# AN ANALYSIS OF THE STABILITY IN MULTIVARIATE CORRELATION STRUCTURES

by

**YUQIAN ZHAO**

A thesis submitted to the University of Birmingham
for the degree of DOCTOR OF PHILOSOPHY

Department of Economics
Birmingham Business School
College of Social Sciences
University of Birmingham
June 2017

# Abstract

Analysing the instability in the multivariate correlation structure, the present thesis starts from assessing in-sample and out-of-sample performances of multivariate GARCH models with or without a structural break. The result emphasizes the importance of correlation change point detection for model fittings. We then propose semi–parametric CUSUM tests to detect a change point in the covariance structures of non–linear multivariate models with dynamically evolving volatilities and correlations. The asymptotic distributions of the proposed statistics are derived under mild conditions. We discuss the applicability of our method to the most often used models, including constant conditional correlation, dynamic conditional correlation (DCC), BEKK, factor, asymmetric DCC/BEKK processes. Our simulations show that, even though the nearly unit root property distorts the size and power of tests, the standardization of the data with conditional standard deviations in multivariate volatility models can correct such distortions. Lastly, concerning classical trimmed issue in change point test, we extend the semi-parametric CUSUM to weighted CUSUM tests, which enhances the power across either ends of a sample. The asymptotic limit of weighted CUSUM tests are also derived. A Monte Carlo simulation study suggests that weighted CUSUM tests exhibit better performances than unweighted ones in finite samples. Regarding empirical applications, we show the absorption ratio is a leading indicator of the financial fragility, and we study global financial contagion effect, also we investigate unexpected events in the U.S. equity market.

To My Parents

# Acknowledgements

I would like to express my special appreciation and thanks to my supervisor Dr. Marco R. Barassi, who has been a tremendous mentor for me. I would like to thank him for encouraging my research and for allowing me to grow as a research scientist. His advice on both research as well as on my life have been invaluable.

I would also like to thank Prof. David Maddison, who awarded me the PhD scholarship. I can never chase my dream without this funding. I thank my second supervisor Mr. Nicholas Housewood, who is a good friend and encourages me during last four years.

I truly thank my advisor Professor Zhenya Liu for his professional training and brilliant comments. He brings me to a high-standard platform and shows me the wise direction in the dark. I would especially like to thank Professor Lajos Horváth from University of Utah, U.S.A., who trains me statistics with high tolerance and patience. Without his supports, I can never complete a thesis in such a level. And I also thank the selfness consultants given by Dr. William Pouliot, Professor Anindya Banerjee. I also want to thank who attended my presentations in Ph.D. workshops for their comments and supports. I thank Yi Liu and Rong Huang for their I.T. supports, also all my colleague and friends for their kind company.

A special thanks to my family. Words can not express how grateful I am to my father and mother for all of the sacrifices that they have made on my behalf. They give me power and enhance my faith, letting me through all the difficulties. Finally I thank my grandfathers and grandmothers for their prayer for me was what sustained me thus far.

# Preface

Data science has become increasingly crucial to many subjects across both natural and social science since the availability of data is explosive in both dimension and frequency. Concomitantly, advanced technology brings us to an age of boosted computational ability, thereby leading to low costs in analysing multivariate data. Being different from univariate data, multivariate analysis can be complicated because of their cross-relationships, especially in the second moment. Modelling a stationary variance covariance has been developed for decades and has reached a relatively efficient and accurate level. However, keeping stationarity in the second moment in the long run is somewhat impractical in such an unstable global economic environment. Therefore, studying instability in the second moment is vital to multivariate data analysis. The present thesis devotes itself to contribute on this topic, providing a survey study and developing multivariate volatility change-point tests with several financial applications.

The pioneering work to model the second moment in univariate financial data was completed by Engle (1982), who proposed the ARCH (Autoregressive Conditional Heteroskedasticity) model to explain the effect of volatility clustering in financial data. The conditional variance in the ARCH model quickly became the prevailing method to measure risk in finance. But empirically, the ARCH model suffers from a severe trade-off between over-fitting and the nature of a long memory process. This limitation is overcome by the GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model proposed by Bollerslev (1986), and later extended to different versions to form a GARCH family.

Nevertheless, univariate GARCH models were empirically inadequate since the multiple risk factor model became essential to asset pricing. In the 1980s, the well-known CAPM model was argued by adding other risk factors such as size (Banz, 1981), leverage effect (Bhandari, 1988) and time-varying covariance (Bollerslev et al., 1988). The multivariate GARCH model then started its lengthy and extensive developments. The detailed literature will be reviewed in Chapter 1.

Then, modelling the second moment of a set of variables in equity markets, bond markets or currency markets usually suggests investment and policy implications. For example, Karolyi (1995) constructed a bivariate system between the United States and Canada based on the framework of the multivariate GARCH-in-mean model and found significant volatility spillover effects in the two countries. Also, the dependency in the second moment can be used to measure systemic risk (Billio et al., 2012). Especially after the Great Recession in 2008, it raised plenty of researches on inspecting the contagion effect and systematic risk. For instance, Gatfaoui (2013) applied the BEKK model to measure systemic risk and spillover effect among the U.S., French and U.K.

In reality, considering the efficiency of modelling a multivariate random vector, economic or financial relationships can never be stable either in the first moment or the second moment, due to shocks (burst of a bubble, policy change and other innovations) or the financial contagion effect in the market. As a result, instability in the second moment influences on the behaviour of multivariate GARCH models, and providing poor modelling. Thus, detecting instability in the second moment is an imperative task in theoretical and empirical multivariate data analysis. With regard to the literature in the detection of structural break/change point, it has cumulated a large amount to test the instability of mean and variance for univariate case, while the study in the multivariate random vector is still at a primitive stage. Existing literature in this area mainly considers instability in the covariance structure, while the instability in co-volatilities can be a result of either volatility changes or correlation changes. Considering the correlation structure plays an intuitively important role in economics and finance, it is motivated and necessary to fulfil this marginal gap. More related literature is discussed in Chapters 3 and 4.

The present thesis is structured as below. Chapter 1 compares representative multivariate GARCH models for their in-sample and out-of-sample performances. By manipulating the correlation structure, a Monte Carlo simulation study assesses the performances of multivariate GARCH models in cases with or without a single structural break in finite samples. Chapter 2 provides an empirical application. By predicting the market conditional covariance matrices, a concentration index - absorption ratio is extracted from the predicted covariance matrix. The analysis shows that absorption ratio can be used as a leading indicator for the financial fragility, providing pre-measurement of systemic risk in the U.S. equity market. Also, the second chapter tries to detect change points in absorption ratio process as a proxy of the instability in the covariance structure.

Chapters 3 and 4 develop change-point tests for testing the instability in the correlation structure. The third chapter generalises the CUSUM tests (Aue et al. 2009) to semi-parametric CUSUM tests to detect change points in multivariate observations with both dynamic evolving variance and correlation structures. The application section provides a study of the global financial contagion effect. Lastly, concerning the trimming issue of change-point tests argued by Andrews (1993), Chapter 4 extends semi-parametric CUSUM tests to semi-parametric weighted CUSUM tests. An application of detecting unexpected events in the U.S. equity market is discussed.

This thesis contributes to the literature in statistics and finance fields: a comparison of representative multivariate GARCH models; change-point detection in time-varying correlation structures; weighted approximation of correlation change-point tests in dynamic correlation structures; finding a leading indicator for systemic risk in equity markets; identifying cross-market financial contagion effects with unknown break dates; detecting systemic events in equity markets.

# Contents

# List of Tables

# List of Figures

# Abbreviations and Notations

**ARCH** ....... Autoregressive Conditional Heteroskedasticity Model

**GARCH** ..... Generalized ARCH Model

**M-GARCH** .. Multivariate GARCH Model

**EWMA** ...... Exponential Weighted Moving Average Model

**VEC** .......... Vector Representative GARCH Model

**DVEC** ........ Diagonal VEC Model

**BEKK** ........ Baba, Engle, Kraft and Kroner GARCH Model

**DBEKK** ...... Diagonal Baba, Engle, Kraft and Kroner GARCH Model

**CCC** .......... Constant Conditional Correlation Model

**OGARCH** ... Orthogonal GARCH Model

**DCC** ......... Dynamic Conditional Correlation Model

**ADCC** ........ Asymmetric Dynamic Conditional Correlation Model

**RDCC** ........ Rotated Dynamic Conditional Correlation Model

**ARDCC** ...... Asymmetric Rotated Dynamic Conditional Correlation Model

**DCCSB** ...... Dynamic Conditional Correlation with Structural Break

**ADCCSB** .... Asymmetric Dynamic Conditional Correlation Model with Structural Break

**RDCCSB** .... Rotated Dynamic Conditional Correlation Model with Structural Break

**ARDCCSB** .. Asymmetric Rotated Dynamic Conditional Correlation Model with Structural Break

**QMLE** ........ Quasi Maximum Likelihood Estimation

**LL** ............ Log Likelihood

**AIC** .......... Akaike Information Criterion

**BIC** ........... Bayesian Information Criterion

**MAFE** ....... Mean Absolute Forecasting Error

**MSFE** ........ Mean Squared Forecasting Error

**PCA** .......... Principal Component Analysis

**AR** ........... Absorption Ratio

**CUSUM** ..... Cumulative Sum

**WCUSUM** ... Weighted Cumulative Sum

**QR** ........... Quantile Regression

**cf** ............. Call For

$\mathcal{M}$ ............ Model

$\sigma$ .............. Conditional Variance Scaler

$\mathbf{\Sigma}$ ............. Conditional Covariance Matrix

$\mathbf{D}$ ............. Diagonal Conditional Variance Matrix

$\mathbf{R}$ ............. Conditional Correlation Matrix

$\mathbf{Q}$ ............. Quasi Conditional Correlation Matrix

$\Lambda$ .............. Diagonal Matrix Composed by Eigenvalues

$vech$ ........... Matrix Stacking Operation

$\circ$ .............. Hadamard Product

$\bigotimes$ ............. Kronecker Product

$\cdot$ ............... Product

$\mathcal{F}$ ............. Filtration

$T$ .............. Sample Size

$K$ .............. Number of Parameters

$d$ .............. Dimensionality

$t$ .............. Time Index

$t^*$ .............. Break Point Location in $t$

$\mathfrak{t}$ .............. Fractional Time Index where $0 < \mathfrak{t} < 1$

$\mathfrak{t}^*$ .............. Break Point Location in $\mathfrak{t}$

$\mathbf{W}$ .............. Standard Brownian Motion

$\mathbf{B}$ .............. Standard Brownian Bridge

$q(\cdot)$ .............. Weight Function

$\mathfrak{D}$ .............. Long Run Covariance

$\Gamma(\cdot)$ .............. Gamma Function

$\mathbb{Z}$ .............. Set of Integers

$\mathbb{R}$ .............. Set of Real Numbers

$O$ .............. Stochastic Boundedness

$o$ .............. Convergence

$O_P$ .............. Stochastic Boundedness in Probability

$o_P$ .............. Convergence in Probability

$sup$ .............. Supremum

$max$ .............. Maximum

$\lfloor\rfloor$ .............. Integer Part

# Chapter 1

# A Comparison of In-sample and Out-of-sample Performance of Multivariate GARCH Models

## 1.1 Introduction

Selecting an appropriate multivariate GARCH model for empirical applications is always a hard task as plenty of them are available for the purpose of modelling historical covariance relationships or forecasting covariance structure. In this paper, we briefly categorise M-GARCH models into two streams according to their model specifications. The first type is directly derived by generalising from the univariate GARCH, which is intuitive, but suffers from problems of positive definite matrix and heavy parameterisation. Another type is relatively parsimonious, which is obtained by linearly or nonlinearly combining dynamic components from the second moment. Survey papers were completed by Bauwens et al. (2006), Engle (2009) and Silvennoinen and Teräsvirta (2009), and for statistical inference, we refer to Francq and Zakoian (2010).

Plenty of covariance models make the pool of candidates, empirical researchers care more about which M-GARCH model is the most appropriate for their particular studies. Among the survey, Engle (2009) provided comparisons between the most popular models on in-sample performance. Laurent et al. (2012) compared the performance of M-GARCH models in the term of their forecasting ability. To extend their contributions, this paper compares in-sample performances of some representative M-GARCH models using series with different correlation structures and may also display a structural break. Also, we particularly assess the forecasting ability of these M-GARCH models in terms of predicting the evolution of the correlation structure.

The candidate models considered begin from those in the early work by Bollerslev et al. (1988), who represented conditional cross-volatilities by using a stacking operator on a matrix version of the GARCH model, obtaining the VEC-GARCH model. However, this model is not well specified in theory as it is difficult to keep the covariance matrix as positive definite. Also, the number of parameters in the VEC model increases exponentially along with increasing dimensionality. To overcome the issue of heavy parameterisation, Ding and Engle (2001) assumed that there exists diagonal coefficient matrices and extended the VEC to the Diagonal VEC (Diag/DVEC) model, although such specification does not help much in a large system. As regard to the issue of ensuring a positive definite covariance matrix, an important milestone is the construction of the BEKK model (Engle and Kroner, 1995), which provides a positive definite covariance matrix under weak conditions. Compared to previous methods, the BEKK model also benefits from slight lighter parametrisation.

Even then, considering the cost and the relatively inefficient computational ability of that era, all models in this first category can hardly be applied to a system with more than four variables without encountering computational problems. Thus, inspired by the idea of the factor model, Engle et al. (1990) applied a common dynamic structure into a conditional covariance matrix and proposed the linear factor GARCH model, where the Factor GARCH is a special case of the BEKK model (cf. Francq and Zakoian, 2010). The dynamics in covariance structure can be caught by a small number of common factors, resulting in a useful model in a larger framework. Developments of Factor GARCH models are discussed by Bollerslev and Engle (1993) and Vrontos et al. (2003).

Tse and Tsui (2002) pointed out that the parameters in generalised covariance models are difficult to interpret, and the total number of parameters is still large. The second category tries to improve this drawback. To model covariance structure, we can model variance and correlation separately, leading to a more intuitive interpretation, as the conditional covariance matrix can be decomposed into a conditional diagonal variance matrix and a conditional correlation matrix. Such a framework was initially proposed by Bollerslev et al. (1990), who introduced the constant conditional correlation (CCC) model, by assuming that the dynamics of a covariance matrix are entirely contributed by conditional variance while the correlation term remains constant over time. This model's merit is its simple construction, and it provides an innovative view for later developments, although the assumption of constant conditional correlations is apparently unrealistic in the real world. As an extension, the assumptions in the Orthogonal GARCH (O-GARCH) (Alexander, 2001) are slightly realistic, allowing the CCC structure to only exist in a nonsingular linear combination of the variables rather than raw data naturally. The main idea of the O-GARCH is that it applies principal component analysis extracting orthogonal factors from the covariance matrix, and simplifies modelling to factor terms. For more details, we can refer to Van der Weide (2002), who proposed Generalized Orthogonal GARCH (GO-GARCH model).

Later, in order to assume more realistic assumptions on the dynamics of the correlation structure, Tse and Tsui (2002) and Engle (2002) proposed the dynamic conditional correlation (DCC) (also see Kroner and Ng (1998)). The model relaxes the assumption of the CCC model and allows a time-varying conditional correlation structure. Comparing these two versions, Tse and Tsui (2002) estimated conditional correlations as a weighted sum of lagged correlations and innovations, while Engle (2002) introduced a standardisation mechanism and let the correlations follow a quasi conditional correlation process. Benefitting from its flexibility and intuition, Engle's DCC model has become more popular.

Nonetheless, the DCC model is valid under some strict assumptions, giving rise to more developments of further DCC-type models in later years. Here we list four developments: 1) To relax the requirement that conditional correlations follow same dynamics which is too strict in empirical studies, Billio and Pelizzon (2003, 2006) proposed to the Flexible DCC model, and even rich dynamics were captured by the Rotated DCC model (Noureldin et al. 2014); 2) To avoid that estimators of the DCC model are inconsistent in large dimensional data, Aielli (2006, 2013) proposed the Consistent DCC model; 3) As the DCC model does not account for the leverage effect, Cappiello et al. (2006) generalise to the Asymmetric DCC model; 4) Lastly, Pesaran and Pesaran (2007) proposed the Student-T DCC model.

With regard to the estimation method, the first generation ARCH and GARCH models were estimated using Maximum Likelihood Estimation (MLE) under the assumption of a Gaussian distributed error terms. Later, to accommodate fat-tail distributed data, the estimation method is generalised to MLE with a student-T distributed residuals (Engle and Bollerslev (1986), Bollerslev (1987) and Nelson and Foster (1995)). This in turn has led to a consistency issue of estimation and developed into the Quasi-Maximum Likelihood Estimation (QMLE). The QMLE provides consistent estimators under mild assumptions. Berkes et al. (2004) suggested a consistent non-Gaussian QMLE with a relatively strong condition on residual. More recently, Fan et al. (2014) proposed a two-step non-Gaussian QMLE and QMLE with heavy-tailed likelihoods which are proved to be more efficient with simulated and empirical data.

Aiming to compare several classical M-GARCH models, we have conducted a Monte Carlo simulation study in this paper. Our findings indicates that DCC-type models outperform to others with their in-sample performances, while as expected, a structural break distorts the model fitting in-sample. In terms of predictive ability, models with heavy parameterisation show inferior forecasting performances. Benefitting from parsimoniously specified conditional correlation, the DCC model shows better ability in predicting covariance (though the EWMA is the most efficient one).

In the next section, we review the model specifications of some popular multivariate GARCH models. Section 1.3 describes a Monte Carlo Simulation study to assess models discussed in Section 1.2. Section 1.4 discusses simulation results and ends with a conclusion.

## 1.2 M-GARCH Model Specifications

Let us assume a $d \times 1$ stochastic vector $\mathbf{y}_t = (y_t(1), y_t(2), \ldots, y_t(d))^\top$, for $E(\mathbf{y})_t = \mathbf{0}$, and stationary in its second moment, then $\mathbf{y}_t$ can be expressed commonly as,

$$\mathbf{y}_t = \mathbf{\Sigma}_t^{\frac{1}{2}} \mathbf{e}_t \tag{1.2.1}$$

where $\mathbf{\Sigma}_t$ is the conditional covariance matrix, and $\mathbf{e}_t$ presents an innovation vector distributed according to Gaussian, i.e. $e_t \sim N(0, I)$, where $I$ is an identity matrix. Denoting the information set up to time $t-1$ as $\mathcal{F}_{t-1}$, the conditional variance and covariance of $\mathbf{y}_t$ equals to,

$$Var(\mathbf{y}_t|\mathcal{F}_{t-1}) = Var_{t-1}(\mathbf{e}_t) = \mathbf{\Sigma}_t^{\frac{1}{2}} Var_{t-1}(\mathbf{e}_t)(\mathbf{\Sigma}_t^{\frac{1}{2}})^\top = \mathbf{\Sigma}_t \tag{1.2.2}$$

M-GARCH models are designed to precisely trace, estimate, and forecast volatilities and cross-volatilities $\mathbf{\Sigma}_t$. In this section, we review nine representative models for modelling the variances $\mathbf{\Sigma}_t$, including generalizing multivariate GARCH models —exponential weighted moving average, VEC, Diagonal VEC, BEKK and Diagonal BEKK model; linear and non-linear M-GARCH models —constant conditional correlation, orthogonal-GARCH, dynamic conditional correlation, asymmetric dynamic conditional correlation and rotated dynamic conditional correlation model.

### 1.2.1 Generalized M-GARCH Models

Following the specification of univariate GARCH models, where the conditional variance is expressed as a function of lagged squared innovations and previous conditional variance, M-GARCH models generalize univiariate GARCH to multivariate cases through vectorization.

#### 1.2.1.1 Exponential Weighted Moving Average

The idea of the exponential weighted moving average is always simple to implement. Rather than an equally-weighted moving average, the EWMA model puts more weights on recent historical data, which is more sensible as the effect of innovations would decay exponentially over time horizon (J.P. Morgan1996). The following presents the EWMA($d_0$, $\lambda_0$) recursive equation, denoting $d_0$ for window length and $\lambda_0$ for exponential decay factor.

$$\mathbf{\Sigma}_t = \lambda_0 \mathbf{\Sigma}_{t-1} + \frac{1-\lambda_0}{1-\lambda_0^{d_0}} \mathbf{y}_{t-1}\mathbf{y}_{t-1}^\top - \frac{1-\lambda_0}{1-\lambda_0^{d_0}} \lambda_0^{d_0-1} \mathbf{y}_{t-d_0-1}\mathbf{y}_{t-d_0-1}^\top \tag{1.2.3}$$

The EWMA model can be simplified as,

$$\mathbf{\Sigma}_t = \lambda_0 \mathbf{\Sigma}_{t-1} + (1-\lambda_0)\mathbf{y}_{t-1}\mathbf{y}_{t-1}^\top \tag{1.2.4}$$

For the sake of avoiding the similarity in unconditional variance, Riskmetrics (1996) suggests the window length $d_0$ between 50 and 250 periods. In the term of exponential decay factor, substantial empirical results suggest a value for the parameter $\lambda_0 = 0.94$ for daily frequency, and $\lambda_0 = 0.97$ for monthly frequency. Because this model is simple and almost computation free, it is the most widely used M-GARCH model in the financial sector. But in theory, the EWMA model suffers from drawbacks: 1. it is hard to maintain the unit weights summation for infinite $d_0$; 2. it is not suitable to fully describe the mean reversion mechanism in second moment; 3. it only guarantees a positive semi-definite conditional covariance matrix, where conditional covariance matrix should be guaranteed as positive definite matrix.

#### 1.2.1.2 VEC (Bollerslev et al., 1988)

Using a stack operator, the VEC model directly vectorizes univariate GARCH model to the multivariate case,

$$vech(\mathbf{\Sigma}_t) = \mathbf{C} + \sum_{i=1}^{q} \mathbf{A}_i vech(\mathbf{y}_{t-i}\mathbf{y}'_{t-i}) + \sum_{j=1}^{p} \mathbf{B}_j vech(\mathbf{\Sigma}_{t-j}) \tag{1.2.5}$$

where $vech(.)$ is the lower triangular stacking operator so that a $d \times d$ matrix is stacked into a $\frac{d\times(d+1)}{2} \times 1$ vector. As the covariance matrix is symmetric in nature, considering the lower triangular elements is enough to analyse $\mathbf{\Sigma}_t$. The intercept vector $\mathbf{C}$ is formed as $\frac{d\times(d+1)}{2}\times 1$, the coefficient matrices $\mathbf{A}_i$ and $\mathbf{B}_j$ are $\frac{d\times(d+1)}{2}\times\frac{d\times(d+1)}{2}$ matrices. Argument $\sigma_{i^*j^*,t}$ located at $i$th row $j$th column of $\mathbf{\Sigma}_t$ represents the conditional covariance between process $i^*$ and $j^*$, and it is estimated by the univariate GARCH model,

$$\sigma_{i^*j^*,t} = \gamma_{i^*,j^*} + \sum_{l=1}^{q}\alpha_l \mathbf{y}_{t-l}(i^*)\mathbf{y}_{t-l}(j^*) + \sum_{k=1}^{p}\beta_k\sigma_{i^*j^*,t-l} \tag{1.2.6}$$

where $\gamma_{i^*,j^*}$, $\alpha_l$ and $\beta_k$ are corresponding parameters in coefficient matrices $\mathbf{C}$, $\mathbf{A}_i$ and $\mathbf{B}_j$. In total, there are $(p+q)(\frac{d(d+1)}{2})^2+\frac{d(d+1)}{2}$ parameters in the VEC model, for example, in the case of a binary VEC (1,1) model, the total number of parameters is $2\times(\frac{2\times(2+1)}{2})^2+\frac{2\times(2+1)}{2} = 21$. Hence, although the VEC model is flexible and easy to understand, it suffers from the limitation of heavy computational cost as the number of parameters increases tremendously along with the dimensionality of the model. Moreover, it is hard not to violate non-negativity constraints and thus obtaining a positive definite covariance matrix in VEC model. The VEC model is estimated by the QMLE, given the conditional log-likelihood function as,

$$L_t(\theta) = -\frac{d}{2}log2\pi - \frac{1}{2}log|\mathbf{\Sigma}_t(\theta)| - \frac{1}{2}\mathbf{y}_t(\theta')\mathbf{\Sigma}_t^{-1}(\theta)\mathbf{y}_t(\theta)$$

where $\theta$ represents relevant set of parameters in the model.

#### 1.2.1.3 Diagonal VEC (Ding and Engle, 2001)

One improvement of the VEC model is the diagonal VEC. This is obtained by imposing an assumption that parameter matrices $\mathbf{A}_i$ and $\mathbf{B}_i$ in Equation (1.2.5) are diagonal matrices. Thus, diagonal transformation is not going to affect the consistency of results, but it reduces the number of parameters substantially. Assuming that the coefficient matrices $\mathbf{A}_i$ and $\mathbf{B}_j$ are diagonal matrices, the number of parameters in the DVEC(1,1) model decreases to

$\frac{(p+q+1)N(N+1)}{2}$. More importantly, it guarantees a positive-definite covariance matrix under some assumptions. Ding and Engle (2001) specified the DVEC (1,1) model as

$$\mathbf{\Sigma}_t = \mathbf{C}\mathbf{C}^\top + \mathbf{A}\mathbf{A}^\top \circ \mathbf{y}_{t-1}\mathbf{y}_{t-1}^\top + \mathbf{B}\mathbf{B}^\top \circ \mathbf{\Sigma}_{t-1} \tag{1.2.7}$$

where $\circ$ is the Hadamard product. The element in $i$th row and $j$th column of $\mathbf{\Sigma}_t$ is estimated by a unvariate GARCH,

$$\sigma_t(i,j) = w_{i,j} + \alpha_{i,j}\mathbf{y}_{t-1}(i)\mathbf{y}_{t-1}(j) + \beta_{i,j}\sigma_{t-1}(i,j) \tag{1.2.8}$$

where $w_{i,j} = (\mathbf{C}\mathbf{C}^\top)_{i,j}$, $\alpha_{i,j} = (\mathbf{A}\mathbf{A}^\top)_{i,j}$, $\beta_{i,j} = (\mathbf{B}\mathbf{B}^\top)_{i,j}$. As long as the coefficient matrices $\mathbf{A}$ and $\mathbf{B}$ are positive definite, the positive definiteness of the covariance matrix can be easily guaranteed by basic lemmas of Hadamard product. Still, DVEC model endures onerous parameters when it comes to a large system. The estimation method is same with the VEC model.

#### 1.2.1.4 BEKK (Engle and Kroner, 1995)

Since a positive definite matrix always can be Cholesky decomposed, Engle and Kroner (Baba, Engle, Kraft and Kroner) proposed the classical BEKK model which is formed as,

$$\mathbf{\Sigma}_t = \mathbf{C}\mathbf{C}^\top + \sum_{j=1}^{q} \mathbf{A}_k^\top \mathbf{y}_{t-k}\mathbf{y}_{t-k}^\top \mathbf{A}_k + \sum_{j=1}^{p} \mathbf{B}_j^\top \mathbf{\Sigma}_{t-j}\mathbf{B}_j \tag{1.2.9}$$

where $\mathbf{C}$ is a $d \times d$ triangle positive definite matrix, $\mathbf{A}_k$ and $\mathbf{B}_j$ are $d \times d$ triangle semi-positive definite matrices. Thus, $\mathbf{\Sigma}_t$ is strict positive definite. In addition, the eigenvalues of matrix $\mathbf{A} \bigotimes \mathbf{A} + \mathbf{B} \bigotimes \mathbf{B}$, where $\bigotimes$ is the Kronecker Product, should be inside of unit circle to maintain covariance stationary for BEKK (Silvennoinen and Teräsvirta, 2009). The BEKK model benefits by its compact parametrization and its delivering consistently positive definite variances.

Also, the coefficient matrices **A** and **B** can be parametrized to other forms. For example, the diagonal BEKK provides relative parsimonious parametrization. Coefficients matrices are redesigned as diagonal matrices, i.e. **A** and **B** become to $\mathbf{A} = diag(a_{11}, a_{22}, \ldots, a_{kk})$ and $\mathbf{B} = diag(b_{11}, b_{22}, \ldots, b_{kk})$. Hence, adopting Hadamard product, the diagonal BEKK (1,1) expresses,

$$\mathbf{\Sigma}_t = \mathbf{C}\mathbf{C}^\top + \mathbf{A}\mathbf{A}^\top \circ \mathbf{y}_{t-1}\mathbf{y}_{t-1}^\top + \mathbf{B}\mathbf{B}^\top \circ \mathbf{\Sigma}_{t-1} \tag{1.2.10}$$

The number of parameters equals to $\frac{(p+q)Kd+d(d+1)}{2}$. It is noticeable that the DBEKK model is a DVEC model with special restrictions. Essentially, any BEKK model can imply a unique VEC model, yet this does not correct conversely (Silvennoinen and Teräsvirta, 2009).

### 1.2.2 Linear and Non-Linear Combined Models

Enlightened by the well-known correlation equation,

$$\rho_{xy} = \frac{cov(xy)}{\sqrt{var(x)}\sqrt{var(y)}} \tag{1.2.11}$$

the conditional covariance matrix can be decomposed into conditional covariance and variance,

$$\mathbf{\Sigma}_t = \mathbf{D}_t\mathbf{R}_t\mathbf{D}_t \tag{1.2.12}$$

where $\mathbf{D}_t = diag(\sigma_{11,t}^{\frac{1}{2}}, \sigma_{22,t}^{\frac{1}{2}}, \ldots, \sigma_{dd,t}^{\frac{1}{2}})$ is the a $d \times d$ diagonal matrix composed by conditional standard deviation of $\mathbf{y}_t$ estimated by univariate GARCH models. $\mathbf{R}_t$ is a $d \times d$ correlation matrix with unit diagonal elements, and off-diagonal elements ranged between $-1$ and $+1$. The specification and estimation of $\mathbf{R}_t$ are discussed in diverse forms.

#### 1.2.2.1 Constant Conditional Correlation (Bollerslev, 1990)

As a first theoretical construction, the constant conditional correlation model assumes that conditional correlation is time-invariant,

$$\boldsymbol{\Sigma}_t = \mathbf{D}_t \mathbf{R} \mathbf{D}_t \tag{1.2.13}$$

Bollerslev (1990) used the GARCH (1,1) model to estimate $\mathbf{D}_t$. Thus, the $m$th diagonal element in $\mathbf{D}_t$ is estimated through,

$$\sigma_t(m) = w_m + \sum_{i=1}^{q} \alpha_{m,i} y_{t-i}^2(m) + \sum_{j=1}^{p} \beta_{m,j} \sigma_{t-j}^2(m) \tag{1.2.14}$$

The conditional covariance matrix will be positive definite as long as both $\mathbf{D}_t$ and $\mathbf{R}$ are positive definite. This model is much lighter in computation, since it is not necessary to estimate parameters for cross coordinates, and the number of parameter only comes from conditional variance estimation. To estimate the CCC model, one first estimates $\mathbf{D}_t$ by QMLE, that is maximising the following,

$$L(\theta) = -\frac{Td}{2}(1 + log(2\pi) - log(T)) - \sum_{t=1}^{T} log|\mathbf{D}_t| - \frac{T}{2} log|\sum_{t=1}^{T} \mathbf{y}_t^* (\mathbf{y}_t^*)^\top|$$

where $\mathbf{y}_t^* = \mathbf{y}_t \cdot \mathbf{D}_t^{-1}$. Then, the constant conditional correlation term $\mathbf{R}$ is estimated by computing the unconditional correlation matrix. The CCC model opens up a shortcut which bypasses using a large number of parameters to estimate the conditional covariance directly. Nevertheless, although it works efficiently in empirical large data set, in the real world, the correlation relationship between two series is hardly found to be stable either in short run or long run, thereby requiring further extensions under more realistic conditions.

#### 1.2.2.2 Orthogonal GARCH (Alexander, 2001)

One of extension models is the orthogonal GARCH model. It admits that the second moment of real financial data is structured as the CCC, but with some transformations. One popular transformation indicates that rather than the $d$-dimensional multivariate vector itself, its principal components follow a CCC structure (Alexander, 2001). The OGARCH model is more flexible than CCC model, and it also reduces further the dimensionality by using less factors.

The transformation on $\mathbf{y}_t$ gives,

$$\mathbf{\Lambda}_t = Var_{t-1}(P \times \mathbf{y}_t) = \mathbf{D}_t \mathbf{R} \mathbf{D}_t \tag{1.2.15}$$

where $\mathbf{\Lambda}_t$ is the conditional covariance matrix of $P \times \mathbf{y}_t$, which follows the CCC structure. Then, the conditional covariance matrix of $\mathbf{y}_t$ can be derived by,

$$\mathbf{\Sigma}_t = P^{-1} \mathbf{D}_t \mathbf{R} \mathbf{D}_t (P^{-1})^\top \tag{1.2.16}$$

Furthermore, the conditional correlation of $\mathbf{y}_t$ shows,

$$Corr(\mathbf{y}_t) = diag(P^{-1}\mathbf{D}_t\mathbf{R}\mathbf{D}_t(P^{-1})^\top)^{-\frac{1}{2}} \cdot P^{-1}\mathbf{D}_t\mathbf{R}\mathbf{D}_t(P^{-1})^\top \cdot diag(P^{-1}\mathbf{D}_t\mathbf{R}\mathbf{D}_t(P^{-1})^\top)^{-\frac{1}{2}} \tag{1.2.17}$$

This specification allows time-varying conditional correlation. But it is still necessary to discuss the definition of matrix $P$. Alexander (2001) specified the inverse matrix of $P$ is a matrix of eigenvalues extracted from the unconditional covariance matrix of $\mathbf{y}_t$. Hence, $P \cdot \mathbf{y}_t$ is principal components for $\mathbf{y}_t$.

#### 1.2.2.3 Dynamic Conditional Correlation (Engle, 2002)

Even though the O-GARCH model considers a time-varying correlation structure through the principal component transformation of the CCC model, it does not solve time-invariant correlation problem of the CCC model intrinsically. Engle (2002) proposed a dynamic conditional correlation model by allowing dynamic evolution of the conditional correlations (also cf. Tse and Tsui, 2002). The conditional covariance given by:

$$\mathbf{\Sigma}_t = \mathbf{D}_t \cdot \mathbf{R}_t \cdot \mathbf{D}_t \tag{1.2.18}$$

where conditional correlation $\mathbf{R}_t$ is a stochastic process. In order to specify and estimate $\mathbf{\Sigma}_t$, Engle (2002) developed a three steps approach: 1. DE-GARCHING; 2. Estimating the Quasi-Correlations; 3. Rescaling the correlations. In the first step, we use univariate GARCH models, e.g. GARCH (1,1), to estimate the conditional variance of $y_t(m)$,

$$\sigma_t(m) = \omega_m + \alpha_m y_{t-1}(m)^2 + \beta_i \sigma_{t-1}(m) \tag{1.2.19}$$

where $\sigma_t^{\frac{1}{2}}(m)$ is the $m$th diagonal element of $\mathbf{D}_t$. Also, in order to analyse the correlation term $\mathbf{R}_t$, the DE-GARCHING process eliminates the effect of conditional variance through,

$$y_t^*(i) = y_t(i)/\hat{\sigma}_t^{\frac{1}{2}}(i) \tag{1.2.20}$$

$\mathbf{y}_t^* = (y_t^*(1), y_t^*(2), \ldots, y_t^*(d))^\top$ denotes the devolatized vector.

The second step designs a quasi conditional correlation process $\mathbf{Q}_t$, where $\mathbf{Q}_t$ can be specified differently, such as the exponential smoothing form,

$$\mathbf{Q}_t = (1-\lambda)\mathbf{y}_{t-1}^*(\mathbf{y}_{t-1}^*)^\top + \lambda \mathbf{Q}_{t-1} \tag{1.2.21}$$

where $\lambda$ is defined in Section 1.2.1.1. The scalar DCC model is specified as,

$$\mathbf{Q}_t = \mathbf{\Omega} + \sum_{i=1}^{N} \theta_{1,i}(\mathbf{y}^*_{t-i}(\mathbf{y}^*_{t-i})^\top) + \sum_{j=1}^{M} \theta_{2,j}(\mathbf{Q}_{t-j}) \tag{1.2.22}$$

where $\mathbf{\Omega}$ is positive definite, coefficients $\theta_{1,i}$ and $\theta_{2,j} \geq 0$. Where the initial value of $\mathbf{Q}_t$ is the unconditional correlation $\bar{R}$, $\bar{R} = \frac{1}{T}\sum_{t=1}^{T} \mathbf{y}^*_t(\mathbf{y}^*_t)^\top$. Therefore, $\mathbf{Q}_t$ is a positive definite matrix. $\mathbf{Q}_t$ can also be specified in a matrix representation,

$$\mathbf{Q}_t = \mathbf{\Omega} + \sum_{i=1}^{N} \mathbf{A}_i \circ (\mathbf{y}^*_{t-i}(\mathbf{y}^*_{t-i})^\top) + \sum_{j=1}^{M} \mathbf{B}_j \circ (\mathbf{Q}_{t-j}) \tag{1.2.23}$$

where $\mathbf{\Omega}$ is a positive definite matrix, and $\mathbf{A}_i$, $\mathbf{B}_j$ are non-negative definite matrices. However, due to the fact that it is hard to keep the diagonal elements of $\mathbf{Q}_t$ as equal to 1 although the average value of them is one, a rescaling is necessary.

$$\mathbf{R}_t = diag(\mathbf{Q}_t)^{-\frac{1}{2}} \cdot \mathbf{Q}_t \cdot diag(\mathbf{Q}_t)^{-\frac{1}{2}} \tag{1.2.24}$$

Engle (2002) suggested using the QMLE to estimate models in step 1 and 2, maximising the following:

$$L(\theta) = -\frac{1}{2}\sum_{t=1}^{T}(d \cdot log(2\pi) + 2log(|\mathbf{D}_t|) + log(|\mathbf{R}_t|) + (\mathbf{y}^*)_t^\top \mathbf{R}_t^{-1} \mathbf{y}^*_t)$$

#### 1.2.2.4 Asymmetric DCC (Cappiello et al., 2006)

Since there is significant evidence of leverage effects (negative shocks earn more weights) in financial data, the DCC model was extended to Asymmetric DCC (ADCC) model by Cappiello et al. (2006), by allowing an asymmetric indicator function into the dynamics of the conditional variance and quasi conditional correlations. In the first step, the original GARCH (1,1) is replaced by any univariate asymmetric models, commonly, the GJR-GARCH model

(Glosten et al., 1993), and the conditional variance of the $m$th process is estimated by,

$$\sigma_t(m) = w_m + \sum_{i=1}^{q} \alpha_{m,i} y_{t-i}^2(m) + \sum_{k=1}^{q} \gamma_{m,k} y_{t-k}^2(m) I_{y_{t-k}<0} + \sum_{j=1}^{p} \beta_{m,j} \sigma_{t-j}^2(m) \tag{1.2.25}$$

where $I_{y_{t-i}} < 0$ with $i \geq 1$ is an identity term which equals to 1 if $y_{t-i}$ was negative, otherwise equals to 0.

$$y_t^*(m) = y_t(m)/\hat{\sigma}_t(m)^{\frac{1}{2}}$$

Then, considering an asymmetric indicator into quasi conditional correlation process,

$$\mathbf{Q}_t = \mathbf{\Omega} + \sum_{i=1}^{N} \theta_{1,i} \mathbf{y}_{t-i}^* (\mathbf{y}_{t-i}^*)^\top + \sum_{k=1}^{O} \theta_{2,k} \boldsymbol{\eta}_{t-k} \boldsymbol{\eta}_{t-k}^\top + \sum_{j=1}^{M} \theta_{3,j} \mathbf{Q}_{t-j} \tag{1.2.26}$$

where $\boldsymbol{\eta}_t = min[\mathbf{y}_t^*, 0]$, and $\mathbf{\Omega}$ is a positive definite matrix. The third rescaling step is the same as the DCC model. The positive definite covariance is easily guaranteed as long as the coefficients $\theta_i$, $i = 1, 2, 3$ are non-negative. The estimation method is the QMLE as described in the DCC model.

#### 1.2.2.5 Rotated DCC (Noureldin et al., 2014)

To improve the flexibility of the DCC model, Noureldin et al (2014) proposed a rotated specification of the DCC model. The idea of the RDCC model is to model the dynamics of the conditional variance and correlations non-simultaneously, thus driving more realistic dynamics. They define the unconditional covariance of $\mathbf{y}_t$ as $\Lambda$, and $\Lambda$ can be decomposed as,

$$Var(\mathbf{y}_t) = \Lambda = P_c \Lambda_c P_c^\top \tag{1.2.27}$$

where $P_c$ is the eigenvectors and $\Lambda_c$ is a diagonal matrix, the elements of which are the eigenvalues of the covariance matrix. The RDCC model also follows the similar three steps estimation. In the first step, DE-GARCHING is completed by adopting rotated $\tilde{\mathbf{y}}_t$,

$$\tilde{\mathbf{y}}_t = \Lambda^{-\frac{1}{2}} \cdot \mathbf{y}_t = P_c \Lambda_c^{-\frac{1}{2}} P_c^\top \mathbf{y}_t \tag{1.2.28}$$

In the second step, keeping the mean-reverting correlation targeting as the unconditional correlation, the quasi conditional correlation is modelled in matrix representation,

$$\tilde{\mathbf{Q}}_t = \mathbf{\Omega} + \sum_{i=1}^{N} \mathbf{A}_i \tilde{\mathbf{y}}_{t-i} \tilde{\mathbf{y}}_{t-i}^\top \mathbf{A}_i^\top + \sum_{j=1}^{M} \mathbf{B}_j \tilde{\mathbf{Q}}_{t-j} \mathbf{B}_j^\top \tag{1.2.29}$$

Where $\mathbf{\Omega}$ is a positive definite matrix. The initial value of $\tilde{\mathbf{Q}}_t$ is the unconditional covariance matrix of $\tilde{\mathbf{y}}_t$. Coefficient matrices $\mathbf{A}_i$ and $\mathbf{B}_i$ are semi-positive definite diagonal matrices. A scalar RDCC model also exists.

$$\tilde{\mathbf{Q}}_t = \mathbf{\Omega} + \sum_{i=1}^{N} \theta_{1,i} \tilde{\mathbf{y}}_{t-i} \tilde{\mathbf{y}}_{t-i}^\top + \sum_{j=1}^{M} \theta_{2,j} \tilde{\mathbf{Q}}_{t-j} \tag{1.2.30}$$

The assumptions on $\mathbf{\Omega}$, $\theta_{1,i}$ and $\theta_{2,j}$ are same with defined in (1.2.22).

Similarly, the asymmetric RDCC model gives,

$$\tilde{\mathbf{Q}}_t = \mathbf{\Omega} + \sum_{i=1}^{N} \theta_{1,i} \tilde{\mathbf{y}}_{t-i} \tilde{\mathbf{y}}_{t-i}^\top + \sum_{k=1}^{O} \theta_{2,k} \tilde{\boldsymbol{\eta}}_{t-k} \tilde{\boldsymbol{\eta}}_{t-k}^\top + \sum_{j=1}^{M} \theta_{3,j} \tilde{\mathbf{Q}}_{t-j} \tag{1.2.31}$$

where $\tilde{\boldsymbol{\eta}}_t = min[\tilde{\mathbf{y}}_t, 0]$, $\mathbf{\Omega}$ is a positive definite matrix, scalar coefficients $\theta_{1,i}$, $\theta_{2,k}$ and $\theta_{3,j}$ are non-negative parameters. Eventually, in the third step, the conditional correlation $\mathbf{R}_t$ is rescaled by,

$$\mathbf{R}_t = diag(\tilde{\mathbf{Q}}_t)^{-\frac{1}{2}} \tilde{\mathbf{Q}}_t diag(\tilde{\mathbf{Q}}_t)^{-\frac{1}{2}} \tag{1.2.32}$$

### 1.2.3 Conditional Correlation Model with Structural Breaks

Since Chow (1960) discussed the issue of structural breaks in linear regression models, the effect of structural breaks on modeling economic or financial data always needs to be addressed. Engle (2009) allowed that a structural break may occur in quasi conditional correlation process. Assuming that a known break occurs at $1 < t^* < T$, and defining $\mathfrak{t} = \frac{t^*}{T}$ with $0 < \mathfrak{t} < 1$, we then can estimate quasi conditional correlation process with dummies at the break point $\mathfrak{t}$. The DCCSB (1,1) model has a quasi conditional correlation process,

$$\mathbf{Q}_t = \mathbf{\Omega}_1 + \mathbf{\Omega_2} \cdot DM_t + (\theta_{1,1} + \theta_{1,2} \cdot DM_t)\mathbf{y}^*_{t-1}(\mathbf{y}^*_{t-1})^\top + (\theta_{2,1} + \theta_{2,2} \cdot DM_t)\mathbf{Q}_{t-1} \quad (1.2.33)$$

where parameters $DM_t$ is a dummy variable which equals to 0 before $t^*$, and $DM_t = 1$ after $t^*$. Thus, $\mathbf{\Omega}_2$, $\theta_{1,2}$ and $\theta_{2,2}$ are additional intercept and coefficients after break happened, either be positive or negative values.

Likewise, in order to test what is the effect in global equity markets after introducing the Euro, Cappiello et al. (2006) introduced a ADCC with structural break model. The quasi conditional correlation process in ADCCSB (1,1,1) model is given by:

$$\begin{aligned}\mathbf{Q}_t = \mathbf{\Omega}_1 + \mathbf{\Omega_2} \cdot DM_t + (\theta_{1,1} + \theta_{1,2} \cdot DM_t)\mathbf{y}^*_{t-1}(\mathbf{y}^*_{t-1})^\top + (\theta_{2,1} + \theta_{2,2} \cdot DM_t)\boldsymbol{\eta}_{t-1}\boldsymbol{\eta}_{t-1}^\top \\ +(\theta_{3,1} + \theta_{3,2} \cdot DM_t)\mathbf{Q}_{t-1}\end{aligned} \quad (1.2.34)$$

The structural break dummy also can be introduced into the ARDCC model, such that the quasi conditional correlation follows,

$$\begin{aligned}\tilde{\mathbf{Q}}_t = \mathbf{\Omega}_1 + \mathbf{\Omega}_2 \cdot DM_t + (\theta_{1,1} + \theta_{1,2} \cdot DM_t)\tilde{\mathbf{y}}_{t-1}\tilde{\mathbf{y}}_{t-1}^\top + (\theta_{2,1} + \theta_{2,2} \cdot DM_t)\tilde{\boldsymbol{\eta}}_{t-1}\tilde{\boldsymbol{\eta}}_{t-1}^\top \\ +(\theta_{3,1} + \theta_{3,2} \cdot DM_t)\tilde{\mathbf{Q}}_{t-1}\end{aligned} \quad (1.2.35)$$

All models are estimated by the QMLE method.

## 1.3 The Model Selection Criteria

In order to compare the performance of M-GARCH models, this section introduces several model selection methods to evaluate the in-sample and out-of-sample performance of rival models. We firstly consider the efficiency of the model, as a model can be of heavy computational burden if it was excessively parametrised. Table 1.3.1 lists the number of parameters estimated in each M-GARCH model. It is clear that for a bivariate model, the full matrix representation of VEC and BEKK models are heavily parametrised, and the number becomes lower by using the corresponding diagonal matrix representation. In linear/nonlinear correlation models, although the number of parameters is still large, most of parameters are estimated for conditional variance in univariate case. Beside, a structural break dummy can double the number of parameters in one model.

To estimate these models, we use the MFE toolbox (Sheppard, 2012) in Matlab. In each model, we apply the BFGS method to obtain the quasi-maximum likelihood estimates.

Since all models are estimated by the same method, we can compare their in-sample performances through evaluating their log-likelihood (LL) values. The best model is given by the model with the highest LL value. If two models $\mathcal{M}_1$ and $\mathcal{M}_2$ gave similar LL values, it is better to apply the log-likelihood ratio test to test the hypothesis,

$$\begin{cases} H_0: \quad \mathcal{M}_1 > \mathcal{M}_2 \\ H_1: \quad reject \quad H_0 \end{cases} \tag{1.3.1}$$

using the chi-square statistics,

$$\chi^2_{df} = -2 \cdot ln(\frac{L(\mathcal{M}_1)}{L(\mathcal{M}_2)}) \tag{1.3.2}$$

where $L(\mathcal{M}_i)$ is the LL value of model $i$, $df$ is the degree of freedom defining as the difference of the number of parameter from two models. The critical values can be found in the $\chi^2$ statistics table.

Table 1.3.1: The Number of Parameters in M-GARCH Models

| Model | Number of Parameters (Bivariate model) |
|---|---|
| EWMA(0.96) | 0 |
| VEC(1,1) | 21 |
| DVEC(1,1) | 9 |
| BEKK(1,1) | 11 |
| DBEKK(1,1) | 7 |
| CCC(1,1) | 7 |
| OGARCH(1,1) | 6 |
| DCC(1,1) | 11 |
| ADCC(1,1,1) | 14 |
| RDCC(1,1) | 9 |
| ARDCC(1,1,1) | 11 |
| ADCCSB(1,1,1) | 28 |
| ARDCCSB(1,1,1) | 22 |

Nonetheless, higher LL statistics are always obtained in over-fitted models. Hence, the information criteria is an alternative method to compensate the LL statistics by penalising overfitting. Popular information criteria are the Akaike information criterion (Akaike, 1987) and the Bayesian information criterion (Schwarz, 1978). There are two version of the AIC: AIC and AICc (Hurvich and Tsai, 1995),

$$AIC = -2LL + 2K \tag{1.3.3}$$

$$AICc = -2LL + 2K + \frac{2K(K+1)}{T-K-1} \tag{1.3.4}$$

where $K$ and $T$ are the number of parameters and the sample size, respectively. The AICc is more appropriate in small samples because the last term in (1.3.4) only exists for a smaller $K/T$. Compared with AIC, the BIC puts more penalty on the number of estimated parameters (Burnham and Anderson, 2004), and it is specified,

$$BIC = -2LL + K \cdot log(T) \tag{1.3.5}$$

As a loss function, opposite to the selection law by LL statistics, the best model is given by the one with the smallest AIC or BIC information criteria.

Regarding the out-of-sample performance, we apply mean absolute forecasting error and mean square forecasting error. Since the present chapter studies the prediction ability of the correlation structure in M-GARCH models, the prediction error only depends on predicted $\hat{\rho}$ and actual correlation $\rho$.

$$MAFE = \frac{1}{\hat{T}} \sum_{i=1}^{\hat{T}} |\hat{\rho}_i - \rho_i| \tag{1.3.6}$$

$$MSFE = \frac{1}{\hat{T}} \sum_{i=1}^{\hat{T}} (\hat{\rho}_i - \rho_i)^2 \tag{1.3.7}$$

where $\hat{T}$ is the number of observations from the out-of-sample. The MAFE is a linear score and the MSFE is a quadratic score. One model has better prediction ability if it delivered smaller MAFE and MSFE values.

## 1.4 A Monte Carlo Simulation Study

In order to assess the performance of M-GARCH models in in-sample and out-of-sample, we design a Monte Carlo simulation study considering different types of generated data sets. We aim to generate data to approximate real financial data. Changing the value of parameters in the data generating process, we have conducted five experiments. For simplicity, we use bivariate models such that $\mathbf{y}_t = (y_{1,t}, y_{2,t})^\top$.

$$\mathbf{y}_t = \mathbf{\Sigma}_t^{\frac{1}{2}} \mathbf{e}_t \tag{1.4.1}$$

where $\mathbf{\Sigma}_t$ is a generated covariance sequence which governs the covariance structure of $\mathbf{y}_t$, and $\mathbf{\Sigma}_t^{\frac{1}{2}}$ is in the Cholesky form. The innovation term $\mathbf{e}_t = (e_{1,t}, e_{2,t})^\top$. Considering the fact of fat-tailed distributed financial data, we generate $\mathbf{e}_t$ by a student-T distributed random vector with degree of freedom 4.

To form $\mathbf{\Sigma}_t$, we define $\sigma_{ij,t}$ as the coordinate in $\mathbf{\Sigma}_t$, tracking the conditional covariance between processes $e_{i,t}$ and $e_{j,t}$.

Thus, $\sigma_{ij,t}$ is the conditional variance with $i = j$, and $\sigma_{ii,t}$ is generated through the GARCH (1,1) model,

$$\sigma_{11,t} = \omega_1 + \alpha_1 e_{1,t-1}^2 + \beta_1 \sigma_{11,t-1} \tag{1.4.2}$$

$$\sigma_{22,t} = \omega_2 + \alpha_2 e_{2,t-1}^2 + \beta_2 \sigma_{22,t-1} \tag{1.4.3}$$

Or, to generate data with financial leverage effect, we also generate conditional variance through the GJR (1,1,1) model (Glosten et al., 1993),

$$\sigma_{11,t} = \omega_1 + \alpha_1 e_{1,t-1}^2 + \gamma_1 I_{t-1} e_{1,t-1}^2 + \beta_1 \sigma_{11,t-1} \tag{1.4.4}$$

$$\sigma_{22,t} = \omega_2 + \alpha_2 e_{2,t-1}^2 + \gamma_2 I_{t-1} e_{1,t-1}^2 + \beta_2 \sigma_{22,t-1} \tag{1.4.5}$$

where $I$ is an asymmetric indicator that $I_{t-1} = 1$ if $\mathbf{e}_{t-1} < 0$, otherwise $I_{t-1} = 0$. The initial value sets as 1.

Meanwhile, the conditional covariance element $\sigma_{ij,t}$ with $i \neq j$ can be obtained by formulating,

$$\sigma_{ij,t} = \sigma_{ii,t}^{\frac{1}{2}} \rho_{ij,t} \sigma_{jj,t}^{\frac{1}{2}} \tag{1.4.6}$$

where $\sigma_{ii,t}$ is generated from (1.4.2), (1.4.3), (1.4.4) and (1.4.5). We then discuss the specification of conditional correlation $\rho_{ij,t}$.

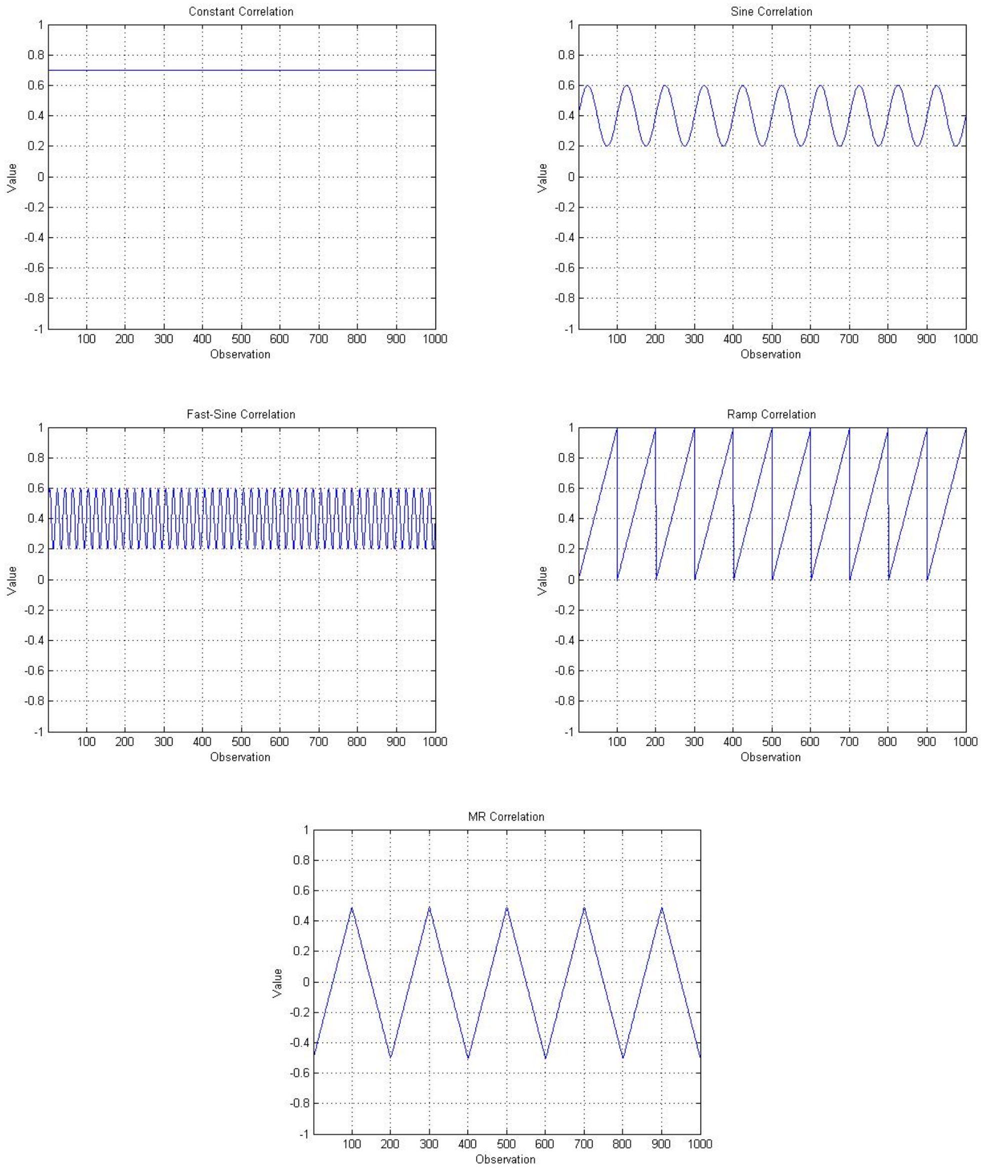
#### 1.4.0.1 Conditional Correlation without a Structural Break

We first consider $\rho_{ij,t}$ is evolving without a structural break. The dynamics of $\rho_{ij,t}$ are formed in scenarios including constant, 'sine', fast 'sine', ramp and a mean reversion process over $-1$ to $+1$.

- 1. Constant: $\rho_{ij,t} = 0.7$
- 2. Sine: $\rho_{ij,t} = 0.4 + 0.2sin(\frac{2\pi t}{100})$
- 3. Fast Sine: $\rho_{ij,t} = 0.4 + 0.2sin(\frac{2\pi t}{20})$
- 4. Ramp: $\rho_{ij,t} = mod(\frac{t}{100})$
- 5. Mean reversion to zero: $\rho_{ij,t} = I \cdot \frac{t}{100}$, where $I = \pm 1$

Visually displays in Figure 1.4.1, with $T = 1000$.

Figure 1.4.1: Conditional Correlations without Structural Breaks




#### 1.4.0.2 Conditional Correlation with a Structural break

Since real financial data are sensitive to market news and prone to structural breaks, we then consider correlation structure with a mean break occurred at the centre of sample, $\mathsf{t} = 0.5$. The conditional correlation $\rho_{ij,t}$ are then generated as,

- 1. Constant:

$$\rho_t = \begin{pmatrix} 0.3 & t \leq \frac{T}{2} \\ 0.7 & t > \frac{T}{2} \end{pmatrix} \tag{1.4.7}$$

- 2. Sine:

$$\rho_t = \alpha + 0.2sin(\frac{2\pi t}{100}) \tag{1.4.8}$$

where,

$$\alpha = \begin{pmatrix} 0.4, t < \frac{T}{2} \\ 0.7, t > \frac{T}{2} \end{pmatrix} \tag{1.4.9}$$

- 3. Fast Sine:

$$\rho_t = \alpha + 0.2sin(\frac{2\pi t}{20}) \tag{1.4.10}$$

where,

$$\alpha = \begin{pmatrix} 0.4, t < \frac{T}{2} \\ 0.7, t > \frac{T}{2} \end{pmatrix} \tag{1.4.11}$$

- 4. Ramp:

$$\rho_t = mod(\frac{t}{50}) \tag{1.4.12}$$

where $t = t_1$ or $t_2$, $1 < t_1 < \frac{T}{2}$, $\frac{T}{2} < t_2 < T$.

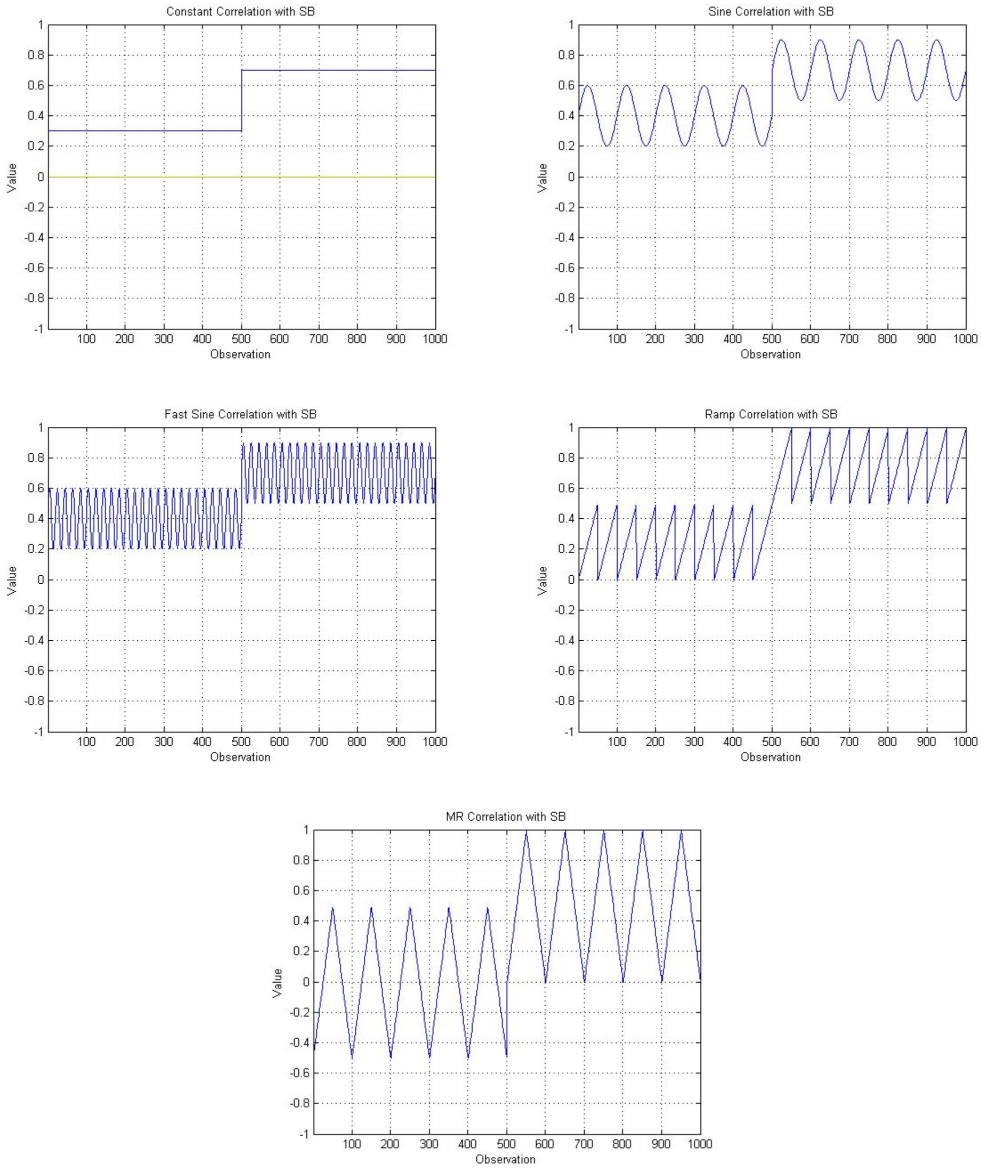
- 5. Mean reverting to zero:

$$\rho_t = I \cdot \frac{t}{50} \tag{1.4.13}$$

where $t = t_1$ or $t_2$, $1 < t_1 < \frac{T}{2}$, $\frac{T}{2} < t_2 < T$; and $I = \pm 1$

Type 1-5 are visually displayed in Figure 1.4.2, with $T = 1000$

Figure 1.4.2: Conditional Correlations with a Structural Break




The simulation is replicated 1000 times, and in each replication we generate 1000 observations such as $T = 1000$, roughly 4 years of daily data in an equity market. The data generating process - DGP firstly burns 200 observation for warming up. As discussed in Section 3, we compare the M-GARCH models through using Log-likelihood, AIC and BIC information criteria for in-sample performance.

To compare the prediction ability of M-GARCH models, we use 700 observations as the training sample to estimate the model, then we use estimators to forecast the remaining 300 observations iteratively with one-step forward. We use MAFE and MSFE statistics to evaluate their forecasting performances.

### 1.4.1 Experiment.1

In the first experiment, to generate conditional variance of $y_{1,t}$, we set $\omega_1 = 0.05$, $\alpha_1 = 0.1$ (low ARCH effect) and $\beta_1 = 0.85$ (high GARCH/persistence level) in (1.4.2). In (1.4.3), let $\omega_1 = 0.05$, $\alpha_2 = 0.4$ (moderate valued ARCH) and $\beta_2 = 0.4$ (moderate valued GARCH) for $y_{2,t}$.

We apply nine M-GARCH models, including EWMA, VEC, DVEC, BEKK, DBEKK, CCC, O-GARCH, DCC and RDCC. The results of in-sample performances are displayed in Table 1.4.1.

According to Table 1.4.1, unsurprisingly, the simple EWMA model shows the poorest result over nine models followed by the covariance models (VECH, VECH-DIAG, BEKK and BEKK-DIAG). The linear correlation models (O-GARCH and CCC) and nonlinear correlation models display superior performances. Among conditional correlation models, the DCC and RDCC models are located at the top of the pyramid. After adjusted for the number of parameters, the RDCC model slightly outperforms to the DCC model in the context of Sine/Fast-Sine correlations.

Table 1.4.1: In-sample Statistics Table for Exp.1

'Constant, Sine, Fast-Sine, Ramp and MR' represent DGP with constant, sine, fast-sine, ramp and mean reverting conditional correlations. 'SUM' records the sum of statistics across five types of conditional correlations. The statistics in bold highlight winners.

| | | EWMA | VEC | DVEC | BEKK | DBEKK | CCC | OGARCH | DCC | RDCC |
|---|---|---|---|---|---|---|---|---|---|---|
| Constant | LL | -2790.4 | -2742.6 | -2751.0 | -2742.4 | -2685.7 | -2659.8 | -2751.4 | **-2642.7** | -2653.3 |
| | AIC | 5582.8 | 5527.2 | 5520.0 | 5506.8 | 5385.4 | 5333.6 | 5514.8 | **5307.4** | 5324.6 |
| | BIC | 5587.7 | 5630.1 | 5564.1 | 5560.7 | 5419.7 | 5367.9 | 5544.2 | **5361.3** | 5368.7 |
| Sine | LL | -3009.4 | -2937.1 | -2944.9 | -2937.9 | -2903.6 | -2887.5 | -2943.4 | **-2878.0** | -2882.8 |
| | AIC | 6020.8 | 5916.2 | 5907.8 | 5897.8 | 5821.2 | 5789.0 | 5898.8 | **5778.0** | 5783.6 |
| | BIC | 6025.7 | 6019.1 | 5951.9 | 5951.7 | 5855.5 | 5823.3 | 5928.2 | 5831.9 | **5827.7** |
| Fast-Sine | LL | -3011.5 | -2946.2 | -2944.5 | -2935.6 | -2906.6 | -2890.5 | -2951.4 | -2878.0 | **-2880.4** |
| | AIC | 6025 | 5934.4 | 5907.0 | 5893.2 | 5827.2 | 5795.0 | 5914.8 | 5796.4 | **5778.8** |
| | BIC | 6029.9 | 6037.3 | 5951.1 | 5947.1 | 5861.5 | 5829.3 | 5944.2 | 5850.3 | **5822.9** |
| Ramp | LL | -2967.5 | -2866.1 | -2851.0 | -2845.3 | -2805.4 | -2837.1 | -2861.4 | -2798.0 | **-2793.1** |
| | AIC | 5781 | 5774.2 | 5720.0 | 5712.6 | 5624.8 | 5688.2 | 5734.8 | 5618 | **5604.2** |
| | BIC | 5785.9 | 5877.1 | 5764.1 | 5766.5 | 5659.1 | 5722.5 | 5764.2 | 5671.9 | **5648.3** |
| MR | LL | -3043.2 | -2994.6 | -2991.1 | -2984.4 | -2957.1 | -2970.8 | -2979.8 | **-2942.5** | -2947.8 |
| | AIC | 6088.4 | 6031.2 | 6000.2 | 5990.8 | 5928.2 | 5955.6 | 5971.6 | **5907.0** | 5913.6 |
| | BIC | 6093.3 | 6134.1 | 6044.3 | 6044.7 | 5962.5 | 5989.9 | 6001.0 | 5960.9 | **5957.7** |
| SUM | LL | -14744.0 | -14486.6 | -14482.5 | -14445.6 | -14258.4 | -14245.7 | -14487.4 | **-14148.4** | -14157.4 |
| | AIC | 29498 | 29183.2 | 29055.0 | 29001.2 | 28586.8 | 28561.4 | 29034.8 | 28406.8 | **28404.8** |
| | BIC | 29522.5 | 29697.7 | 29275.5 | 29270.7 | 28758.3 | 28732.9 | 29181.8 | 28676.3 | **28625.3** |

Generally, the in-sample performance can be strengthened by designing an extensively parameterised model; however, this would cause the problem of overfitting and then provide poor predictions. Hence, a good forecasting model requires both simplicity and accuracy. Table 1.4.2 displays MAFE and MSFE statistics for Experiment 1. The nonlinear correlation model outperforms to the generalised covariance and linear correlation models. The DCC model shows its superior forecasting ability over all models. Also notably, the simplest EWMA model gives relatively better prediction in generalised covariance models. The

BEKK model, as a largely parameterised model, shows the worst forecasting performance.

Table 1.4.2: Out-of-sample Statistics Table for Exp.1

'Constant, Sine, Fast-Sine, Ramp and MR' represent DGP with constant, sine, fast-sine, ramp and mean reverting conditional correlations. 'SUM' records the sum of statistics across five types of conditional correlations. The statistics in bold highlight winners.

| | | EWMA | VEC | DVEC | BEKK | DBEKK | CCC | OGARCH | DCC | RDCC |
|---|---|---|---|---|---|---|---|---|---|---|
| Constant | MSFE | 0.0133 | 0.0184 | 0.0153 | 0.0538 | 0.0175 | 0.0034 | 0.0317 | **0.0050** | 0.0056 |
| | MAFE | 0.0874 | 0.0929 | 0.0909 | 0.1489 | 0.0945 | 0.0433 | 0.1287 | **0.0519** | 0.0596 |
| Sine | MSFE | 0.0329 | 0.0396 | 0.0439 | 0.0658 | 0.0375 | 0.0247 | 0.0530 | **0.0266** | 0.0275 |
| | MAFE | 0.1478 | 0.1556 | 0.1662 | 0.1938 | 0.1561 | 0.1351 | 0.1840 | **0.1360** | 0.1403 |
| Fast-Sine | MSFE | 0.0377 | 0.0445 | 0.0445 | 0.0665 | 0.0399 | 0.0255 | 0.0544 | **0.0283** | 0.0294 |
| | MAFE | 0.1591 | 0.1673 | 0.1676 | 0.1952 | 0.1616 | 0.1363 | 0.1868 | **0.1410** | 0.1486 |
| Ramp | MSFE | 0.0854 | 0.0748 | 0.0716 | 0.0998 | 0.0797 | 0.0905 | 0.0950 | **0.0796** | 0.0837 |
| | MAFE | 0.2375 | 0.2196 | 0.2123 | 0.2508 | 0.2301 | 0.2575 | 0.2499 | **0.2239** | 0.2363 |
| MR | MSFE | 0.0567 | 0.0669 | 0.0695 | 0.1039 | 0.0754 | 0.0950 | 0.0954 | **0.0568** | 0.0659 |
| | MAFE | 0.2014 | 0.2091 | 0.2136 | 0.2593 | 0.2265 | 0.2617 | 0.2590 | **0.1961** | 0.2015 |
| SUM | MSFE | 0.2260 | 0.2442 | 0.2448 | 0.3898 | 0.2500 | 0.2391 | 0.3295 | **0.1963** | 0.2121 |
| | MAFE | 0.8332 | 0.8445 | 0.8506 | 1.0480 | 0.8688 | 0.8339 | 1.0084 | **0.7489** | 0.7863 |

### 1.4.2 Experiment.2

The DGP in experiment 2 generates two series with low ARCH and high GARCH effects in conditional variances. Let $\omega_1 = 0.05$, $\alpha_1 = 0.1$ and $\beta_1 = 0.8$ in (1.4.2); $\omega_2 = 0.05$, $\alpha_2 = 0.1$ and $\beta_2 = 0.8$ in (1.4.3).

Table 1.4.3 exhibits results of the in-sample performance. The results are consistent with discussion in experiment 1; the RDCC model interprets better in-sample performance in most of cases, followed by the DCC model. Linear/nonlinear models overall show better performances than generalised covariance models. Besides, to compare the model fitness in different correlation structures, we found that the time-variant correlation structures rise some difficulties to model fitness, evidenced by 'constant' correlation gain larger LL, smaller AIC and BIC scores. The rich dynamics in conditional correlation certainly results a more complex modelling issue, but it is more close to the real case.

The out-of-sample performance comparisons are documented in Table 1.4.4. In most of cases, the CCC model outperforms to others, also the DVEC and the DCC models show good forecasting abilities, which produce half of forecasting error than the EWMA model. The model with poor predicting performances are the EWMA and the BEKK model, the EWMA preforms badly because the model is not well specified according to the data, and the BEKK deserves such results because it suffers from the problem of overfitting. Among good models, the CCC model is the best choice when the correlation structures are generated as 'constant', 'sine' and 'fast sine' types. Since the conditional correlation process either be a constant or be a regular smooth-variant process, the CCC model benefited with its parsimonious beats other models, even the DCC model. In cases that conditional correlations are generated as the 'Ramp' and 'MR' types, the DVEC and DCC model provide more accurate predictions. The 'SUM' sector show that in general, the DCC model obtains the least MAFE and the DVEC model obtains the least MSFE score. Therefore, the CCC model is a better choice when the correlation patterns are less volatile, but since it is a rare case in real world, we might still consider the DCC or the DVEC model to predict the correlation structures.

Table 1.4.3: In-sample Statistics Table for Exp.2

'Constant, Sine, Fast-Sine, Ramp and MR' represent DGP with constant, sine, fast-sine, ramp and mean reverting conditional correlations. 'SUM' records the sum of statistics across five types of conditional correlations. The statistics in bold highlight winners.

| | | EWMA | VEC | DVEC | BEKK | DBEKK | CCC | OGARCH | DCC | RDCC |
|---|---|---|---|---|---|---|---|---|---|---|
| | LL | -3791.0 | -3754.1 | -3766.0 | -3761.6 | -3740.4 | -3736.6 | -3769.5 | **-3719.4** | -3724.7 |
| Constant | AIC | 7584.0 | 7550.2 | 7550.0 | 7545.2 | 7494.8 | 7487.2 | 7551.0 | **7460.8** | 7467.4 |
| | BIC | 7588.9 | 7653.1 | 7594.1 | 7599.1 | 7529.1 | 7521.5 | 7580.4 | 7514.7 | **7511.5** |
| | LL | -4023.5 | -3983.0 | -3981.2 | -3980.8 | -3972.4 | -3970.3 | -3984.9 | -3970.9 | **-3961.8** |
| Sine | AIC | 8049.0 | 8008.0 | 7980.4 | 7983.6 | 7958.8 | 7954.6 | 7981.8 | 7963.8 | **7985.7** |
| | BIC | 8053.9 | 8110.9 | 8024.5 | 8037.5 | 7993.1 | 7988.9 | 8011.2 | 8017.7 | **5827.7** |
| | LL | -4035.6 | -3989.6 | -3989.1 | -3989.0 | -3975.7 | -3971.5 | -3983.0 | -3972.6 | **-3966.2** |
| Fast-Sine | AIC | 8073.2 | 8021.2 | 7996.2 | 8000.0 | 7965.4 | 7957.0 | 7978.0 | 7967.2 | **7950.4** |
| | BIC | 8078.1 | 8124.1 | 8040.3 | 8053.9 | 7999.7 | 7991.3 | 8007.4 | 8021.1 | **7994.5** |
| | LL | -3961.2 | -3895.8 | -3906.3 | -3894.6 | -3893.1 | -3908.5 | -3912.4 | -3874.2 | **-3870.2** |
| Ramp | AIC | 7804.4 | 7833.6 | 7830.6 | 7811.2 | 7800.2 | 7831.0 | 7836.8 | 7770.4 | **7758.4** |
| | BIC | 7809.3 | 7936.5 | 7874.7 | 7865.1 | 7834.5 | 7865.3 | 7866.2 | 7824.3 | **7802.5** |
| | LL | -4102.2 | -4039.8 | -4039.4 | -4028.7 | -4031.5 | -4052.3 | -4054.9 | **-4029.2** | -4031.1 |
| MR | AIC | 8206.4 | 8121.6 | 8096.8 | 8079.4 | 8077.0 | 8118.6 | 8121.8 | 8080.4 | **8080.2** |
| | BIC | 8211.3 | 8224.5 | 8140.9 | 8133.3 | 8111.3 | 8152.9 | 8151.2 | 8134.3 | **8124.3** |
| | LL | -19853.5 | -19662.3 | -19682.0 | -19654.7 | -19613.1 | -19639.2 | -19704.7 | -19566.3 | **-19554.0** |
| SUM | AIC | 39717 | 39534.6 | 39454.0 | 39419.4 | 39296.2 | 39348.4 | 39469.4 | 39242.6 | **39198.0** |
| | BIC | 39741.5 | 40049.1 | 39674.5 | 39688.9 | 39467.7 | 39519.9 | 39616.4 | 39512.1 | **39418.5** |

Table 1.4.4: Out-of-sample Statistics Table for Exp.2

'Constant, Sine, Fast-Sine, Ramp and MR' represent DGP with constant, sine, fast-sine, ramp and mean reverting conditional correlations. 'SUM' records the sum of statistics across five types of conditional correlations. The statistics in bold highlight winners.

| | | EWMA | VEC | DVEC | BEKK | DBEKK | CCC | OGARCH | DCC | RDCC |
|---|---|---|---|---|---|---|---|---|---|---|
| Constant | MSFE | 0.0104 | 0.0086 | 0.0059 | 0.0677 | 0.0107 | **0.0037** | 0.0089 | 0.0043 | 0.0047 |
| | MAFE | 0.0763 | 0.0720 | 0.0581 | 0.1611 | 0.0776 | **0.0477** | 0.0727 | 0.0489 | 0.0504 |
| Sine | MSFE | 0.0351 | 0.0259 | 0.0281 | 0.0760 | 0.0329 | **0.0249** | 0.0333 | 0.0272 | 0.0289 |
| | MAFE | 0.1536 | 0.1306 | 0.1369 | 0.2035 | 0.1486 | **0.1351** | 0.1498 | 0.1358 | 0.1401 |
| Fast-Sine | MSFE | 0.0375 | 0.0330 | 0.0308 | 0.0740 | 0.0350 | **0.0243** | 0.0344 | 0.0297 | 0.0293 |
| | MAFE | 0.1581 | 0.1485 | 0.1459 | 0.2040 | 0.1531 | **0.1347** | 0.1522 | 0.1431 | 0.1403 |
| Ramp | MSFE | 0.0933 | 0.0683 | **0.0716** | 0.1147 | 0.0806 | 0.0902 | 0.0936 | 0.0726 | 0.0816 |
| | MAFE | 0.2475 | **0.2105** | 0.2153 | 0.2658 | 0.2264 | 0.2570 | 0.2525 | 0.2135 | 0.2339 |
| MR | MSFE | 0.0577 | 0.0605 | **0.0538** | 0.1240 | 0.0745 | 0.0901 | 0.0993 | 0.0601 | 0.0745 |
| | MAFE | 0.1992 | 0.2003 | **0.1892** | 0.2787 | 0.2225 | 0.2553 | 0.2664 | 0.2023 | 0.2255 |
| SUM | MSFE | 0.2340 | 0.1963 | **0.1902** | 0.4564 | 0.2337 | 0.2332 | 0.2695 | 0.1939 | 0.2190 |
| | MAFE | 0.8347 | 0.7619 | 0.7454 | 1.1131 | 0.8282 | 0.8298 | 0.8936 | **0.7436** | 0.7902 |

### 1.4.3 Experiment.3

Experiment 3 generates data with moderate ARCH and GARCH effect in conditional variances. Let $\omega_1 = 0.05$, $\alpha_1 = 0.4$ and $\beta_1 = 0.5$ in (1.4.2), $\omega_2 = 0.05$, $\alpha_2 = 0.4$ and $\beta_2 = 0.5$ in (1.4.3). Results in Table 1.4.5 is also consistent with those of experiment 1 and 2. The RDCC model fits the data better in cases of 'Constant', 'Sine' and 'Fast-Sine' typed conditional correlations. And the best one turns to the DCC model when the data follows the 'Ramp' and 'MR' typed conditional correlation structures. Also considering the computational efficiency and relative accuracy, it is worth to notice that the CCC model produces competitive LL, AIC and BIC as well. Beyond our expectation, the well-specified BEKK model does not show good fitness, this might be due to the conditional correlation processes do not follow a recursive dynamic process.

The results in Table 1.4.6 expresses the linear/nonlinear models can predict better, where the DCC model consistently outperforms to others by obtaining less forecasting errors. This result is consistent with results in experiment 1. The BEKK model gets the highest forecasting errors, followed by the DBEKK and the VEC model, as discussed above, these models are heavily parametrised and deserved poor forecasting abilities, even the most naive and efficient model - EWMA model perform better than them.

Table 1.4.5: In-sample Statistics Table for Exp.3

'Constant, Sine, Fast-Sine, Ramp and MR' represent DGP with constant, sine, fast-sine, ramp and mean reverting conditional correlations. 'SUM' records the sum of statistics across five types of conditional correlations. The statistics in bold highlight winners.

| | | EWMA | VEC | DVEC | BEKK | DBEKK | CCC | OGARCH | DCC | RDCC |
|---|---|---|---|---|---|---|---|---|---|---|
| Constant | LL | -2445.4 | -2428.1 | -2437.8 | -2418.9 | -2400.6 | -2326.5 | -2421.0 | -2311.2 | **-2305.7** |
| | AIC | 4892.8 | 4898.2 | 4893.6 | 4859.8 | 4815.2 | 4667.0 | 4854.0 | 4644.4 | **4629.4** |
| | BIC | 4897.7 | 5001.1 | 4937.7 | 4913.7 | 4849.5 | 4701.3 | 4883.4 | 4698.3 | **4673.5** |
| Sine | LL | -2621.7 | -2613.1 | -2623.8 | -2599.9 | -2588.8 | -2546.8 | -2602.0 | -2549.6 | **-2546.0** |
| | AIC | 5245.4 | 5268.2 | 5265.6 | 5221.8 | 5191.6 | 5107.6 | 5216.0 | 5121.2 | **5110.0** |
| | BIC | 5250.3 | 5371.1 | 5309.7 | 5275.7 | 5225.9 | 5141.9 | 5245.4 | 5175.1 | **5154.1** |
| Fast-Sine | LL | -2629.9 | -2620.9 | -2607.2 | -2597.3 | -2601.3 | -2552.1 | -2594.2 | -2553.3 | **-2546.0** |
| | AIC | 5261.8 | 5283.8 | 5232.4 | 5216.6 | 5216.6 | 5118.2 | 5200.4 | 5128.6 | **5110.0** |
| | BIC | 5266.7 | 5386.7 | 5276.5 | 5270.5 | 5250.9 | 5152.5 | 5229.8 | 5182.5 | **5154.1** |
| Ramp | LL | -2528.2 | -2526.1 | -2518.0 | -2518.4 | -2496.3 | -2495.7 | -2529.5 | **-2460.5** | -2470.7 |
| | AIC | 5058.4 | 5094.2 | 5054.0 | 5058.8 | 5006.6 | 5005.4 | 5071.0 | **4943** | 4959.4 |
| | BIC | 5063.3 | 5197.1 | 5098.1 | 5112.7 | 5040.9 | 5039.7 | 5100.4 | **4996.9** | 5003.5 |
| MR | LL | -2668.2 | -2655.4 | -2662.7 | -2648.5 | -2640.6 | -2614.5 | -2635.8 | **-2591.3** | -2606.7 |
| | AIC | 5338.4 | 5352.8 | 5343.4 | 5319.0 | 5295.2 | 5243.0 | 5283.6 | **5204.6** | 5231.4 |
| | BIC | 5343.3 | 5455.7 | 5387.5 | 5372.9 | 5329.5 | 5277.3 | 5313.0 | **5258.5** | 5275.5 |
| SUM | LL | -12893.4 | -12843.6 | -12849.5 | -12783.0 | -12727.6 | -12535.6 | -12782.5 | **-12465.9** | -12476.1 |
| | AIC | 25796.8 | 25897.2 | 25789.0 | 25676.0 | 25525.2 | 25141.2 | 25625.0 | 25041.8 | **25040.2** |
| | BIC | 25821.3 | 26411.7 | 26009.5 | 25945.5 | 25696.7 | 25312.7 | 25772.0 | 25311.3 | **25260.7** |

Table 1.4.6: Out-of-sample Statistics Table for Exp.3

'Constant, Sine, Fast-Sine, Ramp and MR' represent DGP with constant, sine, fast-sine, ramp and mean reverting conditional correlations. 'SUM' records the sum of statistics across five types of conditional correlations. The statistics in bold highlight winners.

| | | EWMA | VEC | DVEC | BEKK | DBEKK | CCC | OGARCH | DCC | RDCC |
|---|---|---|---|---|---|---|---|---|---|---|
| Constant | MSFE | 0.0158 | 0.0352 | 0.0250 | 0.0516 | 0.0311 | 0.0044 | 0.0373 | **0.0051** | 0.0062 |
| | MAFE | 0.0950 | 0.1351 | 0.1145 | 0.1540 | 0.1245 | 0.0490 | 0.1447 | **0.0514** | 0.0534 |
| Sine | MSFE | 0.0371 | 0.0557 | 0.0536 | 0.0734 | 0.0522 | 0.0290 | 0.0499 | **0.0281** | 0.0312 |
| | MAFE | 0.1564 | 0.1847 | 0.1803 | 0.2084 | 0.1793 | 0.1436 | 0.1807 | **0.1392** | 0.1524 |
| Fast-Sine | MSFE | 0.0409 | 0.0618 | 0.0619 | 0.0747 | 0.0550 | 0.0254 | 0.0487 | **0.0278** | 0.0304 |
| | MAFE | 0.1641 | 0.1953 | 0.1948 | 0.2108 | 0.1847 | 0.1364 | 0.1780 | **0.1397** | 0.1458 |
| Ramp | MSFE | 0.0923 | 0.0780 | 0.0832 | 0.1005 | 0.0877 | 0.0891 | 0.1062 | **0.0761** | 0.0802 |
| | MAFE | 0.2045 | 0.2287 | 0.2152 | 0.2618 | 0.2255 | 0.2581 | 0.2760 | **0.1942** | 0.2202 |
| MR | MSFE | 0.0594 | 0.0814 | 0.0723 | 0.1085 | 0.0777 | 0.0914 | 0.1084 | **0.0569** | 0.0616 |
| | MAFE | 0.2045 | 0.2287 | 0.2152 | 0.2618 | 0.2255 | 0.2581 | 0.2760 | **0.1942** | 0.2094 |
| SUM | MSFE | 0.2455 | 0.3121 | 0.2960 | 0.4087 | 0.3037 | 0.2393 | 0.3505 | **0.1940** | 0.2096 |
| | MAFE | 0.8245 | 0.9725 | 0.9200 | 1.0968 | 0.9395 | 0.8452 | 1.0554 | **0.7187** | 0.7812 |

### 1.4.4 Experiment.4

Experiments 1-3 not consider the leverage effect in volatilities. We then specify conditional variances formed as a GJR(1,1,1) model. Let $\omega_1 = 0.05$, $\alpha_1 = 0.1$, $\gamma_1 = 0.05$ and $\beta_1 = 0.8$ (low ARCH but high GARCH) in (1.4.4); and let $\omega_2 = 0.05$, $\alpha_2 = 0.1$, $\gamma_2 = 0.05$ and $\beta_2 = 0.8$ in (1.4.5). To account for a leverage effect, also combined with results from experiment 1-3, here we only compare nonlinear correlation models, including DCC, ADCC, RDCC and ARDCC.

Table 1.4.7 exhibits in-sample results. First of all, as expected, by introducing an asymmetric term, the ADCC and ARDCC models are superior to the DCC and RDCC models, respectively, indicating that asymmetric models are supposed to be selected as long as the data shows the leverage effect. Secondly, LL statistics show that ADCC and ARDCC models are more or less same. But after adjusting the number of parameters, both the AIC and BIC criteria pointed out that the ARDCC model is slightly better than the ADCC model. Table 1.4.8 presents out-of-sample performances. The DCC model still provides the least error in forecasting. One explanation is that the prediction ability is contributed by two aspects: the explanatory ability and the level of parameterisation. Once the overload of the second aspect was greater than the first aspect, the predictability of the model would get worse. Thus, although ADCC models interpret data better in-sample, the parsimonious DCC model brings more precise predictors.

Table 1.4.7: In-sample Statistics Table for Exp.4

'Constant, Sine, Fast-Sine, Ramp and MR' represent DGP with constant, sine, fast-sine, ramp and mean reverting conditional correlations. 'SUM' records the sum of statistics across five types of conditional correlations. The statistics in bold highlight winners.

| | | DCC | ADCC | RDCC | ARDCC |
|---|---|---|---|---|---|
| Constant | LL | -2441.5 | -2430.8 | -2453.1 | **-2427.7** |
| | AIC | 4905.0 | 4889.6 | 4924.2 | **4877.4** |
| | BIC | 4958.9 | 4958.2 | 4968.3 | **4931.3** |
| Sine | LL | -2681.6 | **-2656.7** | -2673.8 | -2658.4 |
| | AIC | 5385.2 | 5341.4 | 5365.6 | **5338.8** |
| | BIC | 5439.1 | 5410.0 | 5409.7 | **5392.7** |
| Fast-Sine | LL | -2677.8 | -2666.9 | -2681.0 | **-2662.4** |
| | AIC | 5377.6 | 5361.8 | 5380.0 | **5346.8** |
| | BIC | 5431.5 | 5430.4 | 5424.1 | **5400.7** |
| Ramp | LL | -2593.8 | -2581.6 | -2598.8 | **-2574.3** |
| | AIC | 5209.6 | 5191.2 | 5215.6 | **5170.6** |
| | BIC | 5263.5 | 5259.8 | 5259.7 | **5224.5** |
| MR | LL | -2743.8 | **-2717.6** | -2730.5 | -2724.5 |
| | AIC | 5509.6 | **5463.2** | 5479.0 | 5471.0 |
| | BIC | 5563.5 | 5531.8 | 5523.1 | **5524.9** |
| SUM | LL | -13138.5 | -13053.6 | -13137.2 | **-13047.3** |
| | AIC | 26387.0 | 26247.2 | 26364.4 | **26204.6** |
| | BIC | 26656.5 | 26590.2 | 26584.9 | **26474.1** |

Table 1.4.8: Out-of-sample Statistics Table for Exp.4

'Constant, Sine, Fast-Sine, Ramp and MR' represent DGP with constant, sine, fast-sine, ramp and mean reverting conditional correlations. 'SUM' records the sum of statistics across five types of conditional correlations. The statistics in bold highlight winners.

| | | CCC | DCC | ADCC | RDCC | ARDCC |
|---|---|---|---|---|---|---|
| Constant | MSFE | 0.0053 | 0.0054 | **0.0045** | 0.0059 | 0.0053 |
| | MAFE | 0.0548 | 0.0534 | **0.0493** | 0.0553 | 0.0536 |
| Sine | MSFE | 0.0255 | **0.0256** | 0.0305 | 0.0263 | 0.0323 |
| | MAFE | 0.1373 | **0.1330** | 0.1429 | 0.1399 | 0.1475 |
| Fast-Sine | MSFE | 0.0264 | **0.0287** | 0.0309 | 0.0292 | 0.0317 |
| | MAFE | 0.1383 | **0.1412** | 0.1455 | 0.1426 | 0.1489 |
| Ramp | MSFE | 0.0915 | **0.0724** | 0.0744 | 0.0818 | 0.0836 |
| | MAFE | 0.2584 | **0.2137** | 0.2139 | 0.2404 | 0.2444 |
| MR | MSFE | 0.0950 | **0.0582** | 0.0642 | 0.0711 | 0.0770 |
| | MAFE | 0.2636 | **0.1983** | 0.2066 | 0.2281 | 0.2335 |
| SUM | MSFE | 0.2437 | **0.1903** | 0.2045 | 0.2143 | 0.2299 |
| | MAFE | 0.8524 | **0.7396** | 0.7582 | 0.8063 | 0.8279 |

### 1.4.5 Experiment.5

For a deeper look into the effect of conditional correlations, one noticeable characteristics in the above four experiments is that M-GARCH models perform better in data with constant correlations. It reveals that time-varying correlations weaken the goodness-of-fit of M-GARCH models, and things can go even worse if there occurs a structural break in the correlation structure. Segmenting the correlation process into sub-samples according to the break location is a possible solution to increase model fitting in-sample, i.e. considering 'local conditional correlation' or 'change-point in correlation structure'.

In experiment 5, we allow a mean change in conditional correlation (Recall Section 1.4.0.2). The conditional variances are still generated by the GJR(1,1,1) model. Let $\omega_1 = 0.05$, $\alpha_1 = 0.1$, $\gamma_1 = 0.05$ and $\beta_1 = 0.8$ in (1.4.4); and $\omega_2 = 0.05$, $\alpha_2 = 0.1$, $\gamma_2 = 0.05$ and $\beta_2 = 0.8$ in (1.4.5).

It is difficult to compare out-of-sample performance in Experiment 5 because it requires an online sequential change-point detection in conditional correlation structures. Hence, we do not discuss M-GARCH models' predictability with structural breaks; we only discuss for in-sample performances. Experiment 4 suggests that asymmetric models can interpret the leverage effect better in in-samples. Thus, in Experiment 5, we only compare nonlinear asymmetric correlation models containing the following: ADCC, ADCC with structural break dummy (ADCCSB), ARDCC and Asymmetric RDCC with structural break dummy (ARDCCSB).

Table 1.4.9 demonstrates results of the in-sample performance. The ADCCSB model and the ARDCCSB show roughly equivalent performances; however, they significantly outperform corresponding models without structural break dummy. Consistent with previous experiments, the RDCC-type model provides better fitting than the DCC-type model.

Therefore, our results reveal that detecting change-point in conditional correlation structure is crucial to model the covariance of real financial data. Besides, for empirical study with real financial data, detecting structural break in a certain data set also helps to discover intuitive economic or financial meaning before and after the break location.

Table 1.4.9: In-sample Statistics Table for Exp.5

'Constant, Sine, Fast-Sine, Ramp and MR' represent DGP with constant, sine, fast-sine, ramp and mean reverting conditional correlations. 'SUM' records the sum of statistics across five types of conditional correlations. The statistics in bold highlight winners.

| | | ADCC | ADCCSB | ARDCC | ARDCCSB |
|---|---|---|---|---|---|
| Constant | LL | -2429.4 | -2417.5 | -2441.6 | **-2414.6** |
| | AIC | 4886.8 | 4863.0 | 4905.2 | **4851.2** |
| | BIC | 4955.4 | 4931.6 | 4959.1 | **4905.1** |
| Sine | LL | -2665.3 | **-2638.8** | -2658.4 | -2640.7 |
| | AIC | 5358.6 | 5305.6 | 5338.8 | **5303.4** |
| | BIC | 5427.2 | 5374.2 | 5392.7 | **5357.3** |
| Fast-Sine | LL | -2663.4 | -2649.0 | -2668.7 | **-2645.0** |
| | AIC | 5354.8 | 5326.0 | 5359.4 | **5312** |
| | BIC | 5423.4 | 5394.6 | 5413.3 | **5365.9** |
| Ramp | LL | -2577.1 | -2561.5 | -2580.4 | **-2556.5** |
| | AIC | 5182.2 | 5151.0 | 5182.8 | **5135.0** |
| | BIC | 5250.8 | 5219.6 | 5236.7 | **5188.9** |
| MR | LL | -2728.9 | **-2698.3** | -2715.7 | -2709.5 |
| | AIC | 5485.8 | **5424.6** | 5453.4 | 5441.0 |
| | BIC | 5554.4 | **5493.2** | 5507.3 | 5494.9 |
| SUM | LL | -13064.1 | **-12965.1** | -13064.8 | -12966.3 |
| | AIC | 26268.2 | 26070.2 | 26239.6 | **26042.6** |
| | BIC | 26611.2 | 26413.2 | 26509.1 | **26312.1** |

## 1.5 Summary

This chapter reviewed classical M-GARCH models including generalized covariance models, linear and nonlinear correlation models. Aiming to compare their performances from aspects of in-sample fitting and out-of-sample forecasting, we designed a Monte Carlo simulation study to assess these models through fitting criteria and forecasting errors.

Generating data with different types of conditional correlations, the simulation results found that the nonlinear correlation models are better than other two clusters in both of in-sample and out-of-sample. Particularly, without considering the financial leverage effect and structural change, the RDCC model is superior than others in in-sample, and the DCC model shows the best predictability. Once the leverage effect is considered in DGP, asymmetric models help to identify the leverage effect in in-sample, although the symmetric DCC model still gives the most precise predictors. Lastly, detecting structural break and using break dummies to segment correlations improve the fitting of M-GARCH models in in-samples. The issue of structural break in out-of-sample performance requires a sequential change point detection method, and we leave it for later chapters.

# Chapter 2

# The Absorption Ratio: a Leading Indicator of the Financial Fragility

## 2.1 Introduction

Systemic risk is the risk harming the stability of the entire financial system, which is easily to be triggered because of the financial fragility. Compared to risk associated with one of the industries or sectors, systemic risk has a much larger influence, such as the explosion of the Great Recession in 2008. However, there is not a unique framework to study systemic risk, and it is controversial to have a standard method to measure systemic risk. Early studies on systemic risk come mainly from banking sectors because most systemic risk stems from bank runs and quickly spread panic to the whole system. Thus, the strong interlinkages in financial systems could be a sign of financial fragility and then be an intensifier for systemic risk. Following this idea, some literature tried to use the correlation level of the financial market to measure systemic risk, and the results showed that studying the dynamic correlation structure makes a lot of sense in this topic (Kirtzman et al. 2012 and Billio et al. 2012). However, during recent years, rarely papers discussed the role of dynamic correlation structure in the field of preventing systemic risk. This chapter tries to contribute on this gap and aims to investigate whether studying the dynamic correlation structure of a financial system can provide any sort of pre-warning mechanism to financial fragility, and even systemic risk. Also, this chapter provides an application to empirically forecast correlation structure by using the class of models discussed in the first chapter.

We follow the method proposed by Kirtzman et al. (2012), using absorption ratio to express the concentrated level in the correlation structure. By analysing the relationship between the absorption ratio and market returns, we found that absorption ratio can be applied as an early warning indicator for the financial fragility. Hence, the predicted absorption ratio can provide potential indications to policy makers and investors.

This chapter is structured as below. In the next section, we review relevant literature about measuring systemic risk. Section 2.3 discusses two data sets in U.S. equity markets. Then, we investigate whether the absorption ratio of conditional correlations can be used as an early warning indicator in Section 2.4. Section 2.5 provides two methods to forecast absorption ratio for warning financial fragility, and a summary concludes.

## 2.2 Literature Review

The financial fragility is always connected with the systemic risk, while the concept of systemic risk is hard to define. It is a result experienced from historical data, but it is hard to capture or monitor for both policymakers and investors. Existing literature provides many different methods to measure systemic risk. De Bandt and Hartmann (2000) provided a survey literature on systemic risk, and in most of the literature reviewed, the systemic risk is mainly captured by the financial contagion effect (Forbes and Rigobon, 2002). The financial contagion is such an effect that the distress from a financial sector rapidly propagates to the whole system via some amplification mechanisms. Using the concept of financial contagion, systemic risk was then measured by the spillover effect from financial instability to the real economy (Group of ten, 2001).

Although these methods offered an explanatory investigation of systemic risk, they drew relatively less attention to recent years, particularly after the Great Recession in 2008. The huge market loss during the crisis led people to consider more about modelling risk and expected loss in the financial sector. Hence, Billio et al. (2012) defined the systemic risk as any situation that results in a market loss through weakening financial stability and reducing public confidence. Also, a recent survey (Bisias et al. 2012) documented all types of methodologies, and indicated that the level of risk concentration is an useful indicator for the systemic risk.

To measure the level of risk concentration between financial institutions or portfolios, there are three widely used methods in recent literature. Adrian and Brunnermeier (2011) proposed the concept of conditional value-at-risk (CoVaR) which provides the VaR of one financial entity conditional on other financial entities that are under financial distress. Later, Girardi and Ergün (2013) improved the CoVaR by considering a more specific case of financial distress, estimated by multivariate GARCH models. The second method is proposed by Acharya et al. (2011), who proposed systemic expected shortfall to measure the expected loss of a single financial entity conditional on the entire distress environment in such a financial sector. The third method is mainly applied in studying banking systems. Huang et al. (2012) used distressed insurance loss to measure the systemic risk. The idea is that the systemic risk can be measured through knowing how much insurance premium is required for covering the expected loss during a financial distress period. The first two methods discuss how systemic risk from one financial institution or the entire financial sector marginally affects another financial institution, and they provide a way to explore systemic risk and its influence. The last method explains systemic risk through the cover of insurance, which is suitable for the banking system, but not instinctively to measure systemic risk for portfolios. The above streams discussed mainly measures the expected loss of financial institutions during financial distresses, but does not measure the level of systemic risk itself. Billio et al. (2012) argued that an essence of these methods is that they calculate the expected loss for all financial institutions simultaneously during a crisis; however, this has rarely occurred since the rapid development of financial systems. Moreover, since the conditional VaR and expected shortfall are highly related to market volatility, low market volatility during boom and oscillation periods would reduce the effectiveness of these methods. Hence, it is essential to consider the evolution of correlation for early warning the financial fragility, and eventually to systemic risk.

Inspired by the well-known fact that the correlation increases during financial crises, the IMF (2009) reported that conditional correlations can effectively monitor financial stress. Beside market volatility, the covariance structure shows a more intuitive way for the systemic risk. Because principal component analysis provides sufficiently precise pictures of covariance matrix, Pukthuanthong and Roll (2009) simultaneously regressed country returns at time $t$ on principal components extracted from the conditional covariance at $t-1$ and used the average value of the $R^2$ from this regression to measure the integration level of international financial markets.

Similarly, Billio et al. (2012) applied the PCA to examine the integration level among financial institutions and further to indicate the extent of systemic risk. Through monitoring a single principal component, they found that during financial crises, the component explains a greater proportion of total variance because industries are more unified than during normal periods. As an extension, Kirtzman et al. (2011) applied the PCA to a rolling sample and defined the concept of absorption ratio which uses a finite number of components to explain the proportion from the total variation in the system. Later, Melkuev (2014) introduced the absorption ratio into even broader markets, including equity, bond, CDS and currency markets, and then, in order to provide more econometric sense of absorption ratio, he used Granger causality test to show that the absorption ratio is a leading indicator for the market return.

Bisias et al. (2012) defined 'early warning' models as those that can provide forecast power for systemic monitoring. Modelling the serial correlation of objective processes can provide predictions, such as forecasting the systemic risk in the hedge fund industry (Getmansky et al. 2004). Alessi and Detken (2009) used a signalling approach to predict some low-frequency financial variables as early warning indicators. However, since Melkuev (2014) showed that conditional covariance is informative in constructing an early warning indicator, the existed literature barely provide a leading indicator for financial fragility through forecasting the concentration level of market correlation. The present chapter will fill this gap. It is commonly accepted that there is a trade-off between model parametrisation and the accuracy of the forecast. Considering both forecasting efficiency and accuracy, relatively parsimonious models are recommended. The existed simulation study found that the DCC model beats other multivariate GARCH models in out-of-sample performance (Laurent et al., 2012), although the equally weighted moving average model gives a more efficient prediction.

Then, in order to discuss the usage of early warning indicators, several kinds of backing tests are available, for example, the Granger causality test (Billio et al. 2012), among others. More recently, due to the fact that systemic risks are easier to observe, Giglio et al. (2015) pointed out that the quantile regression is an appropriate tool for examining the impact of systemic risk on the tails. Quantile regression provides a flexible method to investigate the impact of one distribution on another distribution, from quantile to quantile. Nonetheless, even though quantile regression can simplify the interlinkage between systemic risk and market states to a pure statistical issue, it bears the limitation of not being easily interpretable.

Distinguishing market states should give a better understanding about how the pre-warning indicator and systemic risk evolve. Billio et al. (2012) used the proportion explained by the components to identify a two-state market. Also, Melkuev (2014) applied a monitoring change-point method on the indicator of integration level to classify the market states. To detect the mean change of a single process, Csörgő and Horváth (1997) provided various types of methods. More recently, Aue et al. (2009) derived an asymptotic framework of cumulative sum statistics to detect changes in the covariance structure. However, as this could deviate from the topic studied here, we may leave it for later works.

In this chapter, we study the correlation concentration level in the U.S. equity market, and use quantile regression and change point detection methods to examine whether the correlation concentration level can be used as a pre-warning indicator, lastly provide a method to forecast correlation concentration level.

## 2.3 The Data in U.S. Equity Markets

To study the early warning role of correlation concentration level to the systemic risk, we study the events of 'dot-com' bubble and the great recession in the U.S. market. In 1990s, the U.S. equity market took roughly 5-10 years to form 'dot-com' bubbles, and commonly consider that the burst of bubble occurred during 1999-2001. Hence, in order to minimize the clash effect from the Asian crisis around 1997, we collect the data set from 01-Jan-1998 to 31-Dec-2001. With regard to the great recession, treating September 2008 as the centre of crisis, we collect the data set between 01-Sep-2006 and 31-Aug-2010.

To measure the market correlation concentration level, it would be better if all market information are included, but it is not feasible to collect every share in the U.S. market because it is unrealistic to analyse the conditional covariance structure of an over ten thousands dimensional data set. As an alternative solution, we collect S&P 500 sector indices as proxies for entire market shares. For each sector, the index is computed as weighted portfolio of representative shares in such a sector. There are ten sectors in total, containing Energy (EN), Financial (FI), Consumer Discretionary (CD), Consumer Staples (CS), Health Care (HC), Information Technology (IT), Industrial (IN), Material (MA), Utilities (UT) and Technology Service (TS).

We then collect the S&P 500 index to be the market index. The raw data are the daily closing price data adjusted for splits, dividends and distributions, and then the log return is utilized to induce stationarity.

$$r_t = log(p_t/p_{t-1})$$

where $r_t$ and $p_t$ denote the log return and adjusted price series, respectively. The data are collected from the Datastream. Table 2.3.1 documents the statistics summary of the log-returns. The Augmented Dickey-fuller test and the KPSS test reveal that all log-returns are stationary. Also, the statistics show that the systemic risk during 2006-2010 is more severe than the period during the 'dot-com' bubble, because of more negative signed values in mean and higher values in kurtosis during the great recession.

Table 2.3.1: The Statistics Summary

The table reports the statistics summary of log-return of indices, where 'MKT' implies the market index, and others are S&P sector indices as denoted above. The mean value with the unit of $1.0 \cdot e^{-3}$.

| Dot-Com Bubble (1998-2001) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | MKT | $r_{EN}$ | $r_{FI}$ | $r_{CD}$ | $r_{CS}$ | $r_{HC}$ | $r_{IT}$ | $r_{IN}$ | $r_{MA}$ | $r_{UT}$ | $r_{TS}$ |
| Mean | 0.0701 | 0.0507 | 0.0891 | 0.1400 | -0.0012 | 0.1644 | 0.1356 | 0.1011 | -0.0257 | -0.0074 | -0.0381 |
| Std | 0.0055 | 0.0068 | 0.0077 | 0.0068 | 0.0053 | 0.0065 | 0.0113 | 0.0062 | 0.0068 | 0.0054 | 0.0072 |
| Skewness | -0.1522 | 0.2410 | 0.1679 | -0.2616 | -0.2026 | -0.1041 | 0.2154 | -0.2940 | 0.2563 | -0.3390 | -0.0320 |
| Kurtosis | 5.2201 | 3.8019 | 4.6214 | 7.3814 | 9.2442 | 5.8764 | 5.3791 | 6.7662 | 5.2257 | 4.9046 | 4.7915 |
| ADF | Reject | Reject | Reject | Reject | Reject | Reject | Reject | Reject | Reject | Reject | Reject |
| KPSS | Not Reject | Not Reject | Not Reject | Not Reject | Not Reject | Not Reject | Not Reject | Not Reject | Not Reject | Not Reject | Not Reject |
| The Great Recession (2006-2010) | | | | | | | | | | | |
| | MKT | $r_{EN}$ | $r_{FI}$ | $r_{CD}$ | $r_{CS}$ | $r_{HC}$ | $r_{IT}$ | $r_{IN}$ | $r_{MA}$ | $r_{UT}$ | $r_{TS}$ |
| Mean | -0.0928 | -0.0449 | -0.3753 | -0.0409 | 0.0162 | -0.0667 | 0.0047 | -0.0827 | -0.0219 | -0.0506 | -0.0903 |
| Std | 0.0073 | 0.0101 | 0.0140 | 0.0082 | 0.0048 | 0.0056 | 0.0076 | 0.0078 | 0.0094 | 0.0066 | 0.0075 |
| Skewness | -0.2029 | -0.3248 | -0.0141 | 0.0616 | 0.1916 | 0.1108 | 0.0325 | -0.3209 | -0.3529 | 0.4667 | 0.3848 |
| Kurtosis | 10.4470 | 12.6354 | 10.0639 | 8.3257 | 13.3060 | 14.5545 | 8.6724 | 7.1882 | 8.0930 | 13.0086 | 11.6099 |
| ADF | Reject | Reject | Reject | Reject | Reject | Reject | Reject | Reject | Reject | Reject | Reject |
| KPSS | Not Reject | Not Reject | Not Reject | Not Reject | Not Reject | Not Reject | Not Reject | Not Reject | Not Reject | Not Reject | Not Reject |

Besides, in order to account for the market return, we also consider other macroeconomic variables in daily frequency. We first collect the federal funds rate, $IR_t$, as the interest rate in the U.S. market. The second variable is the S&P 500 volatility index, $Vol_t$, which is expected to capture the risk attitude of the market. We take the log difference of the $Vol_t$ for the stationarity. The third variable is collected from the FRED database: the TED spread, $Ted_t$, defined as the spread between 3-month Treasury Bill and 3-month LIBOR based on U.S. dollars, which is considered as a market liquidity factor. All macroeconomic variables are ranged between 01-Jan-1998 and 31-Dec-2001 for the dot-com bubble sample, and between 01-Sep-2006 and 31-Aug-2010 for the great recession sample. With regarding to the stationarity of these macroeconomic variables, both of KPSS and ADF tests indicate that the $IR_t$ and $Ted_t$ are non-stationary, the $Vol_t$ is rather stationary. Therefore, we shall apply the first order differences, $\Delta IR_t$ and $\Delta Ted_t$, and the level $Vol_t$ in later analysis.

## 2.4 An Early Warning Indicator - The Absorption Ratio

In this section, in order to investigate whether the correlation concentration level can be seen as an early warning indicator for financial fragility and systemic risk, we introduce the absorption ratio. In the context of our data set, the market correlation concentration level or the market integration level can be traced through computing the absorption ratio from the system of S&P 500 sectors indices.

### 2.4.1 The Absorption Ratio

Let us consider a $d$-dimensional random vector $\mathbf{Y}_t = [y_{1,t}, y_{2,t}, \ldots, y_{d,t}]$. Assuming that $y_{i,t}$ and $y_{j,t}$ for $i \neq j$ are linear independent, and $\mathbf{Y}_t$ has mean $\mu$ and covariance $\mathbf{\Sigma} \in \mathbb{R}^d$.
In order to study the covariance structure of $\mathbf{Y}$, by eigen-decomposing, we let,

$$\begin{aligned} \mathbf{\Sigma} &= \textstyle\sum_{i=1}^{d} \lambda_i \boldsymbol{\nu}_i \boldsymbol{\nu}_i^\top \\ &= \mathbf{\Gamma \Lambda \Gamma}^\top \end{aligned} \tag{2.4.1}$$

where $\boldsymbol{\nu}_i$ is the $i$th egienvector, and $\boldsymbol{\nu}_i^\top \boldsymbol{\nu}_i = 1$ under the standardized scale, and $\lambda_i$ is the $i$th eigenvalue, $\mathbf{\Lambda}$ is a diagonal matrix composed by $\lambda_i$, $1 \leq i \leq d$. The eigen-decomposition can be achieved by principal component analysis, which provides a linear transformation of $\mathbf{Y} \to \mathbf{PC}$, where $\mathbf{PC}$ is a vector of principal components $PC_i$. The principal components are orthogonal and their variances are decreasing gradually, meaning that the corresponding eigenvalues $\lambda_i$ extracted are decreasing ($\lambda_1 > \lambda_2 > \cdots > \lambda_d$).
This is achieved through the mapping of $\mathbf{PC} = \boldsymbol{\nu}' \mathbf{Y}$, and $\mathbf{\Lambda}$ provides the variance for each

principal components, that is,

$$\begin{aligned} Var[PC_i] &= Cov[\boldsymbol{\nu}_i^\top \mathbf{Y}, \boldsymbol{\nu}_i^\top \mathbf{Y}] \\ &= \boldsymbol{\nu}_i^\top Cov[\mathbf{Y}, \mathbf{Y}] \boldsymbol{\nu}_i \\ &= \boldsymbol{\nu}_i^\top \boldsymbol{\Sigma} \boldsymbol{\nu}_i \\ &= \boldsymbol{\nu}_i^\top \boldsymbol{\nu}_i \lambda_i \\ &= \lambda_i \end{aligned}$$

so that the $\lambda_i$ measures the variance of $i$th principal component and $\boldsymbol{\nu}_i$ shows the direction of $\lambda_i$.

Given that the principal components are orthogonal, the total variation of $\mathbf{Y}$ can be expressed as,

$$\Psi = \sum_{i=1}^{d} \lambda_i \tag{2.4.2}$$

and the variation proportion explained by $i$th principal components is,

$$\psi_i = \frac{\lambda_i}{\Psi}$$

Hence, we can define the variation accounted by the first $k$ components $AR_k$ as the absorption ratio with first $k$ principal components,

$$AR_k = \sum_{i=1}^{k} \psi_i, \quad k \leq d \tag{2.4.3}$$

In our case, the $d$-dimensional random vector is $\mathbf{r}_t = [r_{CD,t}, r_{IT,t}, r_{CS,t}, r_{EN,t}, r_{FI,t}, r_{HC,t}, r_{IN,t}, r_{MA,t}, r_{TE,t}, r_{UT,t}]^\top$ and $d = 10$.

To interpret the absorption ratio $AR_k$, according to Equation 2.4.3, we know that $AR_k$ accounts for the concentration level of cumulative eigenvalues, which explains the co-movement of variations in the random vector $\mathbf{r}_t$. Therefore, during a vulnerable period, highly correlated variables or integrated systems are supposed to have high values of $AR_k$ with a finite number of $k \leq d$.

We apply the absorption ratio to track the movement of concentration level of the market correlation structure. Thus, conditional on historical filtration $\mathcal{F}_{t-1}$, the conditional absorption ratio $E(AR_{k,t}|\mathcal{F}_{t-1})$ can be calculated from the estimator of covariance matrix $\hat{\mathbf{\Sigma}}_t$ conditional on filtration $\mathcal{F}_{t-1}$.

We then adopt M-GARCH models to estimate consistently the coefficients of the conditional covariance matrix, using as initial values those of the unconditional covariance matrix $\mathbf{\Sigma}$, so that the estimator of the conditional correlation $\hat{\mathbf{\Sigma}}_t$ can be obtained iteratively.

Considering the fact that the EWMA is widely applied in industry and the DCC is a standard method to model conditional correlations, we are going to apply these two models (which have been discussed in Chapter 1) to estimate the conditional correlations, also because of their superior performance out-of-sample. Figure 2.4.1 and 2.4.2 plot the absorption ratios extracted from ten sector indices for $k = 1, 2, 3, 4$. Compared to the absorption ratio modelled by the EWMA model, the DCC model provides a less noisy ratio, but also less sensitive to market fragility. Regarding absorption ratios with different values of $k$, although higher $k$ are expected to explain a large proportion of the total variation, and lower $k$ acts more sensitive than higher $k$, absorption ratios show roughly same dynamic patterns for all $k$.

During the period of the dot-com bubble, we can see that the absorption ratio starts from a relatively high level, and rebounds a bit around end of 1998, which indicates the systemic risk in the market is quite high. This feature could be a result of the Asian Crisis and its posterior effect. We also observe that the dot-com bubble burst in the end of 2000. Investigating the second data set, we find absorption ratios climbed to relatively high levels in three sub-periods: April 2007 for the subprime mortgage crisis, September 2008 for the bankruptcy of Lehman Brother, and the beginning of 2010 for the worries about European debt crisis. Therefore, the analysis indicates that the high level of absorption ratio implies financial fragility and is followed by the explosive of the systemic risk.

Figure 2.4.1: The Absorption Ratio during the Dot-com Bubble

These two figures plot the absorption ratio $AR_k$ for $k = 1, 2, 3, 4$ during the dot-com bubble, The left sub-figure shows $AR_k$ extracted from conditional correlation matrix modelled by EWMA model, and the right sub-figure shows $AR_k$ modelled by DCC model.

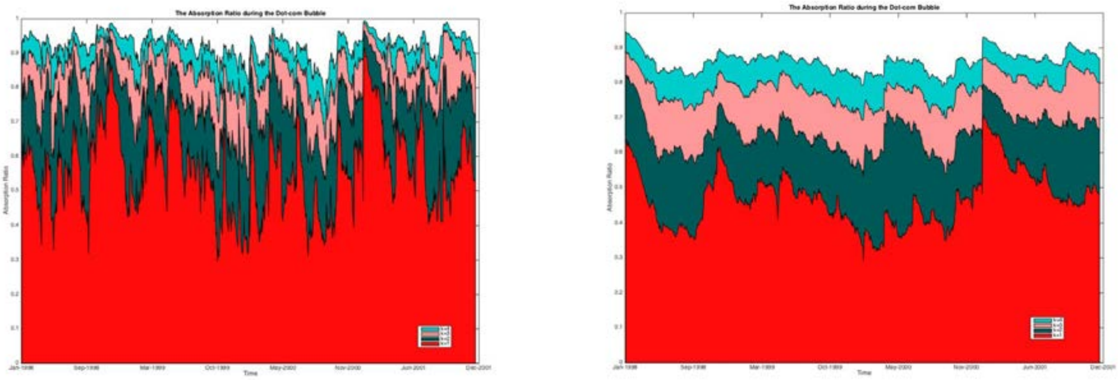
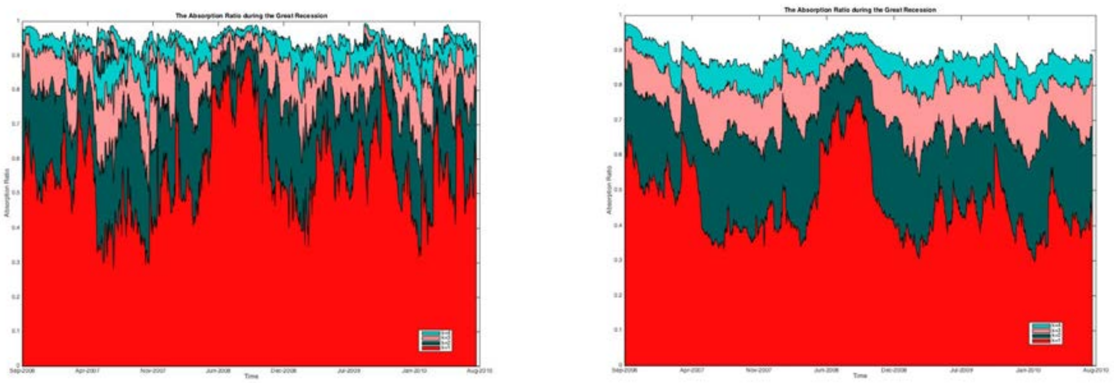


Figure 2.4.2: The Absorption Ratio during the Great Recession

These two figures plot the absorption ratio $AR_k$ for $k = 1, 2, 3, 4$ during the great recession, The left sub-figure shows $AR_k$ extracted from conditional correlation matrix modelled by EWMA model, and the right sub-figure shows $AR_k$ modelled by DCC model.

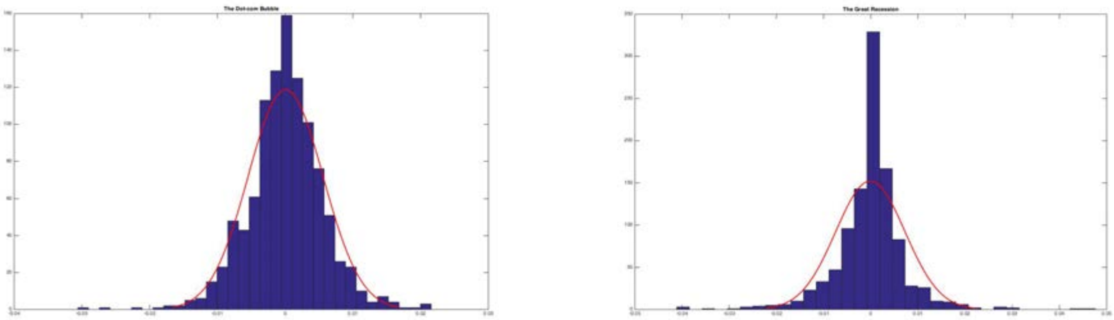


### 2.4.2 A Quantile Regression Analysis

Since we have discussed potential linkages between the absorption ratio and the financial fragility, it is necessary to study how the absorption ratio statistically influences market returns. Considering that the absorption ratio is non-stationary, we take the first order difference which $\Delta AR_t$ interpret the daily growth of market connectedness. A simple linear regression model might be used, to this end, possibly including other control variables avoid missing variables and to strengthen the explanatory power of the model. Beside using the absorption ratio as explanatory variables, we control the market interest rate $IR_t$, market risk attitude factor - $Vol_t$ and liquidity factor $Ted_t$. The stationary and ADF tests suggest that $IR_t$ and $Ted_t$ are non-stationary, thus the daily growth of federal funds rate $\Delta IR_t$ and growth of Ted $\Delta Ted_t$ spread would be used in regressions. In order to have a more detailed picture on the distribution, we apply the Quantile Regression (QR hereafter) methods. The QR provides an approach to regress $\tau$ percentages from the distribution of the market return on the absorption ratio and other control variables, we denote $\tau$ as the quantile index for $0 \leq \tau \leq 1$, as shown in Equation (2.4.4).

Figure 2.4.3 plots the fitted distribution of market returns in two different periods. The plotted distributions also display that the great recession suffer from more severe systemic risk than the dot-com bubble, because the kurtosis in the right sub-figure is much higher so that with fatter-tails compared with the left sub-figure.

Figure 2.4.3: The Fitted Distributions of Market Returns




Because we are aiming to test the pre-warming mechanism of absorption ratio, we then use the growth of absorption ratio as the explanatory variable and use quantile regression coefficient $\hat{\beta}_\tau$ to interpret the quantile behavior of market returns. Thus, the QR model is formulated as,

$$r_t = \beta_0 + \beta_{1,\tau}\Delta IR_t + \beta_{2,\tau}Vol_t + \beta_{3,\tau}\Delta Ted_t + \beta_{4,\tau}\Delta AR_{4,t} + e_t \tag{2.4.4}$$

where $r_t$ is the log market return, and $\beta_\tau$ is an unknown parameter associated with the $\tau^{th}$ quantile. We choose $k = 4$ for absorption ratio because it contains more information about the correlation comovement. Similarly with ordinary least squares method, the quantile regression minimises

$$\sum_{t=1}^{T} \tau\,|e_t| + \sum_{t=1}^{T}(1-\tau)\,|e_t|$$

to estimate the quantile coefficient $\beta_\tau$, where the front part $\tau\,|e_t|$ and the rear part $(1-\tau)\,|e_t|$ are giving asymmetric penalties for under and over predictions, respectively.

Following Fitzenberger et al. (2013), the log return $r$ has a probability distribution function $F(q)$ such that,

$$F(q) = Prob(r \leq q)$$

The $\tau$ quantile of $r_t$ can be defined as the inverse function of $F(\cdot)$

$$\mathcal{Q}_r(\tau) = inf\{q : F(q) \geq \tau\}$$

For any sample of $r = \{r_1, r_2, \ldots, r_n\}$, the $\tau$ sample quantile $\xi\tau$ (analogue of $\mathcal{Q}(\tau)$), can be found through,

$$\underset{\xi \in \mathbb{R}}{min} \sum_i = 1^n \rho_\tau(y_i - \xi(\tau))$$

where $\rho_\tau(z) = z(\tau - I)$, and $I$ is an indicator function equals to one in the case of $z < 0$. Hence, the estimators of the coefficients of quantile regression are:

$$\hat{\mu} = argmin_{\mu \in \mathbb{R}} \sum (r_t - \mu)^2$$

$$\hat{\boldsymbol{\beta}}_\tau = argmin_{\beta \in \mathbb{R}} \sum \rho_\tau(r_t - \mathbf{X}_t^\top \boldsymbol{\beta})^2$$

where $\mathbf{X} = [1, \Delta IR_t, Vol_t, \Delta Ted_t, \Delta AR_{k,t}]^\top$ and $\boldsymbol{\beta} = [\beta_0, \beta_{1,\tau}, \beta_{2,\tau}, \beta_{3,\tau}, \beta_{4,\tau}]^\top$. We use 'quantreg' package in R (Koenker, 2013) to estimate the quantile coefficients. Table 2.4.1 and Table 2.4.2 document the results of quantile regression base on two data sets.

We first investigate the Dot-com Bubble period. According to Table 2.4.1, the simple ordinary least square estimation shows that regardless using the EWMA or DCC method to compute the absorption ratio, control variables volatility index and the growth of term spread are significant but the growth of federal funds rate and absorption ratio do not significantly explain the market return. The results in quantile regression become more interesting. Through the quantile $\tau = 0.05$ to $\tau = 0.2$, the growth of absorption ratios significantly show negative effect on market return, by that means, the more growth on absorption ratio, the less market return could be, which verifies that the high strength of market connections could be a factor of bearish market return. This result is robust by using either the EWMA or

DCC models. However, along with the quantile $\tau$ increasing, the absorption ratio becomes insignificant and the sign changes to positive. Also in lower quantiles, the growth of federal funds rate and term spread positively account for the market rate, and the volatility index show highly positive effects on the market return.

Figure 2.4.1 plots the quantile regression estimations through $\tau = 0.01$ to $\tau = 0.99$ and show the similar results. The sub-plots with quantile variables $\Delta AR\_ewma$ and $\Delta AR\_dcc$ exhibit that the absorption ratio has significant negative effect to the market return if the quantile $\tau$ is small enough.

Table 2.4.1: Estimated Regression Coefficients (The Dot-com Bubble)

This table displays the OLS and the quantile regression estimation results for the sample of the Dot-com Bubble. $***$, $**$ and $*$ indicate the significant level at 99%, 95% and 90% significance levels, respectively.

| | EWMA | | | | | DCC | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Intercept | $\Delta IR_t$ | $Vol_t$ | $\Delta Ted_t$ | $\Delta AR_t(k=4)$ | Intercept | $\Delta IR_t$ | $Vol_t$ | $\Delta Ted_t$ | $\Delta AR_t(k=4)$ |
| OLS | 0.0001 | -0.0015 | $0.9721***$ | $-0.0046*$ | 0.0006 | 0.0001 | -0.0015 | $0.9736***$ | $-0.0044*$ | 0.0378 |
| | (0.0001) | (0.0025) | (0.0349) | (0.0023) | (0.0119) | (0.0001) | (0.0025) | (0.0349) | (0.0023) | (0.0304) |
| Quantile: 0.05 | $-0.0064***$ | $0.0078**$ | $0.9046***$ | $-0.0102*$ | $-0.04853***$ | $-0.0065***$ | $0.0086**$ | $0.8655***$ | $-0.0107**$ | $-0.1608***$ |
| | (0.0003) | (0.0026) | (0.0856) | (0.0058) | (0.0177) | (0.0003) | (0.0039) | (0.0722) | (0.0054) | (0.0450) |
| Quantile: 0.10 | $-0.0053***$ | $0.0095**$ | $0.9234***$ | $-0.0089**$ | $-0.0669***$ | $-0.0052***$ | $0.0097**$ | $0.9251***$ | $-0.0076*$ | $-0.1371**$ |
| | (0.0002) | (0.0046) | (0.0558) | (0.0026) | (0.0189) | (0.0002) | (0.0048) | (0.0596) | (0.0041) | (0.0632) |
| Quantile: 0.20 | $-0.0032***$ | $0.0090**$ | $0.9989***$ | -0.0042 | $-0.0275*$ | $-0.0032***$ | $0.0086***$ | 0.9981 | -0.0054 | -0.0656 |
| | (0.0002) | (0.0011) | (0.0434) | (0.0031) | (0.0150) | (0.0002) | (0.0009) | (0.0539) | (0.0034) | (0.0588) |
| Quantile: 0.30 | $-0.0017***$ | $0.0115**$ | $1.0081***$ | -0.0037 | -0.0001 | $-0.0017***$ | $0.0115***$ | $1.0076***$ | -0.0038 | -0.0348 |
| | (0.0001) | (0.0029) | (0.0486) | (0.0032) | (0.0173) | (0.0002) | (0.0026) | (0.0464) | (0.0031) | 0.0492 |
| Quantile: 0.40 | $-0.0006***$ | $0.0070**$ | $1.0338***$ | -0.0037 | -0.0028 | $-0.0006***$ | $0.0071**$ | $1.0342***$ | -0.0034 | -0.0108 |
| | (0.0001) | (0.0012) | (0.0345) | 0.0024 | (0.0119) | (0.0001) | (0.0034) | (0.0381) | (0.0026) | (0.0249) |
| Quantile: 0.50 | 0.0001 | $0.0046***$ | $1.0239***$ | -0.0030 | 0.0042 | 0.0001 | $0.0046***$ | $1.0206***$ | -0.0029 | 0.0316 |
| | (0.0001) | (0.0015) | (0.0299) | (0.0020) | (0.0111) | (0.0001) | (0.0017) | (0.0298) | (0.0021) | (0.0414) |

Figure 2.4.4: The Quantile Regression Coefficients during the Dot-com Bubble

These two figures plots quantile regression coefficients for $\tau \in [0.01, 0.99]$, the upper sub-figure is obtained from the EWMA absorption ratio, and the lower sub-figure plots results obtained from the DCC absorption ratio. The red solid line is the OLS coefficient and its 95% confidence interval is showed as red dot line. The black dot line is the quantile coefficients, and grey belt is its confidence intervals.

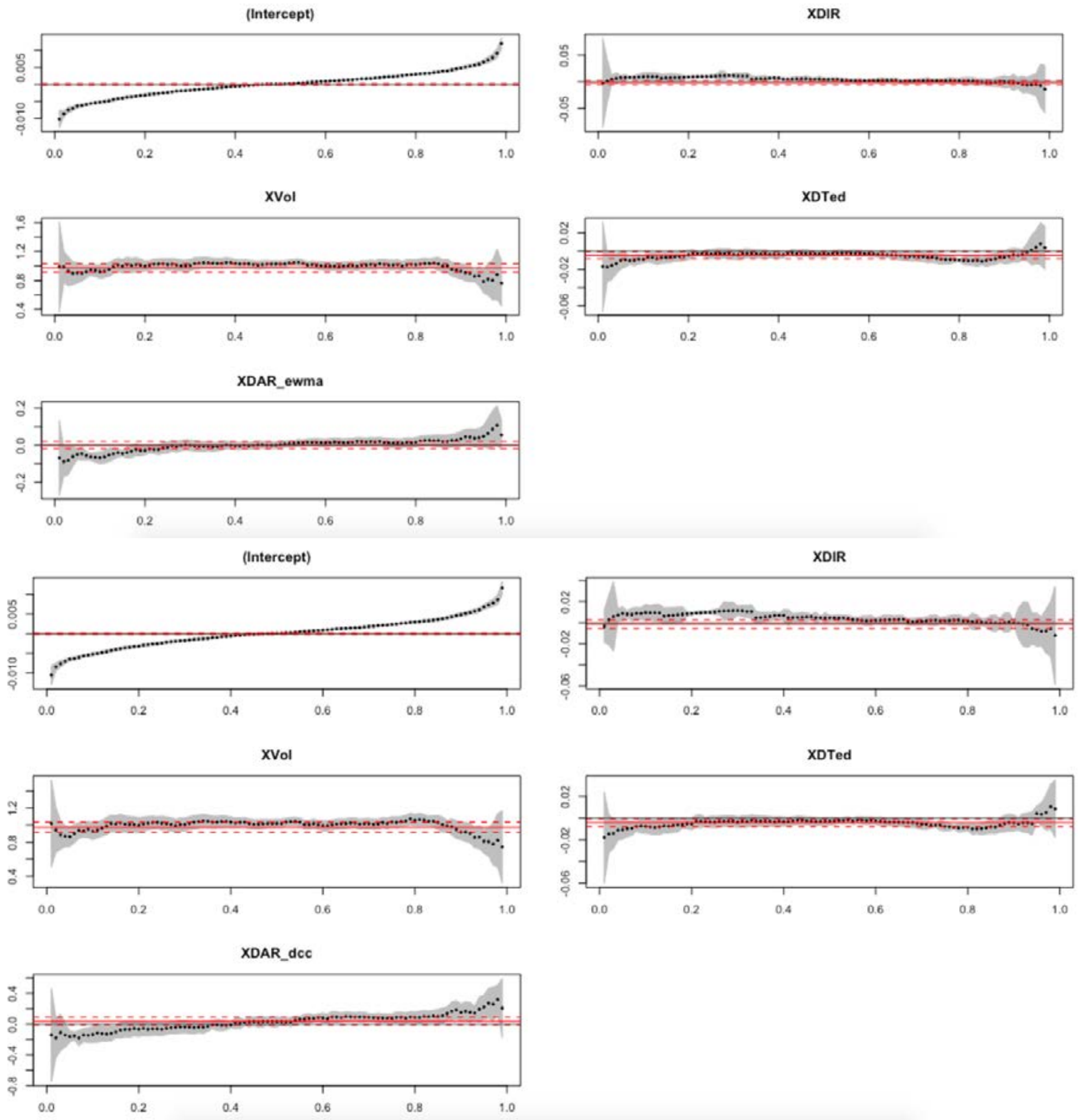


Table 2.4.2: Estimated Regression Coefficients (The Great Recession)

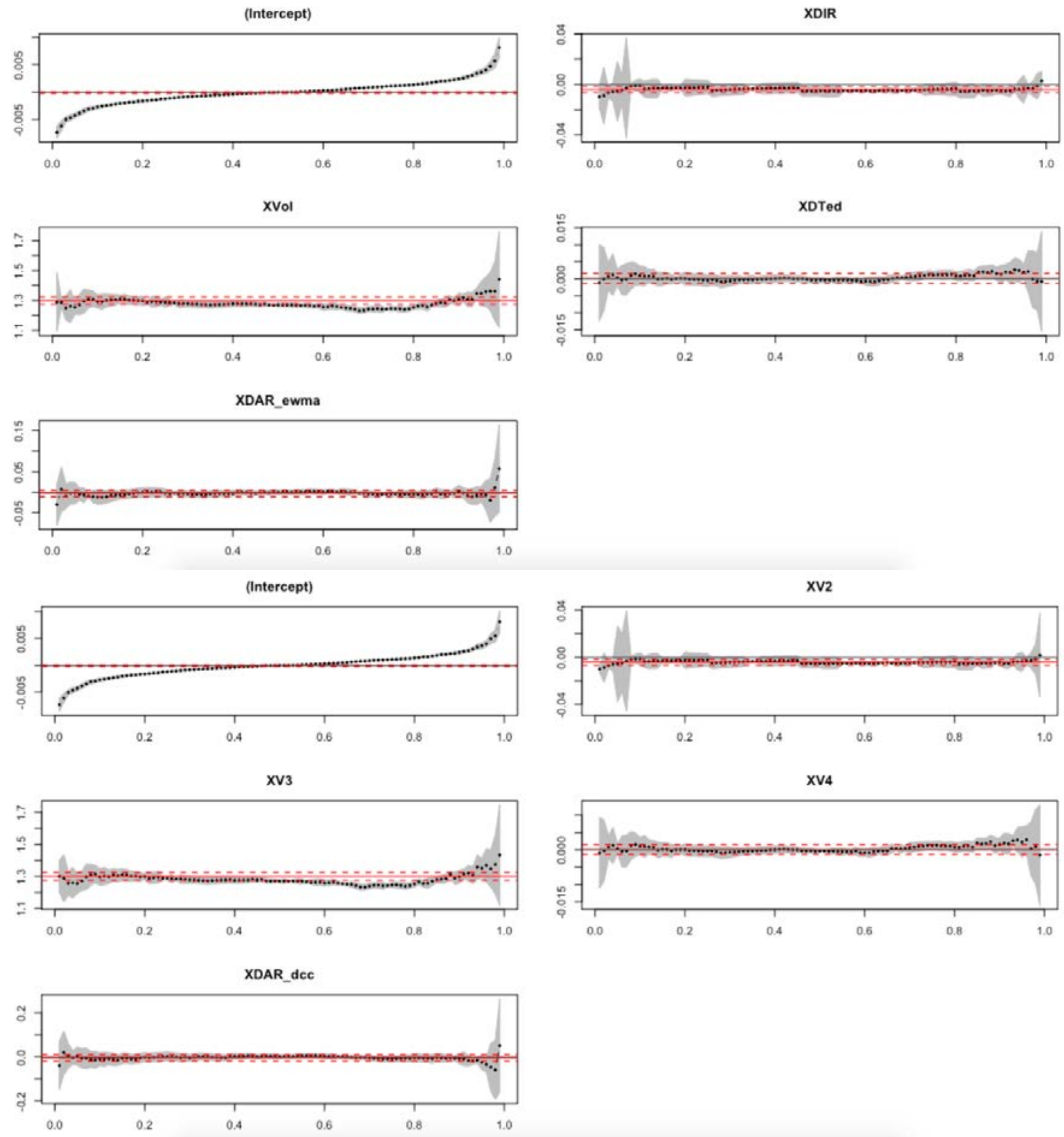
This table displays the OLS and the quantile regression estimation results for the sample of the Great Recession. $***$, $**$ and $*$ indicate the significant level at 99%, 95% and 90% significance levels, respectively.

| | EWMA | | | | | DCC | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Intercept | $\Delta IR_t$ | $Vol_t$ | $\Delta Ted_t$ | $\Delta AR_t(k=4)$ | Intercept | $\Delta IR_t$ | $Vol_t$ | $\Delta Ted_t$ | $\Delta AR_t(k=4)$ |
| OLS | -0.0001 | $-0.0004***$ | $1.300***$ | 0.0001 | -0.0028 | -0.0001 | $-0.0004***$ | $1.300***$ | 0.0001 | -0.0042 |
| | (0.0001) | (0.0015) | (0.0153) | (0.0008) | (0.0049) | (0.0001) | (0.0002) | (0.0153) | (0.0008) | (0.0093) |
| Quantile: 0.05 | $-0.0042***$ | -0.0055 | $1.2545***$ | 0.0004 | $-0.0021*$ | $-0.0043***$ | -0.0056 | $1.2551***$ | 0.0003 | -0.0031 |
| | (0.0002) | (0.0144) | (0.0289) | (0.0014) | (0.0012) | (0.0003) | (0.0196) | (0.0371) | (0.0009) | (0.0024) |
| Quantile: 0.10 | $-0.0027***$ | -0.0009 | $1.2940***$ | 0.0009 | -0.0116 | $-0.0027***$ | -0.0014 | $1.2954***$ | 0.0011 | -0.0101 |
| | (0.0002) | (0.0032) | (0.0278) | (0.0017) | (0.0105) | (0.0002) | (0.0032) | (0.0321) | (0.0015) | (0.0181) |
| Quantile: 0.20 | $-0.0016***$ | -0.0025 | $1.2981***$ | 0.0001 | -0.0007 | $-0.0016***$ | -0.0024 | $1.2987***$ | 0.0001 | -0.0016 |
| | (0.0001) | (0.0041) | (0.0187) | (0.0011) | (0.0071) | (0.0001) | (0.0043) | (0.0115) | (0.0010) | (0.0129) |
| Quantile: 0.30 | $-0.0008***$ | -0.0041 | $1.2781***$ | -0.0006 | -0.0037 | $-0.0008***$ | -0.0041 | $1.2781***$ | -0.0007 | -0.0042 |
| | (0.0001) | (0.0027) | (0.0122) | (0.0009) | (0.0042) | (0.0001) | (0.0027) | (0.0141) | (0.0009) | (0.0102) |
| Quantile: 0.40 | $-0.0003***$ | -0.0031 | $1.2783***$ | 0.0000 | -0.0028 | $-0.0004***$ | -0.0032 | $1.2748***$ | 0.0001 | 0.0037 |
| | (0.0001) | (0.0025) | (0.0113) | (0.0006) | (0.0042) | (0.0001) | (0.0025) | (0.0123) | (0.0007) | (0.0082) |
| Quantile: 0.50 | 0.0001 | -0.0052 | $1.2706***$ | -0.0004 | 0.0001 | -0.0001 | -0.0052 | $1.2692***$ | -0.0004 | 0.0023 |
| | (0.0001) | (0.0025) | (0.0085) | (0.0005) | (0.0001) | (0.0001) | (0.0023) | (0.0053) | (0.0001) | (0.0058) |

We then study the case of the Great Recession. The OLS estimation in Table 2.4.2 reports that the variables of volatility index and the growth of federal funds rate significantly explain the market return in the second sample. However, the patterns are quite different with the dot-com bubble, where the volatility index become dramatically significant, and other control variables and the growth of absorption ratio nearly insignificantly account to the market return.

Figure 2.4.5: The Quantile Regression Coefficients during the Great Recession

These two figures plots quantile regression coefficients for $\tau \in [0.01, 0.99]$, the upper sub-figure is obtained from the EWMA absorption ratio, and the lower sub-figure plots results obtained from the DCC absorption ratio. The red solid line is the OLS coefficient and its 95% confidence interval is showed as red dot line. The black dot line is the quantile coefficients, and grey belt is its confidence interval.




Although the sign of the coefficients are negative for $\tau = 0.05 - 0.3$, the growth of absorption ratio based on the EWMA model only significantly interprets the market return at $\tau = 0.05$

quantile at 10% significance level. Compared with the bubble period, the absorption ratio became less significant, and this is probably because the systemic risk during the Great Recession is more severe so that the control variable - the volatility index explains most of variations of the market return. The explanatory ability of the growth of absorption ratio on the market return can be seen as an compensation in down-side market. Similarly with the previous case, along with the quantile $\tau$ increasing, the explanatory ability of absorption ratio is decreasing and becoming insignificant. Nevertheless, it is sensible that the absorption ratio displays reasonable negative signs in left tailed quantile regressions ($\tau = 0.05 - 0.3$) - the higher absorption ratio during down-side market, the lower value of the market return. Another fact is compared with the absorption ratio computed through the EWMA model, the one from the DCC model seems to be more conservative, which is consistent with previous discussions.

Figure 2.4.5 also considers the right tail of the distribution, i.e. $\tau > 0.5$, but explanatory ability of the absorption ratio on up-ward market is not crucial to study the systemic risk. The rest of features in sub-figures are consistent with Table 2.4.2.

Therefore, as a naturally second moment risk factor, the absorption ratio plays compensative role of volatility during the exposure of systemic risk. It is fair to say that the results from the quantile regression model reveal more detailed inter-linkages between the absorption ratio and the down-side market return. Considering severe systemic risks always starts from a bearish market indication, we suggest that the financial fragility indicator, absorption ratio, might be useful for pre-warning systemic risk.

### 2.4.3 A Change Point Analysis

The absorption ratio is an indicator of the financial fragility, as it can reflect the movement of market return in the bear market. Thus, detecting changes in the mean of the absorption ratio would give a sense of changing of market states. If changes in absorption ratio are consistently found prior to changes happening in the market state, then it can empirically used as an early warning indicator for systemic risk. Following Billio et al. (2012) who use the proportion explained by principal components to identify market status as downward and upward, we apply a more statistical test to detect the mean change of the absorption ratio.

We apply a non-parametric CUSUM test, which is designed for independent identically distributed processes (Csörgő and Horváth, 1997), and further extended to mixing processes by Ling (2007). The CUSUM test constructs a statistics $\Lambda_T$ to distinguish the following null hypothesis

$$H_0\colon\ \mu_1 = \mu_2 = \cdots = \mu_T$$

against the alternative

$$H_1\colon\ \mu_1 = \mu_2 = \cdots = \mu_{t^*} \neq \mu_{t^*+1} = \cdots = \mu_T$$

where $t^*$ is an unknown change point for $1 < t^* < T$. $\mu_t$ is the conditional mean of any stationary process, and $T$ is the sample size.

To detect break point $t^*$ in any random variable $y_t$, we need to demean variable $y_t$, denoting as $\tilde{y}_t = y_t - \frac{1}{T}\sum_1^T y_t$. Then, we have statistics $M_n(t)$ as

$$\left.\begin{aligned} &M_n(k) = n^{-\frac{1}{2}}(S((T+1)\tfrac{k}{T}) - \tfrac{k}{T}S(T)) \quad 0 < k < T \\ &M_n(k) = 0 \quad k = 0 \end{aligned}\right\} \tag{2.4.5}$$

where $S(k) = \sum_{t=1}^{k} \tilde{y}_t$. The limit distribution is given by,

$$\Lambda_T = Sup_{1<t<T}\frac{1}{\sigma}\left|M_n(t)\right| \xrightarrow{D} Sup_{1<t<T}\left|B(t)\right| \tag{2.4.6}$$

where $B(t)$ is the standard Brownian Bridge. The asymptotic critical values can be found from the distribution of Brownian Bridge, refer to Csörgő and Horváth (1997).

As it is designed to test stability versus a single change, we apply the CUSUM test to absorption ratio processes $AR_{k,t}$ for detecting the biggest change point in its mean. The idea is that because the data set spans over the whole crisis period, we see the biggest change in absorption ratio as a start date of systemic risk exposure, and thereby the detection should give implications for policy makers and investors. Thus, we specifically test the hypotheses that,

$$H_0\colon\ AR_{1,1} = AR_{1,2} = \cdots = AR_{1,T}$$

against the alternative

$$H_1\colon\ AR_{1,1} = AR_{1,2} = \cdots = AR_{1,t^*} \neq AR_{1,t^*+1} = \cdots = AR_{1,T}$$

Here, we choose $k = 1$ because lower $k$ generates a more sensitive absorption ratio, Table 2.4.3 reports the break point in absorption ratio $AR_{1,t}$. As we studied above, for each financial crisis, we apply two types of multivariate covariance models –EWMA and DCC for absorption ratios. The statistics $\Lambda_T$ rejects the null hypothesis in every processes, and CUSUM statistics $M_n(t)$ are plotted in Figure 2.4.6. The dates with maximum statistics can be referred as break dates in absorption ratio during the dot-com bubble and the great recession.

Regarding to the dot-com bubble, both the EWMA and DCC absorption ratios experience a mean change on 05-December-2000. This date suggests that the U.S. equity market was highly concentrated, and a small shock in any sector could contaminate to the whole financial system. While during the great recession, the CUSUM test detects inconsistent change locations in EWMA and DCC absorption ratios. The detected change point in the EWMA absorption ratio is on 01-April-2008, which is prior to the break out of the recession in

September, but the counterpart for DCC absorption ratio is on 31-October-2008. This is probably because the absorption ratio in the DCC model are less sensitive to market information, so that the mean change is more likely to be detected after a sizeable structural change in the process. And this also causes the fact that supreme of cumulative sum statistics $\Lambda$ in the DCC absorption ratio is larger than $\Lambda$ in the EWMA absorption ratio.

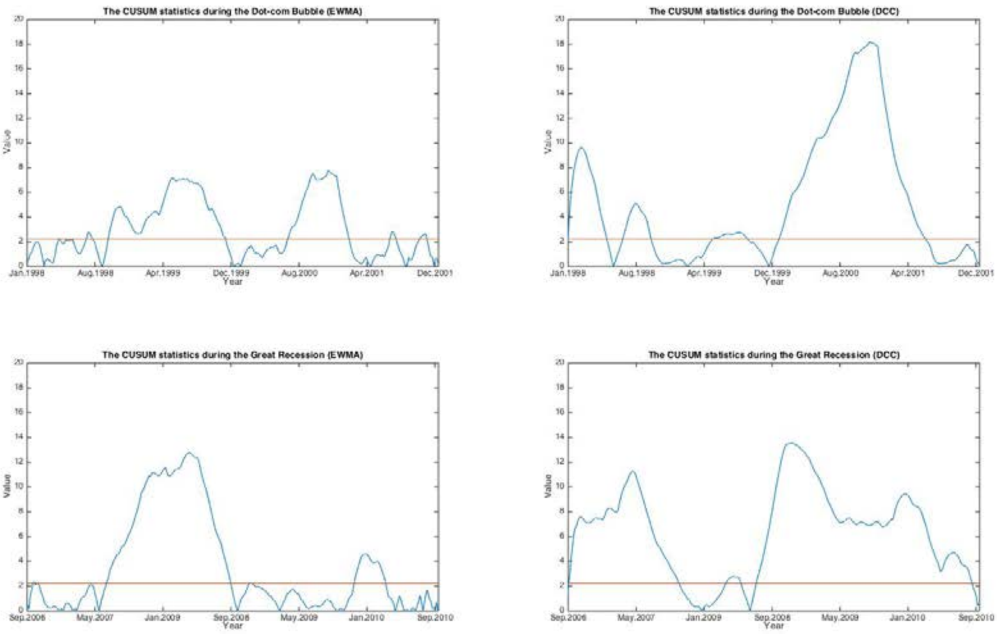
Therefore, as an early warning indicator, the EWMA absorption ratio may indicate more implications for policy making and real trading.

Table 2.4.3: The Change Point Detections on $AR_{1,t}$

| | The Dot-com Bubble | | The Great Recession | |
|---|---|---|---|---|
| | EWMA | DCC | EWMA | DCC |
| Change Date | 05/Dec/2000 | 05/Dec/2000 | 01/Apr/2008 | 31/Oct/2008 |
| $\Lambda_T$ | 7.80*** | 18.20*** | 12.80*** | 13.55*** |

Figure 2.4.6: The CUSUM statistics $M_n(t)$

The blue line is the $M_n(t)$ statistics, and the red line is the critical value under 95% significance level.



## 2.5 Forecasting Absorption Ratio

As discussed in the last section, the absorption ratio can be applied as an early warning indicator of systemic risk, therefore it becomes important to predict absorption ratios for early warning the systemic risk in financial markets. This section provides methods to forecast the absorption ratio. Recall that the absorption ratio $AR_t$ is extracted from the conditional covariance matrix $\hat{\mathbf{\Sigma}}_t$, we then use M-GARCH models to forecast conditional covariance $\hat{\mathbf{\Sigma}}_{t+1}$, and compute the predicted absorption ratio $\hat{AR}_{t+1}$.

$$\hat{AR}_{t+1} = f(\hat{\mathbf{\Sigma}}_{t+1})$$

where $f(\cdot)$ is a function described as (2.4.3) in section 2.4.

Among various types of M-GARCH models, we select two typical models considering the efficiency and accuracy: the EWMA and DCC model. Chapter 1 concluded that the DCC model produces less forecast error than the EWMA, but EWMA is the most efficient method to forecast $\hat{\mathbf{\Sigma}}_{t+1}$. Concerning the accuracy of forecasting, we apply iterated one step a head forecast method with fixed rolling window.

### 2.5.1 Forecasting Models

#### 2.5.1.1 The EWMA Model

The first forecasting model is Exponential Weighted Moving Averaging model (Riskmetrics, 1996). The EWMA model is efficient in computation because it is not necessary to estimate parameters, and the only parameter in the model is a fixed value $\lambda_0 = 0.94$ for daily data, which is derived from numerous empirical works (Riskmetrics, 1996). The model shows as,

$$\hat{\mathbf{\Sigma}}_{t+1} = \sum_{l=0}^{p} \lambda_0 \hat{\mathbf{\Sigma}}_{t-l} + \sum_{l=0}^{q} (1-\lambda_0) \boldsymbol{r}_{t-l} \boldsymbol{r}_{t-l}^{\top} \tag{2.5.1}$$

where $p$ and $q$ is the lag length of the GARCH and ARCH effect. We set the initial value of $\hat{\mathbf{\Sigma}}_t$ as the unconditional covariance $\bar{\mathbf{\Sigma}}_t = \frac{1}{T}\sum_{t=1}^{T}\mathbf{r}_t\mathbf{r}_t^\top$. Thus, the out-of-sample can be predicted based on conditional covariance $\hat{\mathbf{\Sigma}}_{t-l}$, where $\hat{\mathbf{\Sigma}}_{t-l}$ can be obtained from in-sample data.

#### 2.5.1.2 The Dynamic Conditional Correlation Model

The second forecasting model is Dynamic Conditional Correlation model (Engle, 2002). Since the conditional covariance matrix can be decomposed to diagonal conditional standard deviation matrix $\mathbf{D}_t$ and conditional correlation matrix $\mathbf{R}_t$, shown as,

$$\mathbf{\Sigma}_t = \mathbf{D}_t \cdot \mathbf{R}_t \cdot \mathbf{D}_t$$

Thus, the innovation term in conditional correlation $\mathbf{R}_t$ would be de-volatized process $\hat{\epsilon}_t$ obtained from Equation 2.5.2.

$$\hat{r}_t^*(i) = r_t(i)/\sqrt{\hat{\mathbf{\Sigma}}_{i,i,t}} \tag{2.5.2}$$

where $\mathbf{\Sigma}_{i,i,t}$ is the $i$th diagonal element of $\mathbf{D}_t \cdot \mathbf{D}_t$, indicating the conditional variance of $r_t(i)$ ($r_t(i) \in \mathbf{r}_t$). We can use GARCH (p,q) model to estimate,

$$\mathbf{\Sigma}_{i,i,t} = w_i + \sum_{l=1}^{q} a_{i,l}\mathbf{r}_{t-l}^2 + \sum_{l=1}^{p} b_{i,l}\mathbf{\Sigma}_{i,i,t-l}$$

The GARCH (p,q) model can be estimated by the Maximum Likelihood. Then, we set the initial value of $\mathbf{\Sigma}_{i,i,t}$ as unconditional variance $\bar{\mathbf{\Sigma}}_{i,i,t} = \frac{1}{T_1}\sum_{t=1}^{T_1} r_{i,t}^2$, where $T_1$ is the in-sample length. Combined with estimated coefficients, we then estimate the conditional variance $\hat{\mathbf{\Sigma}}_{i,i,t}$.

For the conditional correlation process, assuming there exists a quasi conditional correlation process $\mathbf{Q}_t$.

$$\mathbf{Q}_t = \bar{\mathbf{R}} + \sum_{l=1}^{q} \alpha_l (\hat{\mathbf{r}}_{t-l}^* (\hat{\mathbf{r}}_{t-l}^*)^\top - \bar{\mathbf{R}}) + \sum_{l=1}^{p} \beta_l (\mathbf{Q}_{t-l} - \bar{\mathbf{R}})$$

where unconditional correlation matrix $\bar{R} = \frac{1}{T_1} \sum_{t=1}^{T_1} \hat{\mathbf{r}}_t^* (\hat{\mathbf{r}}_t^*)^\top$. The conditional correlation of entries $i$ and $j$ are $\rho_{i,j,t} = \frac{\mathbf{Q}_{i,j,t}}{\sqrt{\mathbf{Q}_{i,i,t}}\sqrt{\mathbf{Q}_{j,j,t}}}$. Hence, the conditional correlation matrix $R_t$ is,

$$\mathbf{R}_t = diag(\mathbf{Q}_t)^{-\frac{1}{2}} \cdot \mathbf{Q}_t \cdot diag(\mathbf{Q}_t)^{-\frac{1}{2}}$$

Engle (2002) showed that the QMLE consistently estimate the parameters in the DCC model. With estimated parameters, using the unconditional correlation matrix $\bar{\mathbf{R}}_t$ as the initial value of $\mathbf{Q}_t$, we obtain estimated conditional correlation $\hat{\mathbf{Q}}_t$.

Using the estimation from the in-sample, we then iteratively forecast one-step ahead predictors on $\hat{\mathbf{R}}_{t+1}$, and $\mathbf{D}_{t+1}$ is predicted through following Equations:

$$\hat{\mathbf{Q}}_{t+1} = \bar{\mathbf{R}} + \sum_{l=0}^{q} \hat{\alpha}_l (\hat{\mathbf{r}}_{t-l}^* (\hat{\mathbf{r}}_{t-l}^*)^\top - \bar{\mathbf{R}}) + \sum_{l=0}^{p} \beta_l (\hat{\mathbf{Q}}_{t-l} - \bar{\mathbf{R}}) \tag{2.5.3}$$

$$\mathbf{R}_{t+1} = diag(\mathbf{Q}_{t+1})^{-\frac{1}{2}} \cdot \mathbf{Q}_{t+1} \cdot diag(\mathbf{Q}_{t+1})^{-\frac{1}{2}} \tag{2.5.4}$$

$$\hat{\mathbf{\Sigma}}_{i,i,t+1} = \hat{w}_i + \sum_{l=0}^{q} \hat{a}_{i,l} r_{t-l}^2(i) + \sum_{l=0}^{p} \hat{b}_{i,l} \hat{\mathbf{\Sigma}}_{i,i,t-l} \tag{2.5.5}$$

where $\hat{\mathbf{\Sigma}}_{i,i,t+1}$ is the $i$th diagonal element in $\mathbf{D}_{t+1}$. Therefore, the we can forecast conditional covariance matrix $\hat{\mathbf{\Sigma}}_{t+1}$ as,

$$\hat{\mathbf{\Sigma}}_{t+1} = \hat{\mathbf{D}}_{t+1} \cdot \hat{\mathbf{R}}_{t+1} \cdot \hat{\mathbf{D}}_{t+1} \tag{2.5.6}$$

The DCC forecasting model is supposed to be more precise than the EWMA, but it has higher computational costs.

## 2.5.2 The Prediction Results

In order to predict the absorption ratio during the 'dot-com' bubble and the great recession, we split samples into in-sample and out-of-sample. We roll the fixed in-sample window with new observations at $t$, and iterated forecast $t+1$. Each data set contains 1042 observations, we set 522 observations (roughly two years) as in-sample length, and treat the rest 520 observations as out-of-sample. Specifically, in the first data set, we estimate models based on data from 01-January-1998 to 03-January-2000, and forecast the covariance matrix and the absorption ratio between 04-January-2000 and 31-December-2001; in the second data set, the models are estimated through the sample from 01-September-2006 to 02-September-2008, and forecast the absorption ratio between 03-September-2008 and 31-August-2010. The predicted absorption ratios are plotted in Figure 2.5.1 and 2.5.2. From figures, we notice that the features of absorption ratios are consistent with the discussion on Figure 2.4.1 and 2.4.2. If there was a sharp boost in absorption ratio, we identify it gives an early warning to the systemic risk.

Figure 2.5.1: The Predicted Absorption Ratio for the Dot-com Bubble

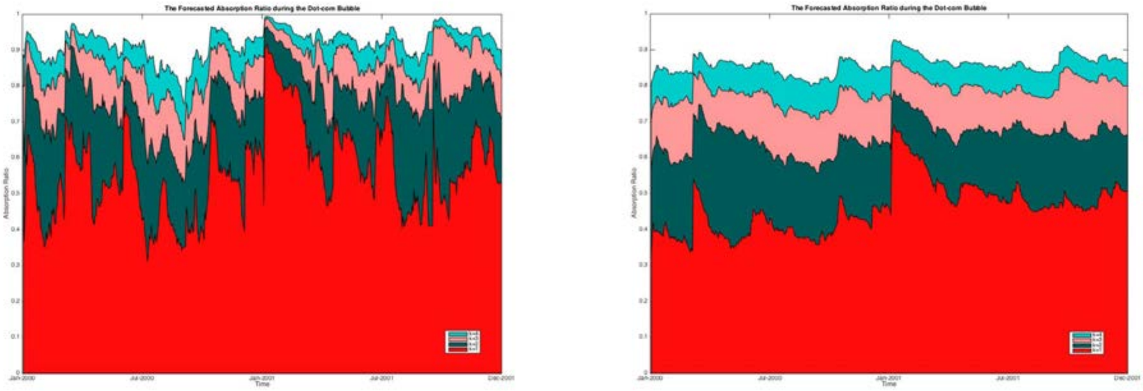


Figure 2.5.2: The Predicted Absorption Ratio for the Great Recession

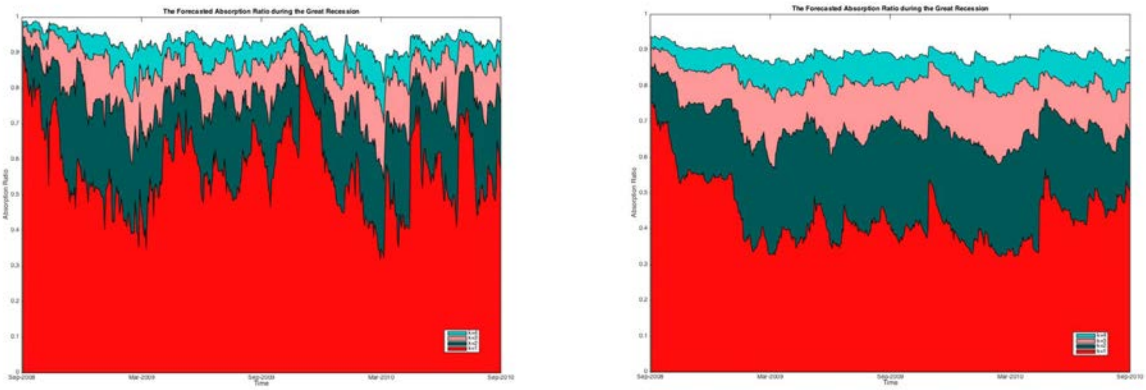


Table 2.5.1 reports the forecast errors between the EWMA and DCC models. Not surprisingly, the DCC model provides relatively more accurate forecasting in terms of mean square forecast error (MSFE) and mean absolute forecast error (MAFE). Meanwhile, the EWMA model produces acceptable forecasts as well. Combined with its superiorities in interpreting down-side market returns and forecasting efficiency, we would suggest to apply the EWMA model to predict absorption ratio for financial fragility, to give more implications in preventing the systemic risk.

Table 2.5.1: The Forecast Error

| | The Dot-com Bubble | | The Great Recession | |
|---|---|---|---|---|
| | EWMA | DCC | EWMA | DCC |
| MSFE | 0.19 | 0.16 | 0.58 | 0.52 |
| MAFE | 0.37 | 0.36 | 0.75 | 0.71 |

## 2.6 Summary

This chapter provided an application of analysing the financial fragility in the U.S. equity market. We empirically showed the correlation concentration level can be used as an early warning indicator of the financial fragility, even for systemic risk, through analysing the absorption ratio by the quantile regression and the change point analysis (also supported by existed literature Kirtzman et al. (2011) and Billio et al. (2012)). Based on predicted conditional covariance, we computed the predicted absorption ratio, to indicate further implementations to policy makers and investors. Detecting the instability in the absorption ratio is a rough proxy for detecting the instability in the covariance structure. More methods about testing for the instability in the covariance structure will be developed in next two chapters.

# Chapter 3

# Change Point Detection in The Conditional Correlation Structure of Multivariate Volatility Models

## 3.1 Introduction

Multivariate models, whether for fitting first moment or second moments, are being more and more widely used in financial applications. As a remarkable application, testing for contagion in global financial markets has been widely re-visited after the recent great recession (See Chiang et al. 2007; Celık 2012; Wang and Nguyen Thi 2013; amongst others). Forbes and Rigobon (2002) defined a contagion effect as the correlation structure of several markets experiences an increase after a certain date. To identify contagion effect, majority of literature consider some financial events as change dates for correlation structure, but this assumption contains biases because the occurrence of change may happen earlier within some sensitive financial indicators or with delay in some macroeconomic indicators. Therefore, to be precise, it is motivated to consider an endogenous date for the change in correlation structures.

A part of previous literature on correlation breaks concentrated on using parametric methods to detect changes on the coefficients of multivariate volatility models (Andreou and Ghysels, 2003; Qu and Perron, 2007). This stream of tests has reasonable statistical properties but suffers from the usual difficulties associated with parametric model selection and the heavy computational cost in large dimensional systems. Later, Aue et al. (2009) proposed a non-parametric CUSUM test for detecting instabilities of volatilities and cross-volatilities of multivariate time series models. They provided an asymptotic CUSUM test for the null of no change also derived its asymptotic distribution, and applied it to several multivariate frameworks such as, multivariate ARMA process and constant conditional correlation GARCH. Their test was modified by Wied et al. (2012) and their test is based on sample correlations.

Although the tests of Aue et al. (2009) and Wied et al. (2012) are of merits from aspects of easy computation and good statistical properties, it is developed under the assumption that the correlation structure remains a constant term. However, in the financial world, conditional correlation structures are seldom constant (Tse and Tsui, 2002; Engle, 2002), which can be captured through standardizing the observations by conditional variances. This paper aims to contribute to the literature on proposing a modification of the non-parametric CUSUM technique of Aue et al. (2009), to detect unknown change points of the dynamic evolving conditional correlations within multivariate volatility models.

In more detail, we propose a two steps semi-parametric test. The first step is to estimate the covariance structure through Multivariate GARCH models (including BEKK, Factor, CCC, DCC and ADCC), thereby obtaining conditional correlations and volatilities. Then, the second step is apply CUSUM test to detect correlation changes in the context of multivariate models with dynamically evolving conditional correlations. In order to discuss the performance of the test under mixing property correlation structure, a Monte Carlo simulation study is taken on data generating process with different extents of dependence. Results present that strong mixing data would deteriorate the performance of CUSUM tests, but this limitation can be controlled in an acceptable level by using M-GARCH models to estimate conditional correlations.

We then use the semi-parametric CUSUM test to detect the correlation breaks in large financial system for identifying the contagion effect. Investigating the dynamic correlations among Latin American, Central East European and East Asian emerging markets with U.S. markets, we find that the contagion effect existed in these three regions during the Great Recession. However, the contamination from United States to East Asian markets is not as strong as to other two regions.

The paper is structured as below. In the next section, we review literature on the aspects of M-GARCH models, relevant existing change point tests and their developments. Section 3.3 discussed the theory of semi-parametric CUSUM test. In section 3.4, we assess the performance of the test with a Monte Carlo simulation study. Section 3.5 provides an empirical application in the context of tests for financial contagion, a summary concludes.

## 3.2 Literature Review

The dynamics of volatilities and cross volatilities of multivariate processes play a crucial role in the understanding of the relationship between economic and especially financial observations. Hence the literature on multivariate volatilities, especially on GARCH-type models is rich. For reviews we refer to Bauwens et al. (2006), Engle (2009), Silvennoinen and Teräsvirta (2009) and Francq and Zakoian (2010). It has been established from an empirical as well as theoretical point of view that non-linear multivariate processes, especially with non-constant conditional, usually outperform linear time series models. Thus generalizations of the conditional correlation model of Bollerslev (1990), including the DCC model (Engle, 2002), and their extensions (cf. Cappiello et al. (2006), Noureldin et al. (2014)), are used to produce accurate forecasts both in-sample and out-of-sample.

However, all these popular models, like every empirical model in econometrics, must also account for changes in their parameters, which might arise as a result of sudden shocks occurring in the economy, market crashes, financial crises or intervention of the policy markers. One of the first and most influential work on change point detection is due to Page (1955), who proposed a cumulative sum method (CUSUM). Since then, the literature has developed both parametric and non-parametric tests, the classical results have been extended to time series (cf. Bai and Perron (1998), Bai and Perron (2003), Qu and Perron (2007), Aue et al. (2006) and Aue et al. (2009)). Due to their optimality properties, likelihood-based parametric tests have been widely used. Due to their simplicity and robustness, non-parametric, especially CUSUM-based approaches have become popular. Csörgő and Horváth (1997) study their asymptotic properties focusing on independent observations. De Pooter and Van

Dijk (2004) used CUSUM test to detect change in the variance of a heteroscedastic process, and their results showed a deterioration in the type I. Also, they pointed out that the size and power became good again when they conducted same test on the standardized residuals of a GARCH model. Aue and Horváth (2013) and Horváth and Rice (2014) reviewed several methods on how to derive asymptotic properties of popular methods when dependence between the observations cannot be neglected.

It is therefore clear why this theoretical issue has become an important concern for applied economists, and the extension of change point tests to a more flexible environment has become a critical research topic for theorists. Hence the literature on this topic has been developing in the last two decades. For example, Horváth and Kokoszka (1997) provided the asymptotic behavior of change point tests for Gaussian long-range dependent data. Antoch et al. (1997) studied the behavior of partial sum change point statistics when data follows an autoregressive order 1 process, establishing its asymptotic theory. Similarly, the change point tests of Quandt (1958, 1960) which was designed for i.i.d processes, was studied by Yao and Davis (1986) and further extended to detect change points in near-epoch dependent sequences by Ling et al. (2007). Inclan and Tiao (1994) were the first to attempt to detect changes in the variance of i.i.d processes using a cumulative sums of squares test. Then, the literature developed further with the contribution of Kim (2000), who showed the consistency of CUSUM tests when applied to a GARCH (1,1) process; and further, Lee et al. (2003) applied the test of Inclan and Tiao (1994) to detect variance change of non-stationary AR(q) process with strongly mixing innovations. Through these developments, most of the change point tests have become applicable to a wider variety of economic or financial processes.

However, in the context of financial data, second moments are usually modeled by ARCH or GARCH models. Kokoszka and Leipus (2000) examined change point tests in processes with dependent volatility. Their findings showed the CUSUM test is valid when applied to short memory ARCH model making feasible to detect changes within certain types of ARCH models in financial data (cf. also Andreou and Ghysels (2002) and Fryzlewicz and Rao (2011)). Andreou and Ghysels (2002) applied the test of Kokoszka and Leipus (2000) to detect multiple changes in the volatility of high frequency stock and foreign exchange data, where the conditional variance is captured by a GARCH model. They found, consistent with results reported by Kokoszka and Leipus (2000) and De Pooter and Van Dijk (2004), that if the data are highly persistent or have long memory, the existence of change points is overestimated and the tests became completely unreliable, especially in case of nearly nonstationary processes.

Change points detection in the second moments is not limited to univariate cases, but it can be extended to the covariance and correlation structure of multivariate models. Early studies on change points detection in the covariance structure were focussed on using model selection criteria and standard stability tests on the parameters of M-GARCH models. For example, Yao (1988), Lavielle and Teyssiere (2006) proposed penalized contrast function to detect multiple changes in covariance structure simultaneously. The procedures however are only designed to be applied to i.i.d and weakly dependent constant conditional correlation (CCC) processes. However, even though these methods can detect change points within the framework of CCC models, it is arguable that allowing for dynamic conditional correlations (DCC) would be more realistic as shown in Tse and Tsui (2002) and Engle (2002). In order to detect conditional correlation change in foreign exchange market, Andreou and Ghysels (2002) modeled multivariate data by DCC model and then detected parameter changes through the test of Bai and Perron (1998). Among the theoretical studies, Aue et al. (2006) provided a strong approximation for the CUSUM statistics in the context of GARCH processes. Aue et al. (2009) constructed a CUSUM statistics for detecting changes in the covariance structure of multivariate sequences and derived their asymptotics. Their method was transferred to study the stability of the correlation matrix by Wied et al. (2012). As mentioned above, assuming constant conditional correlations is often unrealistic, as highly-persistent conditional correlation exists as the norm especially in financial data. Hence, the emergence of the need of detecting changes in dynamically evolving persistent conditional correlation structures. In this paper, deriving the appropriate conditions, we show that the asymptotic theory of CUSUM tests is still valid for the cases where the conditional correlation structure is allowed to be (strongly) persistent (albeit not integrated) and modeled as constant conditional correlation (CCC), dynamic conditional correlation (DCC), BEKK, factor, asymmetric DCC and BEKK processes. Finally we apply the procedure to detect the occurrence of financial contagion as defined in Forbes and Rigobon (2002).

## 3.3 The CUSUM test

Drawing from the previous work of Aue et al. (2009) and Wied et al. (2012), we test the stability of covariance structure of a d-dimensional random vector $(\mathbf{y}_t : t \in \mathbb{Z})$, with $E[|\mathbf{y}_t|^2] = \mu$ and $E[|\mathbf{y}_t|^2] < \infty$, where $|\cdot|$ denotes the Euclidean norm in $\mathbb{R}^d$. In this section, we modify the test of Aue et al. (2009) and extend it to the cases where the observations evolve according to the specifications of BEKK, Factor, CCC, DCC and ADCC GARCH models.

### 3.3.1 Test the stability of time-varying correlation structures

The CUSUM statistics are designed to examine the stability of volatilities and cross-volatilities, but studying just pure correlation relationships sometimes is an issue to assets or other financial variables. Indeed, changes in the correlation structure may be caused by changes in volatility or covariance. To detect changes in the correlation structure, this paper uses de-volatilized data to remove the influence from volatilities. Let $\mathbf{y}_1$, $\mathbf{y}_2$, ..., $\mathbf{y}_n$ denote the observations satisfying

$$E[\mathbf{y}_t] = 0$$

and $\mathbf{y}_t = [y_t(1), y_t(2), \ldots, y_t(d)]^\top$. The conditional variance of $y_t(j)$ given the past is denoted by $\tau_t^2(j)$, i.e. $\tau_t^2(j) = E(y_t^2(j)|\mathcal{F}_{t-1})$, where the $\sigma$-algebra $\mathcal{F}$ is generated by $\{\mathbf{y}_s, s \leq t-1\}$. The de-volatilized observations are denoted by

$$y_t^*(j) = \frac{y_t(j)}{\tau_t(j)}, \quad 1 \leq t \leq n, 1 \leq j \leq d$$

Our paper follows the methodology of the most often used multivariate volatility model

$$\mathbf{y}_t = \mathbf{\Sigma}_t^{\frac{1}{2}} \mathbf{e}_t \tag{3.3.1}$$

where

**Assumption 3.3.1.** $\{\mathbf{e}_t, -\infty < t < \infty\}$ *are independent and identically distributed random vector in* $\mathbb{R}^d$, $E\mathbf{e}_0 = \mathbf{0}$ *and* $E\mathbf{e}_0\mathbf{e}_0^\top = \mathbf{I}_d$.

where $\mathbf{I}_d$ is the $d \times d$ identity matrix.

**Assumption 3.3.2.** $\mathbf{\Sigma}_t \in \mathcal{F}_{t-1}$ *and* $\{\mathbf{\Sigma}_t, -\infty < t < \infty\}$ *is a stationary and ergodic sequence. Hence the conditional covariance matrix of* $\mathbf{y}_t$ *with respect to its past is* $E(\mathbf{y}_t\mathbf{y}_t\top|\mathcal{F}_{t-1}) = \mathbf{\Sigma}_t$.

To avoid degenerate cases we assume that,

**Assumption 3.3.3.** *There is a positive definite matrix* $\mathbf{\Sigma}^0$ *such that* $\mathbf{\Sigma}_t - \mathbf{\Sigma}^0$ *is non-negative definite for all* $t$.

If $\mathbf{\Sigma}_t = \{\sigma_t(k,j), 1 \leq k, j \leq d\}$, then $\tau_t(j) = \sigma_t^{\frac{1}{2}}(j,j)$. It follows from Assumption 3.3.3 that there is a positive constant $\tau_0$ such that $\tau_t(j) \geq \tau_0$ for all $n$ and $1 \leq j \leq d$. It is an immediate consequence of Assumptions 3.3.1 and 3.3.2 that $\mathbf{y}_t$ is composed by stationary and ergodic sequences. The next condition is on the dependence structure of the observations:

**Assumption 3.3.4.** *$E\|\mathbf{y}_0\|^r$ with some $r > 2$ and $\{\mathbf{y}_t, -\infty < t < \infty\}$ is $\beta$-mixing with rate $t^{(-\delta-r)/(r-2)}$ with some $\delta > 0$*

We note that Assumption 3.3.4 can be replaced with the condition that $\mathbf{\Sigma}_t$ is $\beta$-mixing.

Let

$$\rho_t(i,j) = E y_t^*(i) y_t^*(j), \quad 1 \le t \le T, 1 \le i,j \le d$$

be the covariance of the devolatized observations $y_t^*(i)$. In this paper, we wish to test the null hypothesis that

$$H_0\colon\ \rho_1(i,j) = \rho_2(i,j) = \cdots = \rho_T(i,j) \text{ for all } 1 \le i,j \le d$$

against the alternative

$$H_A\colon \text{ there are } 1 < t^* < T \text{ and } 1 \le i,j \le d \text{ such that}$$
$$\rho_1(i,j) = \rho_2(i,j) = \cdots = \rho_{t^*}(i,j) \ne \rho_{t^*+1}(i,j) = \cdots = \rho_T(i,j)$$

under the null hypothesis the covariance matrix of the vector $[y_t^*(1), y_t^*(2), \ldots, y_t^*(d)]^\top$ does not depend on the time $t$ while under the alternative at least one of the elements of the covariance matrix changes at the unknown time $t^*$.

Our procedure is based on two functionals of the CUSUM of the vectors

$$\mathfrak{r} = vech(y_t^*(i) y_t^*(j)), 1 \le i,j \le d$$

Define the partial sum process

$$\mathfrak{s}(t) = \sum_{s=1}^{t} \mathfrak{r}_s$$

Assuming that $H_0$ holds, i.e. the data are stationary we define the long run covariance matrix

$$\mathfrak{D} = \sum_{s=-\infty}^{\infty} E\mathfrak{s}_0\mathfrak{s}_s^\top$$

The normalization in our procedures requires

**Assumption 3.3.5.** $\mathfrak{D}$ *is a nonsingular matrix.*

Following Aue et al. (2009) (cf. Wied et al, 2012) we define two statistics

$$M_T^{(1)} = \frac{1}{T}\max_{1\le t\le T}(\mathfrak{s}(t) - \frac{t}{T}\mathfrak{s}(T))^\top\mathfrak{D}^{-1}(\mathfrak{s}(t) - \frac{t}{T}\mathfrak{s}(T))$$

$$M_T^{(2)} = \frac{1}{T^2}\sum_{t=1}^{T}(\mathfrak{s}(t) - \frac{t}{T}\mathfrak{s}(T))^\top\mathfrak{D}^{-1}(\mathfrak{s}(t) - \frac{t}{T}\mathfrak{s}(T))$$

**Theorem 3.3.1.** *If $H_0$ and Assumptions 3.3.1-3.3.5 hold, then*

$$M_T^{(1)} \xrightarrow{\mathcal{D}} M^{(1)} \tag{3.3.2}$$

$$M_T^{(2)} \xrightarrow{\mathcal{D}} M^{(2)} \tag{3.3.3}$$

where

$$M^{(1)} = \sup_{0\le u\le 1} \textstyle\sum_{i=1}^{\bar{d}} B_i^2(u) \text{ and } M^{(2)} = \sum_{i=1}^{\bar{d}} \int_0^1 B_i^2(u)du \text{ with } \bar{d} = \frac{d(d+1)}{2}$$

and $B_1$, $B_2$, ..., $B_{\bar{d}}$ denote independent Brownian bridges.

The limiting random variables $M^{(1)}$ and $M^{(2)}$ already appeared in Aue et al. (2009), where selected critical values and approximations for moderate and large values of $\bar{d}$ can also be found.

The conditional covariance matrices $\mathbf{\Sigma}_t$ can be written as functionals of the random vectors $\mathbf{y}_s, s \le t-1$. However, since we can observe only $\mathbf{y}_1$, $\mathbf{y}_2$, ..., $\mathbf{y}_T$, first we replace $\tau_t(i)$

with $\bar{\tau}_t(i)$, where $\bar{\tau}_t(i)$ is a function of $\mathbf{y}_1$, $\mathbf{y}_2$, ..., $\mathbf{y}_{t-1}$ only. Also, we consider parametric models in the present paper, so $\tau_t(i)$ as well as $\bar{\tau}_t(i)$ depend on unknown parameters which will be denoted by $\boldsymbol{\theta} \in R^p$. We require that $\bar{\tau}_t(i;\boldsymbol{\theta})$ and $\tau_t(i;\boldsymbol{\theta})$ are close if $t$ is large. This requirement is standard in the estimation of GARCH and similar volatility processes (cf. Francq and Zakoian, 2010):

**Assumption 3.3.6.** *There is a ball $\boldsymbol{\Theta}_0 \subset R^P$ with centre $\boldsymbol{\theta}_0$ such that $\max_{1\le i\le d} sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}_0} |\tau_t(i) - \bar{\tau}_t(i)|$ $= O(a(t))$ a.s. and $ta(t) \to 0$ as $t \to \infty$.*

We estimate $\boldsymbol{\theta}$ with $\hat{\boldsymbol{\theta}}_T$ which is consistent with rate $T^{-1/2}$:

**Assumption 3.3.7.** $\left\|\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0\right\| = O_P(T^{-1/2})$*, where $\boldsymbol{\theta}_0$ denotes the value of the parameter under $H_0$.*

The random functions $\tau_t(i) = \tau_t(i;\boldsymbol{\theta}), 1 \le i \le p$, are smooth functions of $\boldsymbol{\theta}$ in a neighborhood of $\boldsymbol{\theta}_0$:

**Assumption 3.3.8.** *There is a ball $\boldsymbol{\Theta}_0 \subset R^P$ with center $\boldsymbol{\theta}_0$ such that*

$$\left\|\tau_t(i;\boldsymbol{\theta}) - \tau_t(i;\boldsymbol{\theta}_0) - \mathbf{g}_t^\top(i)(\boldsymbol{\theta} - \boldsymbol{\theta}_0)\right\| \le \bar{g}_t \left\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\right\|^2$$

*for all $\boldsymbol{\theta} \in \boldsymbol{\Theta}_0$, where $\{\mathbf{g}_t(i), 1 \le i \le p, \bar{g}_t, -\infty < t < \infty\}$ is a stationary and ergodic sequence with $E\left\|\mathbf{g}_0(i)\right\|^2 < \infty$ and $E\left|\bar{g}_0\right|^2 < \infty$.*

The quasi maximum likelihood method is the most often used technique to estimate parameters of a nonlinear time series model. In the examples we discuss in this paper, the QMLE satisfies Assumption 3.3.6 - 3.3.8.

Now the de-volatized variables

$$\hat{y}_t(i) = \frac{y_t(i)}{\bar{\tau}_t(i;\hat{\boldsymbol{\theta}}_T)}$$

can be computed from the sample. The long run covariance matrix $\mathfrak{D}$ is estimated from the sample by $\hat{\mathfrak{D}}_T$ which satisfies

**Assumption 3.3.9.** $\left\|\hat{\mathfrak{D}}_T - \mathfrak{D}\right\| = o_P(1)$.

We discuss the estimation of $\mathfrak{D}$ in the next subsection. Similarly to $M_T^{(1)}$ and $M_T^{(2)}$ we define

$$\hat{M}_T^{(1)} = \frac{1}{T} \underset{1\le t\le T}{max} (\hat{\mathfrak{s}}(t) - \frac{t}{T}\hat{\mathfrak{s}}(T))^\top \hat{\mathfrak{D}}^{-1}(\hat{\mathfrak{s}}(t) - \frac{t}{T}\hat{\mathfrak{s}}(T))$$

$$\hat{M}_T^{(2)} = \frac{1}{T^2} \sum_{t=1}^{T} (\hat{\mathfrak{s}}(t) - \frac{t}{T}\hat{\mathfrak{s}}(T))^\top \hat{\mathfrak{D}}^{-1}(\hat{\mathfrak{s}}(t) - \frac{t}{T}\hat{\mathfrak{s}}(T))$$

where

$$\hat{\mathfrak{s}}(t) = \textstyle\sum_{s=1}^{t} \hat{\mathfrak{r}}_s, \text{ with } \hat{\mathfrak{r}}_s = vech(\hat{y}_s(i)\hat{y}_s(j), 1 \le i, j \le d).$$

**Theorem 3.3.2.** *If $H_0$ and Assumptions 3.3.1-3.3.9 hold, then*

$$\hat{M}_T^{(1)} \xrightarrow{\mathcal{D}} M^{(1)} \tag{3.3.4}$$

*and*

$$\hat{M}_T^{(2)} \xrightarrow{\mathcal{D}} M^{(2)} \tag{3.3.5}$$

where $M^{(1)}$ and $M^{(2)}$ are defined in Theorem 3.3.1.

### 3.3.2 Examples for time dependent conditional volatilities

In this section, we consider some often used models where Assumption 3.3.1-3.3.8 are satisfied. There are two main methods to establish stationarity and geometric mixing properties of nonlinear time series models. Markov chain theory combined with algebraic geometry can be used to establish the existence and basic properties of the underlying model. The important technical tools are summarized in the reference book of Meyn and Tweedie (1993). This approach was used successfully in Carrasco and Chen (2002) for univariate models and later extended by Boussama et al. (2011) to the multivariate case. Diaconis and Freeman (1999) showed that random recursions satisfying "contraction in average" have a unique stationary solution and the method of the proof gives geometric ergodicity. Applications of the method of Diaconis and Freeman (1999) to nonlinear time series are detailed in Douc et al. (2014). For further multivariate GARCH type models, we refer to the survey paper of Bauwens et al. (2006), Silvennoinen and Teräsvirta (2009), Francq and Zakoian (2010).

**Example 3.3.1.** *(CCC(p,q) model)*

Bollerslev (1990) and Jeantheau (1998) specified the constant conditional correlation model by the following equations:

$$\mathbf{\Sigma}_t = \mathbf{D}_t \mathbf{R} \mathbf{D}_t \tag{3.3.6}$$

$$\mathbf{D}_t = diag(h_t^{\frac{1}{2}}(1), h_t^{\frac{1}{2}}(2), \ldots, h_t^{\frac{1}{2}}(d)), \quad \mathbf{h}_t = (h_t(1), h_t(2), \ldots, h_t(d))^\top \tag{3.3.7}$$

and

$$\mathbf{h}_t = \mathbf{c} + \sum_{l=1}^{q} \mathbf{A}_l (\mathbf{y}_{t-l} \circ \mathbf{y}_{t-l}) + \sum_{j=1}^{p} \mathbf{B}_j \mathbf{h}_{t-j} \tag{3.3.8}$$

where $\circ$ denotes the Hadamard product of vectors (coordinatewise multiplication), $\mathbf{R}$ is a correlation matrix, $\mathbf{C}$ is a vector of positive coordinates, $\mathbf{A}_l$, $1 \leq l \leq q$, $\mathbf{B}_j$, $1 \leq j \leq p$ are matrices with nonnegative elements. Sufficient conditions for the existence of a unique stationary solution and existence of moments are given in Aue et al. (2009), and their method can also be used to prove Assumption 3.3.4. For further discussion we refer to Francq and Zakoian (2010). Francq and Zakoian (2010) also give a detailed account of the estimation of the parameters of an CCC(p,q) sequence by quasi maximum likelihood and Assumption

3.3.6-3.3.8 are established. In addition to the QMLE, the variance targeting estimator also satisfies our assumptions (cf. Francq et al. 2016). Francq and Zakoian (2014) proposes a new method to estimate parameters utilizing the covariance structure of the observations. Their proofs show that Assumptions 3.3.6-3.3.8 hold.

**Example 3.3.2.** *(DCC model)*

Dynamic conditional correlation GARCH models are an extension of Example 3.3.1, obtained by introducing a dynamic for the conditional correlation by replacing $\mathbf{R}$ with random matrices $\mathbf{R}_t$, measurable with respect to the $\sigma$-algebra generated by $\{\mathbf{y}_s, s \leq t-1\}$ i.e. Equation 3.3.6 is replaced with

$$\mathbf{\Sigma}_t = \mathbf{D}_t \mathbf{R}_t \mathbf{D}_t \tag{3.3.9}$$

while recursions 3.3.7 and 3.3.8 still hold. Tse and Tsui (2002) suggested,

$$\mathbf{R}_t = \theta_1 \mathbf{C} + \theta_2 \mathbf{\Psi}_{t-1} + \theta_3 \mathbf{R}_{t-1}$$

where $\mathbf{C}$ is a positive definite matrix, and the weights $\theta_1 > 0$, $\theta_i \geq 0$ $i = 2, 3$ satisfy $\theta_1 + \theta_2 + \theta_3 = 1$, $\mathbf{\Psi}$ is the empirical correlation matrix from innovation terms, and $\mathbf{R}_t$ is the empirical covariance matrix of $\mathbf{y}_{t-1}$, $\mathbf{y}_{t-2}$, $\ldots$, $\mathbf{y}_{t-K}$. Engle (2002) advocated

$$\mathbf{R}_t = (diag(\mathbf{Q}_t))^{-\frac{1}{2}} \mathbf{Q}_t (diag(\mathbf{Q}_t))^{-\frac{1}{2}}$$

where $\mathbf{Q}_t$ is a covariance matrix measurable with respect to $\{\mathbf{y}_s, s < t\}$. A scalar parameterization is

$$\mathbf{Q}_t = \theta_1 \mathbf{C} + \theta_2 \mathbf{y}_{t-1} \mathbf{y}_{t-1}^\top + \theta_3 \mathbf{Q}_{t-1} \tag{3.3.10}$$

where $\mathbf{C}$ is a positive definite matrix and $\theta_1$, $\theta_2$ and $\theta_3$ are weights as above. One can use a matrix representation as well, replacing 3.3.9 with

$$\mathbf{Q}_t = \mathbf{C} + \mathbf{A} \circ \mathbf{y}_{t-1}\mathbf{y}_{t-1}^\top + \mathbf{B} \circ \mathbf{Q}_{t-1} \tag{3.3.11}$$

where $\mathbf{A}$ and $\mathbf{B}$ are non-negative definite matrices. For further results we refer to Francq and Zakoian (2010).

**Example 3.3.3.** *(BEKK(p,q) model)*

Baba, Engle, Kraft and Kroner (cf. Engle and Kroner, 1995) introduced the model where the conditional covariance matrix satisfies,

$$\mathbf{\Sigma}_t = \mathbf{C} + \sum_{j=1}^{q}\sum_{k=1}^{k_0} \mathbf{A}_{j,k}\mathbf{y}_{j,t}\mathbf{y}_{j,t}^\top\mathbf{A}_{j,k}^\top + \sum_{j=1}^{p}\sum_{k=1}^{k_0} \mathbf{B}_{j,k}\mathbf{\Sigma}_{j,t}\mathbf{B}_{j,k}^\top \tag{3.3.12}$$

where $k_0$ is an integer, $\mathbf{C}$, $\mathbf{A}_{j,k}$, $1 \le j \le q$, $1 \le k \le k_0$, and $\mathbf{B}_{j,k}$, $1 \le j \le p$, $1 \le k \le k_0$ are $d \times d$ matrices. The expected covariance $\mathbf{C}$ is positive definite. Boussama et al. (2011) finds conditions which implies that the BEKK model has a unique stationary solution, and it is geometrically $\beta$-mixing. Comte and Lieberman (2003) studied the properties of QMLE while Pedersen and Rahbek (2014) discussed the variance targeting method. For a detailed discussion of the BEKK model, we refer to Francq and Zakoian (2010).

**Example 3.3.4.** *(Factor Model)*

Engle, Ng and Rothschild (1990) defined the conditional covariance matrix $\mathbf{\Sigma}_t$ by,

$$\mathbf{\Sigma}_t = \mathbf{C} + \sum_{j=1}^{p} \lambda_{j,t}\boldsymbol{\beta}_j\boldsymbol{\beta}_j^\top \quad \textit{and} \quad \lambda_t(j) = \omega_j + \alpha_j y_{t-1}(j)^2 + \beta_j\lambda_{t-1}(j)$$

where $\mathbf{C}$ is a positive definite matrix, $\omega_j > 0$, $\alpha_j \ge 0$, $\beta_j \ge 0$, $1 \le j \le p$, and $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots \boldsymbol{\beta}_p$ are linear independent. Francq and Zakoian (2010) points out that the factors model can be written in a BEKK form, so clearly Assumptions (3.3.1 - (3.3.4) are satisfied under mild conditions.

**Example 3.3.5.** *(Asymmetric Models)*

The models in Examples 3.3.1 - 3.3.4 are symmetric, i.e. we have the same dynamic for the conditional covariances for negative and positive coordinates of the observations. There are several extensions to accommodate the asymmetric nature of financial data. The leverage effect in financial data indicates that negative innovations make an additional impact on the dynamic evolution of correlation. Let $\mathbf{y}_t^- = (y_t(1)I\{y_t(1) < 0\}, y_t(2)I\{y_t(2) < 0\}, \ldots, y_t(d)I\{y_t(d) < 0\})^\top$, where $I$ is an indicator which equals to 1 if $y_t(i) < 0$, otherwise equals to 0. Cappiello et al. (2006) introduced an asymmetric version of the DCC model in Example 3.1.4. They replaced Equation 3.3.9 by adding an asymmetric indicator.

$$\mathbf{Q}_t = \theta_1 \mathbf{C} + \theta_2 \mathbf{y}_{t-1}\mathbf{y}_{t-1}^\top + \theta_3 \mathbf{y}_{t-1}^-(\mathbf{y}_{t-1}^-)^\top + \theta_4 \mathbf{Q}_{t-1} \qquad (3.3.13)$$

where $\mathbf{C}$ is positive definite, $\theta_1 > 0$, $\theta_i \geq 0$, $i = 2, 3, 4$ with restriction of $\theta_1 + \theta_2 + \theta_3 + \theta_4 = 1$. The asymmetric version of (3.3.11) is given by

$$\mathbf{Q}_t = \mathbf{C} + \mathbf{A} \circ \mathbf{y}_{t-1}\mathbf{y}_{t-1}^\top + \mathbf{D} \circ \mathbf{y}_{t-1}^-(\mathbf{y}_{t-1}^-)^\top + \mathbf{B} \circ \mathbf{Q}_{t-1}$$

where $\mathbf{C}$ is a positive definite matrix, $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{D}$ are non-negative matrices.

In the asymmetric version of the BEKK model in Example 3.3.3, the conditional covariance matrix satisfies the recursion

$$\mathbf{\Sigma}_t = \mathbf{C} + \sum_{j=1}^{q}\sum_{k=1}^{k_0} \mathbf{A}_{j,k}\mathbf{y}_{j,t}(\mathbf{y}_{j,t}\mathbf{A}_{j,k})^\top + \sum_{j=1}^{q}\sum_{k=1}^{k_0} \mathbf{A}_{j,k}^-\mathbf{y}_{j,t}^-(\mathbf{y}_{j,t}^-\mathbf{A}_{j,k}^-)^\top + \sum_{j=1}^{p}\sum_{k=1}^{k_0} \mathbf{B}_{j,k}\mathbf{\Sigma}_{j,t}\mathbf{B}_{j,k}^\top$$

where $k_0$ is an integer, $\mathbf{C}$, $\mathbf{A}_{j,k}$, $\mathbf{A}_{j,k}^-$, $1 \leq j \leq q$, $1 \leq k \leq k_0$, and $\mathbf{B}_{j,k}$, $1 \leq j \leq p$, $1 \leq k \leq k_0$ are $d \times d$ matrices, and $\mathbf{C}$ is positive definite. Silvennoinen and Teräsvirta (2009) gave a review of the properties of the asymmetric BEKK model.

### 3.3.3 Estimation of the long run covariance matrix $\mathfrak{D}$

We use the kernel estimators

$$\hat{\mathfrak{D}}_T = \sum_{l=-T}^{T} K(\frac{l}{h})\hat{\gamma}_l \tag{3.3.14}$$

where

$$\hat{\gamma}_l = \begin{cases} \frac{1}{T}\sum_{t=1}^{T-l}(\hat{\mathfrak{r}}_t - \bar{\mathfrak{r}}_T)(\hat{\mathfrak{r}}_{t+l} - \bar{\mathfrak{r}}_T)^\top & if \quad 0 \le l < T \\ \frac{1}{T}\sum_{t=-T+1}^{T}(\hat{\mathfrak{r}}_t - \bar{\mathfrak{r}}_T)(\hat{\mathfrak{r}}_{t+l} - \bar{\mathfrak{r}}_T)^\top & if \quad -T \le l < 0 \end{cases}$$

where

$$\bar{\mathfrak{r}}_T = \frac{1}{T}\sum_{s=1}^{T}\hat{\mathfrak{r}}_s$$

There are several choices for the the choice of the kernel $K$, including Bartlett, truncated, Parzen, Tukey-Hanning and quadratic spectral kernels (cf. Andrews (1991) for a review of the properties of kernel functions). The window (smoothing parameter) satisfies $h = h(T)$, $h/T \to \infty$ and $h/T \to 0$. It can be shown that under Assumption 3.3.1 - 3.3.8, the estimator $\hat{\mathfrak{D}}$ is consistent in case of Bartlett, truncated, Parzen, Tukey-Hanning and quadratic spectral kernels. For the choice of $h$ we refer to Andrews (1991).

## 3.4 The Monte Carlo simulations

To assess the performance of semi-parametric CUSUM tests in the models of Examples 3.3.1 - 3.3.5, in this section we conducts a Monte Carlo simulation study to measure the rejection rate under the null and alternative hypothesis in case of finite sample sizes. For simplicity, we consider bivariate observations $\mathbf{y}_t = (y_t(1), y_t(2))^\top$. In the data generating process DGP we specify $\mathbf{e}_t$ as a bivariate standard normal vector and $\mathbf{\Sigma}_t^{1/2}$ of (3.3.1) is in Cholesky form. For each model, we set the initial value $\mathbf{\Sigma}_0$ to be the $2 \times 2$ identity matrix and simple iterations give $\Sigma_t$ for the specified parameter values. Under the alternative a single change occurs at time $t^* = T/2$. The Bartlett kernel

$$K_B(x) = (1 - |x|)I\{|x| \leq x\}$$

and the Newey-West optimal window (smoothing parameter) are used in the definition of $\hat{\mathfrak{D}}_T$. The observations are first demeaned, i.e. the sample mean is removed from the observations. We provide numerical results for both $\hat{M}_T^{(1)}$ and $\hat{M}_T^{(2)}$. Assuming that a change occurred, we estimate the time of change with $argmax\{\hat{\mathfrak{s}}(t) - (t/T)\hat{\mathfrak{s}}(T), 1 \leq t \leq T\}$.

In each experiment, we set number of observations $T = 100$ for a small sample and $T = 1000$ for a large sample, and each simulation is replicated 5000 times. The warming up parameter is 0.2, which implies that the test will burn 200 observations if sample size is 1000.

### 3.4.1 CCC Model

Let $p = q = 1$ in Example 3.3.1. The constant correlation matrix is given by

$$\mathbf{R} = \begin{bmatrix} 1 & \delta \\ \delta & 1 \end{bmatrix}$$

with $\mathbf{c} = (0.01, 0.01)^\top$,

$$\mathbf{A_1} = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} \quad and \quad \mathbf{B_1} = \begin{bmatrix} b & 0 \\ 0 & b \end{bmatrix} \tag{3.4.1}$$

with $a = 0.01$ and $b = 0.7$ in our simulation study. The choice of $\mathbf{A}_1$ and $\mathbf{B}_1$ is motivated by the empirical observation that financial data show low ARCH but high GARCH (persistence) effect. The matrices $\mathbf{A}_1$ and $\mathbf{B}_1$ determine the dynamics of the process, but their values are not crucial after devolatizing was done. Under the null hypothesis $\delta = 0$ and the critical values for $\hat{M}_T^{(1)}$ and $\hat{M}_T^{(2)}$ are reported for sample sizes $T = 100$ and 1000 together with the asymptotic critical values. Table 3.4.1 shows that the asymptotic critical values overestimate the true onces and it is very mild $\hat{M}_T^{(2)}$. Table 3.4.2 shows the empirical rejections under the alternative for $\delta$ changing from 0 to $\delta = 0.2, 0.4, 0.6$ and 0.8. at $t^* = T/2$ ($\delta = 0$ corresponds to the empirical rejection under the null hypothesis). The power os increasing with $\delta$ for both test, $\hat{M}_T^{(2)}$ showing somewhat larger power for $T = 100$ and $\hat{M}_T^{(1)}$ for $T = 1000$. The estimator for $t^*/T$ the fraction of the time of change is accurate and it is getting better for larger $\delta$ and larger sample sizes in Table 3.4.3.

Table 3.4.1: Critical Values for the statistics $\hat{M}_T^{(1)}$, $\hat{M}_T^{(2)}$ statistics (d=2) and the asymptotic critical values computed from the distribution of $M^{(1)}$ and $M^{(2)}$ (d=2)

| | | CCC | | BEKK | | Factor | | DCC | | ADCC | | Asymptotic |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| T | | 100 | 1000 | 100 | 1000 | 100 | 1000 | 100 | 1000 | 100 | 1000 | $\infty$ |
| | 90% | 2.40 | 2.42 | 2.61 | 2.50 | 2.61 | 2.53 | 2.40 | 2.42 | 2.39 | 2.43 | 2.63 |
| $\hat{M}_t^{(1)}$ | 95% | 2.81 | 2.85 | 3.15 | 2.87 | 3.14 | 2.90 | 2.81 | 2.85 | 2.78 | 2.84 | 3.06 |
| | 99% | 3.99 | 3.72 | 4.43 | 3.67 | 4.42 | 3.81 | 3.99 | 3.72 | 4.01 | 3.64 | 3.95 |
| | 90% | 0.78 | 0.79 | 0.87 | 0.78 | 0.90 | 0.83 | 0.78 | 0.79 | 0.79 | 0.80 | 0.83 |
| $\hat{M}_t^{(2)}$ | 95% | 0.95 | 0.95 | 1.05 | 0.94 | 1.09 | 0.98 | 0.95 | 0.95 | 0.95 | 0.95 | 0.97 |
| | 99% | 1.40 | 1.34 | 1.52 | 1.29 | 1.51 | 1.37 | 1.40 | 1.34 | 1.37 | 1.28 | 1.30 |

Table 3.4.2: Empirical rejection rates for $\hat{M}_T^{(1)}$ and $\hat{M}_T^{(2)}$ in the CCC model of Example 3.3.1

| | $\hat{M}_T^{(1)}$ (T=100) | | | | | $\hat{M}_T^{(2)}$ (T=100) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $\delta$ | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 |
| 90% | 0.10 | 0.19 | 0.44 | 0.82 | 0.99 | 0.12 | 0.21 | 0.46 | 0.83 | 0.99 |
| 95% | 0.06 | 0.12 | 0.32 | 0.73 | 0.98 | 0.06 | 0.12 | 0.33 | 0.71 | 0.97 |
| 99% | 0.01 | 0.03 | 0.12 | 0.41 | 0.86 | 0.01 | 0.03 | 0.13 | 0.42 | 0.84 |
| | $\hat{M}_T^{(1)}$ (T=1000) | | | | | $\hat{M}_T^{(2)}$ (T=1000) | | | | |
| $\delta$ | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 |
| 90% | 0.11 | 0.80 | 1.00 | 1.00 | 1.00 | 0.11 | 0.79 | 1.00 | 1.00 | 1.00 |
| 95% | 0.06 | 0.71 | 1.00 | 1.00 | 1.00 | 0.05 | 0.69 | 1.00 | 1.00 | 1.00 |
| 99% | 0.01 | 0.43 | 1.00 | 1.00 | 1.00 | 0.01 | 0.39 | 1.00 | 1.00 | 1.00 |

Table 3.4.3: Estimated of the time of change as a percentage of the observation period in the CCC model of Example 3.3.1 when $t^* = T/2$

| | T=100 | | | | T=1000 | | | |
|---|---|---|---|---|---|---|---|---|
| $\delta$ | 0.2 | 0.4 | 0.6 | 0.8 | 0.2 | 0.4 | 0.6 | 0.8 |
| Mean | 0.52 | 0.50 | 0.49 | 0.48 | 0.50 | 0.50 | 0.50 | 0.50 |
| Median | 0.52 | 0.50 | 0.49 | 0.49 | 0.50 | 0.50 | 0.49 | 0.50 |
| SD | 0.18 | 0.14 | 0.09 | 0.06 | 0.09 | 0.03 | 0.01 | 0.01 |

### 3.4.2 DCC Model

Let $p = q = 1$. We use (3.3.9), where $\mathbf{C}$ is given by,

$$\mathbf{C} = \begin{bmatrix} 1 & \delta \\ \delta & 1 \end{bmatrix} \tag{3.4.2}$$

where $\delta = 0$ under the no change null hypothesis and it changes from 0 to $\delta = 0.2, 0.4, 0.6$ and 0.8 at $t^* = T/2$. $\theta_1 = 1 - \delta_2 - \delta_3$, $\theta_2 = 0.01$ or 0.02 (low and high ARCH effect in quasi conditional correlation process), and $\theta_3 = 0.1$ or 0.9 (low and high persistence). To generate conditional variance matrix $\mathbf{D}_t$, the matrices $\mathbf{A}_1$ and $\mathbf{B}_1$ are defined in (3.4.1) with $a = 0.01$ and $b = 0.7$, and $\mathbf{c} = (0.01, 0.01)^\top$ is same as in Section 3.4.1. The outcome of the Monte Carlo experiment is similar to the results we obtained for the CCC model in Section 3.4.1. Table 3.4.1 compares the finite sample critical values to the asymptotic ones. The asymptotic critical values somewhat larger than the finite sample values but still acceptable according to the columns in Table 3.4.4 when $\delta = 0$. Table 3.4.4 and 3.4.5 show the power and the estimation of the location of the time of change.

Table 3.4.4: Empirical rejection rates for $\hat{M}_T^{(1)}$ and $\hat{M}_T^{(2)}$ in the DCC model of Example 3.3.2

$\delta$ is the change magnitude, which is the argument in constant coefficient matrix **C**. Parameters $a$ is the argument in ARCH coefficient matrix **A**, and $b$ is the argument in GARCH coefficient matrix **B**. The rejection rate is computed under three significance level, 90%, 95% and 99%.

**T=100**

| | $a = 0.01$ & $b = 0.1$ | | | | | | | | | | $a = 0.01$ & $b = 0.9$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | |
| $\delta$ | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 |
| 90% | 0.11 | 0.19 | 0.43 | 0.80 | 0.99 | 0.12 | 0.92 | 0.97 | 1.00 | 1.00 | 0.14 | 0.21 | 0.39 | 0.72 | 0.95 | 0.15 | 0.92 | 0.96 | 1.00 | 1.00 |
| 95% | 0.06 | 0.12 | 0.32 | 0.71 | 0.97 | 0.06 | 0.84 | 0.94 | 0.99 | 1.00 | 0.08 | 0.13 | 0.30 | 0.61 | 0.90 | 0.08 | 0.85 | 0.92 | 0.98 | 1.00 |
| 99% | 0.01 | 0.03 | 0.12 | 0.42 | 0.84 | 0.01 | 0.58 | 0.77 | 0.96 | 1.00 | 0.02 | 0.05 | 0.11 | 0.32 | 0.69 | 0.02 | 0.60 | 0.75 | 0.93 | 1.00 |

| | $a = 0.02$ & $b = 0.1$ | | | | | | | | | | $a = 0.02$ & $b = 0.9$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | |
| $\delta$ | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 |
| 90% | 0.10 | 0.19 | 0.43 | 0.80 | 0.99 | 0.11 | 0.92 | 0.97 | 1.00 | 1.00 | 0.14 | 0.22 | 0.38 | 0.67 | 0.91 | 0.16 | 0.93 | 0.96 | 0.99 | 1.00 |
| 95% | 0.05 | 0.12 | 0.31 | 0.70 | 0.97 | 0.05 | 0.84 | 0.94 | 0.99 | 1.00 | 0.08 | 0.14 | 0.28 | 0.56 | 0.85 | 0.09 | 0.86 | 0.91 | 0.98 | 1.00 |
| 99% | 0.01 | 0.03 | 0.12 | 0.41 | 0.84 | 0.01 | 0.58 | 0.77 | 0.96 | 1.00 | 0.02 | 0.05 | 0.11 | 0.28 | 0.60 | 0.02 | 0.61 | 0.74 | 0.90 | 0.99 |

**T=1000**

| | $a = 0.01$ & $b = 0.1$ | | | | | | | | | | $a = 0.01$ & $b = 0.9$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | |
| $\delta$ | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 |
| 90% | 0.10 | 0.82 | 1.00 | 1.00 | 1.00 | 0.10 | 1.00 | 1.00 | 1.00 | 1.00 | 0.14 | 0.74 | 1.00 | 1.00 | 1.00 | 0.14 | 0.99 | 1.00 | 1.00 | 1.00 |
| 95% | 0.05 | 0.70 | 1.00 | 1.00 | 1.00 | 0.05 | 0.99 | 1.00 | 1.00 | 1.00 | 0.07 | 0.63 | 1.00 | 1.00 | 1.00 | 0.07 | 0.99 | 1.00 | 1.00 | 1.00 |
| 99% | 0.01 | 0.48 | 1.00 | 1.00 | 1.00 | 0.01 | 0.97 | 1.00 | 1.00 | 1.00 | 0.02 | 0.41 | 0.99 | 1.00 | 1.00 | 0.01 | 0.95 | 1.00 | 1.00 | 1.00 |

| | $a = 0.02$ & $b = 0.1$ | | | | | | | | | | $a = 0.02$ & $b = 0.9$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | |
| $\delta$ | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 |
| 90% | 0.10 | 0.81 | 1.00 | 1.00 | 1.00 | 0.10 | 1.00 | 1.00 | 1.00 | 1.00 | 0.17 | 0.69 | 1.00 | 1.00 | 1.00 | 0.17 | 0.99 | 1.00 | 1.00 | 1.00 |
| 95% | 0.04 | 0.69 | 1.00 | 1.00 | 1.00 | 0.04 | 0.99 | 1.00 | 1.00 | 1.00 | 0.09 | 0.58 | 1.00 | 1.00 | 1.00 | 0.09 | 0.98 | 1.00 | 1.00 | 1.00 |
| 99% | 0.01 | 0.47 | 1.00 | 1.00 | 1.00 | 0.01 | 0.97 | 1.00 | 1.00 | 1.00 | 0.02 | 0.37 | 0.98 | 1.00 | 1.00 | 0.02 | 0.94 | 1.00 | 1.00 | 1.00 |

Table 3.4.5: Estimate of the time of change as a percentage of the observation period in the DCC model of Example 3.3.2 when $t^* = T/2$

This table shows statistics - mean, median and standard deviation of detected locations. Estimated change times are computed in percentage.

| | T=100 | | | | | | | | T=1000 | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $a = 0.01$ & $b = 0.1$ | | | | $a = 0.02$ & $b = 0.1$ | | | | $a = 0.01$ & $b = 0.1$ | | | | $a = 0.02$ & $b = 0.1$ | | | |
| $\delta$ | 0.2 | 0.4 | 0.6 | 0.8 | 0.2 | 0.4 | 0.6 | 0.8 | 0.2 | 0.4 | 0.6 | 0.8 | 0.2 | 0.4 | 0.6 | 0.8 |
| Mean | 0.53 | 0.51 | 0.48 | 0.48 | 0.53 | 0.51 | 0.49 | 0.48 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 |
| Median | 0.52 | 0.50 | 0.50 | 0.48 | 0.52 | 0.50 | 0.50 | 0.48 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 |
| SD | 0.18 | 0.15 | 0.09 | 0.06 | 0.18 | 0.15 | 0.10 | 0.06 | 0.09 | 0.03 | 0.01 | 0.01 | 0.09 | 0.03 | 0.01 | 0.01 |
| | $a = 0.01$ & $b = 0.9$ | | | | $a = 0.02$ & $b = 0.9$ | | | | $a = 0.01$ & $b = 0.9$ | | | | $a = 0.02$ & $b = 0.9$ | | | |
| $\delta$ | 0.2 | 0.4 | 0.6 | 0.8 | 0.2 | 0.4 | 0.6 | 0.8 | 0.2 | 0.4 | 0.6 | 0.8 | 0.2 | 0.4 | 0.6 | 0.8 |
| Mean | 0.53 | 0.51 | 0.49 | 0.48 | 0.53 | 0.52 | 0.50 | 0.48 | 0.50 | 0.49 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 |
| Median | 0.51 | 0.50 | 0.48 | 0.48 | 0.51 | 0.50 | 0.49 | 0.49 | 0.49 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 |
| SD | 0.18 | 0.16 | 0.11 | 0.07 | 0.19 | 0.16 | 0.12 | 0.08 | 0.10 | 0.04 | 0.02 | 0.01 | 0.11 | 0.04 | 0.02 | 0.01 |

### 3.4.3 BEKK Model

We use Example 3.3.3 with $p = q = 1$ and $k_0 = 1$. We specify $\mathbf{C}$ by (4.4.4) and let

$$\mathbf{A}_{1,1} = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} \quad \mathbf{B}_{1,1} = \begin{bmatrix} b & 0 \\ 0 & b \end{bmatrix}$$

The choice of $a = 0.01$ represents low ARCH, while $a = 0.02$ stands for high ARCH effect. Similarly, $b$ will be 0.1 (low persistence) or 0.9 (high persistence).

Table 3.4.1 illustrates that the convergence to the limit is slower in the BEKK model than in the other models. However, the BEKK model requires the estimation of larger number of parameters. To achieve good accuracy, one needs larger sample sizes. The power of the tests in the BEKK model is exhibited in Table 3.4.6 and the result of the estimation of $t^*/T$

is sumarized in Table 3.4.7.

Table 3.4.6: Empirical rejection rates for $\hat{M}_T^{(1)}$ and $\hat{M}_T^{(2)}$ in the BEKK model of Example 3.3.3

$\delta$ is the change magnitude, which is the argument in constant coefficient matrix $\mathbf{C}$. Parameters $a$ is the argument in ARCH coefficient matrix $\mathbf{A}$, and $b$ is the argument in GARCH coefficient matrix $\mathbf{B}$. The rejection rate is computed under three significance level, 90%, 95% and 99%.

| **T=100** | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $a = 0.01$ & $b = 0.1$ | | | | | | | | | | $a = 0.01$ & $b = 0.9$ | | | | | | | | | |
| | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | |
| $\delta$ | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 |
| 90% | 0.09 | 0.15 | 0.39 | 0.79 | 0.99 | 0.10 | 0.86 | 0.96 | 1.00 | 1.00 | 0.10 | 0.16 | 0.40 | 0.77 | 0.99 | 0.10 | 0.87 | 0.96 | 1.00 | 1.00 |
| 95% | 0.04 | 0.08 | 0.26 | 0.65 | 0.96 | 0.05 | 0.77 | 0.91 | 0.99 | 1.00 | 0.05 | 0.09 | 0.27 | 0.63 | 0.96 | 0.05 | 0.77 | 0.92 | 0.99 | 1.00 |
| 99% | 0.01 | 0.02 | 0.08 | 0.33 | 0.79 | 0.01 | 0.51 | 0.78 | 0.97 | 1.00 | 0.01 | 0.01 | 0.10 | 0.33 | 0.78 | 0.01 | 0.50 | 0.79 | 0.96 | 1.00 |
| | $a = 0.02$ & $b = 0.1$ | | | | | | | | | | $a = 0.02$ & $b = 0.9$ | | | | | | | | | |
| | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | |
| $\delta$ | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 |
| 90% | 0.09 | 0.15 | 0.39 | 0.79 | 0.99 | 0.09 | 0.86 | 0.96 | 1.00 | 1.00 | 0.10 | 0.16 | 0.40 | 0.77 | 0.99 | 0.10 | 0.87 | 0.96 | 1.00 | 1.00 |
| 95% | 0.04 | 0.08 | 0.26 | 0.65 | 0.96 | 0.04 | 0.77 | 0.91 | 0.99 | 1.00 | 0.04 | 0.09 | 0.27 | 0.63 | 0.96 | 0.05 | 0.77 | 0.95 | 0.99 | 1.00 |
| 99% | 0.01 | 0.02 | 0.08 | 0.33 | 0.78 | 0.01 | 0.50 | 0.78 | 0.97 | 1.00 | 0.01 | 0.02 | 0.10 | 0.33 | 0.78 | 0.01 | 0.50 | 0.79 | 0.96 | 1.00 |
| **T=1000** | | | | | | | | | | | | | | | | | | | | |
| | $a = 0.01$ & $b = 0.1$ | | | | | | | | | | $a = 0.01$ & $b = 0.9$ | | | | | | | | | |
| | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | |
| $\delta$ | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 |
| 90% | 0.09 | 0.78 | 1.00 | 1.00 | 1.00 | 0.10 | 1.00 | 1.00 | 1.00 | 1.00 | 0.09 | 0.78 | 1.00 | 1.00 | 1.00 | 0.11 | 1.00 | 1.00 | 1.00 | 1.00 |
| 95% | 0.04 | 0.71 | 1.00 | 1.00 | 1.00 | 0.05 | 1.00 | 1.00 | 1.00 | 1.00 | 0.05 | 0.69 | 1.00 | 1.00 | 1.00 | 0.05 | 1.00 | 1.00 | 1.00 | 1.00 |
| 99% | 0.01 | 0.51 | 1.00 | 1.00 | 1.00 | 0.01 | 0.98 | 1.00 | 1.00 | 1.00 | 0.01 | 0.49 | 1.00 | 1.00 | 1.00 | 0.01 | 0.97 | 1.00 | 1.00 | 1.00 |
| | $a = 0.02$ & $b = 0.1$ | | | | | | | | | | $a = 0.02$ & $b = 0.9$ | | | | | | | | | |
| | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | |
| $\delta$ | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 |
| 90% | 0.10 | 0.78 | 1.00 | 1.00 | 1.00 | 0.11 | 1.00 | 1.00 | 1.00 | 1.00 | 0.08 | 0.78 | 1.00 | 1.00 | 1.00 | 0.10 | 1.00 | 1.00 | 1.00 | 1.00 |
| 95% | 0.05 | 0.71 | 1.00 | 1.00 | 1.00 | 0.05 | 1.00 | 1.00 | 1.00 | 1.00 | 0.04 | 0.71 | 1.00 | 1.00 | 1.00 | 0.04 | 1.00 | 1.00 | 1.00 | 1.00 |
| 99% | 0.01 | 0.51 | 1.00 | 1.00 | 1.00 | 0.01 | 0.98 | 1.00 | 1.00 | 1.00 | 0.01 | 0.51 | 1.00 | 1.00 | 1.00 | 0.01 | 0.98 | 1.00 | 1.00 | 1.00 |

Table 3.4.7: Estimation of the time of change as a percentage of the observation period in the BEKK model of Example 3.3.3 when $t^* = T/2$

This table shows statistics - mean, median and standard deviation of detected locations. Estimated change times are computed in percentage.

| | T=100 | | | | | | | | T=1000 | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $a = 0.01$ & $b = 0.1$ | | | | $a = 0.02$ & $b = 0.1$ | | | | $a = 0.01$ & $b = 0.1$ | | | | $a = 0.02$ & $b = 0.1$ | | | |
| $\delta$ | 0.2 | 0.4 | 0.6 | 0.8 | 0.2 | 0.4 | 0.6 | 0.8 | 0.2 | 0.4 | 0.6 | 0.8 | 0.2 | 0.4 | 0.6 | 0.8 |
| Mean | 0.53 | 0.51 | 0.49 | 0.48 | 0.53 | 0.52 | 0.49 | 0.48 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.49 | 0.50 | 0.50 |
| Median | 0.52 | 0.50 | 0.49 | 0.49 | 0.52 | 0.50 | 0.49 | 0.49 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 |
| SD | 0.17 | 0.13 | 0.09 | 0.06 | 0.17 | 0.14 | 0.09 | 0.06 | 0.09 | 0.03 | 0.01 | 0.01 | 0.09 | 0.03 | 0.01 | 0.01 |
| | $a = 0.01$ & $b = 0.9$ | | | | $a = 0.02$ & $b = 0.9$ | | | | $a = 0.01$ & $b = 0.9$ | | | | $a = 0.02$ & $b = 0.9$ | | | |
| $\delta$ | 0.2 | 0.4 | 0.6 | 0.8 | 0.2 | 0.4 | 0.6 | 0.8 | 0.2 | 0.4 | 0.6 | 0.8 | 0.2 | 0.4 | 0.6 | 0.8 |
| Mean | 0.53 | 0.51 | 0.49 | 0.48 | 0.53 | 0.51 | 0.49 | 0.48 | 0.50 | 0.49 | 0.50 | 0.50 | 0.50 | 0.50 | 0.49 | 0.50 |
| Median | 0.52 | 0.50 | 0.49 | 0.49 | 0.52 | 0.50 | 0.49 | 0.49 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 |
| SD | 0.18 | 0.14 | 0.10 | 0.06 | 0.18 | 0.14 | 0.10 | 0.06 | 0.09 | 0.03 | 0.01 | 0.01 | 0.09 | 0.03 | 0.02 | 0.01 |

### 3.4.4 Factor Model

Let $p = 2$ in Example 3.3.4. We use $\mathbf{C}$ of (4.4.4) with the same choices of $\delta$ as in Section 3.4.3. Let $\omega = 0.01$, $\alpha_1 = \alpha_2 = 0.01$, $\beta_1 = \beta_2 = 0.1$, $\boldsymbol{\beta}_1 = (0.1, 0.1)^\top$ (low ARCH effect and persistence); $\alpha_1 = \alpha_2 = 0.01$, $\beta_1 = \beta_2 = 0.9$, $\boldsymbol{\beta}_1 = (0.9, 0.9)^\top$ (low ARCH effect and high persistence); $\alpha_1 = \alpha_2 = 0.02$, $\beta_1 = \beta_2 = 0.1$, $\boldsymbol{\beta}_1 = (0.1, 0.1)^\top$ (high ARCH effect and low persistence); and $\alpha_1 = \alpha_2 = 0.02$, $\beta_1 = \beta_2 = 0.9$, $\boldsymbol{\beta}_1 = (0.9, 0.9)^\top$ (high ARCH effect and persistence), factors' conditional variances are generated recursively from $\lambda_0 = 1$. The results are in Tables 3.4.8 and 3.4.9 and they are similar but somewhat better than in the BEKK case. This is not surprising since the factor model can be written in BEKK form.

Table 3.4.8: Empirical rejection rates for $\hat{M}_T^{(1)}$ and $\hat{M}_T^{(2)}$ in the factor model of Example 3.3.4

$\delta$ is the change magnitude, which is the argument in constant coefficient matrix **C**. Parameters $a$ is the argument in ARCH coefficient matrix **A**, and $b$ is the argument in GARCH coefficient matrix **B**. The rejection rate is computed under three significance level, 90%, 95% and 99%.

**T=100**

| | $a = 0.01$ & $b = 0.1$ | | | | | | | | | | $a = 0.01$ & $b = 0.9$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | |
| $\delta$ | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 |
| 90% | 0.10 | 0.16 | 0.43 | 0.78 | 0.99 | 0.09 | 0.86 | 0.96 | 1.00 | 1.00 | 0.09 | 0.17 | 0.38 | 0.75 | 0.98 | 0.09 | 0.87 | 0.96 | 1.00 | 1.00 |
| 95% | 0.05 | 0.09 | 0.30 | 0.66 | 0.96 | 0.05 | 0.76 | 0.91 | 0.99 | 1.00 | 0.04 | 0.09 | 0.25 | 0.61 | 0.95 | 0.05 | 0.77 | 0.91 | 0.99 | 1.00 |
| 99% | 0.01 | 0.03 | 0.11 | 0.35 | 0.80 | 0.02 | 0.52 | 0.78 | 0.96 | 1.00 | 0.01 | 0.02 | 0.09 | 0.33 | 0.75 | 0.01 | 0.55 | 0.77 | 0.96 | 1.00 |

| | $a = 0.02$ & $b = 0.1$ | | | | | | | | | | $a = 0.02$ & $b = 0.9$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | |
| $\delta$ | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 |
| 90% | 0.09 | 0.16 | 0.43 | 0.78 | 0.99 | 0.09 | 0.86 | 0.96 | 1.00 | 1.00 | 0.09 | 0.17 | 0.37 | 0.74 | 0.98 | 0.09 | 0.87 | 0.95 | 0.99 | 1.00 |
| 95% | 0.05 | 0.09 | 0.30 | 0.66 | 0.96 | 0.05 | 0.76 | 0.91 | 0.99 | 1.00 | 0.05 | 0.09 | 0.25 | 0.59 | 0.94 | 0.05 | 0.77 | 0.91 | 0.99 | 1.00 |
| 99% | 0.01 | 0.03 | 0.11 | 0.35 | 0.80 | 0.02 | 0.52 | 0.78 | 0.96 | 1.00 | 0.01 | 0.02 | 0.08 | 0.3172 | 0.72 | 0.01 | 0.54 | 0.77 | 0.96 | 1.00 |

**T=1000**

| | $a = 0.01$ & $b = 0.1$ | | | | | | | | | | $a = 0.01$ & $b = 0.9$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | |
| $\delta$ | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 |
| 90% | 0.10 | 0.79 | 1.00 | 1.00 | 1.00 | 0.10 | 1.00 | 1.00 | 1.00 | 1.00 | 0.09 | 0.76 | 1.00 | 1.00 | 1.00 | 0.10 | 1.00 | 1.00 | 1.00 | 1.00 |
| 95% | 0.06 | 0.70 | 1.00 | 1.00 | 1.00 | 0.05 | 0.99 | 1.00 | 1.00 | 1.00 | 0.05 | 0.68 | 1.00 | 1.00 | 1.00 | 0.05 | 0.99 | 1.00 | 1.00 | 1.00 |
| 99% | 0.01 | 0.49 | 1.00 | 1.00 | 1.00 | 0.01 | 0.98 | 1.00 | 1.00 | 1.00 | 0.01 | 0.46 | 1.00 | 1.00 | 1.00 | 0.01 | 0.97 | 1.00 | 1.00 | 1.00 |

| | $a = 0.02$ & $b = 0.1$ | | | | | | | | | | $a = 0.02$ & $b = 0.9$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | |
| $\delta$ | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 |
| 90% | 0.10 | 0.73 | 1.00 | 1.00 | 1.00 | 0.10 | 0.99 | 1.00 | 1.00 | 1.00 | 0.11 | 0.73 | 1.00 | 1.00 | 1.00 | 0.11 | 0.99 | 1.00 | 1.00 | 1.00 |
| 95% | 0.05 | 0.63 | 1.00 | 1.00 | 1.00 | 0.05 | 0.99 | 1.00 | 1.00 | 1.00 | 0.06 | 0.63 | 1.00 | 1.00 | 1.00 | 0.05 | 0.99 | 1.00 | 1.00 | 1.00 |
| 99% | 0.01 | 0.40 | 1.00 | 1.00 | 1.00 | 0.01 | 0.96 | 1.00 | 1.00 | 1.00 | 0.01 | 0.40 | 1.00 | 1.00 | 1.00 | 0.01 | 0.96 | 1.00 | 1.00 | 1.00 |

Table 3.4.9: Estimation of the time of change as a percentage of the observation period in the factor model of Example 3.3.4 when $t^* = T/2$

This table shows statistics - mean, median and standard deviation of detected locations. Estimated change times are computed in percentage.

| | T=100 | | | | | | | | T=1000 | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $a = 0.01$ & $b = 0.1$ | | | | $a = 0.02$ & $b = 0.1$ | | | | $a = 0.01$ & $b = 0.1$ | | | | $a = 0.02$ & $b = 0.1$ | | | |
| $\delta$ | 0.2 | 0.4 | 0.6 | 0.8 | 0.2 | 0.4 | 0.6 | 0.8 | 0.2 | 0.4 | 0.6 | 0.8 | 0.2 | 0.4 | 0.6 | 0.8 |
| Mean | 0.52 | 0.50 | 0.49 | 0.48 | 0.51 | 0.49 | 0.49 | 0.49 | 0.50 | 0.50 | 0.50 | 0.50 | 0.49 | 0.50 | 0.50 | 0.50 |
| Median | 0.51 | 0.49 | 0.49 | 0.49 | 0.50 | 0.49 | 0.49 | 0.49 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 |
| SD | 0.18 | 0.14 | 0.10 | 0.06 | 0.18 | 0.14 | 0.09 | 0.06 | 0.08 | 0.03 | 0.01 | 0.01 | 0.09 | 0.03 | 0.01 | 0.01 |
| | $a = 0.01$ & $b = 0.9$ | | | | $a = 0.02$ & $b = 0.9$ | | | | $a = 0.01$ & $b = 0.9$ | | | | $a = 0.02$ & $b = 0.9$ | | | |
| $\delta$ | 0.2 | 0.4 | 0.6 | 0.8 | 0.2 | 0.4 | 0.6 | 0.8 | 0.2 | 0.4 | 0.6 | 0.8 | 0.2 | 0.4 | 0.6 | 0.8 |
| Mean | 0.52 | 0.49 | 0.48 | 0.48 | 0.52 | 0.49 | 0.48 | 0.48 | 0.50 | 0.50 | 0.50 | 0.50 | 0.49 | 0.50 | 0.50 | 0.50 |
| Median | 0.52 | 0.49 | 0.49 | 0.49 | 0.51 | 0.49 | 0.49 | 0.49 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 |
| SD | 0.18 | 0.15 | 0.10 | 0.06 | 0.18 | 0.15 | 0.10 | 0.06 | 0.09 | 0.03 | 0.01 | 0.06 | 0.10 | 0.04 | 0.01 | 0.01 |

### 3.4.5 ADCC Model

We use the asymmetric DCC model of (3.3.13) with $p = q = 1$, $\theta_1 = 1 - \theta_2 - \theta_3 - \theta_4$, $\theta_2 = a$, $\theta_3 = a/2$, $\theta_4 = b$ and $\mathbf{C}$ is from 4.4.4. We consider four types of $a$ and $b$ for quasi conditional correlation process: $a = 0.01$, $b = 0.1$ (low ARCH and persistence); $a = 0.01$, $b = 0.9$ (low ARCH and high persistence); $a = 0.02$, $b = 0.1$ (high ARCH and low persistence); $a = 0.02$, $b = 0.9$ (high ARCH and persistence). The matrices $\mathbf{c}$, $\mathbf{A}_1$ and $\mathbf{B}_1$, determining conditional variances, are same with used in Section 3.4.1 and 3.4.2, where $\mathbf{c} = (0.01, 0.01)^\top$, $a = 0.01$ and $b = 0.7$. The results are exhibited in Tables 3.4.10 and 3.4.11 and similar to the outcome of the Monte Carlo experiment for the DCC model in Section 3.4.2.

Table 3.4.10: Empirical rejection rates for $\hat{M}_T^{(1)}$ and $\hat{M}_T^{(2)}$ in the asymmetric DCC model of Example 3.3.5

$\delta$ is the change magnitude, which is the argument in constant coefficient matrix **C**. Parameters $a$ is the argument in ARCH coefficient matrix **A**, and $b$ is the argument in GARCH coefficient matrix **B**. The rejection rate is computed under three significance level, 90%, 95% and 99%.

| **T=100** | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $a=0.01$ & $b=0.1$ | | | | | | | | | | $a=0.01$ & $b=0.9$ | | | | | | | | | |
| | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | |
| $\delta$ | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 |
| 90% | 0.11 | 0.19 | 0.45 | 0.82 | 0.99 | 0.12 | 0.90 | 0.97 | 1.00 | 1.00 | 0.14 | 0.20 | 0.40 | 0.73 | 0.95 | 0.14 | 0.92 | 0.96 | 1.00 | 1.00 |
| 95% | 0.06 | 0.12 | 0.32 | 0.74 | 0.97 | 0.06 | 0.83 | 0.94 | 1.00 | 1.00 | 0.08 | 0.14 | 0.30 | 0.63 | 0.92 | 0.07 | 0.84 | 0.92 | 0.99 | 1.00 |
| 99% | 0.01 | 0.03 | 0.11 | 0.42 | 0.85 | 0.01 | 0.56 | 0.80 | 0.97 | 1.00 | 0.02 | 0.04 | 0.11 | 0.33 | 0.69 | 0.02 | 0.59 | 0.77 | 0.94 | 0.99 |
| | $a=0.02$ & $b=0.1$ | | | | | | | | | | $a=0.02$ & $b=0.9$ | | | | | | | | | |
| | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | |
| $\delta$ | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 |
| 90% | 0.13 | 0.19 | 0.44 | 0.82 | 0.99 | 0.13 | 0.91 | 0.97 | 1.00 | 1.00 | 0.14 | 0.21 | 0.38 | 0.68 | 0.92 | 0.16 | 0.92 | 0.96 | 0.99 | 1.00 |
| 95% | 0.07 | 0.12 | 0.32 | 0.73 | 0.97 | 0.07 | 0.83 | 0.93 | 1.00 | 1.00 | 0.09 | 0.14 | 0.28 | 0.57 | 0.85 | 0.09 | 0.84 | 0.91 | 0.97 | 1.00 |
| 99% | 0.01 | 0.03 | 0.11 | 0.42 | 0.84 | 0.02 | 0.56 | 0.80 | 0.97 | 1.00 | 0.02 | 0.04 | 0.09 | 0.29 | 0.61 | 0.03 | 0.59 | 0.75 | 0.91 | 0.99 |
| **T=1000** | | | | | | | | | | | | | | | | | | | | |
| | $a=0.01$ & $b=0.1$ | | | | | | | | | | $a=0.01$ & $b=0.9$ | | | | | | | | | |
| | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | |
| $\delta$ | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 |
| 90% | 0.10 | 0.81 | 1.00 | 1.00 | 1.00 | 0.10 | 1.00 | 1.00 | 1.00 | 1.00 | 0.14 | 0.72 | 1.00 | 1.00 | 1.00 | 0.13 | 1.00 | 1.00 | 1.00 | 1.00 |
| 95% | 0.05 | 0.71 | 1.00 | 1.00 | 1.00 | 0.05 | 0.99 | 1.00 | 1.00 | 1.00 | 0.07 | 0.61 | 1.00 | 1.00 | 1.00 | 0.06 | 0.99 | 1.00 | 1.00 | 1.00 |
| 99% | 0.02 | 0.50 | 1.00 | 1.00 | 1.00 | 0.01 | 0.98 | 1.00 | 1.00 | 1.00 | 0.02 | 0.43 | 0.99 | 1.00 | 1.00 | 0.01 | 0.96 | 1.00 | 1.00 | 1.00 |
| | $a=0.02$ & $b=0.1$ | | | | | | | | | | $a=0.02$ & $b=0.9$ | | | | | | | | | |
| | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | | $\hat{M}_T^{(1)}$ | | | | | $\hat{M}_T^{(2)}$ | | | | |
| $\delta$ | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 0 | 0.2 | 0.4 | 0.6 | 0.8 |
| 90% | 0.11 | 0.80 | 1.00 | 1.00 | 1.00 | 0.11 | 1.00 | 1.00 | 1.00 | 1.00 | 0.16 | 0.67 | 1.00 | 1.00 | 1.00 | 0.16 | 1.00 | 1.00 | 1.00 | 1.00 |
| 95% | 0.06 | 0.69 | 1.00 | 1.00 | 1.00 | 0.06 | 0.99 | 1.00 | 1.00 | 1.00 | 0.09 | 0.56 | 1.00 | 1.00 | 1.00 | 0.08 | 0.99 | 1.00 | 1.00 | 1.00 |
| 99% | 0.02 | 0.49 | 1.00 | 1.00 | 1.00 | 0.02 | 0.97 | 1.00 | 1.00 | 1.00 | 0.03 | 0.38 | 0.98 | 1.00 | 1.00 | 0.03 | 0.94 | 1.00 | 1.00 | 1.00 |

Table 3.4.11: Estimation of the time of change as a percentage of the observation period in the asymmetric DCC model of Example 3.3.5 when $t^* = T/2$

This table shows statistics - mean, median and standard deviation of detected locations. Estimated change times are computed in percentage.

| | T=100 | | | | | | | | T=1000 | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $a = 0.01$ & $b = 0.1$ | | | | $a = 0.02$ & $b = 0.1$ | | | | $a = 0.01$ & $b = 0.1$ | | | | $a = 0.02$ & $b = 0.1$ | | | |
| $\delta$ | 0.2 | 0.4 | 0.6 | 0.8 | 0.2 | 0.4 | 0.6 | 0.8 | 0.2 | 0.4 | 0.6 | 0.8 | 0.2 | 0.4 | 0.6 | 0.8 |
| Mean | 0.53 | 0.50 | 0.48 | 0.48 | 0.53 | 0.50 | 0.48 | 0.48 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 |
| Median | 0.52 | 0.50 | 0.50 | 0.48 | 0.52 | 0.50 | 0.50 | 0.48 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 |
| SD | 0.18 | 0.14 | 0.09 | 0.06 | 0.18 | 0.14 | 0.09 | 0.06 | 0.09 | 0.03 | 0.01 | 0.01 | 0.09 | 0.03 | 0.01 | 0.01 |
| | $a = 0.01$ & $b = 0.9$ | | | | $a = 0.02$ & $b = 0.9$ | | | | $a = 0.01$ & $b = 0.9$ | | | | $a = 0.02$ & $b = 0.9$ | | | |
| $\delta$ | 0.2 | 0.4 | 0.6 | 0.8 | 0.2 | 0.4 | 0.6 | 0.8 | 0.2 | 0.4 | 0.6 | 0.8 | 0.2 | 0.4 | 0.6 | 0.8 |
| Mean | 0.53 | 0.51 | 0.49 | 0.48 | 0.53 | 0.51 | 0.49 | 0.48 | 0.50 | 0.50 | 0.50 | 0.50 | 0.51 | 0.50 | 0.50 | 0.50 |
| Median | 0.51 | 0.50 | 0.48 | 0.48 | 0.51 | 0.50 | 0.48 | 0.48 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 |
| SD | 0.19 | 0.16 | 0.11 | 0.07 | 0.19 | 0.16 | 0.12 | 0.08 | 0.11 | 0.04 | 0.02 | 0.01 | 0.11 | 0.05 | 0.02 | 0.01 |

## 3.5 Empirical Application: Testing for financial contagion

Forbes and Rigobon (2002) defined that financial contagion effect occurs if the inter-linkages across markets experienced a significant increase after some market events. To investigate the contagion, it is crucial to assess the dynamic evolution on conditional correlations during the period interest. As a standard and successful model for conditional correlations, the M-GARCH family naturally becomes a tool to explain contagions. If the contagion effect existed, a significant increase should be found in the conditional correlations after break dates. The actual change date is hidden and needs to be detected through statistical methods.

In the recent literature one could find analysis and detection of contagion by change point methods. Dimitriou et al. (2013) used a Markov regime switching model to find endogenous date, and analyzed the contagion effect among big five emerging markets. Dungey et al.

(2015) used a structural GARCH model with embedded smooth transition functions to identify the accurate start and finish date of the financial crisis, thereby finding contagions during the crises. Meanwhile, Blatt et al. (2015) applied a parameters change point detection technique (Qu and Perron, 2007) for multivariate system to derive a sequential procedure to find endogenous dates of changes. Our paper contributes this literature with an efficient method of testing and detection using the semi-parametric tests $\hat{M}_T(1)$ and $\hat{M}_T^{(2)}$ to find contagions and estimate their times during the great recession between 2006 and 2010 in global financial markets.

We collect stock market price indexes from emerging markets in three regions: six Latin American markets including Argentina (Argentina MERVAL index), Brazil (Brazil BOVESPA index), Chile (Chile Santiago SE General index), Mexico (Mexico IPC index), Colombia (Colombia IGBC index), Peru (BVL General index); seven Central East European (CEE hereafter) markets including Czech (Prague SEPX index), Estonia (OMX Tallin index), Hungary (Budapest index), Poland (Warsaw General index), Romania (Romania BET index), Slovakia (Slovakia SAX 16 index), Slovenia (Slovenian blue chip index); ten East Asian markets including HongKong (Hang Seng index), Indonesia (IDX composite index), Japan (Nikkei 225 index), South Korea (Korea SE composite index), Malaysia (Malaysia KLCI index), Philippines (Philippine SE index), Singapore (Straits Times index), Taiwan (Taiwan SE weighted index), Thailand (Bangkok S.E.T index), China Mainland(Shanghai S.E. A share index); and the United States (S&P 500 index). The US, as the eye of storm, experienced several financial events, such as sub-prime crisis and the great recession and her relation to the emerging markets changed during the observation period. The data are taken from the Datastream database, and ranged from 1 September, 2006 to 1 September, 2010. For each price index, we calculate its log return process to achieve mean stationarity. We then split the return data into three sub-sets: Latin, CEE and East Asian markets, and in each sub-sets, we consider U.S. market as the source of the global financial contagion effect. In order to find change points in correlation structures of these three data sets, we compute the $\hat{M}_T(1)$ and $\hat{M}_T(2)$ based on the different multivariate volatility models of Examples 3.3.1 - 3.3.5. The critical values used in the analysis are taken from Aue et al. (2009). We find three change points in each sub-set by the binary segmentation method. We use $\hat{M}_T(2)$ to test the significance of the change points we detected, i.e. we test if on the interval $[\hat{t}_1, \hat{t}_{i+1})$ a change point is located with $\hat{t}_0, \hat{t}_4 = T$ and $\hat{t}_i,\ i = 1, 2, 3$ are determined by the binary segmentation. The detection results are documented in Table 3.5.1.

Remarkably, all tests report roughly consistent changing locations for these three regions. The results show that the first change is dated around August 2007 (ceasing activities in the US mortgage debt market), the second changes happens mainly close to September 2008 (the bankruptcy of Lehman Brothers), and the third change occurred after April 2009 (bailout decision made by G20 summit). Basically, these three dates would split the whole sample into four periods: boom, bubble, bust and recovery. More specifically, we can see that Latin American markets react relatively faster than others, particularly in case of the last change. This indicates that the markets in the American region are more closely integrated, and any good news from U.S. financial system produce optimistic expectation for Latin American. With regard to the CEE markets, their reaction dates close to the events mentioned above, but the second change is prior to the events. This reveals that the capital flow from the US has a relatively sensitive impact on this region. Lastly, compared with the two regions mentioned above, the East Asian markets tend to have relatively higher resistance to the US market crash. This remark is supported by results from the estimated correlations in Table 3.5.2. The higher resistance might due to the fact that emerging markets in this region are more correlated with the large markets in this region, such as Japan, Hong Kong and Mainland China.

Table 3.5.1: The estimated change points in the global regions ($*$,$**$ and $***$ indicate significance at 90%, 95% and 99% levels, respectively)

| | CCC | DCC | BEKK | Factor | ADCC |
|---|---|---|---|---|---|
| | | | Latin American Markets | | |
| 1st | 30/Oct/2007 | 30/Oct/2007 | 30/Jul/2007 | 27/Jul/2007 | 26/Oct/2007 |
| Change | (12.71***) | (6.94***) | (9.16***) | (16.15***) | (7.13***) |
| 2nd | 15/Sep/2008 | 15/Sep/2008 | 15/Sep/2008 | 05/Sep/2008 | 15/Sep/2008 |
| Change | (10.26***) | (10.26***) | (9.31***) | (16.27***) | (10.55***) |
| 3rd | 22/May/2009 | 22/May/2009 | 18/Mar/2009 | 20/Mar/2009 | 20/May/2009 |
| Change | (6.66***) | (6.66***) | (7.07***) | (9.41***) | (6.69***) |
| | | | Central East European Markets | | |
| 1st | 22/Jun/2007 | 24/Jun/2007 | 10/Jul/2007 | 24/Jul/2007 | 06/Jul/2007 |
| Change | (10.04***) | (10.04***) | (14.54***) | (17.94***) | (10.68***) |
| 2nd | 04/Jul/2008 | 04/Jul/2008 | 26/Jun/2008 | 26/Aug/2008 | 02/Jun/2008 |
| Change | (10.81***) | (10.81***) | (11.84***) | (16.32***) | (10.99***) |
| 3rd | 02/Jun/2009 | 02/Jun/2009 | 01/Jun/2009 | 01/Jun/2009 | 02/Nov/2009 |
| Change | (8.70***) | (8.70***) | (12.53***) | (22.95***) | (7.46***) |
| | | | East Asian Markets | | |
| 1st | 12/Feb/2008 | 12/Feb/2008 | 13/Mar/2008 | 18/Jan/2008 | 19/Feb/2008 |
| Change | (16.31***) | (16.31***) | (51.40***) | (27.32***) | (16.64***) |
| 2nd | 16/Oct/2008 | 16/Oct/2008 | 17/Sep/2008 | 12/Sep/2008 | 12/Sep/2008 |
| Change | (13.80***) | (25.17***) | (20.56***) | (30.95***) | (20.67***) |
| 3rd | 09/Oct/2009 | 09/Oct/2009 | 05/Oct/2009 | 10/Apr/2009 | 08/Oct/2009 |
| Change | (12.05***) | (12.18***) | (14.59***) | (27.34***) | (13.32***) |

Table 3.5.2: The regional correlation levels and correlation levels with the U.S. market between 2006 and 2010

This table reports the estimated correlation concentration level, and the level of estimated regional correlation with U.S. market, denoting as $\bar{\delta}$ and $\delta_{US}$, respectively.

**Latin American Emerging Markets**

| | CCC | | DCC | | BEKK | | Factor | | ADCC | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\bar{\delta}$ | $\bar{\delta}_{US}$ | $\bar{\delta}$ | $\bar{\delta}_{US}$ | $\bar{\delta}$ | $\bar{\delta}_{US}$ | $\bar{\delta}$ | $\bar{\delta}_{US}$ | $\bar{\delta}$ | $\bar{\delta}_{US}$ |
| Phase 1 | 0.48 | 0.54 | 0.48 | 0.54 | 0.29 | 0.26 | 0.46 | 0.53 | 0.44 | 0.53 |
| Phase 2 | 0.44 | 0.43 | 0.44 | 0.43 | 0.30 | 0.25 | 0.53 | 0.46 | 0.48 | 0.46 |
| Phase 3 | 0.64 | 0.65 | 0.64 | 0.65 | 0.33 | 0.37 | 0.64 | 0.62 | 0.64 | 0.65 |
| Phase 4 | 0.57 | 0.63 | 0.58 | 0.64 | 0.35 | 0.34 | 0.61 | 0.68 | 0.57 | 0.64 |

**Central East European Emerging Markets**

| | CCC | | DCC | | BEKK | | Factor | | ADCC | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\bar{\delta}$ | $\bar{\delta}_{US}$ | $\bar{\delta}$ | $\bar{\delta}_{US}$ | $\bar{\delta}$ | $\bar{\delta}_{US}$ | $\bar{\delta}$ | $\bar{\delta}_{US}$ | $\bar{\delta}$ | $\bar{\delta}_{US}$ |
| Phase 1 | 0.17 | 0.16 | 0.17 | 0.16 | 0.13 | 0.07 | 0.22 | 0.19 | 0.17 | 0.16 |
| Phase 2 | 0.21 | 0.13 | 0.21 | 0.13 | 0.16 | 0.08 | 0.24 | 0.17 | 0.20 | 0.13 |
| Phase 3 | 0.31 | 0.24 | 0.31 | 0.24 | 0.19 | 0.11 | 0.38 | 0.28 | 0.30 | 0.25 |
| Phase 4 | 0.30 | 0.28 | 0.30 | 0.29 | 0.16 | 0.08 | 0.33 | 0.29 | 0.30 | 0.28 |

**East Asian Emerging Markets**

| | CCC | | DCC | | BEKK | | Factor | | ADCC | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\bar{\delta}$ | $\bar{\delta}_{US}$ | $\bar{\delta}$ | $\bar{\delta}_{US}$ | $\bar{\delta}$ | $\bar{\delta}_{US}$ | $\bar{\delta}$ | $\bar{\delta}_{US}$ | $\bar{\delta}$ | $\bar{\delta}_{US}$ |
| Phase 1 | 0.39 | 0.09 | 0.40 | 0.09 | 0.20 | 0.01 | 0.43 | 0.08 | 0.40 | 0.08 |
| Phase 2 | 0.40 | 0.08 | 0.40 | 0.08 | 0.19 | 0.04 | 0.40 | 0.07 | 0.35 | 0.07 |
| Phase 3 | 0.44 | 0.20 | 0.44 | 0.20 | 0.18 | 0.07 | 0.52 | 0.22 | 0.46 | 0.19 |
| Phase 4 | 0.41 | 0.18 | 0.41 | 0.18 | 0.20 | 0.06 | 0.44 | 0.16 | 0.41 | 0.18 |

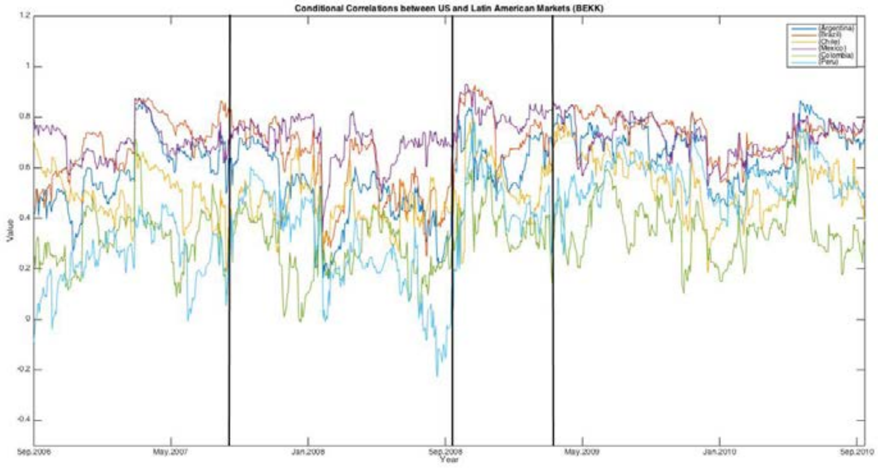
Once the change points are detected, we then use the models in Examples 3.3.1 - 3.3.5 to estimate the relevant parameters in each phase. If there existed an increase on the correlation level between U.S. and emerging markets, we identify it as a contagion effect from U.S. to emerging markets. We use the (empirical) correlation matrix to measure correlation levels. The regional correlation level, also known as the regional integration level, is measured by averaging off diagonal elements of the (empirical) correlation matrix and it is denoted by $\bar{\delta}$. The regional correlation level with U.S. market, denoted as $\bar{\delta}_{US}$, is measured by averaging elements related to U.S. in the (empirical) correlation matrix. Table 3.5.2 reports the results, showing that the contagion effect significantly exists among these three regions, especially after the second change date in 2008.

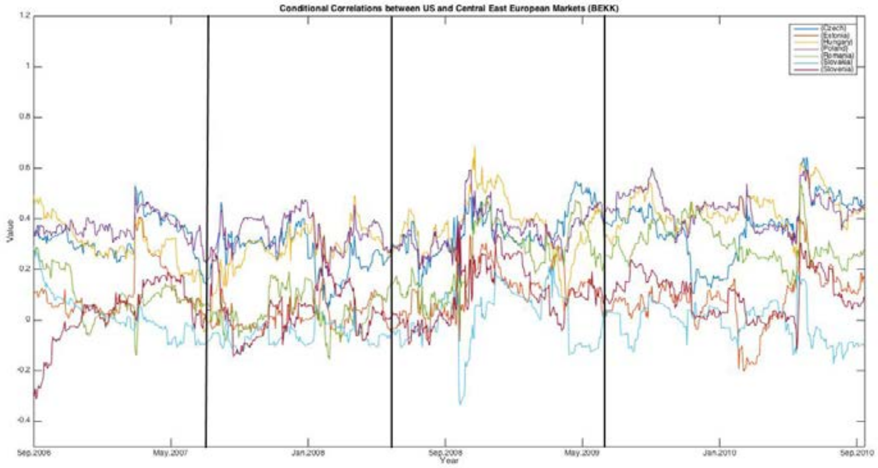
The estimators are consistent for most volatility models, except for the BEKK model which gives lower correlation. Nevertheless, all estimates indicate same features. Firstly, in case of regional integration level, the Latin American and East Asian regions are more concentrated than the CEE area, with $\bar{\delta}$ around 0.4 before the crisis in first two regions, while this value equals 0.2 in CEE markets. This result may be due to that the market liberalization in Central East European emerging markets is not as high as the other two regions, and the CEE markets are more influenced by other developed market in this region, for example, by Germany.

Secondly, regarding the regional linkage with U.S. market, the Latin American region has closest relation to the U.S. market, since before the crisis, $\bar{\delta}_{US}$ is around 0.5 for the Latin American, followed by CEE area with $\bar{\delta}_{US} = 0.15$, and East Asian area shows the lowest correlation with the U.S. market, barely reaching 0.1. Apparently, the U.S. market has weaker influence on the East Asian region. This remark remains true after the crisis.

Thirdly, the evolution of the correlation in the three regions experiences similar patterns: the regional integration level keeps increasing before a decrease in the last phase; the regional linkages with U.S. climb to a peak after September 2008 in all three regions, and then decrease slightly in the last phase for Latin American and East Asian region, except in the CEE markets where the correlation keeps rising. These results show that contagion effects are significant in all data sets, particularly the transmission from U.S. market, and its impact to the Latin American and the CEE regions are more severe than in the East Asian region. To dynamically investigate the contagion effect, we estimate conditional correlation in each region with the U.S. market in all five volatility models discussed in this paper. Since the simulation results show that the BEKK and the DCC model gave good finite sample performances when our semi-parametric CUSUM tests are used, we exhibit the conditional correlations obtained from the BEKK and the DCC models, as an example, in Figure 3.5.1 and 3.5.2. Our analysis is based on the scalar DCC model of (1.2.22). In each region, the estimators for the times of the changes split the conditional correlations into four phases, visually presenting the results.

Figure 3.5.1: The BEKK Conditional Correlations in Three Regions



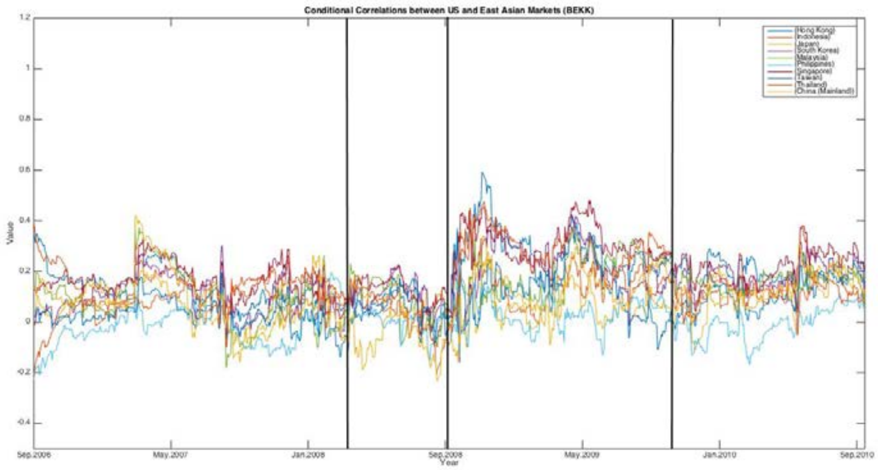
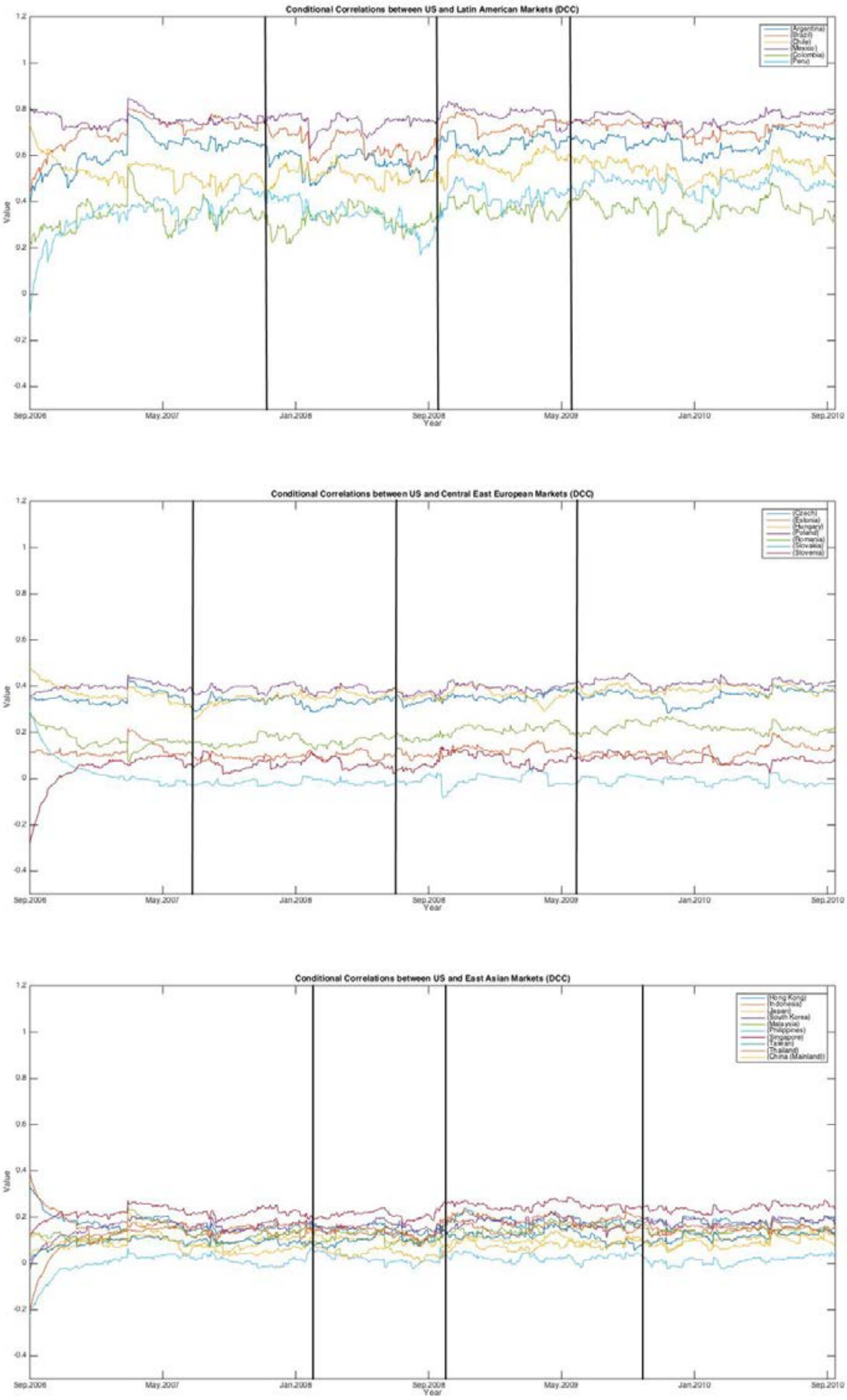



Figure 3.5.2: The DCC Conditional Correlations in Three Regions



## 3.6 Summary

This paper extends the CUSUM change point tests on covariance structure proposed by Aue et al. (2009) to more flexible environments. The CUSUM test can be properly applied to detect changes in dynamically evolving correlation structures modeled by multivariate variance models. Thus, the CUSUM test can detect correlation changes in a system of de-volatized processes estimated by several Multivariate GARCH mdoels. Benefited from asymptotic theory derived by Aue et al. (2009), the limit of CUSUM tests is valid with most of stationary multivariate processes, we discussed examples containing CCC, BEKK, Factor-GARCH, DCC and ADCC models.

Our simulations showed that de-volatizing process by M-GARCH models results good behavior on CUSUM tests. The $\hat{M}_T^{(2)}$ test show better power than $\hat{M}_T^{(1)}$ test. In the term of a middle change, the power of tests is positively correlated with the magnitude of change. Even though, the fact of strong mixing property deteriorates the size and power of CUSUM tests, the simulation found that the BEKK model can relatively overcome the distortion. While as the most fashionable model in financial application, the sustaining tolerance ability of the DCC typed CUSUM tests are relatively lower.

Finally, we analyzed global financial contagion effect with endogenous change locations, we apply the CUSUM tests on stock indexes modeled with BEKK, Factor, DCC and ADCC models. The findings show that contagion effect significantly existed during the great recession in Latin American, Central East European and East Asian markets.

# Chapter 4

# Weighted CUSUM Tests for the Dynamic Correlation Structures

## 4.1 Introduction

Modelling the multivariate covariance structure has been widely studied in the literature, and studying covariance breaks also showed its importances in many financial applications such as investigating financial contagion, measuring systemic risk and managing assets portfolios. Although the family of multivariate volatility models has developed rapidly during last decades fitting all sort of characteristics present in financial data (most contributions are due to R. Engle and T. Bollerslev among others), there is still little research on how in-sample fitting is deteriorated by the occurrence of breaks in the covariance/correlation structure. Hence, testing instability in the covariance/correlation structure has become a more interesting topic in recent years. Related contributions on testing covariance stability have been made by Andreou and Ghysels (2003) and Qu and Perron (2007) who have proposed parametric tests whilst Aue et al. (2009) have been proposing non-parametric tests.

In particular, non-parametric CUSUM tests, benefit from the efficiency in computation, developed further since Aue et al. (2009) derived the asymptotic theory of CUSUM test in the context of covariance. Wied et al. (2012) extended the CUSUM tests to sample correlations with mixing or near-epoch dependence. In Chapter 3 we proposed a semi-parametric CUSUM tests for data allowing dynamically evolving correlation structures.

Nevertheless, self-normalised change-point tests naturally suffer from a common limitation, that is as Andrews (1993) pointed out, that the supremum statistics are always applied to a trimmed proportion of the data, and for this reason, the tests lose power if a break occurred close to either ends of the sample ($T - t^* \to T$ and $t^* \to T$, $T$ is the sample size and $1 < t^* < T$ is the break location).

In order to overcome this limitation, a recent work by Kao et al. (2016), assessed the eigensystem to examine instabilities in the eigenvectors and eigenvalues. Using the strong

Invariance Principal Components for a CUSUM statistics, their tests provide better power to detect early and later change-points in the covariance structure. In this chapter, following Chapter 3, we propose a weighted approximation of the semi-parametric CUSUM statistics to detect changes in the tails of covariance/correlation structures.

To empirically apply these tests in the field of economics or finance, the CUSUM tests provide good detection power if we have a basic conjecture about market status, because we can always collect data over the suspected period so as to place potential changes in the centre of sample. However, we still take a risk of biases in our conjecture, i.e. the tests may omit changes in either ends of the sample. Also, it is not allowed to roll sample observations under certain scenarios. For instance, the CUSUM tests are unreliable to detect the most recent market changes because future data are unavailable; or for event studies, it is common to have a fixed investigation window. Therefore, studying the instability of covariance/correlation structure over the whole sample period is crucial to empirical studies.

Section 4.2 reviews related literature. Section 4.3 introduces two types of potential weight functions, and outlines the weighted CUSUM tests by deriving the asymptotic distributions. In section 4.4, we conduct a simulation study and compare semi-parametric CUSUM with weighted CUSUM tests. Section 4.5 provides an application on detecting systemic events in the U.S. equity market, a summary concludes.

## 4.2 Literature Review

The instability in the covariance/correlation structure has drawn increasing attention since Aue et al. (2009) proposed a well-constructed CUSUM test to detect unknown changes in the covariance structure. The recent survey paper by Aue et al. (2013) and Horváth and Rice (2014) reviewed the development of literature, also Horváth and Rice (2014) considered an averaged CUSUM statistics as a generalization. As mentioned in its discussion (Trapani, 2014), the self-normalised CUSUM retains its power in the trimmed sample $[t_1, t_2]$ (where $t_1$ and $t_2$ are trimming points between observations $[1, T]$) and the test loses power beyond this interval. This problem was initially raised by Andrews (1993), who introduced the weak invariance principle for this issue, and then examined by Andrews (2003). Actually, most self-normalised change-point tests suffer from the trimmed-sample issue, including Wald, Likelihood-ratio or the CUSUM tests (see Perron, 2006).

Because the trimmed-sample issue is originally caused by unbalanced weights in statistics, weight functions can help to overcome such limitation. Szyszkowicz (1994) studied three types of weight functions and derived the asymptotic behavior of these weighted empirical processes. We refer to Csörgő and Horváth (1997) for the basic assumptions and properties of an appropriate weight function, as they already demonstrated that the appropriate weight function should pass what is known as the integral test. The integral test guarantees that the weight function selected is finite, thereby resulting in the convergence of the statistics of change-point test. One weight function $[\mathfrak{t}(1-\mathfrak{t})]^{\alpha}$, for $0 < \mathfrak{t} < 1$ and $0 < \alpha < 1/2$, is widely used in applications.

Later literature has concentrated on developing the weighted approximation of change-point tests. However, theoretically, the weighted approximation for a change-point test is not simply weighting everything on its limit, it changes the converge rate in some cases. Horváth and Rice (2014) discussed the weighted approximation for both self-normalised CUSUM statistics and selected self-normalised CUSUM statistics. They asserted that it is easy to find the limit of the weighted self-normalised CUSUM, while the limit of maximally selected self-normalised CUSUM is hard to define, as the convergence rate can be very slow. By this, they mean that it is necessary to have a very large sample to approximate the distribution of selected self-normalised CUSUM statistics.

To find the limit with slow convergence rate, Darling and Erdös (1957) proved a limit theorem through a technique called Darling and Erdös law. For the details of Darling and Erdös law, we refer to Csörgő and Horváth (1997). Another contribution on the slow convergence is made by Berkes et al. (2004), and they showed the permutation re-sampling provides better convergence rate than the Darling and Erdös law.

With theoretical supports, the weighted approximation is valid for many change-point tests, although with slow rate of convergence. For instance, Hidalgo and Seo (2013) extended the likelihood ratio test proposed by Horváth (1993) to detect changes in the slopes of a regression model without the restriction of "middle" change. Horváth et al. (2016) applied Darling and Erdös law to find the convergence rate of a weighted test detecting regression coefficient multiple changes. To detect changes in the second moment, Gombay et al. (1996) derived the weighted approximation of weighted maximum likelihood test to detect variance change. More recently, Aue et al. (2006) applied weighted CUSUM statistics to detect untrimmed change locations in augmented GARCH processes. Berkes et al. (2009) provided the weighted approximation for the weighted CUSUM test to detect autocorrelation changes in linear processes. For detecting covariance changes, Kao et al (2016) used weak invariance principles to strong invariance principles, and derived the limit of a statistics formed by eigenvectors extracted from covariance matrix, and then used weighted CUSUM tests to detect changes in each of the eigenvectors.

This Chapter aims to extend the semi-parametric CUSUM tests (Chapter 3) to detect correlation changes close to either ends, through the weighted approximation method summarized in Csörgő and Horváth (1997). We applied weight functions mentioned in Szyszkowicz (1994) to the semi-parametric CUSUM test, and then detect changes in the dynamic evolving correlation structure.

Note that, there always exists a trade-off by adding weight function to a change-point test, because a weight function tends to give excessive weight to either ends but to relatively neglect the centre. As a possible solution, Orasch and Pouliot (2004) split the sample into "centre" and "ends", and then applied weighted and un-weighted CUSUM tests to detect mean changes in split sub-samples separately. The trimming parameter can be obtained by solving the inequation between weighted and unweighted functions. However, for reasons concerning the property of continuity and the issue of obtaining unique critical values, we do not combine different typed CUSUM tests in one sample. Our contribution concentrates on proposing weighted CUSUM tests to detect unknown correlation changes with asymptotic unit power across the whole sample. We also discuss the limiting distribution of weighted CUSUM tests and tabulate the relative critical values.

## 4.3 The Semi-parametric Weighted CUSUM Test

In this section, we describe the proposed semi-parametric weighted CUSUM tests. Following the existing literature (Aue et al. 2009; Wied et al. 2012), this chapter tests the stability of dynamic evolving correlation structure of a $d$-dimensional random vector $\mathbf{y}_1$, $\mathbf{y}_2$, ..., $\mathbf{y}_T$, where $\mathbf{y}_t = (y_t(1), y_t(2), \ldots, y_t(d))^\top$ $1 \le t \le T$. $\mathbf{y}_t$ satisfies

$$E[\mathbf{y}_t] = \mathbf{0}$$

and

$$E[|\mathbf{y}_t|^2] < \infty$$

where $|\cdot|$ denotes the Euclidean norm in $\mathbb{R}^d$.

It is customary to use multivariate volatility models to express the vector $\mathbf{y}_t$ such that:

$$\mathbf{y}_t = \mathbf{\Sigma}_t^{\frac{1}{2}} \mathbf{e}_t \tag{4.3.1}$$

**Assumption 4.3.1.** *The innovation* $\mathbf{e}_t$ *is independent and identically distributed, and conditional covariance* $\mathbf{\Sigma}_t$ *is stationary, ergodic and* $\beta$*-mixing.*

To study the dynamics of the correlation structure, it is necessary to obtain de-volatized data $\hat{\mathbf{y}}_t$ by filtering realized conditional variances $\hat{\tau}_t(j)$ for all $1 < j < d$.

**Assumption 4.3.2.** *The estimator* $\hat{\tau}_t(j)$ *is a consistent estimator of* $\tau_t(j)$ *obtained by considering dependence between all coordinates.*

Assumption 4.3.2 is a fairly standard one and can be fulfilled by consistently estimating the relevant coefficients with multivariate GARCH models (See Francq and Zakoian, 2010). This is proved in A.1. We used the dynamic conditional model correlations (DCC) (see Engle, 2002) in the empirical sections. Then, we use the obtained estimator of the variance to get the de-volatized data $\hat{y}_t^*(j)$ as:

$$\hat{y}_t^*(j) = \frac{y_t(j)}{\hat{\tau}_t(j)}, \quad 1 \leq t \leq T, 1 \leq j \leq d \tag{4.3.2}$$

The conditional covariance matrix $\mathbf{\Sigma}_t^*$ of $\mathbf{y}_t^* = (y_t^*(1), y_t^*(2), \ldots, y_t^*(d))^\top$ in this way reduces to the conditional correlation matrix, composed by off-diagonal pair-wise conditional correlation terms, denoted as $\rho_t(i,j)$ for entries $i$ and $j$. Our test aims is to distinguish the dynamics of $\rho_t(i,j)$ under the null hypothesis of no change in the correlation structure from the alternative, such as in:

$$H_0\colon\ \rho_1(i,j) = \rho_2(i,j) = \cdots = \rho_T(i,j) \text{ for all } 1 \leq i,j \leq d$$

versus

$$H_A\colon \text{ there are } 1 < t^* < T \text{ and } 1 \leq i,j \leq d \text{ such that}$$
$$\rho_1(i,j) = \rho_2(i,j) = \cdots = \rho_{t^*}(i,j) \neq \rho_{t^*+1}(i,j) = \cdots = \rho_T(i,j)$$

with the change-point $t^*$ allowing for $t^* \to 0$ or $T - t^* \to 0$.

### 4.3.1 The Weighting Functions

For the purpose of detecting correlation changes which occurred in the beginning or end of sample $(0, \mathfrak{t}_1] \cup [\mathfrak{t}_2, 1)$ for $\mathfrak{t}_1, \mathfrak{t}_2 \in (0, 1)$, we require to put more weights to either ends. The CUSUM test naturally allocates more weights to the centre of sample, as its distribution asymptotically converges to the sum of squared Brownian Bridges (cf. Equation 2.7 and 2.8 in Aue et al. 2009). Consequently, the CUSUM statistics converge only in the interval of $(\mathfrak{t}_1, \mathfrak{t}_2)$, especially for those maximally selected self-normalised CUSUM tests.

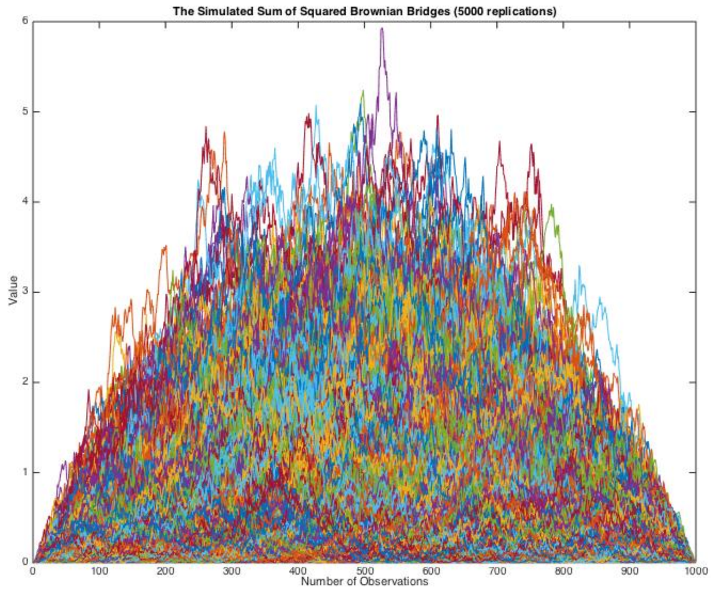
To visually see this argument, Figure 4.3.1 plots the simulated limit of self-normalised CUSUM statistics calculated on the correlations of a bivariate model (dimension $d = 2$),

$$\underset{0 \leq \mathfrak{t} \leq 1}{sup} \sum_{i=1}^{\bar{d}} B_i^2(\mathfrak{t})$$

where $\mathfrak{t}$ is the fraction index on interval $[0, 1]$, and $\mathfrak{t} = \frac{t}{T}$ in sample cases. $B_i$ for $1 \leq i \leq \bar{d}$ denotes independent Brownian Bridge, and $\bar{d} = \frac{d \times (d-1)}{2}$.

We simulate the bivariate model of $T = 1000$ observations, with 5000 replications. As shown, the limit starts from zero and ends at zero, and either the supreme or the maximum is always allocated in the center area of the graph. For finite samples if the actual break $t^*$ satisfies conditions of $t^* \to T$ or $(T - t^*) \to T$, the CUSUM tests will fail to detect any break. Therefore, we aim to give more weight in the sample ends, hunting for suitable weight functions $q(\mathfrak{t})$ for weighting the CUSUM statistics.

Figure 4.3.1: The Simulated Sum of Squared Brownian Bridges




Issues about weighted approximations have been discussed by Szyszkowicz (1994), Gombay et al. (1997) and Csörgő and Horváth (1997), and they assumed

**Assumption 4.3.3.** *A weight function used should be increasing in the neighbourhood of* $\mathsf{t} = 0$*, and decreasing in the neighbourhood of* $\mathsf{t} = 1$

We use the following two weight functions:

$$q_1(\mathsf{t}, \alpha) = (\mathsf{t}(1 - \mathsf{t}))^{\alpha}, \quad 0 \leq \mathsf{t} \leq 1 \tag{4.3.3}$$

$$q_2(\mathsf{t}, \alpha) = (\mathsf{t}(1 - \mathsf{t}) loglog \frac{1}{\mathsf{t}(1 - \mathsf{t})})^{\alpha}, \quad 0 \leq \mathsf{t} \leq 1 \tag{4.3.4}$$

where $\alpha$ is a self-selected parameter.

**Assumption 4.3.4.** *According to the integral test (Csörgő and Horváth, 1997), weight functions* $q(\alpha, \mathsf{t})$ *satisfy that* $\underset{0 \leq \mathsf{t} \leq 1-c}{inf} q(\mathsf{t}, \alpha) > 0$ *and* $\underset{0 \leq \mathsf{t} \leq 1-c}{sup} q(\mathsf{t}, \alpha) < \infty$ *for all* $0 < c < \alpha$*.*

Hence, such a weight function $q(\mathsf{t}, \alpha)$ would not distort the convergence of the CUSUM statistics. Assumption 4.3.4 emphasizes whether a weight function can satisfy the integral test, which requires that the weight function should be finite. To be more specific, for a weight function such that in 4.3.3, the integral test requires that for some $c > 0$:

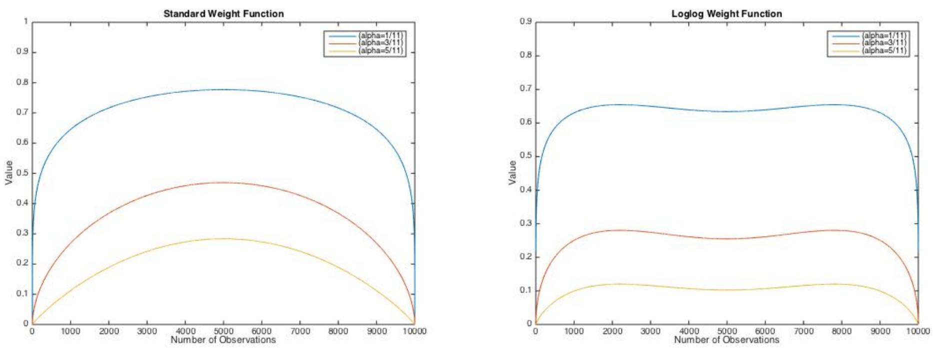
$$I_{0,1}(q_1(\mathsf{t}, \alpha), c) = \int_0^1 \frac{1}{\mathsf{t}(1-\mathsf{t})} e^{-\frac{c \cdot q_1^2(\mathsf{t}, \alpha)}{\mathsf{t}(1-\mathsf{t})}} d\mathsf{t} < \infty$$

and for the weight function 4.3.4, it requires that for some $c > 0$:

$$I_{0,1}(q_2(\mathsf{t}, \alpha), c) = \int_0^1 \frac{1}{\mathsf{t}(1-\mathsf{t}) loglog \frac{1}{\mathsf{t} \cdot (1-\mathsf{t})}} e^{-\frac{c \cdot q^2(\mathsf{t}, \alpha)}{\mathsf{t}(1-\mathsf{t}) loglog \frac{1}{\mathsf{t} \cdot (1-\mathsf{t})}}} d\mathsf{t} < \infty$$

Gombay et al. (1997) suggested and proved that the self-selected parameter $\alpha$ should be chosen from the interval $(0, 1/2)$. We therefore choose three values of $\alpha = 1/11$, $3/11$ and $5/11$. Figure 3.2 plots weight functions 4.3.3 and 4.3.4 with different $\alpha$.

Figure 4.3.2: The plot of weight functions




From Figure 4.3.2, we see that compared with weight function $q_2(\mathsf{t}, \alpha)$, $q_1(\mathsf{t}, \alpha)$ attributes more weights to either ends. Thus, weight function $q_1(\mathsf{t}, \alpha)$ is expected to outperform $q_2(\mathsf{t}, \alpha)$ in detecting changes in either ends, while as a trade-off, it may lose some power in the centre part. Besides, in each sub-plot, we can see that the higher value of $\alpha$, the less weights are obtained in ends, and more weights are distributed in centre area.

Because the weight function is used to standardize the weight of CUSUM statistics, we should consider standardizing the CUSUM statistics by the chosen weight function. The standardized statistics would thus take a shape looking like the reciprocal of the weight functions in Figure 4.3.2. Thus, applying higher value of $\alpha$ is expected to have higher power in ends of the sample, but consequently deteriorate the size as a sacrifice.

### 4.3.2 The Weighted CUSUM Tests

We apply weight functions into CUSUM statistics for detecting covariance changes in interval $(0,1)$. To test the null hypothesis, we use two types of weighted CUSUM tests,

$$W_T^{(1)} = \underset{0<\mathfrak{t}<1}{max} \frac{1}{T \cdot q_i(\mathfrak{t},\alpha)} (\mathfrak{s}(\lfloor T \cdot \mathfrak{t} \rfloor) - \frac{\lfloor T \cdot \mathfrak{t} \rfloor}{T} \mathfrak{s}(T))^\top \mathfrak{D}^{-1} (\mathfrak{s}(\lfloor T \cdot \mathfrak{t} \rfloor) - \frac{\lfloor T \cdot \mathfrak{t} \rfloor}{T} \mathfrak{s}(T))$$

$$W_T^{(2)} = \frac{1}{T^2 \cdot q_i(\mathfrak{t},\alpha)} \sum_{k=1}^{T} (\mathfrak{s}(\lfloor T \cdot \mathfrak{t} \rfloor) - \frac{\lfloor T \cdot \mathfrak{t} \rfloor}{T} \mathfrak{s}(T))^\top \mathfrak{D}^{-1} (\mathfrak{s}(\lfloor T \cdot \mathfrak{t} \rfloor) - \frac{\lfloor T \cdot \mathfrak{t} \rfloor}{T} \mathfrak{s}(T))$$

where

$$\mathfrak{s}(k) = \textstyle\sum_{s=1}^{\lfloor T \cdot \mathfrak{t} \rfloor} \mathfrak{r}_s, \text{ with } \mathfrak{r}_s = vech(\tilde{y}_s(i)\tilde{y}_s(j), 1 \leq i,j \leq d).$$

and $\tilde{y}_s(i)$ are de-meaned observations of $y_s(i)$, as from Equation 4.3.2. The Weight function is selected from the following,

$$q_1(\mathfrak{t},\alpha) = (\tfrac{\lfloor T \cdot \mathfrak{t} \rfloor}{T} \cdot (1 - \tfrac{\lfloor T \cdot \mathfrak{t} \rfloor}{T}))^\alpha.$$

$$\text{and } q_2(\mathfrak{t},\alpha) = (\tfrac{\lfloor T \cdot \mathfrak{t} \rfloor}{T} \cdot (1 - \tfrac{\lfloor T \cdot \mathfrak{t} \rfloor}{T}) \cdot loglog \tfrac{1}{\frac{\lfloor T \cdot \mathfrak{t} \rfloor}{T} \cdot (1 - \frac{\lfloor T \cdot \mathfrak{t} \rfloor}{T})})^\alpha$$

Following the last sub-section, we choose $\alpha$ strictly less than $1/2$ to guarantee the convergence of weighted CUSUM statistics. Nonetheless, the asymptotic distribution of weighted CUSUM tests is different from that of classical CUSUM tests.

Given the linearity of self-normalised weighted CUSUM statistics, the asymptotic distribution of $W_T^{(2)}$ can be easily derived from Equation 2.8 in Aue et al. (2009). However, for the reason that the maximally selected self-normalised weighted CUSUM statistics are converging at a slower rate (Horváth and Rice, 2014), the asymptotics of $W_T^{(1)}$ becomes more complicated.

We first discuss the limit of $W_T^{(2)}$. As explained above, the weighted self-normalised CUSUM statistics is simply derived from unweighted limit using following theorem:

**Theorem 4.3.1.** *Assuming a weight functions $q(\cdot) < \infty$ for some c, with $\alpha < 1/2$, the weighted limit distribution of the CUSUM test $W_T^{(2)}$ converges to the following:*

$$W_T^{(2)} \xrightarrow{\mathcal{D}} W^{(2)} = \sum_{j=1}^{\bar{d}} \int_0^1 \frac{B_j(\mathfrak{t})^2}{q_i^2(\mathfrak{t})} d\mathfrak{t} \tag{4.3.5}$$

with $i = 1, 2$. The proof is provided in appendix A.2.

We then discuss the limit of $W_T^{(1)}$. According to Equation A.2.2 in the proof of Theorem 4.3.1, the continuous mapping theorem, it is not hard to see that,

$$W_T^{(1)} = \underset{\mathfrak{t}_1 \le t \le \mathfrak{t}_2}{sup} \sum_{j=1}^{\bar{d}} \frac{B_j^2(\mathfrak{t})}{q_i^2(\mathfrak{t})} \tag{4.3.6}$$

with $i = 1, 2$, while since $W_T^{(1)}$ only can be found in $[\mathfrak{t}_1, \mathfrak{t}_2]$, and $W_T^{(1)}$ is not continuous at $[0, 1]$, $\underset{\mathfrak{t}_1 \le \mathfrak{t} \le \mathfrak{t}_2}{sup} W_T^{(1)}$ would be unbounded under the $H_0$, so that it is still necessary to trim the sample for detection. One solution is to normalise $W_T^{(1)}$ through suitable normalising terms, as suggested by the Darling and Erdös law (Darling and Erdös, 1956). The normalising constants are,

$$a_T = (2logX)^{\frac{1}{2}}, \quad b_T = 2logX + \frac{1}{2}loglogX - log\Gamma\left(\frac{1}{2}\right)$$

where $\Gamma(\cdot)$ is a gamma function, and we set $X = logT$

$$a_T = (2loglogT)^{\frac{1}{2}}, \quad b_T = 2loglogT + \frac{1}{2}logloglogT - log\Gamma\left(\frac{1}{2}\right) \tag{4.3.7}$$

**Theorem 4.3.2.** *Assuming a weight functions $q(\cdot) < \infty$ for some c. The Darling and Erdös theorem implies that,*

$$P\{a_T \cdot [\sup_{0<t<1} W_T^{(1)}] \leq \mathcal{K} + b_T\} \rightarrow e^{-2e^{-\mathcal{K}}} \tag{4.3.8}$$

*for $\mathcal{K} \in \mathbb{R}$.*

The proof of Theorem 4.3.2 can be found from Appendix A.2. However, the statistics in Theorem 4.3.3 converges to the Extreme Value distribution, and the convergence rate is very slow. Hence, to distinguish the null from the alternative, this limit is only approximated in very large samples.

To test the null hypothesis, we need to apply it to a de-volatized process. Recall that de-volatizing process (4.3.2), together with assumption 4.3.2 maintains the consistency of de-volatized process $\hat{y}_t^*(j)$.

The consistency of long run covariance estimator needs,

**Assumption 4.3.5.** $\left\|\hat{\mathfrak{D}}_T - \mathfrak{D}\right\| = o_P(1)$

The consistency of $\hat{\mathfrak{D}}_T$ has been discussed in literature, the present chapter uses the Bartlett kernel function $K(\cdot)$ and the Newey-West optimal bandwidth $h$.

$$\hat{\mathfrak{D}}_T = \sum_{l=-T}^{T} K(\frac{l}{h})\hat{\gamma}_l \tag{4.3.9}$$

where

$$\hat{\gamma}_l = \begin{cases} \frac{1}{T}\sum_{t=1}^{T-l}(\hat{\mathfrak{r}}_t - \bar{\mathfrak{r}}_T)(\hat{\mathfrak{r}}_{t+l} - \bar{\mathfrak{r}}_T)^\top & if \quad 0 \leq l < T \\ \frac{1}{T}\sum_{t=-T+1}^{T}(\hat{\mathfrak{r}}_t - \bar{\mathfrak{r}}_T)(\hat{\mathfrak{r}}_{t+l} - \bar{\mathfrak{r}}_T)^\top & if \quad -T \leq l < 0 \end{cases}$$

where $\bar{\mathfrak{r}}_T = \frac{1}{T}\sum_{s=1}^{T}\hat{\mathfrak{r}}_s$

Then, we use $\hat{W}_T^{(1)}$ and $\hat{W}_T^{(2)}$ to detect a single change in the correlations.

$$\hat{W}_T^{(1)} = \max_{0<\mathfrak{t}\leq 1}\frac{1}{T\cdot q_i(\mathfrak{t},\alpha)}(\hat{\mathfrak{s}}(\lfloor T\cdot\mathfrak{t}\rfloor) - \frac{\lfloor T\cdot\mathfrak{t}\rfloor}{T}\hat{\mathfrak{s}}(T))^\top\hat{\mathfrak{D}}^{-1}(\hat{\mathfrak{s}}(\lfloor T\cdot\mathfrak{t}\rfloor) - \frac{\lfloor T\cdot\mathfrak{t}\rfloor}{T}\hat{\mathfrak{s}}(T))$$

$$\hat{W}_T^{(2)} = \frac{1}{T^2\cdot q_i(\mathfrak{t},\alpha)}\sum_{k=1}^{T}(\hat{\mathfrak{s}}(\lfloor T\cdot\mathfrak{t}\rfloor) - \frac{\lfloor T\cdot\mathfrak{t}\rfloor}{T}\hat{\mathfrak{s}}(T))^\top\hat{\mathfrak{D}}^{-1}(\hat{\mathfrak{s}}(\lfloor T\cdot\mathfrak{t}\rfloor) - \frac{\lfloor T\cdot\mathfrak{t}\rfloor}{T}\hat{\mathfrak{s}}(T))$$

where

$$\hat{\mathfrak{s}}(k) = \sum_{s=1}^{\lfloor T\cdot\mathfrak{t}\rfloor}\hat{\mathfrak{r}}_s, \text{ with } \hat{\mathfrak{r}}_s = vech(\tilde{y}_s^*(i)\tilde{y}_s^*(j), 1\leq i,j\leq d).$$

and $\tilde{y}_s^*(i)$ is the de-meaned data of $\hat{y}_s^*(i)$.

**Theorem 4.3.3.** *Satisfying assumptions 4.3.1 to 4.3.5, the limits of* $\hat{W}_T^{(1)}$ *and* $\hat{W}_T^{(2)}$ *give as,*

$$P\{a_T\cdot[\sup_{0<\mathfrak{t}<1}\hat{W}_T^{(1)}]\leq\mathcal{K}+b_T\}\to e^{-2e^{-\mathcal{K}}} \tag{4.3.10}$$

$$\hat{W}_T^{(2)}\xrightarrow{\mathcal{D}} W_T^{(2)}$$

Relevant proofs can be referred to A.1.

## 4.4 The Empirical Simulation

This section conducts empirical Monte Carlo simulation study under the null and alternative hypothesis. We first compute critical values for $\hat{W}_T^{(1)}$ and $\hat{W}_T^{(2)}$ with finite samples, as well as asymptotic critical values. Beside the bivariate observations, we also consider higher dimensions up to 10. We then plot the empirical processes under the null hypothesis for the comparison. The empirical rejection rates are provided in the last subsections.

### 4.4.1 Computation of Critical Values

We first compute the critical values in asymptotic and finite empirical samples. Asymptotic critical values for $\hat{W}_T^{(2)}$ are obtained by simulating (4.3.5). Section 4.4.2 demonstrates the empirical process used under the null hypothesis. Table 4.4.1 tabulates critical values of $\hat{W}_T^{(2)}$ with $q_1(\cdot)$ and $q_2(\cdot)$ at 90%, 95% and 99% significance levels, where the value of parameter $\alpha$ spans from 1/11, 3/11 to 5/11. The dimensionality of $\mathbf{y}_t$ is 2. Table 4.4.2 however reports asymptotic critical values in higher dimensions. The statistics with weight function $q_2(\mathfrak{t}, \alpha)$ exhibits larger critical values. Note that the higher value of self-selected parameter $\alpha$, the larger value of critical values.

Table 4.4.1: Critical Values for $\hat{W}_T^{(2)}$ (d=2)

The replication run 5000 times, critical values in finite samples are obtained from empirical critical values, while the infinite case is obtained from asymptotic critical values with observation 10000.

| | | $q_1(\mathfrak{t},\alpha)$ | | | | | $q_2(\mathfrak{t},\alpha)$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\alpha$ | | T=100 | T=200 | T=500 | T=1000 | $T=\infty$ | T=100 | T=200 | T=500 | T=1000 | $T=\infty$ |
| | 90% | 1.13 | 1.11 | 1.05 | 1.07 | 1.14 | 1.30 | 1.28 | 1.21 | 1.24 | 1.33 |
| $\alpha=\frac{1}{11}$ | 95% | 1.36 | 1.31 | 1.25 | 1.27 | 1.34 | 1.58 | 1.51 | 1.44 | 1.48 | 1.60 |
| | 99% | 2.02 | 1.92 | 1.70 | 1.80 | 1.88 | 2.34 | 2.20 | 1.99 | 2.10 | 2.22 |
| | 90% | 2.09 | 2.07 | 1.95 | 1.98 | 2.08 | 3.13 | 3.12 | 2.96 | 3.04 | 3.22 |
| $\alpha=\frac{3}{11}$ | 95% | 2.49 | 2.48 | 2.30 | 2.34 | 2.45 | 3.78 | 3.79 | 3.51 | 3.58 | 3.85 |
| | 99% | 3.87 | 3.51 | 3.06 | 3.16 | 3.33 | 5.64 | 5.30 | 4.72 | 4.84 | 5.02 |
| | 90% | 4.10 | 3.94 | 3.76 | 3.76 | 3.99 | 7.61 | 7.51 | 7.19 | 7.27 | 7.76 |
| $\alpha=\frac{5}{11}$ | 95% | 4.98 | 4.68 | 4.42 | 4.39 | 4.64 | 9.57 | 8.92 | 8.58 | 8.67 | 9.28 |
| | 99% | 7.52 | 6.70 | 6.01 | 6.12 | 6.05 | 13.86 | 12.87 | 11.97 | 12.08 | 12.80 |

Table 4.4.2: Critical Values for $\hat{W}_T^{(2)}$ with $d = 3 - 10$ ($T = \infty$)

| | | $q_1(\mathfrak{t}, \alpha)$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | d=3 | d=4 | d=5 | d=6 | d=7 | d=8 | d=9 | d=10 |
| | 90% | 1.69 | 2.63 | 3.80 | 5.12 | 6.66 | 8.30 | 10.24 | 12.32 |
| $\alpha = 1/11$ | 95% | 1.91 | 2.92 | 4.13 | 5.54 | 7.07 | 8.76 | 10.76 | 12.83 |
| | 99% | 2.37 | 3.51 | 4.77 | 6.22 | 7.85 | 9.71 | 11.83 | 13.88 |
| | 90% | 2.31 | 3.62 | 5.15 | 6.96 | 8.99 | 11.28 | 13.85 | 16.77 |
| $\alpha = 3/11$ | 95% | 2.63 | 3.99 | 5.55 | 7.44 | 9.54 | 11.89 | 14.61 | 17.55 |
| | 99% | 3.36 | 4.75 | 6.43 | 8.44 | 10.62 | 13.20 | 15.80 | 19.11 |
| | 90% | 3.18 | 4.90 | 6.98 | 9.49 | 12.31 | 15.41 | 19.02 | 22.94 |
| $\alpha = 5/11$ | 95% | 3.62 | 5.40 | 7.53 | 10.10 | 12.99 | 16.16 | 19.83 | 23.82 |
| | 99% | 4.46 | 6.43 | 8.63 | 11.44 | 14.36 | 17.71 | 21.60 | 25.58 |

| | | $q_2(\mathfrak{t}, \alpha)$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | d=3 | d=4 | d=5 | d=6 | d=7 | d=8 | d=9 | d=10 |
| | 90% | 2.32 | 3.59 | 5.09 | 6.92 | 8.93 | 11.23 | 13.69 | 16.46 |
| $\alpha = 1/11$ | 95% | 2.64 | 3.94 | 5.52 | 7.38 | 9.48 | 11.87 | 14.45 | 17.26 |
| | 99% | 3.41 | 4.68 | 6.39 | 8.48 | 10.66 | 12.99 | 15.86 | 18.67 |
| | 90% | 5.58 | 8.80 | 12.49 | 16.77 | 21.74 | 27.09 | 33.76 | 40.27 |
| $\alpha = 3/11$ | 95% | 6.46 | 9.62 | 13.66 | 17.84 | 23.12 | 28.54 | 35.30 | 41.99 |
| | 99% | 8.17 | 11.42 | 15.80 | 19.99 | 25.79 | 31.28 | 38.30 | 45.43 |
| | 90% | 13.84 | 21.32 | 30.32 | 40.78 | 52.83 | 66.30 | 81.95 | 98.65 |
| $\alpha = 5/11$ | 95% | 15.76 | 23.38 | 32.65 | 43.67 | 56.12 | 70.18 | 85.73 | 103.51 |
| | 99% | 19.44 | 27.16 | 37.68 | 49.52 | 62.29 | 77.59 | 93.82 | 111.28 |

Table 4.4.3: Critical Values for $\hat{W}_T^{(1)}$ (d=2)

The replication run 5000 times, critical values in finite samples are obtained from empirical critical values, while the infinite case is obtained from asymptotic critical values with observation 10000.

| | | $q_1(\mathrm{t},\alpha)$ | | | | | $q_2(\mathrm{t},\alpha)$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\alpha$ | | T=100 | T=200 | T=500 | T=1000 | $T=\infty$ | T=100 | T=200 | T=500 | T=1000 | $T=\infty$ |
| | 90% | 3.30 | 3.21 | 3.13 | 3.19 | 3.39 | 3.82 | 3.81 | 3.69 | 3.79 | 4.10 |
| $\alpha=\frac{1}{11}$ | 95% | 3.93 | 3.75 | 3.58 | 3.76 | 3.95 | 4.57 | 4.43 | 4.26 | 4.50 | 4.69 |
| | 99% | 5.63 | 5.28 | 4.81 | 4.93 | 5.21 | 6.55 | 6.20 | 5.74 | 5.90 | 6.40 |
| | 90% | 6.23 | 6.09 | 5.65 | 5.66 | 6.10 | 9.22 | 9.17 | 9.00 | 9.28 | 9.86 |
| $\alpha=\frac{3}{11}$ | 95% | 7.92 | 7.24 | 6.53 | 6.56 | 7.05 | 11.23 | 10.89 | 10.50 | 10.66 | 11.40 |
| | 99% | 13.06 | 10.19 | 8.64 | 8.61 | 8.80 | 16.39 | 15.54 | 13.90 | 14.15 | 15.10 |
| | 90% | 15.98 | 15.87 | 14.42 | 14.15 | 11.80 | 23.50 | 23.12 | 22.38 | 22.57 | 24.00 |
| $\alpha=\frac{5}{11}$ | 95% | 22.74 | 21.94 | 18.96 | 18.27 | 13.10 | 29.18 | 27.79 | 26.35 | 26.21 | 28.20 |
| | 99% | 47.87 | 48.17 | 38.90 | 38.85 | 16.50 | 45.44 | 42.14 | 35.04 | 36.99 | 36.10 |

Regarding maximally selected statistics $\hat{W}_T^{(1)}$, we can obtain the asymptotic critical values through simulating (4.3.10) directly. However, as it follows an extreme value distribution, and its critical value is a function of sample size $T$ and significance level $p$, the convergence rate is very low and rejection rate under the alternative should be hampered. Therefore, following Csörgö and Horváth (1997), we use the $\underset{0<\mathrm{t}<1}{sup}\,\hat{W}_T^{(1)}$ as an approximation for (4.3.10), and compute the critical values through simulating

$$\underset{t_1\le t\le t_2}{sup}\sum_{i=1}^{\bar{d}}\frac{B_{(i)}^2(t)}{q(t)} \tag{4.4.1}$$

and tabulate the results as in Table 4.4.3. Table 4.4.4 reports asymptotic critical values in higher dimensions.

Table 4.4.4: Critical Values for $W_T^{(1)}$ with $d = 3 - 10$ ($T = \infty$)

| | | $q_1(\mathfrak{t}, \alpha)$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | d=3 | d=4 | d=5 | d=6 | d=7 | d=8 | d=9 | d=10 |
| | 90% | 5.12 | 7.02 | 9.39 | 11.89 | 14.82 | 17.94 | 21.53 | 25.34 |
| $\alpha = 1/11$ | 95% | 5.78 | 7.82 | 10.30 | 12.74 | 15.74 | 19.23 | 22.88 | 26.76 |
| | 99% | 7.15 | 9.32 | 12.13 | 14.70 | 17.89 | 21.61 | 25.58 | 29.64 |
| | 90% | 8.82 | 12.12 | 15.92 | 20.34 | 25.34 | 30.42 | 36.35 | 42.43 |
| $\alpha = 3/11$ | 95% | 9.87 | 13.29 | 17.32 | 22.00 | 27.26 | 32.39 | 38.37 | 44.62 |
| | 99% | 12.36 | 15.82 | 20.10 | 25.07 | 30.61 | 35.70 | 42.62 | 48.67 |
| | 90% | 16.96 | 22.59 | 29.02 | 36.20 | 44.88 | 53.97 | 64.28 | 75.16 |
| $\alpha = 5/11$ | 95% | 18.65 | 24.67 | 31.11 | 38.75 | 47.99 | 56.90 | 67.40 | 78.56 |
| | 99% | 22.27 | 28.45 | 35.58 | 43.93 | 53.48 | 63.10 | 73.12 | 86.28 |

| | | $q_2(\mathfrak{t}, \alpha)$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | d=3 | d=4 | d=5 | d=6 | d=7 | d=8 | d=9 | d=10 |
| | 90% | 6.17 | 8.48 | 11.27 | 14.33 | 17.94 | 21.82 | 26.96 | 30.72 |
| $\alpha = 1/11$ | 95% | 6.96 | 9.43 | 12.40 | 15.49 | 19.31 | 23.30 | 27.52 | 32.38 |
| | 99% | 8.62 | 11.46 | 14.60 | 17.94 | 21.77 | 26.48 | 31.10 | 35.50 |
| | 90% | 14.93 | 20.46 | 27.64 | 35.26 | 43.94 | 53.26 | 64.35 | 75.90 |
| $\alpha = 3/11$ | 95% | 16.92 | 22.95 | 30.05 | 38.11 | 47.33 | 56.90 | 68.50 | 79.85 |
| | 99% | 21.49 | 27.88 | 35.59 | 44.09 | 52.98 | 63.71 | 76.38 | 88.17 |
| | 90% | 36.67 | 50.87 | 67.64 | 86.15 | 108.31 | 132.52 | 158.54 | 188.65 |
| $\alpha = 5/11$ | 95% | 41.74 | 56.90 | 74.07 | 93.37 | 116.50 | 142.07 | 167.98 | 198.97 |
| | 99% | 52.05 | 69.84 | 88.43 | 107.17 | 131.62 | 162.33 | 189.47 | 219.63 |

### 4.4.2 The Empirical Process

We generate bivariate observations $\mathbf{y}_t = [y_t(1), y_t(2)]^\top$ with (4.3.1), where $\mathbf{e}_t$ is a two dimensional standard normal random vector $\mathbf{e}_t = [e_t(1), e_t(2)]^\top$, and $\mathbf{\Sigma}_t^{\frac{1}{2}}$ is in a Cholesky form. The dynamics of second moments in $\mathbf{y}_t$ is described by conditional covariance matrix $\mathbf{\Sigma}_t$, which can be generated by most of stationary multivariate volatility models, mainly Multivariate GARCH models.

We generate $\mathbf{\Sigma}_t$ through the dynamic conditional correlation (DCC) model.

$$\mathbf{\Sigma}_t = \mathbf{D}_t \mathbf{R}_t \mathbf{D}_t \tag{4.4.2}$$

where $\mathbf{D}_t$ and $\mathbf{R}_t$ describe conditional volatilities and correlations, respectively. $\mathbf{D}_t$ is a diagonal matrix composed by conditional standard deviations $\mathbf{D}_t = diag[\sqrt{h_t(1)}, \sqrt{h_t(2)}]$, where $h_t(i)$ are recursively generated by GARCH (1,1) processes with initial value $h_0(i) = 1$,

$$h_t(i) = \beta_0 + \beta_1 \cdot y_t^2(i) + \beta_2 \cdot h_{t-1}(i)$$

To be more closer to real financial data, we globally set $\beta_0 = 0.1$, $\beta_1 = 0.01$ and $\beta_2 = 0.7$, thereby giving low ARCH effect but high persistence in conditional volatilities. The value on $\beta_i$ $i = 1, 2, 3$ play less importance as $\mathbf{y}_t$ will be standardized by realized conditional variances. $\mathbf{R}_t$ is a symmetric matrix with unit diagonal elements, and it is generated by the process (cf. Engle, 2002).

$$\mathbf{R}_t = (diag(\mathbf{Q}_t))^{-\frac{1}{2}} \mathbf{Q}_t (diag(\mathbf{Q}_t))^{-\frac{1}{2}}$$

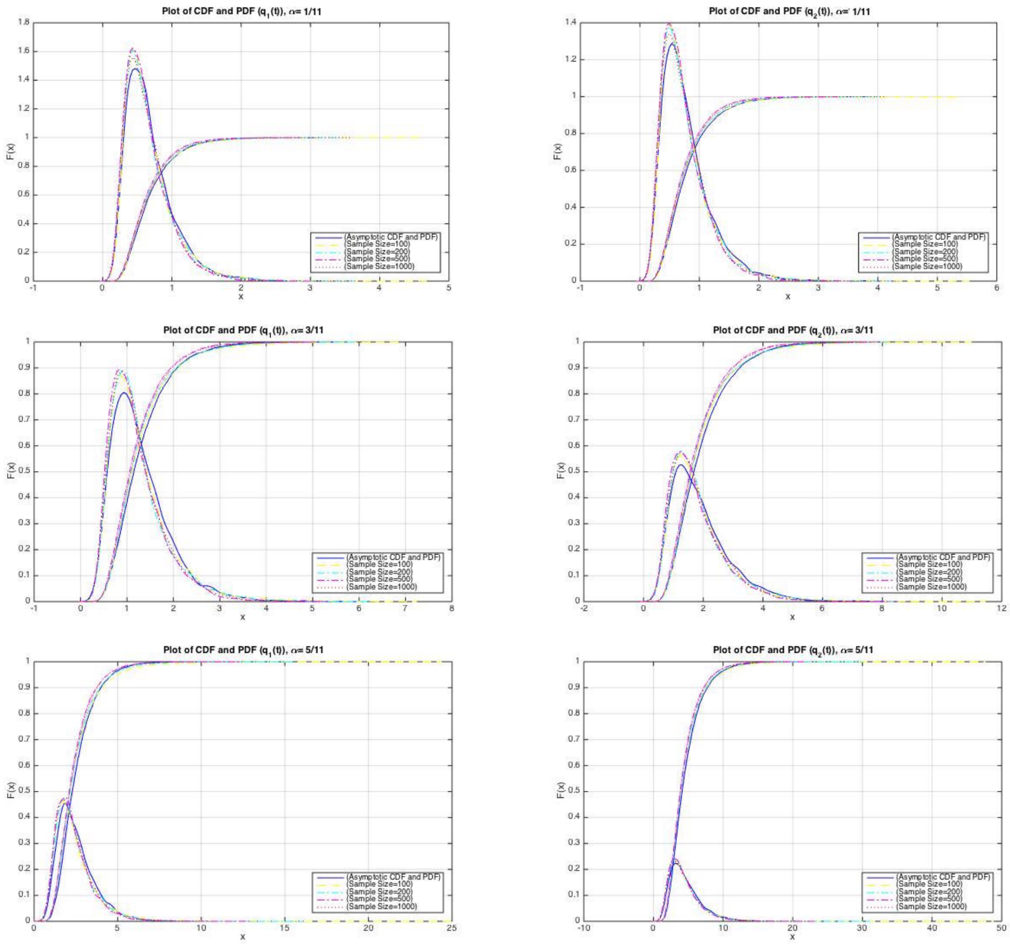
and the quasi conditional correlation matrix $\mathbf{Q}_t$ is recursively generated by

$$\mathbf{Q}_t = \theta_1 \mathbf{C} + \theta_2 \mathbf{y}_{t-1} \mathbf{y}_{t-1}^T + \theta_3 \mathbf{Q}_{t-1} \tag{4.4.3}$$

We set the initial value as identity matrix $\mathbf{Q}_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

where $\mathbf{C}$ is positive definite, and weight coefficients satisfy $\theta_1 > 0$, $\theta_2$, $\theta_3 \geq 0$, with restriction $\theta_1 = 1 - \theta_2 - \theta_3$.

Figure 4.4.1: The Asymptotic and Empirical Process of $\hat{W}_T^{(2)}$ with weight function $q_1(\mathfrak{t}, \alpha)$ and $q_2(\mathfrak{t}, \alpha)$ for $\alpha = 1/11$, $3/11$ and $5/11$.

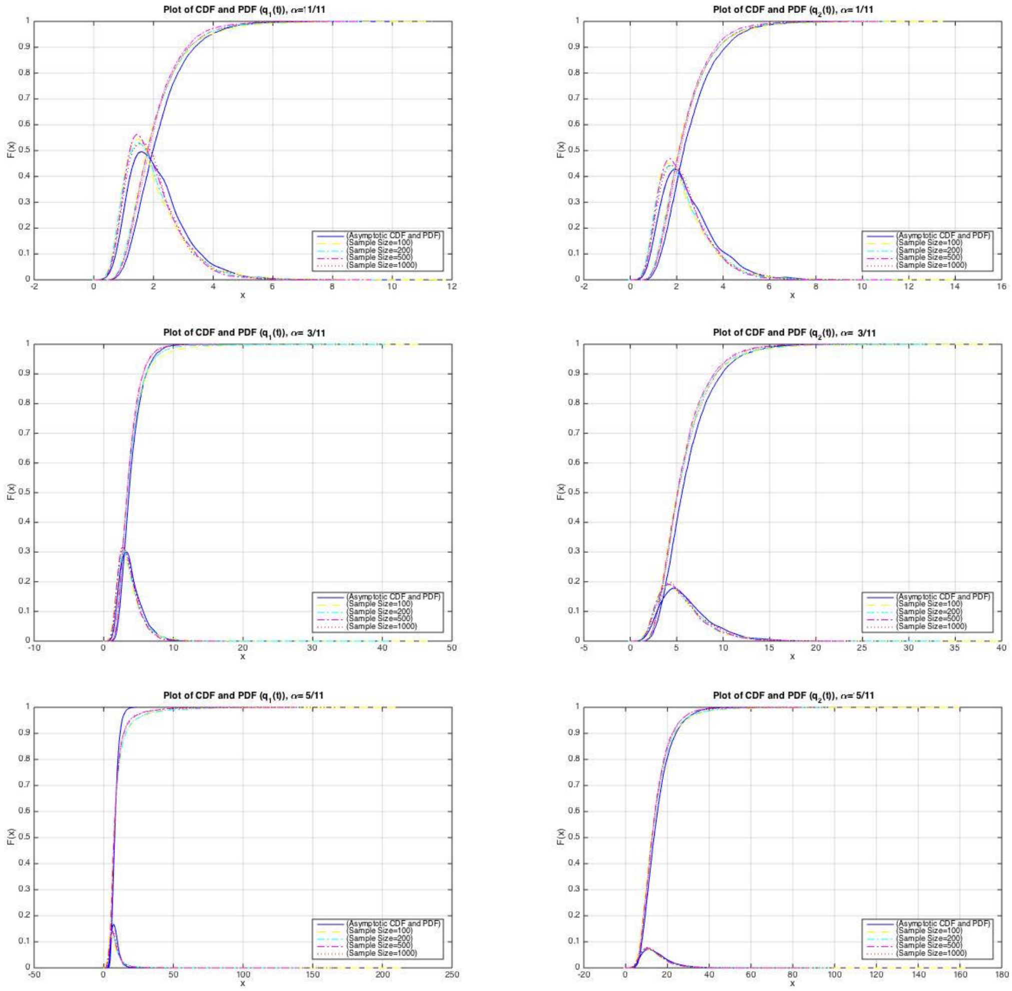



To verify weighted CUSUM statistics $\hat{W}_T^{(1)}$ and $\hat{W}_T^{(2)}$ on empirical process under the null hypothesis, we first generate observations by (4.4.3) with

$$\mathbf{C} = \begin{bmatrix} 1 & \delta \\ \delta & 1 \end{bmatrix} \tag{4.4.4}$$

$\delta$ is the correlation term which will be defined later, and $\theta_2 = 0.01$, $\theta_3 = 0.1$.

Figure 4.4.2: The Asymptotic and Empirical Process of $\hat{W}_T^{(1)}$ with weight function $q_1(\mathfrak{t}, \alpha)$ and $q_2(\mathfrak{t}, \alpha)$ for $\alpha = 1/11$, $3/11$ and $5/11$.




Empirically giving different sample sizes $T = 100, 200, 500$ and $1000$, we first subtract the sample-mean from generated process, and then estimate the conditional covariance matrices through the DCC model, so that diagonal elements from estimated conditional covariance matrix can be used to devolatize the demeaned data. We numerically compute $\hat{W}_T^{(1)}$ and $\hat{W}_T^{(2)}$ based on devolatized processes $\hat{\mathbf{y}}_t^*$. Asymptotic statistics are obtained by simulating (4.3.5) and (4.4.1) with sample size 10000. Both are replicated 10000 times. The warming-up parameter in DGP is 0.2. Then we plot probability density function (PDF) and cumulative distribution function (CDF) for both asymptotic and empirical sequences.

Figure 4.4.1 shows distribution convergence on $\hat{W}_T^{(2)}$ with $\alpha = 1/11, 3/11$ and $5/11$. Generally, empirical distributions fit the theory well. For using weight function $q_1(\mathfrak{t}, \alpha)$, we find that a closer value of $\alpha$ to $1/2$ shows better convergence, i.e. the case $\alpha = 5/11$. The convergence pattern for using weight function $q_2(\mathfrak{t}, \alpha)$ is roughly consistent with weight function $q_1(\mathfrak{t}, \alpha)$, the case of $\alpha = 5/11$ produces better fitting on distribution convergence, and rather slight divergences exist in cases $\alpha = 1/11$ and $\alpha = 3/11$.

Figure 4.4.2 displays the distribution plots for $\hat{W}_T^{(1)}$. Unlike $\hat{W}_T^{(2)}$, the fitting of $\hat{W}_T^{(1)}$ is not as precise, a part from a particular case is fitting right tail of $q_1(\mathfrak{t}, \alpha)$ distribution in case of $\alpha = 5/11$. Not surprisingly, due to the slow convergence rate, the empirical process of maximally selected self-normalised weighted CUSUM slightly deviates from the asymptotic distribution in finite samples. Nonetheless, we can see that generally the empirical distributions converge to the theoretical ones. The size and power of these two tests will be assessed and discussed in next section using a simulation study.

### 4.4.3 The Rejection Rate

Treating semi-parametric CUSUM test (Chapter 3) as a bench-mark test ($\hat{M}_T^{(1)}$ and $\hat{M}_T^{(2)}$), we then assess the performance of weighted CUSUM tests by measuring rejection rates under the null and alternative hypotheses. The simulation is composed by five experiments, and each experiment uses the DGP in Section 4.4.2. - the constant term $\mathbf{C}$ is specified in (4.4.4), where $\delta = 0$ under the null of no change. Under the alternative hypothesis, we assume that a single change occurred at $t^* = \lfloor \mathfrak{t}^* \cdot T \rfloor$, and $\delta$ equals to 0 before $t^*$, and alter to different change magnitudes 0.2, 0.4, 0.6 and 0.8 after $t^*$.

We consider two scenarios of dependence in conditional correlations: weak dependent processes (low ARCH $\theta_2 = 0.005$ and low GARCH $\theta_3 = 0.1$) and strong dependent processes (high ARCH $\theta_2 = 0.01$ and high GARCH $\theta_3 = 0.9$). Observations are generated in a small sample $T = 100$, and a large sample $T = 500$. Every simulation replicates 5000 times. For binary observations, numerical CUSUM-typed statistics can reject $H_0$ if they are greater than critical values in Table 4.4.1 and 4.4.3 (critical values for semi-parametric CUSUM tests are those in Table 3.4.1 in Chapter 3).

#### 4.4.3.1 Experiment I ($\mathsf{t}^* = 0.1$)

We first set the change-point location $\mathsf{t}^*$ at the beginning of sample, $\mathsf{t}^* = 0.1$. Table 4.4.5 reports rejection rates of CUSUM-typed tests. Table 4.4.6 documents the estimated change $\hat{\nu}_T$.

Results can be summarized to following points. Firstly, rejection rates are increasing along with change magnitude $\delta$ increasing, and $\hat{\nu}_T$ also approaches $\mathsf{t}^*$ gradually. Secondly, all tests are fairly correctly sized, and the $\hat{W}_T^{(1)}$ statistic produce slightly lower rejection rates than the counterpart in $\hat{W}_T^{(2)}$ under $H_A$. Thirdly, comparing the CUSUM tests, weighted CUSUM tests outperform to unweighted CUSUM tests, particularly when using the weight function $q_1(\mathsf{t}, \alpha)$. Selecting a higher value of $\alpha$, the power of $\hat{W}_T^{(1)}$ and $\hat{W}_T^{(2)}$ become higher. Lastly, although the strong dependence distorts the performance of CUSUM-type tests, the size and power of tests are still reliable, in both of small and large samples.

The accuracy of the estimated locations is consistent with the result of the rejection rate. The mean, medium and standard deviation of estimated time $\hat{\nu}_T$ show reasonable patterns. Weighted CUSUM tests with higher $\alpha$ converge to $\mathsf{t}^* = 0.1$ more rapidly, although the standard deviation of $\hat{\nu}_T$ is becoming larger. The strong dependence also slightly deteriorates the precise estimated change-point location, however this issue disappears in large enough samples. For example the weighted CUSUM tests with the weight function $q_1(\mathsf{t}, 5/11)$, the mean of $\hat{\nu}_T$ approaches 0.1 for $\delta = 0.8$ in small samples and for $\delta = 0.6$ in the large samples.

Table 4.4.5: Empirical rejection rates for CUSUM-typed tests ($\mathfrak{t}^* = 0.1$)

T=100

| | | | $\hat{M}_T^{(1)}$ | $\hat{W}_T^{(1)}$ | | | | | | $\hat{M}_T^{(2)}$ | $\hat{W}_T^{(2)}$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Coefficients | $\delta$ | | Unweighted | $q_1(t,\frac{1}{11})$ | $q_1(t,\frac{3}{11})$ | $q_1(t,\frac{5}{11})$ | $q_2(t,\frac{1}{11})$ | $q_2(t,\frac{3}{11})$ | $q_2(t,\frac{5}{11})$ | Unweighted | $q_1(t,\frac{1}{11})$ | $q_1(t,\frac{3}{11})$ | $q_1(t,\frac{5}{11})$ | $q_2(t,\frac{1}{11})$ | $q_2(t,\frac{3}{11})$ | $q_2(t,\frac{5}{11})$ |
| | | 90% | 0.11 | 0.10 | 0.09 | 0.10 | 0.10 | 0.09 | 0.08 | 0.12 | 0.10 | 0.09 | 0.09 | 0.10 | 0.09 | 0.09 |
| | 0 | 95% | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.04 | 0.06 | 0.05 | 0.04 | 0.04 | 0.05 | 0.04 | 0.04 |
| | | 99% | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| | | 90% | 0.12 | 0.12 | 0.15 | 0.26 | 0.11 | 0.11 | 0.12 | 0.14 | 0.13 | 0.13 | 0.15 | 0.12 | 0.11 | 0.11 |
| | 0.2 | 95% | 0.06 | 0.06 | 0.08 | 0.13 | 0.05 | 0.06 | 0.07 | 0.07 | 0.07 | 0.06 | 0.07 | 0.06 | 0.05 | 0.06 |
| | | 99% | 0.01 | 0.01 | 0.01 | 0.03 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.02 | 0.01 | 0.01 | 0.01 |
| | | 90% | 0.18 | 0.19 | 0.27 | 0.36 | 0.17 | 0.17 | 0.19 | 0.22 | 0.21 | 0.23 | 0.26 | 0.20 | 0.19 | 0.19 |
| $[\theta_2 = 0.005,\ \theta_3 = 0.1]$ | 0.4 | 95% | 0.09 | 0.11 | 0.17 | 0.21 | 0.09 | 0.10 | 0.12 | 0.14 | 0.13 | 0.13 | 0.15 | 0.12 | 0.11 | 0.11 |
| | | 99% | 0.03 | 0.02 | 0.04 | 0.06 | 0.02 | 0.02 | 0.04 | 0.04 | 0.03 | 0.04 | 0.05 | 0.03 | 0.03 | 0.03 |
| | | 90% | 0.36 | 0.39 | 0.51 | 0.58 | 0.35 | 0.36 | 0.38 | 0.43 | 0.42 | 0.45 | 0.49 | 0.40 | 0.40 | 0.40 |
| | 0.6 | 95% | 0.25 | 0.29 | 0.37 | 0.39 | 0.24 | 0.25 | 0.27 | 0.33 | 0.32 | 0.32 | 0.36 | 0.31 | 0.28 | 0.29 |
| | | 99% | 0.10 | 0.11 | 0.13 | 0.16 | 0.08 | 0.10 | 0.12 | 0.13 | 0.12 | 0.14 | 0.16 | 0.10 | 0.11 | 0.12 |
| | | 90% | 0.63 | 0.67 | 0.78 | 0.82 | 0.61 | 0.62 | 0.63 | 0.72 | 0.70 | 0.73 | 0.76 | 0.69 | 0.68 | 0.68 |
| | 0.8 | 95% | 0.50 | 0.56 | 0.66 | 0.65 | 0.50 | 0.52 | 0.51 | 0.60 | 0.60 | 0.61 | 0.64 | 0.58 | 0.56 | 0.56 |
| | | 99% | 0.31 | 0.34 | 0.35 | 0.32 | 0.29 | 0.29 | 0.29 | 0.36 | 0.35 | 0.37 | 0.41 | 0.33 | 0.34 | 0.34 |
| | | 90% | 0.14 | 0.12 | 0.11 | 0.11 | 0.12 | 0.11 | 0.10 | 0.14 | 0.12 | 0.11 | 0.11 | 0.12 | 0.11 | 0.10 |
| | 0 | 95% | 0.06 | 0.06 | 0.05 | 0.05 | 0.06 | 0.06 | 0.05 | 0.07 | 0.07 | 0.06 | 0.05 | 0.06 | 0.05 | 0.05 |
| | | 99% | 0.02 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.02 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| | | 90% | 0.15 | 0.14 | 0.17 | 0.26 | 0.13 | 0.12 | 0.13 | 0.17 | 0.15 | 0.15 | 0.17 | 0.14 | 0.13 | 0.13 |
| | 0.2 | 95% | 0.07 | 0.07 | 0.09 | 0.12 | 0.06 | 0.07 | 0.07 | 0.09 | 0.08 | 0.07 | 0.08 | 0.07 | 0.06 | 0.06 |
| | | 99% | 0.02 | 0.01 | 0.01 | 0.03 | 0.01 | 0.01 | 0.01 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 |
| | | 90% | 0.20 | 0.21 | 0.26 | 0.34 | 0.19 | 0.18 | 0.19 | 0.24 | 0.22 | 0.23 | 0.25 | 0.21 | 0.21 | 0.20 |
| $[\theta_2 = 0.01,\ \theta_3 = 0.9]$ | 0.4 | 95% | 0.10 | 0.11 | 0.16 | 0.19 | 0.09 | 0.10 | 0.11 | 0.15 | 0.14 | 0.13 | 0.15 | 0.13 | 0.12 | 0.12 |
| | | 99% | 0.03 | 0.03 | 0.03 | 0.05 | 0.03 | 0.03 | 0.03 | 0.04 | 0.04 | 0.04 | 0.04 | 0.03 | 0.04 | 0.03 |
| | | 90% | 0.34 | 0.36 | 0.45 | 0.52 | 0.33 | 0.33 | 0.34 | 0.41 | 0.40 | 0.42 | 0.45 | 0.38 | 0.38 | 0.37 |
| | 0.6 | 95% | 0.22 | 0.25 | 0.31 | 0.32 | 0.21 | 0.23 | 0.24 | 0.30 | 0.30 | 0.29 | 0.32 | 0.28 | 0.26 | 0.26 |
| | | 99% | 0.08 | 0.08 | 0.10 | 0.12 | 0.07 | 0.07 | 0.09 | 0.11 | 0.10 | 0.11 | 0.13 | 0.09 | 0.10 | 0.10 |
| | | 90% | 0.57 | 0.59 | 0.66 | 0.71 | 0.55 | 0.54 | 0.54 | 0.65 | 0.64 | 0.66 | 0.69 | 0.63 | 0.61 | 0.61 |
| | 0.8 | 95% | 0.44 | 0.47 | 0.54 | 0.52 | 0.43 | 0.42 | 0.42 | 0.53 | 0.53 | 0.52 | 0.56 | 0.51 | 0.49 | 0.49 |
| | | 99% | 0.24 | 0.24 | 0.23 | 0.22 | 0.21 | 0.21 | 0.21 | 0.29 | 0.27 | 0.31 | 0.32 | 0.25 | 0.27 | 0.27 |

T=500

| | | | $\hat{M}_T^{(1)}$ | $\hat{W}_T^{(1)}$ | | | | | | $\hat{M}_T^{(2)}$ | $\hat{W}_T^{(2)}$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Coefficients | $\delta$ | | Unweighted | $q_1(t,\frac{1}{11})$ | $q_1(t,\frac{3}{11})$ | $q_1(t,\frac{5}{11})$ | $q_2(t,\frac{1}{11})$ | $q_2(t,\frac{3}{11})$ | $q_2(t,\frac{5}{11})$ | Unweighted | $q_1(t,\frac{1}{11})$ | $q_1(t,\frac{3}{11})$ | $q_1(t,\frac{5}{11})$ | $q_2(t,\frac{1}{11})$ | $q_2(t,\frac{3}{11})$ | $q_2(t,\frac{5}{11})$ |
| | | 90% | 0.10 | 0.09 | 0.09 | 0.11 | 0.10 | 0.10 | 0.10 | 0.10 | 0.09 | 0.09 | 0.09 | 0.09 | 0.09 | 0.09 |
| | 0 | 95% | 0.04 | 0.04 | 0.04 | 0.05 | 0.04 | 0.04 | 0.04 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.04 |
| | | 99% | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| | | 90% | 0.13 | 0.14 | 0.15 | 0.16 | 0.13 | 0.13 | 0.12 | 0.16 | 0.15 | 0.16 | 0.18 | 0.15 | 0.15 | 0.15 |
| | 0.2 | 95% | 0.06 | 0.07 | 0.09 | 0.06 | 0.07 | 0.06 | 0.06 | 0.09 | 0.09 | 0.09 | 0.09 | 0.09 | 0.09 | 0.08 |
| | | 99% | 0.01 | 0.02 | 0.02 | 0.01 | 0.01 | 0.01 | 0.01 | 0.02 | 0.02 | 0.02 | 0.02 | 0.01 | 0.01 | 0.02 |
| | | 90% | 0.35 | 0.40 | 0.51 | 0.49 | 0.35 | 0.35 | 0.33 | 0.41 | 0.41 | 0.45 | 0.50 | 0.40 | 0.41 | 0.41 |
| $[\theta_2 = 0.005,\ \theta_3 = 0.1]$ | 0.4 | 95% | 0.23 | 0.28 | 0.40 | 0.25 | 0.23 | 0.22 | 0.21 | 0.31 | 0.32 | 0.35 | 0.38 | 0.30 | 0.30 | 0.30 |
| | | 99% | 0.08 | 0.11 | 0.18 | 0.02 | 0.08 | 0.08 | 0.06 | 0.13 | 0.12 | 0.14 | 0.16 | 0.11 | 0.11 | 0.12 |
| | | 90% | 0.88 | 0.91 | 0.96 | 0.95 | 0.88 | 0.89 | 0.88 | 0.89 | 0.89 | 0.91 | 0.93 | 0.88 | 0.89 | 0.89 |
| | 0.6 | 95% | 0.78 | 0.85 | 0.93 | 0.86 | 0.79 | 0.79 | 0.79 | 0.82 | 0.83 | 0.86 | 0.89 | 0.81 | 0.82 | 0.82 |
| | | 99% | 0.57 | 0.68 | 0.81 | 0.32 | 0.58 | 0.59 | 0.52 | 0.61 | 0.60 | 0.67 | 0.72 | 0.56 | 0.57 | 0.59 |
| | | 90% | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | 0.8 | 95% | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | | 99% | 0.99 | 0.99 | 1.00 | 0.97 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 1.00 | 1.00 | 0.99 | 0.99 | 0.99 |
| | | 90% | 0.12 | 0.13 | 0.12 | 0.12 | 0.13 | 0.13 | 0.12 | 0.12 | 0.12 | 0.12 | 0.12 | 0.12 | 0.12 | 0.12 |
| | 0 | 95% | 0.06 | 0.07 | 0.06 | 0.05 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 |
| | | 99% | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| | | 90% | 0.16 | 0.17 | 0.18 | 0.17 | 0.16 | 0.16 | 0.15 | 0.18 | 0.17 | 0.18 | 0.20 | 0.17 | 0.17 | 0.17 |
| | 0.2 | 95% | 0.08 | 0.09 | 0.11 | 0.06 | 0.09 | 0.08 | 0.08 | 0.11 | 0.11 | 0.11 | 0.11 | 0.11 | 0.10 | 0.10 |
| | | 99% | 0.02 | 0.02 | 0.03 | 0.01 | 0.02 | 0.02 | 0.01 | 0.03 | 0.02 | 0.03 | 0.02 | 0.02 | 0.02 | 0.02 |
| | | 90% | 0.33 | 0.36 | 0.45 | 0.40 | 0.33 | 0.33 | 0.31 | 0.38 | 0.38 | 0.40 | 0.44 | 0.37 | 0.37 | 0.37 |
| $[\theta_2 = 0.01,\ \theta_3 = 0.9]$ | 0.4 | 95% | 0.21 | 0.25 | 0.33 | 0.19 | 0.22 | 0.21 | 0.20 | 0.27 | 0.28 | 0.30 | 0.33 | 0.27 | 0.26 | 0.26 |
| | | 99% | 0.08 | 0.10 | 0.14 | 0.01 | 0.08 | 0.08 | 0.05 | 0.12 | 0.10 | 0.12 | 0.13 | 0.10 | 0.10 | 0.10 |
| | | 90% | 0.75 | 0.80 | 0.87 | 0.86 | 0.75 | 0.76 | 0.74 | 0.76 | 0.77 | 0.81 | 0.85 | 0.76 | 0.76 | 0.76 |
| | 0.6 | 95% | 0.63 | 0.71 | 0.81 | 0.68 | 0.64 | 0.63 | 0.62 | 0.67 | 0.69 | 0.73 | 0.77 | 0.67 | 0.67 | 0.67 |
| | | 99% | 0.39 | 0.48 | 0.62 | 0.15 | 0.40 | 0.40 | 0.34 | 0.42 | 0.42 | 0.48 | 0.53 | 0.39 | 0.40 | 0.41 |
| | | 90% | 0.99 | 1.00 | 1.00 | 1.00 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 1.00 | 0.99 | 0.99 | 0.99 |
| | 0.8 | 95% | 0.98 | 0.99 | 1.00 | 0.98 | 0.98 | 0.98 | 0.98 | 0.97 | 0.98 | 0.99 | 0.99 | 0.97 | 0.97 | 0.97 |
| | | 99% | 0.93 | 0.96 | 0.98 | 0.81 | 0.93 | 0.93 | 0.92 | 0.92 | 0.92 | 0.95 | 0.96 | 0.91 | 0.91 | 0.92 |

Table 4.4.6: Estimated of the time of change as a percentage of the observation ($t^* = 0.1$)

| T=100 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Coefficients | $\delta$ | | Unweighted | $q_1(t, \frac{1}{11})$ | $q_1(t, \frac{3}{11})$ | $q_1(t, \frac{5}{11})$ | $q_2(t, \frac{1}{11})$ | $q_2(t, \frac{3}{11})$ | $q_2(t, \frac{5}{11})$ |
| $[\theta_2 = 0.005, \theta_3 = 0.1]$ | 0.2 | Mean | 0.52 | 0.51 | 0.46 | 0.37 | 0.52 | 0.50 | 0.47 |
| | | Med | 0.53 | 0.53 | 0.48 | 0.10 | 0.53 | 0.52 | 0.50 |
| | | SD | 0.22 | 0.25 | 0.34 | 0.40 | 0.23 | 0.25 | 0.29 |
| | 0.4 | Mean | 0.45 | 0.43 | 0.36 | 0.29 | 0.45 | 0.43 | 0.39 |
| | | Med | 0.47 | 0.43 | 0.22 | 0.07 | 0.47 | 0.45 | 0.41 |
| | | SD | 0.25 | 0.28 | 0.34 | 0.37 | 0.25 | 0.27 | 0.30 |
| | 0.6 | Mean | 0.35 | 0.31 | 0.24 | 0.19 | 0.34 | 0.31 | 0.28 |
| | | Med | 0.29 | 0.20 | 0.10 | 0.06 | 0.28 | 0.23 | 0.11 |
| | | SD | 0.25 | 0.27 | 0.29 | 0.31 | 0.25 | 0.27 | 0.28 |
| | 0.8 | Mean | 0.22 | 0.19 | 0.13 | 0.10 | 0.21 | 0.19 | 0.16 |
| | | Med | 0.12 | 0.10 | 0.08 | 0.06 | 0.10 | 0.10 | 0.08 |
| | | SD | 0.19 | 0.20 | 0.19 | 0.19 | 0.20 | 0.21 | 0.21 |
| $[\theta_2 = 0.01, \theta_3 = 0.9]$ | 0.2 | Mean | 0.53 | 0.52 | 0.48 | 0.38 | 0.52 | 0.51 | 0.48 |
| | | Med | 0.54 | 0.54 | 0.50 | 0.12 | 0.54 | 0.53 | 0.51 |
| | | SD | 0.21 | 0.24 | 0.33 | 0.40 | 0.22 | 0.24 | 0.29 |
| | 0.4 | Mean | 0.47 | 0.45 | 0.39 | 0.31 | 0.46 | 0.44 | 0.42 |
| | | Med | 0.49 | 0.47 | 0.32 | 0.08 | 0.48 | 0.47 | 0.45 |
| | | SD | 0.24 | 0.27 | 0.33 | 0.37 | 0.24 | 0.27 | 0.30 |
| | 0.6 | Mean | 0.39 | 0.36 | 0.29 | 0.23 | 0.38 | 0.36 | 0.33 |
| | | Med | 0.36 | 0.30 | 0.11 | 0.06 | 0.36 | 0.35 | 0.29 |
| | | SD | 0.24 | 0.27 | 0.31 | 0.33 | 0.25 | 0.26 | 0.29 |
| | 0.8 | Mean | 0.29 | 0.26 | 0.19 | 0.14 | 0.28 | 0.26 | 0.24 |
| | | Med | 0.21 | 0.15 | 0.10 | 0.06 | 0.20 | 0.16 | 0.10 |
| | | SD | 0.22 | 0.23 | 0.24 | 0.25 | 0.22 | 0.23 | 0.25 |

| T=500 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Coefficients | $\delta$ | | Unweighted | $q_1(t, \frac{1}{11})$ | $q_1(t, \frac{3}{11})$ | $q_1(t, \frac{5}{11})$ | $q_2(t, \frac{1}{11})$ | $q_2(t, \frac{3}{11})$ | $q_2(t, \frac{5}{11})$ |
| $[\theta_2 = 0.005, \theta_3 = 0.1]$ | 0.2 | Mean | 0.47 | 0.47 | 0.46 | 0.48 | 0.47 | 0.48 | 0.49 |
| | | Med | 0.48 | 0.46 | 0.45 | 0.42 | 0.48 | 0.48 | 0.49 |
| | | SD | 0.20 | 0.22 | 0.28 | 0.39 | 0.20 | 0.20 | 0.23 |
| | 0.4 | Mean | 0.35 | 0.32 | 0.26 | 0.25 | 0.34 | 0.34 | 0.34 |
| | | Med | 0.31 | 0.26 | 0.14 | 0.10 | 0.31 | 0.32 | 0.32 |
| | | SD | 0.21 | 0.22 | 0.25 | 0.32 | 0.21 | 0.21 | 0.24 |
| | 0.6 | Mean | 0.19 | 0.16 | 0.12 | 0.11 | 0.18 | 0.18 | 0.17 |
| | | Med | 0.13 | 0.11 | 0.10 | 0.09 | 0.12 | 0.11 | 0.10 |
| | | SD | 0.13 | 0.11 | 0.09 | 0.15 | 0.13 | 0.14 | 0.15 |
| | 0.8 | Mean | 0.12 | 0.10 | 0.10 | 0.09 | 0.11 | 0.10 | 0.10 |
| | | Med | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 |
| | | SD | 0.04 | 0.03 | 0.02 | 0.04 | 0.04 | 0.04 | 0.04 |
| $[\theta_2 = 0.01, \theta_3 = 0.9]$ | 0.2 | Mean | 0.48 | 0.48 | 0.48 | 0.51 | 0.49 | 0.49 | 0.50 |
| | | Med | 0.49 | 0.48 | 0.48 | 0.53 | 0.49 | 0.49 | 0.50 |
| | | SD | 0.20 | 0.22 | 0.28 | 0.38 | 0.20 | 0.20 | 0.22 |
| | 0.4 | Mean | 0.38 | 0.36 | 0.33 | 0.31 | 0.38 | 0.38 | 0.39 |
| | | Med | 0.36 | 0.32 | 0.21 | 0.11 | 0.36 | 0.37 | 0.38 |
| | | SD | 0.21 | 0.23 | 0.27 | 0.35 | 0.21 | 0.22 | 0.24 |
| | 0.6 | Mean | 0.25 | 0.21 | 0.16 | 0.14 | 0.24 | 0.23 | 0.23 |
| | | Med | 0.18 | 0.13 | 0.10 | 0.10 | 0.17 | 0.15 | 0.12 |
| | | SD | 0.17 | 0.16 | 0.16 | 0.20 | 0.17 | 0.17 | 0.20 |
| | 0.8 | Mean | 0.15 | 0.13 | 0.10 | 0.09 | 0.14 | 0.13 | 0.12 |
| | | Med | 0.11 | 0.10 | 0.10 | 0.09 | 0.10 | 0.10 | 0.10 |
| | | SD | 0.09 | 0.07 | 0.05 | 0.07 | 0.09 | 0.09 | 0.09 |

#### 4.4.3.2 Experiment II ($\mathfrak{t}^* = 0.3$)

Let the change-point location $\mathfrak{t}^*$ occurs at the front middle of sample, $\mathfrak{t}^* = 0.3$. Table 4.4.7 and 4.4.8 record rejection rates and estimated change locations $\hat{\nu}_t$.

The results show similar patterns with experiment $I$, all tests are correctly sized, and the rejection rates are increasing along with the change point magnitude $\delta$ and the sample size $T$ increasing. However, different from experiment $I$, the power of unweighted CUSUM tests converges to unit faster than weighted CUSUM tests. This is because the fraction change time $\mathfrak{t} = 0.3$ lies within the trimmed point $[\mathfrak{t}_1, \mathfrak{t}_2]$, causing the trimming issue to be less relevant. Thus, the technique of adding weight functions becomes less crucial to detect a middle change. To be specific in the small samples, with low dependent observations, unweighted CUSUM, weighted CUSUM with $q_1(\mathfrak{t}, \frac{1}{11})$ and weighted CUSUM with $q_2(\mathfrak{t}, \alpha)$ show relatively higher power. These results are deteriorated in the case of high dependence, and the power reaches to roughly 0.6 when a big magnitude change occurred. In large samples, regardless of whether in the low dependent or high dependent cases, the power of all tests converge to unit with change magnitude equalling to 0.6. The unweighted test, weighted test with $q_1(\mathfrak{t}, \frac{1}{11})$ and $q_2(\mathfrak{t}, \alpha)$ are more reliable in smaller sized changes.

Within the weighted CUSUM tests, those using the weight function $q_2(\mathfrak{t}, \alpha)$ perform roughly same as the unweighted one, and tests with the weight function $q_1(\mathfrak{t}, \alpha)$ converge to unit power more slowly. Particularly, because higher valued $\alpha$ put more weights in ends, tests with lower valued $\alpha$ show slower convergence rate. Nonetheless, this is no longer a problem when it comes to a large sample even with a moderate magnitude of the change. Besides, considering the estimated change time, the test with the weight function $q_1(\mathfrak{t}, 5/11)$ provides more accurate estimation of change-point locations. Other patterns are same as those described in experiment $I$.

## Table 4.4.7: Empirical rejection rates of CUSUM-typed tests ($\mathbf{t}^* = 0.3$)

T=100

| | | | $\hat{M}_T^{(1)}$ | $\hat{W}_T^{(1)}$ | | | | | | $\hat{M}_T^{(2)}$ | $\hat{W}_T^{(2)}$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Coefficients | $\delta$ | | Unweighted | $q_1(t, \frac{1}{11})$ | $q_1(t, \frac{3}{11})$ | $q_1(t, \frac{5}{11})$ | $q_2(t, \frac{1}{11})$ | $q_2(t, \frac{3}{11})$ | $q_2(t, \frac{5}{11})$ | Unweighted | $q_1(t, \frac{1}{11})$ | $q_1(t, \frac{3}{11})$ | $q_1(t, \frac{5}{11})$ | $q_2(t, \frac{1}{11})$ | $q_2(t, \frac{3}{11})$ | $q_2(t, \frac{5}{11})$ |
| | | 90% | 0.11 | 0.10 | 0.09 | 0.10 | 0.10 | 0.09 | 0.08 | 0.12 | 0.10 | 0.09 | 0.09 | 0.10 | 0.09 | 0.09 |
| | 0 | 95% | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.04 | 0.06 | 0.05 | 0.04 | 0.04 | 0.05 | 0.04 | 0.04 |
| | | 99% | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| | | 90% | 0.18 | 0.17 | 0.16 | 0.15 | 0.17 | 0.15 | 0.14 | 0.20 | 0.18 | 0.17 | 0.17 | 0.18 | 0.17 | 0.16 |
| | 0.2 | 95% | 0.10 | 0.09 | 0.08 | 0.07 | 0.09 | 0.09 | 0.08 | 0.12 | 0.10 | 0.08 | 0.08 | 0.10 | 0.09 | 0.08 |
| | | 99% | 0.02 | 0.02 | 0.01 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.01 | 0.02 | 0.02 | 0.02 |
| | | 90% | 0.38 | 0.36 | 0.32 | 0.25 | 0.36 | 0.34 | 0.30 | 0.41 | 0.38 | 0.36 | 0.35 | 0.38 | 0.36 | 0.35 |
| $[\theta_2 = 0.005,\ \theta_3 = 0.1]$ | 0.4 | 95% | 0.26 | 0.24 | 0.19 | 0.09 | 0.24 | 0.22 | 0.18 | 0.29 | 0.27 | 0.24 | 0.22 | 0.27 | 0.25 | 0.24 |
| | | 99% | 0.10 | 0.07 | 0.02 | 0.01 | 0.08 | 0.06 | 0.04 | 0.11 | 0.09 | 0.08 | 0.07 | 0.08 | 0.09 | 0.08 |
| | | 90% | 0.73 | 0.71 | 0.66 | 0.50 | 0.71 | 0.68 | 0.63 | 0.76 | 0.73 | 0.71 | 0.69 | 0.73 | 0.71 | 0.70 |
| | 0.6 | 95% | 0.61 | 0.59 | 0.49 | 0.23 | 0.59 | 0.55 | 0.47 | 0.64 | 0.61 | 0.57 | 0.54 | 0.61 | 0.58 | 0.57 |
| | | 99% | 0.34 | 0.28 | 0.10 | 0.02 | 0.29 | 0.24 | 0.17 | 0.36 | 0.32 | 0.29 | 0.25 | 0.31 | 0.32 | 0.30 |
| | | 90% | 0.97 | 0.97 | 0.95 | 0.87 | 0.97 | 0.96 | 0.94 | 0.98 | 0.98 | 0.97 | 0.96 | 0.98 | 0.97 | 0.97 |
| | 0.8 | 95% | 0.94 | 0.93 | 0.89 | 0.58 | 0.93 | 0.92 | 0.87 | 0.96 | 0.95 | 0.92 | 0.90 | 0.95 | 0.93 | 0.93 |
| | | 99% | 0.78 | 0.72 | 0.41 | 0.09 | 0.73 | 0.67 | 0.55 | 0.79 | 0.75 | 0.72 | 0.65 | 0.74 | 0.75 | 0.73 |
| | | 90% | 0.14 | 0.12 | 0.11 | 0.11 | 0.12 | 0.11 | 0.10 | 0.14 | 0.12 | 0.11 | 0.11 | 0.12 | 0.11 | 0.10 |
| | 0 | 95% | 0.06 | 0.06 | 0.05 | 0.05 | 0.06 | 0.06 | 0.05 | 0.07 | 0.07 | 0.06 | 0.05 | 0.06 | 0.05 | 0.05 |
| | | 99% | 0.02 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.02 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| | | 90% | 0.20 | 0.18 | 0.17 | 0.16 | 0.18 | 0.17 | 0.15 | 0.21 | 0.19 | 0.18 | 0.18 | 0.19 | 0.18 | 0.18 |
| | 0.2 | 95% | 0.11 | 0.11 | 0.08 | 0.07 | 0.10 | 0.10 | 0.08 | 0.13 | 0.11 | 0.10 | 0.09 | 0.11 | 0.10 | 0.10 |
| | | 99% | 0.03 | 0.02 | 0.01 | 0.02 | 0.02 | 0.02 | 0.02 | 0.03 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 |
| | | 90% | 0.36 | 0.33 | 0.29 | 0.23 | 0.33 | 0.31 | 0.28 | 0.39 | 0.35 | 0.34 | 0.33 | 0.36 | 0.34 | 0.33 |
| $[\theta_2 = 0.01,\ \theta_3 = 0.9]$ | 0.4 | 95% | 0.24 | 0.22 | 0.17 | 0.09 | 0.22 | 0.20 | 0.18 | 0.27 | 0.26 | 0.22 | 0.21 | 0.25 | 0.23 | 0.23 |
| | | 99% | 0.09 | 0.07 | 0.02 | 0.01 | 0.07 | 0.06 | 0.04 | 0.10 | 0.08 | 0.07 | 0.06 | 0.08 | 0.09 | 0.08 |
| | | 90% | 0.64 | 0.62 | 0.55 | 0.41 | 0.62 | 0.59 | 0.55 | 0.67 | 0.64 | 0.62 | 0.61 | 0.64 | 0.63 | 0.61 |
| | 0.6 | 95% | 0.50 | 0.48 | 0.40 | 0.17 | 0.48 | 0.46 | 0.40 | 0.55 | 0.52 | 0.48 | 0.45 | 0.52 | 0.49 | 0.48 |
| | | 99% | 0.26 | 0.21 | 0.06 | 0.02 | 0.22 | 0.19 | 0.14 | 0.29 | 0.25 | 0.23 | 0.19 | 0.25 | 0.25 | 0.24 |
| | | 90% | 0.91 | 0.90 | 0.86 | 0.71 | 0.90 | 0.89 | 0.85 | 0.93 | 0.92 | 0.91 | 0.90 | 0.92 | 0.92 | 0.91 |
| | 0.8 | 95% | 0.83 | 0.82 | 0.73 | 0.39 | 0.82 | 0.79 | 0.74 | 0.88 | 0.86 | 0.82 | 0.79 | 0.85 | 0.83 | 0.82 |
| | | 99% | 0.59 | 0.53 | 0.24 | 0.05 | 0.54 | 0.48 | 0.38 | 0.64 | 0.60 | 0.56 | 0.49 | 0.59 | 0.60 | 0.58 |

T=500

| | | | $\hat{M}_T^{(1)}$ | $\hat{W}_T^{(1)}$ | | | | | | $\hat{M}_T^{(2)}$ | $\hat{W}_T^{(2)}$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Coefficients | $\delta$ | | Unweighted | $q_1(t, \frac{1}{11})$ | $q_1(t, \frac{3}{11})$ | $q_1(t, \frac{5}{11})$ | $q_2(t, \frac{1}{11})$ | $q_2(t, \frac{3}{11})$ | $q_2(t, \frac{5}{11})$ | Unweighted | $q_1(t, \frac{1}{11})$ | $q_1(t, \frac{3}{11})$ | $q_1(t, \frac{5}{11})$ | $q_2(t, \frac{1}{11})$ | $q_2(t, \frac{3}{11})$ | $q_2(t, \frac{5}{11})$ |
| | | 90% | 0.09 | 0.09 | 0.08 | 0.10 | 0.09 | 0.09 | 0.09 | 0.10 | 0.09 | 0.09 | 0.08 | 0.09 | 0.09 | 0.09 |
| | 0 | 95% | 0.03 | 0.04 | 0.04 | 0.04 | 0.04 | 0.03 | 0.04 | 0.05 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 |
| | | 99% | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| | | 90% | 0.42 | 0.43 | 0.40 | 0.22 | 0.43 | 0.42 | 0.39 | 0.43 | 0.42 | 0.42 | 0.42 | 0.42 | 0.42 | 0.41 |
| | 0.2 | 95% | 0.29 | 0.30 | 0.29 | 0.08 | 0.29 | 0.28 | 0.25 | 0.32 | 0.31 | 0.30 | 0.29 | 0.32 | 0.31 | 0.31 |
| | | 99% | 0.12 | 0.13 | 0.11 | 0.01 | 0.12 | 0.12 | 0.08 | 0.14 | 0.13 | 0.13 | 0.11 | 0.13 | 0.12 | 0.13 |
| | | 90% | 0.96 | 0.96 | 0.95 | 0.85 | 0.96 | 0.96 | 0.95 | 0.95 | 0.95 | 0.95 | 0.95 | 0.95 | 0.95 | 0.94 |
| $[\theta_2 = 0.005,\ \theta_3 = 0.1]$ | 0.4 | 95% | 0.92 | 0.93 | 0.92 | 0.62 | 0.92 | 0.91 | 0.89 | 0.92 | 0.91 | 0.91 | 0.90 | 0.91 | 0.91 | 0.91 |
| | | 99% | 0.79 | 0.81 | 0.78 | 0.04 | 0.79 | 0.78 | 0.71 | 0.78 | 0.76 | 0.76 | 0.73 | 0.76 | 0.75 | 0.76 |
| | | 90% | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | 0.6 | 95% | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | | 99% | 1.00 | 1.00 | 1.00 | 0.76 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | | 90% | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | 0.8 | 95% | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | | 99% | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | | 90% | 0.14 | 0.14 | 0.14 | 0.12 | 0.14 | 0.15 | 0.13 | 0.13 | 0.12 | 0.12 | 0.12 | 0.12 | 0.13 | 0.12 |
| | 0 | 95% | 0.07 | 0.08 | 0.07 | 0.05 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 |
| | | 99% | 0.01 | 0.02 | 0.01 | 0.01 | 0.01 | 0.02 | 0.01 | 0.02 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| | | 90% | 0.38 | 0.38 | 0.36 | 0.22 | 0.38 | 0.37 | 0.33 | 0.38 | 0.36 | 0.36 | 0.37 | 0.37 | 0.37 | 0.36 |
| | 0.2 | 95% | 0.26 | 0.27 | 0.27 | 0.08 | 0.26 | 0.25 | 0.22 | 0.28 | 0.28 | 0.27 | 0.26 | 0.27 | 0.27 | 0.27 |
| | | 99% | 0.11 | 0.11 | 0.09 | 0.00 | 0.10 | 0.10 | 0.07 | 0.11 | 0.10 | 0.10 | 0.09 | 0.10 | 0.10 | 0.10 |
| | | 90% | 0.91 | 0.91 | 0.90 | 0.74 | 0.90 | 0.90 | 0.88 | 0.89 | 0.89 | 0.89 | 0.89 | 0.89 | 0.89 | 0.89 |
| $[\theta_2 = 0.01,\ \theta_3 = 0.9]$ | 0.4 | 95% | 0.84 | 0.85 | 0.85 | 0.47 | 0.85 | 0.83 | 0.80 | 0.83 | 0.83 | 0.83 | 0.82 | 0.83 | 0.83 | 0.82 |
| | | 99% | 0.66 | 0.67 | 0.64 | 0.02 | 0.66 | 0.65 | 0.57 | 0.66 | 0.62 | 0.63 | 0.60 | 0.62 | 0.62 | 0.63 |
| | | 90% | 1.00 | 1.00 | 1.00 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | 0.6 | 95% | 1.00 | 1.00 | 1.00 | 0.97 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | | 99% | 0.99 | 0.99 | 0.99 | 0.46 | 0.99 | 0.99 | 0.98 | 0.99 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 |
| | | 90% | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | 0.8 | 95% | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | | 99% | 1.00 | 1.00 | 1.00 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |

Table 4.4.8: Estimated of the time of change as a percentage of the observation ($t^* = 0.3$)

| T=100 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Coefficients | $\delta$ | | Unweighted | $q_1(t, \frac{1}{11})$ | $q_1(t, \frac{3}{11})$ | $q_1(t, \frac{5}{11})$ | $q_2(t, \frac{1}{11})$ | $q_2(t, \frac{3}{11})$ | $q_2(t, \frac{5}{11})$ |
| $[\theta_2 = 0.005, \theta_3 = 0.1]$ | 0.2 | Mean | 0.51 | 0.52 | 0.53 | 0.50 | 0.51 | 0.51 | 0.52 |
| | | Med | 0.50 | 0.51 | 0.53 | 0.49 | 0.50 | 0.50 | 0.51 |
| | | SD | 0.20 | 0.22 | 0.29 | 0.37 | 0.20 | 0.21 | 0.23 |
| | 0.4 | Mean | 0.43 | 0.43 | 0.43 | 0.40 | 0.43 | 0.44 | 0.44 |
| | | Med | 0.38 | 0.37 | 0.34 | 0.30 | 0.39 | 0.40 | 0.41 |
| | | SD | 0.18 | 0.20 | 0.26 | 0.33 | 0.18 | 0.19 | 0.21 |
| | 0.6 | Mean | 0.35 | 0.34 | 0.34 | 0.32 | 0.35 | 0.36 | 0.36 |
| | | Med | 0.31 | 0.30 | 0.30 | 0.28 | 0.31 | 0.32 | 0.33 |
| | | SD | 0.12 | 0.14 | 0.19 | 0.26 | 0.13 | 0.13 | 0.15 |
| | 0.8 | Mean | 0.30 | 0.30 | 0.29 | 0.27 | 0.31 | 0.31 | 0.32 |
| | | Med | 0.30 | 0.30 | 0.28 | 0.27 | 0.30 | 0.30 | 0.30 |
| | | SD | 0.07 | 0.07 | 0.11 | 0.17 | 0.07 | 0.08 | 0.09 |
| $[\theta_2 = 0.01, \theta_3 = 0.9]$ | 0.2 | Mean | 0.52 | 0.53 | 0.54 | 0.51 | 0.52 | 0.52 | 0.52 |
| | | Med | 0.51 | 0.52 | 0.55 | 0.51 | 0.51 | 0.51 | 0.52 |
| | | SD | 0.20 | 0.22 | 0.28 | 0.37 | 0.20 | 0.21 | 0.23 |
| | 0.4 | Mean | 0.46 | 0.46 | 0.46 | 0.43 | 0.46 | 0.46 | 0.47 |
| | | Med | 0.42 | 0.42 | 0.39 | 0.33 | 0.43 | 0.44 | 0.45 |
| | | SD | 0.19 | 0.20 | 0.26 | 0.34 | 0.19 | 0.19 | 0.21 |
| | 0.6 | Mean | 0.39 | 0.38 | 0.38 | 0.37 | 0.39 | 0.40 | 0.40 |
| | | Med | 0.34 | 0.33 | 0.31 | 0.30 | 0.35 | 0.36 | 0.38 |
| | | SD | 0.15 | 0.16 | 0.21 | 0.29 | 0.15 | 0.15 | 0.17 |
| | 0.8 | Mean | 0.34 | 0.33 | 0.32 | 0.30 | 0.34 | 0.35 | 0.36 |
| | | Med | 0.31 | 0.31 | 0.30 | 0.28 | 0.32 | 0.32 | 0.34 |
| | | SD | 0.10 | 0.11 | 0.15 | 0.22 | 0.10 | 0.11 | 0.12 |

| T=500 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Coefficients | $\delta$ | | Unweighted | $q_1(t, \frac{1}{11})$ | $q_1(t, \frac{3}{11})$ | $q_1(t, \frac{5}{11})$ | $q_2(t, \frac{1}{11})$ | $q_2(t, \frac{3}{11})$ | $q_2(t, \frac{5}{11})$ |
| $[\theta_2 = 0.005, \theta_3 = 0.1]$ | 0.2 | Mean | 0.43 | 0.42 | 0.43 | 0.47 | 0.43 | 0.43 | 0.45 |
| | | Med | 0.39 | 0.38 | 0.36 | 0.35 | 0.39 | 0.40 | 0.42 |
| | | SD | 0.16 | 0.17 | 0.22 | 0.32 | 0.16 | 0.16 | 0.18 |
| | 0.4 | Mean | 0.33 | 0.33 | 0.31 | 0.32 | 0.34 | 0.34 | 0.35 |
| | | Med | 0.31 | 0.30 | 0.30 | 0.30 | 0.31 | 0.32 | 0.32 |
| | | SD | 0.07 | 0.07 | 0.08 | 0.17 | 0.08 | 0.08 | 0.09 |
| | 0.6 | Mean | 0.30 | 0.30 | 0.30 | 0.29 | 0.31 | 0.31 | 0.32 |
| | | Med | 0.30 | 0.30 | 0.30 | 0.29 | 0.30 | 0.30 | 0.30 |
| | | SD | 0.03 | 0.02 | 0.02 | 0.07 | 0.03 | 0.03 | 0.04 |
| | 0.8 | Mean | 0.30 | 0.30 | 0.29 | 0.29 | 0.30 | 0.30 | 0.30 |
| | | Med | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 |
| | | SD | 0.01 | 0.01 | 0.01 | 0.02 | 0.01 | 0.01 | 0.01 |
| $[\theta_2 = 0.01, \theta_3 = 0.9]$ | 0.2 | Mean | 0.44 | 0.44 | 0.45 | 0.50 | 0.45 | 0.45 | 0.46 |
| | | Med | 0.40 | 0.40 | 0.38 | 0.38 | 0.41 | 0.42 | 0.43 |
| | | SD | 0.18 | 0.19 | 0.23 | 0.33 | 0.17 | 0.17 | 0.19 |
| | 0.4 | Mean | 0.35 | 0.34 | 0.33 | 0.34 | 0.35 | 0.36 | 0.36 |
| | | Med | 0.31 | 0.31 | 0.30 | 0.30 | 0.32 | 0.32 | 0.33 |
| | | SD | 0.10 | 0.10 | 0.12 | 0.20 | 0.10 | 0.10 | 0.11 |
| | 0.6 | Mean | 0.31 | 0.31 | 0.30 | 0.30 | 0.31 | 0.32 | 0.33 |
| | | Med | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.31 |
| | | SD | 0.04 | 0.04 | 0.04 | 0.08 | 0.04 | 0.05 | 0.05 |
| | 0.8 | Mean | 0.30 | 0.30 | 0.29 | 0.29 | 0.30 | 0.30 | 0.31 |
| | | Med | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 |
| | | SD | 0.02 | 0.01 | 0.02 | 0.04 | 0.02 | 0.02 | 0.02 |

#### 4.4.3.3 Experiment III ($\mathfrak{t}^* = 0.5$)

In this experiment, the change-point location $\mathfrak{t}^*$ is set at the exactly centre of sample, $\mathfrak{t}^* = 0.5$. Table 4.4.9 and 4.4.10 show rejection rates and estimated change-point locations.

Unweighted CUSUM tests outperform all others, and earn the fastest unit power convergence. Compared with experiment $II$, the fraction change-point $\mathfrak{t}$ is even far away from trimmed point $[\mathfrak{t}_1, \mathfrak{t}_2]$, and weighted CUSUM tests are supposed to show their side effect - insufficient weights in the centre. However, this side effect is still in a controllable level. Weighted CUSUM tests, even using the weight function $q_1(\mathfrak{t}, 5/11)$, present good power in large samples, or in the small sample with a sizeable change. All tests show precise $\hat{\nu}$, but when using the test with $q_1(\mathfrak{t}, 5/11)$, the estimated $\hat{\nu}$ has higher standard deviations. Other patterns of estimated change points are consistent with those described in experiment $I$ and $II$.

According to Table 4.4.9, in the small samples, the power of the unweighted CUSUM test $\hat{M}_T^{(1)}$ reach to 0.85 and 0.62 at 99% significance level with weak dependent and strong dependent observations subject to a largest change, respectively. As a maximally selected CUSUM statistics, $\hat{M}_T^{(1)}$ show slightly higher rejection rate than self-normalized CUSUM statistics $\hat{M}_T^{(2)}$ with weak dependent observations, which reaches to 0.82 at 99% significance level. As mentioned above, the weight function $q_2(\mathfrak{t}, \alpha)$ deliver similar performance as unweighted CUSUM tests. Furthermore, the power of weighted CUSUM test with $q_1(\mathfrak{t}, \frac{1}{11})$ reaches to 0.77 and 0.52 with weak and strong dependent observations, respectively for $\hat{W}_T^{(1)}$, and 0.77 and 0.56 for $\hat{W}_T^{(2)}$. In the large samples, both two unweighted CUSUM tests obtain unit power with moderate valued changes, and all tests obtain unit power with sizeable changes.

## Table 4.4.9: Empirical rejection rates of CUSUM-typed tests ($\mathfrak{t}^* = 0.5$)

T=100

| | | | $\hat{M}_T^{(1)}$ | | | $\hat{W}_T^{(1)}$ | | | | $\hat{M}_T^{(2)}$ | | | $\hat{W}_T^{(2)}$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Coefficients | $\delta$ | | Unweighted | $q_1(t, \frac{1}{11})$ | $q_1(t, \frac{3}{11})$ | $q_1(t, \frac{5}{11})$ | $q_2(t, \frac{1}{11})$ | $q_2(t, \frac{3}{11})$ | $q_2(t, \frac{5}{11})$ | Unweighted | $q_1(t, \frac{1}{11})$ | $q_1(t, \frac{3}{11})$ | $q_1(t, \frac{5}{11})$ | $q_2(t, \frac{1}{11})$ | $q_2(t, \frac{3}{11})$ | $q_2(t, \frac{5}{11})$ |
| | | 90% | 0.11 | 0.10 | 0.09 | 0.10 | 0.10 | 0.09 | 0.08 | 0.12 | 0.10 | 0.09 | 0.09 | 0.10 | 0.09 | 0.09 |
| | 0 | 95% | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.04 | 0.06 | 0.05 | 0.04 | 0.04 | 0.05 | 0.04 | 0.04 |
| | | 99% | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| | | 90% | 0.19 | 0.17 | 0.14 | 0.13 | 0.18 | 0.17 | 0.15 | 0.20 | 0.17 | 0.16 | 0.15 | 0.17 | 0.16 | 0.15 |
| | 0.2 | 95% | 0.11 | 0.10 | 0.07 | 0.05 | 0.10 | 0.10 | 0.09 | 0.11 | 0.10 | 0.08 | 0.08 | 0.10 | 0.09 | 0.08 |
| | | 99% | 0.03 | 0.02 | 0.01 | 0.01 | 0.02 | 0.02 | 0.02 | 0.03 | 0.02 | 0.02 | 0.01 | 0.02 | 0.02 | 0.02 |
| | | 90% | 0.47 | 0.44 | 0.35 | 0.19 | 0.46 | 0.44 | 0.41 | 0.48 | 0.44 | 0.41 | 0.38 | 0.45 | 0.43 | 0.42 |
| $[\theta_2 = 0.005,\ \theta_3 = 0.1]$ | 0.4 | 95% | 0.34 | 0.31 | 0.20 | 0.07 | 0.33 | 0.32 | 0.26 | 0.35 | 0.32 | 0.27 | 0.24 | 0.32 | 0.30 | 0.29 |
| | | 99% | 0.13 | 0.10 | 0.02 | 0.01 | 0.11 | 0.10 | 0.07 | 0.14 | 0.11 | 0.09 | 0.07 | 0.11 | 0.12 | 0.11 |
| | | 90% | 0.84 | 0.80 | 0.71 | 0.42 | 0.82 | 0.81 | 0.78 | 0.83 | 0.80 | 0.77 | 0.73 | 0.81 | 0.80 | 0.79 |
| | 0.6 | 95% | 0.72 | 0.68 | 0.51 | 0.14 | 0.71 | 0.70 | 0.65 | 0.72 | 0.69 | 0.61 | 0.56 | 0.69 | 0.66 | 0.66 |
| | | 99% | 0.43 | 0.33 | 0.09 | 0.02 | 0.39 | 0.35 | 0.28 | 0.42 | 0.37 | 0.32 | 0.24 | 0.37 | 0.37 | 0.37 |
| | | 90% | 0.99 | 0.99 | 0.97 | 0.82 | 0.99 | 0.99 | 0.99 | 0.99 | 0.98 | 0.98 | 0.97 | 0.98 | 0.98 | 0.98 |
| | 0.8 | 95% | 0.97 | 0.96 | 0.91 | 0.42 | 0.97 | 0.97 | 0.96 | 0.97 | 0.96 | 0.94 | 0.90 | 0.97 | 0.96 | 0.95 |
| | | 99% | 0.85 | 0.77 | 0.33 | 0.02 | 0.82 | 0.80 | 0.71 | 0.82 | 0.77 | 0.72 | 0.61 | 0.78 | 0.79 | 0.78 |
| | | 90% | 0.14 | 0.12 | 0.11 | 0.11 | 0.12 | 0.11 | 0.10 | 0.14 | 0.12 | 0.11 | 0.11 | 0.12 | 0.11 | 0.10 |
| | 0 | 95% | 0.06 | 0.06 | 0.05 | 0.05 | 0.06 | 0.06 | 0.05 | 0.07 | 0.07 | 0.06 | 0.05 | 0.06 | 0.05 | 0.05 |
| | | 99% | 0.02 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.02 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| | | 90% | 0.20 | 0.18 | 0.15 | 0.13 | 0.19 | 0.18 | 0.16 | 0.21 | 0.18 | 0.17 | 0.16 | 0.18 | 0.17 | 0.16 |
| | 0.2 | 95% | 0.12 | 0.11 | 0.08 | 0.06 | 0.11 | 0.10 | 0.09 | 0.12 | 0.11 | 0.09 | 0.08 | 0.11 | 0.10 | 0.09 |
| | | 99% | 0.03 | 0.02 | 0.01 | 0.01 | 0.02 | 0.02 | 0.02 | 0.03 | 0.03 | 0.02 | 0.02 | 0.03 | 0.03 | 0.03 |
| | | 90% | 0.41 | 0.38 | 0.30 | 0.18 | 0.39 | 0.37 | 0.35 | 0.42 | 0.39 | 0.35 | 0.33 | 0.39 | 0.38 | 0.36 |
| $[\theta_2 = 0.01,\ \theta_3 = 0.9]$ | 0.4 | 95% | 0.29 | 0.26 | 0.17 | 0.06 | 0.28 | 0.26 | 0.22 | 0.30 | 0.27 | 0.23 | 0.20 | 0.27 | 0.25 | 0.24 |
| | | 99% | 0.10 | 0.07 | 0.01 | 0.01 | 0.09 | 0.07 | 0.06 | 0.11 | 0.09 | 0.07 | 0.06 | 0.09 | 0.09 | 0.09 |
| | | 90% | 0.70 | 0.66 | 0.56 | 0.32 | 0.68 | 0.67 | 0.63 | 0.71 | 0.68 | 0.64 | 0.60 | 0.69 | 0.66 | 0.65 |
| | 0.6 | 95% | 0.57 | 0.53 | 0.38 | 0.11 | 0.55 | 0.54 | 0.48 | 0.58 | 0.55 | 0.48 | 0.43 | 0.55 | 0.52 | 0.51 |
| | | 99% | 0.30 | 0.22 | 0.05 | 0.02 | 0.26 | 0.23 | 0.17 | 0.30 | 0.26 | 0.23 | 0.17 | 0.26 | 0.27 | 0.26 |
| | | 90% | 0.93 | 0.92 | 0.85 | 0.59 | 0.93 | 0.92 | 0.91 | 0.94 | 0.92 | 0.90 | 0.87 | 0.93 | 0.92 | 0.91 |
| | 0.8 | 95% | 0.87 | 0.84 | 0.71 | 0.25 | 0.86 | 0.85 | 0.80 | 0.87 | 0.85 | 0.80 | 0.75 | 0.85 | 0.83 | 0.82 |
| | | 99% | 0.62 | 0.52 | 0.17 | 0.01 | 0.57 | 0.54 | 0.45 | 0.62 | 0.56 | 0.50 | 0.42 | 0.56 | 0.57 | 0.56 |

T=500

| | | | $\hat{M}_T^{(1)}$ | | | $\hat{W}_T^{(1)}$ | | | | $\hat{M}_T^{(2)}$ | | | $\hat{W}_T^{(2)}$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Coefficients | $\delta$ | | Unweighted | $q_1(t, \frac{1}{11})$ | $q_1(t, \frac{3}{11})$ | $q_1(t, \frac{5}{11})$ | $q_2(t, \frac{1}{11})$ | $q_2(t, \frac{3}{11})$ | $q_2(t, \frac{5}{11})$ | Unweighted | $q_1(t, \frac{1}{11})$ | $q_1(t, \frac{3}{11})$ | $q_1(t, \frac{5}{11})$ | $q_2(t, \frac{1}{11})$ | $q_2(t, \frac{3}{11})$ | $q_2(t, \frac{5}{11})$ |
| | | 90% | 0.09 | 0.09 | 0.08 | 0.10 | 0.09 | 0.09 | 0.09 | 0.10 | 0.09 | 0.09 | 0.08 | 0.09 | 0.09 | 0.09 |
| | 0 | 95% | 0.03 | 0.04 | 0.04 | 0.04 | 0.04 | 0.03 | 0.04 | 0.05 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 |
| | | 99% | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| | | 90% | 0.55 | 0.54 | 0.48 | 0.26 | 0.55 | 0.56 | 0.54 | 0.53 | 0.51 | 0.50 | 0.49 | 0.52 | 0.52 | 0.52 |
| | 0.2 | 95% | 0.41 | 0.41 | 0.37 | 0.08 | 0.42 | 0.42 | 0.40 | 0.41 | 0.40 | 0.39 | 0.36 | 0.41 | 0.41 | 0.40 |
| | | 99% | 0.21 | 0.21 | 0.16 | 0.01 | 0.22 | 0.22 | 0.18 | 0.21 | 0.18 | 0.18 | 0.16 | 0.18 | 0.18 | 0.20 |
| | | 90% | 0.99 | 0.99 | 0.98 | 0.90 | 0.99 | 0.99 | 0.99 | 0.99 | 0.98 | 0.98 | 0.98 | 0.98 | 0.99 | 0.99 |
| $[\theta_2 = 0.005,\ \theta_3 = 0.1]$ | 0.4 | 95% | 0.98 | 0.98 | 0.97 | 0.68 | 0.98 | 0.98 | 0.97 | 0.97 | 0.97 | 0.96 | 0.95 | 0.97 | 0.97 | 0.97 |
| | | 99% | 0.92 | 0.92 | 0.89 | 0.03 | 0.92 | 0.93 | 0.91 | 0.90 | 0.87 | 0.86 | 0.83 | 0.88 | 0.88 | 0.89 |
| | | 90% | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | 0.6 | 95% | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | | 99% | 1.00 | 1.00 | 1.00 | 0.74 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | | 90% | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | 0.8 | 95% | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | | 99% | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | | 90% | 0.12 | 0.13 | 0.12 | 0.12 | 0.13 | 0.13 | 0.12 | 0.12 | 0.12 | 0.12 | 0.12 | 0.12 | 0.12 | 0.12 |
| | 0 | 95% | 0.06 | 0.07 | 0.06 | 0.05 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 |
| | | 99% | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| | | 90% | 0.47 | 0.47 | 0.43 | 0.24 | 0.48 | 0.48 | 0.46 | 0.47 | 0.45 | 0.44 | 0.43 | 0.46 | 0.46 | 0.46 |
| | 0.2 | 95% | 0.34 | 0.35 | 0.31 | 0.08 | 0.35 | 0.35 | 0.33 | 0.34 | 0.33 | 0.32 | 0.31 | 0.34 | 0.34 | 0.33 |
| | | 99% | 0.16 | 0.17 | 0.14 | 0.01 | 0.17 | 0.18 | 0.15 | 0.17 | 0.15 | 0.15 | 0.13 | 0.15 | 0.15 | 0.16 |
| | | 90% | 0.97 | 0.96 | 0.94 | 0.80 | 0.97 | 0.97 | 0.96 | 0.95 | 0.95 | 0.95 | 0.94 | 0.95 | 0.95 | 0.95 |
| $[\theta_2 = 0.01,\ \theta_3 = 0.9]$ | 0.4 | 95% | 0.92 | 0.92 | 0.90 | 0.51 | 0.92 | 0.93 | 0.92 | 0.92 | 0.91 | 0.90 | 0.88 | 0.92 | 0.92 | 0.91 |
| | | 99% | 0.81 | 0.81 | 0.75 | 0.02 | 0.81 | 0.82 | 0.79 | 0.78 | 0.75 | 0.74 | 0.71 | 0.75 | 0.75 | 0.77 |
| | | 90% | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | 0.6 | 95% | 1.00 | 1.00 | 1.00 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | | 99% | 1.00 | 1.00 | 1.00 | 0.42 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.99 | 1.00 | 1.00 | 1.00 |
| | | 90% | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | 0.8 | 95% | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | | 99% | 1.00 | 1.00 | 1.00 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |

Table 4.4.10: Estimated of the time of change as a percentage of the observation ($t^* = 0.5$)

| T=100 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Coefficients | $\delta$ | | Unweighted | $q_1(t, \frac{1}{11})$ | $q_1(t, \frac{3}{11})$ | $q_1(t, \frac{5}{11})$ | $q_2(t, \frac{1}{11})$ | $q_2(t, \frac{3}{11})$ | $q_2(t, \frac{5}{11})$ |
| $[\theta_2 = 0.005, \theta_3 = 0.1]$ | 0.2 | Mean | 0.53 | 0.54 | 0.56 | 0.56 | 0.53 | 0.54 | 0.55 |
| | | Med | 0.52 | 0.53 | 0.56 | 0.59 | 0.52 | 0.52 | 0.53 |
| | | SD | 0.18 | 0.20 | 0.26 | 0.34 | 0.18 | 0.19 | 0.21 |
| | 0.4 | Mean | 0.50 | 0.50 | 0.52 | 0.52 | 0.50 | 0.51 | 0.51 |
| | | Med | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 |
| | | SD | 0.14 | 0.16 | 0.21 | 0.29 | 0.14 | 0.14 | 0.15 |
| | 0.6 | Mean | 0.48 | 0.48 | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 |
| | | Med | 0.48 | 0.48 | 0.48 | 0.48 | 0.48 | 0.49 | 0.50 |
| | | SD | 0.09 | 0.10 | 0.14 | 0.23 | 0.09 | 0.09 | 0.10 |
| | 0.8 | Mean | 0.48 | 0.48 | 0.48 | 0.48 | 0.48 | 0.48 | 0.48 |
| | | Med | 0.48 | 0.48 | 0.48 | 0.48 | 0.48 | 0.48 | 0.48 |
| | | SD | 0.06 | 0.07 | 0.10 | 0.17 | 0.06 | 0.05 | 0.06 |
| $[\theta_2 = 0.01, \theta_3 = 0.9]$ | 0.2 | Mean | 0.53 | 0.54 | 0.57 | 0.56 | 0.54 | 0.54 | 0.55 |
| | | Med | 0.53 | 0.54 | 0.57 | 0.60 | 0.53 | 0.53 | 0.53 |
| | | SD | 0.18 | 0.20 | 0.26 | 0.34 | 0.18 | 0.19 | 0.21 |
| | 0.4 | Mean | 0.52 | 0.52 | 0.54 | 0.54 | 0.52 | 0.52 | 0.53 |
| | | Med | 0.51 | 0.51 | 0.52 | 0.52 | 0.51 | 0.51 | 0.51 |
| | | SD | 0.15 | 0.17 | 0.22 | 0.30 | 0.15 | 0.16 | 0.17 |
| | 0.6 | Mean | 0.51 | 0.51 | 0.52 | 0.52 | 0.51 | 0.51 | 0.51 |
| | | Med | 0.50 | 0.50 | 0.50 | 0.51 | 0.50 | 0.50 | 0.50 |
| | | SD | 0.12 | 0.13 | 0.17 | 0.25 | 0.12 | 0.12 | 0.12 |
| | 0.8 | Mean | 0.50 | 0.50 | 0.50 | 0.51 | 0.50 | 0.50 | 0.50 |
| | | Med | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 |
| | | SD | 0.09 | 0.10 | 0.13 | 0.20 | 0.08 | 0.08 | 0.08 |

| T=500 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Coefficients | $\delta$ | | Unweighted | $q_1(t, \frac{1}{11})$ | $q_1(t, \frac{3}{11})$ | $q_1(t, \frac{5}{11})$ | $q_2(t, \frac{1}{11})$ | $q_2(t, \frac{3}{11})$ | $q_2(t, \frac{5}{11})$ |
| $[\theta_2 = 0.005, \theta_3 = 0.1]$ | 0.2 | Mean | 0.50 | 0.50 | 0.51 | 0.54 | 0.50 | 0.50 | 0.51 |
| | | Med | 0.50 | 0.50 | 0.50 | 0.51 | 0.50 | 0.50 | 0.50 |
| | | SD | 0.12 | 0.13 | 0.18 | 0.28 | 0.12 | 0.12 | 0.13 |
| | 0.4 | Mean | 0.49 | 0.49 | 0.49 | 0.50 | 0.49 | 0.49 | 0.49 |
| | | Med | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 |
| | | SD | 0.05 | 0.05 | 0.06 | 0.13 | 0.05 | 0.04 | 0.04 |
| | 0.6 | Mean | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 |
| | | Med | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 |
| | | SD | 0.02 | 0.02 | 0.03 | 0.05 | 0.02 | 0.02 | 0.02 |
| | 0.8 | Mean | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 |
| | | Med | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 |
| | | SD | 0.01 | 0.01 | 0.01 | 0.02 | 0.01 | 0.01 | 0.01 |
| $[\theta_2 = 0.01, \theta_3 = 0.9]$ | 0.2 | Mean | 0.51 | 0.50 | 0.52 | 0.57 | 0.50 | 0.50 | 0.51 |
| | | Med | 0.50 | 0.50 | 0.50 | 0.53 | 0.50 | 0.50 | 0.50 |
| | | SD | 0.14 | 0.15 | 0.19 | 0.29 | 0.13 | 0.13 | 0.15 |
| | 0.4 | Mean | 0.49 | 0.49 | 0.49 | 0.51 | 0.49 | 0.50 | 0.50 |
| | | Med | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 |
| | | SD | 0.06 | 0.07 | 0.09 | 0.17 | 0.06 | 0.06 | 0.06 |
| | 0.6 | Mean | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 |
| | | Med | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 |
| | | SD | 0.03 | 0.03 | 0.03 | 0.07 | 0.03 | 0.03 | 0.02 |
| | 0.8 | Mean | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 |
| | | Med | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 |
| | | SD | 0.02 | 0.02 | 0.02 | 0.03 | 0.02 | 0.01 | 0.01 |

#### 4.4.3.4 Experiment IV ($\mathsf{t}^* = 0.7$)

Now consider the case where the change-point location $\mathsf{t}^*$ occurs at the rear middle of sample, $\mathsf{t}^* = 0.7$. Table 4.4.11 and 4.4.12 show rejection rates and estimated change-point locations. Due to that the fraction change-point $\mathsf{t}$ is located within the trimmed point, $[\mathsf{t}_1, \mathsf{t}_2]$, results are similar with those presented for experiment $II$. Weighted CUSUM tests using the weight function $q_2(\mathsf{t}, \alpha)$ provides similar results with unweighted CUSUM tests. Weight functions slightly weaken the power of tests, but such effect is overcome in large samples or in small samples with larger changes. With regard to estimated change-point locations, the test with the weight function $q_1(\mathsf{t}, 5/11)$ provides more accurate estimated change-point time $\hat{\nu}_T$. Other patterns of the rejection rates and the estimated change point locations are consistent with those demonstrated in experiment $II$.

Discussing the results from Table 4.4.11 in some details, in the smaller samples, in the case of a larger change, the unweighted CUSUM tests $\hat{M}_T^{(1)}$ and $\hat{M}_T^{(2)}$ have power 0.5 and 0.48 at 99% significance level with weak dependent series, respectively. Once it comes to the series with strong dependence, the power reduce to 0.22 and 0.24, respectively. Other weighted CUSUM tests performed even worse than the unweighted ones, for instance, the powers of $\hat{W}_T^{(1)}$ and $\hat{W}_T^{(2)}$ with $q_1(\mathsf{t}, \frac{1}{11})$ reach to 0.41 and 0.42 with weakly dependent data, and 0.17 and 0.19 with strongly dependent data. In the larger samples, all tests obtain unit power with a large change, and the unweighted test and weighted test with $q_1(\mathsf{t}, \frac{1}{11})$ and $q_2(\mathsf{t}, \alpha)$ exhibit relatively fast convergence.

Table 4.4.11: Empirical rejection rates for CUSUM-typed tests ($\mathfrak{t}^* = 0.7$)

T=100

| | | | $\hat{M}_T^{(1)}$ | $\hat{W}_T^{(1)}$ | | | | | | $\hat{M}_T^{(2)}$ | $\hat{W}_T^{(2)}$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Coefficients | $\delta$ | | Unweighted | $q_1(t,\frac{1}{11})$ | $q_1(t,\frac{3}{11})$ | $q_1(t,\frac{5}{11})$ | $q_2(t,\frac{1}{11})$ | $q_2(t,\frac{3}{11})$ | $q_2(t,\frac{5}{11})$ | Unweighted | $q_1(t,\frac{1}{11})$ | $q_1(t,\frac{3}{11})$ | $q_1(t,\frac{5}{11})$ | $q_2(t,\frac{1}{11})$ | $q_2(t,\frac{3}{11})$ | $q_2(t,\frac{5}{11})$ |
| | | 90% | 0.11 | 0.10 | 0.09 | 0.10 | 0.10 | 0.09 | 0.08 | 0.12 | 0.10 | 0.09 | 0.09 | 0.10 | 0.09 | 0.09 |
| | 0 | 95% | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.04 | 0.05 | 0.05 | 0.04 | 0.04 | 0.05 | 0.04 | 0.04 |
| | | 99% | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| | | 90% | 0.17 | 0.15 | 0.13 | 0.11 | 0.15 | 0.14 | 0.12 | 0.16 | 0.14 | 0.13 | 0.13 | 0.14 | 0.13 | 0.13 |
| | 0.2 | 95% | 0.09 | 0.08 | 0.06 | 0.05 | 0.08 | 0.08 | 0.06 | 0.09 | 0.08 | 0.06 | 0.06 | 0.08 | 0.07 | 0.07 |
| | | 99% | 0.03 | 0.02 | 0.01 | 0.01 | 0.02 | 0.02 | 0.01 | 0.02 | 0.02 | 0.02 | 0.01 | 0.02 | 0.02 | 0.02 |
| | | 90% | 0.35 | 0.33 | 0.26 | 0.15 | 0.33 | 0.31 | 0.27 | 0.36 | 0.32 | 0.30 | 0.28 | 0.33 | 0.31 | 0.30 |
| $[\theta_2 = 0.005,\ \theta_3 = 0.1]$ | 0.4 | 95% | 0.24 | 0.22 | 0.14 | 0.06 | 0.22 | 0.20 | 0.16 | 0.24 | 0.22 | 0.18 | 0.17 | 0.22 | 0.19 | 0.18 |
| | | 99% | 0.07 | 0.05 | 0.02 | 0.01 | 0.05 | 0.05 | 0.03 | 0.07 | 0.06 | 0.05 | 0.04 | 0.06 | 0.06 | 0.06 |
| | | 90% | 0.64 | 0.61 | 0.51 | 0.29 | 0.61 | 0.58 | 0.52 | 0.64 | 0.60 | 0.57 | 0.54 | 0.60 | 0.58 | 0.57 |
| | 0.6 | 95% | 0.49 | 0.46 | 0.33 | 0.09 | 0.46 | 0.44 | 0.37 | 0.49 | 0.47 | 0.40 | 0.36 | 0.46 | 0.42 | 0.41 |
| | | 99% | 0.22 | 0.16 | 0.04 | 0.01 | 0.18 | 0.15 | 0.09 | 0.23 | 0.19 | 0.16 | 0.12 | 0.18 | 0.19 | 0.18 |
| | | 90% | 0.92 | 0.90 | 0.83 | 0.56 | 0.90 | 0.88 | 0.84 | 0.89 | 0.87 | 0.85 | 0.83 | 0.87 | 0.86 | 0.84 |
| | 0.8 | 95% | 0.82 | 0.79 | 0.65 | 0.20 | 0.79 | 0.76 | 0.69 | 0.79 | 0.76 | 0.70 | 0.66 | 0.76 | 0.73 | 0.71 |
| | | 99% | 0.50 | 0.41 | 0.11 | 0.02 | 0.44 | 0.38 | 0.28 | 0.48 | 0.42 | 0.37 | 0.29 | 0.41 | 0.42 | 0.40 |
| | | 90% | 0.14 | 0.13 | 0.11 | 0.11 | 0.13 | 0.12 | 0.11 | 0.14 | 0.12 | 0.11 | 0.11 | 0.12 | 0.11 | 0.11 |
| | 0 | 95% | 0.07 | 0.07 | 0.05 | 0.05 | 0.07 | 0.06 | 0.05 | 0.08 | 0.07 | 0.05 | 0.05 | 0.07 | 0.06 | 0.06 |
| | | 99% | 0.02 | 0.01 | 0.00 | 0.01 | 0.01 | 0.01 | 0.01 | 0.02 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| | | 90% | 0.16 | 0.15 | 0.12 | 0.11 | 0.14 | 0.13 | 0.11 | 0.17 | 0.14 | 0.13 | 0.12 | 0.14 | 0.13 | 0.13 |
| | 0.2 | 95% | 0.09 | 0.07 | 0.05 | 0.05 | 0.07 | 0.07 | 0.06 | 0.09 | 0.07 | 0.06 | 0.05 | 0.07 | 0.06 | 0.06 |
| | | 99% | 0.02 | 0.01 | 0.01 | 0.01 | 0.02 | 0.01 | 0.01 | 0.02 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| | | 90% | 0.27 | 0.25 | 0.21 | 0.15 | 0.26 | 0.24 | 0.21 | 0.28 | 0.25 | 0.23 | 0.22 | 0.25 | 0.23 | 0.22 |
| $[\theta_2 = 0.01,\ \theta_3 = 0.9]$ | 0.4 | 95% | 0.17 | 0.16 | 0.12 | 0.06 | 0.16 | 0.15 | 0.12 | 0.18 | 0.16 | 0.14 | 0.13 | 0.17 | 0.15 | 0.14 |
| | | 99% | 0.05 | 0.04 | 0.01 | 0.01 | 0.04 | 0.04 | 0.02 | 0.05 | 0.04 | 0.03 | 0.03 | 0.04 | 0.04 | 0.04 |
| | | 90% | 0.42 | 0.39 | 0.31 | 0.19 | 0.40 | 0.36 | 0.32 | 0.43 | 0.39 | 0.36 | 0.35 | 0.39 | 0.37 | 0.36 |
| | 0.6 | 95% | 0.28 | 0.27 | 0.19 | 0.07 | 0.27 | 0.25 | 0.21 | 0.31 | 0.28 | 0.24 | 0.21 | 0.28 | 0.25 | 0.24 |
| | | 99% | 0.11 | 0.08 | 0.02 | 0.01 | 0.09 | 0.07 | 0.05 | 0.11 | 0.09 | 0.08 | 0.06 | 0.09 | 0.09 | 0.09 |
| | | 90% | 0.65 | 0.62 | 0.53 | 0.31 | 0.62 | 0.58 | 0.53 | 0.65 | 0.62 | 0.59 | 0.56 | 0.62 | 0.59 | 0.57 |
| | 0.8 | 95% | 0.49 | 0.47 | 0.34 | 0.12 | 0.47 | 0.44 | 0.37 | 0.51 | 0.48 | 0.43 | 0.40 | 0.48 | 0.44 | 0.43 |
| | | 99% | 0.22 | 0.17 | 0.05 | 0.02 | 0.18 | 0.15 | 0.11 | 0.24 | 0.21 | 0.19 | 0.15 | 0.20 | 0.21 | 0.20 |

T=500

| | | | $\hat{M}_T^{(1)}$ | $\hat{W}_T^{(1)}$ | | | | | | $\hat{M}_T^{(2)}$ | $\hat{W}_T^{(2)}$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Coefficients | $\delta$ | | Unweighted | $q_1(t,\frac{1}{11})$ | $q_1(t,\frac{3}{11})$ | $q_1(t,\frac{5}{11})$ | $q_2(t,\frac{1}{11})$ | $q_2(t,\frac{3}{11})$ | $q_2(t,\frac{5}{11})$ | Unweighted | $q_1(t,\frac{1}{11})$ | $q_1(t,\frac{3}{11})$ | $q_1(t,\frac{5}{11})$ | $q_2(t,\frac{1}{11})$ | $q_2(t,\frac{3}{11})$ | $q_2(t,\frac{5}{11})$ |
| | | 90% | 0.10 | 0.09 | 0.09 | 0.11 | 0.10 | 0.10 | 0.10 | 0.10 | 0.09 | 0.09 | 0.09 | 0.09 | 0.09 | 0.09 |
| | 0 | 95% | 0.04 | 0.04 | 0.04 | 0.05 | 0.04 | 0.04 | 0.04 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.04 |
| | | 99% | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| | | 90% | 0.40 | 0.40 | 0.37 | 0.21 | 0.40 | 0.39 | 0.36 | 0.39 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 |
| | 0.2 | 95% | 0.26 | 0.28 | 0.25 | 0.07 | 0.27 | 0.26 | 0.23 | 0.27 | 0.27 | 0.26 | 0.25 | 0.27 | 0.26 | 0.26 |
| | | 99% | 0.10 | 0.11 | 0.09 | 0.01 | 0.10 | 0.09 | 0.07 | 0.11 | 0.09 | 0.10 | 0.09 | 0.09 | 0.09 | 0.10 |
| | | 90% | 0.95 | 0.95 | 0.94 | 0.78 | 0.95 | 0.94 | 0.93 | 0.93 | 0.93 | 0.93 | 0.92 | 0.93 | 0.93 | 0.92 |
| $[\theta_2 = 0.005,\ \theta_3 = 0.1]$ | 0.4 | 95% | 0.90 | 0.90 | 0.89 | 0.44 | 0.90 | 0.89 | 0.85 | 0.88 | 0.87 | 0.86 | 0.85 | 0.87 | 0.87 | 0.86 |
| | | 99% | 0.72 | 0.74 | 0.69 | 0.01 | 0.72 | 0.71 | 0.62 | 0.69 | 0.65 | 0.66 | 0.62 | 0.64 | 0.64 | 0.65 |
| | | 90% | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | 0.6 | 95% | 1.00 | 1.00 | 1.00 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | | 99% | 1.00 | 1.00 | 1.00 | 0.27 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.99 | 1.00 | 1.00 | 1.00 |
| | | 90% | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | 0.8 | 95% | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | | 99% | 1.00 | 1.00 | 1.00 | 0.97 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | | 90% | 0.12 | 0.13 | 0.12 | 0.12 | 0.13 | 0.13 | 0.12 | 0.12 | 0.12 | 0.12 | 0.12 | 0.12 | 0.12 | 0.12 |
| | 0 | 95% | 0.06 | 0.07 | 0.06 | 0.05 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 |
| | | 99% | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| | | 90% | 0.37 | 0.37 | 0.35 | 0.21 | 0.37 | 0.37 | 0.35 | 0.36 | 0.35 | 0.35 | 0.35 | 0.35 | 0.35 | 0.35 |
| | 0.2 | 95% | 0.24 | 0.26 | 0.25 | 0.07 | 0.25 | 0.24 | 0.21 | 0.25 | 0.25 | 0.25 | 0.24 | 0.25 | 0.25 | 0.25 |
| | | 99% | 0.09 | 0.10 | 0.09 | 0.01 | 0.09 | 0.09 | 0.07 | 0.11 | 0.09 | 0.10 | 0.09 | 0.09 | 0.09 | 0.10 |
| | | 90% | 0.88 | 0.88 | 0.87 | 0.66 | 0.88 | 0.88 | 0.85 | 0.86 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 |
| $[\theta_2 = 0.01,\ \theta_3 = 0.9]$ | 0.4 | 95% | 0.79 | 0.80 | 0.79 | 0.33 | 0.79 | 0.78 | 0.75 | 0.77 | 0.77 | 0.76 | 0.75 | 0.77 | 0.77 | 0.76 |
| | | 99% | 0.58 | 0.60 | 0.55 | 0.01 | 0.58 | 0.58 | 0.48 | 0.56 | 0.53 | 0.54 | 0.51 | 0.53 | 0.52 | 0.53 |
| | | 90% | 1.00 | 1.00 | 1.00 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | 0.6 | 95% | 1.00 | 1.00 | 1.00 | 0.91 | 1.00 | 1.00 | 1.00 | 1.00 | 0.99 | 0.99 | 0.99 | 1.00 | 0.99 | 0.99 |
| | | 99% | 0.98 | 0.99 | 0.98 | 0.13 | 0.98 | 0.98 | 0.96 | 0.97 | 0.96 | 0.96 | 0.95 | 0.96 | 0.96 | 0.96 |
| | | 90% | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | 0.8 | 95% | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | | 99% | 1.00 | 1.00 | 1.00 | 0.80 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |

Table 4.4.12: Estimated of the time of change as a percentage of the observation ($\mathrm{t}^* = 0.7$)

| T=100 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Coefficients | $\delta$ | | Unweighted | $q_1(t, \frac{1}{11})$ | $q_1(t, \frac{3}{11})$ | $q_1(t, \frac{5}{11})$ | $q_2(t, \frac{1}{11})$ | $q_2(t, \frac{3}{11})$ | $q_2(t, \frac{5}{11})$ |
| $[\theta_2 = 0.005,\ \theta_3 = 0.1]$ | 0.2 | Mean | 0.56 | 0.57 | 0.59 | 0.59 | 0.56 | 0.57 | 0.57 |
| | | Med | 0.58 | 0.60 | 0.65 | 0.70 | 0.58 | 0.58 | 0.57 |
| | | SD | 0.19 | 0.21 | 0.26 | 0.33 | 0.19 | 0.20 | 0.21 |
| | 0.4 | Mean | 0.58 | 0.59 | 0.61 | 0.60 | 0.58 | 0.58 | 0.59 |
| | | Med | 0.63 | 0.64 | 0.67 | 0.70 | 0.62 | 0.62 | 0.61 |
| | | SD | 0.16 | 0.18 | 0.22 | 0.29 | 0.16 | 0.17 | 0.18 |
| | 0.6 | Mean | 0.61 | 0.62 | 0.63 | 0.63 | 0.61 | 0.61 | 0.61 |
| | | Med | 0.66 | 0.66 | 0.68 | 0.70 | 0.65 | 0.64 | 0.63 |
| | | SD | 0.13 | 0.13 | 0.16 | 0.24 | 0.12 | 0.12 | 0.13 |
| | 0.8 | Mean | 0.63 | 0.64 | 0.65 | 0.67 | 0.63 | 0.63 | 0.62 |
| | | Med | 0.67 | 0.67 | 0.68 | 0.70 | 0.66 | 0.66 | 0.65 |
| | | SD | 0.10 | 0.10 | 0.12 | 0.17 | 0.10 | 0.10 | 0.10 |
| $[\theta_2 = 0.01,\ \theta_3 = 0.9]$ | 0.2 | Mean | 0.56 | 0.56 | 0.59 | 0.57 | 0.56 | 0.56 | 0.56 |
| | | Med | 0.57 | 0.59 | 0.63 | 0.67 | 0.57 | 0.57 | 0.56 |
| | | SD | 0.19 | 0.20 | 0.26 | 0.33 | 0.19 | 0.19 | 0.21 |
| | 0.4 | Mean | 0.57 | 0.58 | 0.59 | 0.60 | 0.57 | 0.57 | 0.57 |
| | | Med | 0.59 | 0.61 | 0.64 | 0.68 | 0.59 | 0.59 | 0.58 |
| | | SD | 0.18 | 0.19 | 0.24 | 0.31 | 0.18 | 0.18 | 0.19 |
| | 0.6 | Mean | 0.61 | 0.62 | 0.64 | 0.64 | 0.61 | 0.61 | 0.60 |
| | | Med | 0.65 | 0.66 | 0.68 | 0.70 | 0.64 | 0.63 | 0.62 |
| | | SD | 0.15 | 0.16 | 0.20 | 0.27 | 0.15 | 0.15 | 0.16 |
| | 0.8 | Mean | 0.63 | 0.64 | 0.66 | 0.66 | 0.62 | 0.62 | 0.62 |
| | | Med | 0.67 | 0.67 | 0.70 | 0.71 | 0.66 | 0.65 | 0.64 |
| | | SD | 0.13 | 0.14 | 0.17 | 0.24 | 0.13 | 0.13 | 0.14 |

| T=500 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Coefficients | $\delta$ | | Unweighted | $q_1(t, \frac{1}{11})$ | $q_1(t, \frac{3}{11})$ | $q_1(t, \frac{5}{11})$ | $q_2(t, \frac{1}{11})$ | $q_2(t, \frac{3}{11})$ | $q_2(t, \frac{5}{11})$ |
| $[\theta_2 = 0.005,\ \theta_3 = 0.1]$ | 0.2 | Mean | 0.58 | 0.59 | 0.61 | 0.64 | 0.58 | 0.58 | 0.59 |
| | | Med | 0.62 | 0.63 | 0.66 | 0.70 | 0.62 | 0.61 | 0.60 |
| | | SD | 0.15 | 0.16 | 0.20 | 0.29 | 0.15 | 0.15 | 0.16 |
| | 0.4 | Mean | 0.65 | 0.65 | 0.66 | 0.69 | 0.64 | 0.64 | 0.63 |
| | | Med | 0.68 | 0.68 | 0.69 | 0.70 | 0.67 | 0.67 | 0.66 |
| | | SD | 0.08 | 0.08 | 0.09 | 0.15 | 0.08 | 0.08 | 0.09 |
| | 0.6 | Mean | 0.67 | 0.67 | 0.68 | 0.69 | 0.67 | 0.66 | 0.65 |
| | | Med | 0.69 | 0.69 | 0.69 | 0.70 | 0.69 | 0.68 | 0.68 |
| | | SD | 0.05 | 0.05 | 0.04 | 0.06 | 0.05 | 0.05 | 0.06 |
| | 0.8 | Mean | 0.68 | 0.68 | 0.69 | 0.70 | 0.68 | 0.67 | 0.67 |
| | | Med | 0.69 | 0.69 | 0.70 | 0.70 | 0.69 | 0.69 | 0.68 |
| | | SD | 0.03 | 0.03 | 0.02 | 0.03 | 0.03 | 0.04 | 0.04 |
| $[\theta_2 = 0.01,\ \theta_3 = 0.9]$ | 0.2 | Mean | 0.57 | 0.57 | 0.59 | 0.63 | 0.57 | 0.56 | 0.57 |
| | | Med | 0.61 | 0.61 | 0.65 | 0.69 | 0.60 | 0.59 | 0.58 |
| | | SD | 0.16 | 0.17 | 0.21 | 0.30 | 0.16 | 0.16 | 0.17 |
| | 0.4 | Mean | 0.63 | 0.64 | 0.65 | 0.68 | 0.63 | 0.63 | 0.62 |
| | | Med | 0.67 | 0.68 | 0.68 | 0.70 | 0.67 | 0.66 | 0.65 |
| | | SD | 0.10 | 0.10 | 0.12 | 0.18 | 0.10 | 0.10 | 0.10 |
| | 0.6 | Mean | 0.66 | 0.67 | 0.68 | 0.69 | 0.66 | 0.65 | 0.65 |
| | | Med | 0.68 | 0.69 | 0.69 | 0.70 | 0.68 | 0.68 | 0.67 |
| | | SD | 0.06 | 0.06 | 0.06 | 0.08 | 0.06 | 0.06 | 0.07 |
| | 0.8 | Mean | 0.68 | 0.68 | 0.69 | 0.69 | 0.67 | 0.67 | 0.66 |
| | | Med | 0.69 | 0.69 | 0.70 | 0.70 | 0.69 | 0.69 | 0.68 |
| | | SD | 0.04 | 0.03 | 0.03 | 0.04 | 0.04 | 0.04 | 0.05 |

#### 4.4.3.5 Experiment V ($\mathfrak{t}^* = 0.9$)

Lastly, let a change occurred at the end of sample, $\mathfrak{t}^* = 0.9$. Rejection rates and estimated change-point locations are tabulated in Table 4.4.13 and 4.4.14, respectively.

Results show similar patterns with those in experiment $I$. One noticeable point is that compared with detecting an early change ($\mathfrak{t} = 0.1$), weighted CUSUM tests show relatively slower convergence to unit power. This might be due to the fact that there are not enough observations for CUSUM statistics after the change-point $\mathfrak{t}^*$. Besides, the fraction change-point lie outside of the trimmed interval $[\mathfrak{t}_1, \mathfrak{t}_2]$. Thus, weighted CUSUM tests outperform to unweighted CUSUM tests. Among all tests, weighted CUSUM tests with the weight function $q_1(\mathfrak{t}, \frac{5}{11})$ show the fastest convergence rate and provide best detections. Table 4.4.14 displays same pattern with Table 4.4.6 in experiment $I$.

To be specific in Table 4.4.13, in the small samples, the weighted CUSUM tests $\hat{W}_T^{(1)}$ and $\hat{W}_T^{(2)}$ with $q_1(\mathfrak{t}, \frac{3}{11})$ get power 0.44 and 0.40 with weak dependent observations at 90% significance level, respectively, and these results become to 0.23 and 0.21 subject to strong dependent observations. In the large samples with weak dependent observations, the power of weighted CUSUM tests $\hat{W}_T^{(1)}$ and $\hat{W}_T^{(2)}$ with $q_1(\mathfrak{t}, \frac{3}{11})$ reach to 0.98 and 0.87 at 90% significance level, respectively, and these results become to 0.85 and 0.74 subject to strong dependent observations, respectively.

Table 4.4.13: Empirical rejection rates for CUSUM-typed tests ($\mathfrak{t}^* = 0.9$)

T=100

| | | | $\hat{M}_T^{(1)}$ | $\hat{W}_T^{(1)}$ | | | | | | $\hat{M}_T^{(2)}$ | $\hat{W}_T^{(2)}$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Coefficients | $\delta$ | | Unweighted | $q_1(t, \frac{1}{11})$ | $q_1(t, \frac{3}{11})$ | $q_1(t, \frac{5}{11})$ | $q_2(t, \frac{1}{11})$ | $q_2(t, \frac{3}{11})$ | $q_2(t, \frac{5}{11})$ | Unweighted | $q_1(t, \frac{1}{11})$ | $q_1(t, \frac{3}{11})$ | $q_1(t, \frac{5}{11})$ | $q_2(t, \frac{1}{11})$ | $q_2(t, \frac{3}{11})$ | $q_2(t, \frac{5}{11})$ |
| | | 90% | 0.11 | 0.10 | 0.09 | 0.10 | 0.10 | 0.09 | 0.08 | 0.12 | 0.10 | 0.09 | 0.09 | 0.10 | 0.09 | 0.09 |
| | 0 | 95% | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.04 | 0.06 | 0.05 | 0.04 | 0.04 | 0.05 | 0.04 | 0.04 |
| | | 99% | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| | | 90% | 0.19 | 0.19 | 0.20 | 0.18 | 0.17 | 0.16 | 0.14 | 0.20 | 0.18 | 0.18 | 0.18 | 0.18 | 0.17 | 0.16 |
| | 0.2 | 95% | 0.11 | 0.10 | 0.11 | 0.08 | 0.10 | 0.09 | 0.07 | 0.11 | 0.10 | 0.09 | 0.09 | 0.09 | 0.08 | 0.08 |
| | | 99% | 0.02 | 0.03 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.03 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 |
| | | 90% | 0.21 | 0.22 | 0.24 | 0.21 | 0.19 | 0.18 | 0.14 | 0.23 | 0.21 | 0.20 | 0.21 | 0.20 | 0.19 | 0.18 |
| $[\theta_2 = 0.005,\ \theta_3 = 0.1]$ | 0.4 | 95% | 0.11 | 0.11 | 0.13 | 0.09 | 0.10 | 0.10 | 0.08 | 0.14 | 0.13 | 0.11 | 0.11 | 0.12 | 0.10 | 0.10 |
| | | 99% | 0.03 | 0.03 | 0.02 | 0.02 | 0.03 | 0.02 | 0.02 | 0.04 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 |
| | | 90% | 0.28 | 0.29 | 0.32 | 0.31 | 0.26 | 0.24 | 0.21 | 0.30 | 0.28 | 0.28 | 0.28 | 0.27 | 0.26 | 0.24 |
| | 0.6 | 95% | 0.17 | 0.19 | 0.20 | 0.13 | 0.16 | 0.15 | 0.12 | 0.20 | 0.18 | 0.17 | 0.17 | 0.18 | 0.15 | 0.15 |
| | | 99% | 0.05 | 0.05 | 0.04 | 0.02 | 0.04 | 0.04 | 0.03 | 0.06 | 0.06 | 0.06 | 0.05 | 0.05 | 0.05 | 0.05 |
| | | 90% | 0.38 | 0.41 | 0.44 | 0.40 | 0.36 | 0.34 | 0.31 | 0.43 | 0.40 | 0.40 | 0.41 | 0.39 | 0.37 | 0.37 |
| | 0.8 | 95% | 0.25 | 0.27 | 0.29 | 0.19 | 0.24 | 0.22 | 0.18 | 0.29 | 0.28 | 0.26 | 0.26 | 0.26 | 0.24 | 0.23 |
| | | 99% | 0.08 | 0.08 | 0.06 | 0.04 | 0.06 | 0.06 | 0.04 | 0.10 | 0.08 | 0.08 | 0.08 | 0.07 | 0.08 | 0.07 |
| | | 90% | 0.14 | 0.12 | 0.11 | 0.11 | 0.12 | 0.11 | 0.10 | 0.14 | 0.12 | 0.11 | 0.11 | 0.12 | 0.11 | 0.10 |
| | 0 | 95% | 0.06 | 0.06 | 0.05 | 0.05 | 0.06 | 0.06 | 0.05 | 0.07 | 0.07 | 0.06 | 0.05 | 0.06 | 0.05 | 0.05 |
| | | 99% | 0.02 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.02 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| | | 90% | 0.20 | 0.20 | 0.20 | 0.18 | 0.19 | 0.17 | 0.15 | 0.21 | 0.20 | 0.19 | 0.18 | 0.19 | 0.18 | 0.17 |
| | 0.2 | 95% | 0.11 | 0.11 | 0.11 | 0.08 | 0.10 | 0.10 | 0.08 | 0.12 | 0.11 | 0.09 | 0.09 | 0.10 | 0.09 | 0.09 |
| | | 99% | 0.03 | 0.02 | 0.02 | 0.01 | 0.02 | 0.02 | 0.02 | 0.03 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 |
| | | 90% | 0.19 | 0.19 | 0.20 | 0.17 | 0.17 | 0.15 | 0.13 | 0.22 | 0.19 | 0.19 | 0.18 | 0.19 | 0.18 | 0.17 |
| $[\theta_2 = 0.01,\ \theta_3 = 0.9]$ | 0.4 | 95% | 0.09 | 0.10 | 0.10 | 0.07 | 0.09 | 0.08 | 0.07 | 0.12 | 0.10 | 0.09 | 0.09 | 0.10 | 0.09 | 0.08 |
| | | 99% | 0.03 | 0.02 | 0.02 | 0.01 | 0.02 | 0.02 | 0.02 | 0.03 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 |
| | | 90% | 0.20 | 0.20 | 0.21 | 0.19 | 0.18 | 0.17 | 0.15 | 0.22 | 0.19 | 0.19 | 0.19 | 0.19 | 0.17 | 0.16 |
| | 0.6 | 95% | 0.11 | 0.11 | 0.12 | 0.08 | 0.10 | 0.09 | 0.08 | 0.12 | 0.11 | 0.10 | 0.10 | 0.11 | 0.09 | 0.09 |
| | | 99% | 0.03 | 0.03 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.04 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 |
| | | 90% | 0.22 | 0.22 | 0.23 | 0.21 | 0.20 | 0.18 | 0.16 | 0.23 | 0.21 | 0.21 | 0.21 | 0.21 | 0.19 | 0.18 |
| | 0.8 | 95% | 0.13 | 0.13 | 0.14 | 0.09 | 0.12 | 0.11 | 0.09 | 0.14 | 0.13 | 0.11 | 0.11 | 0.13 | 0.11 | 0.10 |
| | | 99% | 0.04 | 0.04 | 0.03 | 0.02 | 0.03 | 0.03 | 0.02 | 0.04 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 |

T=500

| | | | $\hat{M}_T^{(1)}$ | $\hat{W}_T^{(1)}$ | | | | | | $\hat{M}_T^{(2)}$ | $\hat{W}_T^{(2)}$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Coefficients | $\delta$ | | Unweighted | $q_1(t, \frac{1}{11})$ | $q_1(t, \frac{3}{11})$ | $q_1(t, \frac{5}{11})$ | $q_2(t, \frac{1}{11})$ | $q_2(t, \frac{3}{11})$ | $q_2(t, \frac{5}{11})$ | Unweighted | $q_1(t, \frac{1}{11})$ | $q_1(t, \frac{3}{11})$ | $q_1(t, \frac{5}{11})$ | $q_2(t, \frac{1}{11})$ | $q_2(t, \frac{3}{11})$ | $q_2(t, \frac{5}{11})$ |
| | | 90% | 0.10 | 0.09 | 0.09 | 0.11 | 0.10 | 0.10 | 0.10 | 0.10 | 0.09 | 0.09 | 0.09 | 0.09 | 0.09 | 0.09 |
| | 0 | 95% | 0.04 | 0.04 | 0.04 | 0.05 | 0.04 | 0.04 | 0.04 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.04 |
| | | 99% | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| | | 90% | 0.13 | 0.14 | 0.15 | 0.13 | 0.13 | 0.14 | 0.13 | 0.15 | 0.15 | 0.15 | 0.16 | 0.15 | 0.15 | 0.15 |
| | 0.2 | 95% | 0.07 | 0.08 | 0.09 | 0.05 | 0.07 | 0.06 | 0.06 | 0.09 | 0.09 | 0.09 | 0.09 | 0.09 | 0.09 | 0.08 |
| | | 99% | 0.02 | 0.02 | 0.02 | 0.01 | 0.02 | 0.02 | 0.01 | 0.02 | 0.01 | 0.02 | 0.02 | 0.01 | 0.01 | 0.02 |
| | | 90% | 0.28 | 0.30 | 0.35 | 0.29 | 0.28 | 0.28 | 0.25 | 0.32 | 0.31 | 0.33 | 0.36 | 0.31 | 0.31 | 0.31 |
| $[\theta_2 = 0.005,\ \theta_3 = 0.1]$ | 0.4 | 95% | 0.17 | 0.21 | 0.25 | 0.12 | 0.18 | 0.17 | 0.15 | 0.22 | 0.22 | 0.24 | 0.25 | 0.22 | 0.22 | 0.21 |
| | | 99% | 0.05 | 0.06 | 0.09 | 0.01 | 0.05 | 0.05 | 0.03 | 0.08 | 0.07 | 0.09 | 0.09 | 0.06 | 0.06 | 0.07 |
| | | 90% | 0.56 | 0.62 | 0.73 | 0.64 | 0.56 | 0.56 | 0.52 | 0.59 | 0.59 | 0.64 | 0.69 | 0.57 | 0.58 | 0.57 |
| | 0.6 | 95% | 0.39 | 0.47 | 0.61 | 0.32 | 0.40 | 0.38 | 0.35 | 0.45 | 0.46 | 0.50 | 0.54 | 0.45 | 0.45 | 0.44 |
| | | 99% | 0.16 | 0.20 | 0.31 | 0.03 | 0.16 | 0.16 | 0.12 | 0.22 | 0.21 | 0.25 | 0.26 | 0.19 | 0.19 | 0.20 |
| | | 90% | 0.86 | 0.91 | 0.98 | 0.96 | 0.86 | 0.87 | 0.84 | 0.82 | 0.83 | 0.87 | 0.91 | 0.81 | 0.81 | 0.81 |
| | 0.8 | 95% | 0.71 | 0.82 | 0.94 | 0.73 | 0.72 | 0.71 | 0.67 | 0.70 | 0.73 | 0.78 | 0.83 | 0.69 | 0.70 | 0.69 |
| | | 99% | 0.41 | 0.52 | 0.70 | 0.10 | 0.41 | 0.40 | 0.33 | 0.45 | 0.44 | 0.51 | 0.55 | 0.41 | 0.41 | 0.43 |
| | | 90% | 0.12 | 0.12 | 0.12 | 0.11 | 0.13 | 0.13 | 0.11 | 0.13 | 0.12 | 0.12 | 0.12 | 0.12 | 0.12 | 0.12 |
| | 0 | 95% | 0.06 | 0.06 | 0.06 | 0.04 | 0.06 | 0.06 | 0.06 | 0.07 | 0.07 | 0.06 | 0.06 | 0.07 | 0.07 | 0.06 |
| | | 99% | 0.02 | 0.02 | 0.02 | 0.01 | 0.02 | 0.02 | 0.02 | 0.02 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| | | 90% | 0.16 | 0.17 | 0.16 | 0.13 | 0.16 | 0.16 | 0.15 | 0.16 | 0.15 | 0.16 | 0.17 | 0.15 | 0.16 | 0.15 |
| | 0.2 | 95% | 0.07 | 0.09 | 0.09 | 0.05 | 0.08 | 0.07 | 0.07 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 |
| | | 99% | 0.02 | 0.02 | 0.02 | 0.01 | 0.02 | 0.02 | 0.01 | 0.03 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 |
| | | 90% | 0.26 | 0.28 | 0.32 | 0.25 | 0.26 | 0.26 | 0.24 | 0.29 | 0.29 | 0.30 | 0.33 | 0.28 | 0.28 | 0.28 |
| $[\theta_2 = 0.01,\ \theta_3 = 0.9]$ | 0.4 | 95% | 0.15 | 0.18 | 0.22 | 0.11 | 0.16 | 0.14 | 0.13 | 0.19 | 0.20 | 0.21 | 0.23 | 0.19 | 0.19 | 0.19 |
| | | 99% | 0.04 | 0.06 | 0.07 | 0.01 | 0.04 | 0.04 | 0.04 | 0.07 | 0.06 | 0.07 | 0.07 | 0.05 | 0.05 | 0.06 |
| | | 90% | 0.44 | 0.48 | 0.57 | 0.47 | 0.44 | 0.44 | 0.40 | 0.47 | 0.47 | 0.51 | 0.56 | 0.46 | 0.46 | 0.47 |
| | 0.6 | 95% | 0.27 | 0.34 | 0.44 | 0.23 | 0.28 | 0.27 | 0.25 | 0.35 | 0.36 | 0.39 | 0.42 | 0.35 | 0.35 | 0.34 |
| | | 99% | 0.11 | 0.14 | 0.20 | 0.02 | 0.11 | 0.11 | 0.08 | 0.16 | 0.15 | 0.17 | 0.18 | 0.14 | 0.14 | 0.14 |
| | | 90% | 0.67 | 0.75 | 0.85 | 0.78 | 0.67 | 0.68 | 0.64 | 0.69 | 0.70 | 0.74 | 0.79 | 0.68 | 0.68 | 0.68 |
| | 0.8 | 95% | 0.52 | 0.60 | 0.76 | 0.50 | 0.53 | 0.52 | 0.50 | 0.57 | 0.59 | 0.64 | 0.68 | 0.57 | 0.57 | 0.56 |
| | | 99% | 0.27 | 0.34 | 0.47 | 0.04 | 0.27 | 0.27 | 0.21 | 0.34 | 0.33 | 0.37 | 0.41 | 0.31 | 0.31 | 0.32 |

Table 4.4.14: Estimated of the time of change as a percentage of the observation ($t^* = 0.9$)

| T=100 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Coefficients | $\delta$ | | Unweighted | $q_1(t, \frac{1}{11})$ | $q_1(t, \frac{3}{11})$ | $q_1(t, \frac{5}{11})$ | $q_2(t, \frac{1}{11})$ | $q_2(t, \frac{3}{11})$ | $q_2(t, \frac{5}{11})$ |
| $[\theta_2 = 0.005, \theta_3 = 0.1]$ | 0.2 | Mean | 0.63 | 0.65 | 0.70 | 0.72 | 0.63 | 0.64 | 0.65 |
| | | Med | 0.65 | 0.70 | 0.81 | 0.88 | 0.65 | 0.65 | 0.65 |
| | | SD | 0.21 | 0.23 | 0.25 | 0.30 | 0.21 | 0.22 | 0.23 |
| | 0.4 | Mean | 0.64 | 0.66 | 0.71 | 0.73 | 0.64 | 0.65 | 0.65 |
| | | Med | 0.67 | 0.71 | 0.83 | 0.90 | 0.67 | 0.66 | 0.66 |
| | | SD | 0.22 | 0.23 | 0.26 | 0.30 | 0.22 | 0.22 | 0.23 |
| | 0.6 | Mean | 0.66 | 0.69 | 0.74 | 0.76 | 0.67 | 0.67 | 0.68 |
| | | Med | 0.71 | 0.76 | 0.86 | 0.90 | 0.71 | 0.71 | 0.71 |
| | | SD | 0.21 | 0.22 | 0.23 | 0.28 | 0.21 | 0.21 | 0.22 |
| | 0.8 | Mean | 0.69 | 0.72 | 0.78 | 0.80 | 0.70 | 0.70 | 0.71 |
| | | Med | 0.74 | 0.80 | 0.88 | 0.90 | 0.75 | 0.75 | 0.76 |
| | | SD | 0.20 | 0.20 | 0.21 | 0.24 | 0.20 | 0.20 | 0.21 |
| $[\theta_2 = 0.01, \theta_3 = 0.9]$ | 0.2 | Mean | 0.62 | 0.65 | 0.69 | 0.71 | 0.62 | 0.63 | 0.64 |
| | | Med | 0.64 | 0.68 | 0.80 | 0.88 | 0.63 | 0.63 | 0.63 |
| | | SD | 0.21 | 0.22 | 0.25 | 0.31 | 0.21 | 0.22 | 0.23 |
| | 0.4 | Mean | 0.62 | 0.64 | 0.69 | 0.71 | 0.62 | 0.63 | 0.63 |
| | | Med | 0.64 | 0.68 | 0.79 | 0.88 | 0.64 | 0.63 | 0.63 |
| | | SD | 0.21 | 0.23 | 0.26 | 0.31 | 0.22 | 0.22 | 0.23 |
| | 0.6 | Mean | 0.62 | 0.64 | 0.69 | 0.73 | 0.62 | 0.63 | 0.64 |
| | | Med | 0.64 | 0.68 | 0.80 | 0.88 | 0.64 | 0.63 | 0.63 |
| | | SD | 0.21 | 0.23 | 0.26 | 0.30 | 0.22 | 0.22 | 0.23 |
| | 0.8 | Mean | 0.62 | 0.65 | 0.69 | 0.71 | 0.63 | 0.63 | 0.64 |
| | | Med | 0.65 | 0.69 | 0.80 | 0.88 | 0.64 | 0.63 | 0.63 |
| | | SD | 0.21 | 0.22 | 0.26 | 0.31 | 0.21 | 0.21 | 0.23 |

| T=500 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Coefficients | $\delta$ | | Unweighted | $q_1(t, \frac{1}{11})$ | $q_1(t, \frac{3}{11})$ | $q_1(t, \frac{5}{11})$ | $q_2(t, \frac{1}{11})$ | $q_2(t, \frac{3}{11})$ | $q_2(t, \frac{5}{11})$ |
| $[\theta_2 = 0.005, \theta_3 = 0.1]$ | 0.2 | Mean | 0.55 | 0.56 | 0.60 | 0.64 | 0.55 | 0.55 | 0.56 |
| | | Med | 0.55 | 0.57 | 0.63 | 0.81 | 0.55 | 0.55 | 0.55 |
| | | SD | 0.19 | 0.21 | 0.26 | 0.35 | 0.19 | 0.20 | 0.22 |
| | 0.4 | Mean | 0.63 | 0.65 | 0.71 | 0.76 | 0.63 | 0.63 | 0.64 |
| | | Med | 0.65 | 0.69 | 0.81 | 0.89 | 0.64 | 0.63 | 0.63 |
| | | SD | 0.20 | 0.21 | 0.23 | 0.28 | 0.20 | 0.20 | 0.21 |
| | 0.6 | Mean | 0.71 | 0.75 | 0.81 | 0.86 | 0.71 | 0.72 | 0.72 |
| | | Med | 0.75 | 0.81 | 0.88 | 0.90 | 0.75 | 0.75 | 0.77 |
| | | SD | 0.17 | 0.17 | 0.15 | 0.16 | 0.17 | 0.18 | 0.18 |
| | 0.8 | Mean | 0.78 | 0.81 | 0.86 | 0.89 | 0.78 | 0.78 | 0.79 |
| | | Med | 0.83 | 0.86 | 0.89 | 0.90 | 0.84 | 0.85 | 0.87 |
| | | SD | 0.14 | 0.12 | 0.09 | 0.07 | 0.14 | 0.14 | 0.15 |
| $[\theta_2 = 0.01, \theta_3 = 0.9]$ | 0.2 | Mean | 0.54 | 0.54 | 0.59 | 0.63 | 0.54 | 0.54 | 0.56 |
| | | Med | 0.53 | 0.54 | 0.61 | 0.77 | 0.53 | 0.53 | 0.54 |
| | | SD | 0.19 | 0.21 | 0.26 | 0.35 | 0.19 | 0.19 | 0.21 |
| | 0.4 | Mean | 0.60 | 0.62 | 0.67 | 0.73 | 0.60 | 0.60 | 0.61 |
| | | Med | 0.62 | 0.65 | 0.75 | 0.88 | 0.61 | 0.61 | 0.61 |
| | | SD | 0.20 | 0.22 | 0.25 | 0.30 | 0.20 | 0.20 | 0.21 |
| | 0.6 | Mean | 0.67 | 0.70 | 0.77 | 0.82 | 0.67 | 0.68 | 0.68 |
| | | Med | 0.71 | 0.75 | 0.86 | 0.90 | 0.70 | 0.70 | 0.70 |
| | | SD | 0.19 | 0.19 | 0.20 | 0.22 | 0.19 | 0.19 | 0.20 |
| | 0.8 | Mean | 0.73 | 0.76 | 0.83 | 0.88 | 0.73 | 0.74 | 0.74 |
| | | Med | 0.79 | 0.83 | 0.89 | 0.90 | 0.79 | 0.80 | 0.82 |
| | | SD | 0.17 | 0.16 | 0.14 | 0.13 | 0.17 | 0.17 | 0.18 |

## 4.5 An Application: Detecting Unexpected Events in the U.S. Equity Market

In this section, we apply the change-point tests to detect the most harmful systemic events in the U.S. equity market during last three individual year. Systemic events, normally referring to events causing instabilities of entire financial system, draw controversial discussion in literature as there is no standard method to measure systemic risk.

A survey paper on systemic risk has been completed by Bisias et al. (2012), one important method to measure the level of systemic risk is using conditional correlation, and this becomes more popular since IMF (2009) reported that the evolution of correlation can be used to monitor market stress. Empirically, Billio et al. (2012) adopted Granger causality network model to several financial markets, and specified that the high concentrated level of conditional correlations always takes higher risks of a market drop (Also see Kritzman et al., 2011).

Thus, for the purpose of knowing market systemic events, it is important to detect the instability of correlation structure of entire market. However, it may be argued that the instability of correlation can also be a result of decreasing of correlations. While, according to knowledges from behavioral finance, bull news have diverse reactions in sectors, but bear news are more likely to impact all sectors immediately. Hence, a sudden structural change usually is an increment in market correlation structure.

To model the correlation structure of the U.S. stock market, it is unrealistic to collect all equity shares, as by doing so it would result in an over ten thousands dimensional system. Therefore, we use ten Standard & Poor (S&P) 500 sector indexes as proxies for entire equity shares, including Energy (EN), Financial (FI), Consumer Discretionary (CD), Consumer Staples (CS), Health Care (HC), Information Technology (IT), Industrial (IN), Material

(MA), Utilities (UT) and Technology Service (TS) sectors. The adjusted closed price data is collected for each index, ranged over 01-January-2014 to 31-December-2016. We then split it into three sub-samples for trading year 2014, 2015 and 2016, respectively, such that each sub-sample contains 261 observations. To ensure stationarity, we compute log return series for each sector index. The data source is Datastream.

We apply the semi-parametric CUSUM and weighted CUSUM with weight function $[t \cdot (1 - t)]^{5/11}$ to detect correlation changes. In both tests, we estimate the conditional covariance through the DCC model, and then devolatize sector indexes. The long run covariance is estimated by adopting Bartlett kernel and the Newey-West optimal bandwidth.

Figure 4.5.1 plots semi-parametric CUSUM and weighted CUSUM statistics in the year of 2014, 2015 and 2016. The maximally selected CUSUM and $\hat{W}_T^{(1)}$ statistics can be found at the peak point. As a 10 dimensional multivariate data set, we use asymptotic critical values at 95% significance level documented in Table 4.4.4, and Table 2 in Aue et al. (2009) for CUSUM test. In left sub-figures, CUSUM statistics reject the null hypothesis in year of 2014, 2015 and 2016, and detected dates mainly allocated around the period between June and August. Right sub-figures display weighted CUSUM statistics. The $\hat{W}_T^{(1)}$ test rejects the null hypothesis in year of 2014 and 2016. While during 2015, the test rejects the null at 90% but cannot reject at 95% significance level. Different with the CUSUM test, there is no particular restriction on the range of detected dates, and weighted CUSUM statistics are nearly equal weighted in the sample.

Figure 4.5.1: The statistics of semi-parametric CUSUM and weighted CUSUM tests

Left three sub-figures plot CUSUM statistics, and right three sub-figures plot weighted CUSUM statistics. The red line is the critical values at 95% significance level when $d = 10$. We refer CUSUM critical values from Aue et al. (2009), and weighed CUSUM critical values from Table 4.4.4

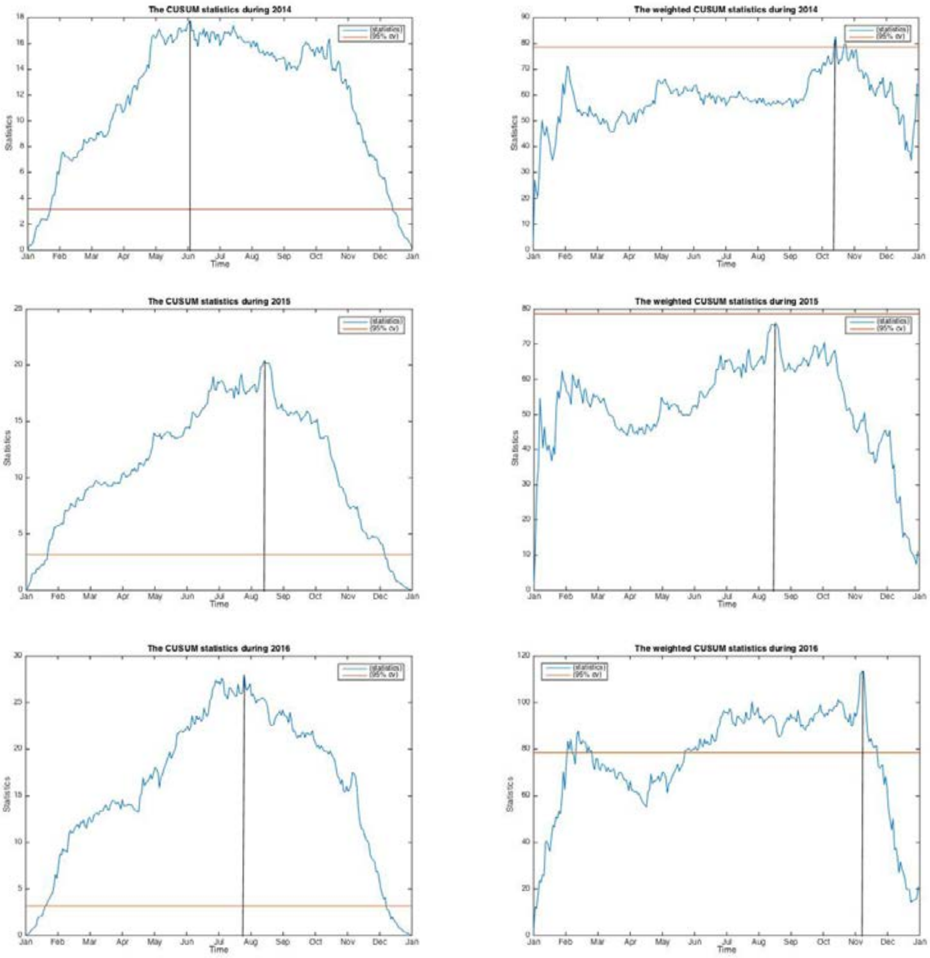



Table 4.5.1 lists market events occurred on detected change days. During 2014, the CUSUM test suggests the biggest correlation change date is 4th June, but there was no specific market event occurred, and the market status during the first half of 2014 actually bulled. The change date detected by weighted CUSUM test is 13th October, which gives more sense, as the U.S. market experienced an one-day plunge event: Dow Jones Industrial Average (DJIA), S&P 500 and tech-laden Nasdaq Composite (TNC) indexes dropped extensively within one trading day. The real situation was worse than just one-day plunge, market records show

downward market during October wiped out all gains from early months in this bull year. This is due to the market fear of Ebola, as well as sour news from European economies.

In year of 2015, CUSUM and weighted CUSUM tests indicate similar correlation change date, around the middle of August. This period can be seen as a start of 2015-16 stock market selloff. The first drop of this selloff was 18th August, and became to a flash crash in 24th August, then followed with several one-day plunges until 2016. This is due to fears about Chinese stock market turbulence, and continued with more fears about U.K. European Union membership referendum in 2016.

It is quite surprising that the CUSUM test did not detect 23rd June - the date of U.K. referendum as the correlation structural break date in 2016, despite the fact that CUSUM statistics on 23rd June displays very high as well in Figure 4.5.1, this is probably because compared with the impact on stock markets, the result of referendum instantly shocked global foreign exchange markets. The actual detected date is 26th July, when the U.S. market experienced another one-day plunge. This drop is because the supply of crude oil exceeds its demand, and oil traders are not insensitive to rebalance market timely. On the other hand, the weighted CUSUM test indicates that the biggest break occurred during the 45th presidential election, this is probably because of a jump in the second moment caused by an extensive short-term reversion. Figure 4.5.2 displays such a pattern.

Based on discussions above, all detected events in the U.S. market experience down side or high volatile market statuses. These events are commonly caused by unanticipated market fears about economy environment, which give rise to temporary market oscillations or systemic effect. Because the CUSUM-typed tests are designed to detect non-smoothing changes, and mainly detect short-term jumps in correlation structure, a further analysis on posterior event is necessary, i.e. to investigate whether the change leads to a structural break or just a transient change.

Table 4.5.1: Change Dates vs Market Events

Semi-parametric CUSUM and semi-parametric weighted CUSUM tets

| | | |
|---|---|---|
| 01-Jan-2014 to 31-Dec-2014 | | |
| | Date | Event |
| CUSUM | 4-Jun-2014 | None |
| Weighted CUSUM | 13-Oct-2014 | One-day Plunge (Fear of Ebola):<br>The Dow Jones Industrial Average declined 0.7%;<br>The Standard & Poor 500 decreased 1.2%;<br>The tech-laden Nasdaq Composite Index declined more than 2%. |
| 01-Jan-2015 to 31-Dec-2015 | | |
| | Date | Event |
| CUSUM | 13-Aug-2015 | The Start of 2015-2016 stock market selloff |
| Weighted CUSUM | 16-Aug-2015 | The Start of 2015-2016 stock market selloff |
| 01-Jan-2016 to 31-Dec-2016 | | |
| | Date | Event |
| CUSUM | 26-Jul-2016 | One-day Plunge (Crude Oil Depreciation):<br>The Dow Jones Industrial Average declined 0.4%;<br>The Standard & Poor 500 decreased 0.3%;<br>The tech-laden Nasdaq Composite Index declined 0.1%. |
| Weighted CUSUM | 8-Nov-2016 | 45th United States presidential election:<br>Global market dropped sharply;<br>the U.S. market experienced a big reversal. |

Figure 4.5.2: The Impact of Presidential Election on Stocks


Source: http://money.cnn.com/2016/11/09/investing/dow-jones-trump-wins-election/

## 4.6 Summary

Concerning the common trimmed issue in testing instability of correlation structure, the present chapter extends semi-parametric CUSUM tests to semi-parametric weighted CUSUM tests. We apply two types of weight functions: $q_1(\alpha, \mathsf{t}) = [\mathsf{t} \cdot (1 - \mathsf{t})]^{\alpha}, \quad 0 \leq \mathsf{t} \leq 1$ and $q_2(\alpha, \mathsf{t}) = [\mathsf{t} \cdot (1 - \mathsf{t}) \cdot loglog \frac{1}{\mathsf{t} \cdot (1-\mathsf{t})}]^{\alpha}$, with $0 \leq \mathsf{t} \leq 1$ and $0 < \alpha < 1/2$. The limits of weighted CUSUM tests are derived under the null and alternative hypothesis. The empirical Monte Carlo simulation study emphasizes that semi-parametric weighted CUSUM tests perform better than semi-parametric CUSUM tests in either ends. To detect a centred correlation change, weighted tests show good power as well, although the power of unweighted tests converge to unit relatively faster. In the end, we provide an application of detecting systemic events in the U.S. market cross over the year of 2014, 2015 and 2016, and the weighted CUSUM test (using $q_1(t, 5/11)$) exhibits more sensible results: the most harmful systemic events in the year of 2014, 2015 and 2016 are the Fear of Ebola, the 2015-2016 stock market selloff and the 45th United States presidential election, respectively.

# Chapter 5

# Conclusion

Multivariate models are widely used in financial applications. The development of technology and the increased computational ability, together with the availability of data at higher frequencies, have made more feasible the modelling and the estimation of systems of larger dimensions. However, the efficiency of estimations is still a challenged problem as structural breaks often occur in real world. Detecting a structural break or change point in the first moment has been widely studied, while the instability in the second moment is still a preliminary period. The milestone in this area is completed by Aue et al. (2009), who derived the asymptotic theory to detect covariance changes in multivariate vectors.

Considering changes in covariance structure can be originated by either variance changes or correlation changes, also correlation structure provides crucial implementations in finance, the present thesis marginally contributed literature about detecting changes in the correlation structure. The thesis started to discuss this issue from a native question: 'Whether a change point in the correlation structure deteriorates the efficiency of multivariate covariance models?', then showed an empirical evidence that how important of analysing correlation structure in finance, lastly proposed change point detection methods in the correlation structure.

Specially, in the first chapter, we compared in-sample and out-of-sample performances of most commonly used multivariate GARCH models, extending existed works completed by Engle (2009) and Laurent et al. (2012). The simulation results indicated the optimal model in in-sample and out-of-sample, respectively; also even though the answer is clear, we found that correlation change deteriorates models' in-sample performances. In order to show the importance of analysing correlation structure to finance, Chapter 2 conducted an empirical work. Followed Billio et al. (2012), we showed that the absorption ratio, a concentration ratio extracted from conditional correlation matrices, can be used as a leading indicator of the financial fragility.

Chapter 3 and 4 completed the most of contribution of this thesis. In the third chapter, we extended CUSUM tests (Aue et al. 2009) into time-varying correlation context, and proposed semi-parametric CUSUM tests to detect change points in correlation structure, and derived their asymptotic distributions. The simulation results implied semi-parametric CUSUM tests can overcome the distortion caused by the nearly unit-root property. The fourth chapter adjusted the limitation of semi-parametric CUSUM test in either sample ends, by selecting suitable weighting functions for semi-parametric CUSUM statistics. The weighted semi-parametric CUSUM tests showed more comprehensive and robust detection abilities in finite samples.

Practically, these tests statistically segment real data according to the evolution of correlation structure, thereby providing more efficient estimations and reliable interpretation in in-sample. The third and fourth chapters provided relevant applications. One is to identify the financial contagion effect in global equity markets during the great recession. Another is to detect unexpected events in the U.S. equity market. More applications are worth to explore in the financial market.

This thesis mainly concentrates on the issue of instability in the historical correlation structure, but correlation change points in the out-of-sample is also important, even provides more implementation views in applications. It sounds unlikely to predict change point in out-of-sample, however, we can sequentially detect a correlation change point by absorbing new observations. The difficulty in this topic is how to design statistics in a fast convergence rate within very finite observations. Sequentially detecting change point in the correlation structure can be a topic for post-doctoral researches.

# Appendix A

## A.1 PROOFS OF THEOREM 3.3.1 - 3.3.2

We start with the weak convergence of the process $\mathfrak{s}(t)$, $0 \le t \le T$.

**Lemma A.1.1.** *If $H_0$ and Assumption 3.3.1-3.3.5 are satisfied, then we have*

$$T^{-1/2}(\mathfrak{s}(Tu) - E\mathfrak{s}(Tu)) \overset{D^{\bar{d}}[0,1]}{\rightarrow} \mathbf{W}_{\mathfrak{D}(u)}$$

*where $\mathbf{W}_{\mathfrak{D}}(u)$, $0 \le u \le 1$ is a Brownian motion in $R^{\bar{d}}$ with covariance matrix $\mathfrak{D}$, i.e. $\mathbf{W}(u)$ is Gaussian with $E\mathbf{W}(u) = 0$ and $E\mathbf{W}_{\mathfrak{D}}(u)\mathbf{W}_{\mathfrak{D}}^{\top}(\nu) = min(u, \nu)\mathfrak{D}$.*

*Proof.* It follows from Assumption 3.3.1-3.3.4 that $y_t^*(i)y_t^*(j)$ is also stationary and $\beta$-mixing with the same rate as of $\mathbf{y}_t$. Also, since Assumption 3.3.3 implies that $\tau_t(i) \ge \tau_0$ we get that

$$E\left|y_t^*(i)y_t^*(j)\right|^{r/2} \le \frac{1}{\tau_0^2}(E\left|y_t^*(i)\right|^r E\left|y_t^*(j)\right|^r)^{1/2} < \infty$$

via the Cauchy-Schwartz inequality and the moment condition in Assumption 3.3.4. Hence, the result of Ibragimov (1962) (cf. also Rio, 2000) implies the lemma.

*Proof of Theorem 3.3.2.* Lemma A.1 implies that

$$T^{-1/2}(\mathfrak{s}(Tu) - \frac{\langle Tu \rangle}{T}\mathfrak{s}(T)) \overset{D^{\bar{d}}[0,1]}{\rightarrow} \mathbf{W}_{\mathfrak{D}}(u) - u\mathbf{W}_{\mathfrak{D}}(1) \qquad \text{(A.1.1)}$$

Checking the covariance structure, one can easily verify that

$$\left\{\mathfrak{D}^{-1/2}(\mathbf{W}_{\mathfrak{D}}(u) - u\mathbf{W}_{\mathfrak{D}}(1)), 0 \le u \le 1\right\} \stackrel{\mathcal{D}}{=} \{(B_1(u), B_2(u), \dots, B_{\bar{d}}(u)), 0 \le u \le 1\} \quad \text{(A.1.2)}$$

where $B_1, B_2, \dots, B_{\bar{d}}$ are independent Brownian Bridges. Hence Theorem 3.3.1 follows from (A.1.1) and (A.1.2) via the continuous mapping theorem.

*Proof of Theorem 3.3.2* It follows from the definition of $\hat{y}_t(i)$ that

$$\hat{y}_t(i)\hat{y}_t(j) - y_t^*(i)y_t^*(j) = a_{t,1}(i,j) + \cdots + a_{t,8}(i,j)$$

where

$$a_{t,1}(i,j) = y_t(i)y_t(j)\Big(\frac{1}{\bar{\tau}_t(i;\hat{\boldsymbol{\theta}}_T)} - \frac{1}{\tau_t(i;\hat{\boldsymbol{\theta}}_T)}\Big)\Big(\frac{1}{\bar{\tau}_t(j;\hat{\boldsymbol{\theta}}_T)} - \frac{1}{\tau_t(j;\hat{\boldsymbol{\theta}}_T)}\Big),$$

$$a_{t,2}(i,j) = y_t(i)y_t(j)\Big(\frac{1}{\bar{\tau}_t(i;\hat{\boldsymbol{\theta}}_T)} - \frac{1}{\tau_t(i;\hat{\boldsymbol{\theta}}_T)}\Big)\Big(\frac{1}{\tau_t(j;\hat{\boldsymbol{\theta}}_T)} - \frac{1}{\tau_t(j;\hat{\boldsymbol{\theta}}_0)}\Big),$$

$$a_{t,3}(i,j) = y_t(i)\Big(\frac{1}{\bar{\tau}_t(i;\hat{\boldsymbol{\theta}}_T)} - \frac{1}{\tau_t(i;\hat{\boldsymbol{\theta}}_T)}\Big)\frac{y_t(j)}{\tau_t(i)},$$

$$a_{t,4}(i,j) = y_t(i)y_t(j)\Big(\frac{1}{\tau_t(i;\hat{\boldsymbol{\theta}}_T)} - \frac{1}{\tau_t(i;\boldsymbol{\theta}_0)}\Big)\Big(\frac{1}{\bar{\tau}_t(j;\hat{\boldsymbol{\theta}}_T)} - \frac{1}{\tau_t(j;\hat{\boldsymbol{\theta}}_T)}\Big),$$

$$a_{t,5}(i,j) = y_t(i)y_t(j)\Big(\frac{1}{\tau_t(i;\hat{\boldsymbol{\theta}}_T)} - \frac{1}{\tau_t(i;\boldsymbol{\theta}_0)}\Big)\Big(\frac{1}{\tau_t(j;\hat{\boldsymbol{\theta}}_T)} - \frac{1}{\tau_t(j;\boldsymbol{\theta}_0)}\Big),$$

$$a_{t,6}(i,j) = \frac{y_t(i)}{\tau_t(i)}y_t(j)\Big(\frac{1}{\bar{\tau}_t(j;\hat{\boldsymbol{\theta}}_T)} - \frac{1}{\tau_t(j;\hat{\boldsymbol{\theta}}_T)}\Big),$$

$$a_{t,7}(i,j) = y_t(i)\Big(\frac{1}{\tau_t(i;\hat{\boldsymbol{\theta}}_T)} - \frac{1}{\tau_t(i;\boldsymbol{\theta}_0)}\Big)\frac{y_t(j)}{\tau_t(j)},$$

and

$$a_{t,8}(i,j) = \frac{y_t(i)}{\tau_t(i)}y_t(j)\Big(\frac{1}{\tau_t(j;\hat{\boldsymbol{\theta}}_T)} - \frac{1}{\tau_t(j;\boldsymbol{\theta}_0)}\Big).$$

Since $\bar{\tau}_t(i;\hat{\boldsymbol{\theta}}_T)\tau_0 > 0$, by Assumptions 3.3.6 and 3.3.7 we have on account of the mean value theorem that

$$T^{-1/2} \max_{1\le t\le T} \sum_{s=1}^{t} |a_{s,1}| = O_P(1)T^{-1/2} \sum_{t=1}^{T} |y_t(i)y_t(j)|\, a^2(t)$$

We can assume without loss of generality that $a_t$ is non-increasing as $t \to \infty$. Using again Assumption 3.3.6 we can find a sequence $a_T$ such that $T^{-1/2}a_T \to 0$ and $T^{1/2}a(a_T) \to 0$ and therefore

$$\begin{aligned} T^{-1/2}\textstyle\sum_{t=1}^{T} |y_t(i)y_t(j)|\, a^2(t) &\le T^{-1/2}\textstyle\sum_{t=1}^{a_T} |y_t(i)y_t(j)|\, a^2(t) + \textstyle\sum_{t=a_T+1}^{T} |y_t(i)y_t(j)|\, a^2(t) \\ & O_P(T^{-1/2}a_T + T^{1/2}a^2(a_T)) \\ & o_P(1) \end{aligned} \tag{A.1.3}$$

where we used the ergodic theorem that

$$\frac{1}{L}\sum_{t=1}^{L} |y_t(i)y_t(j)| \to E\,|y_0(i)y_0(j)| \quad a.s. \quad (L\to\infty)$$

Since by Assumption 3.3.4, $E\,|y_0(i)y_0(j)| \le (Ey_0^2(i)Ey_0^2(j))^{1/2} < \infty$. Putting together Assumption 3.3.6-3.3.8, we conclude via two term Taylor expansion and the mean value theorem that

$$T^{-1/2} \max_{1\le t\le T} \sum_{s=1}^{t} |a_{s,2}| = O_P(T^{-1/2}) \sum_{t=1}^{T} |y_t(i)y_t(j)|\, a_t(t) \left[ \|\mathbf{g}_t(j)\| \left\|\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}\right\| + \bar{g}_t \left\|\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0\right\|^2 \right]$$

Following the proof of (A.1.3), one can show that

$$\begin{aligned} T^{-1/2}\textstyle\sum_{t=1}^{T} |y_t(i)y_t(j)|\, a(t)\, \|\mathbf{g}_t(j)\| \left\|\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0\right\|^2 &= O_P(1)\tfrac{1}{T}\textstyle\sum_{t=1}^{T} |y_t(i)y_t(j)|\, a(t)\, \|\mathbf{g}_t(j)\| \\ &= o_P(1) \end{aligned}$$

Since,

$$E\,|y_t(i)y_t(j)|\,\|\mathbf{g}_t(j)\| \le (E\,|y_t(i)y_t(j)|^2\,E\,\|\mathbf{g}_t(j)\|^2)^{1/2} \le (E\,|y_t^4(i)y_t^4(j)|)^{1/4}(E\,\|\mathbf{g}_t(j)\|^2)^{1/2} < \infty$$

the same arguments give

$$\begin{aligned} T^{-1/2}\textstyle\sum_{t=1}^T |y_t(i)y_t(j)|\,a(t)\bar{g}_t \left\|\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0\right\|^2 &= O_P(1)\tfrac{1}{T^{3/2}}\textstyle\sum_{t=1}^T |y_t(i)y_t(j)|\,a(t)\bar{g}_t \\ &= O_P(1)(\tfrac{1}{T^{1/2}}\max_{1\le t\le T}\bar{g}_t)\tfrac{1}{T}\textstyle\sum_{t=1}^T |y_t(i)y_t(j)|\,a(t) \\ &= o_P(1) \end{aligned}$$

Similarly,

$$T^{-1/2}\max_{1\le t\le T}\sum_{s=1}^t |a_{s,3}| = O_P(1)T^{-1/2}\sum_{t=1}^T |y_t(i)y_t(j)|\,a(t) = o_P(1)$$

By symmetry,

$$T^{-1/2}\max_{1\le t\le T}\sum_{s=1}^t |a_{s,l}| = o_P(1), \quad l = 4,5,6$$

Assumption 3.3.8 implies that

$$\begin{aligned} &T^{-1/2}\max_{1\le t\le T}\left|\textstyle\sum_{s=1}^t a_{s,7} - \sum_{s=1}^t \frac{y_t(i)}{\tau_t^2(i)}\frac{y_t(i)}{\tau_t(j)}\mathbf{g}_s(i)^\top(\boldsymbol{\theta}_0 - \hat{\boldsymbol{\theta}}_T)\right| \\ &= O_P(1)T^{-1/2}\textstyle\sum_{s=1}^T |y_t(i)y_t(j)|\,\bar{g}_t \left\|\boldsymbol{\theta}_0 - \hat{\boldsymbol{\theta}}_T\right\|^\top \\ &= O_P(1)(T^{-1/2}\max_{1\le t\le T}\bar{g}_t)\tfrac{1}{T}\textstyle\sum_{s=1}^T |y_t(i)y_t(j)| \\ &= o_P(1) \end{aligned}$$

Using again the ergodic theorem and Assumption 3.3.4, we conclude

$$\begin{aligned} &T^{-1/2}\max_{1\le t\le T}\left|(\textstyle\sum_{s=1}^t \frac{y_t(i)}{\tau_t^2(i)}\frac{y_t(j)}{\tau_t(j)}\mathbf{g}_s(i) - \frac{t}{T}\sum_{s=1}^t \frac{y_t(i)}{\tau_t^2(i)}\frac{y_t(j)}{\tau_t(j)}\mathbf{g}_s(i))^\top(\boldsymbol{\theta}_0 - \hat{\boldsymbol{\theta}}_T)\right| \\ &= O_P(1)\tfrac{1}{T}\max_{1\le t\le T}\left\|\textstyle\sum_{s=1}^t \frac{y_t(i)}{\tau_t^2(i)}\frac{y_t(j)}{\tau_t(j)}\mathbf{g}_s(i) - \frac{t}{T}\sum_{s=1}^t \frac{y_t(i)}{\tau_t^2(i)}\frac{y_t(j)}{\tau_t(j)}\mathbf{g}_s(i)\right\| \\ &= o_P(1) \end{aligned}$$

Since $E\left|y_0(i)y_0\right|\left\|\mathbf{g}_s(i)\right\| < \infty$, Hence, we obtain that

$$T^{-1/2} \max_{1\le t\le T} \left|\sum_{s=1}^{t} a_{s,7} - \frac{t}{T}\sum_{s=1}^{T} a_{s,7}\right|$$

and by the same arguments

$$T^{-1/2} \max_{1\le t\le T} \left|\sum_{s=1}^{t} a_{s,8} - \frac{t}{T}\sum_{s=1}^{T} a_{s,8}\right|$$

We prove that

$$\begin{aligned} T^{-1/2} \max_{1\le t\le T} |(&\sum_{s=1}^{t} \hat{y}_s(i)\hat{y}_s(j) - \frac{t}{T}\sum_{s=1}^{T} \hat{y}_s(i)\hat{y}_s(j)) \\ -(&\sum_{s=1}^{t} y_s^*(i)y_s^*(j) - \frac{t}{T}\sum_{s=1}^{T} y_s^*(i)y_s^*(j))| = o_P(1) \end{aligned}$$

and therefore the result follows from Theorem 3.3.1.

# A.2 PROOFS OF THEOREM 4.3.1 - 4.3.3

## A.2.1 Proofs of Theorem 4.3.1 and 4.3.2

According to Aue et al. (2009), we can refer to the proof of

$$\frac{1}{\sqrt{T}}\sum_{j=1}^{\lfloor T\cdot \mathsf{t}\rfloor}(vech[\mathbf{y}_j\mathbf{y}_j^T - E[\mathbf{y}_j\mathbf{y}_j^T]) \overset{D}{\to} W_{\Sigma}(t) \quad (T\to\infty) \tag{A.2.1}$$

where $W_{\Sigma}(\mathsf{t}) : \mathsf{t} \in [0,1]$ is a d-dimensional Brownian motion.

Now, we consider to apply weight functions into the sum of $\mathbf{y}_t$. Since weight functions $q_1(\mathsf{t},\alpha)$ and $q_2(\mathsf{t},\alpha)$ with $0 < \alpha < 1/2$ are still converge, by adding one of weighting functions $q(\mathsf{t})$. Recall weight functions in Equations 4.3.3 and 4.3.4.

If $q(\mathsf{t}) < \infty$ for some positive $c$, the weighted version of $\frac{1}{\sqrt{T}}\sum_{j=1}^{\lfloor T\cdot \mathsf{t}\rfloor}\mathbf{y_j}$ should converge to a d-dimensional weighted Brownian motion.

$$\frac{1}{\sqrt{T}\cdot q(\mathsf{t})}\sum_{j=1}^{\lfloor T\cdot \mathsf{t}\rfloor}\mathbf{y_j} \overset{D_d[0,1]}{\to} \frac{W_{\Gamma}(\mathsf{t})}{q(\mathsf{t})} \quad (T\to\infty)$$

Similarly, the weighted version of covariance function would converge to a d-dimensional weighted Brownian motion as well.

$$\frac{1}{\sqrt{T}\cdot q(\mathsf{t})}\sum_{j=1}^{\lfloor T\cdot \mathsf{t}\rfloor}(vech[\mathbf{y}_j\mathbf{y}_j^\top - E[\mathbf{y}_j\mathbf{y}_j^\top]) \overset{D}{\to} \frac{W_{\Sigma}(\mathsf{t})}{q(\mathsf{t})} \quad (T\to\infty) \tag{A.2.2}$$

Then, we can discuss the convergence of weighted CUSUM statistics. Because the long-run covariance absolutely converges (also discussed in Aue et al. 2009), the weighted CUSUM statistics $\hat{W}_T^{(2)}$ converges if following CUSUM statistics converged.

$$\frac{1}{\sqrt{T}\cdot q(\mathsf{t})}\sum_{j=1}^{\lfloor T\cdot \mathsf{t}\rfloor}(vech[\mathbf{y}_j\mathbf{y}_j^T] - E[vech[\mathbf{y}_j\mathbf{y}_j^T]) \overset{D}{\to} \frac{W_{\Sigma}(\mathsf{t})}{q(\mathsf{t})} \quad (T\to\infty)$$

then, it yields,

$$\frac{1}{\sqrt{T}\cdot q(\mathfrak{t})}(\sum_{j=1}^{\lfloor T\cdot\mathfrak{t}\rfloor} vech[\mathbf{y}_j\mathbf{y}_j^T] - \mathfrak{t}\sum_{j}^{T}[vech[\mathbf{y}_j\mathbf{y}_j^T]) \xrightarrow{D} (\frac{W_\Sigma(\mathfrak{t})}{q(\mathfrak{t})} - \frac{t\cdot W(1)}{q(\mathfrak{t})}) \quad (T\rightarrow\infty) \quad \text{(A.2.3)}$$

Furthermore, since Csörgő and Horváth (1981) showed that $(W(\mathfrak{t}) - \mathfrak{t}(W(1)))$ and $B(\mathfrak{t})$ are same distributed, the weighted CUSUM statistics have following limit.

$$\frac{1}{\sqrt{T}\cdot q(\mathfrak{t})}(\sum_{j=1}^{\lfloor T\cdot\mathfrak{t}\rfloor} vech[\mathbf{y}_j\mathbf{y}_j^T] - \mathfrak{t}\sum_{j}^{T} vech[\mathbf{y}_j\mathbf{y}_j^T]) \xrightarrow{D} \frac{B_\Sigma(\mathfrak{t})}{q(\mathfrak{t})} \quad (T\rightarrow\infty) \quad \text{(A.2.4)}$$

Combined with the long run covariance's convergence, we obtain the asymptotic limits shown as Equations A.2.4. Therefore, Theorem 4.3.1 is proved.

*Proof of Theorem 4.3.2* We refer Csörgö and Horváth (1997) to prove Theorem 4.3.2. Denoting $\hat{\mathfrak{s}}(\lfloor T\cdot\mathfrak{t}\rfloor) - \mathfrak{t}\hat{\mathfrak{s}}(T)$ as $\mathcal{S}$, we have $W_T^{(1)} = \max_{0\leq\mathfrak{t}\leq1}\frac{1}{T\cdot q(\mathfrak{t})}\mathcal{S}^\top\mathfrak{D}^{-1}\mathcal{S}$. Concerning the consistency of long run covariance estimator, for estimator $\hat{\mathfrak{D}}$ satisfies the condition that,

$$\sup_{0\leq\mathfrak{t}\leq1}\left\|\hat{\mathfrak{D}} - \mathfrak{D}\right\| = o_P[(loglogT)^{-\frac{1}{2}}]$$

with restrictions on $m$ for $\frac{\sqrt{loglogT}}{m}\rightarrow 0$ and $m\cdot logT\sqrt{\frac{loglogT}{T}}\rightarrow 0$, there also exists

$$\hat{W}_T^{(1)} = \max_{0\leq\mathfrak{t}\leq1}\frac{1}{T\cdot q(\mathfrak{t})}\mathcal{S}^\top\hat{\mathfrak{D}}^{-1}\mathcal{S}.$$

According to the Darling and Erdös law (Csörgö and Horváth, 1997), defining that

$$M(\mathfrak{t}) = \sum\frac{B_{(i)}^2(\mathfrak{t})}{q(\mathfrak{t})}$$

then it gives,

$$P\{a_T \cdot \sup_{\frac{1}{T}<\mathfrak{t}<1-\frac{1}{T}} M(\mathfrak{t}) \leq \mathcal{K} + b_T\} \to e^{-2e^{-\mathcal{K}}} \tag{A.2.5}$$

where $a_T$ and $b_T$ is defined in (4.3.7). Then, as we know that,

$$(2loglogT)^{-\frac{1}{2}} \sup_{\frac{1}{T}\leq\mathfrak{t}\leq 1-\frac{1}{T}} M(\mathfrak{t}) \xrightarrow{P} 1 \tag{A.2.6}$$

If there exists a truncated point $u(T) = exp((logT)^{\Delta})/T$ with some $0 < \Delta < 1$. Then the limit in beginning and end of sample can be extended to,

$$(2loglogT)^{-\frac{1}{2}} \sup_{\frac{1}{T}\leq\mathfrak{t}\leq u(T)} M(\mathfrak{t}) \xrightarrow{P} \delta^{\frac{1}{2}} \tag{A.2.7}$$

and

$$(2loglogT)^{-\frac{1}{2}} \sup_{1-u(T)\leq\mathfrak{t}\leq 1-\frac{1}{T}} M(\mathfrak{t}) \xrightarrow{P} \delta^{\frac{1}{2}} \tag{A.2.8}$$

Because $\mathbf{y}_t$ can be approximated by zero mean i.i.d Gaussian sequence, it is easily to have

$$\sup_{\frac{1}{T}\leq\mathfrak{t}\leq 1-\frac{1}{T}} \left|W_T^{(1)}(\mathfrak{t}) - M(\mathfrak{t})\right| = o_P((loglogT)^{\frac{1}{2}})$$

By applying iterated logarithm, we then have corresponding limits for $W_T^{(1)}$,

$$(2loglogT)^{-\frac{1}{2}} \sup_{\frac{1}{T}\leq\mathfrak{t}\leq 1-\frac{1}{T}} W_T^{(1)}(\mathfrak{t}) \xrightarrow{P} \tau \tag{A.2.9}$$

$$(2loglogT)^{-\frac{1}{2}} \sup_{\frac{1}{T}\leq\mathfrak{t}\leq u(T)} W_T^{(1)}(\mathfrak{t}) \xrightarrow{P} \delta^{\frac{1}{2}}\tau \tag{A.2.10}$$

and

$$(2loglogT)^{-\frac{1}{2}} \sup_{1-u(T)\leq\mathfrak{t}\leq 1-\frac{1}{T}} W_T^{(1)}(\mathfrak{t}) \xrightarrow{P} \delta^{\frac{1}{2}}\tau \tag{A.2.11}$$

Next, we define $\xi(T)$ and $\eta(T)$ as the location of maximum $M(\mathfrak{t})$ and $W_T^{(1)}(\mathfrak{t})$ on $[\frac{1}{T}, 1-\frac{1}{T}]$.

Equation A.2.6 - A.2.11 indicate that,

$$\sup_{0 \leq \mathsf{t} \leq 1} M(\mathsf{t}) = M(\xi(T))$$

and

$$\sup_{0 \leq \mathsf{t} \leq 1} W_T^{(1)}(\mathsf{t}) = W_T^{(1)}(\eta(T))$$

Thus, there exists,

$$P[u(\delta) \leq \xi(T), \eta(T) \leq 1 - u(\delta)] = 1 \tag{A.2.12}$$

thus, for $T \to \infty$ and $\delta \to 0$

$$\frac{1}{\tau} \sup_{u(T) \leq \mathsf{t} \leq 1-u(T)} W_T^{(1)}(\mathsf{t}) - \sup_{u(T) \leq \mathsf{t} \leq 1-u(T)} M(\mathsf{t}) = o_P((loglogT)^{\frac{1}{2}}) \tag{A.2.13}$$

(A.2.12) and (A.2.13) further imply that

$$\frac{1}{\tau} \sup_{\frac{1}{T} \leq \mathsf{t} \leq 1-\frac{1}{T}} W_T^{(1)}(\mathsf{t}) - \sup_{\frac{1}{T} \leq \mathsf{t} \leq 1-\frac{1}{T}} M(\mathsf{t}) = o_P((loglogT)^{\frac{1}{2}}) \tag{A.2.14}$$

Also because that

$$\sup_{0 < \mathsf{t} \leq \frac{1}{T}} W_T^{(1)} = o_P((loglogT)^{\frac{1}{2}})$$

and

$$\sup_{1-\frac{1}{T} \leq \mathsf{t} < 1} W_T^{(1)} = o_P((loglogT)^{\frac{1}{2}})$$

we eventually have,

$$\frac{1}{\tau} \sup_{\mathsf{t}<1} W_T^{(1)}(\mathsf{t}) - \sup_{\frac{1}{T} \leq \mathsf{t} \leq 1-\frac{1}{T}} M(\mathsf{t}) = o_P((loglogT)^{\frac{1}{2}}) \tag{A.2.15}$$

combined with Darling and Erdös theorem, refer to Equation A.2.5, Theorem 4.3.2 is proved.

# Bibliography


[1] Acharya, V. and Pedersen, L.H. and Philippon, T. and Richardson, M.: Measuring systemic risk. *Unpublished Working Paper*. New York University (2011).

[2] Adrian, T. and Brunnermeier, M.: CoVaR *National Bureau of Economic Research* (2010).

[3] Aielli, G. P.: Consistent estimation of large scale dynamic conditional correlations. *Unpublished paper* Department of Statistics, University of Florence, (2006).

[4] Aielli, G. P.: Dynamic conditional correlation: on properties and estimation. *Journal of Business & Economic Statistics* **31**(2013), 282–299.

[5] Akaike, H.: Factor analysis and AIC. *Psychometrika* **52**(1987), 317–332.

[6] Alexander C.: Orthogonal GARCH. In Mastering Risk. Financial Times–Prentice Hall: London. 2001, 21–38.

[7] Alessi, l. and Detken, C.: Real time'early warning indicators for costly asset price boom/bust cycles: a role for global liquidity *ECB working paper* (2009).

[8] Anderson, T.W.: *The Statistical Analysis of Time Series*. Wiley, 1971

[9] Andreou, E. and Ghysels, E.: Detecting multiple breaks in financial market volatility dynamics. *Journal of Applied Econometrics* **17** (2002), 579–600

[10] Andrews, D.W.K.: Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica: Journal of the Econometric Society* (1991), 817–858

[11] Andrews, D.W.K.: Tests for parameter instability and structural change with unknown change point. *Econometrica* **61** (1993), 821–856

[12] Andrews, D.W.K.: Tests for parameter instability and structural change with unknown change point: A corrigendum. *Econometrica* (2003), 395–397.

[13] Andreou, E. and Ghysels, E.: Tests for breaks in the conditional co-movements of asset returns. *Statistica Sinica* (2003), 1045–1073.

[14] Antoch, J. and Hušková, M. and Prášková, Z.: Effect of dependence on statistics for determination of change. *Journal of Statistical Planning and Inference* **60**(1997), 291–310.

[15] Aue, A. and Hörmann, S.and Horváth, L. and Hušková, M.: Dependent functional linear models with applications to monitoring structural change. *Statistica Sinica* **24** (2014), 1043–1073

[16] Aue, A. and Hörmann, S. and Horváth, L. and Reimherr, M.: Break detection in the covariance structure of multivariate time series models. *The Annals of Statistics* **37** (2009), 4046–4087

[17] Aue, A. and Horváth, L.: Structural breaks in time series. *Journal of Time Series Analysis* **34** (2013), 1–16

[18] Aue, A. and Berkes, I. and Horváth, L.: Strong approximation for the sums of squares of augmented GARCH sequences. *Bernoulli* **12**(2006), 583–608.

[19] Bai, J. and Perron, P.: Estimating and testing linear models with multiple structural changes. *Econometrica* (1998), 47–78.

[20] Bai, J. and Perron, P.: Computation and analysis of multiple structural change models. *Journal of applied econometrics* **18**(2003), 1–22.

[21] Banz R. W.: The relationship between return and market value of common stocks. *Journal of financial economics* **9**(1981), 3-18.

[22] Bauwens, L. and Laurent, S. and Rombouts, J. V.: Multivariate GARCH models: a survey. *Journal of applied econometrics* **21**(2006), 79–109.

[23] Berkes, I. and Horváth, L. and Kokoszka, P.: Testing for parameter constancy in GARCH (p, q) models. *Statistics & probability letters* **70**(2004), 263–273.

[24] Berkes, I. and Horváth, L. and Husková, M. and Steinebach, J.: Applications of permutations to the simulations of critical values. *Nonparametric Statistics***16** (2004), 197–216.

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

[26] Berkes, I. and Liu, W. and Wu, W.B.: Komlós-Major-Tusnády approximation under dependence. *Annals of Probability* **42** (2014), 794–819

[27] Berkes I and Philipp W.: Approximation thorems for independent and weakly dependent random vectors. *The Annals of Probability* (1979), 29–54

[28] Bernstein, S.: Sur l'extension du théorème limite du calcul des probabilités aux sommes de quantités dépendantes. *Mathematische Annalen* **97** (1927), 1–59

[29] Bhandari L. C.: Debt/equity ratio and expected common stock returns: Empirical evidence. *The journal of finance* **43**(1988), 507-528.

[30] Billingsley, P.: *Convergence of Probability Measures*. Wiley, New York, 1968

[31] Billio, M. and Pelizzon, L.: Volatility and shocks spillover before and after EMU in European stock markets. *Journal of Multinational Financial Management* bf 13(2003), 323–340.

[32] Billio, M. and Getmansky, M. and Pelizzon, L.: Time-varying risk exposure of hedge funds. *working paper* University of Massachusetts, Boston, MA (2006).

[33] Billio, M. and Getmansky, M. and Lo, A. W. and Pelizzon, L.: Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics* **104**(2012), 535–559

[34] Bisias, D. and Flood, M. and Lo, A. W. and Valavanis, S.: A survey of systemic risk analytics. *US Department of Treasury, Office of Financial Research* (2012).

[35] Blatt, D. and Candelon, B. and Manner, H.: 'Detecting contagion in a multivariate time series system: An application to sovereign bond markets in Europe'. *Journal of Banking & Finance* **59** (2015), 1–13

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

[37] Bollerslev, T.: A conditionally heteroskedastic time series model for speculative prices and rates of return. *The review of economics and statistics* (1987), 542–547.

[38] Bollerslev T. and Engle R. F. and Wooldridge J. M.: A capital asset pricing model with time-varying covariances. *Journal of political Economy* **96**(1988), 116-131.

[39] Bollerslev T.: Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH approach. *Review of Economics and Statistics* **72**(1990), 498–505

[40] Bollerslev T and Engle R.F.: Common persistence in conditional variances. *Econometrica* **61**(1993), 167–186.

[41] Bollerslev, T. and Engle R.F. and Wooldridge J.: A capital asset pricing model with time varying covariances. *Journal of Political Economy* **96**(1998), 116–131.

[42] Boussama, F. and Fuchs, F. and Stelzer, R.: 'Stationarity and geometric ergodicity of BEKK multivariate GARCH models'. *Stochastic Processes and their Applications* **121** (2011), 2331–2360

[43] Bradley, R. C., *Introduction to strong mixing conditions.* Kendrick Press, Heber City, UT, 2007

[44] Breiman, L.: *Probability.* Addison Wesley, 1968

[45] Brockwell P.J. and Davis R.A.: *Time series: theory and methods*, 2nd ed. Springer, New York, 1991.

[46] Burnham, K. P. and Anderson, D. R.: Multimodel inference understanding AIC and BIC in model selection. Sociological methods & research **33**(2004), 261–304.

[47] Cappiello, L.and Engle, R.F. and Sheppard, K.: Asymmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial econometrics* **4**(2006), 537–572.

[48] Carrasco, M. and Chen, X.: Mixing and moment properties of various GARCH and stochastic volatility models. *Econometric Theory* **18**(2002), 17–39.

[49] Celık, S.: 'The more contagion effect on emerging markets: The evidence of DCC-GARCH model'. *Economic Modelling* **29**(2012), 1946–1959

[50] Chiang, T.C. and Jeon, B.N. and Li, H.: 'Dynamic correlation analysis of financial contagion: Evidence from Asian markets'. *Journal of International Money and Finance*, **26**(2007), 1206–1228

[51] Chow, G. C.: Tests of equality between sets of coefficients in two linear regressions. *Econometrica: Journal of the Econometric Society* (1960), 591–605.

[52] Csörgö, M. and Horváth, L.: *Limit Theorems in Change-point Analysis.* Wiley, 1997.

[53] Comte, F. and Lieberman, O.: 'Asymptotic theory for multivariate GARCH processes'. *Journal of Multivariate Analysis*, **84** (2003), 61–84

[54] Darling, D.A. and Erdös, P.: A limit theorem for the maximum of normalized sums of independent random variables. *Duke Mathematical Journal* **23**(1956), 143–155.

[55] De Bandt, O. and Hartmann, P.: Systemic risk: a survey. *ECB working paper* (2000).

[56] De Pooter, M. and Van Dijk, D.: Testing for changes in volatility in heteroskedastic time series-a further examination. *Technical report, Econometrics Institute Research Papers* (2004).

[57] Diaconis, P. and Freedman, D.: Iterated random functions. *SIAM Review* **47**(1999), 45–76.

[58] Dimitriou, D. and Kenourgios, D. and Simos, T.: Global financial crisis and emerging stock market contagion: A multivariate FIAPARCH–DCC approach. *International Review of Financial Analysis* **30**(2013), 46–56.

[59] Ding Z. and Engle R.F.: Large scale conditional covariance matrix modeling, estimation and testing. *Working Paper FIN-01-029* NYU Stern School of Business (2001).

[60] Douc, R. and Moulines, E. and Stoffer, D.: *Nonlinear time series: theory, methods and applications with R examples*. CRC Press, Boca Raton, 2014.

[61] Domowitz I. and Hakkio C. S.: Conditional variance and the risk premium in the foreign exchange market. *Journal of international Economics* **19**(1985), 47-66.

[62] Dungey, M. and Milunovich, G. and Thorp, S. and Yang, M.: Endogenous crisis dating and contagion using smooth transition structural GARCH. *Journal of Banking & Finance* **58**(2015), 71–79.

[63] Eberlein E.: On strong invariance principles under dependence assumptions. *The Annals of Probability* **14** (1986), 260–270.

[64] Engle, R. F.: Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society* (1982), 987–1007.

[65] Engle, R. F.: Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics* **20**(2002), 339–350.

[66] Engle, R. F.: *Anticipating correlations: a new paradigm for risk management.* Princeton University Press, 2009.

[67] Engle, R. F. and Bollerslev, T.: Modelling the persistence of conditional variances. *Econometric reviews* **5**(1986), 1–50.

[68] Engle, R. F. and Ng, V. K. and Rothschild, M.: Asset pricing with a factor-ARCH covariance structure: Empirical estimates for treasury bills. *Journal of Econometrics* **45**(1990), 213–237.

[69] Engle, R. F. and Kroner, K. F.: Multivariate simultaneous generalized ARCH. *Econometric Theory* **11**(1995), 122–150.

[70] Fan, F. and Qi, L. and Xiu, D.: Quasi-maximum likelihood estimation of GARCH models with heavy-tailed likelihoods *Journal of Business & Economic Statistics* **32**(2014) 178–191.

[71] Fitzenberger, B. and Koenker, R. and Machado, J.: Economic applications of quantile regression. *Springer Science & Business Media* (2013).

[72] Forbes, K. J. and Rigobon, R.: No contagion, only interdependence: measuring stock market comovements *The journal of Finance* **57**(2002), 2223–2261.

[73] Francq, C. and Zakoian, J.M.: *GARCH Models: structure, statistical inference and financial applications*. Wiley, 2010.

[74] Francq, C. and Zakoian, J.M.: Comment on "Quasi-Maximum Likelihood Estimation of GARCH Models with Heavy Tailed Likelihoods" by J. Fan, L. Qi et D. Xiu. *Journal of Business & Economic Statistics* **32**(2014), 198-201.

[75] Francq, C. and Horváth, L. and Zakoian, J.M.: Variance targeting estimation of multivariate GARCH models. *Journal of Financial Econometrics* **14**(2016), 353-382.

[76] Fryzlewicz, P. and Rao, S. S.: Mixing properties of ARCH and time-varying ARCH processes, *Bernoulli* **17**(2011), 320–346.

[77] Gatfaoui H.: Translating financial integration into correlation risk: A weekly reporting's viewpoint for the volatility behavior of stock markets. *Economic Modelling* **30**(2013), 776-791.

[78] Getmansky, M. and Lo. A. W. and Markarov.: An econometric model of serial correlation and illiquidity in hedge fund returns. *Journal of Financial Economic* **74**(2004), 529–609

[79] Giglio, S. and Kelly, B. T. and Pruitt, S.: Systemic risk and the macroeconomy: an empirical evaluation. *National Bureau of Economic Research working paper* (2015).

[80] Girardi, G. and Ergün, A.: Systemic risk measurement: Multivariate GARCH estimation of CoVaR. *Journal of Banking & Finance* **37**(2013), 3169–3180.

[81] Glosten, L. R. and Jagannathan, R. and Runkle, D. E.: On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of finance* **48**(1993), 1779–1801.

[82] Gombay, E. and Horváth, L. and Husková, M.: Estimators and tests for change in variances. *Statistics & Risk Modeling* **14**(1996), 145–160.

[83] Group of ten, Report on consolidation in the financial sector: Chapter III. Effects of consolidation on financial risk. *working paper, International Monetary Fund* (2001).

[84] Hidalgo, J. and Seo, M. H.: Testing for structural stability in the whole sample. *Journal of Econometrics* **175**(2013), 84–93.

[85] Horváth, L.: The maximum likelihood method for testing changes in the parameters of normal observations. *The Annals of statistics* (1993), 671–680.

[86] Horváth, L. and Rice, G.: Extensions of some classical methods in change point analysis (with discussions). *TEST* **23**(2014), 219–290.

[87] Horváth, L., Pouliot, W., and Wang, S.: Detecting at-Most-m Changes in Linear Regression Models. *Journal of Time Series Analysis* (2016).

[88] Huang, X. and Zhou, H. and Zhu, H.: Systemic risk contributions. *Journal of Financial Services Research* **42**(2012), 55–83

[89] Hurvich, C. M. and Tsai, C.: Model selection for extended quasi-likelihood models in small samples. *Biometrics* (1995), 1077–1084.

[90] Ibragimov, I. A.: Some limit theorem for stationary processes. *Theory of Probability and Its Applications* **7**(1962), 349–382.

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

[92] International Monetary Fund (IMF).: *Global financial stability report: responding to the financial crisis and measuring systemic risks* (2009).

[93] Jeantheau, T.: Strong consistency of estimators for multivariate ARCH models. *Econometric Theory* **14**(1998), 70–86.

[94] Kao, C. and Trapani, L. and Urga, G.: Testing for instability in covariance structures. *Bernoulli* (2016).

[95] Karolyi G. A.: A multivariate GARCH model of international transmissions of stock returns and volatility: The case of the United States and Canada. *Journal of Business & Economic Statistics* **13**(1995), 11-25.

[96] Kim, S. and Cho, S. and Lee, S.: On the cusum test for parameter changes in GARCH (1, 1) models. *Communications in Statistics-Theory and Methods* **29**(2000), 445–462.

[97] Koenker, R.: Quantreg: quantile regression. *R package version.* **5**(2013).

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

[99] Kroner FK. and Ng VK.: Modelling asymmetric comovements of asset returns. *The Review of Financial Studies* **11**(1998), 817–844.

[100] Kritzman, M. and Li, Y. and Page, S. and Rigobon.: Principal components as a measure of systemic risk. *Journal of Portfolio Management* **37**(2011), 41–56.

[101] Laurent, S. and Rombouts, J. VK. and Violante, F.: On the forecasting accuracy of multivariate GARCH models. *Journal of Applied Econometrics* **27**(2012), 934–955.

[102] Lavielle, M. and Teyssiere, G.: Detection of multiple change-points in multivariate time series. *Lithuanian Mathematical Journal* **46**(2006), 287–306.

[103] Lee, S. and Na, O. and Na, S.: On the cusum of squares test for variance change in non-stationary and nonparametric time series models. *Annals of the Institute of Statistical Mathematics* **55**(2003), 467–485.

[104] Ling, S.: Testing for change points in time series models and limiting theorems for NED sequences. *The Annals of Statistics* **35**(2007), 1213–1237.

[105] Melkuev, D.: Asset return correlations in episodes of systemic crises. *Master Thesis* University of Waterloo (2014).

[106] Meyn, S.P. and Tweedie, R.L.: *Markov Chains and Stochastic Stability.* Springer, London, 1993.

[107] Nelson, D. B. and Foster, D. P.: Filtering and forecasting with misspecified ARCH models II: making the right forecast with the wrong model. *Journal of Econometrics***67** (1995), 303–335.

[108] Noureldin, D. and Shephard, N. and Sheppard, K.: Multivariate rotated ARCH models. *Journal of Econometrics* **179**(2014), 16–30.

[109] Orasch, M. and Pouliot, W.: Tabulating weighted sup-norm functionals used in change-point analysis. *Journal of Statistical Computation and Simulation* **74**(2004), 249–276.

[110] Page, E.S.: A test for a change in a parameter occurring at an unknown point. *Biometrika* **42**(1955), 523–527.

[111] Pedersen, R.S. and Rahbek, A.: Multivariate variance targeting in the BEKK–GARCH model. *The Econometrics Journal* **17**(2014), 24–55.

[112] Pesaran, B. and Pesaran, M. H.: Modelling volatilities and conditional correlations in futures markets with a multivariate t distribution. *Working paper* (2007).

[113] Pukthuanthong, K. and Roll, R.: Global market integration: an alternative measure and its application. *Journal of Business and Economic Statistics* **94**(2014), 214–232.

[114] Qu, Z. and Perron, P.: Estimating and testing structural changes in multivariate regressions. *Econometrica* **75**(2007) 459–502.

[115] Quandt, R. E.: The estimation of the parameters of linear regression system obeying two separate regimes. *Journal of the American Statistical Association* **53**(1958), 873–880.

[116] Quandt, R. E.: Tests of hypothesis that linear regression system obeys two separate regimes. *Journal of the American Statistical Association* **55**(1960), 324–330.

[117] Rio, R.: Théorie asymptotique des precessus aléatoires faiblement dépendants. *Mathmátique & Applications* **31** Springer-Verlag, Berlin, 2000.

[118] Riskmetrics, TM., JP Morgan Technical Document. *New York, JP Morgan*, 1996.

[119] Schwarz, G.: Estimating the dimension of a model. *The annals of statistics* **6**(1978), 461–464.

[120] Silvennoinen, A. and Teräsvirta, T.: Multivariate GARCH models. *Handbook of financial time series* (2009), 201–229.

[121] Szyszkowicz, B.: Weak convergence of weighted empirical type processes under contiguous and changepoint alternatives. *Stochastic processes and their applications* **50**(1994), 281–313.

[122] Tse, Y. K. and Tsui, A. K. C.: A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *Journal of Business & Economic Statistics* **20**(2002), 351–362.

[123] van der Weide R.: GO-GARCH: a multivariate generalized orthogonal GARCH model. *Journal of Applied Econometrics* **17**(2002), 549–564.

[124] Vrontos I.D. and Dellaportas P. and Politis D.N.: A full-factor multivariate GARCH model. *Econometrics Journal* **6**(2003), 311–333.

[125] Weiss, A. A.: Asymptotic theory for ARCH models: estimation and testing. *Econometric theory* **2**(1986), 107-131.

[126] Wied, D. and Krämer, D and Dehling, H.: Testing for a change in correlation at an unknown point in time using an extended functional delta method. *Econometric Theory* **28**(2012), 570–589.

[127] Yao, Y.: Estimating the number of change-points via Schwarz'criterion. *Statistics & Probability Letters* **6**(1988), 181–189.

[128] Yao, Y. and Davis, R. A.: The asymptotic behavior of the likelihood ratio statistic for testing a shift in mean in a sequence of independent normal variates. *Sankhyā: The Indian Journal of Statistics, Series A* (1986), 339–353.