12-2012

Asymptotics for the Arc Length of a Multivariate Time Series and Its Applications as a Measure of Risk

Tharanga Wickramarachchi
Clemson University, twickra@g.clemson.edu

Asymptotics for the Arc Length of a Multivariate Time Series and Its Applications as a Measure of Risk.

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Mathematical Sciences

by
Tharanga Wickramarachchi
December 2012

Accepted by:
Dr. Colin Gallagher, Committee Chair
Dr. Robert Lund
Dr. Peter Kiessler
Dr. Jim Brannan

Abstract

The necessity of more trustworthy methods for measuring the risk (volatility) of financial assets has come to the surface with the global market downturn This dissertation aims to propose sample arc length of a time series, which provides a measure of the overall magnitude of the one-step-ahead changes over the observation time period, as a new approach for quantifying the risk. The Gaussian functional central limit theorem is proven under finite second moment conditions. With out loss of generality we consider equally spaced time series when first differences of the series follow a variety of popular stationary models including autoregressive moving average, generalized auto regressive conditional heteroscedastic, and stochastic volatility. As applications we use CUSUM statistic to identify changepoints in terms of volatility of Dow Jones Index returns from January, 2005 through December, 2009. We also compare asset series to determine if they have different volatility structures when arc length is used as the tool of quantification. The idea is that processes with larger sample arc lengths exhibit larger fluctuations, and hence suggest greater variability.

Dedication

To my beloved parents and brother.

Acknowledgments

I would like to express my gratitude to many individuals who helped me in many ways during this work.

First and my foremost, I am indebted to my advisor Dr. Colin Gallagher for his guidance and support. I thank him for giving me the confidence to tackle difficult problems in my dissertation. His knowledge and experience in statistics helped me immensely to grow up as a researcher. I also highly appreciate the strength he gave me at every step of my graduate student life.

I thank Dr. Robert Lund, mainly for coming up with a precious research idea that paved the way to my success as a researcher. The time and energy he offered to shape my work up, when it was badly needed, is always remembered.

I would also like to thank the Department of Mathematical Sciences for providing me financial support during my stay at Clemson.

Finally, I thank my parents and brother for giving me unconditional love and support during my entire graduate life. Without their support I couldn't have done this.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

When the recent economic crisis hit markets world wide, most of the stock indices and asset prices started to fluctuate rapidly. As a result, financial markets became hard to predict. Being able to predict the behavior of stocks, plays a significant role when one invests in stock markets. Rapid fluctuations indicate the instability of respective stocks and bring a huge risk on investments which involve them. Latest bad developments in financial sector, raised eye brows regarding the methods that have been using to measure the risk involved with an asset series for a quite some time.

Since the volatility of an asset series and the risk involved with it are positively related, a measurement for volatility will always give an idea about the risk. Therefore modeling and measuring volatility of financial assets is important to risk managers and is a necessary component of derivative pricing.

In finance practitioners tend to focus on log returns, which is essentially the first difference of log prices. That is, if asset price at time t is denoted by P_t, $Y_t = \ln P_t - \ln P_{t-1}$ represents log returns. It is known that most of the asset returns are relatively small and a Taylor expansion of natural logarithm around "one" gives

$$Y_t = \ln\left(P_t/P_{t-1}\right) = \ln\left(1 + \frac{P_t - P_{t-1}}{P_{t-1}}\right).$$

Based on the above form, clearly log returns, Y_t can be approximated by $\left(P_t - P_{t-1}\right)/P_{t-1}$, which is the percent return at time t. The famous Black-Scholes option pricing formula is developed based on the assumption that $\{\ln(P_t)\}$ follows a Brownian motion. Log returns are the common base for most of the commonly used volatility models. In the field, log-returns are considered as the baseline transformation since those tend to show a set of common characteristics, which are known as $stylized\ facts$ that asset prices do not reveal.

1. Leptokurtosis: asset returns have a density with heavier tails than those of the normal distribution.

2. Persistence: large (small) absolute returns tend to be followed by large (small) absolute returns.

For more on stylized facts, see Taylor (2005).

Main results discussed in chapter 2 will be verified for three classes of univariate and multivariate versions of volatility models. Proper definitions of those models and certain conditions they have to satisfy to hold some nice properties are stated in chapter 3. Therefore without going in the direction of modeling volatility, we will directly move on to methods of measuring it.

Volatility is typically quantified in terms of fluctuations of the investment asset return, often in terms of sample variances. Even though the asymptotic quantification of the sample variances largely depends on finite fourth moments it is well known that most of the financial series do not satisfy the above requirement.

There are two non-parametric volatility measures commonly used in the area of financial time series, namely squared values and absolute values of log returns, $\{Y_t\}$. Asymptotic theory had been developed for partial sums $n^{-1}\sum_{t=1} Y_t^2$ and/or $n^{-1}\sum_{t=1}^{n} |Y_t|$ as $n \to \infty$ and widely employed in statistical inference for these quantities. In order to prove the Functional Central Limit Theorem (FCLT) we always make an assumption on the existence of process moments to some order. For example, if log returns follow a stationary generalized autoregressive conditional heteroscedastic (GARCH) process, Berkes et al. (2004a) proved a FCLT for Y_t^2 under the assumption that $E|Y_0|^{8+\delta} < \infty$ using weak dependence concepts of Doukhan and Louhichi (1999). As stated in Berkes et al. (2008), a result in Doukhan and Wintenberger (2007) implies a FCLT when $E\left[|Y_0|\right]^{4+\delta} < \infty$ for some $\delta > 0$. Giraitis et al. (2007) deduce a FCLT for the partial sums processes of $\{Y_t^2\}$ when $E\left[Y_0^4\right] < \infty$. In fact if log returns follow a GARCH process we can deduce a FCLT for absolute values under finite second moment conditions $\left(E\left[Y_0\right]^2 < \infty\right)$ using a FCLT proved for $\{|Y_t|^{\delta}\}$ inBerkes et al. (2008), where $\delta > 0$. Given the leptokurtosis of returns, it is desirable to prove a FCLT under relaxed moment conditions. In fact, empirical evidence indicates that many asset return series have finite second moments, but infinite fourth moments (e.g., Cont (2001)).

As an effort to improve quantifying methods by overcoming previously stated higher moment condition requirements, we propose the sample arc length of a time series as a new tool. Most importantly, we prove the functional central limit theorem (FCLT) under finite second moment conditions on first differences of the series.

Let $\{X_{t_j}\}_{j=1}^{n}$ be a univariate time series observed at the time points $t_1 < t_2 < \cdots < t_n$. The k-step-ahead sample arc length, which is denoted by A_n, is

$$A_n^k = \sum_{j=2}^{n} \sqrt{(t_j - t_{j-k})^2 + \left(X_{t_j} - X_{t_{j-k}}\right)^2}.$$

With out loss of generality, for the simplicity we study the one step ahead arc lengths setting $t_i - t_{i-1} \equiv 1$. Under this simplification we can re-write one step ahead sample arc length as

$$A_n^1 = \sum_{j=2}^{n} \sqrt{1 + Y_t^2},$$

where $Y_t = X_t - X_{t-1}$, is the the first difference of the series $\{X_{t_j}\}_{j=1}^{n}$. In this setup, the sample arc length is a natural measure of the overall magnitude of the one-step-ahead series changes over the observation period and A_n/n measures the average magnitude of the one-step ahead changes of the series X_t. We can easily extend the definition of the arc length for a univariate series to a multivariate time series. Suppose $\{\mathbf{X}_t\}$ represents a $d-variate$ series given by $\mathbf{X}_t = (X_{1,t}, \ldots, X_{d,t})'$. The one step ahead sample arc length is now defined by,

$$\sum_{j=2}^{n} \sqrt{1 + (X_{1,t} - X_{1,t-1})^2 + \cdots + (X_{d,t} - X_{d,t-1})^2}.$$

By observing the definition of sample arc lengths of $\{X_t\}$ and comparing it with squared returns $\{Y_t^2\}$ and absolute values $|Y_t|$, we can say though the quantity we propose react the variability as same as the latter two the arc length provides a different measure of log prices than squared and absolute returns.

We can point out several advantages of using arc length as a tool for measuring volatility. First the arc length can be easily defined for unequally space time series. Though the results we discuss here are only derived for equally spaced time series, they can be extended to unequally spaced series as well. Second, our limit theory results hold for most of the parametric models for volatility that are commonly used in practice, including linear processes, GARCH processes, and the stochastic volatility models of Davis and Mikosch (2009). Third, arc length methods handle multivariate case well. Importantly, for multivariate series, components do not need to follow a common model for our results to hold. For example, one component could follow an autoregressive moving-average (ARMA) process and another could be generated by a GARCH recursion. Not only that but also the independence of those components is not required. Our results hold even when two or more components are driven by the same innovation sequence.

The rest of the dissertation proceeds as follows. In Chapter 2 we state our main results, which are the FCLT theorem for the sample arc length under different scenarios. Chapter 3 discusses classes of model that follow main results. In addition to volatility models we also discuss some of their probabilistic features that are being used in the context. Chapter 4 dedicates for applications of sample arc length as measure of risk. We illustrate how the asymptotic theory can be applied to real world data, utilizing arc length as the tool to identify volatility shifts and compare assets in terms of risk. In order to show how well arc length behaves, we compare arc length results with those obtained using squared and absolute values. Some brief simulations show how the methods perform for finite samples. In Section 5.1 we prove all theorems stated in Chapter 2 under general dependence assumptions on the log return processes. Section 5.2 verifies conditions for Theorems stated in Chapter 2, when log-returns follow models listed in Chapter 3.

Chapter 2

Results

In this section we state our main results; the functional central limit theorems for sample arc lengths under finite second moment conditions. We will mainly present conditions that can be easily checked and hold for most of the volatility models commonly used in practice. The asymptotics we state here can be applied for linear, ARMA, GARCH and stochastic volatility type processes.

Let $\{\mathbf{X}_t\}$ be an observable time series and $\{\mathbf{Y}_t\}$ be the first difference of it. We assume that $\{\mathbf{Y}_t\}$ is strictly stationary, but $\{\mathbf{X}_t\}$ is not necessarily stationary. The partial sums for mean corrected arc length up to time np are

$$S_n(p) = n^{-1/2} \sum_{1 \le i \le np} (\eta_i - E[\eta_0]), \qquad p \in [0,1]\,,$$

where $\eta_t = \sqrt{1 + ||\mathbf{Y}_t||^2}$ and $||\cdot||$ is the usual Euclidean norm.

The central limit theorem for these partial sums is derived assuming moment conditions of η_t. Most of the time series models are written in terms of a sequence of random innovations $\mathbf{Z_t}$. The sufficient conditions for the weakly convergence of partial sums of such models are given below.

Suppose that $\{\mathbf{Y}_t\}$ is causal in terms of the innovation sequence $\{\mathbf{Z}_t\}$:

$$\mathbf{Y}_t = g(\mathbf{Z}_t, \mathbf{Z}_{t-1}, \ldots) \tag{2.1}$$

for some function g. To prove an arc length central limit theorem, some assumptions on the dependence structure of $\{\mathbf{Z}_t\}$ need to be imposed. Let $\mathcal{F}_k = \sigma(\mathbf{Z}_k, \mathbf{Z}_{k-1}, \ldots)$ and observe that $\mathbf{Y}_t \in \mathcal{F}_t$.

Assumptions will be phrased in terms of the relationship between $\mathcal{G}_{t+h} = \sigma(\mathbf{Z}_{t+h}, \mathbf{Z}_{t+h+1}, \ldots)$ and $\mathcal{F}_t$. For processes satisfying (2.1), the two assumptions below will be sufficient for our work. Define the mth order truncation

$$\mathbf{Y}_t^{(m)} = g(\mathbf{Z}_t, \mathbf{Z}_{t-1}, \ldots, \mathbf{Z}_{t-m}, \mathbf{0}, \mathbf{0}, \ldots). \tag{2.2}$$

Assumption 1 *The process* $\{\mathbf{Z}_t\}$ *is* ϕ*-mixing with* $\sum_{h=1}^{\infty} \phi_h^{1/2} < \infty$. *Here,*

$$\sup_{A \in \mathcal{G}_{t+h}, B \in \mathcal{F}_t} |P(A|B) - P(A)| \leq \phi_h.$$

Assumption 2 *With* $Q_m = ||\mathbf{Y}_t||^2 - ||\mathbf{Y}_t^{(m)}||^2$, $\sum_{m=0}^{\infty} E[|Q_m|]^{1/2} < \infty$.

Let $\mathcal{D}[0,1]$ denote all real-valued functions on the domain $[0,1]$ that are right-continuous and have left-hand limits, equipped with the usual Skorohod topology. Weak convergence in this space is denoted by $\stackrel{\mathcal{D}[0,1]}{\to}$. Our first result is now stated; its proof is presented in the Chapter 5.

Theorem 1 *Suppose that* $\{\mathbf{X}_t\}$ *is a series with stationary first differences* $\{\mathbf{Y}_t\}$ *satisfying (2.1) and Assumptions 1 and 2. If* $E[||\mathbf{Y}_t||^2] < \infty$, *then the sum in* $\tau^2 = \mathrm{Var}(\eta_0) + 2\sum_{k=1}^{\infty} \mathrm{Cov}(\eta_0, \eta_k)$ *is absolutely convergent and* $\{S_n(p)\} \stackrel{\mathcal{D}[0,1]}{\to} \tau\{W(p)\}$, *where* $\{W(p)\}_{p=0}^{1}$ *is a standard Brownian motion.*

In practice, it is not reasonable to assume that all components in a multivariate series follow the same type of a model. For example we may find cases that one component follows a GARCH type model while another component follows an ARMA type model. In order to accommodate such situations, assume that i^{th} component of $\mathbf{Y}_t$, denoted by $Y_{i,t}$, satisfies

$$Y_{i,t} = g_i(\varepsilon_{i,t}, \varepsilon_{i,t-1}, \ldots). \tag{2.3}$$

Define the mth order truncation by $Y_{i,t}^{(m)} = g_i(\varepsilon_{i,t}, \varepsilon_{i,t-1}, \ldots, \varepsilon_{i,t-m}, 0, 0, \ldots)$. While an assumed model might satisfy both (2.1) and (2.3), $\varepsilon_{i,t} \neq Z_{i,t}$ in most typical cases.

Assumption 3 *The process* $\{\varepsilon_t\}$ *is* ϕ*-mixing with* $\sum_{h=1}^{\infty} \phi_h^{1/2} < \infty$.

Assumption 4 *For* $Q_{m,i} = Y_{i,t}^2 - \left(Y_{i,t}^{(m)}\right)^2$, *one has*

$$\sum_{m=0}^{\infty} \sqrt{E[|Q_{m,i}|]} < \infty, \quad i = 1, 2, \ldots, d.$$

As condition (2.3) is a special case of (2.1) and Assumption 4 implies Assumption 2, the following result is obtained.

Corollary 1 *Suppose that $\{\mathbf{X}_t\}$ has stationary first differences $\{\mathbf{Y}_t\}$ satisfying (2.3) and Assumptions 3 and 4. If $E[||\mathbf{Y}_t||^2] < \infty$, then the sum in $\tau^2 = \mathrm{Var}(\eta_0) + 2\sum_{k=1}^{\infty} \mathrm{Cov}(\eta_0, \eta_k)$ is absolutely convergent and $\{S_n(p)\} \overset{\mathcal{D}[0,1]}{\to} \tau\{W(p)\}$, where $\{W(p)\}_{p=0}^{1}$ is a standard Brownian motion.*

Remark 1 *In Corollary 1, the innovations $\{\varepsilon_t\}$ can have highly dependent components. For example, when $d = 2$, the result holds even if $\varepsilon_{1,t} = \varepsilon_{2,t}$.*

Remark 2 *The main point of Corollary 1 is that only marginal models for the components of the vector-valued process are needed to obtain a FCLT for multivariate arc lengths. Under general conditions, if the FCLT holds for each component arc length, then a FCLT for arc lengths of the vector process also holds.*

The asymptotic distribution of the sample arc length can be found when there exists long-range dependence which are also called long memory processes, using the central limit theorems stated above for stationary processes. When innovations are independent and identically distributed (iid), we can relax Assumption 2 to accommodate some type of long memory processes. Though the result given below is stated for univariate series, it can be extended to multivariate series. Let $\{X_t\}$ be a univariate time series with $Y_t = X_t - X_{t-1}$ satisfying

$$Y_t = g(Z_t, Z_{t-1}, ...), \tag{2.4}$$

where $\{Z_t\}$ is iid. Define $Y_t^{(m)} = g(Z_t, \ldots, Z_{t-m}, 0, 0, \ldots)$ and $c_m = E[|Y_0^2 - (Y_0^{(m-1)})^2|]$. The following result is proven in the Appendix.

Theorem 2 *Suppose that $\{X_t\}$ has stationary first differences $\{Y_t\}$ satisfying (2.4). If $E[Y_t^2] < \infty$ and $\sum_{m=1}^{\infty} \sqrt{c_m/m} < \infty$, then the sum in $\tau^2 = \mathrm{Var}(\eta_0) + 2\sum_{k=1}^{\infty} \mathrm{Cov}(\eta_0, \eta_k)$ is absolutely convergent and $\{S_n(p)\} \overset{\mathcal{D}[0,1]}{\to} \tau\{W(p)\}$, where $\{W(p)\}_{p=0}^{1}$ is a standard Brownian motion.*

More often in applications, it may be desirable to compare the risk of two asset series. The result we state below focuses on the joint asymptotic distribution of the sample arc lengths. Even though it is given for univariate components of a multivariate series, it can be easily extended to any set components.

Let $\{\mathbf{X}_t\}$ be a d-dimensional series with stationary first differences $\{\mathbf{Y}_t\}$ satisfying (2.1). The sample component arc lengths involve

$$\eta_{i,t} = \sqrt{1 + Y_{i,t}^2}, \quad i = 1, 2, \ldots, d. \tag{2.5}$$

For $i = 1, 2, \ldots, d$, let

$$S_n^{(i)}(p) = n^{-1/2} \sum_{1 \le t \le np} (\eta_{i,t} - E[\eta_{i,0}]), \quad p \in [0, 1]. \tag{2.6}$$

Theorem 3 *Suppose that $\{\mathbf{X}_t\}$ has stationary first differences $\{\mathbf{Y}_t\}$ satisfying (2.3) and Assumptions 3 and 4. If $E[||\mathbf{Y}_t||^2] < \infty$, then for $\mathbf{S}_n(p) = (S_n^{(1)}(p), \ldots, S_n^{(d)}(p))'$, the sum in $\tau_i^2 = \mathrm{Var}(\eta_{i,0}) + 2\sum_{k=1}^{\infty} \mathrm{Cov}(\eta_{i,0}, \eta_{i,k})$ is absolutely convergent for $i = 1, 2, \ldots, d$ and $\{\mathbf{S}_n(p)\} \Rightarrow \mathbf{M}\{\mathbf{W}(p)\}$, where $\Rightarrow$ denotes weak convergence, $\{\mathbf{W}(p)\}_{p=0}^1$ is a d-dimensional standard Brownian motion, and $\mathbf{M}$ is a $d \times d$ matrix with ith diagonal component τ_i^2.*

Proofs for all the theorems stated above will be given in Appendix 5.1. In Chapter 3 we will discuss classes of univariate and multivariate models commonly used when volatility is modeled in practice.

Chapter 3

Volatility Models

In this section we will consider some of the univariate and multivariate models that are being commonly discussed in literature. Let $\{\mathbf{P}_t\}_{t=1}^n$ be asset prices and $\{\mathbf{X}_t\}_{t=1}^n$ be the log transformed prices $\mathbf{X}_t = \ln(\mathbf{P}_t)$. The volatility models given below are defined in terms of the log returns $\{\mathbf{Y}_t\} = \{\ln(\mathbf{P}_t) - \ln(\mathbf{P}_{t-1})\}$.

3.1 Volatility Models - Examples

The conditions stated in Theorem 1 can be verified for all the models presented below. Proofs are provided in Section 2 of Chapter 5. Model types we discuss in the context coupled with their proper definitions are given below.

3.1.1 Univariate models

First the arc length of $\{\ln(P_t)\}$ for univariate series is taken into account.

Example 1 *If $\{P_t\}$ follows geometric Brownian motion or $X_t = \ln(P_t)$ is a causal ARIMA$(p, 1, q)$ or ARMA(p, q) series with independent and identically distributed zero-mean innovations $\{\varepsilon_t\}$ having a finite variance, then $\{Y_t\}$ is a linear process in that*

$$Y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j},$$

with $\psi_0 = 1$ and $|\psi_j| \leq c\gamma^j$ for some constants $c > 0$ and $0 \leq \gamma < 1$ (see Brockwell and Davis (1991)).

To make things more clear and organized we provide the definitions of ARIMA$(p, 1, q)$, ARMA(p, q) and geometric Brownian motion below.

Definition 1 *(Brockwell and Davis (2002))*
$\{X_t\}$ is an ARMA(p, q) process if X_t is a stationary and if for every t

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}, \tag{3.1}$$

where $\varepsilon_t \sim WN\left(0, \sigma^2\right)$ and the polynomials $\phi(\cdot) = (1 - \phi_1 z - \cdots - \phi_p z^p)$ and $\theta(z) = (1 + \theta_1 z + \cdots + \theta_p z^p)$ have no common factors.

Definition 2 *(Brockwell and Davis (2002))*
*$\{X_t\}$ is an ARIMA$(p, 1, q)$ process if $Y_t = X_t - X_{t-1}$ is a causal **ARMA**(p, q) process.*

Definition 3 *(Ross (2007))*
If $\{Y_t\}$ is a Brownian motion process, then the process $\{X_t\}$ defined by,

$$X_t = \exp\{Y\}$$

is called a geometric Brownian motion.

Example 2 *Suppose that $\{Y_t\}_{t=-\infty}^{\infty}$ follows a stationary GARCH(p, q) model Bollerslev (1986) in that*

$$\begin{aligned} Y_t &= \sigma_t \varepsilon_t, && (3.2) \\ \sigma_t^2 &= \omega + \sum_{1 \leq i \leq p} \alpha_i Y_{t-i}^2 + \sum_{1 \leq j \leq q} \beta_j \sigma_{t-j}^2, && (3.3) \end{aligned}$$

where

$$\omega > 0, \quad \alpha_i \geq 0, \quad 1 \leq i \leq p, \quad \beta_j \geq 0, \quad 1 \leq j \leq q. \tag{3.4}$$

The conditions assumed in Theorem 2.1 can now be verified using arguments similar to those in Berkes et al. (2008).

Example 3 *Suppose that* $\{Y_t\}$ *follows the stochastic volatility process of Davis and Mikosch (2009); that is,*

$$Y_t = \sigma_t \varepsilon_t, \quad t \geq 1, \quad \text{where} \quad \sigma_t = \exp\left\{\sum_{j=0}^{\infty} \psi_j \nu_{t-j}\right\}.$$

Here, $\{\sigma_t\}_{t=1}^{\infty}$ *is independent of the iid noise sequence* $\{\varepsilon_t\}_{t=1}^{\infty}$*. Also* $\{\psi_j\}_{j=0}^{\infty}$ *is a sequence of absolutely summable deterministic coefficients with* $\psi_0 = 1$ *and* $\{\nu_t\}_{t=-\infty}^{\infty}$ *is an iid sequence of zero mean random variables with a finite variance. This model is more general than the stochastic variance model of Harvey et al. (1994). We assume that the linear process* $\sum_{j=0}^{\infty} \psi_j \nu_{t-j}$ *is generated from an ARMA recursion so that* $|\psi_j| \leq c\gamma^j$ *for some* $c > 0$ *and* $\gamma \in [0, 1)$*. Without loss of generality, we assume that* $c < 2$ *(one can always increase* γ *towards unity to enforce this supposition). To use Theorem 1, one needs* $E[Y_0^2] < \infty$*. This entails assuming* $E[e^{t\nu_0}] < \infty$ *for all* t *with* $|t| \leq 2$*.*

3.1.2 Multivariate models

The univariate models above can be extended to d-dimensional settings in various ways. More information about the multivariate models and the verification of conditions of Theorem 2.1 can be found in Section 5.2. Throughout, $\{\mathbf{Y}_t\}$ is a d-variate series with components $\mathbf{Y}_t = (Y_{1,t}, \ldots, Y_{d,t})'$.

Example 4 *Suppose that each component of* $\{\mathbf{Y}_t\}$ *follows any of the univariate models previously considered. By Corollary 1, the sample arc lengths of* $\{\mathbf{X}_t\}$ *satisfy a FCLT whenever the random shock vector sequence satisfies Assumption 3.*

Example 5 *Let* $\{\mathbf{Y}_t\}$ *be a stationary and causal multivariate AR(*p*) process satisfying*

$$\mathbf{Y}_t - \mathbf{\Phi}_1 \mathbf{Y}_{t-1} - \cdots - \mathbf{\Phi}_p \mathbf{Y}_{t-p} = \mathbf{Z}_t, \tag{3.5}$$

where $\{\mathbf{Z}_t\}$ *is d-variate white noise (the components need not be independent). We can rewrite the equation 3.5 in the more compact form given by*

$$\Phi(B)\mathbf{Y}_t = \mathbf{Z}_t, \quad \{\mathbf{Z}_t\} \sim WN(\mathbf{0}, \Sigma),$$

where

$$\Phi(z) = \mathbf{I} - \Phi_1 z - \cdots - \Phi_p z^p.$$

Then the sample arc lengths of $\{\mathbf{X}_t\}$ *satisfy a FCLT.*

The modern literature has now developed several types of multivariate GARCH(p,q) processes. Bauwens et al. (2006) surveys the topic. Our next example considers the CCC GARCH(p,q) process proposed in Bollerslev (1990).

Example 6 *A d-variate series* $\{\mathbf{Y}_t\}$ *is called a CCC-GARCH*(p,q) *process if it satisfies*

$$\mathbf{Y}_t = \mathbf{H}_t^{1/2}\mathbf{Z}_t \tag{3.6}$$

and

$$\mathbf{H}_t = \mathbf{D}_t\mathbf{R}\mathbf{D}_t = (\rho_{i,j}\sqrt{\sigma_{i,i,t}\sigma_{j,j,t}}), \tag{3.7}$$

where $\mathbf{D}_t = Diag(\sigma_{1,1,t}^{1/2},\ldots,\sigma_{d,d,t}^{1/2})$ *and* $\sigma_{i,i,t}$ *follows a univariate GARCH*(p,q) *model; i.e.,*

$$\sigma_{i,i,t} = \omega_i + \sum_{1\le j\le p}\alpha_{i,j}Y_{i,t-j}^2 + \sum_{1\le j\le q}\beta_{i,j}\sigma_{i,i,t-j}, \tag{3.8}$$

and $\mathbf{R} = (\rho_{i,j})_{i,j=1}^d$ *is a symmetric positive definite matrix with* $\rho_{i,i} = 1$ *for all* $i = 1,\ldots,d$. *Here,* $\mathbf{Z}_t = (Z_{1,t},\ldots,Z_{d,t})'$ *is a d-variate random vector with* $E[\mathbf{Z}_t] \equiv \mathbf{0}$ *and* $\mathrm{Var}(\mathbf{Z}_t) \equiv \mathbf{I}_d$, *where* $\mathbf{I}_d$ *is the d-dimensional identity matrix. Then the sample arc lengths of* $\{\mathbf{X}_t\}$ *satisfy a FCLT.*

As for multivariate GARCH processes, the modern literature contains several multivariate stochastic volatility models. The model below is more general than that proposed by Harvey et al. (1994) and studied in Asai et al. (2006).

Example 7 *For an iid sequence* $\{\mathbf{Z}_t\}$, *suppose* $\{\mathbf{Y}_t\}$ *obeys* $\mathbf{Y}_t = \mathbf{H}_t^{1/2}\mathbf{Z}_t$, *with* $\mathbf{H}_t^{1/2} = Diag(\sigma_{1,t},\ldots,\sigma_{d,t})$. *Also let* $\sigma_{i,t} = \exp\{W_{i,t}\}$, *where the d-variate process* $\{\mathbf{W}_t\}$ *follows a causal vector autoregression. Then the sample arc lengths of* $\{\mathbf{X}_t\}$ *satisfy a FCLT.*

3.2 Properties of Volatility Models

In this section we discuss some of the properties those models need to be satisfied in order to hold conditions for Theorem 1. We will discuss some stationary conditions and causality conditions for both univariate and multivariate versions of models here.

We start our review with the definition of strictly stationarity and weak stationarity of a sequence.

Definition 4 *Brockwell and Davis (2002)*

(a) Strictly stationary sequence

A sequence $\{x_t : t \in \mathbb{Z}\}$ *is strictly stationary if* $(x_1, \ldots, x_n)$ *and* $(x_{1+h}, \ldots, x_{n+h})$ *have the same joint distributions for all integers* h *and* $n > 0$*.*

(b) Weakly stationary (stationary) sequence

A sequence $\{x_t\}$ *is weakly stationary if,*

(i) $E[x_t] = \mu_x(t)$ *is independent of* t*,*

and

(ii) $\gamma_x(t+h, t) = \mathrm{Cov}(x_{t+h}, x_t)$ *is independent of* t *for each* h*.*

For a GARCH(p, q) process the necessary and sufficient conditions for the existence of a strictly stationary solution can be found in Berkes et al. (2008). Before stating those conditions we will represent the squared processes $\left(Y_t^2\right)$ and $\left(\sigma_t^2\right)$ of GARCH(p, q) process in a form of a *stochastic recurrence equation*, given by

$$\mathbf{Q}_t = \mathbf{A}_t \mathbf{Q}_{t-1} + \mathbf{B}_t, \tag{3.9}$$

where

$$\mathbf{Q}_t = \left(\sigma_{t+1}^2, \ldots, \sigma_{t-q+2}^2, Y_t^2, \ldots, Y_{t-p+2}^2\right)^{'},$$

and defining $\mathbf{\Lambda_t} = \left(\beta_{\mathbf{1}} + \alpha_{\mathbf{1}}\varepsilon_{\mathrm{t}}^2, \beta_2, \ldots, \beta_{\mathrm{q}-1}\right) \in \mathbb{R}^{\mathbf{q-1}}$, $\mathbf{\Psi_t} = \left(\varepsilon_{\mathrm{t}}^2, 0, \ldots, 0\right) \in \mathbb{R}^{\mathbf{q-1}}$, $\alpha\left(\alpha_{\mathbf{2}}, \ldots, \alpha_{\mathbf{p-1}}\right)$;

$$\mathbf{A_t} = \begin{bmatrix} \mathbf{\Lambda}_t & \beta_q & \alpha & \alpha_p \\ \mathbf{I_{q-1}} & 0 & 0 & 0 \\ \mathbf{\Psi_t} & 0 & 0 & 0 \\ 0 & 0 & \mathbf{I_{p-2}} & 0 \end{bmatrix}.$$

In $\mathbf{A_t}$, an identity matrix of size j is defined by $\mathbf{I_j}$.A norm of a matrix $\mathbf{A}$ of size $d \times d$ is given as

$$||\mathbf{A}|| = \sup\left\{||\mathbf{Ax}||_{\mathbf{d}} : \mathbf{x} \in \mathbb{R}^{\mathbf{d}} \textbf{ and } ||\mathbf{x}||_{\mathbf{d}} = \mathbf{1}\right\},$$

where $||.||_d$ is the Euclidean norm in $\mathbb{R}^d$. The vector $\mathbf{B_t}$ of equation 3.9 is defined as $\mathbf{B_t} = (\alpha_\mathbf{0}, \mathbf{0}, \ldots, \mathbf{0})^{'}$. We also say a solution to (3.2) and (3.3) is nonanticipative if Y_t is independent of $\sigma(\{\varepsilon_\mathrm{j}, \mathrm{j} > \mathrm{t}\})$ (See Berkes et al. (2008)). A theorem formulated based on the result of Bougerol and Picard (1992) for the strictly stationarity of a GARCH(p, q) process is given in Berkes et al. (2008) and we state it as follows.

Theorem 4 *Suppose that (3.4) holds, $E \log ||\mathbf{A_0}|| < \infty$, and $\{\varepsilon_\mathrm{n}, \mathrm{n} \in \mathbb{Z}\}$ are independent, and identically distributed random variables. Then GARCH(p, q) process as defined in 3.2 and 3.3 has a unique, nonanticipative, stationary and ergodic solution if and only if*

$$\gamma L = \inf_{0 \leq n < \infty} \frac{1}{n+1} E \log ||\mathbf{A_0 A_1} \ldots \mathbf{A_n}|| < \mathbf{0}.$$

Note that γL is the top Lyapunov exponent of the sequence $\mathbf{A_n}$.

Berkes et al. (2008) recovers a result of Bollerslev (1986) and Bougerol and Picard (1992) that shows a strictly stationary solution to (3.2) and (3.3) when $E[Y_0^2] < \infty$ exists if and only if

$$\sum_{i \leq p} \alpha_i E\varepsilon_0^2 + \sum_{\mathrm{i} \leq \mathrm{q}} \beta_\mathrm{i} < 1.$$

Note that the stochastic volatility model defined in Example 3, will have a stationary solution when the ARMA recursion that generates the linear process has a stationary solution.

In addition to stationarity, we assume that ARMA processes stated in examples are causal. Now we present the definition of a cuasal ARMA process.

Definition 5 *(Brockwell and Davis (1991)) An ARMA(p, q) process defined in (3.1) is said to be causal (or more specifically to be a causal function of $\{\varepsilon_\mathrm{t}\}$) if there exist a sequence of constants $\{\Psi_j\}$ such that $\sum_{j=0}^{\infty} |\Psi_j| < \infty$ and*

$$Y_t = \sum_{j=0}^{\infty} \Psi_j \varepsilon_{\mathrm{t-j}}, \quad \mathrm{t} = 0, \pm 1, \ldots \tag{3.10}$$

Following theorem given in Brockwell and Davis (1991) gives the necessary and sufficient condition for an ARMA process to be causal.

Theorem 5 *Let $\{Y_t\}$ be an ARMA(p, q) process for which the polynomials $\phi(\cdot)$ and $\theta(\cdot)$ have no common zeroes. Then $\{Y_t\}$ is causal if and only if $\phi(z \neq 0)$ for all $z \in \mathbb{C}$ such that $|z| \leq 1$.*

In multivariate setting, stationarity and causality of an AR(p) process is also defined in a similar fashion.

Definition 6 *Brockwell and Davis (2002) The d-variate series* $\{mathbf X_t\}$ *is (weakly) stationary if*

- $\mu_{\mathbf{X}}(t)$ *is independent of* t*, and*

- $\Gamma_{\mathbf{X}}(t+h)$ *is independent of* t *for each* h

where

$$\mu_{\mathbf{X}}(t) = E[\mathbf{X}_t] = \begin{bmatrix} \mu_{t1} \\ \mu_{t2} \\ \vdots \\ \mu_{td} \end{bmatrix}$$

and

$$\mathbf{\Gamma}_{\mathbf{X}}(t+h,t) = \begin{bmatrix} \gamma_{11}(t+h,t) & \dots & \gamma_{1d}(t+h,t) \\ \vdots & \ddots & \vdots \\ \gamma_{d1}t+h,t & \dots & \gamma_{dd}(t+h,t) \end{bmatrix}.$$

Here for all $i,j = 1,2,\dots,d$*,* $\gamma_{ij}(t+h,t) = \mathrm{Cov}\,(X_{t+h,i}, X_{t,j})$*.*

As defined for a univariate series, we can define a causal multivariate AR(p) process and provide the condition it need to satisfy to be a causal process as follows.

Definition 7 *Brockwell and Davis (2002) An AR(p) process defined in (3.5) is causal, or a causal function of* $\{\mathbf{Z}_t\}$*, if there exist matrices* Ψ_j *with absolutely summable components such that*

$$\mathbf{Y}_t = \sum_{j=0}^{\infty} \Psi_j \mathbf{Z}_{t-j} \quad for\ \ all\ \ t.$$

Causality is equivalent to the condition det $\Phi(z) \neq 0$ *for all* $z \in \mathbb{C}$ *such that* $|z| \leq 1$*.*

Some real world applications and simulation results, using sample arc lengths as the tool of measuring the volatility are presented in next chapter. We also compare arc length results with squared and absolute values.

Chapter 4

Applications

Sample arc length can be used in may applications that involves with risk and volatility. Detecting volatility shifts, comparing and clustering time series for similarities in terms of risk are some of those, we discuss in this section. In addition to real world scenarios, we also present some simulations studies to illustrate how well sample arc length responds for infinite fourth moment conditions and variety model classes compared to squared and absolute values.

4.1 Changepoint Detection

Detecting time points that changes the market volatility is highly important in finance. Specifically the goal here is to identify points where the volatility changes from high to low or vise versa. We can find some parametric and non-parametric methods that have been developed regarding this subject in literature. Berkes et al. (2004b) propose Gaussian likelihood methods to detect changepoints in GARCH(p, q) processes while Mercurio and Spokoiny (2004) use methods based on locally adaptive volatility estimate.

On the other hand non-parametric methods are proposed based on cumulative sum (CUSUM) or moving sum procedures. Suppose that $\{V_t\}_{t=1}^{n}$ is any given sequence. The CUSUM statistics is defined by

$$C_n = \max_{1\leq k\leq n}\left|\sum_{1\leq i\leq k} V_i - \frac{k}{n}\sum_{1\leq i\leq n} V_i\right|.$$

Suppose that the partial sums of $\{V_t\}$ satisfy a Gaussian FCLT:

$$\left\{ n^{-1/2} \sum_{t=1}^{np} (V_t - E[V]) \right\} \overset{D[0,1]}{\to} \tau_V \{W(p)\},$$

where $\{W(p)\}_{p=0}^{1}$ is a standard Brownian motion. Here, $\tau_V := \lim_{n\to\infty} n\mathrm{Var}(n^{-1}\sum_{t=1}^{n} V_t)$ is n times the long-run variance of the stationary series $\{V_t\}$. By the continuous mapping theorem,

$$\frac{C_n}{\hat{\tau}_V n^{1/2}} \overset{\mathcal{D}}{\to} \sup_{0\le p\le 1} |B(p)|,$$

where $\{B(p)\}_{p=0}^{1}$ is a Brownian bridge and $\hat{\tau}_V$ is any consistent estimator of τ_V.

For a finite n, one estimates τ_V with

$$\gamma_V(0) + 2\sum_{h=1}^{n-1}(1 - h/n)\gamma_V(h),$$

where $\gamma_V(h)$ is the autocovariance of $\{V_t\}$ at lag h. Unfortunately, the usual estimator

$$\widehat{\gamma}_V(h) = \frac{1}{n}\sum_{t=1}^{n-h}(V_t - \overline{V})(V_{t+h} - \overline{V})$$

is biased, markedly so for large h. Hence, the sum is truncated at the greatest integer less than or equal to $\sqrt[3]{n}$:

$$\hat{\tau}_V^2 = \left[\hat{\gamma}_V(0) + 2\sum_{h=1}^{\left[\sqrt[3]{n}\right]} (1 - h/n)\hat{\gamma}_V(h) \right].$$

This estimator is consistent and avoids excessive bias with large lags; Berkes et al. (2009) discusses this and other truncation schemes.

We consider Dow Jones Index (DWJ) daily closing values from January 1, 2005 through December 31, 2009 to detect changepoints in volatility if there exists any. Figure 4.1 shows the daily log prices (left) and daily log-returns (right). After applying the base transformation to generate log-returns the market behavior becomes more evident. This time period is particularly chosen since it represents a time interval before and after the rescission hit the US markets. The impact of the economic crisis can be clearly seen through the high volatility of the market in the last quarter of 2008. Visual inspection of the figure 4.1 suggests several changepoints in volatility mainly in the second half of the series.

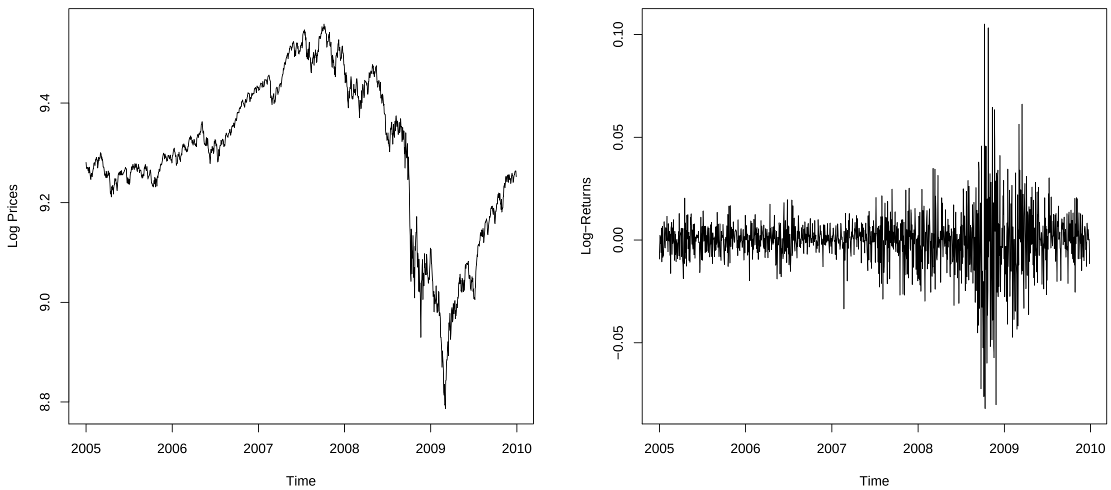

Figure 4.1: *Log-Transformed Prices (Left) and Log-Returns (Right) of Dow Jones Index values.*

In order to accept of deny this suggestion statistically, the CUSUM tests were applied to the sample arc lengths $v_t = \sqrt{1+y_t^2}$, absolute returns $v_t = |y_t|$, and squared returns $v_t = y_t^2$. In all cases, we focused on values of $\hat{\tau}_V^{-1} n^{1/2} C_n$ and any number of it larger than the 95th percentile of $\max_{0\le p\le 1} |B(p)|$ suggests statistically significant volatility changes. The time where C_n is maximized is the estimated changepoint time.

Figure 4.2 presents CUSUM test statistics for the series from January 2005 through December 2009, in which arc length is considered as the tool of quantifying volatility. Here, $v_t = \sqrt{1+y_t^2}$ and y_t is the observed log return. The horizontal line marks the 5% significance threshold. The maximum value of the test statistic is found as 2.321, which is highly significant (p-value $= 4.181 \times 10^{-5}$). The corresponding point, that is considered as the time where volatility shifts significantly is identified as July 23, 2008.

We now divide the series into two about the changepoint recorded on July 23, 2008. This results two segments of series, the first consists of 894 observation recorded from January 1, 2005 — July 23, 2008 and the second consists of 364 observations from July 24, 2008 — December 31, 2009. Then we apply the CUSUM test separately on both series to determine if there exist any changepoints. Figure 4.3 plots those test statistics and the 5% significance thresholds.

In order to determine if any multiple changepoints are present, we subsegment the series and sequentially apply the CUSUM test to the subsegments. Even though this is not the true multiple

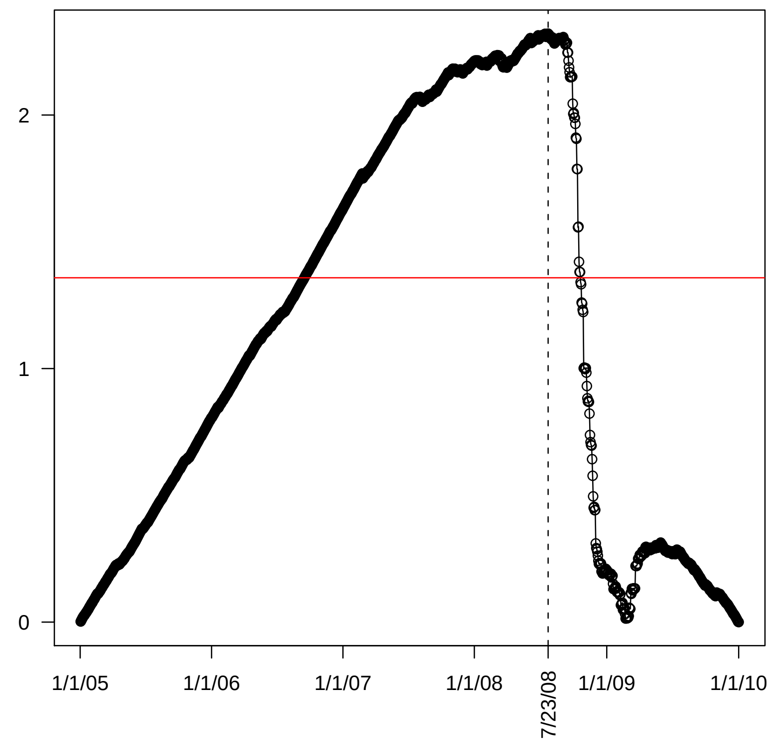

Figure 4.2: *DWJ CUSUM statistics from Jan 1, 2005 — Dec 31, 2009;* $v_t = \sqrt{1 + y_t^2}$.

changepoint identification scheme, this type of methods often perform well. This segmentation and application of CUSUM test is continued until no statistically significant changepoint can be detected in those subsegments. The results of this process are presented in Table 4.1.

We also applied the CUSUM test to both absolute returns $v_t = |y_t|$ and the squared returns $v_t = y_t^2$ to compare those results with ones we observed using arc length. Table 4.1 shows all those those statistically significant changepoints. Both arc length and squared returns indicates four changepoints on exactly same time points. But interestingly arc length produces slightly smaller p-value than square returns provides. This is an indication that arc length may be more powerful than squared returns here. On the other hand CUSUM test based on absolute returns suggests seven changepoints locates at completely different dates compared to time points arc length and squared returns indicate. This behavior might be a result of the rapid fluctuations of absolute returns unlike arc length or squared returns when quantity is near zero. In fact, for large y_t, arc length and absolute

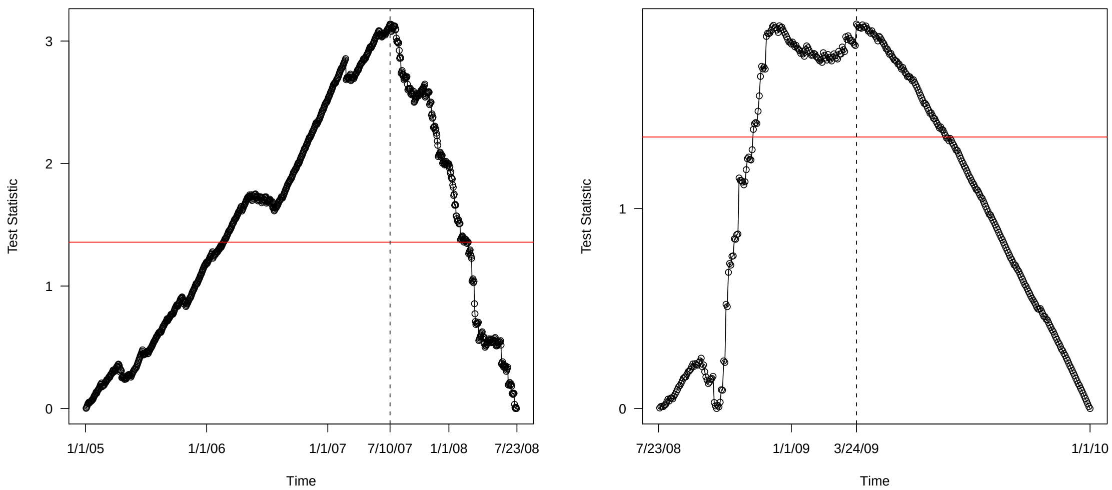

Figure 4.3: *CUSUM statistics for sub-sample 1 (left) and sub-sample 2 (right).*

CUSUM summands are essentially the same.

Table 4.1: Changepoints and p-values for arc length, squared returns, and absolute returns.

Changepoint Date	Arc Length	Squared Returns	Absolute Returns
7/25/2006			0.009
2/27/2007			0.037
6/5/2007			0.0232
7/10/2007	5.551×10^{-9}	5.554×10^{-9}	
10/30/2007			2.770×10^{-8}
7/23/2008	4.181×10^{-5}	4.212×10^{-5}	
9/15/2008			1.878×10^{-5}
3/24/2009	0.00123	0.00124	
4/22/2009			0.0216
7/16/2009	0.0007	0.0007	
7/24/2009			0.0227

We proved the FCLT for arc length under assumption of finite second moments. In contrast, this result hold for squared returns only under finite fourth moment conditions. When fourth moments are infinite, the FCLT for squared returns falls apart and the $\sqrt{n}$ normalization for the squared return CUSUM statistic does not lead to a weak limit. The tail behavior of the underlying distribution determines the correct normalization. In practice it is difficult to infer the tail behavior by observing a set of data. Hence practitioners often misapply the $\sqrt{n}$ normalization though heavy tails are present.

A simulation study is conducted to demonstrate how the power of the CUSUM test varies depend on the moment conditions We draw n iid series from a Pareto distribution with a shape parameter of three and a unit scale parameter. These conditions make sure that the series has finite second moments but infinite fourth moments. In order to introduce an artificial changepoint, we placed a scale shift at the center for each series n. That is, $\{Z_t\}$ is a simulated as iid Pareto and

$$Y_t = \begin{cases} Z_t & \text{for} \quad t \leq n/2 \\ 1.1Z_t & \text{for} \quad t > n/2 \end{cases} .$$

Ten thousand independent series are generated and the empirical power of the CUSUM test is computed. The results are given in Figure 4.4 and clearly it shows that both arc length and absolute returns outperform squared returns.

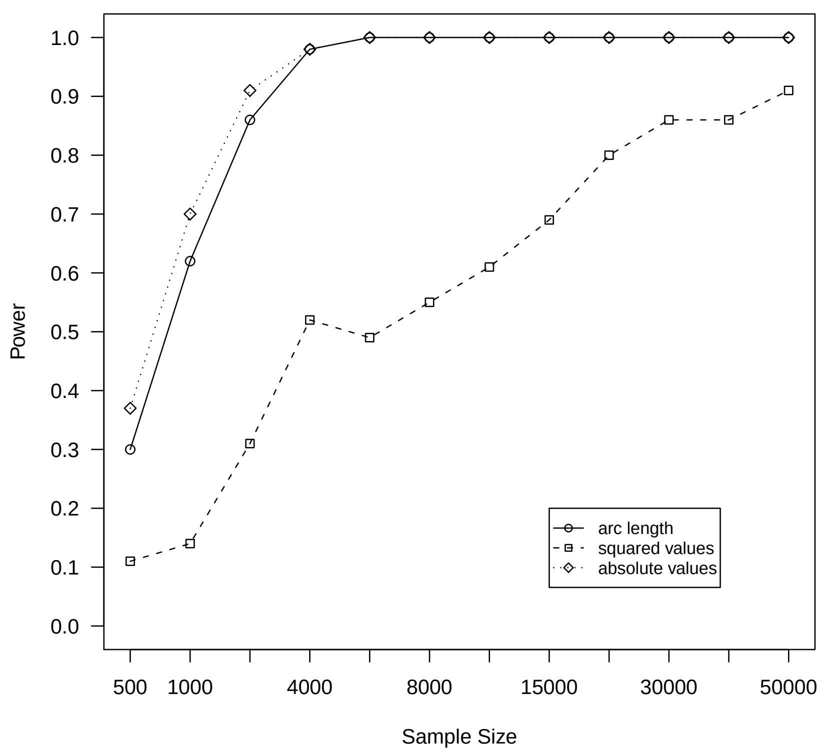

Figure 4.4: *Power of the CUSUM tests for Pareto data.*

We also carried out another simulation study to compare the Type I error of CUSUM test considering four different model types. Specifically, we consider log returns are generated from the following processes: $\{Y_t\}$ is iid $N(0,1)$; $\{Y_t\}$ is a stationary Gaussian ARMA$(1,1)$ series with AR coefficient 0.3 and moving-average coefficient 0.2; $\{Y_t\}$ is a stationary GARCH$(1,1)$ series with $\omega = 0.01$, $\alpha = 0.3$, and $\beta = 0.2$; Stochastic volatility with $\ln(\sigma_t)$ being a stationary Gaussian ARMA $(1,1)$ series with AR coefficient 0.3 and moving-average coefficient 0.2. Ten thousand independent series of length $n = 250$ were generated from each of the above processes and the empirical probability of Type I error with test size $\alpha = 0.05$ was computed. The length $n = 250$ was chosen since it roughly corresponds to one year of daily stock prices. In all models, the iid sequences driving the model errors were generated from the normal family. The simulation results in Table 4.2 indicate that all CUSUM tests are conservative. This phenomenon has been observed in other studies — CUSUM tests based on asymptotic quantiles coming from the supremum of the Brownian bridge tend to be conservative Robbins et al. (2011).

Table 4.2: Probability of Type I error for the CUSUM tests.

Model Type	Arc Length	Squared Returns	Absolute Returns
IID Normal$(0,1)$	0.0316	0.0296	0.0334
ARMA$(1,1)$	0.0370	0.0336	0.0392
GARCH$(1,1)$	0.0346	0.0332	0.0431
Stochastic Volatility	0.0255	0.0127	0.0273

The empirical power of the CUSUM test is also computed based on a simulation of ten thousand series of size $n = 250$. This study is conducted based on the same four scenarios stated above to investigate how arc length perform compared to square returns and absolute returns. We multiplied the simulated data by 1.5 for $t > n/2$ to make sure there is a changepoint at time $n/2$. Table 4.3 displays respective power values. Based on all the findings, it appears that arc length methods outperform methods based on squared returns. Arc length methods also perform comparatively well against absolute returns methods.

Table 4.3: Power of the CUSUM tests.

Model Type	Arc Length	Squared Returns	Absolute Returns
IID Normal$(0,1)$	0.9594	0.9560	0.9457
ARMA$(1,1)$	0.8718	0.8584	0.8557
GARCH$(1,1)$	0.6287	0.6106	0.7106
Stochastic Volatility	0.1241	0.0348	0.1426

4.2 Stock Risk Comparisons

In this subsection we demonstrate how Theorem 3 can be implemented to compare the risk of two series asset prices. Though we compare only individual stocks, the same procedure will be used when two set of portfolios are compared. As a result of the Theorem 3 we state the following result as a corollary.

Corollary 2 *Suppose the $d = 2$ dimensional series $\{\mathbf{X}_t\}$ has first differences $\{\mathbf{Y}_t\}$ satisfying (2.3) and Assumptions 3 and 4. If $E[||\mathbf{Y}_t||^2] < \infty$, then $S_n^{(1)}(1) - S_n^{(2)}(1)$ has an asymptotic normal distribution with zero mean and variance $\tau^2 = \mathrm{Var}(\eta_t) + 2\sum_{k=1}^{\infty} \mathrm{Cov}(\eta_0, \eta_k)$, where $\eta_t = \sqrt{1 + Y_{1,t}^2} - \sqrt{1 + Y_{2,t}^2}$.*

As one of the applications we consider British Petroleum (BP), Exxon Mobil (Exxon), and Royal Dutch Shell (Shell) stock prices from January 01, 2006 — December 31, 2006. The year 2006 is chosen because CUSUM test is unable to detect any changepoint via any method when it is applied to above three return series.

Figure 4.5 displays log-return for BP, Exxon, and Shell in year 2006. It is hard to see a clear difference between three plots by the naked eye. In order to make a statistical conclusion we used Corollary 2 and compare them pairwise for any differences in volatility. The results are listed in Table 4.4. Making our belief true, it appears that none of the companies have significantly different volatilities at the 5% level of significance; BP and Exxon, in fact, had highly similar volatilities.

Table 4.4: Risk comparison of oil companies.

Comparison	P-Value
BP vs Exxon	0.950
Exxon vs Shell	0.175
BP vs Shell	0.069

As another application we now consider three internet technology companies Google, Intel, and Apple from January 01, 2007 — December 31, 2007. The CUSUM test indicates these three return series are free of changepoints for year 2007 via any method. Time plots corresponding to three companies are shown in Figure 4.6.

Applying the Corollary 2 we compared them pairwise to investigate if there exist any statistical difference in terms of volatility. Visually, both Google and Intel seems to behave similarly, while Apple indicates a different behavior compared to other two. The results based on the hypothesis test

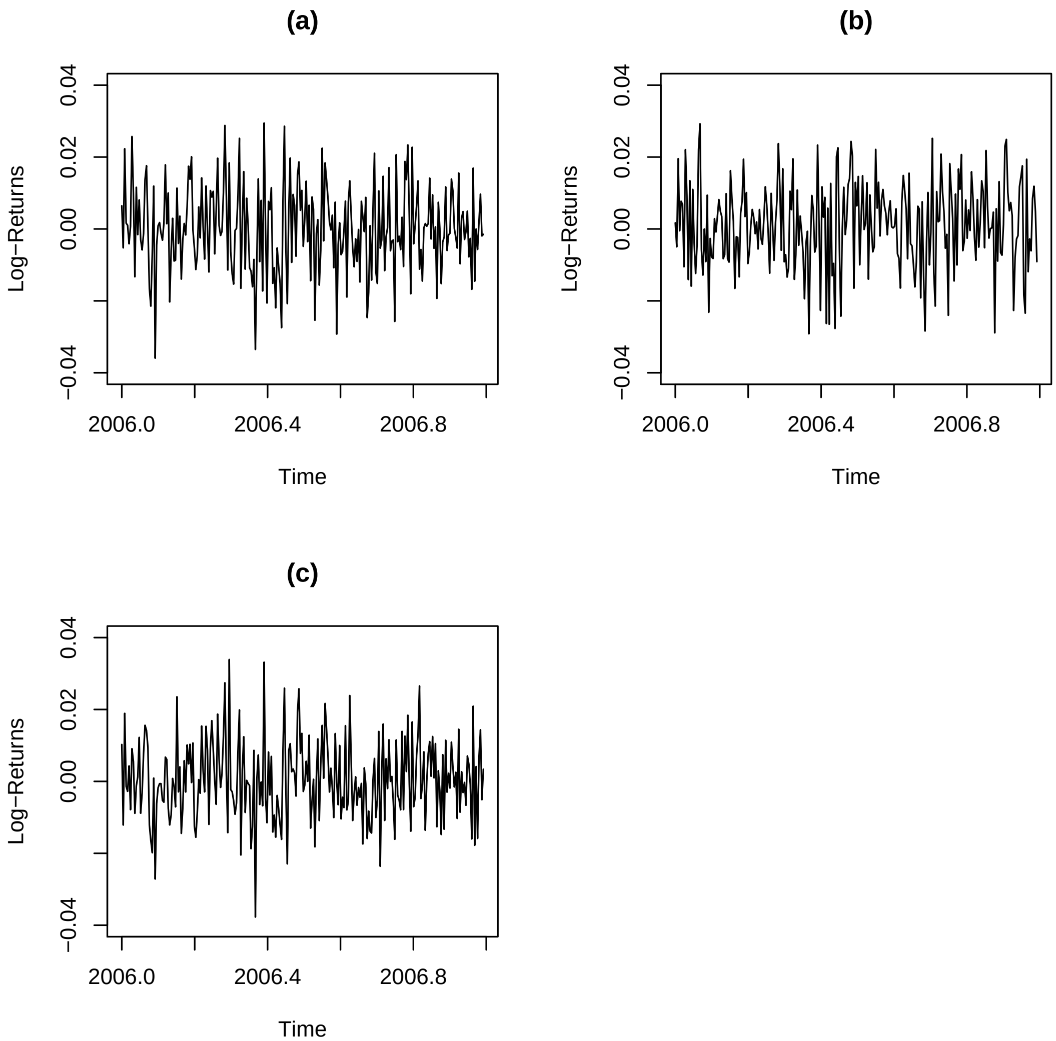

Figure 4.5: *Log Returns of (a) BP, (b) Exxon, and (c) Shell from 1/1/2006 — 12/31/2006.*

are listed below in Table 4.5. The p-values confirm that Google and Intel have similar volatilities while that of Apple is significantly different.

Table 4.5: Risk comparison of internet technology companies.

Comparison	P-Value
Google vs Intel	0.4176
Google vs Apple	0.00002
Intel vs Apple	0.00087

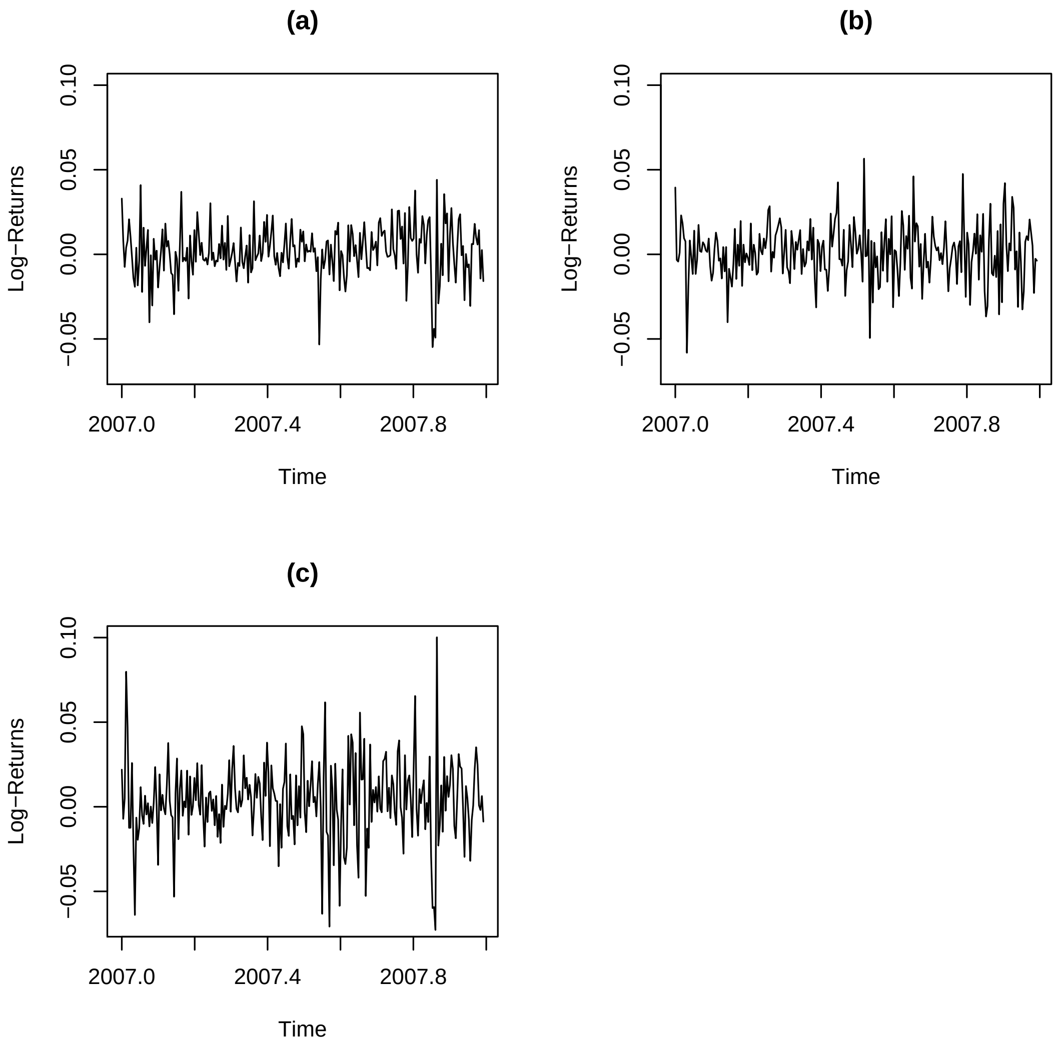

Figure 4.6: *Log Returns of (a) Google, (b) Intel, and (c) Apple from 1/1/2007 — 12/31/2007.*

4.3 Discussion and Future Work

In this work we accomplished several things worth mentioning. We introduce sample arc length as tool for quantifying the magnitude of k-step-ahead changes of a stationary time series. We also present general asymptotic theory for the sample arc length for stationary data under finite second moment conditions. Lastly we verify general conditions required to make sure Gaussian asymptotics hold for popular models including, linear, ARMA, GARCH and stochastic volatility type processes.

On the other hand arc length can be applied on any time scale. It can also be adapted to non-equally spaced time series and unlike other tools, arc length is a natural measure for multivariate time series.

When asymptotics for sample arc length are discussed we also proved the FCLT theorem for long memory processes with independent and identically distributed (iid) innovations, it does not cover the general family of these type of models. It is still a challenging and an open problem to relax the iid innovation assumption. This could lead the arc length to a measure of risk free of model assumptions.

Though we used the arc length to detect changepoints exist in a historical series, it has not been applied in identifying changes occur at real time. Adopting the arc length to detect changepoints at real time stand as another avenue of extending this work in the future.

Chapter 5

Proofs

Here we prove the results in this dissertation.

5.1 Proof of the Main Theorems in Chapter 2

This section proves the theorems stated in Section 2. We start with a lemma that is repeatedly used below.

Lemma 1 *Given two positive random variables U and V with finite variances,*

$$E[(\sqrt{1+U}-\sqrt{1+V})^2] \leq E[|U-V|].$$

Proof(Lemma 1) *Observe that*

$$\begin{aligned} \left(\sqrt{1+U}-\sqrt{1+V}\right)^2 &\leq \left|\left(\sqrt{1+U}-\sqrt{1+V}\right)\left(\sqrt{1+U}+\sqrt{1+V}\right)\right| \\ &\leq |U-V|. \end{aligned}$$

Taking expectations proves the lemma. Theorem 20.1 of Billingsley (1968) is used as the foundation when proving both Thorem 1 and 3 of this work. Hence we state Billingsley's functional central limit theorem beow.

Theorem 6 *Billingsley (1968) Suppose that $\{\varepsilon_n\}$ is φ mixing with $\sum \varphi_n^{1/2} < \infty$ and that the $\eta_n = f(\ldots, \varepsilon_{n-1}, \varepsilon_n, \varepsilon_{n+1} \ldots)$ have mean 0 and finite variance. Suppose further that there ex-*

ist random variables of the form $\eta_{ln} = f(\varepsilon_{n-l}, \ldots, \varepsilon_n, \ldots, \varepsilon_{n+l})$ *such that* $\sum \nu_l^{1/2} < \infty$, *where* $\nu_l = E\left\{|\eta_0 - \eta_{l0}|^2\right\}$. *Then the series*

$$\sigma^2 = E\left\{\eta_0^2\right\} + 2\sum_{k=1}^{\infty} E\left\{\eta_0\eta_k\right\}$$

converges absolutely; if $\sigma^2 > 0$ *and* X_n *is defined by* $X_n = \frac{1}{\sigma\sqrt{n}} S_{[nt]}(\omega)$, *then*

$$X_n \xrightarrow{D} W$$

where $S_n = \eta_1 + \ldots + \eta_n$ *and* W *is a standard Brownian motion.*

Now we can prove Theorem 1.

Proof(Theorem 1)

Let $\eta_t^{(m)} = \sqrt{1 + ||\mathbf{Y}_t^{(m)}||^2}$, *where* $\mathbf{Y}_t^{(m)}$ *is given in (2.2). Since finite variances of* $\{\mathbf{Y}_t\}$ *and Assumption 1 hold, the result follows from Theorem 21.1 of Billingsley (1968) if we can show that*

$$\sum_{m=1}^{\infty} E[|\eta_0 - \eta_0^{(m)}|^2]^{1/2} < \infty. \tag{5.1}$$

Letting $U = ||\mathbf{Y}_0||^2$ *and* $V = ||\mathbf{Y}_0^{(m)}||^2$, *Lemma A.1 and Assumption 2 complete our proof:*

$$\sum_{m=1}^{\infty} E[|\eta_0 - \eta_0^{(m)}|^2]^{1/2} \leq \sum_{m=1}^{\infty} E[|Q_m|]^{1/2} < \infty,$$

Proof(Theorem 2)

Suppose that $\{Y_t\}$ *satisfies (2.4) and set*

$$\xi_t = \sqrt{1 + Y_t^2} - E\left[\sqrt{1 + Y_0^2}\right], \quad \xi_t^{(m)} = \sqrt{1 + (Y_t^{(m)})^2} - E\left[\sqrt{1 + (Y_t^{(m)})^2}\right].$$

The result follows from the invariance principle given in F. et al. (2006) if we can establish that $\sum_{m=1}^{\infty} m^{-1/2}\{E\left[E[\xi_0|\mathcal{F}_{-m}]^2\right]\}^{1/2} < \infty$, *where* $\mathcal{F}_t = \sigma(Z_j, j \leq t)$. *Since* $\{Z_t\}$ *is iid,* $E[\xi_0^{(m-1)}|\mathcal{F}_{-m}] = 0$ *and*

$$\begin{aligned} E\left[E[\xi_0|\mathcal{F}_{-m}]^2\right] &= E\left[E[\xi_0 - \xi_0^{(m-1)}|\mathcal{F}_{-m}]^2\right] \\ &\leq E[(\xi_0 - \xi_0^{(m-1)})^2] \\ &\leq E[(\eta_0 - \eta_0^{(m-1)})^2] \end{aligned}$$

$$\leq \quad E[|Y_0^2 - (Y_0^{(m-1)})^2|],$$

where the last line follows from Lemma A.1. The above bounds give

$$\sum_{m=1}^{\infty} m^{-1/2} \left\{E\left[E[\xi_0|\mathcal{F}_{-m}]^2\right]\right\}^{1/2} \leq \sum_{m=1}^{\infty} (c_m/m)^{1/2},$$

which is finite by assumption.

Cramer-Wold device for process convergence Davidson (1994) plays a major role in the proof of Theorem 3 and we state it below to make the proof clear.

Theorem 7 *Davidson (1994) Let $\mathbf{X}_n \in \mathcal{D}^d$ be a an d-vector of random elements. $\mathbf{X}_n \xrightarrow{\mathcal{D}} \mathbf{X}$, where $P(\mathbf{X} \in \mathcal{C}^d) = 1$, if and only if $\lambda^{'}\mathbf{X}_n \xrightarrow{\mathcal{D}} \lambda^{'}\mathbf{X}$ for every fixed λ with $\lambda^{'}\lambda = 1$*

where $\mathcal{C} = \mathcal{C}[0,1]$ is the space of continuous real valued functions on $[0,1]$.

Proof(Theorem 3)

We prove the theorem for $d = 2$; arguments for higher dimensions are similar. We use the Cramer-Wold device for process convergence (Theorem 29.16 of Davidson (1994)). Let a_1 and a_2 be real numbers and consider $a_1 S_n^{(1)}(p) + a_2 S_n^{(2)}(p)$, where $S_n^{(i)}(p)$ is as given in (2.6) for $i = 1, 2$. Let $\eta_{i,t}$ be as in (2.5) and define $\eta_{i,t}^{(m)} = \sqrt{1 + \left(Y_{i,t}^{(m)}\right)^2}$ and

$$\xi_t = a_1(\eta_{1,t} - E[\eta_{1,t}]) + a_2(\eta_{2,t} - E[\eta_{2,t}]).$$

We now show that $\{\xi_t\}$ satisfies a FCLT.

Using Theorem 21.1 of Billingsley (1968), it is enough to show that

$$\sum_{m=1}^{\infty} E[(\xi_0 - \xi_0^{(m)})^2]^{1/2} < \infty, \tag{5.2}$$

where

$$\xi_0^{(m)} = a_1(\eta_{1,0}^{(m)} - E[\eta_{i,t}]) + a_2(\eta_{2,0}^{(m)} - E[\eta_{i,t}]).$$

We have the following approximations:

$$\begin{aligned}\sum_{m=1}^{\infty} E[(\xi_t - \xi_t^{(m)})^2]^{1/2} &\leq \sqrt{2}(a_1^2 E[(\eta_{1,0} - \eta_{1,0}^{(m)})^2])^{1/2} \\ &+ \sqrt{2}(a_2^2 E[(\eta_{2,0} - \eta_{2,0}^{(m)})^2])^{1/2}\end{aligned}$$

$$
\begin{aligned}
&\le \sqrt{2}|a_1|E\left[|Y_{1,0}^2-(Y_{1,0}^{(m)})^2|\right]^{1/2}\\
&+ \sqrt{2}|a_2|E\left[|Y_{2,0}^2-(Y_{2,0}^{(m)})^2|\right]^{1/2},
\end{aligned}
$$

where the last inequality follows from Lemma A.1. Condition (5.2) now follows from Assumption 4. Applying Theorem 21.1 from Billingsley (1968) shows that

$$
\tau^2 = \mathrm{Var}\left(a_1 y_{1,0} + a_2 y_{2,0}\right) + 2\sum_{k=1}^{\infty}\mathrm{Cov}\left(a_1 y_{1,0} + a_2 y_{2,0}, a_1 y_{1,k} + a_2 y_{2,k}\right)
$$

is absolutely summable and that

$$
a_1 S_n^{(1)}(p) + a_2 S_n^{(2)}(p) \stackrel{\mathcal{D}[0,1]}{\to} \tau\{W(p)\},
$$

where $\{W(p)\}_{p=0}^{1}$ *is a standard Brownian motion.*

Using properties of covariance and standard arguments, $\tau\{W(p)\}$ *follows the same law as* $(a_1, a_2)\mathbf{M}\{\mathbf{W}(p)\}$, *where* $\{\mathbf{W}(p)\}_{p=0}^{1}$ *has components which are independent standard Brownian motions. The result now follows from the Cramer-Wold device.*

5.2 Proof of the Examples in Section 3

5.2.1 Univariate Models

This section provides detailed proofs when log returns follow an ARMA, or a GARCH process, or a Stochastic volatility model. Examples are numbered according to the numbers appear in the paper.

Proof(Example 1)

Let $Y_0^{(m)} = \sum_{j=0}^{m}\psi_j\varepsilon_{-j}$. *Notice that*

$$
\begin{aligned}
|Q_m| &= |Y_0^2 - (Y_0^m)^2|\\
&= |\left(Y_0 - Y_0^m\right)\left(Y_0 + Y_0^m\right)|\\
&= |\left(Y_0 - Y_0^m\right)\left(Y_0 - Y_0^m + 2Y_0^m\right)|\\
&= |\left(Y_0 - Y_0^m\right)^2 + 2\left(Y_0 - Y_0^m\right)Y_0^m|
\end{aligned}
$$

$$\leq \quad |\left(Y_0 - Y_0^m\right)^2| + 2|\left(Y_0 - Y_0^m\right) Y_0^m|$$

Using the iid assumption on $\{\varepsilon_{\mathrm{t}}\}$ *and the fact that all expectations are positive we have*

$$\begin{aligned}
(\mathrm{E}|Q_m|)^{1/2} &\leq \left(\mathrm{E}\left(Y_0 - Y_0^m\right)^2 + 2\mathrm{E}|\left(Y_0 - Y_0^m\right) Y_0^m||\right)^{1/2}, \\
&\leq \left(\mathrm{E}\left(Y_0 - Y_0^m\right)^2\right)^{1/2} + \sqrt{2}\left(\mathrm{E}|Y_0 - Y_0^m|\mathrm{E}|Y_0^m|\right)^{1/2}, \\
&\leq \left(\mathrm{E}\left(Y_0 - Y_0^m\right)^2\right)^{1/2} + \sqrt{2}\left(\mathrm{E}|Y_0 - Y_0^m|\mathrm{E}|Y_0|\right)^{1/2}.
\end{aligned}$$

For $\mathrm{Var}(\varepsilon_0) = \sigma_\varepsilon^2$

$$\begin{aligned}
E\left(Y_0 - Y_0^m\right)^2 &= E\left[\left(\sum_{j=m+1}^{\infty} \psi_j \varepsilon_{-\mathrm{j}}\right)\left(\sum_{k=m+1}^{\infty} \psi_j \varepsilon_{-\mathrm{k}}\right)\right] \\
&= \sum_{j=m+1}^{\infty} \sum_{k=m+1}^{\infty} \psi_j \psi_k E\left[\varepsilon_{-\mathrm{j}} \varepsilon_{-\mathrm{k}}\right] \\
&= \sum_{j=m+1}^{\infty} \psi_j^2 E \varepsilon_{-\mathrm{j}}^2 \\
&= \sum_{j=m+1}^{\infty} \psi_j^2 \sigma_\varepsilon^2 \\
&\leq \sum_{j=m+1}^{\infty} c^2 \gamma^{2j} \sigma_\varepsilon^2 \\
&= c^2 \sigma_\varepsilon^2 \frac{\gamma^{2(m+1)}}{(1-\gamma^2)} \\
\sum_{m=1}^{\infty} \left(E(Y_0 - Y_0^m)^2\right)^{1/2} &\leq c\sigma_\varepsilon \frac{\gamma^2}{\left(1-\gamma\right)\left(1-\gamma^2\right)^{1/2}}
\end{aligned}$$

Also

$$\begin{aligned}
E|Y_0 - Y_0^m| &= E|\sum_{j=m+1}^{\infty} \psi_j \varepsilon_{-\mathrm{j}}| \\
&\leq \sum_{j=m+1}^{\infty} |\psi_j| E|\varepsilon_0| \\
&\leq \left(E\left(\varepsilon_0^2\right)\right)^{1/2} \sum_{j=m+1}^{\infty} |\psi_j|
\end{aligned}$$

$$
\begin{aligned}
&\leq \sigma_\varepsilon \sum_{j=m+1}^{\infty} c\gamma^j \\
&= c\sigma_\varepsilon \frac{\gamma^{m+1}}{(1-\gamma)} \\
\sum_{m=1}^{\infty} \left(E|Y_0 - Y_0^m|\right)^{1/2} &\leq c^{1/2}\sigma_\varepsilon^{1/2} \frac{\gamma}{\left(1-\gamma^{1/2}\right)\left(1-\gamma\right)^{1/2}}
\end{aligned}
$$

Since $E[|Y_0|] \leq E[|\varepsilon_0|] \sum_{j=0}^{\infty} |\psi_j| < \infty$, it follows that $\sum_{m=1}^{\infty} E[|Q_m|]^{1/2} < \infty$. By Theorem 1, a FCLT holds for the sample arc lengths of $\{X_t\}$.

Proof(Example 2)

We follow the arguments in Berkes et al. (2008)'s Theorem 2.1. Without loss of generality, we may take $p = q$. The division lemma implies that every positive integer m can be expressed as $m = p\ell + r$, where ℓ and r are integers satisfying $\ell \geq 0$ and $0 \leq r \leq p-1$.

Define an mth order truncation of the stationary solution of a GARCH(p,q) series at time zero, say $Y_0^{(m)}$, via

$$
\left(Y_0^{(m)}\right)^2 = \omega\varepsilon_0^2 \left(\sum_{k=1}^{\ell} \sum_{1 \leq l_1, \ldots, l_{k-1} \leq p} \prod_{i=1}^{k-1} \left(\beta_{l_i} + \alpha_{l_i} \varepsilon^2_{-l_1-l_2-\cdots-l_i}\right) \right)
$$

Note that ℓ is implicitly defined. Since $Y_0^2 \geq (Y_0^{(m)})^2$, we have

$$
Y_0^2 - (Y_0^{(m)})^2 = \omega\varepsilon_0^2 \sum_{k=\ell+1}^{\infty} \sum_{1 \leq l_1, \ldots, l_{k-1} \leq p} \prod_{i=1}^{k-1} \left(\beta_{l_i} + \alpha_{l_i} \varepsilon^2_{-l_1-l_2-\cdots-l_i}\right).
$$

Now

$$
E[Y_0]^2 - E[(Y_0^{(m)})]^2 \leq \frac{\omega E[\varepsilon_0^2] c^\ell}{1-c},
$$

where $c = [(\alpha_1 + \cdots + \alpha_p)\, E\varepsilon_0^2 + (\beta_1 + \cdots + \beta_q)] < 1$, since the GARCH process is strictly stationary.

It now follows that

$$
\begin{aligned}
\sum_{m=1}^{\infty} E[|Q_m|]^{1/2} &\leq \sum_{\ell=1}^{\infty} \left(\frac{\omega E[\varepsilon_0^2] c^\ell}{1-c} \right)^{1/2} \\
&= \left(\frac{\omega\sigma_\varepsilon^2}{1-c} \right)^{1/2} \sum_{\ell=1}^{\infty} c^{\ell/2}
\end{aligned}
$$

$$= \left(\frac{\omega\sigma_\varepsilon^2}{1-c}\right)^{1/2} \frac{c^{1/2}}{\left(1-c^{1/2}\right)} < \infty.$$

Hence, the FCLT for sample arc lengths of $\{\ln(P_t)\}$ *follows from Theorem 2.1.*

Proof(Example 3)

Define the m-th order truncation

$$Y_0^{(m)} = \varepsilon_0 \exp\left(\sum_{j=0}^{m} \psi_j \nu_{-j}\right).$$

The above assumptions guarantee that $E[(Y_0^{(m)})^2]$ *is bounded in* m. *With* $\alpha_0 = \sum_{j=0}^{\infty} \psi_j \nu_{-j}$ *and* $\alpha_0^{(m)} = \sum_{j=0}^{m} \psi_j \nu_{-j}$, *we obtain*

$$\begin{aligned}
\left|Y_0^2 - \left(Y_0^{(m)}\right)^2\right| &= \left|\varepsilon_0^2 \exp 2\alpha_0 - \varepsilon_0^2 \exp 2\alpha_0^{(\mathrm{m})}\right| \\
&= \varepsilon_0^2 \exp 2\alpha_0^{(\mathrm{m})} \left|\exp 2\alpha_0 - 2\alpha_0^{\mathrm{m}} - 1\right| \\
&= \left(Y_0^{(m)}\right)^2 \left|\exp\left(2\alpha_0 - 2\alpha_0^m\right) - 1\right|
\end{aligned}$$

Taylor expanding the exponential function about zero gives

$$\left|\exp\left(2\alpha_0 - 2\alpha_0^{(m)}\right) - 1\right| \leq 2\left|\alpha_0 - \alpha_0^{(m)}\right| \exp\left(2\delta_m\right),$$

where

$$|\delta_m| < |\alpha_0 - \alpha_0^m| = \left|\sum_{j=m+1}^{\infty} \psi_j \nu_{-j}\right|.$$

These bounds and the fact that $\exp\{2|a-b|\} \leq \exp\{2(a-b)\} + \exp\{2(b-a)\}$ *for all* a, b, *provide*

$$\begin{aligned}
E[|Q_m|] &\leq E[(Y_0^{(m)})^2] E\left[\left(\exp\{2|\alpha_0 - \alpha_0^m|\}\right) 2|\alpha_0 - \alpha_0^m|\right], \\
&\leq E[(Y_0^{(m)})^2] E\left[\left(\exp\left(2\alpha_0^m - 2\alpha_0\right) + \exp\left(2\alpha_0 - 2\alpha_0^m\right)\right) 2|\alpha_0 - \alpha_0^m|\right],
\end{aligned}$$

and the Cauchy-Schwarz inequality now yields

$$E[|Q_m|] \leq 2E[\left(Y_0^{(m)}\right)^2] \left(\left(M_m(4) + M_m(-4)\right) E\left[\left(\alpha_0 - \alpha_0^m\right)^2\right]\right)^{1/2}, \tag{5.3}$$

where $M_m(t)$ is the moment generating function of $\alpha_0 - \alpha_0^{(m)}$. For m_0 sufficiently large and $j > m_0$, $4|\psi_j| < 2$. Both $M_m(4)$ and $M_m(-4)$ are bounded for $m > m_0$.

For $m \leq m_0$, we have

$$E\left[\left|Y_0^2 - (Y_0^{(m)})^2\right|\right] \leq 2E[Y_0]^2 = C_1^2 < \infty. \tag{5.4}$$

Also notice that

$$\begin{aligned}
E\left[(\alpha_0 - \alpha_0^m)^2\right] &= E\left[\left(\sum_{j=m+1}^{\infty} \psi_j \nu_{-j}\right)\left(\sum_{k=m+1}^{\infty} \psi_k \nu_{-k}\right)\right], \\
&= E\left[\sum_{j=m+1}^{\infty}\sum_{k=m+1}^{\infty} \psi_j \psi_k \nu_{-j}\nu_{-k}\right], \\
&= \sum_{j=m+1}^{\infty}\sum_{k=m+1}^{\infty} \psi_j \psi_k E\left[\nu_{-j}\nu_{-k}\right], \\
&= \sum_{j=m+1}^{\infty} \psi_j^2 E\left[\nu_{-j}^2\right], \\
&= \sigma_\nu^2 \sum_{j=m+1}^{\infty} \psi_j^2.
\end{aligned}$$

Now causality gives,

$$\begin{aligned}
E\left[(\alpha_0 - \alpha_0^m)^2\right] &\leq c^2\sigma_\nu^2 \sum_{j=m+1}^{\infty} \gamma^{2j}, \\
&= c^2\sigma_\nu^2 \frac{\gamma^{2(m+1)}}{(1-\gamma^2)}
\end{aligned} \tag{5.5}$$

Combining (5.3), (5.4), and (5.5) illuminates finite constants C_1 and C_2 such that

$$\sum_{m=1}^{\infty} E[|Q_m|]^{1/2} \leq m_0 C_1 + C_2 \sum_{m>m_0} \gamma^{(m+1)/2} < \infty. \tag{5.6}$$

Theorem 2.1 now shows that the partial sums of arc length for $X_t = \ln(P_t)$ satisfy a FCLT.

5.2.2 Multivariate Models

Note that $\mathbf{Y}_t = (Y_{1,t}, \ldots, Y_{d,t})'$ is a d-variate series at time t. Detailed proofs for all mutivariate models listed in section 3.2 can be given as follows.

Proof(Example 5)

We verify the conditions for Theorem 2.1 when $d = 2$; arguments for higher dimensions are similar.

Causality allows us to write the solution to (3.5) as

$$\mathbf{Y}_t = \sum_{j=0}^{\infty} \mathbf{\Psi}_j \mathbf{Z}_{t-j},$$

where $\sum_{j=0}^{\infty} |\mathbf{\Psi}_j| < \infty$ in a component by component sense. For $\mathbf{\Psi}_j$, we denote the element in column ℓ of row k by $\mathbf{\Psi}_{k,\ell,j}$. Equation (3.5) gives

$$\begin{bmatrix} Y_{1,t} \\ Y_{2,t} \end{bmatrix} = \sum_{j=0}^{\infty} \begin{bmatrix} \psi_{1,1,j} Z_{1,t-j} + \psi_{1,2,j} Z_{2,t-j} \\ \psi_{2,1,j} Z_{1,t-j} + \psi_{2,2,j} Z_{2,t-j} \end{bmatrix}.$$

As in proof of Example 1, for $i = 1, 2$, define

$$Y_{i,0}^{(m)} = \sum_{j=0}^{m} \psi_{i,1,j} Z_{1,-j} + \psi_{i,2,j} Z_{2,-j}.$$

Now we define Q_m in a similar manner as it has been defined for univariate models.

$$\begin{aligned}
|Q_m| &= \left| Y_{1,0}^2 + Y_{2,0}^2 - (Y_{1,0}^{(m)})^2 - (Y_{2,0}^{(m)})^2 \right|, \\
&\leq \left| Y_{1,0}^2 - \left(Y_{1,0}^{(m)}\right)^2 \right| + \left| Y_{2,0}^2 - \left(Y_{2,0}^{(m)}\right)^2 \right| \\
&= \sum_{i=1}^{2} |Y_{i,0}^2 - \left(Y_{i,0}^m\right)^2| \\
&= \sum_{i=1}^{2} | \left(Y_{i,0} - Y_{i,0}^m\right) \left(Y_{i,0} + Y_{i,0}^m\right) | \\
&= \sum_{i=1}^{2} | \left(Y_{i,0} - Y_{i,0}^m\right) \left(Y_{i,0} - Y_{i,0}^m + 2Y_{i,0}^m\right) | \\
&= \sum_{i=1}^{2} | \left(Y_{i,0} - Y_{i,0}^m\right)^2 + 2 \left(Y_{i,0} - Y_{i,0}^m\right) Y_{i,0}^m| \\
&\leq \sum_{i=1}^{2} | \left(Y_{i,0} - Y_{i,0}^m\right)^2 | + 2| \left(Y_{i,0} - Y_{i,0}^m\right) Y_{i,0}^m|
\end{aligned}$$

The Cauchy-Schwarz inequality gives

$$
\begin{aligned}
E[|Q_m|] &\leq \sum_{i=1}^{2} E\left[\left| \left(Y_{i,0} - Y_{i,0}^{m}\right)^2 \right| + 2\left| \left(Y_{i,0} - Y_{i,0}^{m}\right) Y_{i,0}^{m} \right| \right], \\
&= \sum_{i=1}^{2} \left(E\left[\left(\Delta_i^{(m)}\right)^2 \right] + 2E\left[|Y_{i,0}^{m}||\Delta_i^{(m)}| \right] \right), \\
&\leq \sum_{i=1}^{2} \left(E\left[\left(\Delta_i^{(m)}\right)^2 \right] + 2\left(E\left[Y_{i,0}^2\right] E\left[(\Delta_i^{(m)})^2 \right] \right)^{1/2} \right),
\end{aligned}
$$

where $\Delta_i^{(m)} = Y_{i,0} - Y_{i,0}^{(m)}$. *Observe that for* $i = 1, 2$,

$$
\begin{aligned}
E\left[(\Delta_i^{(m)})^2 \right] &= E\left[\left(Y_{i,0} - Y_{i,0}^{(m)} \right)^2 \right] \\
&= E\left[\left(\sum_{j=m+1}^{\infty} \psi_{i,1,j} Z_{1,-j} + \psi_{i,2,j} Z_{2,-j} \right) \left(\sum_{k=m+1}^{\infty} \psi_{i,1,k} Z_{1,-k} + \psi_{i,2,k} Z_{2,-k} \right) \right] \\
&= E\Bigg[\sum_{j=m+1}^{\infty} \sum_{k=m+1}^{\infty} \psi_{i,1,j}\psi_{i,1,k} Z_{1,-j} Z_{1,-k} + \sum_{j=m+1}^{\infty} \sum_{k=m+1}^{\infty} \psi_{i,2,j}\psi_{i,2,k} Z_{2,-j} Z_{2,-k} + \\
&\quad 2 \sum_{j=m+1}^{\infty} \sum_{k=m+1}^{\infty} \psi_{i,1,j}\psi_{i,2,k} Z_{1,-j} Z_{2,-k} \Bigg] \\
&= \sum_{j=m+1}^{\infty} \psi_{i,1,j}^2 E\left[Z_{1,-j}^2\right] + \sum_{j=m+1}^{\infty} \psi_{i,2,j}^2 E\left[Z_{2,-j}^2\right] + 2 \sum_{j=m+1}^{\infty} \psi_{i,1,j}\psi_{i,2,j} E\left[Z_{1,-j} Z_{2,-j}\right] \\
&= \sum_{j=m+1}^{\infty} \left(\psi_{i,1,j}^2 \sigma_{1,1} + \psi_{i,2,j}^2 \sigma_{2,2} + 2\psi_{i,1,j}\psi_{i,2,j}\sigma_{1,2} \right),
\end{aligned}
$$

where $\sigma_{l,k} = \mathrm{Cov}(Z_{l,0}, Z_{k,0})$.

Causality implies the bound $|\psi_{i,k,j}| \leq c\gamma^j$ *for some constants* $c > 0$ *and* $0 < \gamma < 1$. *This bound holds uniformly in* i *and* k. *Using this gives*

$$
\begin{aligned}
E\left[(\Delta_i^{(m)})^2 \right] &\leq \left(\sigma_{1,1} + \sigma_{2,2} + 2\sigma_{1,2} \right) \sum_{j=m+1}^{\infty} \gamma^{2j} \\
&= \left(\sigma_{1,1} + \sigma_{2,2} + 2\sigma_{1,2} \right) \frac{\gamma^{2(m+1)}}{1-\gamma^2}. \qquad (5.7)
\end{aligned}
$$

Hence for all $i = 1, 2$

$$\begin{aligned}\sum_{m=1}^{\infty} E\left[(\Delta_i^{(m)})^2\right]^{1/2} &\leq \frac{\left(\sigma_{1,1}+\sigma_{2,2}+2\sigma_{1,2}\right)^{1/2}}{(1-\gamma^2)^{1/2}} \sum_{m=1}^{\infty} \gamma^{(m+1)}, \\ &= \frac{\left(\sigma_{1,1}+\sigma_{2,2}+2\sigma_{1,2}\right)^{1/2}\gamma^2}{(1-\gamma^2)^{1/2}(1-\gamma)},\end{aligned}$$

and

$$\begin{aligned}\sum_{m=1}^{\infty} E\left[(\Delta_i^{(m)})^2\right]^{1/4} &\leq \frac{\left(\sigma_{1,1}+\sigma_{2,2}+2\sigma_{1,2}\right)^{1/4}}{(1-\gamma^2)^{1/4}} \sum_{m=1}^{\infty} \gamma^{(m+1)/2}, \\ &= \frac{\left(\sigma_{1,1}+\sigma_{2,2}+2\sigma_{1,2}\right)^{1/4}\gamma}{(1-\gamma^2)^{1/4}\left(1-\gamma^{1/2}\right)}.\end{aligned}$$

Also based on the same arguments we made above it can be shown that

$$\begin{aligned}E\left[Y_{i,0}^2\right] &= E\left[\left(\sum_{j=0}^{\infty}\psi_{i,1,j}Z_{1,-j}+\psi_{i,2,j}Z_{2,-j}\right)\left(\sum_{k=0}^{\infty}\psi_{i,1,k}Z_{1,-k}+\psi_{i,2,k}Z_{2,-k}\right)\right] \\ &= E\Bigg[\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\psi_{i,1,j}\psi_{i,1,k}Z_{1,-j}Z_{1,-k}+\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\psi_{i,2,j}\psi_{i,2,k}Z_{2,-j}Z_{2,-k}+ \\ &\quad 2\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\psi_{i,1,j}\psi_{i,2,k}Z_{1,-j}Z_{2,-k}\Bigg] \\ &= \sum_{j=0}^{\infty}\psi_{i,1,j}^2E\left[Z_{1,-j}^2\right]+\sum_{j=0}^{\infty}\psi_{i,2,j}^2E\left[Z_{2,-j}^2\right]+2\sum_{j=0}^{\infty}\psi_{i,1,j}\psi_{i,2,j}E\left[Z_{1,-j}Z_{2,-j}\right] \\ &= \sum_{j=0}^{\infty}\left(\psi_{i,1,j}^2\sigma_{1,1}+\psi_{i,2,j}^2\sigma_{2,2}+2\psi_{i,1,j}\psi_{i,2,j}\sigma_{1,2}\right),\end{aligned}$$

and as a result of the causality, we get

$$\begin{aligned}E\left[Y_{i,0}^2\right] &\leq \left(\sigma_{1,1}+\sigma_{2,2}+2\sigma_{1,2}\right)\sum_{j=0}^{\infty}\gamma^{2j} \\ &= \frac{\left(\sigma_{1,1}+\sigma_{2,2}+2\sigma_{1,2}\right)}{1-\gamma^2} < \infty.\end{aligned}$$

Now using all the pieces showed above, it follows that $\sum_{m=1}^{\infty} E[|Q_m|]^{1/2} < \infty$. *Theorem 2.1 now allows us to infer that the sample arc lengths of the log-price vectors satisfy a FCLT.*

Proof(Example 6)

For ease of presentation, we restrict attention to $d = 2$, the arguments being similar for larger d. Then $\mathbf{Y}_t = (Y_{1,t}, Y_{2,t})'$ satisfies

$$Y_{1,t}^2 + Y_{2,t}^2 = \sigma_{1,1,t} Z_{1,t}^2 + \sigma_{2,2,t} Z_{2,t}^2 + 2\sigma_{1,2,t} Z_{1,t} Z_{2,t},$$

where

$$\sigma_{1,2,0} = \rho_{1,2}\sqrt{\sigma_{11,0}\sigma_{2,2,0}}.$$

Lemma 2.1 of Berkes et al.(2008) shows that the unique stationary solution of $\sigma_{i,i,0}$ for $i = 1, 2$ can be expressed as

$$\sigma_{i,i,0} = \omega_i \left(\sum_{k=1}^{\infty} \sum_{1 \le l_1, \ldots, l_{k-1} \le p} \prod_{j=1}^{k-1} \left(\beta_{il_j} + \alpha_{il_j} Z_{i,-l_1-l_2-\cdots-l_j}^2 \right) \right).$$

Strict stationarity ensures that for each i,

$$c_i = \left[(\alpha_{i,1} + \cdots + \alpha_{i,p}) E[\varepsilon_{i,0}^2] + (\beta_{i,1} + \cdots + \beta_{i,p}) \right] < 1.$$

Also note that $E[\sigma_{i,i,0}] = (1 - c_i)^{-1}\omega_i < \infty$ for each i.

Mimicking the univariate GARCH arguments, define an mth order truncation as

$$\left(Y_{1,0}^{(m)}\right)^2 + \left(Y_{2,0}^{(m)}\right)^2 = \sigma_{1,1,0}^{(m)} Z_{1,0}^2 + \sigma_{2,2,0}^{(m)} Z_{2,0}^2 + 2\sigma_{1,2,0}^{(m)} Z_{1,0} Z_{2,0},$$

where

$$\begin{aligned} \sigma_{i,i,0}^{(m)} &= \omega_i \left(\sum_{k=1}^{\ell} \sum_{1 \le l_1, \ldots, l_{k-1} \le p} \prod_{j=1}^{k-1} \left(\beta_{il_j} + \alpha_{il_j} \varepsilon_{i,-l_1-l_2-\cdots-l_j}^2 \right) \right); \ i = 1, 2, \\ \sigma_{1,2,0}^{(m)} &= \rho_{1,2}\sqrt{\sigma_{11,0}^{(m)}\sigma_{2,2,0}^{(m)}}. \end{aligned}$$

Here, ℓ is such that $m = p\ell + r$, where ℓ and r are integers satisfying $\ell \ge 0$ and $0 \le r \le p - 1$. As in the univariate arguments, we again take $p = q$ in the GARCH representations of $\{\sigma_{i,i,t}\}_{i=1,2}$ as this can be done without loss of generality.

Observe that

$$\begin{aligned} Q_m &= Y_{1,0}^2 + Y_{2,0}^2 - (Y_{1,0}^{(m)})^2 - (Y_{2,0}^{(m)})^2 \\ &= \left(\sigma_{1,1,0} - \sigma_{1,1,0}^{(m)}\right) Z_{1,0}^2 + \left(\sigma_{2,2,0} - \sigma_{2,2,0}^{(m)}\right) Z_{2,0}^2 + \left(\sigma_{1,2,0} - \sigma_{1,2,0}^{(m)}\right) Z_{1,0} Z_{2,0}. \end{aligned}$$

Hence,

$$\begin{aligned} \sum_{m=1}^{\infty} E[Q_m]^{1/2} &\leq \sum_{m=1}^{\infty} \left(E\left[\sigma_{1,1,0} - \sigma_{1,1,0}^{(m)}\right] E\left[Z_{1,0}^2\right] \right)^{1/2} \\ &+ \sum_{m=1}^{\infty} \left(E\left[\sigma_{2,2,0} - \sigma_{2,2,0}^{(m)}\right] E\left[Z_{2,0}^2\right] \right)^{1/2} \\ &+ \sum_{m=1}^{\infty} \left(E\left[\sigma_{1,2,0} - \sigma_{1,2,0}^{(m)}\right] E\left[Z_{1,0} Z_{2,0}\right] \right)^{1/2}. \end{aligned} \tag{5.8}$$

The first two sums on the right hand side of (5.8) are summable by the arguments in Example 2. To handle the third term, consider $\sigma_{1,2,0} - \sigma_{1,2,0}^{(m)}$*. As shown in Berkes et. al(2008),* $\sigma_{i,i,0} - \sigma_{i,i,0}^{(m)} > 0$ *for each* $m > 0$*. Hence,*

$$\begin{aligned} \sigma_{1,2,0} - \sigma_{1,2,0}^{(m)} &= \rho_{1,2} \left(\sqrt{\sigma_{1,1,0}\sigma_{2,2,0}} - \sqrt{\sigma_{1,1,0}^{(m)}\sigma_{2,2,0}^{(m)}} \right) \\ &\leq \rho_{1,2} \left(\sigma_{1,1,0}\sigma_{2,2,0} - \sigma_{1,1,0}^{(m)}\sigma_{2,2,0}^{(m)} \right)^{1/2} \\ &\leq \rho_{1,2} \left\{ \left[\sigma_{1,1,0} \left(\sigma_{2,2,0} - \sigma_{2,2,0}^{(m)} \right) \right]^{1/2} + \left[\sigma_{2,2,0}^{(m)} \left(\sigma_{1,1,0} - \sigma_{1,1,0}^{(m)} \right) \right]^{1/2} \right\} \\ &\leq \rho_{1,2} \left\{ \left[\sigma_{1,1,0} \left(\sigma_{2,2,0} - \sigma_{2,2,0}^{(m)} \right) \right]^{1/2} + \left[\sigma_{2,2,0} \left(\sigma_{1,1,0} - \sigma_{1,1,0}^{(m)} \right) \right]^{1/2} \right\}. \end{aligned}$$

Applying the Cauchy-Schwarz inequality gives

$$\begin{aligned} E\left[\sigma_{1,2,0} - \sigma_{1,2,0}^{(m)}\right] &\leq \rho_{1,2} \left(E\left[\sigma_{1,1,0}\right] E\left[\sigma_{2,2,0} - \sigma_{2,2,0}^{(m)}\right] \right)^{1/2} \\ &+ \rho_{1,2} \left(E\left[\sigma_{2,2,0}\right] E\left[\sigma_{1,1,0} - \sigma_{1,1,0}^{(m)}\right] \right)^{1/2}. \end{aligned} \tag{5.9}$$

Equation (5.9) shows existence of finite constants C_1 *and* C_2 *such that*

$$\sum_{m=1}^{\infty} \left(E\left[\sigma_{1,2,0} - \sigma_{1,2,0}^{(m)}\right] E\left[Z_{1,0} Z_{2,0}\right] \right)^{1/2} \leq C_1 \sum_{m=1}^{\infty} \left(E\left[\sigma_{2,2,0} - \sigma_{2,2,0}^{(m)}\right] \right)^{1/4}$$

$$+ \quad C_2 \sum_{m=1}^{\infty} \left(E\left[\sigma_{1,1,0} - \sigma_{1,1,0}^{(m)} \right] \right)^{1/4}.$$

Note that for $i = 1, 2$,

$$\sum_{m=1}^{\infty} E\left[\sigma_{i,i,0} - \sigma_{i,i,0}^{(m)} \right]^{1/4} \leq \frac{\omega_i^{1/4} c_i^{1/4}}{(1-c_i)^{1/4}(1-c_i^{1/4})}.$$

As $E[Z_{i,0}^2] < \infty$, $E[\sigma_{i,i,0}] < \infty$ *for* $i = 1, 2$ *and* $E[Z_{1,0}Z_{2,0}] < \infty$, *we have* $\sum_{m=1}^{\infty} E[Q_m]^{1/2} < \infty$. *By Theorem 2.1, the arc lengths of the log prices* $\{\mathbf{X}_t\}$ *satisfy the FCLT when the log-returns,* $\{\mathbf{Y}_t\}$, *follow the CCC-GARCH model.*

Proof(Example 7)

Definition B.2 1 *The d-variate series* $\{\mathbf{Y}_t\} = (Y_{1,t}, \ldots, Y_{d,t})'$ *follows a MSV if it satisfies*

$$\mathbf{Y}_t = \mathbf{H}_t^{1/2}\mathbf{Z}_t, \tag{5.10}$$

$$\mathbf{H}_t^{1/2} = Diag(\sigma_{1,t}, \ldots, \sigma_{d,t}),$$

$$\ln(\sigma_{\mathbf{t}}) = \mu + \mathbf{\Phi} \circ \ln(\sigma_{t-1}) + \mathbf{V}_t, \tag{5.11}$$

$$\begin{pmatrix} \mathbf{Z}_t \\ \mathbf{V}_t \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \mathbf{P}_\varepsilon & 0 \\ 0 & \mathbf{\Sigma}_\nu \end{pmatrix} \right], \tag{5.12}$$

where $\mathbf{Z}_t = (Z_{1,t}, \ldots, Z_{d,t})'$ *and* $\mathbf{V}_t = (\nu_{1,t}, \ldots, \nu_{d,t})'$ *are random vectors,* $\sigma_t = (\sigma_{1,t}, \ldots, \sigma_{d,t})'$ *is a d-dimensional vector of unobserved volatilities,* μ *and* $\mathbf{\Phi}$ *are* $d \times 1$ *parameter vectors, the operator* $\circ$ *denotes the Schur element-by-element product,* $\Sigma_\nu = \{\sigma_{\nu,i,j}\}_{i,j=1}^d$ *is a positive definite covariance matrix, and* $\mathbf{P}_\varepsilon$ *is the correlation matrix.*

Since Equation 5.11 has a causal stationary solution, we can express $\sigma_{\mathbf{t}}$ *as*

$$\sigma_t = \exp\left(\mu(\mathbf{I} - \mathbf{\Phi})^{-1} + \sum_{j=0}^{\infty} \mathbf{\Psi}_j \circ \mathbf{V}_{t-j} \right), \tag{5.13}$$

where $\mathbf{\Psi}_j = (\psi_{1,j}, \ldots, \psi_{d,j})'$ *is a sequence of deterministic vectors with* $\mathbf{\Psi}_0 = (1, \ldots, 1)'$. *The next example considers a more general model.*

For an iid sequence $\{\mathbf{Z}_t\}$, suppose $\{\mathbf{Y}_t\}$ obeys

$$\mathbf{Y}_t = \mathbf{H}_t^{1/2}\mathbf{Z}_t,$$

with $\mathbf{H}_t^{1/2} = Diag(\sigma_{1,t}, \ldots, \sigma_{d,t})$. Also let $\sigma_{i,t} = \exp\{W_{i,t}\}$, where the d-variate process $\{\mathbf{W}_t\}$ follows a causal vector autoregression. We now verify the conditions in Theorem 2.1 assuming $d = 2$; arguments are similar for higher dimensions.

As in Example 5, we invoke causality to write

$$\mathbf{W}_t = \sum_{j=0}^{\infty} \mathbf{\Psi}_j \varepsilon_{t-j}.$$

We further assume an iid $\{\varepsilon_t\}$. Causality implies that $|\psi_{i,k,j}| \leq c\gamma^j$ and, as before, one can take $c < 2$ without loss of generality. Our bivariate system can be written as

$$Y_{1,t} = \sigma_{1,t}Z_{1,t}, \quad \sigma_{1,t} = \exp\left(\sum_{j=0}^{\infty} \psi_{1,1,j}\varepsilon_{1,\mathrm{t-j}} + \psi_{1,2,\mathrm{j}}\varepsilon_{2,\mathrm{t-j}}\right),$$

$$Y_{2,t} = \sigma_{2,t}Z_{2,t}, \quad \sigma_{2,t} = \exp\left(\sum_{j=0}^{\infty} \psi_{2,1,j}\varepsilon_{1,\mathrm{t-j}} + \psi_{2,2,\mathrm{j}}\varepsilon_{2,\mathrm{t-j}}\right).$$

As in Example 3, moment conditions will be assumed. Suppose that the moment generating function of ε_t exists in an appropriate neighborhood of zero; specifically, assume

$$E[\exp\left(t_1\varepsilon_{1,0} + \mathrm{t}_2\varepsilon_{2,0}\right)] < \infty$$

for $|t_1| \leq 2$ and $|t_2| \leq 2$. Let

$$\mathrm{Y}_{1,t}^{(m)} = Z_{1,t}\exp\left(\sum_{j=0}^{m} \psi_{1,1,j}\varepsilon_{1,\mathrm{t-j}} + \psi_{1,2,\mathrm{j}}\varepsilon_{2,\mathrm{t-j}}\right).$$

and define $Y_{2,t}^{(m)}$ similarly. As in the proof of Example 5,

$$|Q_m| \quad = \quad \left|Y_{1,0}^2 + Y_{2,0}^2 - (Y_{1,0}^{(m)})^2 - (Y_{2,0}^{(m)})^2\right|,$$

$$\leq \quad \left|Y_{1,0}^2 - \left(Y_{1,0}^{(m)}\right)^2\right| + \left|Y_{2,0}^2 - \left(Y_{2,0}^{(m)}\right)^2\right|.$$

To verify Assumption 2.2, note that

$$E[|Q_m|] \leq E\left|Y_{1,0}^2 - \left(Y_{1,0}^{(m)}\right)^2\right| + E\left|Y_{2,0}^2 - \left(Y_{2,0}^{(m)}\right)^2\right|$$

and

$$E[|Q_m|]^{1/2} \leq E\left[|Y_{1,t}^2 - (Y_{1,t}^{(m)})^2|\right]^{1/2} + E\left[|Y_{2,t}^2 - (Y_{2,t}^{(m)})^2|\right]^{1/2} \tag{5.14}$$

Note that for all $i = 1, 2$

$$E\left[|Y_{i,t}^2 - (Y_{i,t}^{(m)})^2|\right] = E\left[Z_{i,t}^2 \exp\left(2\sum_{j=0}^{m} \psi_{i,1,j}\varepsilon_{\mathrm{i,t-j}} + \psi_{\mathrm{i,2,j}}\varepsilon_{\mathrm{i,t-j}}\right)\right],$$

and observe that each term on the right hand side of (5.14) is summable (in m) by the argument presented in Example 3. Applying Theorem 2.1 shows that the sample arc lengths of log prices satisfy a FCLT.

Bibliography

Asai, M., McAleer, M., and Yu, J. (2006). Multivariate stochastic volatility: A review. *Econometric Review*, (25):145–175.

Bauwens, L., Laurent, S., and Rombouts, J. V. K. (2006). Multivariate garch models: A survey. *Journal of Applied Econometrics*, (21):79–109.

Berkes, I., Gombay, E., and Horváth, L. (2009). Testing for changes in the covariance structure of linear processes. *Journal of Statistical Planning and Inference*, (139):2044–2063.

Berkes, I., Hörmann, S., and Horváth, L. (2008). The functional central limit theorem for a family of garch observations with applications. *Statistics and Probability Letters*, (78):2725–2730.

Berkes, I., Horváth, L., and Kokoszka, P. (2004a). Probabilistic and statistical properties of garch processes. in: Asymptotic methods in stochastics. in: Fiels inst. commun. *Amer. Math. Soc.*, (44):409–425.

Berkes, I., Horváth, L., and Kokoszka, P. (2004b). Testing for parameter consistency in garch(p, q) models. *Statistics & Probability Letters*, (70):263–273.

Billingsley, P. (1968). *Convergence of probability Measures.* Wiley, New York.

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

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

Bougerol, P. and Picard, N. (1992). Stationarity of garch processes and some nonnegative time seriesl. *J. Econometrics*, (52):115–127.

Brockwell, P. J. and Davis, R. A. (1991). *Time series: Theory and Methods.* Springer Science+Business Media LLC., second edition.

Brockwell, P. J. and Davis, R. A. (2002). *Introduction to Time Series and Forecasting.* Springer-Verlag New York., second edition.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, (1):223–236.

Davidson, J. (1994). *Stochastic Limit Theory; An Introduction for Econometricians.* Oxford University Press Inc., New York.

Davis, R. A. and Mikosch, T. (2009). *Handbook of Financial Time Series.* Springer-Verlag, Berlin Heidelberg.

Doukhan, P. and Louhichi, S. (1999). A new weak dependence condition and applications to moment inequalities. *Stochastic Processes and their Applications*, (84):313–342.

Doukhan, P. and Wintenberger, O. (2007). An invariance principle for weakly dependent stationary general models. *Probab. Math. Statist.*, (27):45–73.

F., M., Peligrad, M., and Utev, S. (2006). Recent advances in invariance principles for stationary sequences. *Probability Surveys*, 34:1–36.

Giraitis, L., Leipus, R., and Surgailis, D. (2007). *Recent Advances in ARCH Modeling. In: Long Memory in Economics.* Springer.

Harvey, A. C., Ruiz, E., and Shephard, N. (1994). Multivariate stochastic variance models. *Review of Economic Studies*, (61):247–264.

Mercurio, D. and Spokoiny, V. (2004). Statistical inference for time-inhomogeneous volatility models. *Annals of Statistics*, (32):577–602.

Robbins, M., Gallagher, C., Lund, R., and Aue, A. (2011). Mean shift testing in correlated data. *Journal of Time Series Analysis*, (32):Page numbers TBA.

Ross, S. M. (2007). *Introduction to Probability Models.* Elsevier Inc., nineth edition.

Taylor, S. J. (2005). *Asset Price Dynamics, Volatility, and Prediction.* Princeton University Press, Princeton and Oxford.