consider a case where both the conditional mean and the conditional variance functions are changing. This should be done carefully due to the curse of dimensionality problem.

Acknowledgements

The authors greatly acknowledge Pan African University Institute for Basic Sciences, Technology, and Innovation (PAUSTI) in collaboration with African Development Bank and African Union for funding this research.

The USD/KSH exchange rate data set used to support the findings of this study is available from The Central Bank of Kenya https://www.centralbank.go.ke.

References

[1] B. Chen and Y. Hong, “Detecting for smooth structural changes in garch models,” Econometric Theory, vol. 32, no. 3, pp. 740-791, 2016.

[2] T. Bollerslev, R. F. Engle, and D. B. Nelson, “Arch models,” Handbook of econometrics, vol. 4, pp. 2959-3038, 1994.

[3] C. Inclan and G. C. Tiao, “Use of cumulative sums of squares for retrospective detection of changes of varinace,” Journal of the American Statistical Association, vol. 89, no. 427, pp. 913-923, 1994.

[4] G. Chen, Y. K. Choi, and Y. Zhou, “Nonparametric estimation of structural change points in volatility models for time series,” Journal of Econometrics, vol. 126, no. 1, pp. 79-114, 2005.

[5] G. J. Ross, “Modelling financial volatility in the presence of abrupt changes,” Physica A: Statistical Mechanics and its Applications, vol. 392, no. 2, pp. 350-360, 2013.

[6] S. Aminikhanghahi and D. J. Cook, “A survey of methods for time series change point detection,” Knowledge and information systems, vol. 51, no. 2, pp. 339-367, 2017.

[7] C. Gerstenberger, “Robust wilcoxon-type estimation of change-point location under shortrange dependence,” Journal of Time Series Analysis, vol. 39, no. 1, pp. 90-104, 2018.

[8] J. Chen and A. K. Gupta, Parametric statistical change point analysis: with applications to genetics, medicine, and finance. Springer Science & Business Media, 2011.

[9] E. Brodsky and B. S. Darkhovsky, Nonparametric methods in change point problems, vol. 243. Springer Science & Business Media, 2013.

[10] S. Lee, O. Na, and S. Na, “On the cusum of squares test for variance change in nonstationary and nonparametric time series models,” Annals of the Institute of Statistical Mathematics, vol. 55, no. 3, pp. 467-485, 2003.

[11] A. Sansó, J. Carrion, and V. Aragó, “Testing for changes in the unconditional variance of financial time series,” Revista de Economía Financiera, 2004, vol. 4, p. 32-52, 2020.

[12] I. Berkes, E. Gombay, L. Horváth, and P. Kokoszka, “Sequential change-point detection in garch (p, q) models,” Econometric theory, vol. 20, no. 6, pp. 1140-1167, 2004.

[13] A. Koubková, Sequential change-point analysis. PhD thesis, 2006.

[14] C. Perna and M. Sibillo, Mathematical and statistical methods in insurance and finance. Springer, 2008.

[15] M. Rosenblatt, “A central limit theorem and a strong mixing condition,” Proceedings of the National Academy of Sciences of the United States of America, vol. 42, no. 1, p. 43, 1956.

[16] E. A. Nadaraya, “On estimating regression,” Theory of Probability & Its Applications, vol. 9, no. 1, pp. 141-142, 1964.

[17] G. S. Watson, “Smooth regression analysis,” Sankhyā: The Indian Journal of Statistics, Series A, pp. 359-372, 1964.

[18] J. Fan and Q. Yao, “Efficient estimation of conditional variance functions in stochastic regression,” Biometrika, vol. 85, no. 3, pp. 645-660, 1998.

[19] H. Oodaira and K.-i. Yoshihara, “Functional central limit theorems for strictly stationary processes satisfyinc the strong mixing condition,” in Kodai Mathematical Seminar Reports, vol. 24, pp. 259-269, Department of Mathematics, Tokyo Institute of Technology, 1972.

[20] P. Hall and C. Hyde, Martingale limit theory and its application. Academic Press, 1980.

[21] A. W. Van der Vaart, Asymptotic statistics, vol. 3. Cambridge university press, 2000.

[22] S. Schwaar, Asymptotics for change point tests and change point estimators. PhD thesis, 2017.

[23] M. Csorgo and L. Horváth, Limit theorems in change-point analysis. John Wiley & Sons Chichester, 1997.