UNIVERSITA' DEGLI STUDI DI MESSINA

Department of Economics

PhD in: ECONOMICS, MANAGEMENT AND STATISTICS

XXXII Cycle

Coordinator: Prof. Edoardo OTRANTO

_____

On modelling the multivariate Realized

Kernel financial time series

PhD Student:

Dott.ssa Alessandra COSTA

Tutor:

Illustrious Professor

Edoardo OTRANTO

_____

ACADEMIC YEAR 2018/2019

Sommario

Introduction and Objectives of the Work

It is widely known that financial time series are characterized by very complex patterns and dynamics, that have to be accounting in estimation and forecasting analysis. Multivariate GARCH models have been and continue to be the most widely used time series models, both with parametric and non-parametric specifications: major examples are BEKK of Engle and Kroner (1995), the CCC of Bollerslev (1990) and the DCC of Engle (2002). Those models allow for volatility and correlations to change over time, and one of their main advantage is their easily implementation in many areas. More recently, the growing availability of high frequency data allows the development of different ex-post unbiased estimators of daily and intraday volatility, and this stimulated the development of specific models directly fitted on realized measures. Thus, although many multivariate variants of existing models could be applied on realized measures too, other specific models were adopted. Just to mention a few, the Realized GARCH of Hansen, Huang and Shek (2010), the Multiplicative Error Model of Engle (2002) and HEAVY models. Moreover, when multiple assets are involved, other issues arise (asynchronous trading and microstructure noise). Previous considerations underline the need for parametrizations and models able to guarantee the positiveness of variance covariance matrix at each point in time. As a natural consequence, the best choice is to adopt a Wishart distribution that automatically generates PSD matrices. Secondly, since the number of parameters to estimate clearly are proportionally to the number of assets included in the analysis, the estimation process could become infeasible. Thus, to overcome the curse-of-dimensionality problem that could arise, it is necessary to define several constraints in the models to estimate, however guaranteeing a high level of flexibility in the model parametrization. In view of this background, this thesis is structured as follow:

1. Chapter 1 introduces some preliminary concepts about volatility and its stylized facts, high frequency data and realized measures;
2. Chapter 2 presents a class of models inspired by the consideration of simplicity, feasibility and computational speed in high dimensional environment.

Among them, a new parametrization that combine a multiplicative structure with dynamic correlations and variances, the DCC-CAW-MEM, directly fitted on the realized variance covariances matrices. After exploiting its characteristics, its estimation process (ML approach and QML interpretation), we also consider other parametrizations that have been found to be fit well in high dimensional context. Some of them rely on the same multiplicative error structure of DCC-CAW-MEM and require the maximization of the same likelihood function, even if each of them presents a different structure and has a different degree of flexibility.

3. Chapter 3 shows the empirical application of all the models discussed in the previous chapter, both in the in-sample and out of sample scenarios.

Chapter 1

A Survey of Realized Volatility and High Frequency Data

Abstract

In this chapter, I briefly review the concept of volatility, its main features and problems that can arise during its measurement. Then I introduce the topic of high frequency (HF) data, by focusing on the relationship between HF data and the microstructure noise. Consequently, I present the most common ex-post volatility estimator, the Realized Variance (RV) (Andersen et al, 2003). Then I extend the concept to the multivariate context, by considering more than two assets through the introduction of realized covariance estimators. Both in the univariate and in the multivariate setting, I will start from the introduction of a diffusion continuous time framework of analysis (Brownian motion).

1.Introduction

Volatility represents an interesting topic in financial time series, because it plays a central role in different area[1]. Although volatility is, in principle, a clear concept, its statistical treatment is not simple, because it is a latent variable which has to be estimated. This explains why it has received a lot of attention in literature, that has moved into two directions: the development of new models with richer formulations, that should result in higher estimation and forecasting performances, and a better classification of the existing models. With respect to the model development process, models range from the basic historical volatility ones (random walk, moving average, etc.) to the more complex and well-known ARCH ang GARCH models and their extension. Comprehensive reviews of literature can be found in Bauwens et al (2006) and in Silvennoinen and Terasvirta (2009). Referring to ranking among models, there is no consensus about which class of model performs better, and such conclusion comes from a set of mixed literature findings, even if many GARCH models still perform well in many applications. More recently, the availability of intraday data has led to the rise of a growing interest in new ex-post volatility measures. In fact, realized volatility measures solve some issues of the traditional methods that rely on daily data (usually, daily squared returns).

High frequency data (HF) introduces the possibility to add further assumptions for multivariate process high frequency variables. The availability of a great amount of information has led to an increase in the accuracy of the forecasts and to an advantage over traditional low frequency methodologies, which require data over a long period

---

[1] Poon and Granger (2003) stated that volatility forecasting is an important tool in financial risk management, while Andersen et al (2003) found three preliminarily applications of it: a) pure forecasting applications; 2) financial risk management for banks and trading houses; 3) generic applications outside financial field, but related to real economy. Relative to the last application of volatility forecasting, a lot of evidence can be found in literature. Just to mention a few, Coshall et al (2009) applied univariate volatility models to the UK tourism demand to the country's most popular international destinations, using twelve quarterly time series collected from the UK Office for National Statistics over the period 1977-2008, while Moran et al (2017) used multivariate time series analysis for modelling daily mean-mortality series of intensive car society units (ICUs) from Australian adult patient database, also confirming the validity of classical financial time series patterns. Moreover, Escribano et al (2018) modelled the volatility of electricity prices of Nord Pool Spot AS through a multiplicative component log-GARCH-X. Obviously, other relevant examples can be founded, but the selected ones help to evidence how much volatility model techniques are widely used in all interesting areas.

of time. At the same time, there could be gains from financial processes and from the successful management of risks and trading costs.

Moreover, one of the most important aspect when we talk about high frequency data is the sampling frequency. Even if using low frequency data measure could lead to imprecise estimates and forecast, computing the estimates to a very high frequency introduce some biases in the estimation itself, due to the contamination of microstructure noise. This has led many researchers to focus their research on ex-post estimation of daily variance and covariance of assets, using information based on intra-daily price registrations and on the development of unbiased estimators and alternative approaches.

1.2 Volatility

In everyday language, volatility refers to the fluctuations observed in a certain phenomenon over time, while in economics it is used to describe the variability of a time series.

More formally, in financial economics, the term volatility is defined as the standard deviation of a continuous time process.

Depending on data availability and on the uses of model estimates, volatility models can be designed in discrete time or continuous time, even if it is clear that in many liquid financial markets, transactions take time at very short interval. As a consequence, it is natural to think about the priced and returns series of financial assets as arising through discrete observations from a continuous time diffusion process. Although many discrete models have been shown to be not consistent with the corresponding continuous modes, they are still used and preferred because of their flexibility in the estimation process. In this thesis I refer to volatility as the level of uncertainty associated to a financial asset, and it has been proved to be characterized by certain stylized facts, that are crucial for the correct specification of all the volatility models.

1.2.1 Fat tails

It is a well-known point that the distribution of financial time series normally exhibits fatter tails than the normal distribution. Assuming that the fourth order moment of a GARCH(1,1) exists, Bollerslev stated that many financial time series exhibit a kurtosis index greater than the standardized value of 3. Afterward, He and Terasvirta (1999) generalized the same results for the GARCH(*p,q*).

1.2.2 Volatility Clustering

Volatility clustering was firstly introduced by Mandelbrot (1963), who reported "… in other words large changes tend to be followed by large changes- of either sign- and small changes tend to be followed by small changes…".
Lux and Marchesi (2000) attribute volatility clustering to the agent switching between fundamentalist and chartist strategies, while Kirchler and Huber (2007) suggested as source of volatility clustering the presence of heterogeneous information, and by assuming that traders learn from observed prices[2].
As reported by Aldridge (2009) current high volatility does not typically revert to lower volatility levels instantaneously, but higher volatility typically persists for different periods (this is also known as ARCH or GARCH effect). Thus, clustering is strictly related to the persistence or long memory of volatility. Especially in high frequency context, there is a lot of evidence of a near unit root in the conditional variances. Bollerslev (1998) suggests to use a visual inspection for determining the time series properties of volatility, by referring both to the absolute returns and square returns in the correlogram analysis.

1.2.3 Leverage effect

The leverage effect, also known as volatility asymmetry (Black, 1972), typically refers to the negative correlation between volatility and past returns. More precisely, this

---

[2] For further details, see Smith, V. L. (1982), Markets As Economizers Of Information: Experimental Examination Of The "Hayek Hypothesis". Economic Inquiry, 20: 165-179. doi:10.1111/j.1465-7295.1982.tb01149.x

effect implies that an increase in asset prices is accompanied by declining volatility, and it is asymmetric, so a moderate reduction of an asset price determines a larger increase of financial volatility.

1.3 High Frequency Data

In the last years, the growing availability of high frequency data has represented one of the most attractive field of research in financial econometrics. With financial and technological developments, the possibility to have tick by tick data has increased the interest of researchers in their management, to better understand the structure of financial markets.

Andersen et al (2003) highlighted the improvements high frequency data produces in terms of predictive performances respect to the standard procedures that rely on daily data alone. This, in principle, led to obvious advantages in financial applications, especially in situation where prices arise from a continued time diffusion process. In other words, we could expect that more observation we have, the more precise the estimations will be. Unfortunately, some problems arise. It is obvious that except for most liquid assets, the continuous price is not available, and this implies the existence of a measurement error that must be taken into account. Secondly, when the sampling frequency increases, it also will increase the probability of occurrence of spurious correlations. Therefore, in practical applications, we have to deal with a trade-off between accuracy of estimates and forecasts and measurement errors, in order to reduce bias and distortions in statistical inference. In this section I analyze the trade-off between microstructure noise and sampling scheme, by introducing some sources of microstructure noise and analyzing the different sampling schemes available in financial literature.

Falkenberry (2002) synthetized the nature of the problems arising from the creation and the management of high frequency database. Among others (different nature of error types, intraday seasonal patterns in tick frequency, treatment of time, etc.), the main sources of microstructure noise are:

- Asynchronous nature of tick data;
- Bid-ask bounce.

1.3.1 Asynchronous trading

In an ideal market, trading is synchronized, that means that all the buyers and sellers meet at the same time and conclude transactions, whose prices are recorded in the same moment. In practice, this not happens, and transactions take place in different moments. The effect of non-synchronous trading has firstly been investigated by Atchison et al (1987), who examined its effect on autocorrelation among assets. The effects of asynchronous trading arise when time series are wrongly measured. Generally, reported daily returns only reflect the last trade that takes place. Such scenario means that, in high frequency context, one incorrectly assumes daily returns as an equally space time series over a 24-hours interval. Scholes and Williams (1977) demonstrated that this non synchronous trading induces lag cross correlations and lag serial cross correlations into returns. Most importantly, Tsay (2003) evidenced this effect in terms of portfolio construction, since it leads to biased correlations among assets and inaccuracy in assessing risks in management strategies.

1.3.2 Bid-ask spread

The bid-ask spread, often considered as trading cost, is due to the difference between the price at which the buyer wants to buy and the price at which the seller wants to sell.

Following Stoll (1989), the transaction cost from which bid-ask spread arises comes from under processing costs, inventory holding costs and adverse information costs. In fact, in order to maintain liquidity, many exchanges use market makers who stand ready to buy or sell whatever the public wishes to buy or to sell, but they usually buy at bid price and sell at higher ask price. As a convention among market makers, bid-ask spreads usually take few discrete values, while in practice the bid-ask spread is time varying and related to both volume and volatility. Such result is due to transaction prices which are free to move between the ask and the bid, in a manner that they can

produce different log returns, even if the economic values of the assets is still the same. This scenario is known as bid-ask bounce, that is the bid-ask spread that introduce negative lag-1 serial correlation in asset returns, that has been exploited and modelled by Roll (1984).

1.3.3 Univariate Setting: Quadratic Variation and Integrated Variance

Now I am presenting the framework on which my analysis, in univariate setting, is based. As mentioned before, since many models have been formulated in a discrete time framework, it is natural to start from observation available at equally spaced discrete points in time. Let's think to a discrete process

$$Y \equiv y(t)$$

$$with\ t = 1,2,\dots,T$$

where $y(t)$ is assumed to arises from a continuous process, with the following conditional moments (mean and variance)

$$\mu_{t|t-1} = E(y_t|F_{t-1})$$

and

$$\sigma_t^2|F_{t-1} = Var\ [y_t|F_{t-1}] = E(y_t - \mu_t)^2|F_{t-1}$$

where F_{t-1} is the set of information available at time t-1, that determines the differences between the conditional and unconditional moments.

As previously noted, most of the main applications in volatility modelling have been developed in the context of financial economics, thus it is natural to think about prices and returns. If we assume that p(t) is the univariate process of the logarithmic price evolving in the continuous time over a discrete internal [0,T], Back (1991) and Andersen et al (2003), and if the return process does not allow for arbitrage and a finite instantaneous mean, the asset price belongs than the log prices belong to the class of semi-martingale and it can be written as the sum of a drift component A(t) and a local martingale M:

$$p(t) = p(0) + A(t) + M(t)$$

Moreover, the latter can be splitted in two components: a realization of a continuous process, $\Delta A(t)$ and $\Delta M(t)$ and a jump component. With the initial assumption that $\Delta M(0) \equiv A(0) \equiv 0$

$$r(t) = p(0) + A_{(t)}^c + \Delta A(t) + M_{(t)}^c + \Delta M(t)$$

This decomposition is useful for the qualitative characterization of the asset return process. Since the return over [t-h, t] is defined as $p(t) - p(t-h)$, this implies that:

$$r(t) = p(t) + p(t-h) = A(t) + M(t)$$

Similarly, we can refer to the multivariate case involving of returns, by the upper-case letter R.

Until now I have considered a discrete time framework but, as noted before, it is natural to think of prices and return arising from a continuous diffusion process in the form of:

$$\delta p(t) = \mu(t)\delta(t) + \sigma(t)\delta W(t)$$

$$with \ t \geq 0$$

where $W(t)$ is a Wiener process, $\mu(t)$ is a finite variation stochastic process and $\sigma(t)$ is a strictly positive cadlàg and it is the instantaneous volatility of the process. From this characterization, it follows that over the interval [t-h, t] if we consider a unit variation, h=1, we get

$$r(t) = p(t) - p(t-1) = \int_{t-1}^{t} \mu(s)\delta s + \int_{t-1}^{t} \sigma(s)\delta(W)s$$

$$with \ 0 \leq s \leq t \leq T$$

The similitudes between the discrete and continuous characterizations are clear: the conditional mean and variance process are replaced by the corresponding integrated realizations of means and variances. Because the variation of the diffusion path of martingale is not affected by the drift component, intuitively its evolution is related to the evolution of the coefficient σ.

Formally, the return variation is equal to the so-called quadratic variation:

$$QV(t) = \int_{t-h}^{t} \sigma^2(s)\delta s$$

The quadratic variation process could be approximated by cumulating cross products of high frequency returns. As noted by Andersen et al (2003), the quadratic variation is the main determinant of return covariance matrix, especially for shorter horizons, since the variation induced by genuine innovation, represented by the martingale component, locally is an order of magnitude larger than the return variation caused by changes in the conditional mean.

The diffusive sample path variation over [t-h, t] is also known as the integrated variance IV.

$$IV(t) = \int_{t-h}^{t} 6^2(s)\delta s$$

Under the simplifying assumptions of absence of microstructure noise and measurement error, the quadratic and integrated variation coincide.

1.3.4 Realized Variance (RV) Estimator

Given the previous characterization, the researchers developed different estimators of the integrated variance. The most widely used is the realized variance (RV) estimator. I firstly examine the ideal context in which the absence of measurement errors made ex post volatility became observable. Andersen et al. (2003) suggested to estimate the realized variance as the sum of squared intraday returns, sampled at sufficiently high frequency. Considering $t - h + \frac{j}{n}$ a partition of $[t - h, t]$ interval, thus the realized variance will be defined as

$$RV_t = \sum_{h=1}^{T} \left(Y_{t-h+\frac{j}{n}} - Y_{t-h} \right)^2$$

It has been shown to be an unbiased estimator of ex-ante volatility, and in case of absence of microstructure noise it is also an unbiased estimator of the integrated variance. Moreover, several authors refer to realized volatility as the square root of RV.

Many authors have started to individuate an appropriate framework for the estimation and the prediction of conditional variance of financial assets, in order to accommodate for the entire set of information in intraday data. Based on the results of Jacod and Protter (1998), Barndorff-Nielsen (2002) derived the asymptotic distribution of the realized variance as normal with zero mean and unit variance. As sampling frequency increases, the realized variance converges in probability to the quadratic variation (QV)

$$RV(t, h, n) \to QV(t, h) \quad as\ n \to \infty$$

This is the formal link between RV and return quadratic variation, that directly follows from the theory of semi martingales and implies that as sampling frequency increases, RV provides a consistent measure of the IV. If this is true, in theory one could reduce the measurement error by increasing the sampling frequency. However, given the arguments introduce in Section 1.3, returns shouldn't be sampled to often to avoid microstructure noise.

1.3.5 Sampling frequency and Market Microstructure Noise

As stated above, even if high frequency data represent a source of continuous information, there is a well-known bias between sampling frequency and microstructure noise.

Andersen et al. (2003) highlighted how much crucial is the choice of the sampling frequency, considering high frequency data on deutschmark and yen returns against the dollar. By simulated volatility signature at varies sampling intervals, they individuated as optimal the 20-minute return sampling interval (k=20), that contemporaneously measures the microstructural bias and the sampling error.

Aït-Sahalia et al. (2005) and Zhang et al. (2009) proved that RV is not a reliable estimator for the true variation of the returns in case of microstructure noise, since as sampling frequency increases, the noise becomes progressively more dominant. Consequently, literature recommends to use a realized variance estimator constructed by summing up intraday returns at some lower sampling frequency, usually 10, 15 or 20 minutes.

About this, it is possible to refer to the sparse sampling scheme, and it has been suggested by Andersen et al (2001). When sampling frequency is lower than the -1 min, sparse RV, (RV^s) will be:

$$RV_t^s = \sum_{i=1}^{n_t^s} Y_{t,i}^s$$

It has been shown that the bias due to the noise is reduced when $n_t^s < n$, even if the total variance increases. Although sparse sampling tries reducing the microstructure noise in RV calculation, it potentially suffers from a loss of information and leads to a reduction in the efficiency of the estimator. For this reason, Bandi and Russel (2008) proposed to select the optimal sampling frequency by minimizing MSE (mean square

error). The authors presented a methodology similar to Bei et al. (2000), that is able to choose optimal sampling frequency by minimizing the expected square distance between the estimator (RV) and its theoretical counterpart (QV) as summarized by conditional MSE.

At the same time, Zhang et al. proposed an alternative method to benefit from using infrequently data. Instead of arbitrarily selecting a subsample, the authors suggested to select a number of sub-grids of the original grid of observation time $\mathcal{G} = \bigcup_{k=1}^{k} \mathcal{G}(k)$, where $\mathcal{G}^{(k)} \cap \mathcal{G}^l \neq \emptyset$ (when k $\neq$ l) and the natural way to select the kth subgrid is to start from t_{k-1} and then pick every kth sample point until T, that means:

$$\mathcal{G}^k = \{t_{k-1}, t_{k-1+k}, t_{k-1+2k}, \dots, t_{k-1+n_k k}\}$$

Given the realized volatility $RV_t^{(k)} = \sum_{i=1}^{n} Y_{t,i}^2$, Zhang et al. averaged the estimator, and then:

$$RV_t^{(avg)} = \frac{1}{k} \sum_{k=1}^{k} RV_T^{(k)} - \frac{\overline{n_t}}{n} RV_T^{(all)}$$

Fixing T and using the observation within the interval [0,T] asymptotically as $n \to \infty$ and $n/k \to \infty$, n_k hasn't to be the same across k, we could define we get

$$\overline{n_t} = \frac{1}{k} \sum_{k=1}^{k} n_k = \frac{n - k + 1}{k}$$

The estimator is called Two Time Scales Estimator (TTSE), obtained through a combination of single grid and multiple grids.

Under the assumption of IID noise structure (the microstructure noise has zero mean and is IID random variable independent of the price process and the variance of v_t is O(1)), they showed that

$$n^{1/6}(RV_t^{ZMA} - IV) \xrightarrow{\mathcal{L}} 8c - 2((E\varepsilon^2)^2 + c\eta^{2T})^{\frac{1}{2}} N(0,1)$$

and the converge is stable in law. For equidistant sampling and regular allocation to grids, $\eta^2 = \frac{4}{3}\sigma^2 T$, the variance is equal to $8c - 2(E\varepsilon^2)^2 + c\eta^2 T = 8c - 2(E\varepsilon^2)^2 + c\frac{4}{3}\sigma^4 T^2$.

In their paper, Zhang et al. demonstrated that the optimal choice of c becomes:

$$c_{opt} = \left(\frac{12(E\varepsilon^2)}{T^2\sigma^4}\right)^{1/3}$$

As the authors pointed out, it is also possible to estimate c to minimize the actual asymptotic variance from data in the present time period, or $0 \leq t \leq T$. With a bias-type adjustment, the estimator will be unbiased

$$RV_t^{adj} = \left(1 - \frac{\overline{n}}{n}\right)^{-1} RV_t$$

The difference between the two estimators is of order $Op\left(\overline{n}/n\right) = Op\left(k^{-1}\right)$.

Moreover, referring to the sample schemes, McAleer and Medeiros (2008) individuated four sampling schemes:

1. Calendar time sampling (CTS), where the intervals are equidistant in calendar time, that's $d_{i,nt} = \frac{1}{n_t}$ for all i.

2. The transaction time sampling (TrTS), where prices are recorder every n-th transaction;

3. Business time sampling (BTS), where the sampling times are chosen such that $IV_{it} = \frac{IV}{n_t}$;

4. The tick time sampling (TkTS), where prices are recorder with each transaction.

As pointed out by Oomen (2006), the choices of both sampling and calendar schemes are crucial for modelling and forecasting. Oomen showed that from theorical and empirical analysis, transaction time sampling is superior to the common approach of CTS, and it leads to a better minimization in MSE of the realized variance, yielding more accurate variance estimates.

Furthermore, Griffin and Oomen (2008) proposed a comparison of the effect of tick time and transaction time sampling on RV calculations.

1.3.6 Kernel-Based Estimator

Motivated by the some of the issues mentioned before, in literature other estimators have been proposed for modeling the microstructure noise, such as the Kernel based estimators of QV, in presence of autocorrelation caused by microstructure noise.

Zhou (1996) was the first to use kernel function to deal with the problem of microstructure noise in high frequency data, proposing the following estimator:

$$RV_t^K = RV_t + 2\sum_{k=1}^{K} \frac{n}{n-k}\hat{\gamma}_k$$

Where $\hat{\gamma}_k = \left(\frac{n}{n-k}\right)\sum_{j=1}^{n-h} r_{t,j}r_{t,j+1}$, RV_t is the all RV and $k(\cdot)$ is the Kernel function and n is the sum of intraday observations used in RV computation.

Even if this estimator is unbiased, it is not consistent when we move to continuous. Extending Zhan's work, Hansen and Lunde (2005) analyzed the implication of microstructure noise on RV estimates, and by using the Down Jones Industrial Average, they revealed some features of microstructure noise, with important implication for RV estimation.

Given the time dependency and the correlation with the returns of the efficient price, the authors introduced the RV_{AC} estimator, that uses the first order autocorrelation to bias correct the RV. Starting from Zhou's estimator of 1996, the authors derived its properties allowing for non-constant volatility and non-Gaussian microstructure noise. Afterwards they compared the two estimators, demonstrating that RV_{AC} performs better, as summarized by a greater reduction of MSE.

1.3.7 Multivariate Setting: Quadratic Covariation

In this section I focus on the analysis of methodologies used for producing ex-post measures of the covariance between assets.

Before turning into the discussion about them, I have to introduce several assumptions about the underlying data generation process. As in the univariate setting, I focus on a continuous time diffusion model, supposing that I have n stocks whose log price process is Y_t^i. As evidenced in Section 1.3.3 the continuous model can be summarized as:

$$Y_t = \int_{t-1}^{t} \mu(s)\delta s + \int_{t-1}^{t} \Theta(s)\delta(W)s$$

where μ is a m by 1 vector of drift process and Θ is the spot co-volatility, while W is a vector of the Brownian motion.

In this the quadratic variation can be defined using matrix notation, as:

$$QC = plim \sum (Y_{tj} - Y_{tj-1})(Y_{tj} - Y_{tj-1})'$$

Following advances in the realized co-volatility literature, we are interested in modelling and forecasting the daily ex-post covariation, through appropriate realized covariance estimators.

The most used estimator of quadratic covariation is the so-called Realized Covariance (RC), introduce in financial literature by Barndorff-Nielsen & Shephard (2004). The authors showed that, if we can observe the log price processes at discrete points in time over an interval [0,T], with T=1, and if we have j intraday observations, the RC estimator is the sum given in the definition of quadratic variation, so that:

$$RC_t = \sum_{j=1}^{n}(Y_{tj} - Y_{tj-1})(Y_{tj} - Y_{tj-1})'$$

Also in this case, under the assumption of no microstructure noise, it has been shown that RC is an unbiased estimator of the variance covariance matrix Σ_t. Even in theory, the estimation will be simple, in practice some problems persist (asynchronous trading and microstructure noise).

1.3.8 CholCov Estimator

Since we rely on high frequency data, it is necessary to individuate an appropriate and consistent estimator of realized variance covariance matrix, by considering that in a multivariate setting, besides the accuracy of the estimator, the literature has extensively studied the need of a positive definite estimators[3]. Since our object of interest is the spot covariance matrix over the unit interval

$$ICOV = \int_0^1 \Sigma(s)\delta(s)$$

---

[3] For a complete survey about the estimation of both low and high dimensional covariance matrix and related estimation problems see Wei Biao Wu and Han Xiao (2011).

In this work for its computation, I will refer to Boudt et al. (2016) estimator, the CholCov, that has been found to be a consistent ex-post estimator of the variance covariance matrix, both to microstructure noise and asynchronous trading (after some refinement procedures). Starting from the ex-post covariation of log-prices, the estimator relies on a Cholesky decomposition[4] of the spot covariance matrix into symmetric positive definite square matrices. By the Cholesky factorization, the spot covariance matrix $\Sigma(s)$ can be decomposed into a lower triangular matrix $H(s)$ with unit diagonal elements and a diagonal matrix $G(s)$, so that

$$\Sigma(s) = H(s)G(s)H(s)'$$

that means, with a matrix representation:

$$H(s) = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ h_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ h_{n1} & 0 & \cdots & 1 \end{bmatrix} \; and \; G(s) = \begin{bmatrix} g_{11(s)} & 0 & \cdots & 0 \\ 0 & g_{22(s)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & g_{dd(s)} \end{bmatrix}$$

And, as reported by the authors, the link of each element of H and G matrices to the elements of the spot covariance matrix $\Sigma(s)$ are represented by the following links:

$$g_{kk} = \Sigma_{kk} - \sum_{m=1}^{k-1} h_{km}^2 g_{mm}$$

$$h_{kl} = \frac{1}{g_{ll}} \left(\Sigma_{kl} - \sum_{m=1}^{k-1} h_{km} h_{lm} g_{mm} \right)$$

The usefulness of this decomposition is the possibility to use as much data as possible and hold for any PSD matrix, that is true for the covariance matrix. Moreover, the authors demonstrated that its computational effort is compensated by an efficiency

---

[4] To be more precise, the Cholesky decomposition has widely been used in the volatility context (see Chiriac and Voev (2011) and Palandri (2009).

gain. if compared to other robust estimators (think about the multivariate realized kernel of Barndorff-Nielsen et al. (2011)), even if in case of microstructure noise the authors suggest to better refine the estimation procedure, by using integrated variance robust estimators.

Realized variance covariances matrices used in the empirical application of Chapter 3 has been computed through the CholCov estimator, in order to obtain a series of PSD matrices as input of our estimates. Moreover, for better ensuring the reliability of the inferential results, some econometricians suggest to use a robust procedure for getting robust standard errors (lets' think about the sandwich estimator or the Moore—Penrose pseudo-inverse, when the number of stocks is greater than the number of observations).

Chapter 2

A Class of Multivariate Models for Realized Conditional Variance and Covariance Matrices

Abstract:

Starting from the Engles's (1982) paper about ARCH model, many models have been developed in order to better modelling financial volatility and its stylized fats. Even if ARCH and GARCH models are still the most used models in many practical applications, recent attention has been devoted to the volatility measurement on the basis of ultra-high frequency data.

Despite the effectiveness of GARCH models, the new frontiers in analyzing volatility is represented by the Multiplicative Error Model, the MEM (Engle, 2002; Engle and Gallo, 2006; Brownlees et al., 2011), in which volatility is modelled as a product of a time-varying factor and a positive random variable, ensuring positiveness, without resorting to logs.

Moreover, in high frequency context, another task of great importance is the modelling of variance covariance matrix. Although the RC estimator introduced before, the new frontiers in financial literature are represented by dynamic parametrizations for time-varying daily variance covariances matrices. In this spirit we will refer to formulations consistent with dynamic conditional correlation typical of GARCH volatility models.

2.1 Introduction

Even if GARCH models were originally defined in a univariate setting, their formulations have been easily adapted to the multivariate context.

It is a well-known fact that financial time series are characterized by complex patterns and that an important role is played by the co-movements among assets. Thus, considering only financial modelling in univariate setting could led to a misspecification of the models, accounting for co-movements among assets has become a relevant topic in financial literature. In this field of research there has been a flourishing literature and many models have been developed. Multivariate GARCH models have been and still are the most widely used time series to correctly represent the covariance matrix. Let's think about BEKK model of Baba, Engle, Kraft e Kronecker (1995), the Conditional Constant Correlation of Bollerslev (1986), the Dynamic Conditional Correlation of Engle (2002).

Moreover, respect to the traditional setting, when moving from univariate to multivariate framework of analysis, practitioners have to deal with two additional issues, that have been different exploited in research: the imposition of some additional restrictions and conditions to ensure the positiveness of the variance-covariance matrices.

To this regard, different distributional choices have been taken into consideration, in order to obtain a PSD variance covariance matrix, such as the exponential and the gamma distribution, and the multivariate Wishart distribution.

About the latter, one example is the Conditional Autoregressive Wishart (CAW) model of Golosnoy et al. (2012), where the realized variance covariance matrix has been modelled trough the combination of a Wishart distribution and a BEKK time-varying dynamics. In the same spirit, Bauwens et al. (2016) explored the potential benefits of combining high and low frequency data, in the spirit of the Midas Data Sampling (MIDAS) filter, by separately accounting the long run level of covolatilities and their short run dynamics, through CAW-MIDAS.

In addition, the HF data has increased the interest of researchers to model the series of realized covariances, that could be obtained by exploiting all the information available.

Thus, in this chapter I introduce the models used in the chapter 3, for the empirical application. Even if here the interest is not to offer a complete survey of the models, I will highlight the interesting features of each model used in the empirical study. Moreover, by combining some innovative key elements founded in the financial literature, I introduce a DCC-CAW-MEM model in high dimensional context. More precisely, I combined a consistent DCC-GARCH type parametrization with a vector Multiplicative Error Model (MEM) of Cipollini et al (2012), through a multivariate Wishart distribution. The time varying parametrization allowed to extract the co-movement among assets, by using all the information guaranteed by a high frequency multiplicative error structure. Moreover, the additional Wishart distribution allowed to use a Quasi Maximum Likelihood approach, that simplify the optimization.

Finally, before turning into discussion, it is important to note that in many practical applications, the multivariate modelling is translated into much simpler multiple univariate procedures, after appropriate series transformations. Moreover, instead of concentrating on univariate formulation, I will directly describe their multivariate extensions, that will be used in the empirical application in Chapter 3.

2.2 General Framework

The general framework proposed here is the one illustrated in the Section 1.3.7. For the model specification, let's start from the Wishart distribution[5]. Suppose that C_t is a sequence of positive definite symmetric (PDS) realized covariance matrices of daily realized kernel volatility of n assets, conditioned on the available set of information F_{t-1}. By assumption, the conditional distribution of C_t, given the past history of the process F_{t-1}, consisting of C_t for T ≤ t-1, follow a Wishart distribution:

$$C_t | F_{t-1} \sim W_n \left(v, \frac{S_t}{v} \right)$$

where v is the degrees of freedom parameter and $\frac{S_t}{v}$ is a PDS scale matrix of order n.

From the properties of Wishart distribution[6], it follows that:

$$E(C_t | F_{t-1}) = E_{t+1}(C_t) = S_t$$

whit S_t the conditional covariance matrix, whose i,j-th element is the covariance between assets i and j.

The previous equations define the baseline CAW (Conditional Autoregressive Wishart) model with a time-varying scale matrix.

Our idea is to use a decomposition of the scale matrix S_t, in order to separately account for volatility and correlations. In the spirit of Bauwens, Storti and Violante (2012) and Bauwens, Braione and Storti (2016), I firstly decomposed the scale matrix in term of standard deviation and correlations, so that:

---

[5] Several authors use the Wishart distribution, because it allows to model each element of the correlation matrix, when more than two elements vary and it guarantee positive definite draws. Gourieroux, Jasiak and Sufana (2009) use a Wishart process of realized covariance matrices, a WAR (p) or Wishart autoregressive process, while Bonato, Caporin and Ranaldo (2012) proposed a block structure of WAR model, in order to reduce the number of parameters. Afterwards, an autoregressive inverse Wishart process has been proposed, that can directly be applied in portfolio optimization problem.

[6] For further details see Muirhead (1982) and Anderson (1984)

$$S_t = D_t R_t D_t$$

where R_t is a correlation matrix and D_t is a diagonal matrix given by the conditional standard deviations:

$$D_t = \begin{bmatrix} \sqrt{S_{11,t}} & 0 & \cdots & 0 \\ 0 & \sqrt{S_{22,t}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{S_{nn,t}} \end{bmatrix}$$

The specifications used for their temporal dynamics are explained in the following sub-chapter.

2.1 Univariate Volatility Modelling

To begin I opted for a much simpler univariate specification for the conditional variances, that exclude interaction term, in order to simplify statistical inference.

For the dynamics of the scale matrix the choice differs from the traditionally adopted models: despite the previous literature focuses on GARCH (1,1) model, my specification focuses on MEM (1,1) type process.

In fact, in their traditional setting, GARCH models use daily returns (usually squared returns) to extract information about the current level of volatility but this is not best approach in contexts where volatility changes rapidly. In theory, they can be used to estimate realized covariance, as in Engle and Rangel (2008) but, since they rely on daily observed returns, in principle they provide less precise estimates and forecasts of variances and covariances than measures based on intraday data.

Thus, I adopted the high frequency counterpart of GARCH model, the so-called Multiplicative Error Model (MEM) of Engle and Gallo (2006), that is based on the MEM structure proposed by Engle (2002).

Let's call RV_t the realized volatility of an asset and let F_{t-1} be the information set available at time t for forecasting. In the univariate framework of analysis, RV_t is modelled as the product of time varying conditional expectation S_tt and a positive random variable ε_t:

$$RV_t = S_t \varepsilon_t$$

$$\text{with } \mathcal{E}_t|F_{t-1} \sim Gamma\ (\vartheta, \frac{1}{\vartheta})$$

and S_t evolving according to a parameter vector θ:

$$S_t = S(\theta, F_{t-1})$$

In the simplest case, the conditional expectation is specified as:

$$S_t = \omega + \gamma RV_{t-1} + \delta S_{t-1}$$

Where α takes into account the contribution of recent observations. Stationary requires that all the coefficient in the conditional expectation equation of non-negative and that the sum is less than 1, thus ω≥0, γ ≥0, δ ≥0 and (ω+ γ + δ) <1.

The previous assumptions on S_t and $\mathcal{E}_t$, if the Gamma distribution assumption of Engle and Gallo (2006) is adopted, allows to define the following conditional moments:

$$E(RV_t|F_{t-1}) = S$$
$$Var\ (RV_t|F_{t-1}) = 6^2 S_t^2$$

Moreover, as for GARCH model, this implies that the conditional mean is equal to

$$S = \frac{\omega}{1 - \gamma - \delta}$$

2.2 DCC-CAW-MEM

After designed the volatility model to be used, the second stage of analysis consisted in designing the dynamic process of the conditional correlation matrix R_t. In order to realize this, I resorted the multivariate GARCH models widely used in econometrics. In a preliminary analysis, simply referring to correlation, it is obvious that one could separately compute the correlation coefficient, among series, through the classical ratio between covariance and the product of the variances among assets:

$$p_{i,j,t+1} = \frac{6_{i,j,t+1}^2}{6_{i,t+1} 6_{j,t+1}}$$

However, financial econometric developed models that directly account for correlations.

The latter play an important role in financial markets, because they allow to better exploit the market structure and, contemporaneously, they show how much important is the diversification in portfolio strategies.

Investors usually prefer portfolio with low risks and high returns and portfolio optimization theory suggests diversification as one of the best solutions to achieve redditivity and reduce risks.

Evidence of the importance of correlations has further been investigated in literature. King and Wadhwani (1990) found that the contagion coefficients increased during and immediately after the crash in response to rise of volatility but declined as volatility decreases and nothing in the model implies that contagion coefficients are necessarily constant. Goetzman, Li and Rouwenhorst (2005) distinguished two sources of diversification benefits obtained from international correlations: the variation in average correlation and the variation of the investment opportunity set faced by traders, and developed a test for structural changes in matrix correlation. The authors also conducted an element by element test and a more general test about the average correlation differences, and both of them rejected the hypothesis of constancy of the correlation structure between periods and across countries. Lin, Engle and Ito investigated the contemporaneous correlations of returns and volatilities of stock indices between Tokyo and New York, by fitting two separate GARCH models on daytime and overnight returns, using intraday data. Afterward, many researchers focused on testing the hypothesis of constancy of conditional correlations. Tse (2000) introduced a test based on Lagrange multiplier (LM) approach, while Bera and Kim (2002) developed a form of information matrix (IM) test for assessing the time varying structure of correlations, after fitting a bivariate GARCH model. Then, through a Monte Carlo experiment, the authors rejected the null hypothesis of constancy in all cases. In this spirit, the analysis of volatility and co-movements among assets have been the central point of a branch of econometrics and different models have been derived, such as VAR and VECM. However, the GARCH family models have been extensively used respect to the other models.

Starting from ARCH and GARCH models the BEKK model of Baba, Engle, Kraft and Kroner (1995) and the Dynamic Conditional Correlation (DCC) of Engle (2002) and Tse and Tsui (2002) have been proposed, just to name a few. They rely on non-linear

combinations of univariate GARCH models and by construction, ensure a more parsimonious parametrization of the model.

In this spirit, the general formulation that we adopt consists in the following equations:

$$Q_t = (I_n - AA' - BB') + AC_{t-1}A' + BQ_{t-1}B'$$

$$R_t = Q_t^{*-1/2}Q_tQ_t^{*-1/2}$$

where A and B are $n \times n$ matrices of parameter, while Q_t is a PSD definite matrix that represent the correlation driving process at time t. Moreover, $Q_t^{*-1/2}$ is a diagonal matrix. It is important to note that the last equation is necessary and required for practical computation, since it allows to transform Q_t in a correlation matrix, since its diagonal elements are not necessarily equal to 1.

Thus, if I consider the previous equations, I have defined the Realized DCC model of Bauwens et al. (2012), where the realized covariance is included into the conditional covariance matrix by extending the classical DCC of Engle (2002).

Moreover this parametrization, together with the n univariate MEM model for modeling volatility, defines a new model: the DCC-CAW-MEM.

The DCC-CAW-MEM combines three different assumptions: the multivariate Wishart distribution, a consistent dynamic conditional correlation estimator and a multiplicative error model for modelling the variance covariance specification.

In the remain of this session I draw the pros and cons of this approach.

My idea was to separately model the element of the variance covariance matrix in a way that permits such kind of flexibility and does not prevent to estimate the model parameters in large dimension.

As evidenced before, despite of using a classical GARCH(1,1) model for modeling volatility, I adopted the high frequency counterpart, by resorting the scalar vector MEM of Cipollini et al. (2012), for which the estimation is made equation by equation and stationarity is subjected to the condition that the characteristic roots of the polynomial are outside the unit circle, hence $\omega + \gamma + \delta < 1$, so that, for the diagonal matrix D, the *ii* element is represented by:

$$S_{ii,t} = \omega_i + \gamma_i RV_{ii,t-1} + \delta_i S_{ii,t-1}$$

Moreover, the model could be easily extended to incorporate other explanatory variables, let's name X_{t-l}, for l=1,2,…,L variables, so that, traditional assuming that $RV_t = S_t \varepsilon_t$, we get:

$$S_t = \omega + \gamma RV_{t-1} + \delta S_{t-1} + \sum_{l=1}^{L} \theta_l X_{t-l}$$

And extra features can be included, without requiring relevant changes in model structuration. The use of MEM model for the univariate modeling of volatility also open to the possibility of adopting other specification of the error term ε_t. Even if in my case, I adopted a Gamma distribution, every other distribution with probability density function (pdf) over a positive interval $[0,+\infty)$, with unit mean and unknown constant variance, such as Log-normal, Weibull or exponential, or mixture of them[7].

As to mention a few, Engle and Gallo (2006) favoured a Gamma for modelling non negative valued time series while Bauwens et al. (2010) adopted a Weibull distribution. Other important characteristic of my model is the dynamic parametrization of the conditional correlation process. As introduce before, I used the Re-DCC od Engle (2012), where the correlation matrix is constructed through a factorization of Qt matrix, but it's useful to note that by imposing some parameters restriction, it's possible to shift to other conditional correlation dynamics. Since I relied on high dimensionality context, and I wanted to obtain both a realistic and parsimonious parametrization, in order to overcome the curse of dimensionality problem, I choose a scalar specification of our Qt matrix.

Thus:

$$Q_t = \bar{S}(1 - \alpha - \beta) + \alpha C_{t-1} + \beta Q_{t-1}$$

Where $\bar{S}$ is the is the sample covariance matrix of standardized residuals obtained from the univariate MEM estimates.

Moreover, the parametrization of out model allowed to get a consistent estimation of all the parameters through a Quasi Maximum Likelihood (QML) approach. Using the probability density function (pdf) of the Wishart distribution and the decomposition of S_t, into the diagonal matrix of standard deviations and the time varying correlation matrix, allowed to define the log likelihood of each observation, called l_t, in the following way:

---

[7] Moreover, Ghysels, Gourieroux and Jasiak (2004) used an exponential model with gamma heterogeneity, in modelling a two-factor model for duration.

$$l_t = -\frac{v}{2}log|S_t| - \frac{v}{2}tr(S_t^{-1}C_t)$$

And if the univariate equations for the conditional variances do not include spillover term and each conditional variance depends only on own lagged variances, thus the log-likelihood function can be generally written, as the sum of the single univariate log-likelihood functions, such that:

$$L = -\frac{v}{2}\sum_{t=1}^{T}[log|S_t| + tr(S_t^{-1}C_t)]$$

The main advantage of this formulation is that the optimization could be achieved through n different optimization procedures, each for one observation in the time series. Additionally, taking advantage of the model structure, when the number of assets rapidly increases, it is possible to use three steps procedures:

1. Step 1 foresees the estimation of the univariate parameters of volatility, through a MEM (1,1) process;
2. Step 2 provides the estimation of the dynamic process of Qt based on the variance covariance matrix of standardized univariate residuals;
3. In step 3, the parameters of the correlation equation will be estimated from the maximization of the single log-likelihood contributions.

It is important to note that QML interpretation of coefficients guarantees a robust statistical inference and the estimation is consistent even if the underlying distribution is not a Wishart, and it allows to easily compute the standard errors[8].

After I have assessed the main properties of the DCC-CAW-MEM, it is useful to made a brief comparison with other models existing in literature. More specifically my model showed some similitudes with the Conditional Autoregressive Range (CARR) model of Chou (2005) and with the models illustrated in Bauwens, Braione and Storti (hereafter, BBS) (2016). With the first, my model shared the choice of the univariate specification for volatility estimator, even though the one used in Chou is strictly related to the ACD model of Engle and Russel (1998).

---

[8] See Noureldin et al. (2012)

At the same time, Chou used a one-step estimation process, but this is fitted on range volatility measure and it takes the following form:

$$L(\alpha_i, \beta_j, R_1, R_2, \ldots, R_t) = \sum_{t=1}^{T} [\log(\lambda_t) + R_t/\lambda_t]$$

With the model class of BBS, my model shared the dynamic structure of correlations and the one step estimation likelihood-based approach.

Finally, it is useful to remember that my model, in the formulation adopted in this work, does not take into account spillover effect and in the univariate volatility model, I do not adopt a full representation for parameter matrices A and B. In fact, even if it is possible to adopt a full A matrix, in order to have some preliminarily information about correlation signs among assets, I will expect a lot of zero coefficients, because off the high dimensional framework of analysis.

Furthermore, in my specification, I adopted a classical DCC (1,1) model, even though it should be extended, by introducing a parameter g for asymmetry, in the spirit of the AG-DCC model of Cappiello et al. (2006), or as in Silvennoinen and Terasvirta (2012), through a smooth transition specification in modeling conditional correlations of asset returns, that allowed the conditional correlations varying from one state to another as a function of a transition variable .

2.3 Alternative parametrizations: BEKK dynamics

As introduced in the general framework of Section 2.1, given C_t the series of realized variance covariance matrix, for which:

$$E(C_t|F_{t-1}) = E_{t+1}(C_t) = S_t$$

It is possible to nest an alternative and competitive parametrization, without passing through the factorization of S_t. I will refer to this model as ReBEKK.

Instead of using a DCC-GARCH type recursion, it is possible to adopt a BEKK-type framework. Without using the classical BEKK model of Baba, Engle, Kraft and Kroner (1995), I will refer to the multiplicative BEKK model of Hafner and Lindton (2010) that, through covariance targeting, introduced a first order scalar BEKK,

directly fitted to the series of realized variance covariance matrix, with the following form:

$$S_t = \bar{S}(1 - \alpha - \beta) + \alpha C_{t-1} + \beta S_{t-1}$$

Where α and β are positive scalars and $\alpha + \beta < 1$.

Moreover, since I wanted to increase comparability of my model, it was possible to add another parametrization, based on the decomposition of S_t into a long run covariance matrix and a short run component, respectively $L = MM'$ and S_t^*, such that:

$$S_t = M_t S_t^* M_t'$$

In this case, M_t is obtained through a Cholesky factorization of long run matrix L. The latter is estimated by the non-parametric smoother of Nadaraya-Watson, used in the estimation process for all the element of realized variance covariance matrices. For the short run component, I restored the multiplicative scalar BEKK, where, differently as in the first parametrization, I imposed the mean reversion condition in its dynamics, through I_n, so that:

$$S_t^* = (1 - \alpha - \beta)I_n + \alpha C_{t-1}^* + \beta S_{t-1}^*$$

where $C_t^* = M_t^{-1} C_t (M_t^{-1})'$ and it is a variance covariance matrix purged of its long run component.

I will refer to this model as Np-ReBEKK.

For both BEKK parametrization, the estimation process was carried out through a one-step loglikehood-based approach, previously used:

$$L = \sum_{t=1}^{T} [log|S_t| + tr(S_t^{-1} C_t)]$$

2.4 Covariance Targeting

Before turning into the presentation of the other model used in this work, I have to discuss one another feature of the new parametrization offered by DCC-CAW-Mem, that is also at the basis of the estimation procedure of one of the BEKK parametrization of Section 2.3: the covariance targeting.

To understand its usefulness, consider that, even if I imposed a scalar specification of my model, the number of parameters to be estimated was proportional to the number of assets, so that the estimation became cumbersome, if a high number of assets is included in the model design. Hence, in order to exploit the advantages of working with a great number of assets with intraday data, I needed a quick and simple way to guarantee a feasible estimation in the maximum likelihood framework. For this reason, both in DCC-CAW-MEM and in BEKK parametrizations, I resorted to targeting approach, that consists in pre-estimating the constant intercept matrix in the model specification.

More specifically, in the DCC-CAW-MEM

$$Q_t = \bar{S}(1 - \alpha - \beta) + \alpha C_{t-1} + \beta Q_{t-1}$$
$$R_t = Q_t^{*-1/2} Q_t Q_t^{*-1/2}$$

Targeting is imposed in Q_t structure ensures that the conditional correlation will be identical to the sample unconditional correlation of the data. This is strictly achieved by computing the sample covariance matrix of standardized residuals ε_t that come from the n univariate MEM (1,1) estimation, so that:

$$\bar{S} = \hat{E}[\varepsilon_{t-1}\varepsilon'_{t-1}] = T^{-1} \sum_{t=!}^{T} \varepsilon_{t-1}\varepsilon'_{t-1}$$

Moreover, also the first order scalar BEKK model is based on the covariance targeting, since its recursion has the following form:

$$S_t = \bar{S}(1 - \alpha - \beta) + \alpha C_{t-1} + \beta S_{t-1}$$

Where $\bar{S}$ is the unconditional long run variance covariance matrix. In this case, targeting implies the use of a sample covariance estimator for S and the maximization of the likelihood function with respect to the parameters α and β (maximization is made conditionally on the estimates of the long run covariance)[9].

---

[9] For further details, see Pedersen, R. S., & Rahbek, A. (2014)

2.5 Additional models

For the completeness of my analysis, I included one additional model for realized volatility, that combines two different specifications: the Heterogeneous Autoregressive model for the Realized Variance (HAR-RV) of Corsi (2009) and the dynamic structure of a DCC of Engle (2002).

In the original formulation of 2003, Corsi allowed for heterogeneity in the traders' time horizons, by introducing an additive cascade of partial volatility at the different time-points: daily, weekly and monthly, even if more other component could be easily added.

Despite this model formally does not belong to the category of long memory models, Andersen, Bollerslev and Diebold (2007) and Wang, Bauwens and Hsiao (2013) have found that it provides a very good fit for most volatility series. Thus, the HAR-RV model could be considered a genuine long memory model, and in contrast to the other long memory model, it could be easily estimated through OLS procedures. Moreover, instead of considering the DCC of Engle separately from the HAR-RV, I decided to combine them, by using a two-steps estimation log-likelihood-based approach.

In particular, the estimation required:

1. in step 1, a multivariate DCC model was directly run on realized kernel time series, in order to obtain the correlation coefficient among series;

2. in step 2, the standardized residuals ε_t of DCC model were used to reconstruct the volatility series X_t that constitute the inputs of the OLS model, that allowed to find the HAR parameters:

$$X_{t+1} = \beta_0 + \beta_d X_t + \beta_w X_{t-5,t} + \beta_m X_{t-22,t} + \zeta_{t+1}$$

With ζ is a random error.

Finally, as baseline model, the vMEM of Cipollini et al. (2012) has been estimated, through an equation by equation estimation approach.

As for the univariate MEM model examined in the Section 2.1, the basic parametrization is the following:

$$RV_t = S_t \odot \varepsilon_t$$

$$with\ \varepsilon_t|F_{t-1}\sim Gamma\ (\vartheta,\frac{1}{\vartheta})$$

where $\odot$ is the Hadamard product and also in this case $S_t = S(F_{t-1}, \theta)$. The random variable is defined as $k*1$ vector $\varepsilon_t|F_{t-1} \backsim D(1, \Sigma)$. For the previous conditions:

$$E(RV_t|F_{t-1}) = S_t$$
$$Var\ (RV_t|F_{t-1}) = S_t S_t' \odot \Sigma$$

As for the univariate framework, in the multivariate context the conditional expectation is defined as:

$$S_t = \omega + \sum_{l=1}^{L} [\gamma_l RV_{t-l} + \delta_l S_{t-l}]$$

In my study, in order to estimate this recursive form $\omega, \gamma\ and\ \delta$ are scalar parameters. Also in this case, during the estimation procedure, the usual constraints for positiveness and stationarity are imposed $\omega >0, \gamma > 0, \delta > 0$ and $(\omega + \gamma + \delta) < 1$.

2.6 Estimation Procedures Concluding Remarks

All the parametric models have been estimated through a maximum likelihood-based approach: excluding the HAR-DCC model and the vMEM, the estimation is done by a unique likelihood function, based on the Wishart distribution made on the series of realized variance covariance matrices Ct.

The log-likelihood function has been defined as:

$$-\frac{v}{2} \sum_{t=1}^{T} [log|S_t| + tr(S_t^{-1} C_t)]$$

and, as it was shown, it is very general, in the sense that it can be used, in case of absence of microstructure noise, whatever will be the formulation for volatility and correlation dynamics. For those models, I decided to use a one-step approach in order to simplify the inference and to easily compute standard error s and model selection criteria. Moreover, in the case of non-parametric formulation, as for Np-ReBEKK model, the estimation requires two steps: the computation of the kernel smoother for accounting the long run co-movements in variance covariance matrices, and then the short run component is estimated through the previously defined ML approach. In

addition, the HAR-DCC model is the only one model that has been estimated through a two steps procedure, by firstly estimating the DCC parameters, followed by the estimation of HAR coefficient, through the standardized residuals. Finally, the vMEM has been estimated by an equation by equation approach, even if, instead of using a GMM estimator suggested by Cipollini (2012), I followed a ML approach, by adapting the univariate estimation procedure of Brownlees et al (2011).

Chapter 3

Evidence from Real Data

Abstract:

This chapter presents the empirical analysis conducted on a twenty realized kernel financial time series, and the corresponding daily series of realized variance covariance matrices.

Firstly, I will present the dataset being used and the cleaning procedures adopted to obtain it. Then, after providing the full sample estimation results by comparing the models in term of information criteria, I finally presented the theoretically framework for financial one-step ahead forecasts and their comparisons.

3.1 Dataset

In carrying out my empirical analysis, I considered a dataset of annualized realized kernel of volatility for twenty (20) financial indices, provided by the Oxford Man Institute[10], from January 3, 2000 to December 5, 2017.

As stated by Bandorff, Nielsen, Hense, Lunde and Sheppard (2002), realized kernel is a consistent estimator of realized volatility, and it is also robust respect to microstructure noise of the markets.

The included stocks are: AEX Index, All Ordinaries , BVSP BOVESPA Index , Down Jones Index , CAC 40 , FTSE 100 , FTSE MIB , FSTE STRAIT TIME Index, DAX, Hang Seng Index , IBEX 35, Nasdaq 100, Korea Composite Stock Price Index, IPC Mexico, Nikkei 225, NIFTY 50, Russel 2000, SP 500 Index, Swiss Stock Market Index and EUROSTOXX 50.

The original database contained 21 assets, but from a preliminarily analysis, it resulted that the percentage of missing values ranged between 2.45% of CAC40 and 16,78% of S&P/TSX Composite index (GSPTSE).

Moreover, from a more careful examination, it has been possible to note that for GSPTSE index, there were two years of missing values (the series starts on May 02,2002), so that I decided to drop it from the dataset. Furthermore, since I relied on standard measures and not on heavy data, in order to get a complete dataset, I deleted overnight returns (favoring the estimation accuracy). Then I choose, as treatment of missing values, a linear interpolation[11] and, as a special case, in presence of a missing value at the first observation, I replaced it using the forward value.[12]

Additionally, for the estimation process, I had to preliminarily compute the series of realized variance covariance matrices. The daily realized covariance matrices have been computed using the CholCov estimator presented in Section 1.3.8

For a simpler read of the results, the following table reports abbreviations and the full name of the assets included in the study.

---

[10] The database used in this thesis refers to the 0.2 version, that was retired and now at the following link *https://realized.oxford-man.ox.ac.uk/data/download* it's possible to download a more recent and bigger (31 assets) library, that rely on a different cleaning and econometric procedures.
[11] See Lu, Z. & Hui, Y.V. Ann Inst Stat Math (2003) 55: 197. https://doi.org/10.1007/BF02530494
[12] No additive cleaning procedures are used.

Symbol	Asset Name
AEX	AEX Index
AORD	All Ordinaries
BVSP	BVSP BOVESPA Index
DJI	Down Jones Index
FCHI	CAC 40
FTSE	FTSE 100
FTSEMIB	FTSE MIB
FTSTI	FSTE STRAIT TIME Index
GDAXI	DAX
HSI	Hang Seng Index
IBEX	IBEX 35
IXIC	Nasdaq 100
KS11	Korea Composite Stock Price Index (KOSPI)
MXX	IPC Mexico
N225	Nikkei 225
NSEI	NIFTY 50
RUT	Russel 2000
SPX	SP 500 Index
SSMI	Swiss Stock Market Index
STOXX	EUROSTOXX 50

Table 1: Data symbols. The table shows the list of symbols (column 1) used to identify the corresponding financial time series indices (column 2).

Given the framework of analysis presented in the previous chapter, the aim of this empirical analysis is to evaluate both the in-sample and the out-of- sample performance of the models presented before.

Just for a reminder, we report a list of all the models presented in the previous sections.

	Model	N. parameters	Parametric	Semiparametric
1)	**vMEM**	(4*n)	✓	
2)	**DCC-CAW-MEM**	(4*n)+2	✓	
3)	**ReBEKK**	[n(n+1)/2]+2	✓	
4)	**Np-ReBEKK**	3		✓
5)	**HAR-DCC**	(4*n)+2	✓	

Table 2: List of analyzed models. This table lists the estimated models with each own number of parameters, as well as the specific parametrization.

Preliminarily I will report the results of the full sample estimation and, afterward, I will focus on the one step ahead out of sample forecasts. For the latter, the full dataset, from January 3, 2000 to December 5, 2017 is divided into two different periods:

- Period I is the in-sample set for t=1, …, 4161. It corresponds to the period January 03, 2000 to December 04, 2015;
- Period II is the out of sample set for the remaining 522 observations in the dataset, from December 05, 2015 to December 05, 2017, that represents the last observation included in the sample.

Before presenting the full sample results, I will summarize the descriptive statistics for all the time series, distinguishing the full sample, the in sample and the out of sample sets of analysis. On the basis of the kurtosis, it is possible to note that time series properly exhibit fat-tail properties and there is also a departure from the Gaussian distribution, since the skewness is different from 0 and there is positive excess kurtosis.

Symbol	Mean	Median	Min	Max	Std	Skewness	Kurtosis
ESTIMATION SAMPLE: JANUARY 03, 2000 TO DECEMBER 05, 2015 (4161 OBSERVATIONS)							
AEX	15,35	12,67	2,80	99,92	9,53	2,46	12,64
AORD	9,59	8,11	1,72	62,62	5,61	2,73	15,58
BVSP	22,41	20,12	4,37	145,18	10,86	3,44	24,15
DJI	13,98	11,65	3,09	151,65	9,20	3,69	29,15
FCHI	16,83	14,68	3,18	107,09	9,56	2,44	14,11
FTSE	12,60	10,41	3,15	90,59	7,86	2,55	14,39
FTSEMIB	15,68	13,57	2,83	104,15	9,02	2,24	12,90
FTSTE	10,65	9,33	3,84	83,59	5,21	3,03	22,12
GDAXI	18,35	15,47	3,38	127,31	11,12	2,63	15,54
HSI	14,01	11,94	3,29	106,03	8,13	3,73	28,21
IBEX	17,22	15,92	3,23	95,11	9,02	1,95	11,43
IXIC	15,76	12,50	3,19	129,64	10,53	2,52	14,17
KS11	16,12	13,78	3,39	127,78	9,79	2,38	14,63
MXX	11,42	9,76	2,55	87,10	6,17	3,00	20,30
N225	15,03	13,41	4,00	98,32	8,02	3,17	22,08
NSEI	15,91	13,18	2,54	159,60	10,59	3,60	27,94
RUT	14,19	12,00	2,54	127,24	8,87	3,14	20,89
SPX	14,22	11,79	2,43	153,19	9,46	3,34	25,11
SSMI	12,67	10,36	4,17	83,95	7,44	2,79	15,12
STOXX	17,65	15,07	0,28	165,90	10,56	2,92	21,46
FORECASTING SAMPLE: DECEMBER 06, 2015 TO DECEMBER 05, 2017 (522 OBSERVATIONS)							
AEX	10,83	9,34	2,98	55,99	5,82	2,25	12,45
AORD	8,81	7,79	2,95	39,58	3,97	2,51	14,36
BVSP	14,03	13,04	4,32	66,00	5,72	2,54	17,99
DJI	7,65	6,21	2,11	36,18	4,71	2,46	10,90
FCHI	11,92	10,61	3,18	55,75	5,80	2,21	12,50
FTSE	8,01	7,00	3,38	28,50	3,39	1,89	8,35
FTSEMIB	12,79	11,27	3,68	46,82	6,10	1,78	7,57
FTSTE	10,09	9,50	4,10	38,97	3,85	1,83	21,73
GDAXI	11,98	10,75	2,92	46,90	5,99	1,66	7,45
HSI	9,34	8,48	4,26	42,11	3,40	3,42	27,35
IBEX	14,61	12,71	4,79	109,59	7,61	4,86	51,80
IXIC	8,28	7,11	2,49	32,35	4,34	2,21	9,25
KS11	7,10	6,65	3,05	43,57	2,79	5,88	65,63
MXX	10,19	9,36	5,35	54,94	4,14	5,31	48,54
N225	11,61	9,06	3,11	74,59	8,30	3,05	16,49
NSEI	8,52	7,73	3,93	40,82	3,61	3,47	24,67
RUT	9,28	8,48	3,83	31,44	3,92	1,94	8,31
SPX	7,41	6,06	1,88	29,52	4,51	2,26	9,16
SSMI	9,73	8,60	3,89	37,86	4,08	2,16	10,46
STOXX	12,70	11,06	2,58	118,13	7,83	5,45	65,96
FULL SAMPLE: JANUARY 03 2000 TO DECEMBER 05,2017 (4683 OBSERVATIONS)							

AEX	14,84	12,28	2,80	99,92	9,30	2,50	13,07
AORD	9,50	8,06	1,72	62,62	5,46	2,77	16,04
BVSP	21,47	19,37	4,32	145,18	10,75	3,34	23,72
DJI	13,27	10,98	2,11	151,65	9,04	3,63	29,03
FCHI	16,29	14,13	3,18	107,09	9,34	2,47	14,45
FTSE	12,09	9,86	3,15	90,59	7,63	2,63	15,11
FTSEMIB	15,36	13,33	2,83	104,15	8,79	2,27	13,18
FTSTE	10,59	9,97	3,84	83,59	4,91	3,24	25,00
GDAXI	17,64	14,88	2,92	127,31	10,86	2,66	15,95
HSI	13,49	11,45	3,29	106,03	7,88	3,80	29,39
IBEX	16,93	15,61	3,23	109,59	8,91	2,16	13,65
IXIC	14,93	11,70	2,49	129,64	10,30	2,56	14,60
KS11	15,11	12,64	3,05	127,78	9,70	2,39	14,54
MXX	11,28	9,67	2,55	87,10	5,99	3,15	21,68
N225	14,65	13,00	3,11	98,32	8,12	3,06	20,72
NSEI	15,09	12,38	2,54	159,60	10,32	3,64	28,71
RUT	13,64	11,50	2,54	127,24	8,60	3,21	21,83
SPX	13,46	11,09	1,88	153,19	9,29	3,30	25,08
SSMI	12,34	10,10	3,89	83,95	7,20	2,86	15,88
STOXX	17,10	14,57	0,28	165,90	10,41	3,03	22,77

Table 3: Descriptive statistics of realized kernel financial time series. The table reports descriptive statistics of the realized kernel time series of the stocks used in the empirical applications. The three panel report the statistics for the in-sample, the out-of-sample and the dull period, respectively.

Moreover, Figure 1 and Figure 2 show a representative example of time series plots of the realized daily log returns and the annualized realized kernel volatility, respectively, for two representatives series included in the sample: the AORD and STOXX50 indices. The choice was due to their range of variability, respectively the lowest and the highest (for the full sample set of 4683 observations). The plot of all the financial time series is reported in the Appendix A, which also reports the correlation matrix among time series. Even if the correlation coefficients were high, the statistical significance of correlations is assessed only for KS11, MXX and NSEI indices.
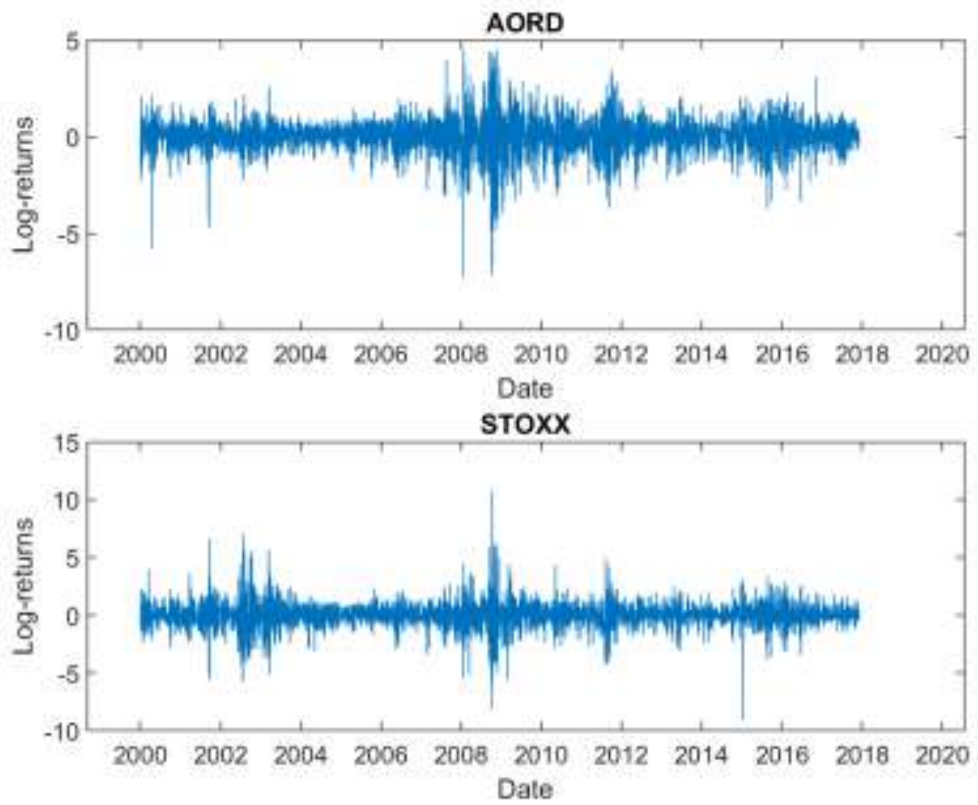
Figure 1. Daily Log-returns of AORD and STOXX assets (full sample).
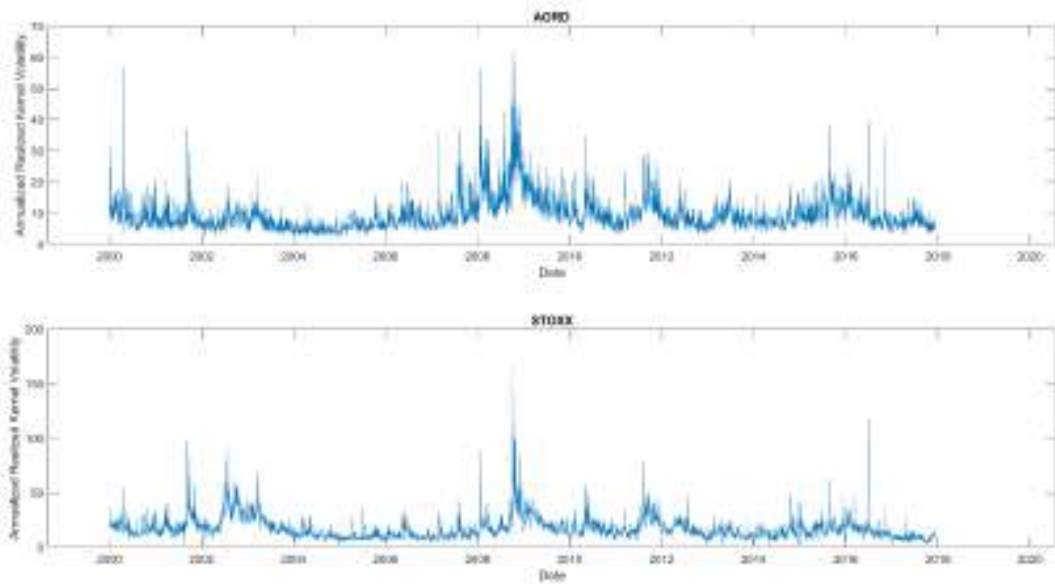

Figure 2. Annualized Realized Kernel of AORD and STOXX assets (full sample).

A simple visual inspection for all the assets showed volatility clustering and persistence. All the estimation was made through MATLAB software. Even if there are some specific toolboxes that have been designed for modelling financial time series, both in univariate and in multivariate setting (such as MFE toolbox of Sheppard), most of the available functions refers to traditional GARCH models and were not suitable for my case, thus every function used in this thesis derived from a personal coding effort. Additionally, since the estimation requires numerical procedures to maximize the maximum likelihood functions, it is important to be aware they can lead to multiple local maxima and also different algorithms could lead to different parameter estimates and, consequently to different inferential results. Moreover, numerical procedures are strictly related to the choice of initial parameters, so that, in order to bypass the related parameters estimation uncertainty, all the results have been confirmed through the Globalsearch and Multistart Matlab's algorithms. Furthermore, all the models have been estimated used the interior point algorithm and, for increasing the computational speed, the Hessian has not been estimated at each interaction, but standard errors have been computed with the robust sandwich covariance estimator, computed at the final parameters estimates.

3.2 Full sample results

To evaluate the relative performance of my models in fitting the data, I will present, in this section, the full sample estimation results for each model.

For the baseline vMEM model, results are reported below. Table 4 shows all the parameters estimated, equation by equation.

	ω		γ		δ		ϑ	
AEX	0,0462		0,2922	***	0,6516	***	5,5979	***
	(0,12)		(0,05)		(0,09)		(1,87)	
AORD	0,0393	***	0,2908	*	0,6599	***	6,5092	***
	(0,01)		(0,15)		(0,19)		(1,87)	
BVSP	0,0598		0,2814		0,6582		4,0866	
	(1,57)		(2,15)		(0,70)		(7,34)	
DJI	0,0500		0,2989	***	0,6411	***	5,6263	***
	(0,04)		(0,08)		(0,04)		(0,11)	
FCHI	0,0740		0,4003	***	0,5156	***	4,4299	***
	(0,17)		(0,08)		(0,09)		(1,07)	
FTSE	0,0533		0,2538		0,6829		5,0026	***
	(0,32)		(1,11)		(0,84)		(0,81)	
FTSEMIB	0,0787	***	0,3483	***	0,5628	***	3,7030	***
	(0,02)		(0,04)		(0,03)		(0,04)	
FTSTI	0,0758		0,3498	***	0,5645	**	6,5017	***
	(0,28)		(0,03)		(0,28)		(1,68)	
GDAXI	0,0706		0,2873		0,6414		3,4880	***
	(0,23)		(0,85)		(0,64)		(0,01)	
HSI	0,0629		0,3056		0,6210	*	4,8165	***
	(0,40)		(0,74)		(0,38)		(0,85)	
IBEX	0,0652		0,3586	***	0,5758	***	5,5601	***
	(0,15)		(0,09)		(0,05)		(0,21)	
IXIC	0,0836	***	0,3450	***	0,5612	***	3,7332	***
	(0,02)		(0,04)		(0,03)		(0,18)	
KS11	0,0919		0,3258	***	0,5721	***	3,2768	***
	(0,08)		(0,09)		(0,02)		(0,07)	
MXX	0,0536	***	0,3540	***	0,5824	***	6,4319	***
	(0,02)		(0,00)		(0,01)		(0,57)	
N225	0,0530		0,3137		0,6231	**	5,1084	***
	(0,04)		(0,38)		(0,31)'		(1,04)	
NSEI	0,0571	***	0,4625	***	0,4704	***	8,4672	***
	(0,01)		(0,13)		(0,16)		(0,83)	
RUT	0,0638		0,3195		0,6066		4,6912	***
	(0,34)		(0,46)		(0,82)		(0,96)	
SPX	0,0597		0,3469	**	0,5835	***	5,3043	***
	(0,12)		(0,16)		0(0,07)		(0,97)	
SSMI	0,0410	*	0,2667	***	0,6823	***	6,0040	***
	(0,05)		(0,15)		(0,08)		(2,10)	
STOXX	0,0699	*	0,4007	***	0,5194	***	4,1403	***
	(0,04)		(0,03)		(0,01)		(0,17)	

Table 4: vMEM parameters' estimate. Coefficients (standard error in parenthesis) of the vMEM estimates. Stars are put in correspondence of small p-values: *(p-value<10%); **(p-value< 5%); ***(p-value< 1%). Full period: January 03, 2000 to December,05 2017. Number of observations: 4683.

It might be noticed that the estimated coefficient values are within a similar range of variability, providing evidence that volatility dynamics have a certain degree of homogeneity. Moreover, for all the assets, the γ_i coefficients usually vary between 0.25 and 0.35, which makes the dynamics more reactive to past realizations of the process than what is typically implied by GARCH (1,1) approach[13]. As the same time δ_i coefficients range between 0.50 and 0.70. Furthermore, vMEM evidences a high level of persistence, calculated as $(\gamma + \delta)$, ranging between 0.8978 (KS11) and 0.95 (AORD). Excluding BVSP, FTSE, GDAXI and RUT, for which all coefficient values are not significant, the most of all the other assets coefficient, the parameters are significant at 1% level. Additionally, Table 5 reports the Ljung Box p-values up to lag 22, for each estimation process. As it can be easily noticed, autocorrelation still remains a problem for the baseline vMEM model.

---

[13] Even if we don't report the GARCH(1,1) estimates, the results are available upon request. Moreover, for GARCH processes γ_i are between 0.07and 0.10, while δ_i parameters are greater than those estimated by vMEM (they are in mean around 0.89)

	Ljung-Box Lag 1	Ljung-Box Lag 5	Ljung-Box Lag 10	Ljung-Box Lag 22
AEX	0,000	0,000	0,000	0,000
AORD	0,014	0,000	0,000	0,000
BVSP	0,000	0,000	0,000	0,000
DJI	0,000	0,000	0,000	0,000
FCHI	0,221	0,000	0,000	0,000
FTSE	0,000	0,000	0,000	0,000
FTSEMIB	0,001	0,000	0,000	0,000
FTSTI	0,061	0,000	0,000	0,000
GDAXI	0,000	0,000	0,000	0,000
HSI	0,008	0,000	0,000	0,000
IBEX	0,000	0,000	0,000	0,000
IXIC	0,000	0,000	0,000	0,000
KS11	0,000	0,000	0,000	0,000
MXX	0,256	0,000	0,000	0,000
N225	0,000	0,000	0,000	0,000
NSEI	0,132	0,000	0,000	0,000
RUT	0,000	0,000	0,000	0,000
SPX	0,000	0,000	0,000	0,000
SSMI	0,000	0,000	0,000	0,000
STOXX	0,506	0,000	0,000	0,000

Table 5: Ljung Box. Ljung Box p-values up to lag 22. Full period: January 03, 2000 to December,05 2017. Number of observations: 4683.

Among the models presented in this chapter, the DCC-CAW-MEM represents the new parametrization that I introduced in the financial econometric literature.

As mentioned before, this model combines three different elements: the high frequency data, through the multiplicative error model (MEM), the Wishart distribution (as a generalization of the multivariate gamma density function) and the Re-DCC of Engle (2012) for the dynamics of conditional volatility in the decomposition of variance-covariance matrix.

The results of the estimation are reported in Table 6.

	ω		γ		δ		ϑ	
AEX	0,0801	***	0,1747	***	0,6442	***	5,8960	***
	(0,000)		(0,000)		(0,000)		(0,000)	
AORD	0,1515	***	0,3681	***	0,4067	***	5,9106	***
	(0,0000)		(0,0000)		(0,0000)		(0,0010)	
BVSP	0,0943	***	0,2258	***	0,5668	***	5,9804	***
	(0,0000)		(0,0000)		(0,0000)		(0,0001)	
DJI	0,0564	***	0,3261	***	0,4352	***	6,0466	***
	(0,0000)		(0,0000)		(0,0000)		(0,0002)	
FCHI	0,0967	***	0,2167	***	0,5723	***	6,0173	***
	(0,0000)		(0,0000)		(0,0000)		(0,0000)	
FTSE	0,1891	***	0,2108	***	0,5798	***	6,0534	***
	(0,0000)		(0,0000)		(0,0000)		(0,0000)	
FTSEMIB	0,0797	***	0,1869	***	0,6269	***	5,9527	***
	(0,0000)		(0,0000)		(0,0000)		(0,0005)	
FTSTI	0,1281	***	0,0003	***	0,6920	***	5,8088	***
	(0,0000)		(0,0000)		(0,0000)		(0,0103)	
GDAXI	0,0876	***	0,1748	***	0,6333	***	6,0341	***
	(0,0000)		(0,0000)		(0,0000)		(0,0000)	
HSI	0,0868	***	0,1971	***	0,6101	***	5,9976	***
	(0,0000)		(0,0000)		(0,0000)		(0,0000)	
IBEX	0,2034	***	0,2142	***	0,5551	***	5,9211	***
	(0,0000)		(0,0000)		(0,0000)		(0,0005)	
IXIC	0,0875	***	0,0002	***	0,7920	***	7,5593	***
	(0,0000)		(0,0000)		(0,0000)		(0,0000)	
KS11	0,1161	***	0,0000	***	0,6930	***	5,8022	***
	(0,0000)		(0,0000)		(0,0000)		(0,0003)	
MXX	0,0647	***	0,3001	***	0,4837	***	5,8916	***
	(0,0000)		(0,0000)		(0,0000)		(0,0000)	
N225	0,0872	***	0,1919	***	0,6139	***	5,9699	***
	(0,0000)		(0,0000)		(0,0000)		(0,0000)	
NSEI	0,0589	***	0,3081	***	0,5100	***	5,9414	***
	(0,0000)		(0,0000)		(0,0000)		(0,0000)	
RUT	0,1174	***	0,2293	***	0,5732	***	5,9191	***
	(0,0000)		(0,0000)		(0,0000)		(0,0002)	
SPX	0,1924	***	0,2919	***	0,4895	***	6,0447	***
	(0,0000)		(0,0000)		(0,0000)		(0,0000)	
SSMI	0,1840	***	0,1965	***	0,5913	***	6,0838	***
	(0,0000)		(0,0000)		(0,0000)		(0,0000)	
STOXX	0,0000	***	0,3019	***	0,4444	***	6,2602	***
	(0,0000)		(0,0000)		(0,0000)		(0,0000)	

	α		β	
	0,0287	***	0,7253	***
	(0,0000)		(0,0000)	

Table 6: DCC-CAW-MEM parameters' estimates. Coefficients (standard error in parenthesis) of the DCC-CAW-MEM estimates. Stars are put in correspondence of small p-values: *(p-value<10%); **(p-value< 5%); ***(p-value< 1%). Full period: January 03, 2000 to December,05 2017. Number of observations: 4683.

It is interesting to highlight that in comparison to the previous parametrization, there are some changes in the parameter values: both γ_i and δ_i values, in general, decrease, but there is a market gain in significance for all the coefficients estimated (they are all significant at 1% level).

As can be easily noticed, the mean level of persistence $(\gamma + \delta)$ decreases and now it ranges between 0.70 (FTSTI) and 0.82 (AEX). Moreover, the estimation results into two additional parameters, referring to the DCC-type dynamic for conditional correlation, that are, respectively, 0.287 and 0.7253. Both are statistically significant at 1% level.

Others three models have been estimated.

As discussed before and as synthesized in Table 3, two different BEKK parametrizations have been adopted. The first one, the scalar ReBEKK, directly nested a BEKK-type recursion for modelling the series of realized variance covariance matrices S_t, while the second formulation, the NpReBEKK split the estimation of this matrix into the estimation of the short run S_t^* matrix, modelled with a BEKK recursion, and the long run component L, through a one side smoother function.

In the first case, the estimation results in 212 parameters, while for the second specification, even though I really do not know how many parameters the estimation procedure involves in the long run component, the estimation finally results in 3 parameters: the smoother kernel and the usually correlation parameters. In the following Table 7 I briefly report only the main parameters associated to the models and their relative standard errors, but obviously the full parameter estimates are available upon request.

	Kernel	Correlation	
		α	β
ReBEKK		0.1344***	0.8154***
		(0.0000)	(0.0000)
NpBEKK	0.0350	0.1059***	0.7765***
		(0.0000)	(0.0000)

Table 7: BEKK parametrizations' estimates. Coefficients (standard error in parenthesis) of the BEKK model. For realized BEKK we only report the correlation coefficients. For the Np-ReBEKK I report the one side Nadaraya-Watson kernel estimator and the correlation parameters. Stars are put in correspondence of small p-values: *(p-value<10%); **(p-value< 5%); ***(p-value< 1%). Full period: January 03, 2000 to December,05 2017. Number of observations: 4683.

Finally, the other parametrization exploited in this thesis is the HAR-DCC model. As stated in Section 2.5, differently from the other parametrizations, for this model we decided to use a two-step estimation procedure. Thus, from the first step, the correlation parameters are estimated though a maximum likelihood approach[14] , while the HAR parameters are consequently fitted on the standardized residuals obtained in the first step. Parameters are reported in Table 8. As it can be seen, all the parameter estimates are significant at 1% level.

---

[14] For the details of the estimation we refer to Engle (2002).

	β_0		β_d		β_w		β_m	
AEX	1,9783	***	0,0619	***	0,4163	***	0,4583	***
	(0,0005)		(0,0005)		(0,0006)		(0,0006)	
AORD	2,3926	***	0,1748	***	0,2051	***	0,5781	***
	(0,0003)		(0,0003)		(0,0003)		(0,0003)	
BVSP	1,0603	***	0,1441	***	0,1424	***	0,6796	***
	(0,0004)		(0,0004)		(0,0005)		(0,0005)	
DJI	5,6878	***	0,1757	***	0,2888	***	0,4335	***
	(0,0004)		(0,0004)		(0,0005)		(0,0006)	
FCHI	1,3927	***	-0,0452	***	0,3834	***	0,5707	***
	(0,0013)		(0,0013)		(0,0018)		(0,0020)	
FTSE	1,0387	***	-0,0259	***	0,3905	***	0,5494	***
	(0,0019)		(0,0019)		(0,0027)		(0,0024)	
FTSEMIB	0,9647	***	0,0585	***	0,2461	***	0,5960	***
	(0,0022)		(0,0022)		(0,0028)		(0,0032)	
FTSTI	0,6903	***	0,0791	***	0,0616	***	0,8227	***
	(0,0012)		(0,0012)		(00,0014)		(0,0016)	
GDAXI	0,9656	***	-0,0405	***	1,2419	***	-0,2466	***
	(0,0003)		(0,0003)		(0,0003)		(0,0003)	
HSI	2,4508	***	0,1932	***	0,3937	***	0,3876	***
	(0,0001)		(0,0001)		(0,0001)		(0,0001)	
IBEX	0,9161	***	0,1208	***	0,4199	***	0,4368	***
	(0,0004)		(0,0004)		(0,0004)		(0,0004)	
IXIC	1,5946	***	0,1585	***	0,3785	***	0,4088	***
	(0,0005)		(0,0005)		(0,0006)		(0,007)	
KS11	0,8653	***	0,2828	***	0,2352	***	0,4601	***
	(0,0002)		(0,0002)		(0,0002)		(0,0002)	
MXX	1,5335	***	0,2463	***	0,2007	***	0,5169	***
	(0,0004)		(0,0004)		(0,0004)		(0,0005)	
N225	1,0138	***	0,1318	***	0,3650	***	0,4321	***
	(0,0013)		(0,0013)		(0,0017)		(0,0019)	
NSEI	0,6327	***	0,0908	***	0,5149	***	0,3459	***
	(0,0013)		(0,0013)		(0,0015)		(0,0016)	
RUT	1,1981	***	0,3621	***	0,0316	***	0,5359	***
	(0,0008)		(0,0008)		(0,0009)		(0,0011)	
SPX	1,1632	***	0,2647	***	0,2332	***	0,4675	***
	(0,0004)		(0,0004)		(0,0005)		(0,0005)	
SSMI	1,6238	***	0,1366	***	0,6318	***	0,1505	***
	(0,0005)		(0,0005)		(0,0006)		(0,0006)	
STOXX	1,0054	***	0,0467	***	0,3010	***	0,4986	***
	(0,0025)		(0,0025)		(0,0036)		(0,0043)	

	α	β
	0,0169***	0,9437***
	(0,0000)	(0,0000)

Table 8: HAR-DCC parameters' estimates. Coefficients (standard error in parenthesis) of the HAR-DCC estimates. Stars are put in correspondence of small p-values: *(p-value<10%); **(p-value< 5%); ***(p-value< 1%). Full period: January 03, 2000 to December,05 2017. Number of observations: 4683.

Furthermore, Table 9 reports the value of log-likelihoods for all the models whose estimation is based on the general log-likelihood function $\sum_{t=1}^{T}[log|S_t| + tr(S_t^{-1}C_t)]$ enounced in Sections 2.3 and 2.6, that are the DCC-CAW-MEM, the ReBEKK and Np-ReBEKK.

For the fully parametrical models I also report the AIC and the BIC values, that are not shown for the only one non parametric specification. This is due to the fact that in the information criteria computations, the penalty term is strictly related to the whole number of parameters estimated, but I do not know how much parameters the estimation of long run covariance matrix L has involved.

Models	Num. Pars.	LogLik	AIC	BIC
1)DCC-CAW-MEM	82	-42927	86018	86547
2)ReBEKK	212	-43421	87265	88633
3)Np-ReBEKK	-	-4026,2	-	-

Table 9: Log-likelihood and information criteria of multiplicative models. Only multiplicative models (parametric and non-parametric) are reported here. Maximized log-likelihood values are reported in the second column. Information criteria are computed as follows: AIC= -2*LogLik+2*Np and BIC= -2*LogLik+Np*log(T). LogLik is the loglikelihood value obtained through numerical optimization procedures, Np is the number of parameters and T is the number of observations. Full period: January 03, 2000 to December,05 2017. Number of observations: 4683.

Moreover, since the estimation procedure of vMEM is obtained through an equation by equation approach, so that I have to deal with n likelihoods, and because off the HAR-DCC model has been estimated through a two steps approach, for those models, the comparison is made through two classical loss-functions, MSE and MAE.

Referring to Table 9, as it can be easily noticed, in the multiplicative parameter class, the maximum likelihood values are not comparable, since the parametrizations rely on a very different number of parameters, thus the only possible comparison is made through the information criteria. Whatever information criterion (AIC or BIC) is used, the DCC-CAW-MEM outperforms the ReBEKK model.

Furthermore, Table 10 reports two specific loss functions values: the Mean Square Error (MSE) and the Mean Absolute Error (MAE).

Hansen and Lunde (2006) proved that those functions are robust and allow an unbiased model ranking. Additionally, the use of MAE is due to the fact that, as suggested by

Andersen et al (1999), this criterion is more robust than RMSE, because less susceptible to outliers.

	Models	MSE	MAE
1)	vMEM	30,00	3,32
2)	DCC-CAW-MEM	106,45	7,68
3)	ReBEKK	207,52	10,68
4)	Np-ReBEKK	291,89	14,94
5)	HAR-DCC	2371,18	23,81

Table 10: Full sample Loss Functions. Full sample performances in terms of the mean square errors and mean absolute errors of the model estimated. The specific loss functions are calculated as follows: $MSE = T^{-1}\sum_{T=1}^{T} e_i^2$ and $MAE = T^{-1}\sum_{T=1}^{T}|e_i|$, with e_i forecasting error. Full period: January 03, 2000 to December,05 2017. Number of observations: 4683.

On the basis of those criteria, among the multiplicative models, the DCC-CAW-MEM has the best performance (MSE= 106.45 and MAE= 7.68), even if the lowest values, in absolute terms, are achieved by the vMEM, but it is important to remember that in the specification adopted in this study, it has only diagonal parameter matrices, thus, differently from the other models, it does not take into account correlations among assets.

3.3 Forecasting results

To determine whether the better in-sample performances of the DCC-CAW-MEM goes along with forecasting gains, I performed an out-of-sample one step ahead prediction for each model.

As mentioned in Section 3.1 I recursively estimated the parameters on the first 4161 daily observations (that is using data from January 03, 2000 to December 04, 2015) and then used the remaining 522 observations as forecast sample.

Afterward, I assessed the out-of-sample performance of the models through the Model Confidence Set (MCS) of Hansen and Lunde (2011), that allowed me to made a comparison without choosing a reference model.

3.3.1 Model Confidence Set

The Model Confidence Set (MCS) of Hansen and Lunde is one of the most popular way to assess the statistical predictive performances among different econometric models. Let M^0 be the initial collection of models for which we compute the one-step ahead predictions. The MCS approach is based on an iterative procedure, that, at each evaluation step, combines an equivalent test δ_M and an elimination rule e_M. The equivalent test δ_M uses the relative performance variable $d_{ij,t} = L_{i,t} - L_{j,t}$, defined as difference between loss values of alternative models, i and j, so that models are ranked in term of expected loss $\mu_{ij} \equiv E(d_{ij,t})$, and the alternative i is preferred to the alternative j if $\mu_{ij} < 0$.

Accordingly to the MCS procedure, the equivalence test is applied to the set $M = M^o$ and if d_{Mij} is rejected, it means that there is evidence that the objects in M are "not equally good" and the elimination rule is used to individuate the models with the poorest performance from M. this procedure, that is repeated until δ_M is not rejected, allow to individuate the set M^* of surveying models. Thus, at a given confidence level, chosen by the users, the MCS contains the single model or the set of models with better

forecasting performances. Moreover, as required by the procedure, the MCS procedure requires one or more loss measures. Regarding the choice of loss function, in order to guarantee a comparability with the in-sample performances of the models, I used the same loss functions previously selected: the Mean Square (Forecast) Error (MSE) and the Mean Absolute (Forecast) Error (MAE), and I computed the MCS at two confidence levels (75 % and 90%). Whatever is the confidence level used, the MCS approach only includes, in the final "better quality" models set, the DCC-CAW-MEM, that outperforms respect to its competitive multiplicative models (ReBEKK and Np-ReBEKK) and to the baseline vMEM and the two-step HAR-DCC model.

Moreover, for the multiplicative models that rely on the Wishart distribution assumption, the MCS has also been computed through the QLIKE loss function, computed as $QLIKE = [log|S_t| + tr(S_t^{-1} C_t)]$. As can be noticed from the p-values reported below, Table 11, in this case the Np-ReBEKK model outperforms compared to the DCC-CAW-MEM, but this is not surprising, since it separately models the long run covariance matrix, in a non-parametric way.

	vMEM	DCC-CAW-MEM	ReBEKK	Np-ReBEKK	HAR-DCC
	PANEL A: 90% MCS				
MSE	0,000	**1,000**	0,000	0,000	0,000
MAE	0,000	**1,000**	0,000	0,000	0,000
QLIKE		0,012	0,000	**1,000**	
	PANEL B: 75% MCS				
MSE	0,000	**1,000**	0,000	0,000	0,000
MAE	0,000	**1,000**	0,000	0,000	0,000
QLIKE		0,015	0,000	**1,000**	

Table 11: Forecasting Evaluation
This Table reports the forecasting results, in terms p-values obtained from the MCS approach, based on different loss functions: the MSE, the MAE and the QLIKE.

These findings suggest to extend the DCC-CAW-MEM parametrization, separately modelling the secular variance covariance matrix, in a similar way already seen for the Np-ReBEKK model.

Conclusions

The present thesis tries to contribute to the existing econometric literature, dealing with the estimation and the forecasts of variances, covariances and correlations.

My work is heavily based on the most recent finding in financial econometric, and evidences the important role played by realized measures in developing more robust estimators and in stimulating the development of new models.

To summarize, an initial part of the work is devoted to the presentation of basic volatility concepts and to the analysis of high frequency context, by introducing the realized measures and in particular, realized variance and realized covariance matrix estimators. They have been found to guarantee more precise results, especially when they are used together with models directly designed to deal with daily and intraday data, instead of the traditional GARCH class models.

In view of this background, I introduced different model parametrizations that can be used in large dimensional systems, with tens of assets. I decided to consider some multiplicative models, based on the Wishart distribution, but with different structures, ranging from mean reverting models to others that are computationally intensive but more realistic, in large systems. Among the others (vMEM, ReBEKK, Np-ReBEKK and HAR-DCC), I introduced a new parametrization, the DCC-CAW-MEM, that is directly fitted on the daily series of realized variance covariance matrices, and both allow for a dynamic parametrization through the Engle's ReDCC dynamic and incorporate as much information as possible through the multiplicative error model component.

Obviously, the estimation of all the models has required the definition of several constraints, for ensuring the accuracy of the parameters and relative standard errors, without losses in efficiency of the obtained results.

Main message here is that building a multivariate time series model requires an intensive effort and computational complexity has to been carefully considered.

Thus, different econometric techniques have been considered in the model design (from the CholCov estimator of daily realized variance covariance matrix and the covariance targeting procedure, to the robust sandwich estimator and model

confidence set approach) and other computational tasks, most of which arising from numerical optimization techniques, have required different solutions.

In conclusion, this model is will be improved and many others parametrizations could be introduced, but it was suitable for my goals.

Appendix A

a.1) Plots of Daily Log-Returns and Annualized Realized kernel volatility

For brevity, in Chapter 3, when the dataset has been presented, in that Section I only reported the daily log-returns and annualized realized kernel volatility for the two assets that presented the lowest and the highest range of variability, given the full sample period, with all the 4683 observations. For completeness, in the following pages I report the same plot for all the assets included in the sample.

a.2) Correlation Matrix

This Appendix reports the correlation matrix among series. As stated in Chapter 3, even if the correlations coefficients among asset are high, the statistical significance is assessed only for KS11, MXX and NSEI indices.

a.1) Daily Log-Returns: plots

AEX

AORD

BVSP

DJI

FCHI

FTSE

FTSEMIB

FTSTI

GDAXI

HSI

IBEX

IXIC
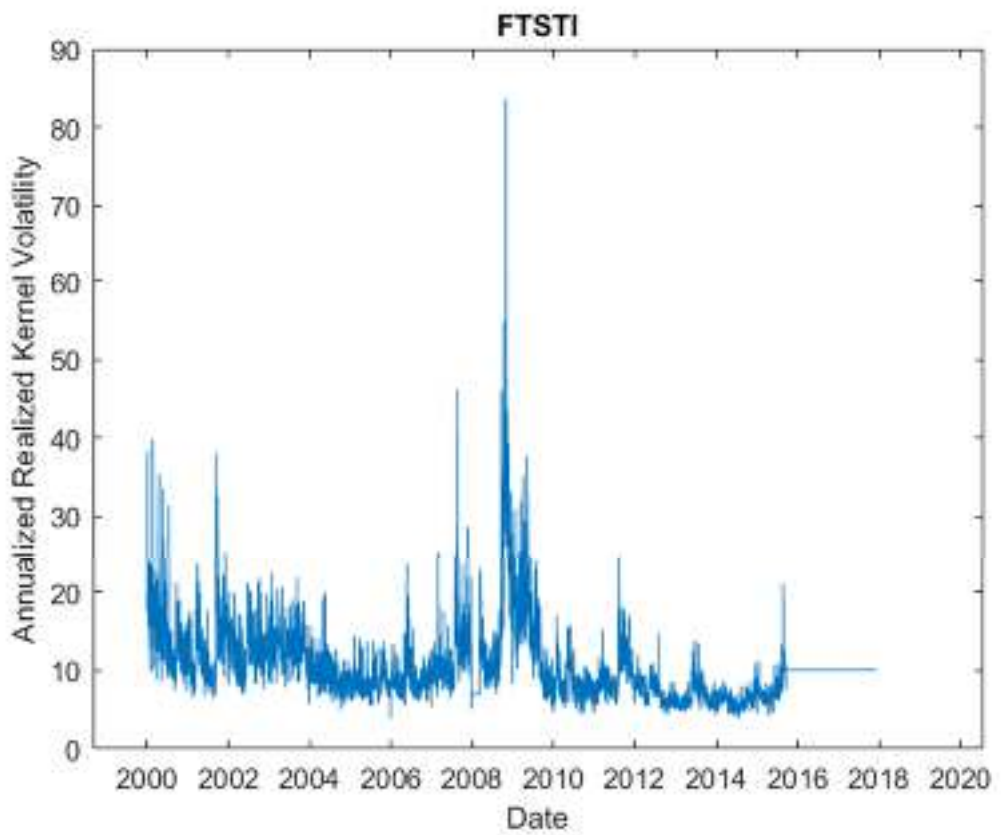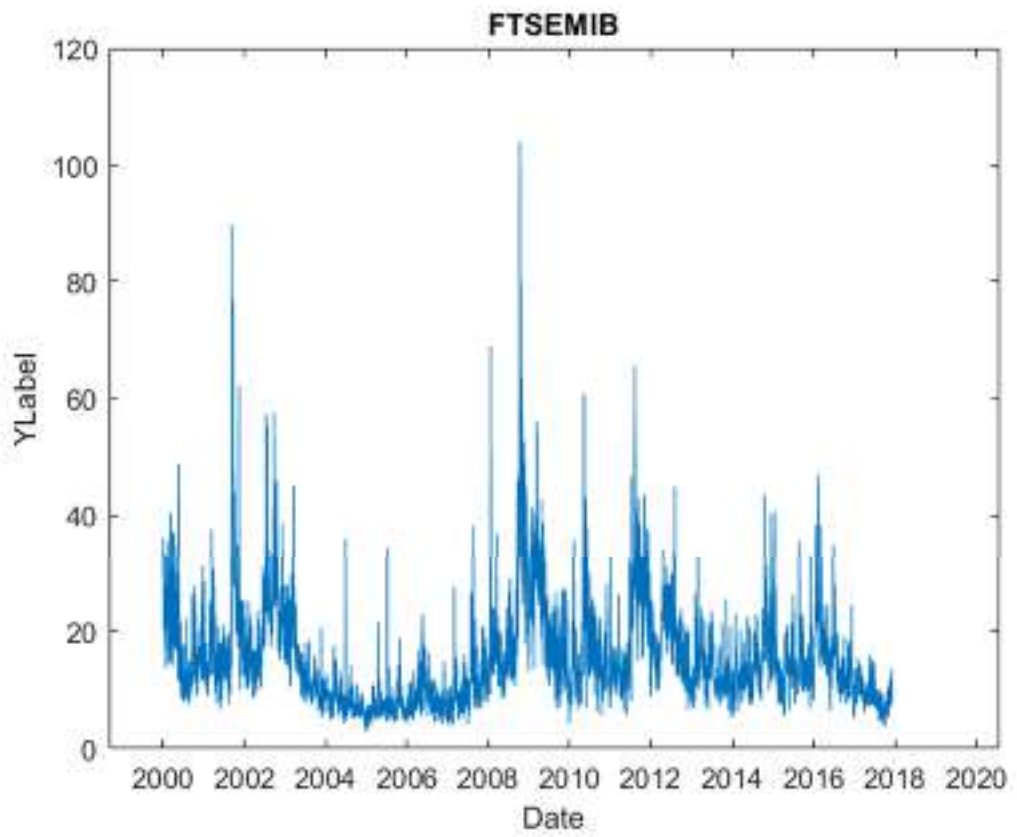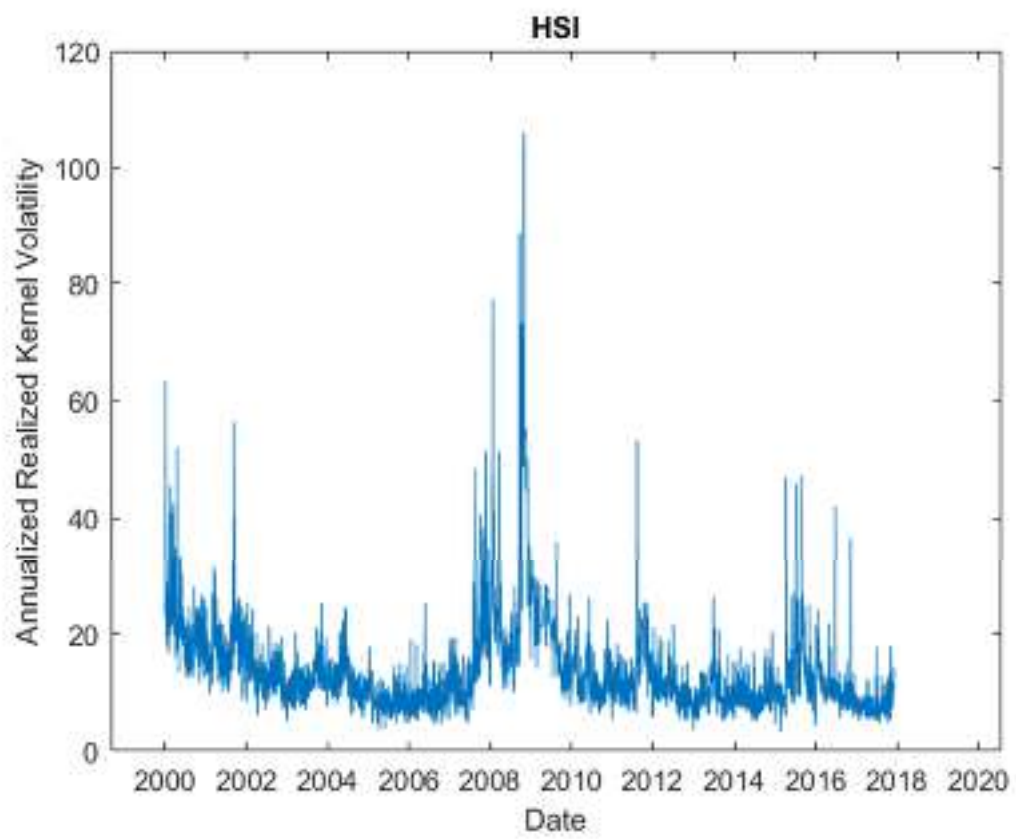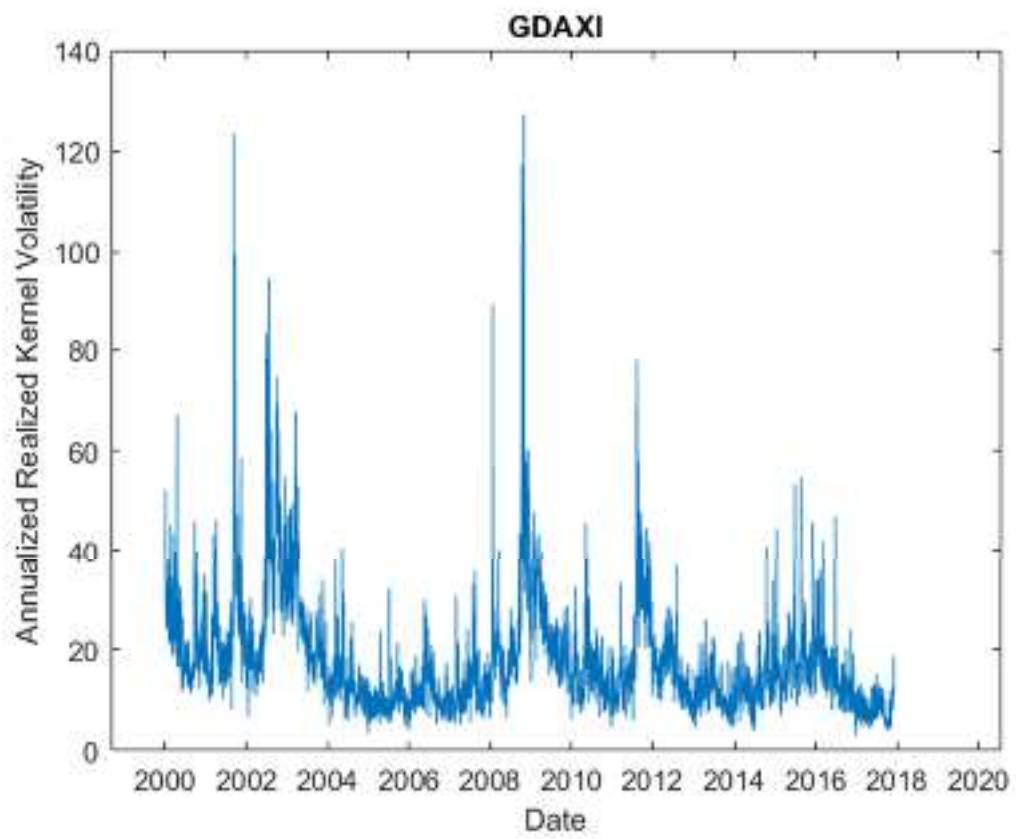
KS11

MXX

N225
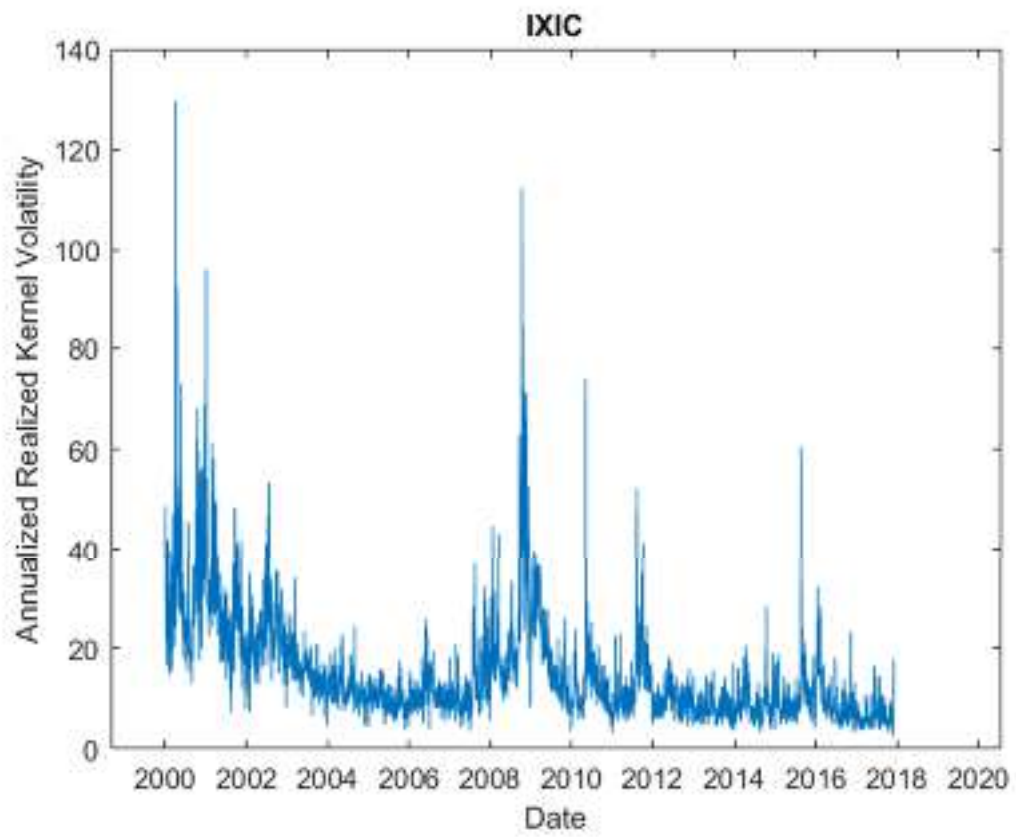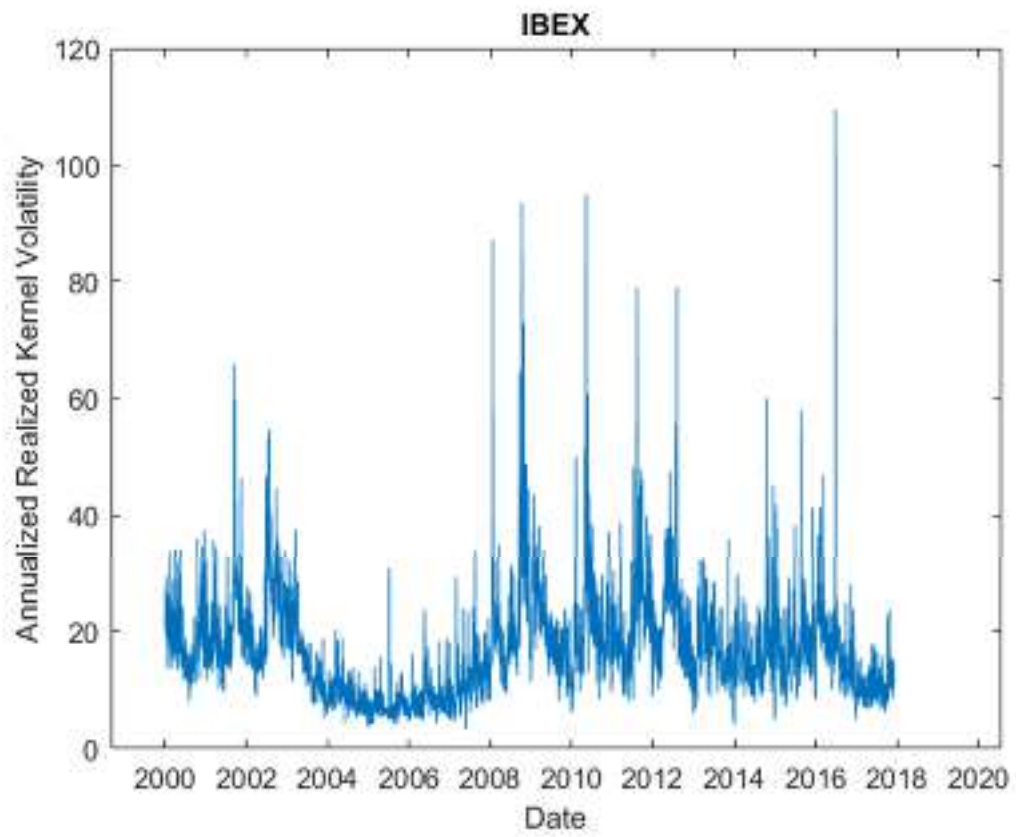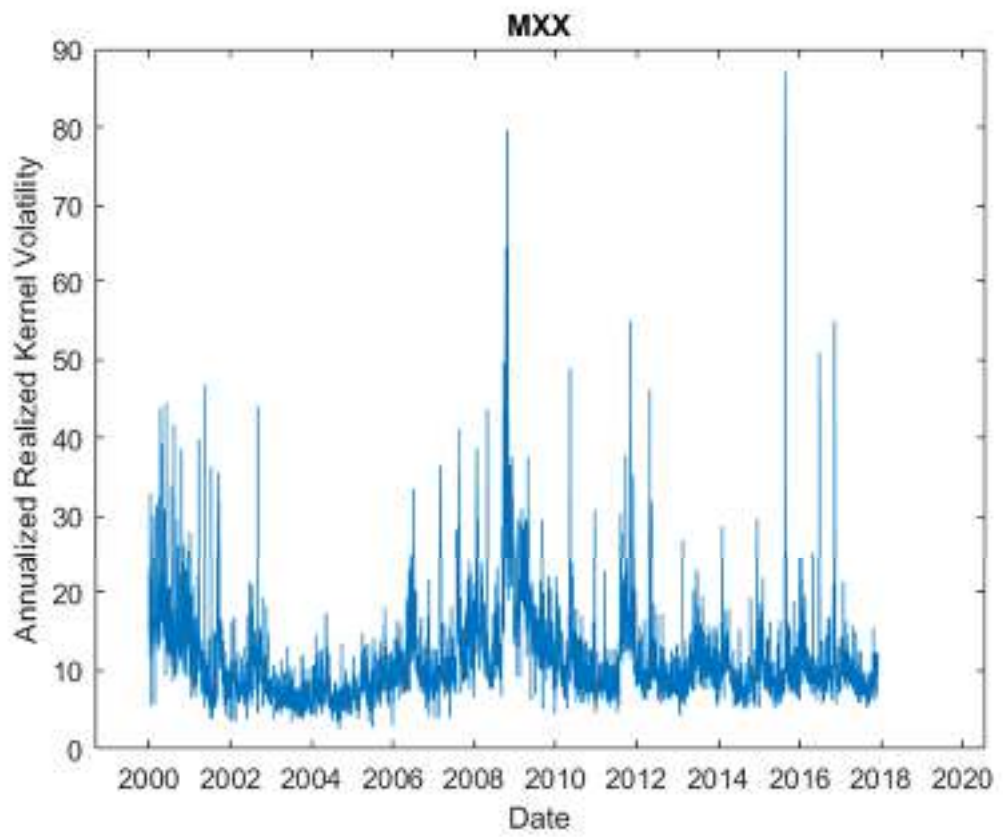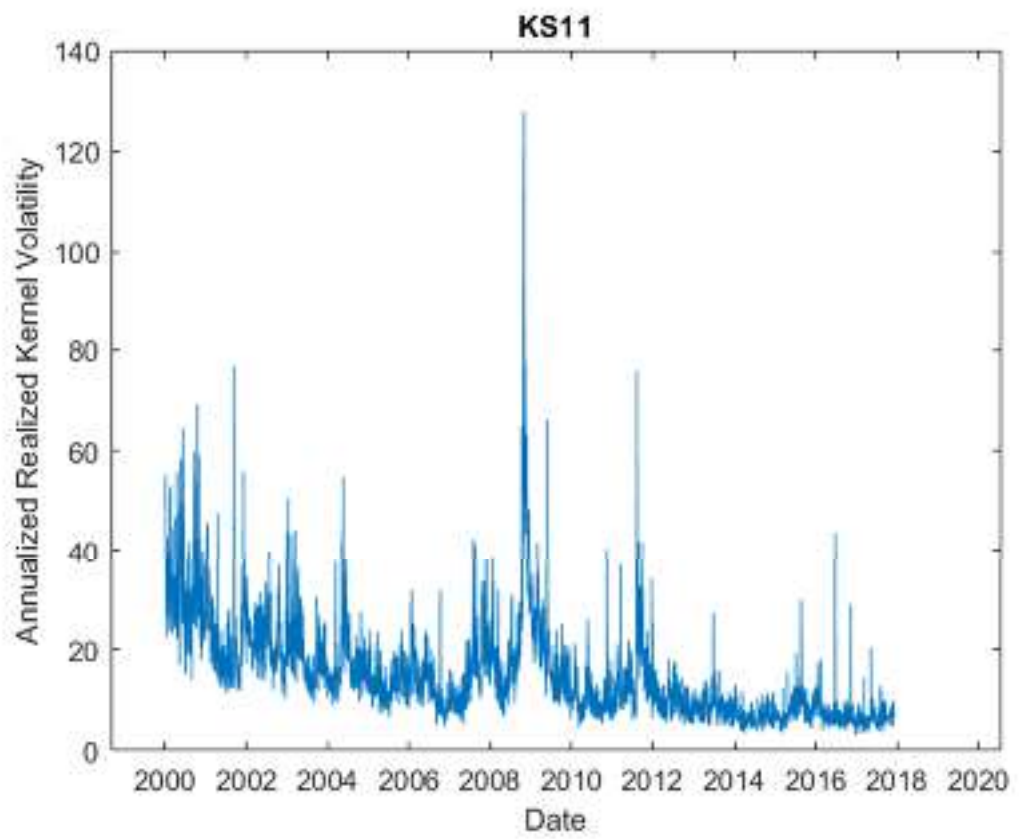

NSEI

RUT

SPX
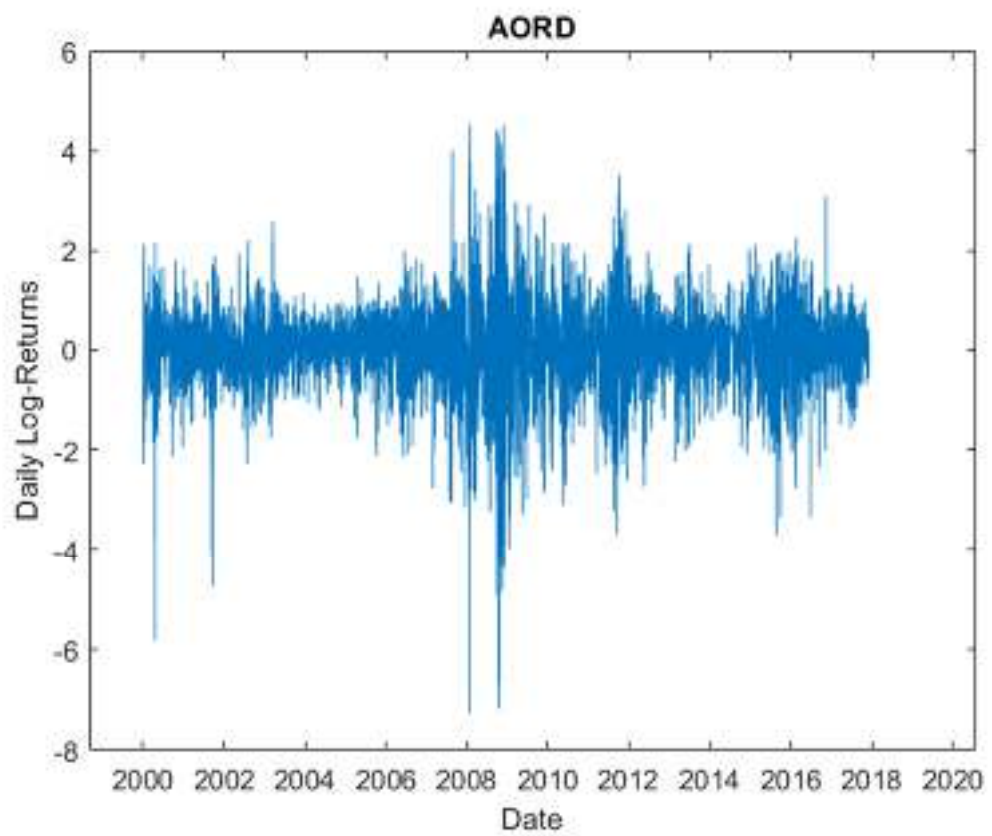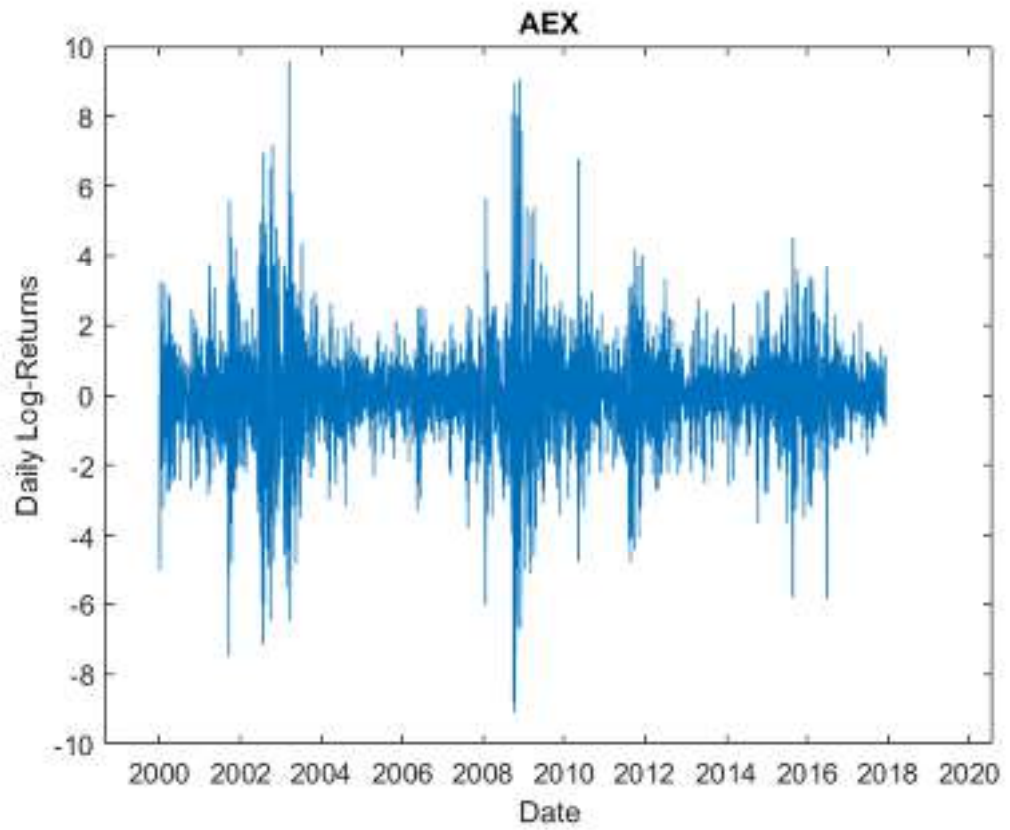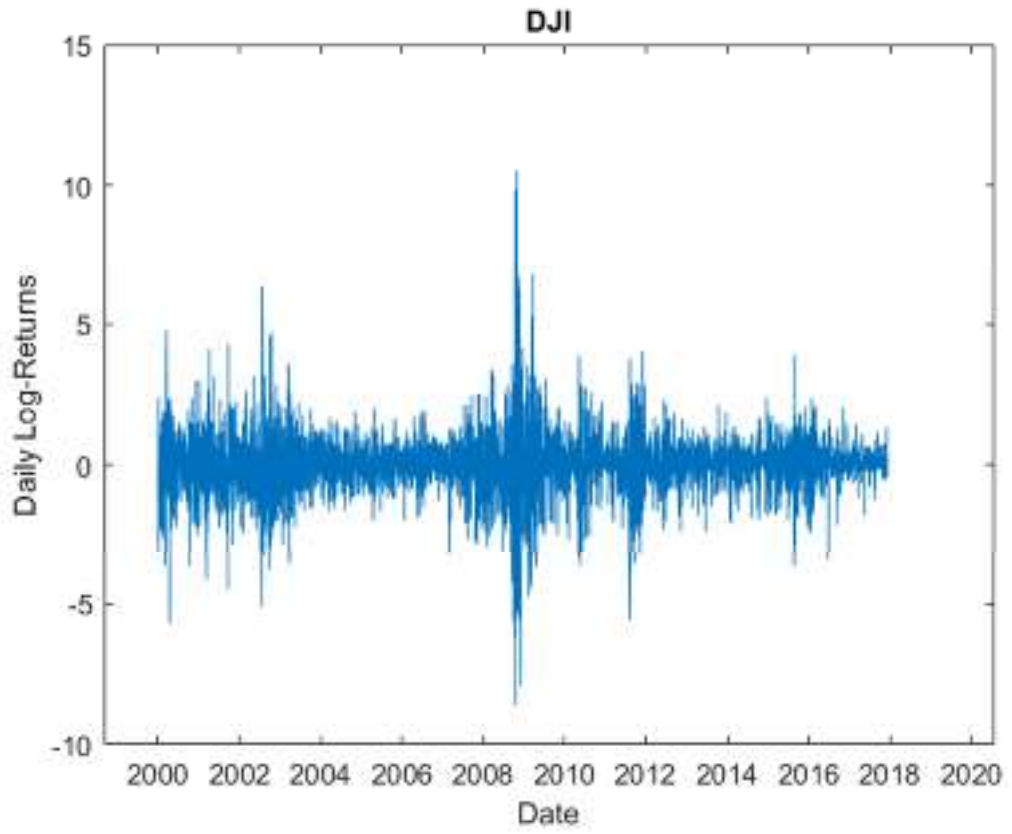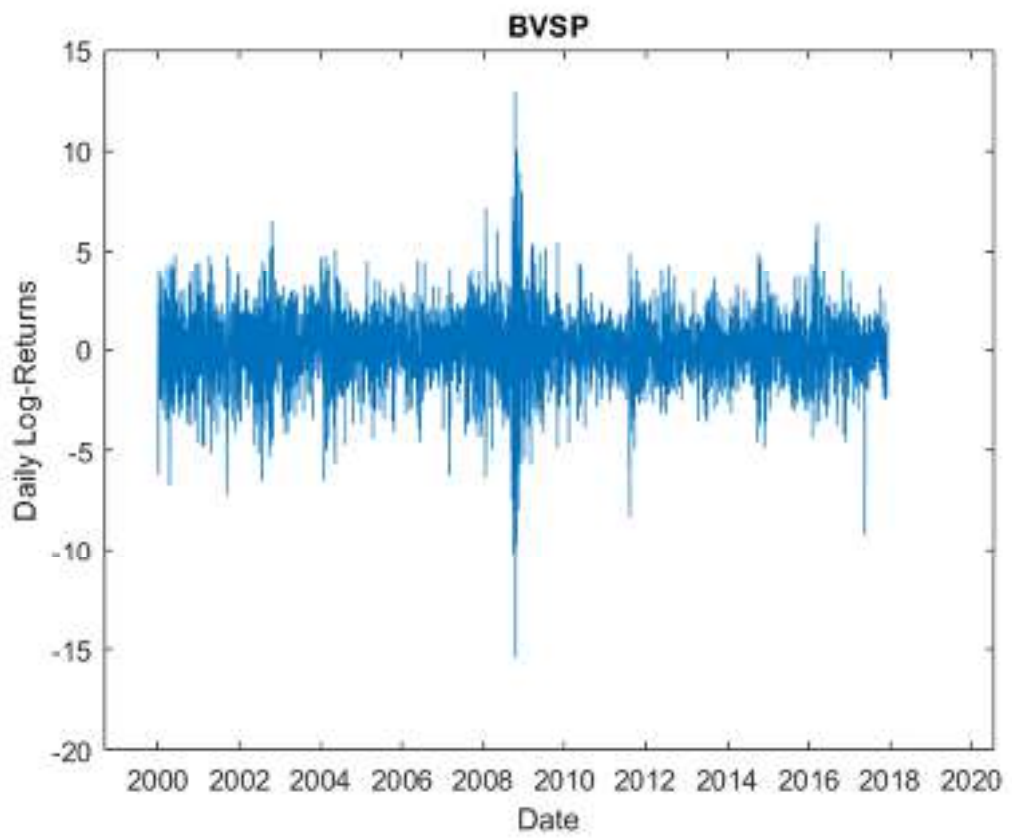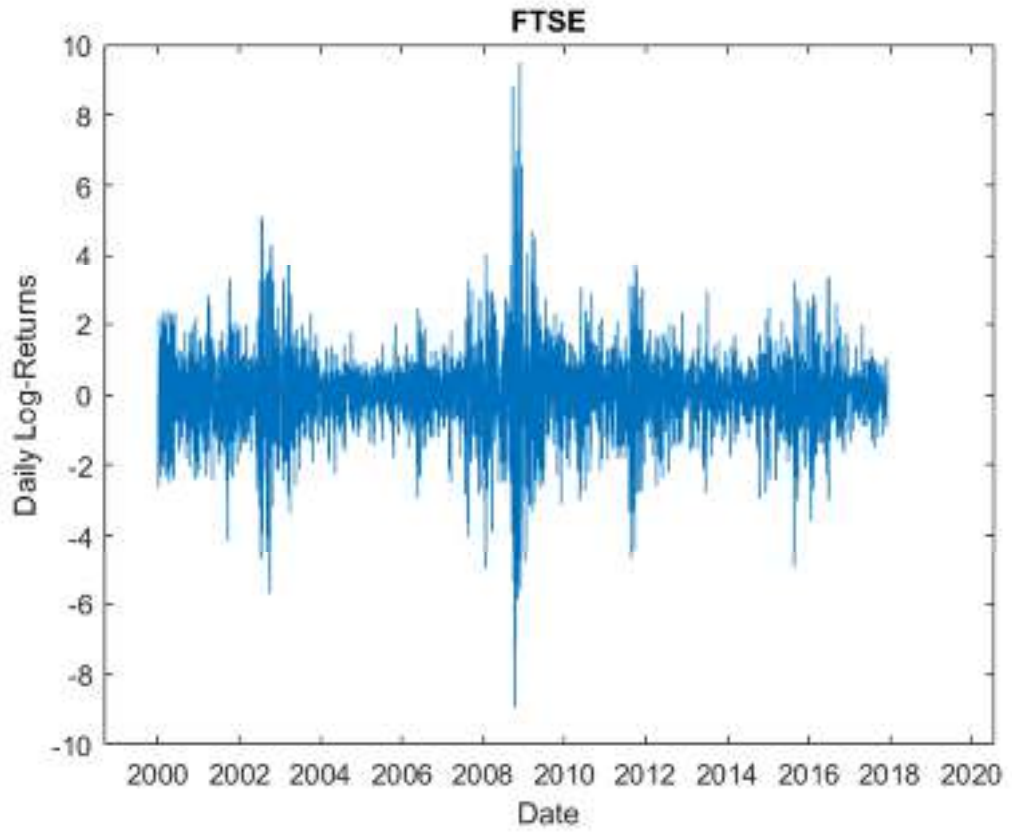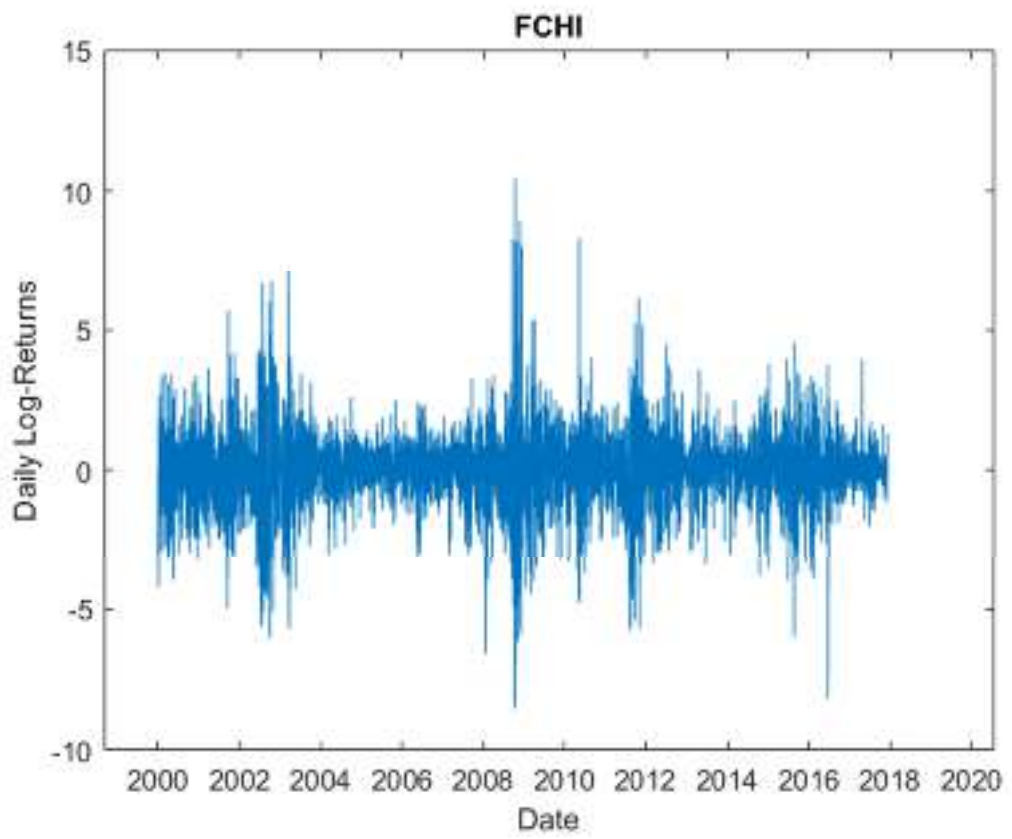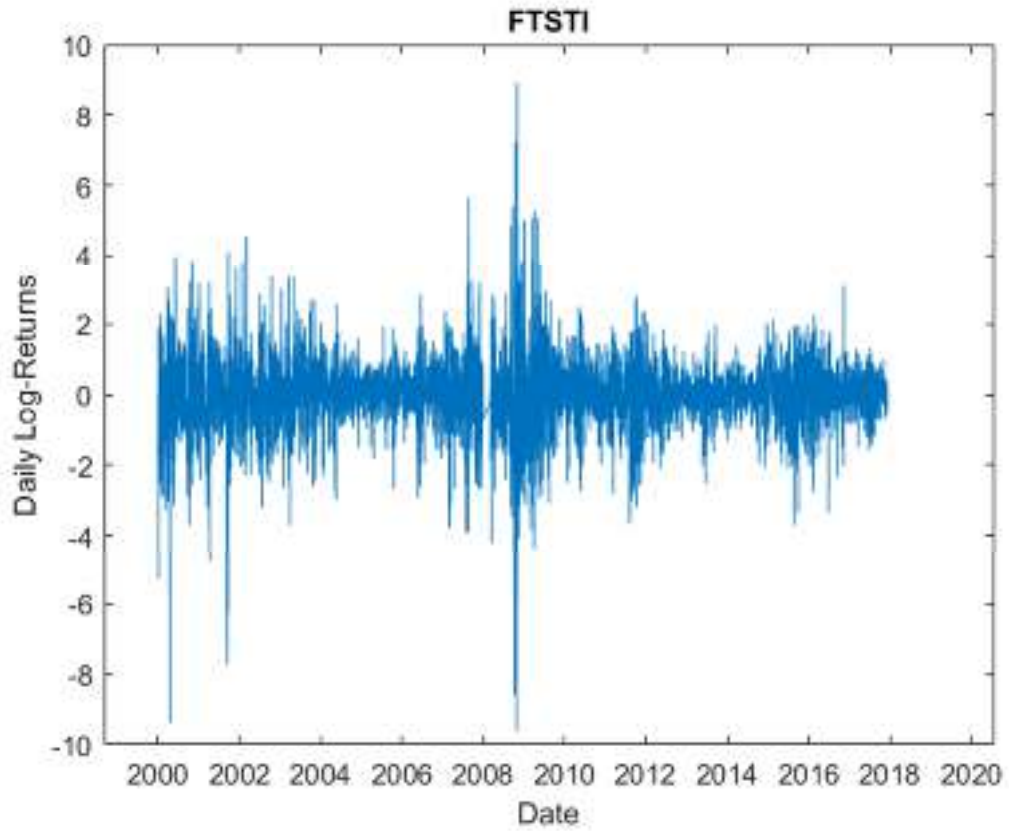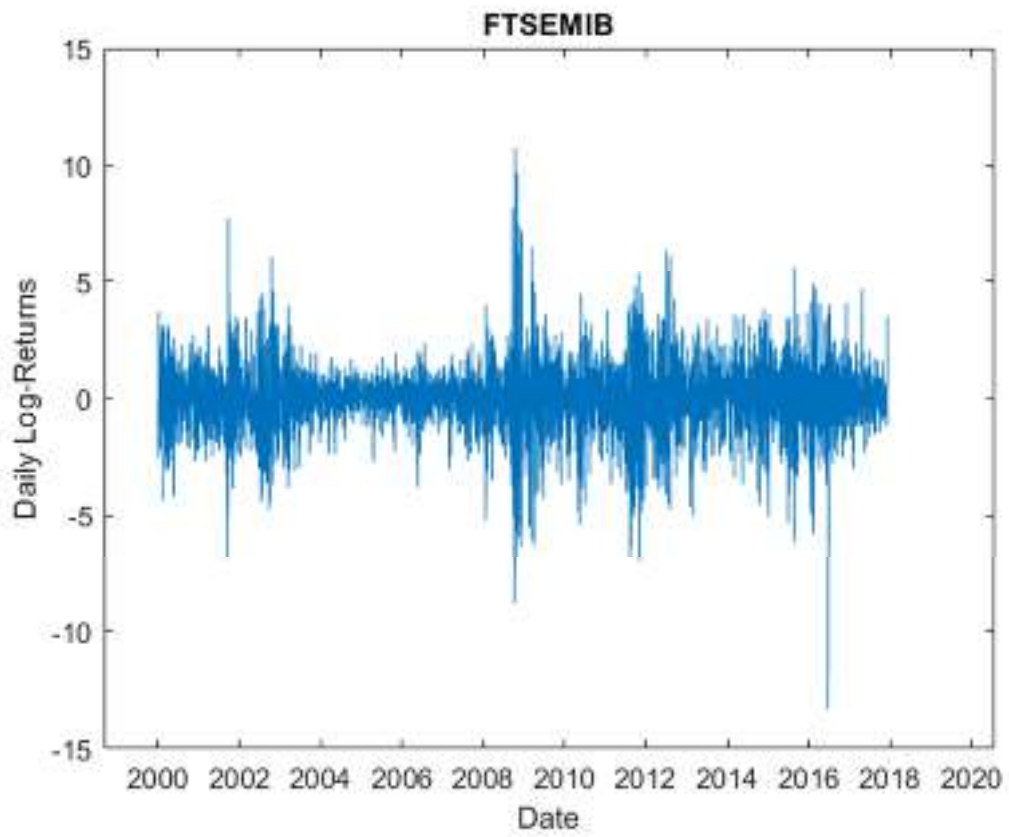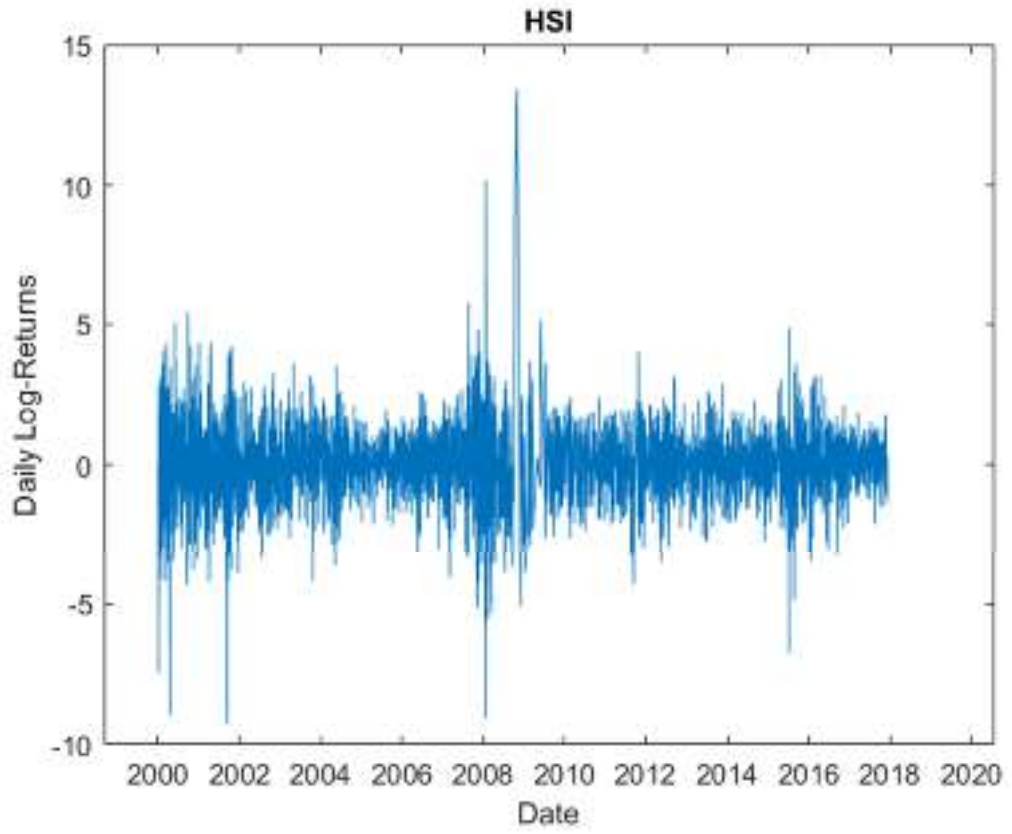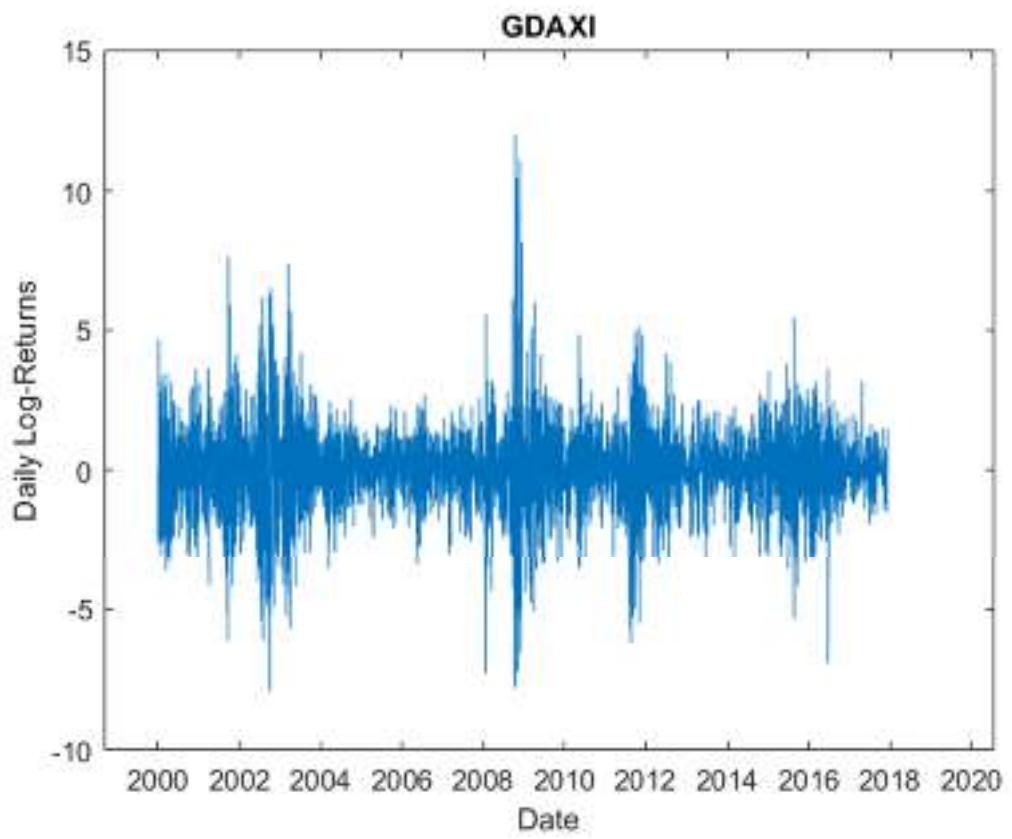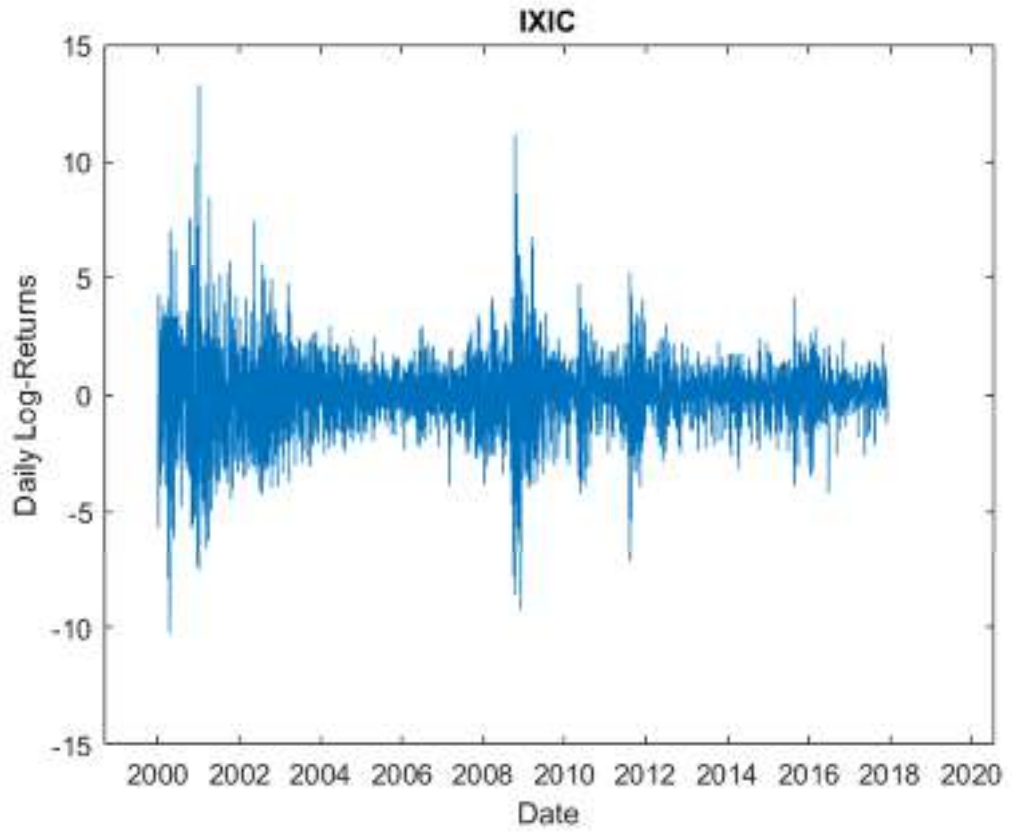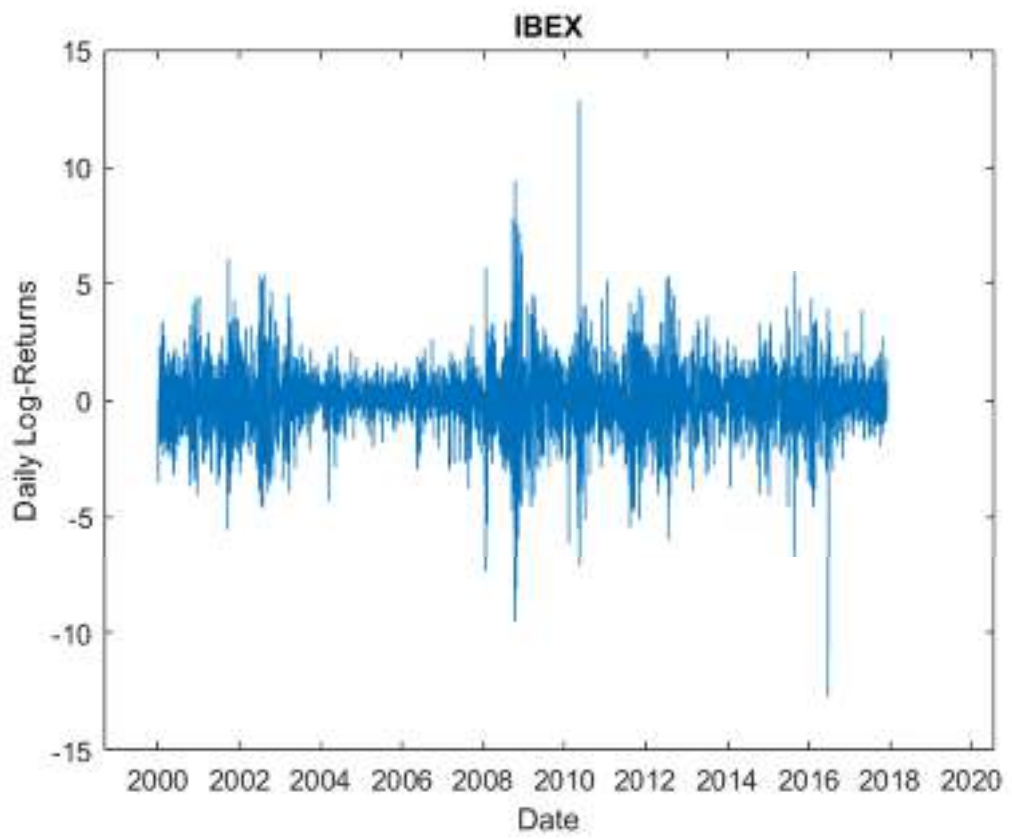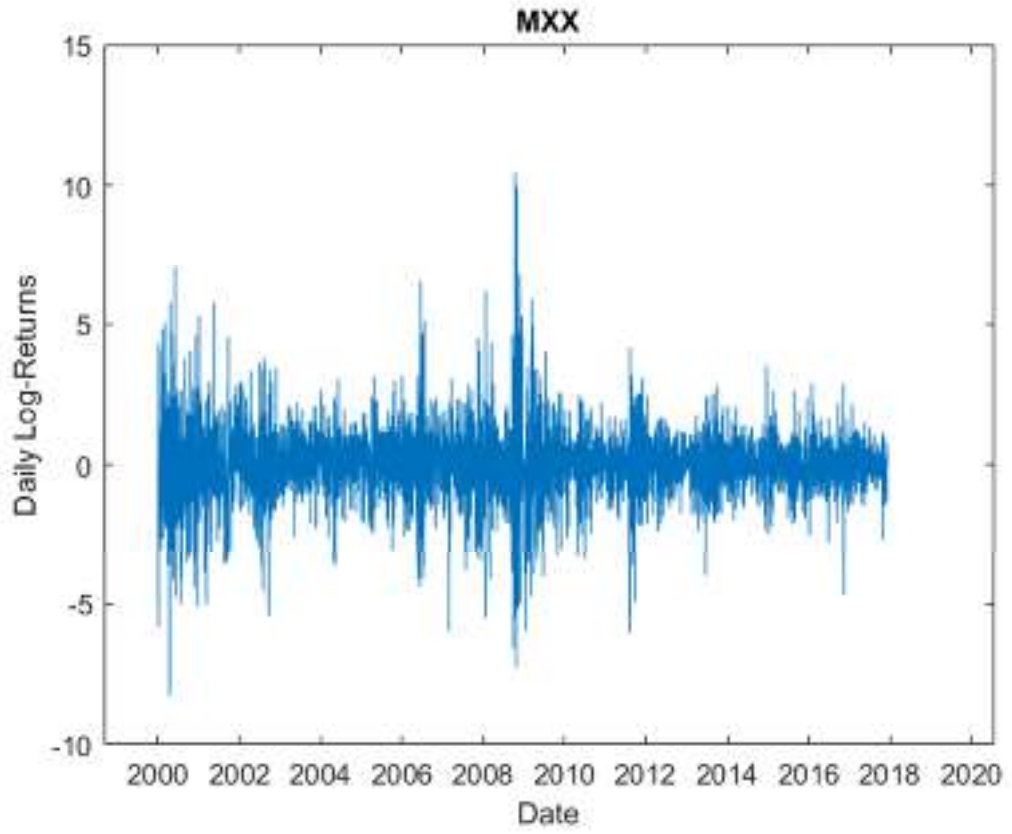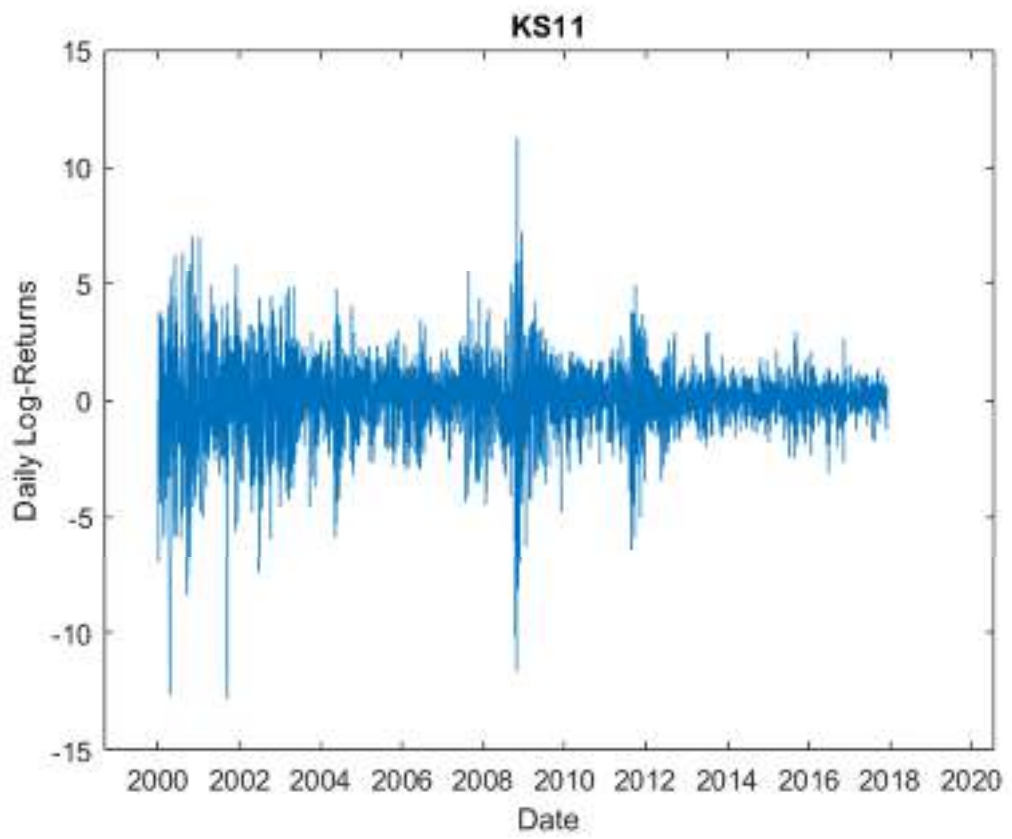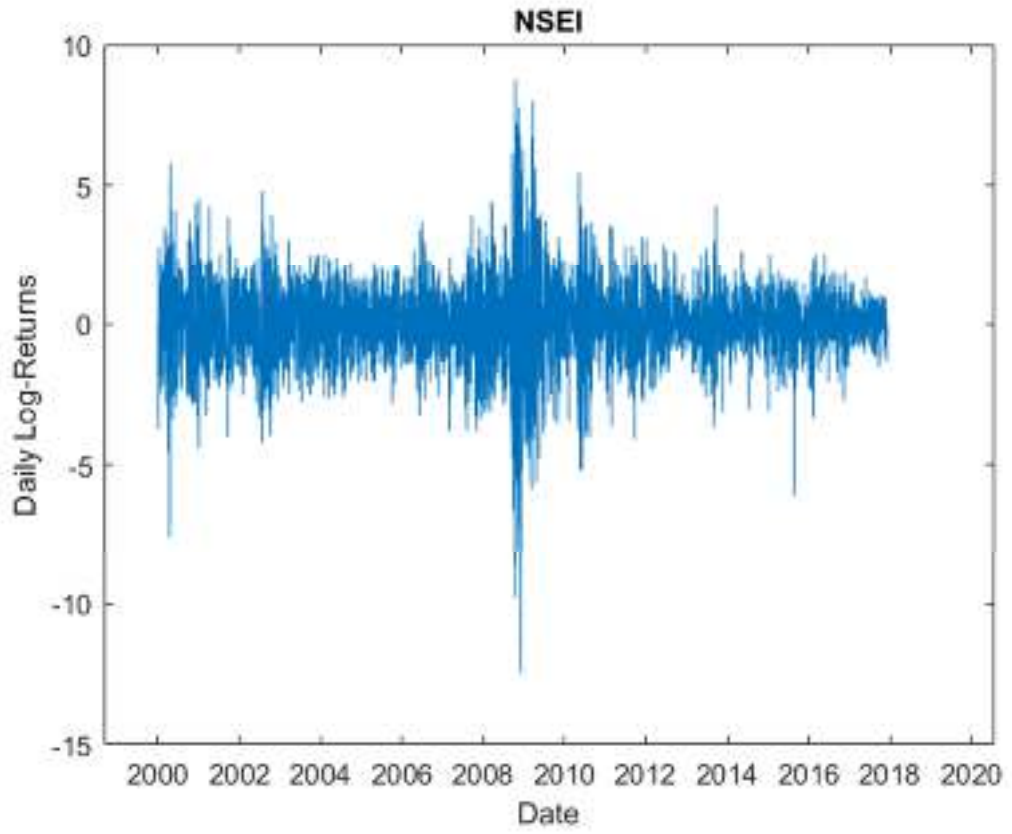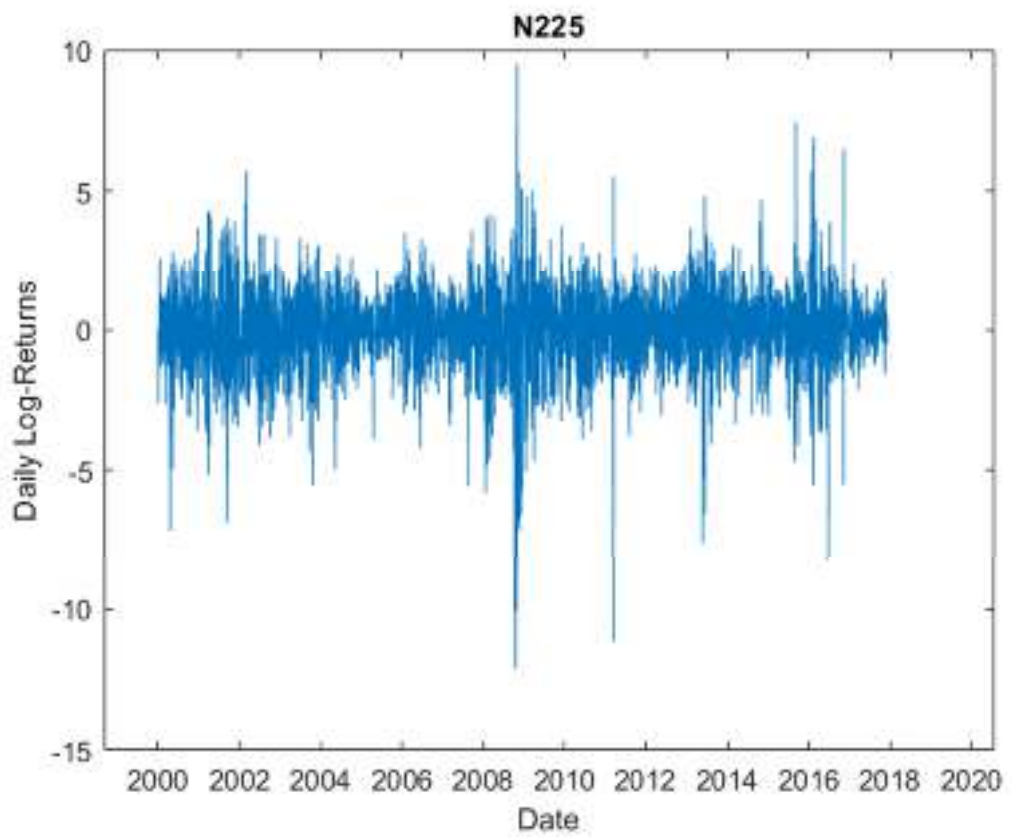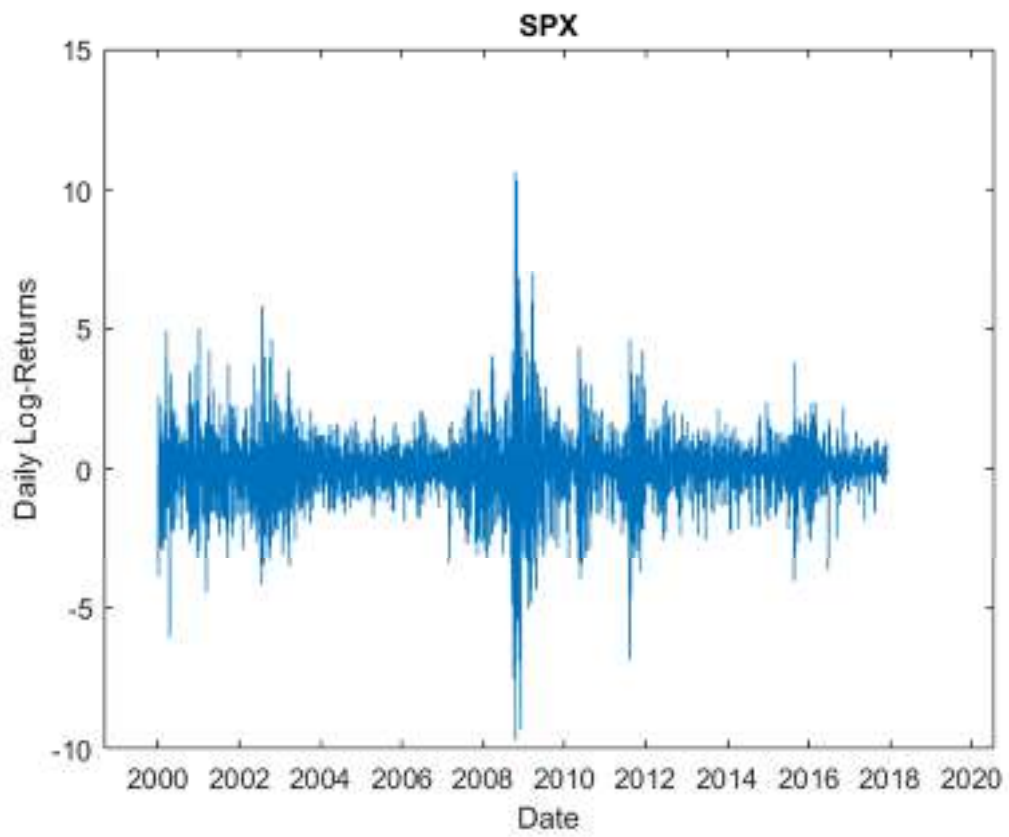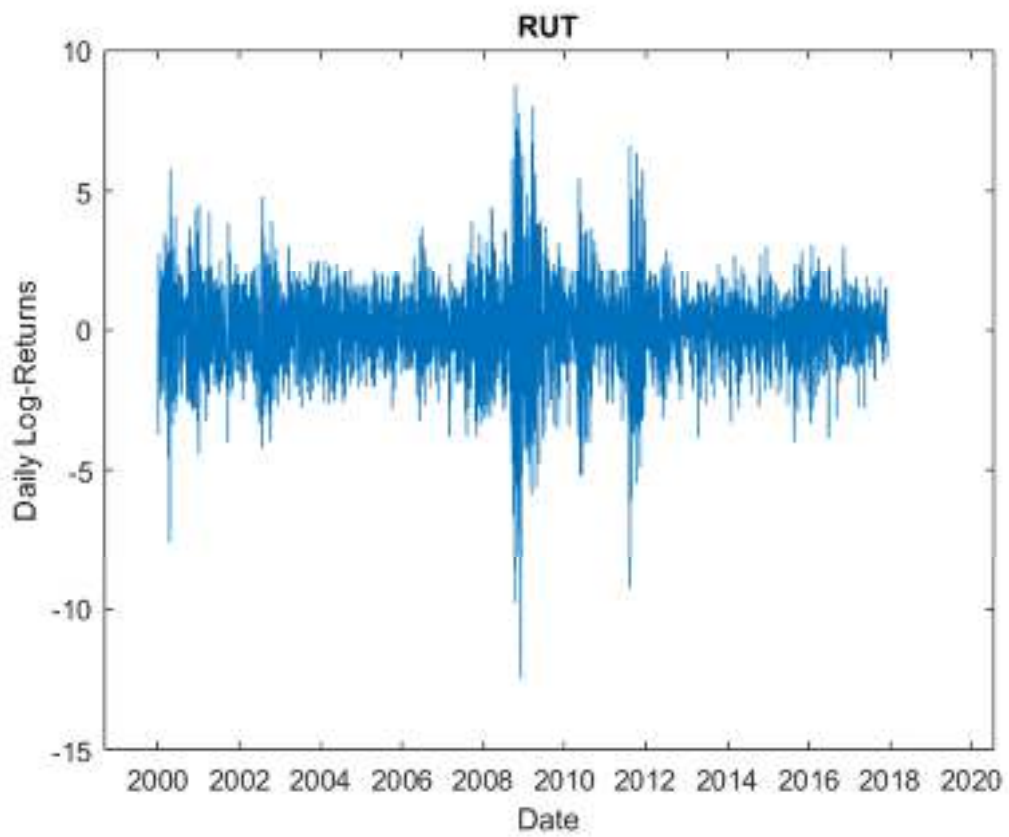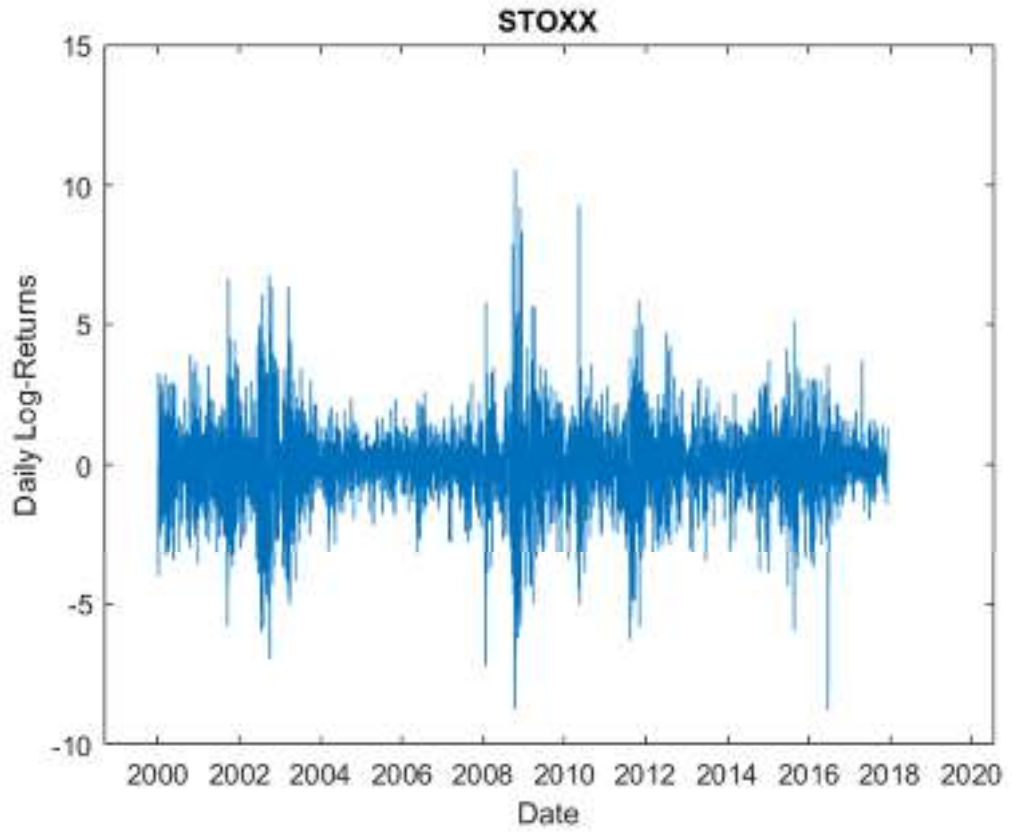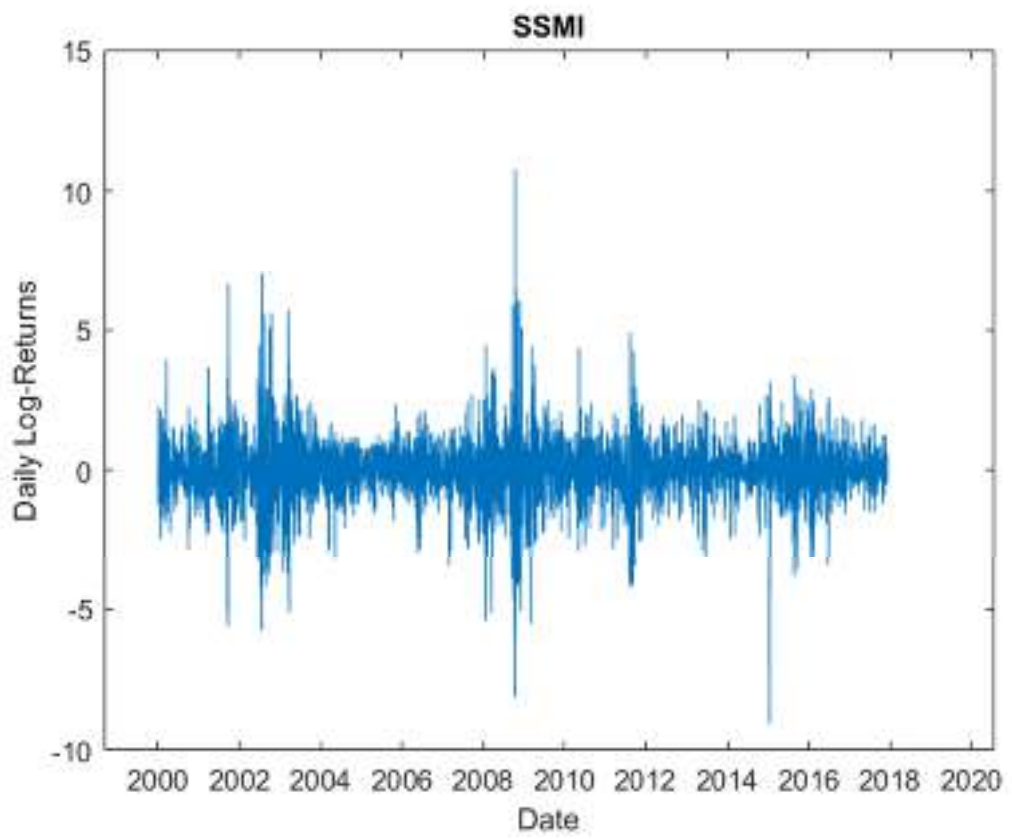
a.1) Annualized Realized Kernel Volatility: plots

BVSP

DJI

FCHI

FTSE

FTSEMIB

FTSTI

GDAXI

HSI

IBEX

IXIC

KS11

MXX

N225

NSEI

RUT

SPX

82

SSMI

STOXX

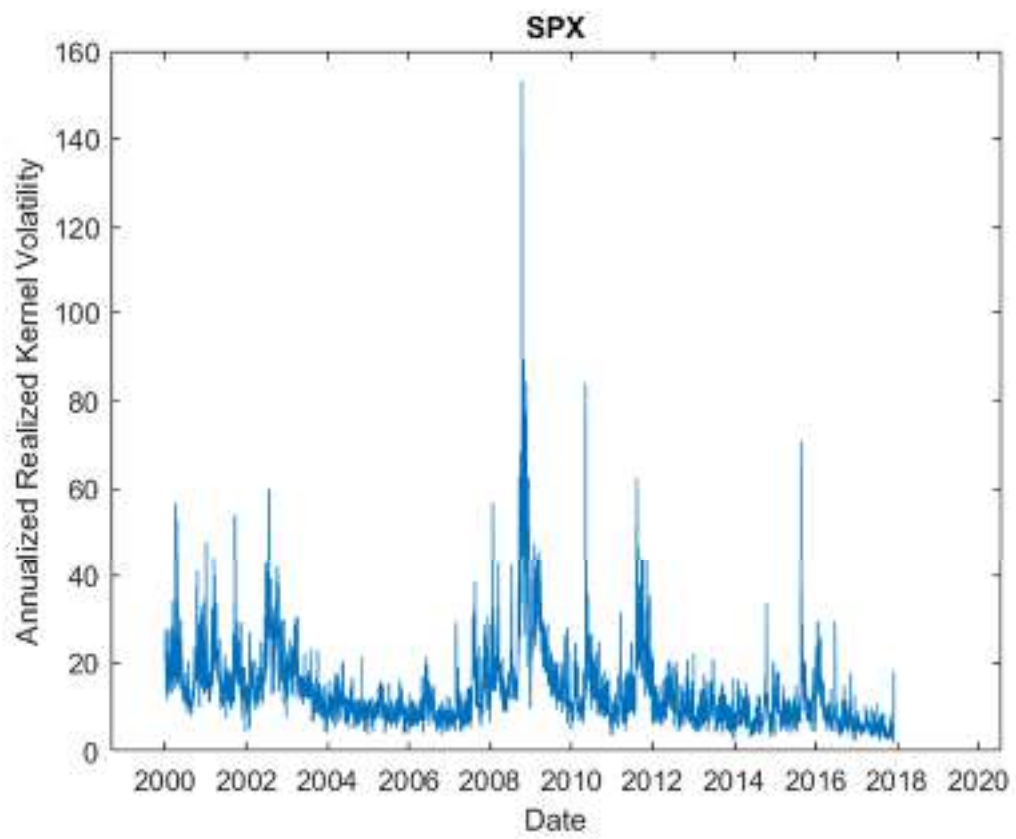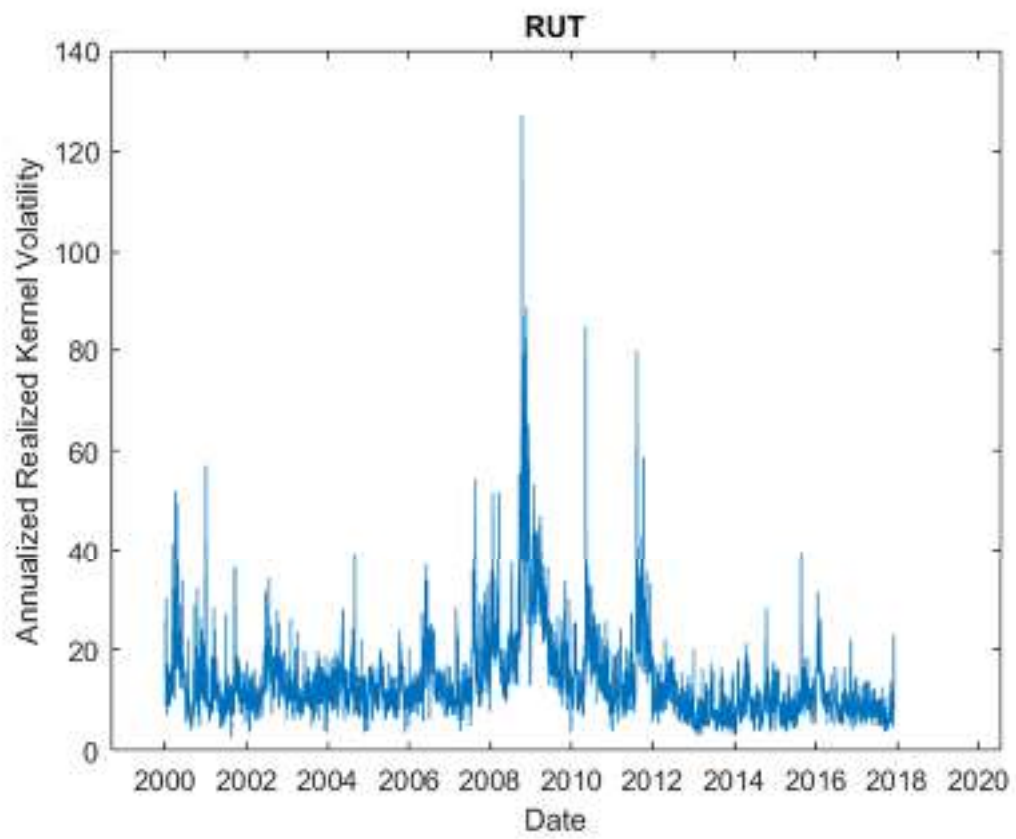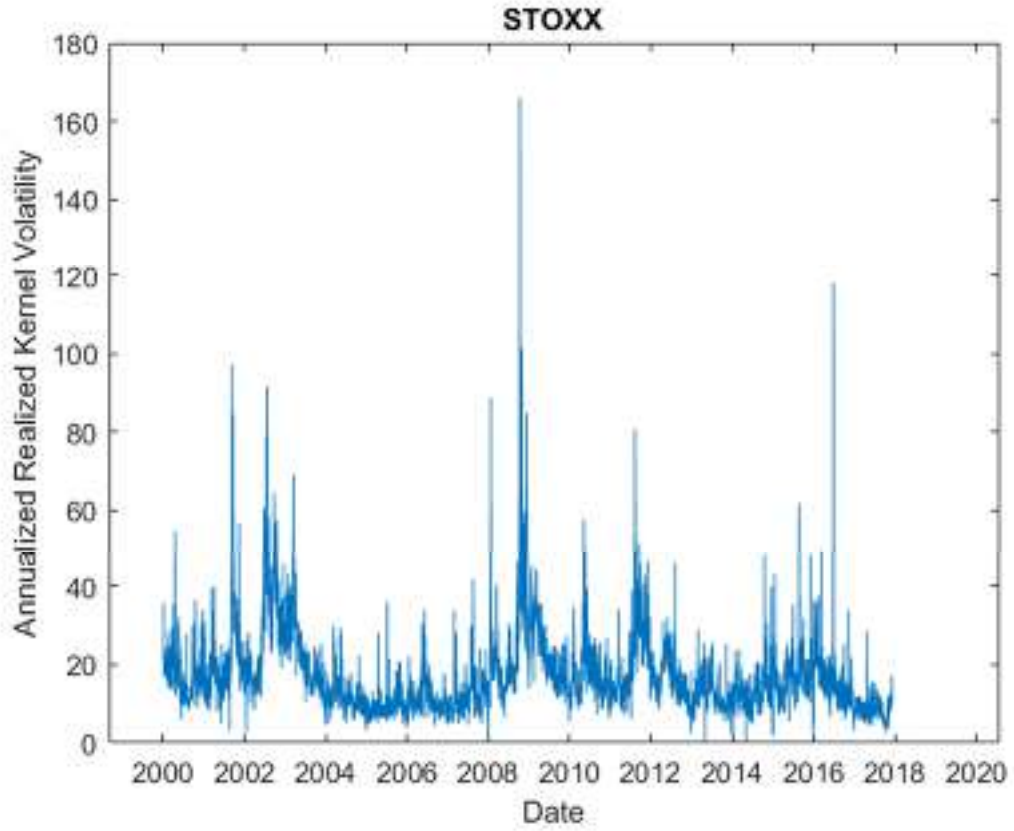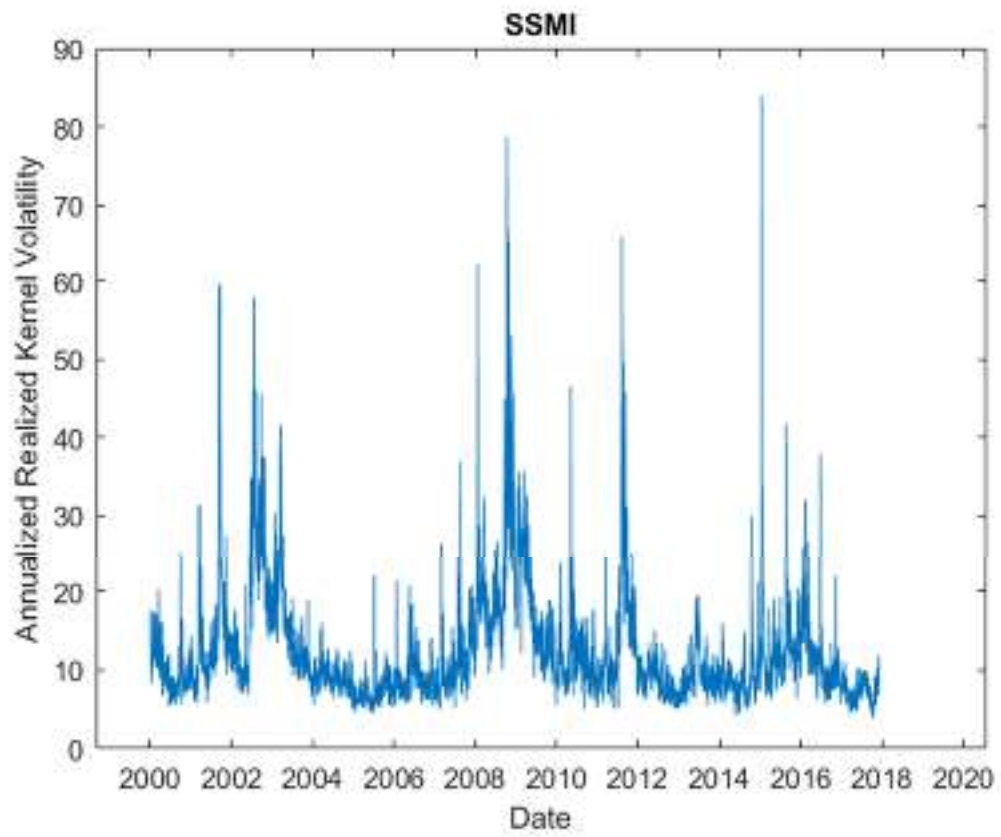|	AEX	AORD	BVSP	DJI	FCHI	FTSE	FTSEMIB	FTSTI	GDAXI	HSI	IBEX	IXIC	KS11	MXX	N225	NSEI	RUT	SPX	SSMI	STOXX
AEX	1,00																			
AORD	0,58	1,00																		
BVSP	0,60	0,54	1,00																	
DJI	0,80	0,64	0,73	1,00																
FCHI	0,95	0,60	0,60	0,81	1,00															
FTSE	0,92	0,61	0,63	0,82	0,91	1,00														
FTSEMIB	0,85	0,58	0,57	0,74	0,91	0,83	1,00													
FTSTI	0,58	0,48 *	0,56	0,63	0,55	0,63	0,50 *	1,00												
GDAXI	0,93	0,51	0,58	0,78	0,93	0,90	0,84	0,58	1,00											
HSI	0,55	0,59	0,57	0,64	0,57	0,62	0,53	0,68	0,54	1,00										
IBEX	0,79	0,55	0,50 *	0,68	0,86	0,75	0,88	0,37 *	0,76	0,46 *	1,00									
IXIC	0,64	0,45 *	0,57	0,78	0,67	0,72	0,60	0,60	0,67	0,60	0,53	1,00								
KS11	0,50 *	0,36 *	0,49 *	0,58 *	0,51 *	0,58 *	0,42	0,66	0,56	0,64	0,33 *	0,68	1,00							
MXX	0,52	0,61	0,62	0,68	0,56	0,58	0,55 *	0,51 *	0,49 *	0,57	0,51 *	0,57	0,45 *	1,00						
N225	0,57	0,54	0,52	0,62	0,56	0,59	0,51	0,58	0,57	0,59	0,46 *	0,57	0,56	0,49 *	1,00					
NSEI	0,34 *	0,44 *	0,51	0,44	0,32	0,37	0,26 *	0,42 *	0,30 *	0,49 *	0,20 *	0,25 *	0,35 *	0,39 *	0,41 *	1,00				
RUT	0,66	0,65	0,70	0,86	0,67	0,70	0,64	0,58	0,61	0,61	0,56	0,68	0,50 *	0,68	0,52	0,55	1,00			
SPX	0,80	0,62	0,73	0,98	0,82	0,84	0,74	0,64	0,79	0,65	0,68	0,84	0,61	0,68	0,62	0,43 *	0,87	1,00		
SSMI	0,92	0,62	0,61	0,80	0,90	0,88	0,81	0,56	0,87	0,55	0,74	0,57	0,47 *	0,52	0,58	0,41 *	0,68	0,79	1,00	
STOXX	0,94	0,58	0,60	0,80	0,96	0,90	0,88	0,53	0,94	0,51	0,85	0,63	0,48 *	0,53	0,57	0,32 *	0,65	0,80	0,89	1,00

Appendix B

The Wishart distribution

Intuitively, the Wishart distribution is the multivariate extension of the gamma distribution, although most econometricians use the Wishart as a matrix generalization of the Chi-Square distribution.

More formally, for easy in notation, I define the multivariate Gamma function as:

$$\Gamma_{n(v/2} = \Gamma^{n(n-1)/4} \prod_{i=1}^{4} \left[\frac{v/2 + (1-i)}{2} \right]$$

Thus, considering $C \sim W_n(\Sigma, v)$, where v is the degrees of freedom parameter and Σ denote a PSD matrix (in our context this is the variance covariance matrix from a multivariate normal distribution, then the Wishart distribution probability density function is designed as:

$$\frac{|C|^{\frac{v-n-1}{2}} exp\left[-\frac{1}{2} tr(\Sigma^{-1}C) \right]}{2^{vn/2}|\Sigma|^{v/2}\Gamma_{n(v/2)}}$$

Where $|\Sigma|$ is the determinant of the matrix, tr is the trace (that is the sum of the diagonal element of the matrix). Note that v >n-1 is the condition for the invertibility of Σ.

Moreover, since in the multivariate time series analysis, I often need to make inference about the inverse of covariance matrix, let's think about the inverse of our matrix, defined as $B = C^{-1}$. In this case, if $\psi = \Sigma^{-1}$, the inverse of Wishart will have this associated pdf:

$$f(B|\psi) = \frac{|\psi|^{v/2}|B|^{(v+n+1/2)} exp\left[-\frac{1}{2} tr(\psi B^{-1}) \right]}{2^{(vn)/2}\Gamma_{n(v/2)}}$$

I refer to B as the inverted Wishart distribution.

Link with X^2 distribution

If $X_1,\ldots,X_n$ are independent observation, the sum of independent square observations has a Chi square distribution with v degree of freedoms.

When X_i are random vector rather than real valued variables, one possible way to generalize the above sum of squared observations is through the following quadratic form $C = \sum_{i=1}^{n} X_i X_i'$. Essentially, this representation of S is often use for designing the basic Wishart distribution, when v is integer valued and this is equivalent to the Wishart generalization obtained through multivariate gamma distribution, when v is real valued.

Invariance and decomposition of C

The Wishart distribution has two relevant properties: it is invariant under marginalization (without changes in the degree s of freedom parameter), and it is invariant under transformation $C \rightarrow AC'A$.

Overview of models

Throughout the thesis several models are described. Here I briefly report their structure, and when necessary, I provide some further details. The notation used is the same as in the content of each chapter.

The vector Multiplicative Error Model (vMEM) of Cipollini et al. (2012)

The vMEM of Cipollini et al. (2012) is fitted on annualized realized kernel time series. It assumes a multiplicative structure, in the modelling of realized volatility RV. It's described by the following equations:

$$RV_t = \mu_t \odot \varepsilon_t$$

$$with \ \varepsilon_t | F_{t-1} \sim Gamma \ (\vartheta, \frac{1}{\vartheta})$$

and conditional expectation

$$\mu_t = \omega + \sum_{l=1}^{L} [\gamma_l RV_{t-l} + \delta_l \mu_{t-l}]$$

with $\omega, \gamma \ and \ \delta$ scalar parameters.

DCC-CAW-MEM

The DCC-CAW-MEM is the new parametrization, based on the Wishart distribution, a MEM univariate volatility framework and dynamic conditional correlations and covariances. It is directly fitted on the series of realized variance covariance matrices, computed through the CholCov estimator of Boudt (2016) applied on log-returns. It is defined through the following equations:

$$C_t | F_{t-1} \sim W_n \ (v, {S_t}/{v})$$
$$E(C_t | F_{t-1}) = E_{t+1}(C_t) = S_t$$

$$S_t = D_t R_t D_t$$

$$diag\{D_{ii,t}\} = S_{ii,t} = \omega_i + \gamma_i RV_{ii,t-1} + \delta_i S_{ii,t-1}$$

$$Q_t = \bar{S}(1 - \alpha - \beta) + \alpha C_{t-1} + \beta Q_{t-1}$$

$$\bar{S} = \hat{E}[\varepsilon_{t-1}\varepsilon'_{t-1}] = T^{-1}\sum_{t=!}^{T}\varepsilon_{t-1}\varepsilon'_{t-1}$$

$$R_t = Q_t^{*-1/2} Q_t Q_t^{*-1/2}$$

$$Q_t^* = \{diag\,(Qt)\}$$

Where $\bar{S}$ is the sample covariance matrix of standardized residuals ε obtained from MEM specification. Regularization of Q_t through R_t is needed to obtain a well-defined correlation matrix with unit elements on the diagonal.

ReBEKK model

The ReBEKK model is based on the multiplicative model of Hafner and Lindton (2010). This model is sufficiently described by the following equations:

$$C_t|F_{t-1} \sim W_n\,(v\,,{}^{S_t}/_v)$$

$$E(C_t|F_{t-1}) = E_{t+1}(C_t) = S_t$$

$$S_t = \bar{S}(1 - \alpha - \beta) + \alpha C_{t-1} + \beta S_{t-1}$$

Where α and β are positive scalars and $\alpha + \beta < 1$.

NpReBEKK model

Differently from the previous BEKK parametrization, the NpReBEKK model relies on a decomposition of the variance covariance matrix though a long run component and short one. The first is modeled through the smooth one side kernel function of Nadaraya-Watson and the short run is modeled through a multiplicative scalar BEKK.

$$C_t|F_{t-1} \sim W_n\,(v\,,{}^{S_t}/_v)$$

$$E(C_t|F_{t-1}) = E_{t+1}(C_t) = S_t$$
$$S_t = M_t S_t^* M_t'$$

$$L = MM'$$
$$S_t^* = (1 - \alpha - \beta)I_n + \alpha C_{t-1}^* + \beta S_{t-1}^*$$

where $C_t^* = M_t^{-1} C_t (M_t^{-1})'$ and M is a $\{n(n+1)/2\}$ lower triangular matrix of parameters so that L is a mean reverting constant covariance matrix. Moreover, the Nadaraya-Watson estimator is computed as:

$$L_t(T) = \frac{\sum_{t=1}^{T} K_h \left(\frac{t}{T} - T \right) C_t}{\sum_{t=1}^{T} K_h \left(\frac{t}{T} - T \right)}$$

With $T \in [0,1]$, $K_h(.) = (1/h)K(./h)$ and k(.) is the one side Gaussian kernel function, and h the optimal bandwidth parameter.

HAR-DCC model

The HAR-DCC model combines the estimation of a multivariate DCC model of Engle, fitted on realized kernel time series, from which the standardized residuals are computed and used to reconstruct the volatility series X_t , inputs of the HAR-RV model of Corsi (2009).

$$X_{t+1} = \beta_0 + \beta_d X_t + \beta_w X_{t-5,t} + \beta_m X_{t-22,t} + \zeta_{t+1}$$

With ζ is a random error.

References

1. Aielli, G. P., 2006. Consistent estimation of large scale dynamic conditional correlations. University of Florence Working Paper.

2. Aldridge, 2009. High Frequency Trading: A Practical Approach Guide to Algorithmic Strategies and Trading System. John Wiley and Sons, pp.106

3. Amendola, Braione, Candila and Storti,2020. A Model Confidence Set approach to the combination of multivariate volatility forecasts. International Journal of Forecasting, ISSN 0169-2070.

4. Amendola, Braione, Candila et al.,2020. A Model Confidence Set approach to the combination of multivariate volatility forecasts. International Journal of Forecasting.

5. Andersen, Bollerslev, Diebold & Ebens, Heiko, 2001. The Distribution of Stock Return Volatility. Journal of Financial Economics. 61. 43-76.

6. Andersen, Bollerslev, Diebold & Labys, 2000. Great Realizations. Risk 13 (3), 105-108

7. Andersen, Torben G. and Benzoni, Luca, Realized Volatility, 2008. FRB of Chicago Working Paper No. 2008-14. Available at SSRN: https://ssrn.com/abstract=1092203 or http://dx.doi.org/10.2139/ssrn.1092203

8. Atchison, Butler and Simonds, 1987. Nonsynchronous Security Trading and Market Index Autocorrelation. The Journal of Finance, 42(1), 111.

9. Baba, Y., R.F. Engle, D. Kraft, and K. Kroner,1990. Multivariate simultaneous generalized ARCH.

10. Bandi, F., & Russell, J. , 2008. Microstructure Noise, Realized Variance, and Optimal Sampling. The Review of Economic Studies, 75(2), 339-369.

11. Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. and Shephard, N., 2009. Realized kernels in practice: trades and quotes. The Econometrics Journal, 12: C1-C32.

12. Bauwens, Braione & Storti, Giuseppe, 2016. Forecasting Comparison of Long-Term Component Dynamic Models for Realized Covariance Matrices. Annales d'économie et de statistique.

13. Bauwens, Laurent & Rombouts, Jeroen, 2003. Multivariate GARCH models: a survey. Journal of Applied Econometrics.

14. Bauwens, Storti and Violante, Francesco, 2012. Dynamic conditional correlation models for realized covariance matrices. CORE Discussion Papers 2012060.

15. Becker, Ralf & Clements, Adam & O'Neill, Robert., 2018. A Multivariate Kernel Approach to Forecasting the Variance Covariance of Stock Market Returns. Econometrics. 6. 7. 10.3390/econometrics6010007.

16. Bera and Kim,2002. Testing constancy of correlation and other specifications of the BGARCH model with an application to international equity returns. Journal of Empirical Finance,Volume 9, Issue 2,Pages 171-195,ISSN 0927-5398

17. Billio & Caporin & Gobbo, 2006. Flexible Dynamic Conditional Correlation Multivariate GARCH Models for Asset Allocation. Applied Financial Economics Letters. 2. 123-130.

18. Black, 1976. Studies in stock price volatility changes. Proceedings of the 1976 Meeting of the American Statistical Association, Business and Economics Statistics Section, 177-181.

19. Bollerslev, T., R. Engle, and J. Wooldridge, 1988.A Capital Asset Pricing Model with Time Varying Covariances. Journal of Political Economy96, 116-131

20. Bollerslev, Tim, 1990. Modelling the Coherence in Short-run Nominal Exchange Rates: A Multivariate Generalized ARCH Model, The Review of Economics and Statistics, MIT Press, vol. 72(3), pages 498-505.

21. Bonato, Caporin & Ranaldo, 2012. Forecasting Realized (Co)Variances with a Bloc Structure Wishart Autoregressive Model, Working Papers on Finance 1211, University of St. Gallen, School of Finance

22. Boudt, Kris & Laurent, Sébastien & Lunde, Asger & Quaedvlieg, Rogier. 2014. Positive Semidefinite Integrated Covariance Estimation, Factorizations and Asynchronicity. SSRN Electronic Journal. 10.2139/ssrn.2383871.

23. Brownlees and Gallo, 2006. Financial econometric analysis at ultra-high frequency: Data handling concerns. Computational Statistics & Data Analysis, 51(4):2232–2245.

24. Brownlees, Christian T. and Cipollini, Fabrizio and Gallo, Giampiero M.,2011. Multiplicative Error Model.

25. Callot, Laurent & Kock, Anders & Medeiros, Marcelo, 2016. Modeling and Forecasting Large Realized Covariance Matrices and Portfolio Choice. Journal of Applied Econometrics.

26. Caporin and McAleer, 2013. Ten Things You Should Know About the Dynamics Conditional Correlation Representation. Working Papers in Economics 13/21, University of Canterbury, Department of Economics and Finance

27. Cappiello, Engle and Sheppard, 2006. Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns. Journal of Financial Econometrics, Volume 4, Issue 4, Pages 537-572.

28. Chiriac and Voev, Valeri, 2011. Modelling and forecasting multivariate realized volatility, Journal of Applied Econometrics, 26, issue 6, p.922-947.

29. Chou, Ray Yeutien, 2005. Forecasting Financial Volatilities with Extreme Values: The Conditional Autoregressive Range (CARR) Model. Journal of Money, Credit and Banking, vol. 37, no. 3

30. Cipollini, Fabrizio, Engle, Robert and Gallo, Giampiero, 2013. Semiparametric Vector MEM, Journal of Applied Econometrics, 28, issue 7, p. 1067-1086.

31. Corsi, 2009. A Simple Approximate Long-Memory Model of Realized Volatility. Journal of Financial Econometrics, Society for Financial Econometrics, vol. 7(2), pages 174-196, Spring.

32. Corsi, Fulvio & Mittnik, Stefan & Pigorsch, Christian & Pigorsch, Uta., 2008. The Volatility of Realized Volatility. Econometric Reviews. 27. 46-78. 10.1080/07474930701853616.

33. Coshall, John., 2009. Combining volatility and smoothing forecasts of UK demand for international tourism. Tourism Management. 30. 495-511.

Demir, Serdar & Toktamis, Oniz. (2010). On The Adaptive Nadaraya-Watson Kernel Regression Estimators. Hacettepe Journal of Mathematics and Statistics. 39. 429-437.

34. Engle, R. , 2002. New frontiers for arch models. J. Appl. Econ., 17: 425-446. doi:10.1002/jae.683

35. Engle, R. F., & Kroner, K. F. ,1995. Multivariate Simultaneous Generalized ARCH. Econometric Theory, 11, 122-150.

36. Engle, Robert & Colacito, Riccardo, 2006. Testing and Valuing Dynamic Correlations for Asset Allocation. Journal of Business & Economic Statistics. 24. 238-253. 10.1198/073500106000000017.

37. Engle, Robert & Kelly, Bryan., 2011. Dynamic Equicorrelation. Journal of Business & Economic Statistics. 30. 10.2139/ssrn.1354525.

38. Engle, Robert & Ledoit, Olivier & Wolf, Michael., 2017. Large Dynamic Covariance Matrices. Journal of Business & Economic Statistics. 37. 10.1080/07350015.2017.1345683.

39. Engle, Robert F., and Victor K. Ng., 1993. Measuring and Testing the Impact of News on Volatility. The Journal of Finance, vol. 48, no. 5, 1993, pp. 1749–1778.

40. Engle, Robert F.,2002a. Dynamic Conditional Correlation - a Simple Class of Multivariate GARCH Models. UCSD Economics Discussion Paper No. 2000-09.

41. Engle, Robert, Gallo, Giampiero and Velucchi, Margherita, 2012. Volatility Spillovers in East Asian Financial Markets: A Mem-Based Approach. The Review of Economics and Statistics, 94, issue 1, p. 222-223.

42. Escribano and Sucarrat, 2018. Equation-by-Equation Estimation of Multivariate Periodic Electricity Price Volatility.

43. Eustaquio, John & Mapa, Dennis & Mindanao, M.C. & Paz, N.I.,2015. Forecasting time-varying correlation using the DCC model. 64. 53-74.

44. Falkenberry, 2002). High Frequency Data Filtering: A Review of the issues associated with maintaining and cleaning a high frequency financial database. Tick Data Inc.

45. Fresoli, Diego & Ruiz, Esther., 2015. The uncertainty of conditional returns, volatilities and correlations in DCC models. Computational Statistics & Data Analysis. 100. 10.1016/j.csda.2015.03.017.

46. Ghysels, Eric, Gourieroux, Christian and Jasiak, Joann, 2004. Stochastic volatility duration models. Journal of Econometrics, 119, issue 2, p. 413-433.

47. Glosten, Lawrence R, Jagannathan, Ravi and Runkle, David E, 1993. On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. Journal of Finance, 48, issue 5, p. 1779-1801,

48. Goetzman, William, Li, Lingfeng and Rouwenshorst, K., 2005.Long Term Global Market Correlations. The Journal of Business, 78, issue 1, p. 1-38

49. Golosnoy, Vasyl, Gribisch, Bastian and Liesenfeld, Roman, 2012. The conditional autoregressive Wishart model for multivariate stock market volatility. Journal of Econometrics, 167, issue 1, p. 211-223.

50. Gourieroux, C. & Jasiak, J. & Sufana, R., 2009. The Wishart Autoregressive process of multivariate stochastic volatility. Journal of Econometrics, Elsevier, vol. 150(2), pages 167-181.

51. Granger, Clive & Poon, Ser-Huang., 2001. Forecasting Financial Market Volatility: A Review. SSRN Electronic Journal. 10.2139/ssrn.268866.

52. Granger, Clive & Poon, Ser-Huang., 2003. Forecasting Volatility in Financial Markets: A Review. Journal of Economic Literature. 41. 478-539.

53. Gregoriou, 2009. Stock Market Volatility

54. Gregory H. Bauer, Keith Vorkink, 2011. Forecasting multivariate realized stock market volatility. Journal of Econometrics,Volume 160, Issue 1,2011,Pages 93-101.

55. Hamilton, J. D., 1994. " Time series analysis". Princeton: Princeton University Press

56. Hansen and Lunde, 2005. A forecast comparison of volatility models: does anything beat a GARCH (1, 1). Journal of Applied Econometrics, 20(7).

57. Hansen, P., & Lunde, A., 2006. Realized Variance and Market Microstructure Noise. Journal of Business & Economic Statistics,24(2), 127-161.

58. Hansen, Peter & Lunde, Asger & Nason, James., 2005. Model confidence sets for forecasting models.

59. Huang, Xin and Tauchen, George, 2005. The Relative Contribution of Jumps to Total Price Variance.Journal of Financial Econometrics, 3, issue 4, p. 456-499

60. Huang, Zhuo & Liu, Hao & Wang, Tianyi., 2015. Modeling long memory volatility using realized measures of volatility: A realized HAR GARCH model. Economic Modelling. 52. 10.1016/j.econmod.2015.10.018.

61. Jim E. Griffin & Roel C. A. Oomen ,2008. Sampling Returns for Realized Variance Calculations: Tick Time or Transaction Time? Econometric Reviews, 27:1-3, 230-253.

62. Jin, Xin & Maheu, John., 2010. Modeling Realized Covariances and Returns. University of Toronto, Department of Economics, Working Papers. 11. 10.1093/jjfinec/nbs022.

63. Karanasos, M., Xu, Y., & Yfanti, S., 2018. Stylized Facts for Extended HEAVY / GARCH models and MEM: the importance of asymmetries, power transformations, long memory, structural breaks and spillovers.

64. King & Wadhwani, 1990. Transmission of Volatility Between Stock Markets. Review of Financial Studies. 3.

65. Kirchler, Huber, 2007. Fat tails and volatility clustering in experimental asset markets. Journal of Economic Dynamics and Control,Volume 31, Issue 6,Pages 1844-1874

66. Kostrov, Alexander & Tetereva, Anastasija, 2019. Forecasting realized correlations: a MIDAS approach.

67. Lin, Robert F.Engle, Ito, 1994. Do Bulls and Bears Move Across Borders? International Transmission of Stock Returns and Volatility, The Review of Financial Studies Vol. 7, No. 3 I, pp. 507-538

68. Lux, T., Marchesi, M., 2000. Volatility clustering in financial markets, A microsimulation of interacting agents. International Journal of Theoretical and Applied Finance 3, 675-702.

69. Mandelbrot, 1963. The Variation of Certain Speculative Prices, The Journal of Business, University of Chicago Press, vol. 36, pages 394-394.

70. McAleer & Medeiros, 2008. Realized Volatility: A Review. Econometric Reviews. 27.

71. Moran, J. & Solomon, P., 2017. Volatility in High-Frequency Intensive Care Mortality Time Series: Application of Univariate and Multivariate GARCH Models. Open Journal of Applied Sciences. 07. 385-411.

72. Naimoli, Antonio & Storti, Giuseppe, 2019. Heterogeneous component multiplicative error models for forecasting trading volumes. MPRA Paper 93802, University Library of Munich, Germany.

73. Nelson, Daniel B., 1991. Conditional Heteroskedasticity in Asset Returns: A New Approach. Econometrica, vol. 59, no. 2, pp. 347–370.

74. Oomen, R., 2006. Properties of Realized Variance under Alternative Sampling Schemes. Journal of Business & Economic Statistics, 24(2), 219-237.

75. Pakel, C., Shephard, D., Sheppard, K. and Engle, R.F., 2014. Fitting Vast Dimensional Time-Varying Covariance Models. New York University, Mimeo, New York.

76. Palandri, Alessandro., 2006. Consistent Realized Covariance for Asynchronous Observations Contaminated by Market Microstructure Noise. SSRN Electronic Journal. 10.2139/ssrn.2727826.

77. Pedersen, R. S., & Rahbek, A., 2014. Multivariate variance targeting in the BEKK-GARCH model. The Econometrics Journal, 17(1), 24–55.

78. Qiu, Y., Zhang, X., Xie, T., & Zhao, S., 2019. Versatile HAR model for realized volatility: A least square model averaging perspective. Journal of Management Science and Engineering.

79. Qiu, Zhang, Xie and Zhao,2019. Versatile HAR model for realized volatility: A least square model averaging perspective. Journal of Management Science and Engineering,Volume 4, Issue 1,Pages 55-73.

80. Sanfelici and Mancino, 2008. Estimating Covariance via Fourier Method in the Presence of Asynchronous Trading and Microstructure Noise. Department of Economics, Parma University (Italy), Economics Department Working Papers.

81. Silvennoinen & Teräsvirta, 2012. Modelling conditional correlations of asset returns: A smooth transition approach. CREATES Research Papers 2012-09, Department of Economics and Business Economics, Aarhus University.

82. Sylvain Barde, 2015. A fast algorithm for finding the confidence set of large collections of models. Studies in Economics 1519, School of Economics, University of Kent.

83. Taylor, 2008. Modelling Financial Time Series

84. Tse and Tsui, 2002. "A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations". Journal of Business & Economic Statistics 20 (3), 351-362

85. Tse, 2000. A test for constant correlations in a multivariate GARCH model, Journal of Econometrics,Volume 98, Issue 1,Pages 107-127,ISSN 0304-4076

86. Wang, Bauwens and Hsiao, 2013. Forecasting a Long Memory Process Subject to Structural Breaks. CAFE Research Paper No. 13.01.

87. Wilson, Andrew & Ghahramani, Zoubin., 2010. Generalised Wishart Processes. Proceedings of the 27th Conference on Uncertainty in Artificial Intelligence, UAI 2011.

88. Wu and Xiao, 2011. Covariance Matrix Estimation in Time Series.

89. Yu, Philip & Li, Wai & Ng, Fo Chun., 2015. The Generalized Conditional Autoregressive Wishart Model for Multivariate Realized Volatility. Journal of Business & Economic Statistics. 00. 10.1080/07350015.2015.1096788.

90. Zhang, L., Mykland, P., & Aït-Sahalia, Y., 2005. A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data. Journal of the American Statistical Association,100(472), 1394-1411.

91. Zhou, Jiayuan & Jiang, Feiyu & Zhu, Ke & Li, Wai.,2019. Time series models for realized covariance matrices based on the matrix-F distribution