SAPIENZA
UNIVERSITÀ DI ROMA

Sapienza University of Rome

Department of Methods and Models for Economics, Territory and Finance

PhD in Models for Economics and Finance

THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Copula Models for Financial Time Series

Thesis Advisor

Prof. Brunero Liseo

Candidate

Valentina Bruno

2022/2023

Ai miei Nonni che mi hanno insegnato tanto,

a Nonna Nunzia, al suo sorriso contagioso e alla sua pazienza infinita,

alla determinazione e forza di Nonna Mariolina,

a mio Nonno Gianni che saltava nel corridoio sentendomi entrare dopo ogni esame urlando 30 !!!

ai miei genitori ed ai loro sacrifici che non dimenticherò mai,

a mio fratello, piccolo partner in crime, che tifa sempre per me,

al mio fidanzato che sa come convogliare le mie 1000 idee verso la meta sostenendomi sempre,

a zio Isidoro che mi chiamava prima e dopo ogni esame,

ad Eleonora, che dal 1993 mi tiene per mano,

a Letizia per avermi fatto credere in me quando non ci credevo abbastanza

ad Alice perché mi capisce con uno sguardo,

a Daniele che mi ha permesso di credere nei miei sogni

al Professor Liseo per tutto l'aiuto e per la sua ineguagliabile disponibilità

GRAZIE!

Copula model for financial time series

Valentina Bruno

May 3, 2023

Contents

Introduction

The present thesis has two objectives:
a. to provide a clustering methodology with a consequent study of the contagion that allows an efficient diversification of the risk;
b. to build a model for the estimation of market risk that adapts to financial data and their peculiarities more efficiently than the usual models built through the use of Normal distribution. A sample of the nine most capitalized stock indices has been studied, such as: Dow Jones, S&P500, Nasdaq 100, FTSE 100, Nikkei 225, SSE Composite, SZSE Component, Euronext 100, HANG SENG. The chosen time horizon is about ten years, more precisely from January 2, 2012 to October 11, 2022. The work is divided into three chapters, of which the first chapter sets out the theoretical foundations on which the analysis under discussion is based, analytically presenting the empirical properties of financial data, such as: the presence of non-independent and identically distributed yield series; the significant correlation of returns squared; the near zero value of the above conditional expected values; the characteristic leptocurtic tails and the presence of clusters for the extreme values. The importance of an accurate estimate of rare events is underlined and the extreme values theory is dealt with, focusing on the distribution of generalized extreme values (GEV), on the maximum domain of attraction, on the "Block Maxima" method and over threshold exceedances. We proceed by providing the definition and explaining the mathematical properties of the copula function; we describe, moreover, various types of copulas including the skew t copula, the Vine copula and the elliptic copulas, to this last category belong the Gaussian copula and the Student t copula. In addition, the calibration methods of the parameters of the function under consideration, such as the Maximum Likelihood Method (ML), the Margin Inference Functions Method (IFM) and the Method of Maximum Canonical Likelihood (CML). A quick description of the sample being analysed is also provided. The second

chapter presents an analysis of the behavior of time series in risky scenarios in order to enable the implementation of performing portfolio diversification strategies. Through a methodology that is articulated in four different phases: to find the model that is well adapted to the data in analysis, to measure the dependence of tail, to create the matrix of dissimilarity, to construct of the clusters that allow the study of the phenomenon of the contagion in extreme events. Three different thresholds were chosen to study the correlation of the nine indices in risky scenarios and in stable market conditions. The third and final chapter explains accurately all the operations carried out that have finally led to the construction of the desired model, starting from the daily closing prices of the nine indices constituting the portfolio. The combined use of extreme value theory and copulas: t , skew t and Vine leads to a market risk modelling approach that stands out from traditional risk management models. They assume conditional normality for logarithmic returns on financial assets or risk factors despite empirical evidence that yield distributions are characterized by leptocurtic tails. The main objective of this study was to obtain a model consistent with this empirical evidence. Finally, the obtained portfolio index returns are simulated and the Value at Risk is calculated with relative back-testing to test the goodness of the presented model.

Chapter 1

Theoretical notions

1.1 Abstract

The following chapter is aimed at the explanation of the theoretical foundations that will be applied in the following practical elaboration . It first presents and explains the empirical properties of the financial data, below is proposed what is the main theme of this paper, this is the importance of estimating rare events with further study of the theory of extreme values. Since the application model object of this analysis is obtained by combining the theory of extreme values and copulas, it was considered more than appropriate to provide a theoretical explanation of the tool used for the multivariate modelling phase, explaining also the possible calibration methods.

1.2 Empirical properties of financial data

Financial time series have common characteristics that can be traced back to the following list of "stylized facts". This expression indicates a collection of empirical observations, and inferences drawn from these observations, which seem to apply to most daily time series of changes in the risk factor, such as logarithmic returns of shares, indices, foreign exchange transactions and commodity prices. These observations are so deeply rooted in econometric experience that they have been elevated to the state of "facts". They often persist even for longer time intervals, such as weekly or monthly returns, or for shorter time intervals, such as intra-daily returns.

1.2.1 Stylized facts

Below a list of the "Stylized Facts" that characterize the financial historical series :
(1) historical yield series are not independent and identically distributed (i.e.d), even if they show minimal serial correlation;
(2) absolute or square yield series show significant serial correlation;
(3) expected conditional yields are close to zero;
(4) volatility appears to vary over time;
(5) yield series are characterised by heavy or leptocytotic tails;
(6) extreme yield values appear in clusters.
Let $X_1, ..., X_n$ be the series of returns and it is assumed that the latter have been calculated as a logarithmic price difference $(S_t)_{t=0,1,...,n}$, so $X_t = ln(S_t/S_{t-1})$,$t = 1, ..., n$.

Volatility Clustering

Empirical evidence concerning the first two stylized facts is shown in figure 1.1, where (a) it shows 2608 logarithmic returns for the DAX stock index over a ten-year horizon, from January 2, 1985 to December 30, 1994, a period that includes both the 1987 stock market crash and the reunification of Germany in 1989. Figures (b) and (c) show simulated i.e.d. data series originating from a Normal model and a t model of Student, respectively. In both cases the parameters of the model have been set to adapt it to the real yield data using the maximum likelihood method under the assumption that the data under examination are independent and identically distributed. In the case of normal, this means that they were simply simulated i.e.d. data with distribution $N(\mu, \sigma^2)$ where

$$\mu = \bar{X} = n^{-1} \sum_{i=1}^{n} X_i$$

and

$$\sigma^2 = n^{-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

In the case of the Student t likelihood is maximised numerically and the parameter of estimated degrees of freedom is $v = 3,8$. It is evident from the figures that the data simulated by a normal distribution are clearly very

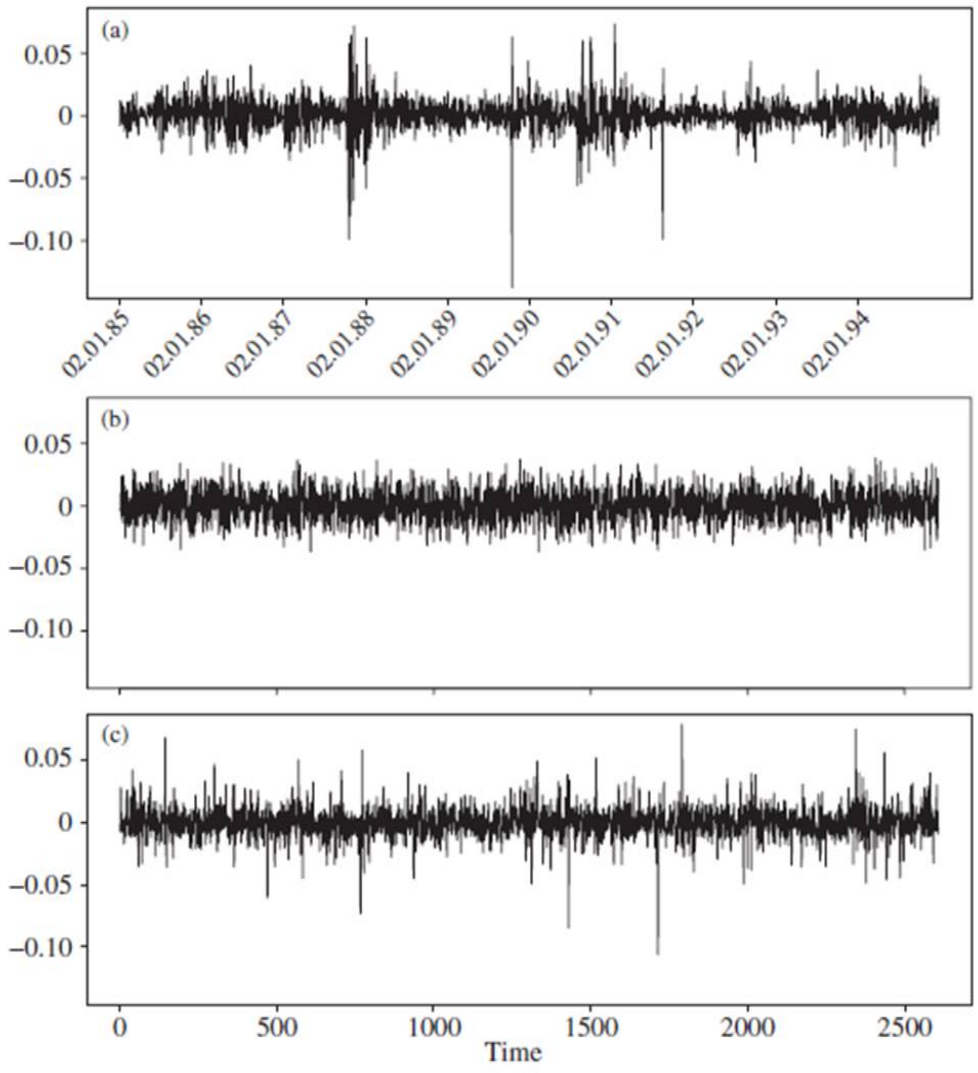

Figure 1.1: Real and simulated returns of the DAX stock index

different from the DAX yield and do not show the same range of extreme events, while the t model of Student can generate extreme values comparable to real data. A more accurate observation reveals that real returns show a phenomenon known as "volatility clustering", which is not present in the simulated series. The term "Volatility clustering" indicates the tendency for extreme yield values to be followed by other extreme values, although not necessarily of the same sign.

Figure 1.2 presents the correlograms of the pure data and their absolute values for all three previously presented cases, namely: real logarithmic yields of the Dax index, simulated i.e.d. returns obtained using a Normal model in

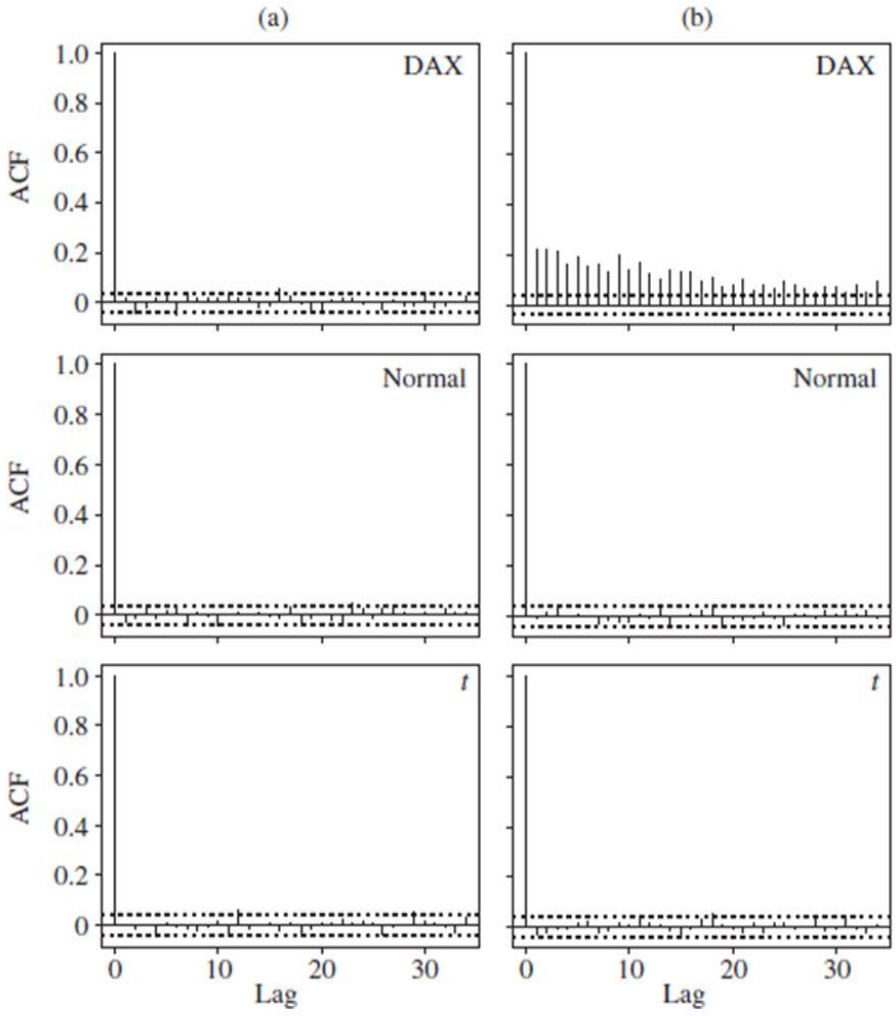

Figure 1.2: Autocorrelation's functions

the first case and with a Student t model in the second. The correlogram graphically shows estimates of serial correlation. Although there is little evidence of serial correlation in the raw data for all three datasets, the absolute values of the real financial data seem to demonstrate the presence of serial dependence or "volatility clustering". Note that more than 5 % of the estimated correlations are outside the dashed lines, indicating potential serial correlation. This serial dependence in absolute returns should be equally evident in squared yield values and seems to confirm the presence of volatility clustering. We conclude that, although there is no evidence against the hypothesis of independence and identical distribution for truly i.e.d., there is strong evidence against this hypothesis for the DAX yield data examined. Moreover, if there are serial dependencies in financial performance data, then the question

arises: to what extent can this dependence be used to make predictions about the future? This is the argument of the third and fourth stylized facts. It is very difficult to predict performance in the subsequent period based solely on historical data. This difficulty in predicting future returns is part of the evidence of the well-known assumption of efficient financial markets that prices react quickly to reflect all available market information. In empirical terms the lack of predictability of returns is shown by the lack of a serial correlation in the historical series of raw or pure returns. For some data samples it is sometimes possible to observe evidence of correlation to the first lag. A minimal positive correlation to the first lag might suggest that there is a clear tendency for a yield with a particular sign, positive or negative, to be followed in the following period by a yield with the same sign. However, this does not appear in the DAX index data examined, which suggests that our best estimate for tomorrow's returns based on observations available to date is zero. Volatility is often formally modelled as the conditional standard deviation of financial returns calculated on historical data, and although the conditional expected values are consistently close to zero. Volatility clustering suggests that conditional standard deviations are constantly changing in a partially predictable way. If we know that yields have been very large in recent days, because of the excitement of the market, then we have reason to believe that the distribution from which tomorrow's yield is calculated should have great variance.

Leptocurtic Distribution

Normal distribution has frequently been shown to be a bad model for daily, weekly and even monthly yields. This statement can be confirmed by using several well-known normality tests, including the Q-Q plot, performed with a normal distribution, as well as a number of formal numerical tests. The Q-Q plot, where the two Q's in the name stand for quantile, is a standard visual tool that shows the relationship between the empirical quantiles of data and the theoretical quantiles of a reference distribution. A lack of linearity in the Q-Q plot is interpreted as evidence against the assumed reference distribution. The common numerical tests we have mentioned include those of Jarque and Bera, Anderson and Darling, Shapiro and Wilk, and D'Agostino. Historically, following the empirical analysis resulting from the application of the tests listed above, it was noted that daily financial returns showed a higher

curtosis than would have been consistent with a normal distribution hypothesis, The distribution of such data has been termed leptocurtic, which means that it proves to be narrower than the normal distribution in the center, but has longer and heavier tails than the latter.Further empirical analyses often suggest that the distribution of daily, weekly, etc financial returns.. have tails that decay slowly according to a power law, unlike the faster exponential decay of the tails of a normal distribution. This means that we tend to see far more extreme values than would be expected in that dataset.

1.3 Why estimate rare events?

In recent decades, empirical evidence has shown that currency crises, stock market collapses and bank failures are not so rare globally. For traditional models of risk measurement it has become of fundamental importance to be able to take into account also these so-called "extreme" phenomena, and as such they fall beyond the normal range of available observations. The Extreme Values Theory (EVT) offers a solid theoretical foundation to build quantitative models suitable to describe "extreme" financial events. For international financial supervisory authorities, as well as for risk managers of financial institutions, it is essential to be able to rely on estimation models capable of considering the impact on the value of the portfolios of catastrophic events. That is, events that occur with very low probability but the effect of which can be particularly severe and therefore catastrophic for financial portfolios. Events with such characteristics can occur in many real financial situations, for example:

•in market risk management, where the daily determination of the Value at Risk (VarR) of trading book losses due to adverse market movements is addressed;

•credit and operational risk management, where the objective is to determine the capital requirement to protect losses arising from deterioration in the credit quality and default of the portfolio assets or from unforeseen operational problems;

•in insurance risk management. A typical problem in this area is the assessment of the premium required for products that offer protection against catastrophic losses such as excess-of-loss reinsurance deals concluded with

primary insurers.

So whenever it is of interest to study the tail of the distribution of returns or losses of a portfolio of financial assets , for the management of market, credit, operational and insurance risks, the extreme values theory (EVT) is a valuable tool for the correct estimation of risk as theoretically superior to traditional methodologies. The latter inevitably end up by dangerously underestimating the risk of tail, that is, the risk associated with extreme but not so unlikely events, as assumed by the normal distribution hypothesis. In such a context, the extreme values theory makes a valuable contribution to the prediction of the dimension by which a rare event occurs. Such theory allows us to obtain an optimal estimate of the tail using a generalized distribution even when the scarce historical data available do not allow to make hypotheses on the form of the underlying distribution.

1.4 The Extreme Values Theory (EVT)

The extreme values theory is a branch of probability that has given many important results by describing the behavior of maximum and minimum samples , statistics of higher order , such as the k-th major value in a sample, and sample values exceeding high thresholds. In this case the interest in this theory is focused on the application of its results to develop models that best approximate the extreme behaviors of financial risk factors. We chose to focus on the two main types of model for extreme values:

•The most traditional models are the Block Maxima Models: they are models suitable for the largest observations collected from large samples of independent and identically distributed observations.

•A more modern and powerful group of models are those that measure threshold exceedances : such models are suitable for all large observations exceeding a certain high level. They are generally considered to be the most useful for practical applications ,this is due to their more efficient use of data on extreme results. Over-threshold models can be incorporated into an elegant process framework that directs simultaneously their occurrence over time and the magnitude of losses exceeding the threshold. This is the so-called peak over-threshold model (POT Model). The POT model serves as a starting point for developing more dynamic descriptions of the occurrence

and magnitude of extreme values.

1.4.1 Maxima: Generalized distribution of extreme values

It is considered a sequence of independent and identically distributed random variables $(X_i)_{i\in_N}$ which represent financial losses. They may have different interpretations, such as operating losses, insurance losses, and losses in a credit portfolio over fixed time intervals. The assumption of independence was later set aside and the random variables are considered to form a stationary historical series of dependent losses. They could be (negative) profits from a single security investment , an index or an investment portfolio. The role of the generalized extreme value distribution (GEV) in the theory of extreme values is similar to that of the Normal distribution in the central limit theory for sums of random variables. Assuming that the random variables $X_1, X_2 \ldots$ are independent and identically distributed with a finite variance and writing

$$S_n = X_1 + X_2 + \cdots + X_n \tag{1.1}$$

as the sum of the first n variables, the standard version of the Central Limit Theorem (CLT) states that the properly normalized sum $(S_n - a_n)/b_n$ converges in the distribution to the distribution of a standard normal for n going to infinity. Proper normalization uses sequences of normalized constants (a_n) e (b_n) defined as follows:

$$(a_n) = nE(X_1) \tag{1.2}$$

and

$$(b_n) = \sqrt{nvar(X_1)} \tag{1.3}$$

In arithmetic terms we have:

$$\lim_{x\to\infty} P\left(\frac{S_n - a_n}{b_n} \leq x\right) = \Phi(x) \tag{1.4}$$

with $x \in R$, where Φ is the CDF of a standard normal distribution.

The Classical extreme values theory concerns limited distributions for normalized maxima. Here we denote the maximum value of n random variables i.e.d. $X_1, X_2, ..., X_n$ as:

$$M_n = max(X_1, ..., X_n) \tag{1.5}$$

Definition : Generalized distribution of extreme values

The definition of the generalized extreme value distribution is given by

$$H_\xi(x) = \begin{cases} \exp(-(1+\xi x)^{-1/\xi}), & \xi \neq 0, \\ \exp(-e^{-x}), & \xi = 0, \end{cases} \tag{1.6}$$

where $1 + \xi n > 0$. A triparametric family is obtained by defining

$$H_{\xi,\mu,\sigma} = H_{\xi(x-\mu)/\sigma} \tag{1.7}$$

with the location parameter $\mu \in R$ and the scale parameter $\sigma > 0$. The parameter ξ is known as the shape parameter of the generalized distribution of extreme values (GEV). The distribution of extreme values is said to be generalized because the parametric form includes three types of distribution that are known with different names depending on the value assumed by ξ: when $\xi > 0$ the distribuzioen is known as Fréchet's distribution; when $\xi = 0$ it is called the Gumbel's distribution; finally when $\xi < 0$ is called the Weibull's distribution. We also note that for a fixed value of x we have that the

$$\lim_{\xi \to 0} H_\xi(x) = H_0(x) \tag{1.8}$$

For this reason, the parameterization present in the definition of the distribution of extreme values presented above is continuous in ξ, which facilitates the use of this distribution in statistical modelling.

The figure 1.3 shows the distribution function and the density of the generalized distribution of extreme values (GEV) for the three cases $\xi = 0,5$; $\xi = 0$, $\xi = -0,5$, which correspond respectively to the Fréchet, Gumbel and Weibull distributions. It is noted that the Weibull distribution is a distribution with "short tail" and limited right field of existence. The distributions of Gumbel and Fréchet have an unlimited field of existence on the right, but the decay of the tail of Fréchet's distribution is much slower than that of Gumbel's

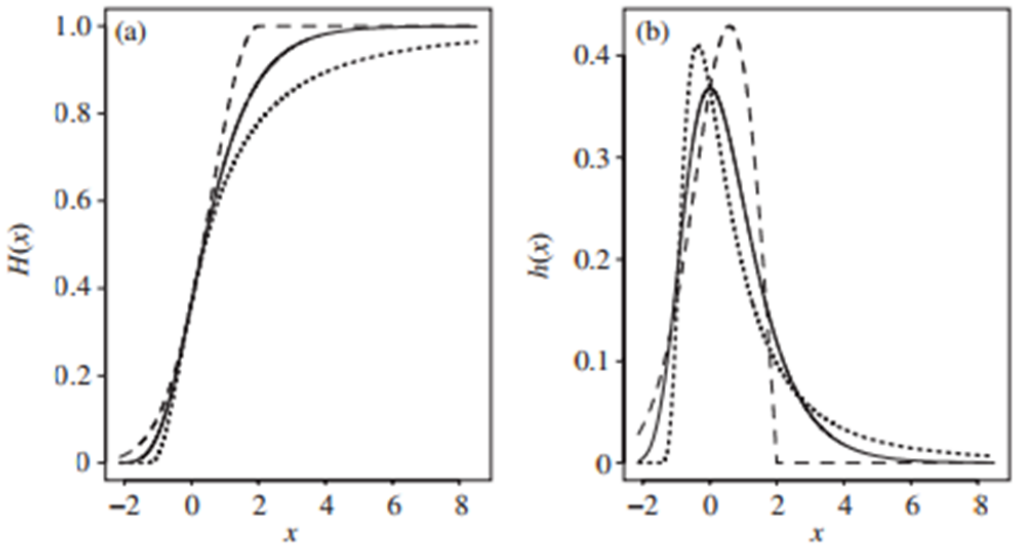

Figure 1.3: Distribution Function and Density of the GEV : Fréchet distribution with $\xi = 0,5$ represented by the dotted line; Gumbel distribution with $\xi = 0$ represented by the solid line; Weibull distribution with $\xi = -0,5$ represented by the dashed line (McNeil,Frey,Embrechts(2015))

distribution. Suppose that the maxima M_n of independent and identically distributed random variables converges for $n \to \infty$ under appropriate standardisation. Recalling that $P((M_n \leq x) = F^n(x)$ it can be observed that this convergence means that there are sequences of real constants (d_n) and (c_n), where $c_n > 0$ for each n, therefore:

$$\lim_{n\to\infty} P((M_n - d_n)/c_n \leq x) = \lim_{n\to\infty} F^n(c_n x + d_n) = H(x) \tag{1.9}$$

for certain nondegenerative H(x) distribution functions. The role of generalized extreme-value distribution in the study of maxima is formalized by the following definition and theorem:

Definizione : Maximum domain of attraction

If $\lim_{n\to\infty} P((M_n - d_n)/c_n \leq x) = \lim_{n\to\infty} F^n(c_n x + d_n) = H(x)$ for certain non degenerative distribution functions of H, then it is said that F is in the maximum attraction domain of H , so you can write that $F \in MDA(H)$.

Theorem: Fisher-Tippet, Gnedenko

If $F \in MDA(H)$ for certain nondegenerative distribution functions of H , so H must necessarily be a distribution of the type H_ξ , that is a generalized extreme value distribution (GEV).

Observations:

1) If a convergence of normalized maxima occurs, the type of limit distribution is uniquely determined, even if the position and scale parameters μ and σ of the limit depend on the exact normalization sequences chosen; this is guaranteed by the so-called convergence to the type theorem. It is always possible to choose these sequences so that the limit appears in the standard form H_ξ.
2)Nondegenerative distribution function means a limit distribution that is not concentrated on a single point.

1.4.2 Maximum Domain of Attraction

For most applications it is sufficient to note that essentially all common continuous distributions inherent in the statistical or actuarial sciences are $MDA(H_\xi)$ for certain values of ξ. In this section we consider the problem of the underlying distributions that limit the maxima.

The Fréchet case:

Distributions leading to the limit of Fréchet $H_\xi(x)$ for $\xi > 0$ have a particularly elegant characterization involving slowly or regularly different functions.

Definition: Slowly variable and regularly different functions

(i) A positive and "Lebesgue-measurable" function L on $(0, \infty)$ is slowly variable in ∞ , that is $L \in R_0$ if:

$$\lim_{x \to \infty} \frac{L(tx)}{L(x)} = 1 \qquad t > 0, \tag{1.10}$$

(ii)A positive and "Lebesgue-measurable" function h on $(0, \infty)$ is regularly

variable in ∞ with the index $\rho \in R$ if :

$$\lim_{x\to\infty} \frac{L(tx)}{L(x)} = t^{\rho} \qquad t > 0, \tag{1.11}$$

Slowly variable functions are functions that change relatively slowly with respect to power functions for large x values , an example is the logarithm: $L(x) = ln(x)$. Regularly variable functions are functions that can be represented by power functions multiplied by slowly variable functions, as an example: $h(x) = x^p L(x)$ for some $L \in R_0$

Theorem: Fréchet MDA, Gnedenko

For $\xi > 0$,

$$F \in MDA(H_\xi) \Longleftrightarrow \bar{F}(x) = x^{1/\xi} L(x) \tag{1.12}$$

for some $L \in R_0$, where $\bar{F}(x)$ is the survival function.

This means that the distributions that give rise to the Fréchet case are distributions with tails that are regularly variable functions with a negative variation index. Their tails essentially decay as a function of power and decay speed $\alpha = 1/\xi$ is often referred to as the distribution tail index. These distributions are the most studied distributions in extreme value theory and are of particular interest to financial applications because they are heavy distributions with infinite moments higher. If X is a non-negative random variable whose distribution function F is an element of $MDA(H_\xi)$ for $\xi > 0$, then it is demonstrable that $E(X^k) = \infty$ for $k > 1/\xi$. If, for some small $\epsilon > 0$, the distribution is in $MDA(H_{(1/2)+\epsilon})$, it is an infinite variance distribution and if the distribution is in $MDA(H_{(1/4)+\epsilon})$, then it is a distribution with infinitely fourth moment.

The Gumbel case:

The characterization of distributions in this class is more complicated than the Fréchet class. Distributions in this class have tails that have essentially exponential decay. A positive random variable with element distribution function of $MDA(H_0)$ with finite moments of any positive order, that is $E(Xk) < \infty$ for every $k > 0$. However, there is great variety in distribution tails in this class, for example, both normal and log-normal distributions belong to

the Gumbel class. The normal distribution has a thin tail , while the lognormal distribution has much heavier tails and it is necessary to collect a lot of data from the lognormal distribution before being able to distinguish the behavior of its tail from that of a distribution belonging to the Fréchet class. In financial modeling it is often mistakenly assumed that the only interesting models for financial returns are the powerful tail distributions of the Fréchet class. The Gumbel class is also interesting because it contains many distributions with tails much heavier than normal, although these are not regularly variable power tails. Examples are generalized hyperbolic and hyperbolic distributions (with the exception of the particular boundary case which is Student's t). Other distributions with $MDA(H_0)$ elements include gamma disributions, chi-square, standard Weibull (to be distinguished from the Weibull special case of the GEV distribution), the Benktander distributions type I and II (which are known distributions for actuarial losses) and the Gumbel distribution itself.

The Weibull case:

This is perhaps the least important case for financial modelling, at least in terms of market risk, as distributions in this class all have well-defined right thresholds. While all potential financial and insurance losses are, in practice, limited, we will continue to foster models that have infinite support for modeling potential losses. An exception might be in credit risk modeling, where it is useful to study probability distributions over the $[0, 1]$ range. One characterization of the Weibull class is as follows.

Theorem: Weibull MDA, Gnedenko

For $\xi < 0$, $F \in MDA(H_\xi) \iff x_F < \infty$ and $\bar{F}(x_F - x^{-1}) = x^{1/\xi} L(x)$ for each $L \in R_0$.

1.4.3 Maxima of strictly stationary time series

The standard theory dealt with in the previous section concerns the maxima of the sequences of independent and identically distributed random variables. Bearing in mind the financial time series and their characteristics, looking briefly at the theory for calculating the maxima of strictly stationary time series, we note that the same types of limit distributions can be applied.

In this section we denote with $(X_i)_{i\in N}$ a strictly stationary time series with stationary distribution F and we denote with $(\check{X}_i)_{i\in Z}$ the associated independent and identically distributed process, that is a process with white noise with the same distribution function F. Set $M_n = max(X_1, ..., X_n)$ and $\check{M}_n = max(\check{X}_1, ..., \check{X}_n)$ indicate respectively the maxima of the original historical series and of that i.e.d.

For many processes $(X_i)_{i\in N}$ it can be shown that there is a real number θ in a range of $(0, 1]$ such that:

$$\lim_{n\to\infty} P(\frac{\check{M}_n - d_n}{c_n} \leq x) = H(x) \tag{1.13}$$

for a value of H(x) limited and not degenerative if and only if

$$\lim_{n\to\infty} P(\frac{M_n - d_n}{c_n} \leq x) = H^\theta(x) \tag{1.14}$$

For such processes this θ value is known as the "extreme index" of the process, not to be confused with the tail index of distributions in the Fréchet class. A formal definition is more technical, but the basic ideas behind (1.13) and (1.14) are easy to explain. For processes with an extreme index, the normalized maxima converge in distribution as long as the maxima of the associated i.i.d. processes converge in distribution i.e., provided that the underlying distribution F is in $MDA(H_\xi)$ for some ξ. Moreover, since $H^\theta_\xi(x)$ can be easily verified as a distribution of the same type of $H_\xi(x)$, the limit distribution of the normalized maxima of dependent series is an GEV distribution with exactly the same ξ parameter as the limit for independent and identically distributed data; only the position and scale of the distribution can change. Writing $u = c_n x + d_n$ we observe that, for n quite large ,(1.13) and (1.14) imply that:

$$P(M_n \leq u) \approx P^\theta(\check{M}_n \leq u) = F^{n\theta}(u) \tag{1.15}$$

so for large values of u the probability distribution of the maximum value of n observations from the historical series with extreme index θ can be approximated by the distribution of the maximum of $n\theta < n$ observations of the associated i.i.d. series. Somehow, $n\theta$ can be seen as an approximation of the number of independent cluster observations in n observations and θ is often interpreted as the reciprocal of the mean cluster size. Not all strictly stationary processes have an extreme index but, for the types of time series

processes that interest us for financial modeling, there is generally an extreme index. Basically, we just have to distinguish between cases where $\theta = 1$ and cases where $\theta < 1$: for the first case there is no high-level cluster trend and the maxima calculated on large historical series samples behave exactly like the maxima calculated on samples of similar amplitude i.i.d. For the second case, we need to be aware of the tendency of extreme values to form clusters.
• The rigorous white noise processes, independent and identically distributed random variables have extreme index $\theta = 1$;
• ARMA processes with white noise Gaussian have $\theta = 1$. However, if the distribution function has values in $MDA(H_\xi)$ for $\xi > 0$, then $\theta < 1$;
• ARCH and GARCH processes have $\theta < 1$. The latter is particularly relevant for our financial applications, as the ARCH and GARCH processes provide good approximation models for many time series of financial returns.

1.4.4 The method "Block Maxima"

Construction of a generalized extreme distribution: Suppose you have data from the F distribution of an unknown underlying, which we assume to be in the attraction domain of a distribution of extreme values H_ξ for some ξ. If the data are realizations of i.i.d. variables, or variables of a process with an extreme index like GARCH, the implication of the theory is that the true distribution of the n-th maximum block M_n can be approximated to fairly large n values from a three-parameter GEV distribution $H_{\xi,\mu,\sigma}$. We use this idea by applying the distribution GEV $H_{\xi,\mu,\sigma}$ to the data of the nth block M_n. Of course we need repeated observations of the nth maximum block, and suppose the data can be divided into m blocks of size n. This method makes more sense when there are natural ways to block data. In fact, it originates from hydrology, where, for example, the daily measurement of water levels can be divided in annual blocks and annual maxima are recorded. Similarly, we will consider financial applications where daily returns data are divided into annual or semi-annual blocks and the maximum daily losses included in those blocks are analysed. We denote the maximum block of the j-th block with M_{nj}, so our data is $M_{n1}, ..., M_{nm}$. The generalized extreme value distribution (GEV) can be applied using various methods, including the maximum likelihood. To apply this methodology it is necessary to assume that the size of the nth block is large enough so that, regardless of whether the underlying data are dependent or not, observations of block maxima can be

considered independent. An alternative is the method of weighted moments of probability.

In this case writing $h_{\xi,\mu,\sigma}$ for the density of the GEV distribution the "log-likelihood" is easily calculated as follows:

$$l(\xi,\mu,\sigma;M_{n1},...,M_{nm}) = \sum_{i=1}^{m} lnh_{\xi,\mu,\sigma}(M_{ni}) =$$

$$-mln\sigma - (1+\frac{1}{\xi})\sum_{i=1}^{m} ln(1+\xi\frac{M_{ni}-\mu}{\sigma}) - \sum_{i=1}^{m}(1+\xi\frac{M_{ni}-\mu}{\sigma})^{-1/\xi}$$

which must be maximized based on parameter constraints: $\sigma > 0$ and $1+\xi(M_{ni}-\mu)/\sigma > 0$, $\forall i$. Although this is a problem of irregular likelihood, due to the dependence of the parameter space on the data values, the coherence and asymptotic efficiency of the MLE results can be established for the case where $\xi > -1/2$ using the results in Smith (1985). In determining the number and size of the blocks (m and n, respectively), there is necessarily a compromise: on the one hand, a large value of n leads to a higher accurate approximation of block distribution of maxima via a GEV distribution and at low polarization in parameter estimates. A large value of m gives more data in blocks of maxima for ML estimation and leads to low variance in parameter estimates. Also note that, in the case of dependent data, it may be advisable to use larger blocks than in the case where the data are i.i.d.; the dependency generally has the effect of slowing convergence to the GEV distribution , since the actual sample size is $n\theta$, which is less than n.

Performance levels and stress losses: The GEV model can be used to analyze two related quantities that describe the presence of stress events: on the one hand we can estimate the size of a stress event that happens with a given frequency, the so called problem of performance levels. On the other hand, we can estimate the frequency of a stress event that has a given size, the so called return period problem.

Definition: level of performance:

H denotes the distribution function of the true distribution of the n-th block maximum, the k performance level of the n-th block maximum is $r_{n,k} =$

$q_{1-\frac{1}{k}}(H)$; that is the quantile $1-\frac{1}{k}$ of H.

The performance level of the k element of the n-th block can be interpreted roughly as that level being exceeded on mean once every k n-th blocks. Using our model we would estimate a return level as follows:

$$\hat{r}_{n,k} = H^{-1}_{\hat{\xi},\hat{\mu},\hat{\sigma}}(1-\frac{1}{k}) = \hat{\mu} + \frac{\hat{\sigma}}{\hat{\xi}}((-ln(1-\frac{1}{k}))^{-\hat{\xi}} - 1) \tag{1.16}$$

Definition: return period:

Let H be the distribution function of the real distribution of the n-th block of maximum. The event return period $M_n > u$ is $k_{n,u} = 1/\bar{H}(u)$. Note that the $k_{n,u}$ return period is defined in such a way that the return level of the $k_{n,u}$ n-th block is u. In other words, in the $k_{n,u}$ of the n-block you would expect to see a single block where the u level was passed. If there was a strong tendency for extreme values to cluster, you would expect to have more values that exceed that level within that block. Assuming that H is the distribution function of a generalized distribution of extreme values and using the parameterized model you can estimate the return period with $\hat{k}_{n,u} = 1/\bar{H}_{\hat{\xi},\hat{\mu},\hat{\sigma}}(u)$. Note that both the parameters $\hat{k}_{n,u}$ and $\hat{r}_{n,u}$ are simple features of the estimated parameters of the GEV distribution. In addition to calculating point estimates for these quantities, confidence intervals reflecting the error in parameter estimates should be established of the GEV distribution. A good method is to base these confidence intervals on the likelihood ratio. To do this, the GEV distribution was reparametrized in terms of the amount of interest.

1.4.5 Excess of threshold

The block of maxima method described in section 1.4.4 has as its main defect the waste of data: in order to carry out such analyses only the maximum losses in big blocks are preserved. For this reason it has been almost completely replaced in practice by methods based on threshold exceedances, where all extremely significant data exceeding a certain level are used.

Generalised distribution of Pareto

The main distribution model for exceedances from thresholds is the Generalised Pareto Distribution (GPD).

Definition: The distribution function of the Generalized Pareto Distribution (GPD) is given by

$$G_{\xi,\beta(x)} = \begin{cases} 1 - (1 + \xi x/\beta)^{-1/\xi}), & \xi \neq 0, \\ 1 - exp(-x/\beta), & \xi = 0, \end{cases} \tag{1.17}$$

where $\beta > 0$, and $x \geq 0$ when $\xi \geq 0$ and $0 \leq x \leq -\beta/\xi$ when $\xi < 0$. The parameters ξ and β refer, respectively, to the parameters of form and scale.

Like the GEV distribution, the GPD is generalized in the sense that it contains a number of special cases: when $\xi > 0$ the distribution function $G_{\xi,\beta}$ is that of an ordinary distribution of Pareto with $\alpha = 1/\xi$ and $\kappa = \beta/\xi$; when $\xi = 0$ we have an exponential distribution; when $\xi < 0$ we have a Pareto distribution of type II with short tail. Also, as in the case of the GEV distribution, fixed the x the parametric form is continuous in ξ, so

$\lim_{\xi \to_0} G_{\xi,\beta}(x) = G_{0,\beta}(x)$

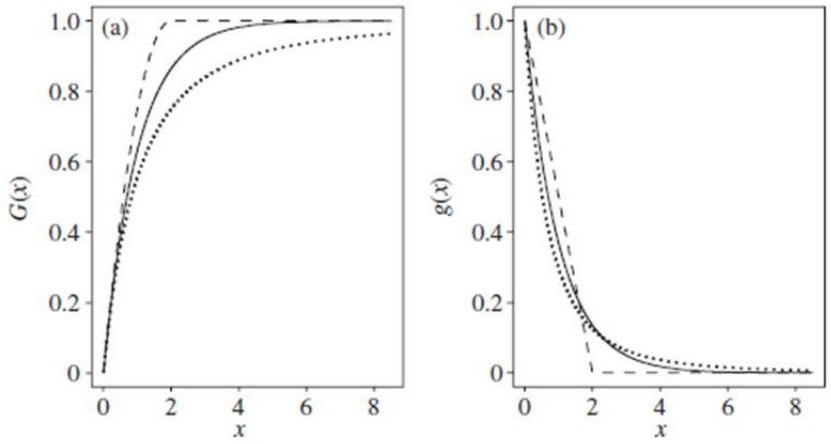

Figure 1.4: Distribution function and Pareto generalized distribution density (GPD)

The distribution function and the Pareto generalized distribution density (GPD) for various values of ξ and $\beta = 1$ are shown in the figure: specifically on the left is indicated with the (a) the GPD distribution function in three cases: the solid line corresponds to the case where $x_i = 0$ (exponential distribution); the dotted line corresponds to the case where $x_i = 0,5$ (Pareto distribution); and the dashed line to the case where $x_i = -0,5$ (Pareto distribution of the second type). The scale parameter β is equal to 1 in all cases. While right (b) indicates the corresponding density. In terms of attraction domain we

have $G_{\xi,\beta} \in MDA(H_\xi)$ for every $\xi \in \Re$. Note that for $\xi > 0$ and $\xi < 0$, this statement easily follows the characterizations in the theorems of Fréchet and Weibull. In the case of heavy tails, $\xi > 0$, it can be easy to verify that $E(X^k) = \infty$ for $k \geq 1/\xi$. The mean of the GPD is defined by placing $\xi < 1$ and is:

$$E(X) = \beta/(1-\xi) \tag{1.18}$$

The role of GPD in extreme value theory is like a natural model for distributions that exceed a given threshold. This concept is defined below along with the mean function of excesses which will also play an important role in theory.

Definition: distribution of exceedances over the u threshold

Let X be a random variable with a distribution function F.

The excess distribution over the u threshold has the following distribution function:

$$F_u(x) = P(X - u \leq x | X > u) = \frac{F(x+u) - F(u)}{1 - F(u)} \tag{1.19}$$

for $0 \leq x < x_F - u$, where $x_F \leq \infty$ is the last point right of F.

Definition:mean function of excesses

The mean function of the excesses of a random variable X with finite mean is given by

$$e(u) = E(X - u | X > u) \tag{1.20}$$

The F_u surplus distribution function describes the distribution of losses that exceed the u threshold. The mean excess function $e(u)$ expresses the mean of F_u as a function of u. In survival analysis the surplus distribution function is more commonly known as the residual life distribution function, it expresses the probability that, for example, an electrical component that worked for u unit of time breaks in the period of time $(u, u + x]$. The mean function of excesses is known as the mean residual life function and indicates the expected residual life of components with different ages.

Theorem:Pickands-Balkema-de Haan

You can find a positive and measurable function $\beta(u)$ such that

$$\lim_{u \to x_F} \sup_{0 \leq x < x_F - u} |F_u(x) - G_{\xi,\beta(u)}(x)| = 0 \tag{1.21}$$

if and only if $F \in MDA(H_\xi)$, $\xi \in \Re$

Thus distributions for which normalized maxima converge to a GEV distribution constitute a series of distributions for which the distribution of exceedances converges to the generalized distribution of Pareto (GPD) when the threshold is raised. In addition, the form parameter of the limiting GPD for excesses is the same as the limiting GEV distribution form parameter for maxima. We have already stated in the previous section that essentially all the continuous distributions of frequently used statistics belong to $MDA(H_\xi)$ for some values of ξ, hence the Pickands-theoremBalkema-de Haan turns out to be a widespread finding that essentially states that GPD is the canonical distribution for modeling losses in excess of high thresholds.

Modelling of excess losses

We use the Pickands-Balkema-de Haan theorem assuming that we are dealing with a distribution of losses $F \in MDA(H_\xi)$ so that, for some high threshold u properly chosen, we can model F_u from a generalized distribution of Pareto. We formalize this with the following assumption.

Assumption 1.1: Let F be a distribution of loss with last right point x_F and suppose that for some high thresholds u we have $F_u(x) = G_{\xi,\beta(x)}$ for $0 \leq x < x_F - u$ and some $\xi \in \Re$ and $\beta > 0$. This is clearly an idealization, as in practice the excess distribution is generally not exactly GPD, but Assumption 1.1 is used to perform the calculations in the following sections.

The methodology. Given the following losses $X_1, ..., X_n$ from F, a random number N_u will exceed the threshold set u. It is useful to rewrite this data as follows $\check{X}_1, \ldots, \check{X}_{N_u}$ for each of these exceedances the amount of $Y_j = \check{X}_j - u$ of losses exceeding the threshold is calculated. The goal is to estimate the parameters of a GPD model by applying this distribution to the N_u excess losses. There are several ways to adapt the GPD, such as the use of maximum likelihood (ML) or probability-weighted moments (PWM). The previous method is more commonly used and is easy to implement if you assume that

the exceedances are independent, since the total density function will then be given by a marginal density product of the GPD. By typing $g_{\xi,\beta}$ for the GPD density, the log-likelihood can be easily calculated as follows:

$$lnL(\xi,\beta;Y_1,...,Y_{N_u}) = \sum_{j=1}^{N_u} lng_{\xi,\beta(Y_j)} = -N_u ln\beta - (1+\frac{1}{\xi})\sum_{j=1}^{N_u} ln(1+\xi\frac{Y_j}{\beta})$$

which shall be maximised according to the following parameter constraints: $\beta > 0$ and $1+\xi Y_j/\beta > 0$ for each j. Solving the maximization issue produces a model GPD $G_{\hat{\xi},\hat{\beta}}$ for distributing the surplus F_u.

Excesses from the highest thresholds: From the model adapted to the distribution of exceedances exceeding the u threshold, you can easily deduce a model for the distribution of exceedances exceeding any threshold. You have, therefore, the following lemma:

Lemma(1.2):

Under assumption 1.1 we have that $F_v(x) = G_{\xi,\beta+\xi(v-u)}(x)$for each threshold $v \geq u$ Proof: We use the (1.21) and the GPD distribution function to deduce that

$$\bar{F}_v(x) = \frac{\bar{F}(v+x)}{\bar{F}(v)} = \frac{\bar{F}(u+(x+v-u))}{\bar{F}(u)}\frac{\bar{F}(u)}{\bar{F}(u+(v-u))} = \frac{\bar{F}_u(x+v-u)}{\bar{F}_u(v-u)}$$

$$= \frac{\bar{G}_{\xi,\beta(x+v-u)}}{\bar{G}_{\xi,\beta(v-u)}} = \bar{G}_{\xi,\beta+\xi(v-u)}(x)$$

So the excess distribution on the higher thresholds remains a GPD with the same parameter ξ, but with a scale parameter that grows linearly with the v threshold. Provided that $\xi < 1$, the mean function of excesses is given by:

$$e(v) = \frac{\beta+\xi(v-u)}{1-\xi} = \frac{\xi v}{1-\xi} + \frac{\beta-\xi u}{1-\xi} \tag{1.22}$$

where $u \leq v < \infty$ if $0 \leq \xi < 1$ and $u \leq v \leq u - \beta/\xi$ if $\xi < 0$

The linearity of the mean surplus function (1.22) in v is commonly used as a diagnostic tool for data accepting a GPD model for distributing exceedances.

It forms the basis for the following simple graphical method for choosing an appropriate threshold.

Plot of the mean sample surplus:

For positive value loss data $X_1, ..., X_n$ is defined as the function of the mean of sample exceedances so that it is an empirical estimator of the function of the mean of excesses. The estimator is given by:

$$e_n(v) = \frac{\sum_{i=1}^{n}(X_i - v)I_{\{}X_i > v\}}{\sum_{i=1}^{n} I_{\{}X_i > v\}} \tag{1.23}$$

To study the mean function of exceedances is generally constructed the mean plot of exceedances

$\{(X_{i,n}, e_n(X_{i,n})) : 2 \leq i \leq n\}$, where $X_{i,n}$ denotes the statistics of i-th order. If the data supports a GPD model on a high threshold, then (1.23) suggests that this plot should become increasingly "linear" for values higher than v. A linear upward trend indicates a GPD model with a positive form parameter of ξ; a horizontal trend plot indicates a GPD with a shape parameter of approximately zero, or, in other words, an exponential distribution of exceedances; a linear downward trend indicates a GPD with negative form parameter.

These are the ideal situations but in practice you need some experience to read the plots of the mean surplus. Even for data which are in themselves distributed according to a generalized Pareto model, the plot of the sample mean of excesses is rarely perfectly linear, particularly for the right end of the distribution, where we are measuring a small number of large exceedances. In fact, we often do not consider the few final points, as they can distort the graphic representation a lot. If we have evidence that the mean surplus plot becomes linear then we could choose as our threshold u a value towards the beginning of the linear section of the plot.

Tail modelling and tail risk measures

In this section we will explain how the GPD model for excess losses is used to estimate the tail of the underlying F loss distribution and the associated risk measures. To make the necessary theoretical calculations, let us take

Assumption 1.1 again.
tail probability and risk measures: First we observe that under the assumption 1.1 we have for $x \geq u$,

$$\bar{F}(x) = P(X > u)P(X > x|X > u) = \bar{F}(u)P(X - u > x - u|X > u)$$
$$= \bar{F}(u)\bar{F}_u(x - u) = \bar{F}(u)(1 + \xi\frac{x-u}{\beta})^{-1/\xi} \tag{1.24}$$

that if $F(u)$ is known it gives us a formula for tail probabilities. This formula can be reversed to obtain a high quantile of the underlying distribution, which is interpreted as the VaR. For $\alpha \geq F(u)$ you have that the VaR is equal to:

$$VaR_\alpha = q_\alpha(F) = u + \frac{\beta}{\xi}((\frac{1-\alpha}{\bar{F}(u)})^{-\xi} - 1) \tag{1.25}$$

It's good at this point, for a better understanding, to give a definition of the Expected Shortfall:

Expected Shortfall: For a L loss with $E(|L|) < \infty$ and distribution function F_L, the ES at an alpha confidence level $\alpha \in (0, 1)$ is defined as:

$$ES_\alpha = \frac{1}{1-\alpha}\int_\alpha^1 q_u(F_L)du \tag{1.26}$$

where $q_u(F_L) = F_L^{\leftarrow}(u)$ is the quantile function of F_L. The condition $E(|L|) < \infty$ ensures that the integral in (1.26) is well defined. From the definition, the Expected Shortfall is linked to the VaR by the following expression:

$$ES_\alpha = \frac{1}{1-\alpha}\int_\alpha^1 VaR_u(L)du \tag{1.27}$$

Returning to the previous treatment, after giving the definition of Expected Shortfall, it is possible to say that assuming that $\xi < 1$ the associated Expected Shortfall can be easily calculated from (1.26) and (1.27). You get:

$$ES_\alpha = \frac{1}{1-\alpha}\int_\alpha^1 q_x(F)dx = \frac{VaR_\alpha}{1-\xi} + \frac{\beta - \xi u}{1-\xi} \tag{1.28}$$

It is important to note that assumption 1.1 and Lemma 1.2 imply that losses in excess respect to the VaR_α have a GPD distribution that meets $F_{VaR_\alpha} = G_{\xi,\beta+\xi(VaR_\alpha - u)}$. The ES estimator in (1.1) can also be obtained by adding the mean of this distribution to the VaR_α, i.e., $ES_\alpha = VaR_\alpha + e(VaR_\alpha)$, where $e(VaR_\alpha)$ is given by (1.22). It is interesting to examine how the ratio of the two risk measures behaves for the large values of the α quantile. It is easily calculated from (1.25) and (1.27) that:

$$lim_{\alpha \longrightarrow 1}\frac{ES_\alpha}{VaR_\alpha} = \begin{cases} (1-\xi)^{-1}, & 0 \leq \xi < 1, \\ 1, & \xi < 0, \end{cases} \tag{1.29}$$

so the ξ form parameter of the GPD actually determines the ratio when we are far enough away from the tail values.

Estimation in practice: Note that under assumption 1.1, tail probabilities, VaR, and Expected Shortfalls are all given by formulas such as $g(\xi, \beta, \bar{F}(u))$. Assuming we have built a GPD model for excess losses above a u threshold, we estimate these quantities first by replacing ξ and β in formulas (1.24) - (1.28) with estimated values. Of course, we also need an estimate of $\bar{F}(u)$ and we use for this purpose the simple empirical estimator N_u/n . In doing this, we implicitly assumed that there is a sufficient proportion of sample values above the u threshold to estimate $\bar{F}(u)$ reliably. However, we hope to get more than an empirical method using some sort of extrapolation based on the GPD model for the most extreme tail probabilities and for risk measures. For the tail probabilities, proposed for the first time by Smith (1987), an estimator of form is obtained:

$$\hat{F}(x) = \frac{N_u}{n}(1 + \hat{\xi}\frac{x-u}{\hat{\beta}})^{-1/\hat{\xi}} \tag{1.30}$$

(where $\hat{\xi}, \hat{\beta}$ are the estimated values) that's important to note is only valid for $x \geq u$. For $\alpha \geq 1 - N_u/n$ a similar estimate is obtained for VaR_α and for ES_α from (1.26) and from (1.29). Of course we also want to get confidence

intervals. If we used the likelihood approach to estimate ξ and β, then it is quite easy to get confidence intervals for $g(\xi, \beta, N_u/n)$ which account for uncertainty in ξ and β (estimated) but overlook the uncertainty in N_u/n as an estimate of $\bar{F}(u)$.

1.5 The Copulas functions

Empirical evidence has shown the inadequacy of normal distribution to model the actual distribution of returns on financial assets in particular on two points:
1) empirical marginal distributions are asymmetric and with fat tails;
2)the assumption of normal multivariate distribution does not take into account the possibility of extreme joint movements of returns on portfolio financial assets.

In other words, the real empirical dependency structure between financial assets moves away from that described by the Gaussian distribution. The copulas functions, used in the financial field, are a valuable tool to implement in a more realistic way efficient algorithms for simulating the actual distributions of financial asset returns. These mathematical functions are able to model the dependency structure between the marginals regardless of the distributional form of the marginals themselves, appropriately describing how the joint distribution is combined with its marginal distributions. Therefore, the copula functions allow us to overcome the problem of the multivariate cumulative distribution function (CDF) and the estimation of its parameters, dividing the problem into two phases:
1) the determination of the marginal CDF, $F_1, ..., F_n$, describing the probabilistic structure of the individual risk factors, and the subsequent estimation of the relevant parameters from the available data, using common statistical methods suitable for this purpose;
2) the determination of the dependency structure between the random variables $X_1, ..., X_n$, through the use of a suitable copulation function for this purpose.

The goodness of the choice both of the analytical representations of the marginal distributions and of the copula function that unites them must be verified with techniques of "backtesting". In this way, it is possible to choose

the analytical forms of the marginals and the copula function that best represent the empirical profit and loss distribution of the portfolio derived from empirical real data. In conclusion, it is possible to construct a multivariate distribution with different marginals and with a dependency structure represented by an appropriate copula function. The copula function is therefore a mathematical function that combines marginal probabilities in a joint distribution (Joe,1997; Nelsen,1999). Through the copula functions it is possible to separate the dependency and the behaviour of the marginals relative to the assets that compose the financial portfolio. It is therefore evident that the phase of selecting and calibrating the appropriate function copula from real financial data is a crucial aspect of the problem.

1.5.1 Definition and mathematical properties of the copula function

Underlining the potential of the analytical instrument of the copula function, its analytical properties are briefly described here. A n-dimensional copula function, C ,is a multivariate cumulative distribution function (CDF) with uniformly distributed margins in the [0.1] range that satisfies the following properties (Sklar, 1959):
i) $C : [0,1]^n \longrightarrow [0,1]$;
ii) C is limited and not decreasing;
iii)C has marginal C_i such that $C_i(u) = C(1, ..., 1, u, 1, ..., 1)$.

It is evident from the above definition that if $F_1, \ldots, F_n$ are functions of univariate distributions, then $C(F_1(x_1), \ldots, C(F_n(x_n)$ is a multivariate CDF with marginals $F_1, \ldots, F_n$ since $U_i = F_i(X_i)$, with $i = 1, \ldots, n$ is a uniform random variable $U(0,1)$. Thus, copulas functions can be used to combine marginal distributions into a multivariate distribution, proving to be a valuable tool for building and simulating multivariate distributions. Copulas functions are also unique, meaning that for a given multivariate distribution with continuous marginal distributions there is only one copula function representing it. The following theorem is known as Sklar's theorem and is in the financial field the most important theorem on copulas functions as used in all practical applications.

The Sklar's Theorem:

If F is an n-dimensional CDF with continuous margins, $F_1, ..., F_n$, then F has the following unique copulas representation:

$$F(x_1, ..., x_n) = C(F_1(x_1), ..., F_n(x_n)) \tag{1.31}$$

From the Sklar theorem (in eq 1.31) it descends that , for continuous multivariate distribution functions, the univariate margins and the multivariate dependence structure can be separated, where the latter can be represented by an appropriate copula function. From Sklar's theorem descends the following Scaillet corollary (2004):

Scaillet's Corollary:

If F is an n-dimensional CDF with continuous margins, $F_1, \ldots, F_n$ and copula C (which satisfies the equation 1.31). So for every $u = (u_1, \ldots, u_n)$ in $[0, 1]^n$:

$$C(u_1, ..., u_n) = F(F_1^{-1}(u_1), ..., F_n^{-1}(u_n)) \tag{1.32}$$

where F_i^{-1} is the generalized inverse of F_i. The density of the F multivariate CDF can be written as follows:

$$f(x_1, ..., x_n) = c(F_1(x_1), ..., F_n(x_n)) \prod_{i=1}^{n} f_i(x_i) \tag{1.33}$$

In equation (1.31) $c(F_1(x_1), ..., F_n(x_n)$ and $f(x_1, ..., x_n)$ are the density functions of the copula and marginal respectively. From Scaillet's corollary (2004) it follows that the dependency structure built into the copula can be recovered from knowledge of the joint distribution F and its marginal F_i. We conclude this paragraph by presenting some simple examples of copulas functions. A first example is the Gumbel copula function, expressed by the following equation:

$$C_\beta^{Gu}(u_1, ..., u_n) = \exp[(-\log u_1)^{1/\beta} + ... + (-\log u_n)^{1/\beta}]^\beta \tag{1.34}$$

where $0 < \beta \leq 1$ is the parameter that expresses the dependence between the marginal components; the case $\beta = 1$ describes the condition of independence between random variables, while $\beta \longrightarrow 0$ gets perfect dependence. Another very simple example is given by the copula of independent random variables, which takes the following analytical form:

$$C_{ind}(u_1, ..., u_n) = u_1 \cdot ... \cdot u_n \tag{1.35}$$

A further example is the Farlie-Gumbel-Morgenstern copula (FGM), which, in the bivariate case, is defined through the following analytical expression:

$$C(u_1, u_2) = u_1 u_2[1 + \alpha(1 - u_1)(1 - u_2)], \qquad -1 \leq \alpha \leq 1 \tag{1.36}$$

1.5.2 The elliptical copulas: the Gaussian copula and the Student t copula

The class of elliptical distributions provides useful examples of multivariate distributions as they share the same tractability properties as the normal multivariate distribution and, at the same time, they allow to model multivariate extreme events and forms of dependence other than normal. Elliptical copulas are simply the copulas of elliptical distributions. The simulation of elliptical distributions is easy to perform. Thus, as a result of Sklar's theorem, the simulation of elliptical copulas is also simple. The copula of the normal multivariate distribution is the normal or Gaussian copula. Thus, a random vector $X = (X_1, ..., X_n)$ is normal multivariate if and only if:
i) univariate marginal distributions, $F_1, ..., F_n$, are normal;
ii) the dependency structure between these marginals is described by a single copula function, C (the normal copula), such that:

$$C_R^{Ga}(u_1, ..., u_n) = \Phi_R(\varphi^{-1}(u_1), ..., \varphi^{-1}(u_n)) \tag{1.37}$$

where Φ_R is the standard CDF multivariate with linear correlation matrix R and ϕ^{-1} is the inverse of the normal CDF univariate standard. Normal multivariate is traditionally adopted in risk management applications to simulate the distribution of the n risk factors that affect the value of the trading portfolio (market risk) and the n systematic factors that drive the value of

a counterparty's credit quality index (credit risk). For $n = 2$, the equation (1.37) can be rewritten as follows:

$$C_R^{Ga}(u,v) = \int_{-\infty}^{\phi^{-1}(u)} \int_{-\infty}^{\phi^{-1}(v)} \frac{1}{2\pi(1-R_{12}^2)^{1/2}} exp\{-\frac{s^2 - 2R_{12}st + t^2}{2(1-R_{12}^2)}\} dsdt \tag{1.38}$$

where R_{12} is, simply, the linear correlation coefficient of the two random variables. The normal bivariate copula, (1.38), is not endowed with dependence in the upper tail if $R_{12} < 1$. Also, since the elliptical distributions are symmetrical, the dependence coefficients of the upper and lower tail are equal. Thus, normal copulas are not endowed with dependence of the lower tail. The Student's t copula is the copula of the multivariate Student t distribution. Assuming that X is a vector with standardized n-variate Student t distribution with v degrees of freedom and with covariance matrix $\frac{v}{v-2}R$ (for $v > 2$), X can be represented as follows:

$$X = \frac{\sqrt{v}}{\sqrt{S}} Y \tag{1.39}$$

In the (1.39) $S \sim \chi_v^2$ is independent from the random vector $Y \sim N_n(0, R)$. The copula of the vector Y is the copula Student t with v degrees of freedom that can be represented analytically as follows:

$$C_{v,R}^t(u) = t_{v,R}^n(t_v^{-1}(u_1), ..., t_v^{-1}(u_n)) \tag{1.40}$$

where

$R_{i,j} = \Sigma_{ij}/\sqrt{\Sigma_{ii}\Sigma jj}$ per $i, j \in \{1, ..., n\}$; $t_{v,R}^n$ is the multivariate CDF of the random vector $\sqrt{v}Y/\sqrt{S}$, where the random variable $S \sim \chi_v^2$ is independent of the random vector Y (which has a normal n-dimensional distribution with mean vector 0 and covariance matrix R); t_v represents the marginal distributions of $t_{v,R}^n$. For n=2, the t-Student copula takes the following analytical form:

$$C_{v,R}^t(u,v) = \int_{-\infty}^{t_v^{-1}(u)} \int_{-\infty}^{t_v^{-1}(v)} \frac{1}{2\pi(1-R_{12}^2)^{1/2}} \{1 + \frac{s^2 - 2R_{12}st + t^2}{v(1-R_{12}^2)}\}^{-(v+2)/2} dsdt$$

where R_{12} is the linear correlation coefficient of the bivariate Student t distribution ($n = 2$) with v degrees of freedom, if $v > 2$. Embrechts, Lindskog

and McNeil (2001) showed that, unlike the Gaussian copula, the bivariate Student's t copula has superior tail dependence. Such dependence is, as you should expect, increasing in R_{12} and decreasing in v. Thus, Student's t copula would seem more suitable (than the Gaussian copula) to simulate events similar to stock market collapses or joint insolvencies of multiple counterparties in a portfolio of credit assets. We can obtain the density function of the Student's t copula (Bouye et al., 2000) through the following equation:

$$c(u_1, ..., u_n; R, v) = \frac{\Gamma((v+n)/2)[\Gamma(v/2)]^n(1+\omega^T R^{-1}\omega)^{-(v+2)/2}}{|R|^{1/2}\Gamma(v/2)[\Gamma(v+1)/2]^n \prod_{i=1}^{n}(1+\omega_i^2/v)^{-(v+1)/2}} \tag{1.41}$$

where $\omega = (\omega_1, ..., \omega_n)^T = (t_v^{-1}(u_1), ..., t_v^{-1}(u_n))$. Student's t copula represents the dependency structure implied in a multivariate Student's t distribution. We highlight how this type of copula function has aroused a lot of operational interest in the financial field especially in a context of modeling multivariate financial data, for example the daily logarithmic price changes of securities. For example Mashal and Zeevi(2002) as well as Breymann et al.(2003) have shown that the adaptation to empirical data of Student's t copula is generally higher than that of normal copula (the latter structure of dependence of the normal multivariate distribution). One reason for this is the student's t ability to grasp the correlation between extreme values, which is often seen in financial phenomena. On the other hand, normal copula is only a particular case of Student's t copula and precisely when the number of degrees of freedom goes to infinity. In addition, unlike normal copula, Student's t copula has tail dependence on both tails (Embrechts, McNeil and Straumann,2001), while the Gaussian copula has no tail dependence if the value of Pearson's linear correlation coefficient is different from ± 1. For these reasons, Student's t copula is capable of generating more extreme events than normal copula, for example the events of insolvency, extreme events by definition.

1.5.3 Vine copula

With the availability of massive multivariate data comes a need to develop flexible multivariate distribution classes. The copula approach allows to construct marginal models for each variable separately and join them with a dependence structure characterized by a copula. The class of multivariate copulas was limited for a long time to elliptical (including the Gaussian and

t-copula) and Archimedean families (such as Clayton and Gumbel copulas). Both classes are rather restrictive with regard to symmetry and tail dependence properties. The class of vine copulas overcomes these limitations by building a multivariate model using only bivariate building blocks. This gives rise to highly flexible models that still allow for computationally tractable estimation and model selection procedures.

Since vine copulas are built out of bivariate copulas we now discuss properties of bivariate copulas. To investigate the dependence properties we consider several dependence measures. Since the Pearson correlation $\rho(X_1, X_2) = Cor(X_1, X_2)$ is not invariant with respect to monotone transformations of the margins, it is more useful to consider invariant dependence measures such as Kendall's τ and Spearman's ρ. In particular, Spearman's rank correlation is defined as the Pearson correlation of the random variables $F_1(X_1)$ and $F_2(X_2)$, i.e., $\rho_s(X_1, X_2) = Cor(F_1(X_1), F_2(X_2))$. Another popular measure invariant to marginal transformations is Kendall's τ , defined as

$\tau(X_1, X_2) = P((X_{11}-X_{21})(X_{12}-X_{22}) > 0)-P((X_{11}-X_{21})(X_{12}-X_{22}) < 0)$ where (X_{11}, X_{12}) and (X_{21}, X_{22}) are independent and identically distributed copies of (X_1, X_2). Since $\tau(X_1, X_2)$ and $\rho(X_1, X_2)$ are invariant with regard to margins they depend only on the underlying copula. More specifically it holds

$$\tau(X_1, X_2) = 4 \int_0^1 \int_0^1 C(u_1, u_2) dC(u_1, u_2) - 1$$

$$\rho(X_1, X_2) = 12 \int_0^1 \int_0^1 u_1 u_2 dC(u_1, u_2) - 3$$

While the catalogue of bivariate parametric copula families is large, this is not the case for $d > 2$. The motivation for vine copula models was to find a way to construct multivariate copulas using only bivariate copulas as building blocks.

Regular vine copulas and distributions

A regular vine distribution for a d-dimensional random vector$X = (X_1, \ldots, X_d)$ is specified by the triplet $(F,\ V,\ B)$ with:
1. Marginal distributions: $F= (F_1, \ldots, F_d)$ is a vector of continuous marginal distribution functions of the random variables $(X_1, \ldots, X_d)$
2. Regular vine tree sequence: V is an R-vine tree sequence on d elements.

3. Bivariate copulas: The set $B = C_e|e \in E_i; i = 1, \ldots, d-1$, where C_e is a bivariate copula with density c_e . Here E_i is the edge set of tree T_i in the R-vine tree sequence V.
4. Relation between R-vine tree sequence and the set of bivariate copulas: for each $e \in E_i$, $i = 1, \ldots, d-1$, $e = (a, b)$, $C_e(\cdot, \cdot)$ is the copula associated with the conditional distribution of $X_{C_{e,a}}$ and $X_{C_{e,b}}$ given X_{D_e}.

Existence of a regular vine distribution

Assume that (F, V, B) satisfy the properties (1)-(3), then there is a valid d-dimensional distribution F with density $f_{1,\ldots,d}(x_1, \ldots, x_d) = f_1(x_1) \times \cdots \times f_d(x_d) \times \prod_{i=1}^{d-1} \prod_{e \in E_i} X_{C_{e,a}} X_{C_{e,a}} \mid D_e(F_{C_{e,a}} \mid D_e(X_{C_{e,a}} \mid X_{D_e}), F_{C_{e,b}} \mid D_e(X_{C_{e,b}} \mid X_{D_e}))$

such that for each $e \in E_i$, $i = 1, \ldots, d-1$, with $e = a, b$ we have for the distribution function of $X_{C_{e,a}}$ and $X_{C_{e,b}}$ given X_{D_e} $F_{C_{e,a}C_{e,b}} \mid D_e(X_{C_{e,a}}, X_{C_{e,b}} \mid X_{D_e}) = C_e(F_{C_{e,a}} \mid D_e(X_{C_{e,a}} \mid X_{D_e}), F_{C_{e,b}} \mid D_e(X_{C_{e,b}} \mid X_{D_e}))$

Further the one-dimensional margins of F are given by $F_i(x_i)$, $i = 1, \ldots, d$. If all margins are standard uniform, we call the resulting distribution a regular Vine copula.

1.5.4 Skew-t copula

If X_i, $i \in 1, 2, ..., p$ are continuous random variables, the density $f(x_1, x_2, \ldots, x_p)$ of their joint distribution can be presented through a copula density $c(u_1, u_2, \ldots, u_p)$ and marginal densities $f_i(x_i)$:

$$f(x_1, x_2, \ldots, x_p) = c(F_1(x_1), \ldots, F_p(x_p)) \cdot f_1(x_1) \cdot ... \cdot f_p(x_p)$$

A copula (A.Azzalini, 2014) is called skew-t copula, if its density function is:

$$c_{p,v}(u; \mu, \Sigma, \alpha) = \frac{g_{p,v}[G_{1,v}^{-1}(u_1; 0, \Sigma_{1,1}, \alpha_1), \ldots, G_{1,v}^{-1}(u_p; 0, \Sigma_{p,p}, \alpha_p); \mu, \Sigma, \alpha]}{\prod_{i=1}^{p} g_{1,v}[G_{1,v}^{-1}(u_i; \mu_i, \Sigma_{i,i}, \alpha_i); \mu_i, \Sigma_{i,i}, \alpha_i]}$$

where: $g_{p,v}(\cdot; \mu, \Sigma, \alpha) : R^p \to R$ is the density function of the p-variate skew-t distribution and function $G_{1,v}^{-1}(u_i; \mu_i, \sigma_{i,i}, \alpha_i) : R^1 \to I$, $i \in \{1, \ldots, p\}$ denotes the inverse of the univariate $t_{1,v}$ distribution function.

Multivariate skew $t_{p,v}$ copula can be presented of the form:

$$C_{t,v}(u;\mu,\Sigma,\alpha) = 2\int_{-\infty}^{A}\cdots\int_{-\infty}^{B} t_{p,v}(x;\mu,\Sigma)\cdot T_{v,p}\big[\alpha^T W^{-1}(x-\mu)(\frac{v+p}{Q+v})^{1/2}\big]dx$$

where

$$A = G_{1,v}^{-1}(u_1;\mu_1,\sigma_{1,1},\alpha_1)$$
$$B = G_{1,v}^{-1}(u_p;\mu_p,\sigma_{p,p},\alpha_p)$$
$$Q = (x-\mu)^T\Sigma^{-1}(x-\mu)$$
$$W = \delta_{i,j}\sqrt{\sigma_{i,j}}$$

1.5.5 Estimation of copula function parameters

In this section, we present some methods to estimate the parameters of some of the most widely used copula functions in the financial field.

The Maximum Likelihood method (ML):

Let's say f is the density function of the joint distribution F. Analytically you have:

$$f(x_1,...,x_n) = c(F_1(x_1),...,F_n(x_n))\prod_{i=1}^{n} f_i(x_i) \tag{1.42}$$

In equation (1.42) f_i is the univariate density function of the marginal distribution F_i and c is the density function of the copula function calculated through the following equation:

$$c(u_1,...,u_n) = \frac{\delta C(u_1,...,u_n)}{\delta u_1 ... \delta u_n} \tag{1.43}$$

Suppose you have a set of T empirical data related to n log-returns on financial assets, $X = \{(x_1^t,...,x_n^t)\}_{t=1}^{T}$. Let's say $\delta = (\delta_1,...,\delta_n,\alpha)$ is the vector of parameters to be estimated where δ_i, with $i = 1,...,n$ is the vector of the parameters of the marginal distribution F_i and α is the vector of the parameters of the copula. The log-likelihood function is expressed as follows:

$$l(\delta) = \sum_{t=1}^{T} lnc(F_1(x_1^t;\delta_1),...,F_n(x_n^t;\delta_n);\alpha) + \sum_{t=1}^{T}\sum_{i=1}^{n} lnf_i(x_i^t;\delta_i) \tag{1.44}$$

The maximum likelihood estimator (ML) $\hat{\delta}$ of the parameter vector δ is the one that maximizes the equation (1.44), in analytical terms: $\hat{\delta} = argmaxl(\delta)$.

The method Inference Functions for Margins(IFM):

According to Joe and Xu's IFM (Inference Functions for Margins) method (1996), marginal distribution parameters are estimated separately from copula parameters. In other words, the estimation process is divided into the following two steps:
i) estimate the parameters δ_i, for $i = 1, \ldots, n$ of the marginal distributions F_i through the maximum likelihood method. Analytically:

$$\hat{\delta}_i = argmaxl^i(\delta_i) = argmax \sum_{t=1}^{T} lnf_i(x_i^t; \delta_i)$$

where l^i is the log-likelihood function of the marginal distribution F_i ;
ii)estimate the vector of the copula parameters, α, after obtaining the estimates of the marginal parameters in the previous phase. Analytically:

$$\hat{\alpha} = argmax\ l^c(\alpha) = argmax \sum_{t=1}^{T} \log c(F_1(x_1^t; \hat{\delta}_1), ..., F_n(x_n^t; \hat{\delta}_n); \alpha)$$

where l^c is the log-likelihood function of the copula.

The Method Canonical Maximum Likelihood(CML):

The CML (Canonical Maximum Likelihood) method differs from the IFM methot: it makes no assumption about the parametric form of marginal distributions. The estimation process is carried out through the following two steps:
i) transform the set of $(x_1^t, ..., x_n^t)$, with $t = 1, \ldots, T$, into uniform determinations $(\hat{u_1^t}, \ldots, \hat{u_n^t})$ using empirical distributions (in other words determinations are generated by empirical copula);
ii) estimate the copula's parameters as follows:

$$\hat{\alpha} = argmax \sum_{t=1}^{T} \log c(\hat{u_1^t}, ..., \hat{u_n^t}; \alpha)$$

It is possible, for example, to estimate the parameter, R, of the Gaussian copula function with the CML method or the IFM method as follows (Durrleman,Nikeghbali and Roncalli, 2000):

$$\hat{R}_{IFM/CML} = \frac{1}{T}\sum_{t=1}^{T} \varsigma_t^T \varsigma_t$$

where $\varsigma_t(\Phi^{-1}(u_1^t), ..., \Phi^{-1}(u_n^t))$.

The following recursive procedure (Bouyè et al., 2000) is used to estimate the R parameter of the t_v-Student copula:
i) assume that $\hat{R}_1$ is the IFM/CML estimator of the R parameter of the Gaussian copula;
ii)

$$\hat{R_{m+1}} = \frac{1}{T}(\frac{v+n}{v})\sum_{t=1}^{T} \frac{\varsigma_t^T \varsigma_t}{1 + \frac{1}{v}\varsigma_t \hat{R}_m^{-1} \varsigma_t^T}$$

con $m = 1, 2, ...$ in cui $\varsigma_t(t_v^{-1}(u_1^t), ..., t_v^{-1}(u_n^t))$;
iii) the point ii) must be repeated until $\hat{R_{m+1}} = \hat{R}_m$. At this point , the IFM/CML estimator of the R Student's t_v copula parameter is $\hat{R}_{IFM/CML} = \hat{R}_\infty$. Mashal and Zeevi (2002) suggested using the following algorithm to estimate the v and R parameters of the Student's t_v copula:
a) turns the $(x_1^t, ..., x_n^t)$ dataset, with $t = 1, \ldots, T$, into uniform determinations $(\hat{u_1^t}, ..., \hat{u_n^t})$ using empirical marginal distributions;
b)estimate $\hat{R}$ using the non-parametric estimator τ of Kendall:
$\hat{R}_i j = sin(\frac{\pi}{2}\hat{\tau}_{ij})$,$i, j = 1, ..., n$
c)run a numerical search for $\hat{v}$, for example

$$\hat{v} = arg \max_{v\in(2,\infty]} [\sum_{t=1}^{T} log(c(u_1^t, ..., u_n^t; v, , \hat{R}))]$$

where:

$$c(u_1, ..., u_n; v, R) = \frac{\Gamma((v+n)/2)[\Gamma(v/2)]^{n-1}(1 + y' R^{-1} y)^{-(v+n)/2}}{|R|^{1/2}[\Gamma((v+1)/2)]^n \prod_{i=1}^{n}(1 + y_i^2/v)^{(v+1)/2}}$$

e

$$y = (y_1, ..., y_n) = (t_v^{-1}(u_1), ..., t_v^{-1}(u_n))$$

Nonparametric estimation:

Through the methods described above it is possible to estimate the parameters of a given copula function. Deheuvels (1979) introduces a technique for constructing empirical copula (or Deheuvels copula) from sample data. The empirical copula is the copula of the empirical multivariate distribution. Assuming that $\{x_1^{(t)}, ..., x_n^{(t)}\}$ are the order statistics and that $\{r_1^t, ..., r_n^t\}$ are the rank statistics (with $t = 1, \ldots, T$) of the sample data set, we have that $x_i^{(r_i^t)} = x_i^t$, with $i = 1, ..., n$.
Each function: $\hat{C}(\frac{t_1}{T}, ..., \frac{t_n}{T}) = \frac{1}{T}\sum_{t=1}^{T}\prod_{i=1}^{n} 1_{[r_i^t \leq t_i]}$, defined on the lattice $\ell = \{(\frac{t_1}{T}, ..., \frac{t_n}{T}) : 1 \leq i \leq n; t_i = 0, ..., T\}$ is an empirical copula. The density function (Nelsen, 1998) of the empirical copula has the following analytical expression:

$$\hat{c} = \{(\frac{t_1}{T}, ..., \frac{t_n}{T}) = \sum_{i_1=1}^{2} ... \sum_{i_n=1}^{2} (-1)^{i_1+...+i_n} \hat{C}(\frac{t_1 - i_1 + 1}{T}, ..., \frac{t_n - i_n + 1}{T})$$

1.5.6 Calibration methods of the copula function parameters

In this paragraph we present some methods of calibration of the parameters from a given analytical rappresection of the copula function. The first step is to select the type of copula that best suits the empirical data. The method described by Deheuvels (1979) is able to select the Archimedean copula that best suits the real data. The Archimedean copula (see Mc Neil et al. 2005) has the analytical representation described by equation(1.45). Then, to select the most suitable copula, simply identify the generator: ψ. In the bivariate case ($n = 2$), Genest and Rivest(1993) defined a univariate function, K, which is linked to the Archimedean copula generator through the following analytical expression:

$$K_\psi(z) = z - \frac{\psi(z)}{\psi'(z)} \tag{1.45}$$

A non-parametric estimate of the above equation is given by the following equation:

$$\hat{K}(z) = \frac{1}{T}\sum_{t=1}^{T} 1_{[\delta_t \leq z]} \tag{1.46}$$

If we choose a parametric representation for the generator, ψ, then the parameter, α of the selected Archimedean copula is estimated using, for example, Kendall's τ estimate:

$$\tau = \binom{T}{2}^{-1} \sum_{i<j} sign[(x_1^i - x_1^j) \cdot (x_2^i - x_2^j)] \tag{1.47}$$

The parameter α can also be estimated using the IFM method or the CML method. Using α, a parametric estimate of the equation (1.47) can be easily obtained. All of the above steps are repeated for different ψ. Frees and Valdez (1998) proposed to use a QQ-plot between the equations (1.45) and (1.46). Optimal copula can also be selected by minimizing the distance based on the L^2 norm between equations (1.45) and (1.46) as in Durrleman, Nikeghbali and Roncalli (2000). Analytically:

$$d_2(\hat{K}, K) = \int_0^1 [K(z) - \hat{K}(z)]^2 dz \tag{1.48}$$

The method described here can also be used to graphically estimate the α parameter of a given Archimedean copula.

Selection of the most appropriate copula using empirical copulation

Let's say $\{C_k\}_{1 \leq k \leq K}$ is the set of available copulas. We choose the C_k copula that minimizes the following distance, expressed in the following equation, based on the L^n discrete norm, between the C_k copula and the empirical copula:

$$\bar{d}_n(\hat{C}, C_k) = (\sum_{t_1=1}^{T} ... \sum_{t_n=1}^{T} [(\hat{C})(\frac{t_1}{T}, ..., \frac{t_n}{T}) - \hat{C}_k(\frac{t_1}{T}, ..., \frac{t_n}{T})]^2)^{1/2} \tag{1.49}$$

The distance (1.49) can also be used to estimate the $\delta \in \Theta$ parameter's vector of a given copula $C(u; \delta)$ as follows:

$$\hat{\delta} = \arg\min_{\delta \in \Theta} (\sum_{u \in \ell} [\hat{C}(u) - C(u; \delta)]^2)^{1/2} \tag{1.50}$$

1.5.7 Simulation algorithms

In this paragraph, we present some algorithms to simulate random determinations $(u_1, \ldots, u_n)$ from certain types of copula function C. From the definition of copula it follows that these random variables are determinations of correlated random variables and evenly distributed in (0,1). So in order to simulate random variables $(x_1, \ldots, x_n)$ from a multivariate distribution F with marginal data F_i , $i = 1, \ldots, n$ and copula C, we need to reverse each u_i using the marginal distributions : $x_i = F_i^{-1}(u_i)$, $i = 1, ..., n$.

Simulation of the Gaussian copula

To generate random determinations of normal copula, we can use the procedure explained in this paragraph. If the R matrix is positive, then there are A matrices of size $n \times n$ such that $R = AA^T$. It is also assumed that the random variables $Z_1, \ldots, Z_n$ are standard independent variables. In addition, the random vector $\mu + AZ$, where $Z = (Z_1, \ldots, Z_n)^T$, and the vector $\mu \in R^n$ is distributed in a multinormal way with medium vector μ and covariance matrix R. The A matrix can be easily determined through the Cholesky decomposition of the R. The result of this decomposition is the unique lower triangular matrix, such that $LL^T = R$. It is therefore possible to generate random determinations of the n-dimensional normal copula by implementing the following algorithm:
i) find the Cholesky A breakdown of the R;
ii) matrix simulate n standard normal independent random determinations

$$z = (z_1, \ldots, z_n)^T$$

iii) fix $x = Az$;
iv) determine the components $u_i = \phi(x_i)$, $i = 1, \ldots, n$; where the vector $(u_1, \ldots, u_n)T$ is a random determination of the n-dimensional gaussian copula, C_R^{Ga}.

Simulation of Copula t_v by Student

To simulate random determinations of Student's t copula, $C_{v,R}^t$, you can use the following algorithm based on the (1.41):
i) find the Cholesky decomposition A of the correlation matrix R;

ii) simulate n independent random determinations $z = (z_1, \ldots, z_n)$ from the standard normal distribution;
iii) simulates a random determination, s, from the distribution χ^2_v , independent of z;
iv) determine the vector $y = Az$;
v) perform the transformation $x = \frac{\sqrt{v}}{\sqrt{s}} \cdot y$;
vi) determine the components $u_i = t_v(x_i)$, $i = 1, \ldots, n$; the resulting vector is:

$$(u_1, \ldots, u_n)^T \sim C^t_{v,R}$$

1.6 Description of the sample

The analysis covered by this report is carried out on a sample of data represented by the daily closing prices of the nine most capitalised stock indices, acquired through the Bloomberg platform, in order to obtain the most representative sample of the stock market as a whole. The time horizon is from 2 January 2012 to 23 September 2022, for a total of about 10 years. The stock indices examined are: Dow Jones, S&P500, Nasdaq 100, FTSE 100, Nikkei 225, SSE Composite, SZSE Component, Euronext 100, HANG SENG. Below is a brief description of these stock indices and the daily returns of each index over the time frame analysed are shown:
i)**Dow Jones**: whose full name is Dow Jones Industrial Average, is the oldest stock index in history, as well as the most well-known of the New York Stock Exchange (NYSE), created by Charles Dow, father of technical analysis and founder of the Wall Street Journal and Edward Jones, American financial statistician. The index is calculated by weighing the price of the top 30 Wall Street stocks. The choice to limit its composition to only 30 Blue Chips has meant that over time, the index has lost much of its importance because it can no longer reflect the entire trend of the American stock market.

ii) **S&P500:** was created by Standard & Poor's in 1957 and follows the trend of a stock basket of the 500 most capitalized US companies. Part of this basket are the shares of large companies traded at the New York Stock Exchange (Nyse), the American Stock Exchange (Amex) and the Nasdaq. The weight attributed to each holding is directly proportional to the market value. This is the most significant index of the entire American market and

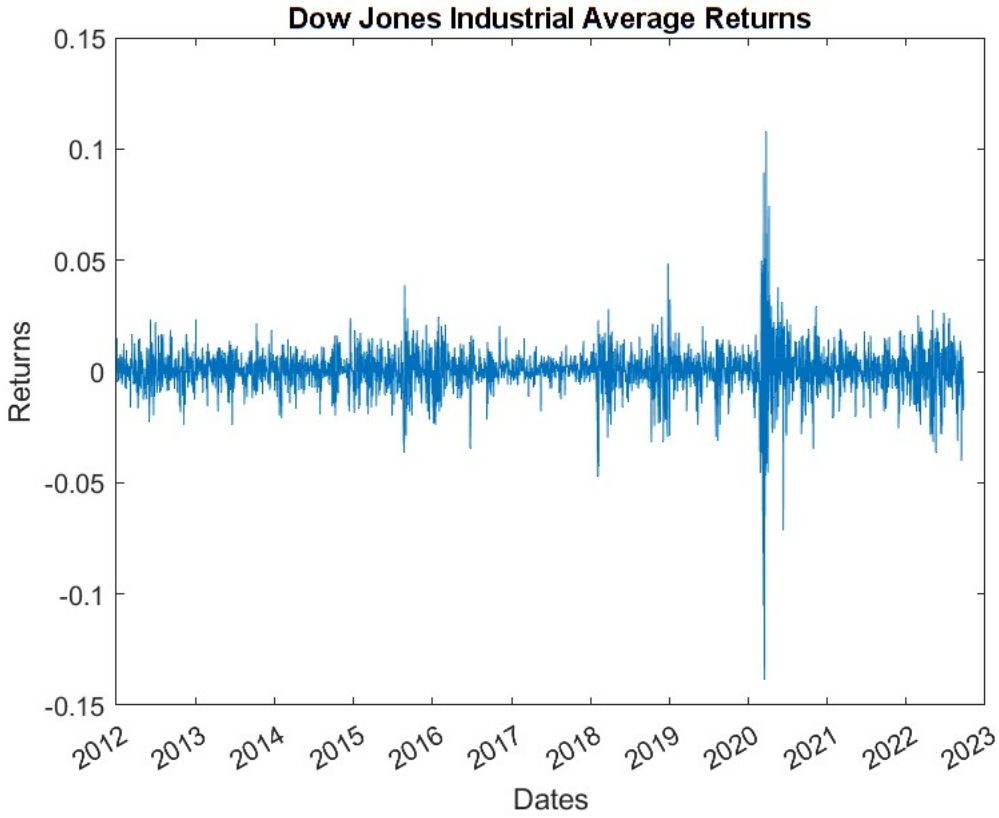

Figure 1.5: Dow Jones Industrial Average Returns

has now surpassed by importance the same Dow Jones, basket of 30 major American stocks. The index was established and continues to be managed by S&P Dow Jones Indices, a joint venture controlled by McGraw Hill Financial.

iii) **Nasdaq 100:** is a stock market index of the top 100 non-financial companies listed on the NASDAQ stock market. It is a weighted index, the weight of the various companies that compose it is based on their market capitalisation, with some rules to take into account the influences of the major components. It does not include financial corporations, and includes some foreign companies. These two factors differentiate it from the S& P500 index.

iv) **FTSE 100:** is a stock index of the 100 most capitalized companies listed on the London Stock Exchange. The index has been quoted since 3 January 1984 at an initial level of 1000. FTSE stands for 'Financial Times Stock Exchange'. The index is managed by the FTSE Group, a now independent company that was originally established as a joint venture between the Financial Times and the London Stock Exchange. The 100 companies

in the index represent about 80% of the market capitalisation of the entire London Stock Exchange. The components the index is established quarterly; the largest companies in the FTSE 250 index are promoted if their market capitalisation places them among the top 90 companies in the FTSE 100. At the end of 2006, for example, the threshold for entry was 2,9 billion. As of 29 December 2006, the six largest companies in the index were BP, Royal Dutch Shell, HSBC, Vodafone, Royal Bank of Scotland and glaxosmithkline, each worth over 60 billion.

v)**Nikkei 225:** The index contains the 225 stocks of the largest companies listed on the Tokyo Stock Exchange (TSE) by capitalization, which are updated annually. The Nikkei 225 began to be calculated from 7 September 1950, retroactively from 16 May 1949 and from January 2010, is updated every 15 seconds. The index peaked on 29 December 1989 at 38957.44 , during the Japanese speculative bubble. It is an index weighted on the stock price, large-cap type, representing the entire market. Therefore, there are no specific weights for the different economic sectors to which a single security belongs: all shares have the same weight based on a nominal value equal to 50 yen.

vi)**SSE Composite:** The Shanghai Stock Exchange Composite Index comprises all A and B shares listed on the Shanghai Stock Exchange, for a total of 872 companies. It's a capitalization-weighted index, created on December 19, 1990, with a base value of 100.

vii)**SZSE Component:** is the main index of the Shenzhen Stock Exchange (SZSE), as well as one of the most well-known stock indices of the Chinese stock market A-share. The SZSE Component Index is designed to represent the performance of the multi-tier Shenzhen stock market, providing market participants with a global benchmark and effective investment tools in an innovative and fast-growing market. The SZSE Component index is weighted with float-free capitalisation and comprises the 500 largest and most liquid A-share shares listed and traded on the Shenzhen Stock Exchange. The index was created on 20 July 1994 with a base value of 1000. On 20 May 2015 the number of constituent securities increased from 40 to 500.

viii)**Euronext 100:** is the largest capitalised and most actively traded stock index on the Euronext stock exchange. Its composition is reviewed

quarterly; companies listed on the Euronext 100 belong to the nations of France, the Netherlands, Portugal, Belgium and Luxembourg; these companies must market annually at least 20% of their issues (calculated on a roundabout basis). It is considered one of the leading benchmarks in the European stock market.

ix) **HANG SENG:** is a capitalization-weighted stock market index. It is used to record and monitor the daily changes of the major companies in the Hong Kong stock market and is the main indicator of the global market performance of the aforementioned square. It consists of 50 companies representing about 58% of the capitalisation of the Hong Kong Stock Exchange. It was launched on 24 November 1969 and is currently compiled and operated by Hang Seng Indexes Company Limited, a wholly owned subsidiary of Hang Seng Bank, one of the largest registered and listed banks in Hong Kong in terms of market capitalisation.

Above we presented only the returns of the Dow Jones index, in the Appendix are presented the returns of all the other eight indices.

Chapter 2

Contagion and clustering among financial time series

2.1 Abstract

The objective of this work is to analyse the behaviour of financial time series in risky scenarios in order to enable the implementation of performing portfolio diversification strategies. The data used are the daily prices of the nine most capitalised stock indices: the Dow Jones, S&P500, Nasdaq 100, FTSE 100, Nikkei 225, SSE Composite, SZSE Component, Euronext 100, HANG SENG. Below is a methodology for grouping financial time series according to association in the tail of their distribution. The work takes as reference the article of F. Durante, R. Pappadà and N. Torelli (2013) and has been extended, however, also taking into account a quantification of the uncertainty present on clustering, based on the posterior distributions of the bivariate measures of dependence. The methodology for obtaining the posterior distribution takes as a reference the work in C.Grazian and B.Liseo (2017).

2.2 The methodology

Let $(x_{it})_{t=1,...,T}$ be a matrix of d financial time series ($i = 1, 2, ..., d$) representing the returns of different assets or stock indices. In order to determine suitable groups among the considered time series according to their pairwise tail dependence, we need to proceed in different steps:
1. Fit a suitable time series model in order to produce series of indipendent residuals

2. Create a pseudo-data matrix using the empirical CDF on each single marginal
3 Select a suitable way to measure the association between time series in the tails, i.e. determine a criterion under which we may assign a number to the strength of the (positive) dependence between the time series in a given tail region of their domain.
4. For each pair of pseudo-data, we determine the posterior distribution of the Spearman's correlation coefficient ρ based on the simulation of a posterior sample,using the methodology described in Schennach, (2005) and C.Grazian and B.Liseo (2017).
5. For each of these ρ_{ij}, we simulate 1000 realizations from its posterior distribution.
6.We calculate a Dissimilarity Matrix.
7. Apply a suitable cluster algorithm, according to the general characteristics of the above introduced dissimilarity measure.

Each step will be explained in detail in the rest of the section.

2.2.1 Fit a suitable time series model:

Time series models constitute a standard tool in order to describe the link among different univariate time series in a flexible way (see Patton 2012, and the references). Here, before applying our clustering procedure, we assume that the time series $(x_{it})_{t=1,...,T}$ are generated by the stochastic process $(X_t, \mathcal{F}_t)$ such that, for $i = 1, ..., d$, one has

$$X_{it} = \mu_i(Z_{t-1}) + \sigma_i(Z_{t-1})\epsilon_{it} \tag{2.1}$$

where Z_{t-1} depends on $\mathcal{F}_{t-1}$, the available information up to time $t-1$, the innovations ϵ_{it} are distributed according to a distribution function F_i, for each t and may be jointly dependent for the same time point t. In other words, we will assume that the conditional mean and variance of each univariate time series are modelled using some parametric specification that allows for a wide variety of models for the conditional mean (e.g., ARMA models, vector autoregressions, and others) and for the conditional variance (e.g., ARCH, GARCH, GJR-GARCH, etc.). Moreover, the innovations ϵ_{it} are assumed to have a constant conditional distribution F_i (with mean 0 and variance 1,

for identification) such that, for every t, the joint distribution function of $(\epsilon_{1t}, ..., \epsilon_{dt})$ can be expressed in the form $C(F_1, ..., F_d)$ for some copula C. As known (see, e.g.,Jaworski et al. 2010, 2013), the copula C is exactly the function that captures the dependence properties of the time series. Both the rank-invariant measures of association (like Kendall's τ, Spearman's ρ and their conditional versions) and the tail dependence coefficients are based on the calculation (in a parametric or non-parametric way) of the respective copula. Moreover, the estimation of the model may be done in two steps: first, we fit the marginal distributions, then we estimate the copula among them. Following these considerations, our procedure to deal with time series can be so summarized:
i. We fit an appropriate ARMA-GARCH model to each univariate time series. The choice of this univariate model is made by classical model selection procedures (e.g., Bayesian Information Criteria) and the goodness of fit may be verified by classical tests of omoscedasticity and uncorrelatedness of the residuals (Patton 2012, 2013). Notice that different models (with different parameters) can be estimated for each univariate time series.
ii. Using the parametric models estimated in previous step, we construct the estimated standardized residuals, for each $i = 1, ..., d$ by

$$\hat{\epsilon}_{it} = \frac{x_{it} - \hat{\mu}_i(Z_{t-1})}{\hat{\sigma}_i(Z_{t-1})} \tag{2.2}$$

iii. The estimated standardized residuals are converted to the estimated probability integral transform variables $z_{it} = F_i(\epsilon_{it})$, where F_i may be estimated from a parametric model (Gaussian, Student t, etc.) or by using the empirical distribution function.

2.2.2 Measuring the tail dependence

One possible way to consider such a kind of dependence is to restrict to a conditional version of the classical Pearson correlation coefficient, as done for instance by Longin and Solnik (2001), Malevergne and Sornette (2006). However, in analogy with the classical correlation coefficient (see, for instance, Embrechts et al. 2002), Pearson's correlation coefficient is often an inappropriate dependence measure since, firstly, it measures only linear dependence, secondly, it is not invariant to a change of the univariate margins, and thirdly,

it is very sensitive to outliers (Schmid and Schmidt 2007). In order to overcome these pitfalls of linear correlations, we suggest to use the Spearman's correlation coefficient ρ_S. In fact, we recall that ρ_S is invariant under rank transformation of the marginals and only depends on the copula of the involved random variables. Moreover, it is more robust than linear correlation, as stressed for instance by Croux and Dehon (2010). Since we are interested in the tail of the time series, we will use as a measure of tail dependence a suitable conditional version of ρ_S, i.e. we calculate ρ_S conditional on the fact that the time series are below a given threshold. Specifically, given two random variables (X_i, X_j) and a threshold $\alpha \in (0, 1)$ representing the risky level, we are interested in the Spearman's correlation $\rho_S^{ij}(\alpha)$ of the conditional distribution of $(X_i, X_j) \| (X_i, X_j) \in T_\alpha^{ij}$, where $T_\alpha^{ij} = (-\infty, q_\alpha(X_i)] \times (-\infty, q_\alpha(X_j)]$ is a set of non-zero probability and $q_\alpha(X_i)$ is the $\alpha-$quantile of X_i for every i. As known from Dobric et al. (2007), Durante and Jaworski (2010) (compare also with Bernard et al. 2013), $\rho_S^{ij}(\alpha)$ can be calculated via:

$$\rho_S^{ij}(\alpha) = 12 \int_{[0,1]^2} C_{T_\alpha^{ij}}(u, v) du dv - 3 \tag{2.3}$$

where $C_{T_\alpha^{ij}}$ is the (threshold)copula associated with the conditional distribution function of $(X_i, X_j) \| (X_i, X_j) \in T_\alpha^{ij}$.

At the end of this procedure,in the work of F. Durante et al (2013) it is obtained a value $\rho_S^{ij}(\alpha)$ representing the association between time series i and j when both markets are experiencing severe losses.

In the present work, instead, we simulate the posterior distribution of the coefficient $\rho_S^{ij}(\alpha)$ for three different values of α.

The output of this computation consists of 108 vectors (we have 36 different pairs of indeces and for each of them we compute the ρ's for three different thresholds) of the type:

$$\rho_1^{ij}(\alpha), \rho_2^{ij}(\alpha), ..., \rho_h^{ij}(\alpha), ..., \rho_M^{ij}(\alpha)$$

where i and j represents the assets or indices; α are the threshold representing the risky level and M is the size of the posterior simulation.

2.2.3 Posterior distribution of the Spearman's correlation coefficient ρ

The method (C.Grazian and B.Liseo ,2017) is based on the simulation of a posterior sample weighted in terms of the Bayesian exponentially tilted empirical likelihood (Schennach, 2005). No assumption is made on the copula structure. The central tool in this approach is the empirical likelihood (Owen,2001); it is adopted an approximate Bayesian approach based on the use of a pseudolikelihood, along the lines of Mengersen et al. (2013). It is used a partially specified model where the prior distribution is explicitly elicited only on the quantity of interest. Its approximate posterior distribution is obtained via the use of the Bayesian exponentially tilted empirical likelihood approximation of the marginal likelihood of the quantity of interest, illustrated in Schennach (2005). This approximation of the true "unknown" likelihood function hopefully reduces the potential bias for incorrect distributional assumptions. The method's goal is the estimation of a given functional of interest of C, say $\phi(C)$. In this respect, it is adopted a semiparametric Bayesian strategy for estimating $\phi(C)$ where the parameter of interest is the particular functional ϕ for which we derive an approximated posterior distribution

$$\pi(\phi|x) \propto \pi(\phi)\hat{L}(\phi; x)$$

where $\hat{L}(\phi; x)$ is a nonparametric approximation of the likelihood function for ϕ. Empirical likelihood has been introduced by Owen (2001); it is a way of producing a nonparametric likelihood for a quantity of interest in an otherwise unspecified statistical model. Schennach (2005) proposed an exponentially tilted empirical likelihood which can also be interpreted as a semiparametric Bayesian procedure. Assume that our dataset is composed of n independent replicates $(x_1, ..., x_d)$ of some random vector X with distribution F and corresponding density f. Rather than defining the usual likelihood function in terms of f, the Bayesian exponentially tilted empirical likelihood is constructed with respect to a given quantity of interest, say ϕ, expressed as a functional of F, i.e. $\phi(F)$, and then a sort of profile likelihood of ϕ is computed in a nonparametric way. More precisely, consider a given set of generalized moment conditions of the form

$$E_F(h(X, \phi)) = 0$$

where $h(\cdot)$ is a known function, and ϕ is the quantity of interest. The resulting Bayesian exponentially tilted empirical likelihood $L_{BEL}(\phi; x)$ is defined as

$$L_{BEL}(\phi; x) = \prod_{i=1}^{n} p_i^*(\phi)$$

where $(p_1^*(\phi), \ldots, p_n^*(\phi))$ is the solution of

$$max_{p_1,\ldots,p_n} \sum_{i=1}^{n} (-p_i log p_i)$$

under the constraints $0 \leq p_i \leq 1$,$\sum_{i=1}^{n} p_i = 1$ and $\sum_{i=1}^{n} h(x_i, \phi) p_i = 0$ Here has beed replaced the empirical likelihood with the exponentially tilted empirical likelihood proposed by Schennach (2005), in order to guarantee a solid Bayesian justification of the procedure. In this particolar case the only moment we use is the expression of a non-parametric consistent of the Spearman's ρ , namely $\hat{\rho_n} = \frac{1}{n} \sum_{i=1}^{n} (\frac{12}{n^2-1} R_i Q_i) - 3\frac{n+1}{n-1}$, where R_i and Q_i are defined in figure 2.1 .

The figure 2.1 shows the approximate Bayesian semiparametric copula algorithm for the Spearman's ρ (Grazian and Liseo, 2017). The Spearman's ρ between X and Y is the correlation coefficient among the transformed variables $U = F_X(X)$ and $V = F_Y(Y)$ or, in a copula language $\rho = 12 \int_0^1 \int_0^1 (C(u, v) - uv) du dv = 12 \int_0^1 \int_0^1 C(u, v) du duv - 3$.

2.2.4 Create a dissimilarity measure

For each of the M $\rho_S^{ij}(\alpha)$ realizations we create a Dissimilarity Matrix that collects the information about the pairwise tail dependence among the posterior distribution values.

Then for each pair (i,j), for $i, j = 1, 2, \ldots, d$, we may define the M dissimilarity matrices $\Delta^h = (\Delta_{ij}^h)$, for h=1, ...,M ,whose elements are given by:

$$\Delta_{ij}^h = \sqrt{2(1 - \rho_h^{ij}(\alpha))} \tag{2.4}$$

Figure 2.1: ABSCop algorithm Spearman's ρ (C.Grazian and B.Liseo, 2017)

Given a sample of n pseudo-observations assumed from an unknown copula function

$$u = \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \\ \cdots & \cdots \\ u_{n1} & u_{n2} \end{pmatrix},$$

1: **for** $b = 1, \cdots, B$ **do**
2: Draw $\rho^{(b)}$ from its prior distribution, for example $\rho^{(b)} \sim \mathcal{U}nif(-1, 1)$
3: Compute a nonparametric estimate of the Spearman's ρ:

$$\hat{\rho}_n = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{12}{n^2 - 1} R_i Q_i \right) - 3 \frac{n+1}{n-1},$$

where $R_i = \sum_{k=1}^{n} \mathbb{I}(u_{1k} \leq u_{1i}), Q_i = \sum_{k=1}^{n} \mathbb{I}(u_{2k} \leq u_{2i}), \quad i = 1, \ldots, n.$
4: Compute $L_{BEL}(\rho^{(b)}; u) = \omega_b$
5: **end for**
6: **return** A weighted sample of size B of values for ρ, where the weights are defined as the L_{BEL}, given the nonparametric estimate $\hat{\rho}_n$.
7: Sample with replacement $(\rho^{(b)} \omega_b)$, $b = 1, \cdots, B$.

Output: a sample of size B of values approximately from the posterior distribution of ρ.

2.2.5 Cluster building

For each dissimilarity matrix we apply a Partition Around Medoids (PAM) cluster algorithm. The PAM algorithm works for a fixed number K of groups. We implemented the algorithm with different values, namely K=3 and K=4. The final output will be M different clustering and the results were synthesized through the probability that the values linked to the indices i and j are in the same cluster or not. It is worth noticing that different clustering structure will be obtained with different values of α; this highlights the impact of the choice of the threshold α in the portfolio strategy.

2.3 Application to real data

To illustrate our approach we analyze the log-returns of the nine most capitalized stock indices: the Dow Jones, the S&P500, the Nasdaq 100, the FTSE 100, the Nikkei 225, the SSE Composite, the SZSE Component, the Euronext 100, the HANG SENG in the period from January 2, 2012 to October 11, 2022. The data were downloaded from Bloomberg and are formed by the daily closing prices of the nine stock indices specified above. We calculated the log-returns from stock prices and fit ARMA-GARCH models to each univariate time series with Student-t distributed errors to account for heavy tails. The choice of this univariate model is made by classical model selection procedure, in this specific case the Akaike Information Criterion (AIC): the most appropriate model will be the one with the lowest AIC value. By using the procedure explained in section 2.2.1 we derive hence the sample $(z_t^1, ..., z_t^N)$ on $[0, 1]$.

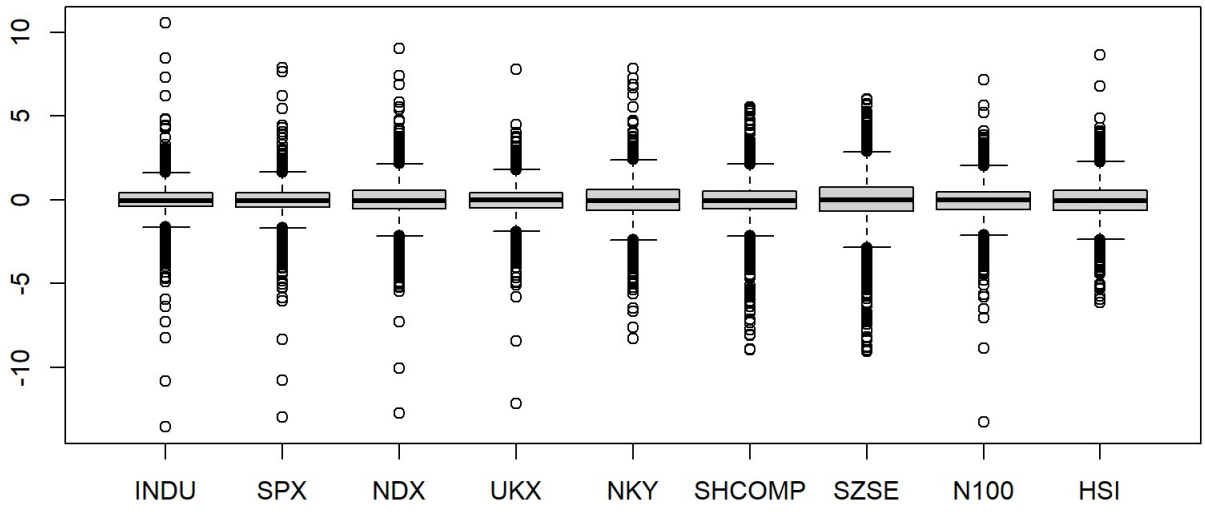

Figure 2.2: ARMA GARCH(1,1) residuals of the nine indexes

Moreover, the innovations ϵ_{it} are assumed to have a constant conditional distribution F_i (with mean 0 and variance 1, for identification) such that, for every t, the joint distribution function of $(\epsilon_{1t}, ..., \epsilon_{dt})$ can be expressed in the form $C(F_1, ..., F_d)$ for some copula C.

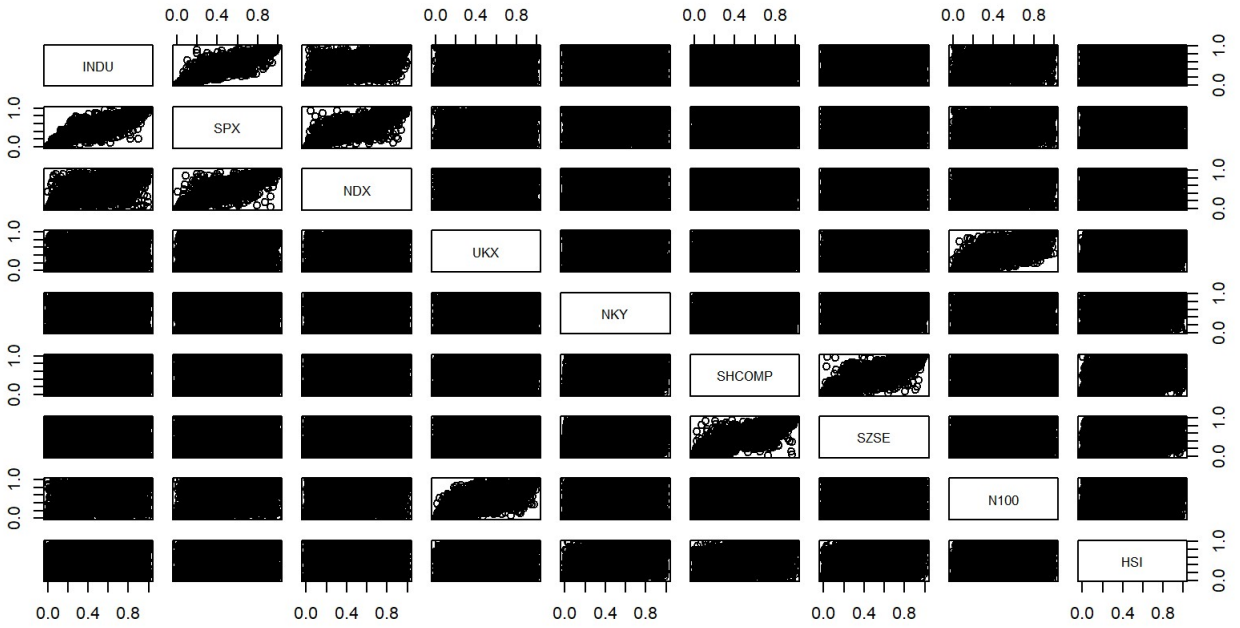

Figure 2.3: Bivariate Plots of Residuals in terms of pseudo-data

In order to restrict our analysis to extreme observations we fix a threshold α denoting the "degree" of risk of the scenario we are considering. We set three different values of α : 0,01 ; 0,25 ; 0,1 . For each α we compute $d(d-1)/2 = 9(8)/2 = 36$ posterior distributions of the Spearman's correlation coefficient ρ associated with the pairs $(z_t^i, z_t^j), i \neq j, i, j = 1, ..., 9$ conditional to the fact that (z_t^i, z_t^j) takes values on $[0, \alpha]^2$ and denote them by

$$\rho_1^{ij}(\alpha), \rho_2^{ij}(\alpha), ..., \rho_h^{ij}(\alpha), ..., \rho_M^{ij}(\alpha)$$

In the figure 2.4 we show the histogram of the values of the posterior distribution of ρ for the couple Dow Jones and S&P500, calculated for the value of α : 0,01. In the Appendix we present the complete series of the 108 histograms that graphically represents the ρ posterior distributions for each couple of indices analyzed for the three values fo α .

For each α value we obtain 1.000 realizations of the ρ and, for each of them, we create a dissimilarity matrix that collects the information about the pairwise tail dependence among the posterior distribution values. The starting point for our clustering procedure is a distance defined through the Spearman's correlation matrix. We apply the PAM cluster algorithm directly

to the matrix $\Delta = (\Delta_{ij})$,for $j = 1, \ldots, 9$ with

$$\Delta_{ij} = \sqrt{2(1 - \rho_h^{ij}(\alpha))}$$

and we select two values for K : 3 and 4.

At the end of our analysis the results were summarised by the following six tables in figure 2.5 and 2.4, one for each combination of the three values of α and two values of K. The following tables show the probability for each pair of indices to be in the same cluster for a specific level α . We can study the correlation of the financial time series analyzed in this way: if the probability for each pair of indices has a high value only for $\alpha = 0,01$ it implies that the two indices are related in very risky scenarios, if otherwise the probability is high even with $\alpha = 0,25$ it means that the two indices have a high correlation and therefore a common behavior . If the probability is high with $\alpha = 0.25$ and low with $\alpha = 0.01$ it means that when the market is stable the two indices have a common behavior between them, and when the market is risky they adopt a different behavior. This is very important for investors and for those who make portfolio choices because they must choose which stocks to invest in order to diversify portfolio risk. For example from our results we can see that the Dow Jones and the S&P500 have a high correlation; infact for K=3 the probability to be in the same cluster are equal to 1 for $\alpha = 0,25$, equal to 0,84 for $\alpha = 0,1$ and equal to 0,60 for $\alpha = 0,01$.This correlation between the Dow Jones and the S&P500 is present also for the K=4. So we can state that the two indices have a common behavior not only in risky scenarios. Completely opposite are the results obtained for the Dow Jones and FTSE 100 index : for K=3 the probability to be in the same cluster are equal to 0,02 for $\alpha = 0,25$, equal to 0,08 for $\alpha = 0,1$ and equal to 0,44 for $\alpha = 0,01$. So we can conclude that the two indices have a common behavior in risky scenarios and a different behavior when the market is stable. In fact, it may happen that the dependence between two assets changes according to different "crisis" periods both the assets are experiencing.

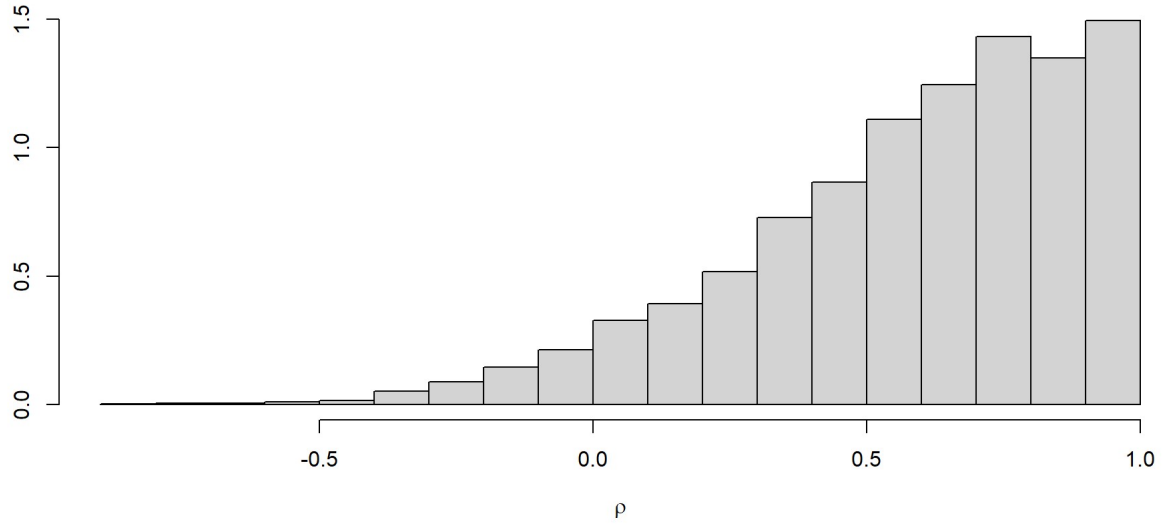

Figure 2.4: posterior distribution of ρ for Dow Jones and S&P500 (α= 0,01)

2.4 Conclusions

We have presented a methodology to analyze the phenomenon of financial contagion (Durante and Foscolo 2013) between financial time series in risky scenarios. From a practical point of view to use this type of methodology the first important thing to establish is the value of α in coherence with desired investment strategy. The main steps of our method are:
1.Fit a suitable copula-based time series model,
2.Measuring the tail dependence through the estimation of its posterior distribution,
3.Create M dissimilarity measure,
4.Apply a cluster algorithm
This analysis has shown how this methodology can be applied in portfolio selection. We have shown how important it is to consider the behavior of financial time series in risky scenarios to diversify risk. Figures 2.7 and 2.8 show in blue the Markowitz's Efficient Frontier calculated on the portfolio made up of all nine indices examined in this work, while in red the efficient frontier calculated on three and four indices respectively belonging to a different cluster. In Figure 2.7 the three indices chosen, one for each cluster, are: S&P 500, SZSE Component, Hang Seng. In Figure 2.8 the four indices are: Dow Jones, Nikkei225, SZSE Component, Hang Seng. The two graphs

show that in very risky scenarios the diversification of the portfolio applied according to the method of clustering presented in this paper, represented by the efficient frontier in red, enables portfolio returns to be higher than the total portfolio of the nine indices. In fact, as the risk increases, the efficient frontier in red tends to have higher returns than the efficient frontier in blue.

K = 3									
α = 0,01	INDU	SPX	NDX	UKX	NKY	SHCOMP	SZSE	N100	HSI
INDU	1,00	0,60	0,42	0,44	0,08	0,27	0,25	0,41	0,07
SPX	0,60	1,00	0,48	0,40	0,07	0,24	0,25	0,38	0,16
NDX	0,42	0,48	1,00	0,38	0,15	0,27	0,25	0,33	0,23
UKX	0,44	0,40	0,38	1,00	0,23	0,25	0,26	0,44	0,22
NKY	0,08	0,07	0,15	0,23	1,00	0,35	0,38	0,27	0,47
SHCOMP	0,27	0,24	0,27	0,25	0,35	1,00	0,39	0,27	0,31
SZSE	0,25	0,25	0,25	0,26	0,38	0,39	1,00	0,28	0,24
N100	0,41	0,38	0,33	0,44	0,27	0,27	0,28	1,00	0,21
HSI	0,07	0,16	0,23	0,22	0,47	0,31	0,24	0,21	1,00

K = 3									
α = 0,25	INDU	SPX	NDX	UKX	NKY	SHCOMP	SZSE	N100	HSI
INDU	1,00	1,00	1,00	0,02	0,15	0,00	0,00	0,02	0,02
SPX	1,00	1,00	1,00	0,02	0,15	0,00	0,00	0,02	0,02
NDX	1,00	1,00	1,00	0,02	0,15	0,00	0,00	0,02	0,02
UKX	0,02	0,02	0,02	1,00	0,77	0,00	0,00	0,98	0,57
NKY	0,15	0,15	0,15	0,77	1,00	0,07	0,07	0,76	0,46
SHCOMP	0,00	0,00	0,00	0,00	0,07	1,00	1,00	0,00	0,40
SZSE	0,00	0,00	0,00	0,00	0,07	1,00	1,00	0,00	0,40
N100	0,02	0,02	0,02	0,98	0,76	0,00	0,00	1,00	0,57
HSI	0,02	0,02	0,02	0,57	0,46	0,40	0,40	0,57	1,00

K = 3									
α = 0,1	INDU	SPX	NDX	UKX	NKY	SHCOMP	SZSE	N100	HSI
INDU	1,00	0,84	0,74	0,08	0,30	0,14	0,17	0,06	0,19
SPX	0,84	1,00	0,76	0,08	0,31	0,14	0,16	0,05	0,19
NDX	0,74	0,76	1,00	0,08	0,35	0,18	0,21	0,05	0,18
UKX	0,08	0,08	0,08	1,00	0,27	0,20	0,25	0,79	0,53
NKY	0,30	0,31	0,35	0,27	1,00	0,25	0,24	0,24	0,33
SHCOMP	0,14	0,14	0,18	0,20	0,25	1,00	0,72	0,21	0,10
SZSE	0,17	0,16	0,21	0,25	0,24	0,72	1,00	0,28	0,08
N100	0,06	0,05	0,05	0,79	0,24	0,21	0,28	1,00	0,53
HSI	0,19	0,19	0,18	0,53	0,33	0,10	0,08	0,53	1,00

Figure 2.5: Correlation for K = 3: green color represents positive correlation and red color represents the negative one

K = 4									
α = 0,01	INDU	SPX	NDX	UKX	NKY	SHCOMP	SZSE	N100	HSI
INDU	1,00	0,52	0,31	0,35	0,04	0,19	0,18	0,30	0,03
SPX	0,52	1,00	0,36	0,30	0,04	0,16	0,17	0,26	0,08
NDX	0,31	0,36	1,00	0,30	0,08	0,19	0,18	0,24	0,14
UKX	0,35	0,30	0,30	1,00	0,16	0,17	0,19	0,35	0,13
NKY	0,04	0,04	0,08	0,16	1,00	0,25	0,26	0,20	0,33
SHCOMP	0,19	0,16	0,19	0,17	0,25	1,00	0,28	0,20	0,18
SZSE	0,18	0,17	0,18	0,19	0,26	0,28	1,00	0,23	0,12
N100	0,30	0,26	0,24	0,35	0,20	0,20	0,23	1,00	0,11
HSI	0,03	0,08	0,14	0,13	0,33	0,18	0,12	0,11	1,00

K = 4									
α = 0,25	INDU	SPX	NDX	UKX	NKY	SHCOMP	SZSE	N100	HSI
INDU	1,00	1,00	0,99	0,01	0,04	0,00	0,00	0,00	0,01
SPX	1,00	1,00	1,00	0,01	0,04	0,00	0,00	0,00	0,01
NDX	0,99	1,00	1,00	0,01	0,04	0,00	0,00	0,01	0,01
UKX	0,01	0,01	0,01	1,00	0,24	0,00	0,00	0,99	0,34
NKY	0,04	0,04	0,04	0,24	1,00	0,01	0,01	0,24	0,19
SHCOMP	0,00	0,00	0,00	0,00	0,01	1,00	1,00	0,00	0,18
SZSE	0,00	0,00	0,00	0,00	0,01	1,00	1,00	0,00	0,18
N100	0,00	0,00	0,01	0,99	0,24	0,00	0,00	1,00	0,34
HSI	0,01	0,01	0,01	0,34	0,19	0,18	0,18	0,34	1,00

K = 4									
α = 0,10	INDU	SPX	NDX	UKX	NKY	SHCOMP	SZSE	N100	HSI
INDU	1,00	0,79	0,66	0,04	0,16	0,09	0,14	0,04	0,12
SPX	0,79	1,00	0,71	0,05	0,16	0,09	0,13	0,02	0,12
NDX	0,66	0,71	1,00	0,05	0,23	0,11	0,17	0,04	0,11
UKX	0,04	0,05	0,05	1,00	0,16	0,10	0,20	0,74	0,31
NKY	0,16	0,16	0,23	0,16	1,00	0,11	0,12	0,12	0,18
SHCOMP	0,09	0,09	0,11	0,10	0,11	1,00	0,64	0,12	0,02
SZSE	0,14	0,13	0,17	0,20	0,12	0,64	1,00	0,20	0,03
N100	0,04	0,02	0,04	0,74	0,12	0,12	0,20	1,00	0,31
HSI	0,12	0,12	0,11	0,31	0,18	0,02	0,03	0,31	1,00

Figure 2.6: Correlation for K = 4 : green color represents positive correlation and red color represents the negative one

Legend								
INDU	SPX	NDX	UKX	NKY	SHCOMP	SZSE	N100	HSI
Dow Jones	S&P500	Nasdaq 100	FTSE 100	Nikkei 225	SSE Composite	SZSE Component	Euronext 100	HANG SENG

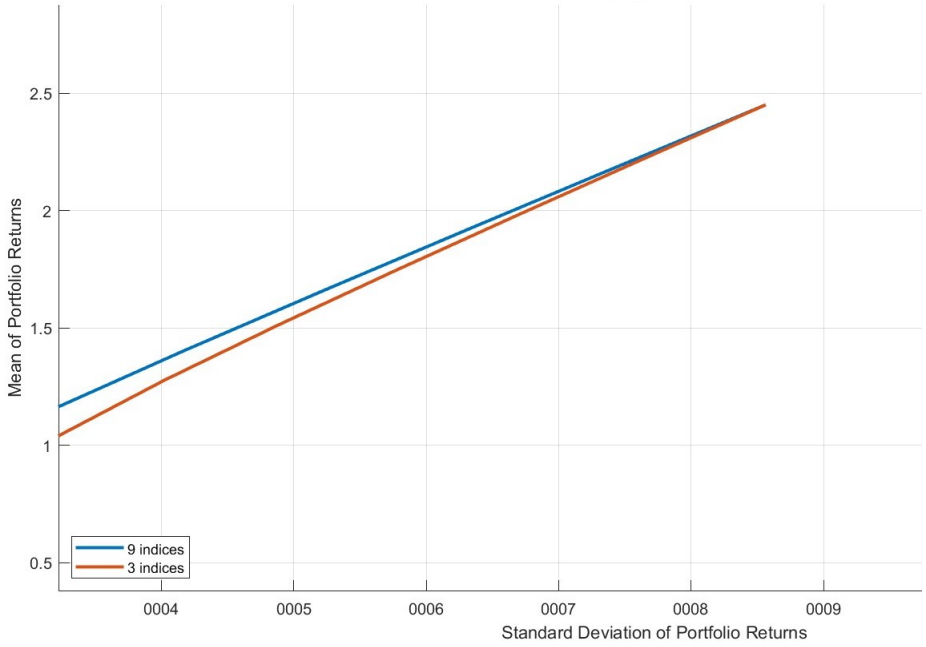

Figure 2.7: Efficient Frontier 9 indices and 3 indices

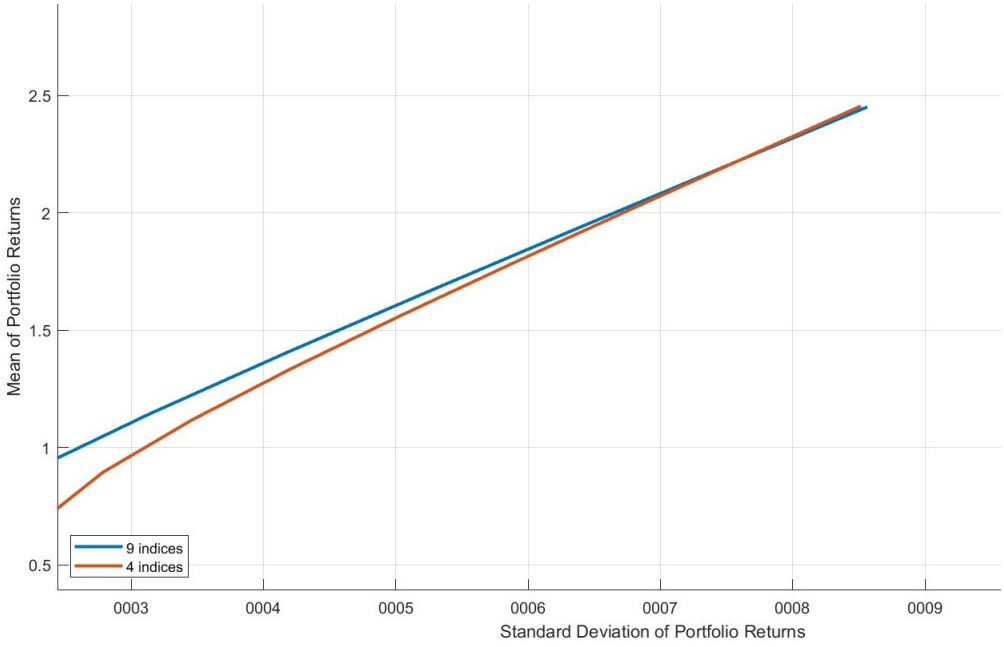

Figure 2.8: Efficient Frontier 9 indices and 4 indices

Chapter 3

Application EVT and Copulas

3.1 Abstract

The following chapter describes the possible methodologies for calculating stock indices, and describes the nine indices that make up the portfolio under study: Dow Jones, S&P500, NASDAQ 100, FTSE 100, Nikkei 225, SSE Composite,the SZSE Component, the Euronext 100, the HANG SENG. The time horizon is from 2 January 2012 to 9 September 2022, for a total of about 10 years. Then we proceed to the explanation of the techniques used for the construction of the models that sees the combined use of the theory of extreme values and copulas: t , skew t and Vine, through a two-step approach: the first of univariate modeling and the second of multivariate modeling. Finally, the obtained portfolio index returns are simulated and the Value at Risk is calculated with relative back-testing to test the goodness of the presented model. The back-test is made on one month period: from 12 September 2022 to 11 October 2022.

3.2 Stock Indices

The stock indices are the synthesis of the value of the basket of the stocks that they represent, and the movements of the index are a good approximation of the variation in the time of the valorization of the securities included in the portfolio. There are different methods of calculating indices, depending on the weighting that is attributed to the shares in the basket. We can therefore distinguish between:

i)Equally weighted indices: these are characterised by equal weighting factors for all the securities making up the index. It does not matter the capitalization of the included companies, because all the stocks of the index have the same weight;
ii) Price weighted indices: in this case the weight associated with each stock varies according to its price (if the price of one stock increases more than the others, automatically it also increases its weight within the index). They are very easy to calculate because they are given by the simple sum of the prices of the securities that make up the index. These indices, however, have the disadvantage of not correctly reflecting the performance of the entire portfolio: in fact, the most "expensive" securities are represented more, regardless of the number of shares present and the size of the company;
iii)Value weighted indices: These solve the problems of the previous ones as the weight of each security is proportional to its stock market capitalization. In contrast to other calculation methodologies, in this case indices are adjusted and adjusted as a result of corporate transactions such as fractionations, groupings, payment of extraordinary dividends, divisions, free allocations or new issues for payment. Most major world indices are therefore calculated using the value weighted methodology.

3.3 Construction of the model

The combined use of extreme value theory and copulas leads to a market risk modelling approach that differs from traditional risk management models. They assume conditional normality for logarithmic returns on financial assets or risk factors despite empirical evidence that yield distributions are characterized by leptocurtic tails. Below is an application of this approach to market risk modelling that characterises the behaviour during financial and economic crises of the nine most highly capitalised stock indices, namely: Dow Jones, S&P500, Nasdaq 100, FTSE 100, Nikkei 225, SSE Composite, SZSE Component, Euronext 100, HANG SENG. The analysis consists of two phases, the first of univariate modeling and the second of multivariate modeling. In the first phase we estimate a probability distribution for each variable using a non-parametric smoothing technique for the central part of the distribution. We apply the extreme value theory to better characterize the extreme values found in the upper and lower tails. When this first phase is over we

will have nine separate univariate models , one for each stock index that constitutes the portfolio under consideration. In the multivariate phase, we tie these models apart with copulas to get a portfolio view, which will allow me to better analyze the data. We chose to use the three different copulas: t copula, skew t copula and Vine copula. Standard distributions often require that all univariate and multivariate marginal distributions are of the same type and only allow for highly symmetric dependence structures. These characteristics are rarely satisfied in applications. For this reason we have also used in our model the Vine copula (C.Czado and T.Nagler, 2021). By decoupling the univariate description of individual variables from the multivariate description of the dependency structure, copulas offer significant theoretical and computational advantages over traditional risk management techniques. Once these two steps have been completed and the copulas have been calibrated, we use the Monte Carlo simulation to calculate the Value at Risk of the equally weighted portfolio over a one-month timeframe.

3.4 Univariate modelling

To carry out the following analysis, it was considered appropriate to use the daily closing prices of the nine most capitalised equity indices: Dow Jones, S&P500, Nasdaq 100, FTSE 100, Nikkei 225, SSE Composite, SZSE Component, Euronext 100, HANG SENG. The relative price changes of each index are illustrated graphically below. The initial level of each index has been normalized to the unit, dividing the first value, or the closing price of January 2, 2012 for itself and then all the values to follow for the first value, to facilitate the comparison of performance between the various stock indices under consideration.

From the prices it was possible to calculate the returns of each index. For the sake of simplicity, we have chosen to report the graphs related to a single index, the Dow Jones, in order to streamline the treatment, but it is important to specify that all the steps explained below for the Dow Jones have been applied to all nine indices examined and constituting our portfolio. And all the graphs presented here for the Dow Jones are in the appendix for all the indices. To produce a series of independent and identically distributed observations, a first order (AR 1) autoregressive model has been adapted to

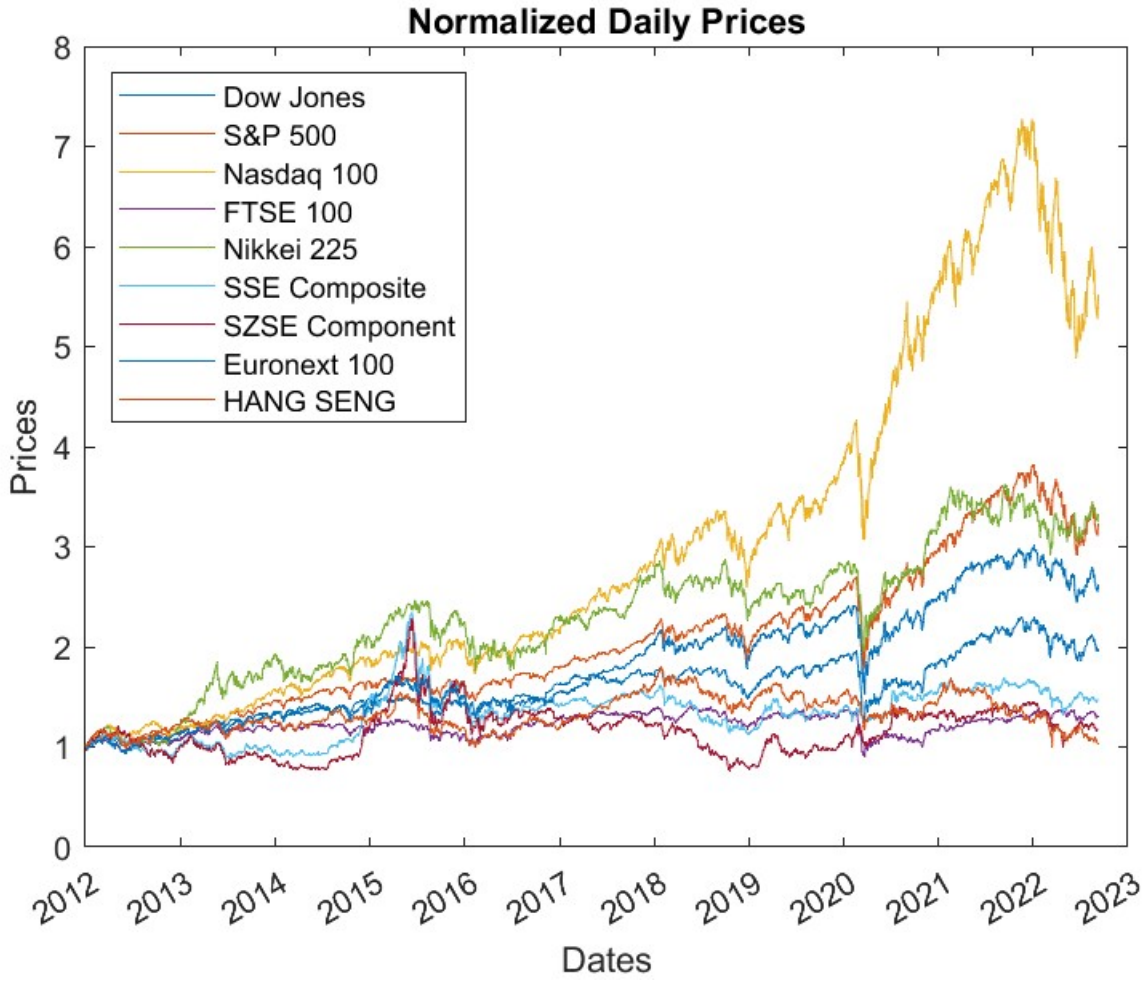

Figure 3.1: Monthly Prices Normalized of the nine indices

the conditional mean returns of each stock index

$$r_t = c + \theta\ r_{t-1} + \epsilon_t$$

and an asymmetric Garch (Generalized Autoregressive Conditional Heteroschedasticity) model for conditional variance

$$\sigma_t^2\ = k + \alpha\sigma_{t-1}^2\ + \phi\epsilon_{t-1}^2\ + \psi\ [\epsilon_{t-1} < 0\ ]\epsilon_{t-1}^2$$

The residuals of each index were modeled from a standardized Student t . These residuals represent the underlying with mean equal to 0 and variance equal to 1 independent and identically distributed on which the extreme values theory of the empirical Cumulative Distribution Function (CDF) estimate is based.

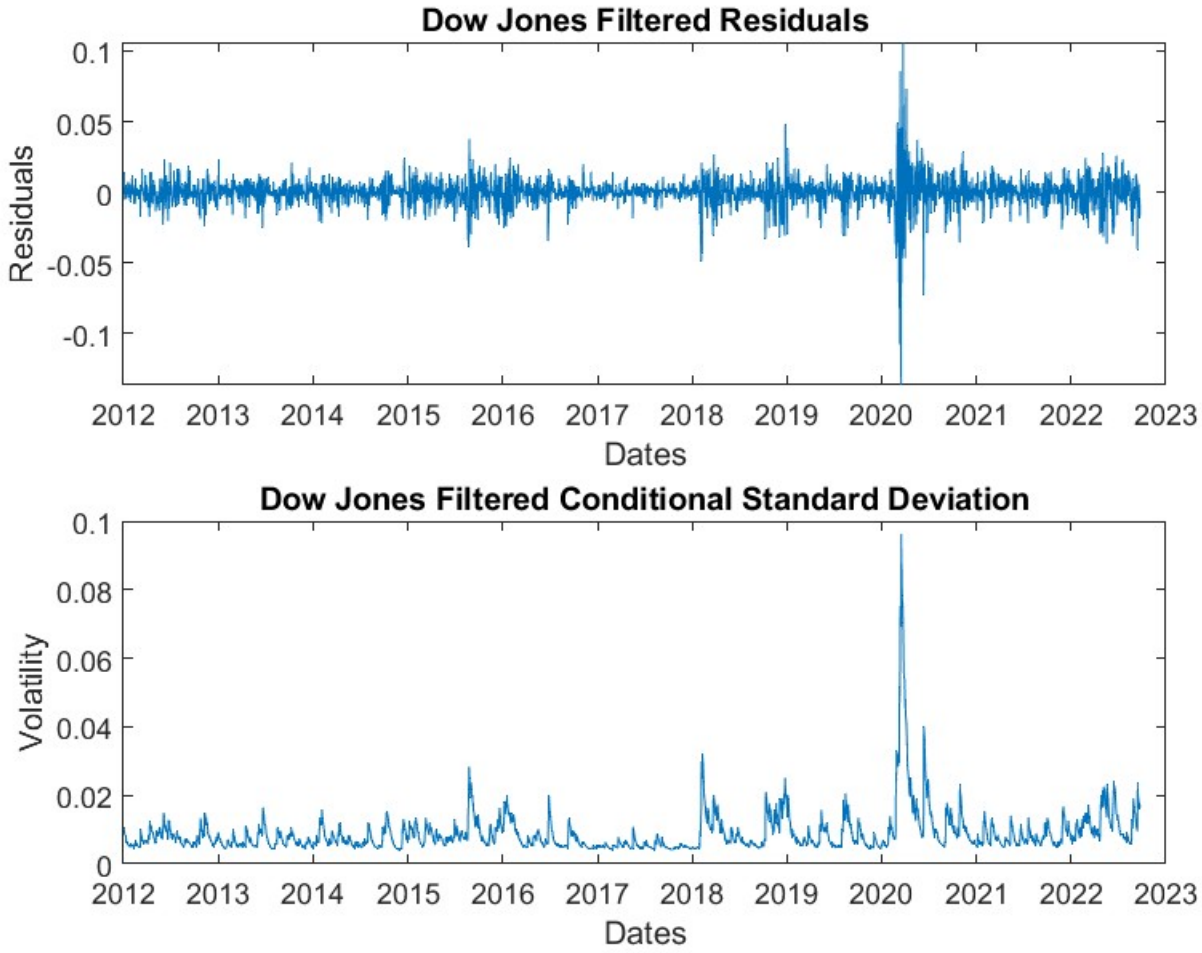

Figure 3.2: Residuals and Volatility Dow Jones

The correlograms presented in figure 3.3 show the autocorrelation function (ACF) of : returns, squared returns, residuals and squared residuals of the Dow Jones Industrial Average stock index. It is possible to note that standardized residuals respond well to our needs, in fact because the autocorrelation defines the degree of dependence between the values under consideration, if we look at the correlogram, we see taht there isn't evidence of serial dependence between the residuals.

Once the data is filtered, the goal is to find a probability distribution to model the daily movements of each index. It is not assumed that the data come from a normal distribution or from any other simple parametric distribution, since the goal is to find a more flexible empirical distribution that allows the data to speak for itself.

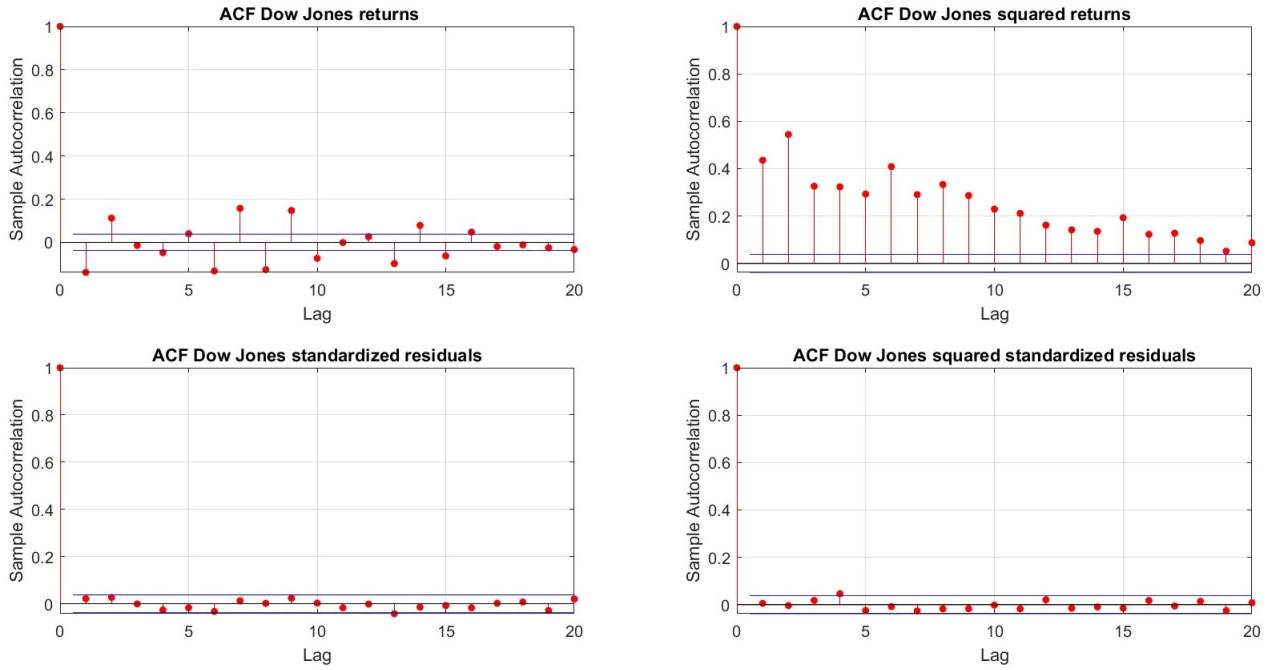

Figure 3.3: Autocorrelation function Dow Jones

•To estimate the CDF within the distribution we use the Kernel Gaussian , which works well within the distribution, where most of the data is located, but not when applied to tails;

•For tails it was chosen to use a generalized pareto distribution (GPD) that will provide a reasonable model of the most extreme observations, large losses and large gains. For risk management, it is essential to accurately characterise tails.

So you get the empirical cumulative distribution function (CDF) for each index, below is the Dow Jones CDF.

Finally, we proceed to the evaluation of the generalised Pareto distribution (GPD) before repeating the steps explained so far for each index in the portfolio. It was chosen to use a Matlab function (Machine Learning Toolbox gpfit) by which empirical CDF data is used to find parameters for the GPD in the tails of the curve, it compares ,then, the empirical CDF with the fitted Generalized pareto CDF and you immediately notice that it is a good approximation.

When the first phase, that is, univariate modeling is completed, there will

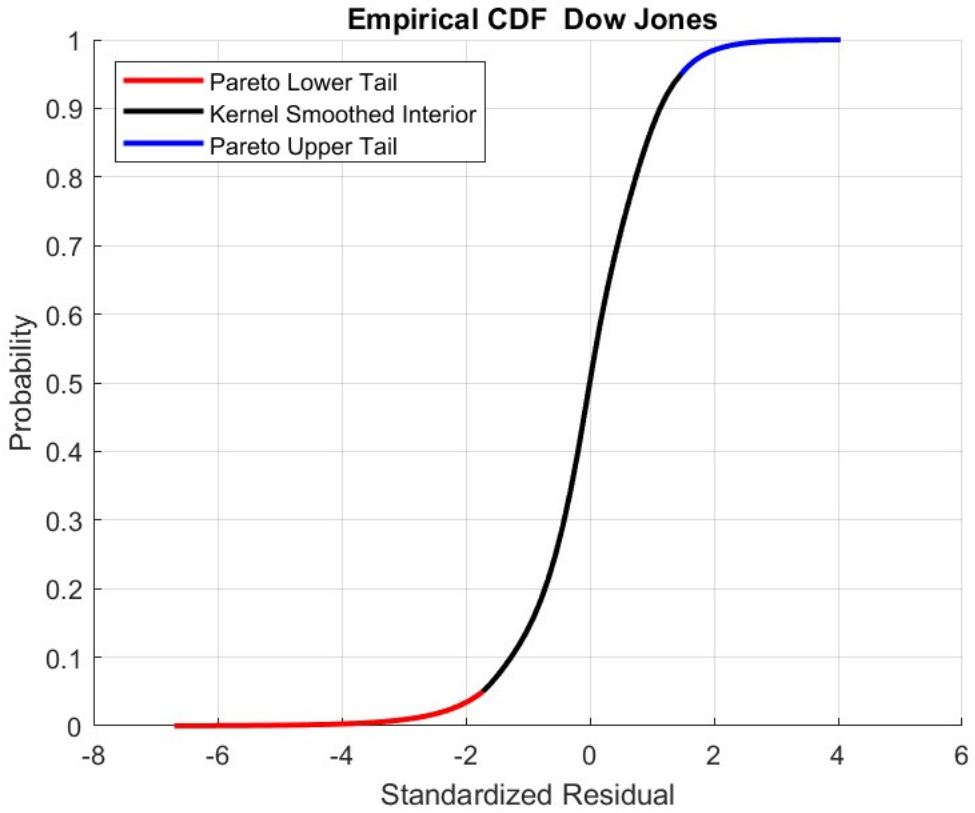

Figure 3.4: Empirical CDF Dow Jones

be nine separate univariate models describing the distribution of gains and losses over the chosen time interval. However, it is necessary for the purpose of this applicative analysis to bind these separate models together, and copulas has been chosen for this purpose.

3.5 Multivariate modelling

Since a copula is a multivariate probability distribution whose individual variables are uniformly distributed, it is now possible to use newly derived univariate distributions to transform the individual data of each index into a uniform scale, which is the module needed to build a copula. This step is very important because it has the calibration of the copulas: t , skew t and Vine. But first it is good to understand which type of copula is most suitable for our purposes. Having common movements in extremes is a widespread phenomenon in the real world, for example if the Canadian index is down 30% today, you can be reasonably sure that the US market has also suffered

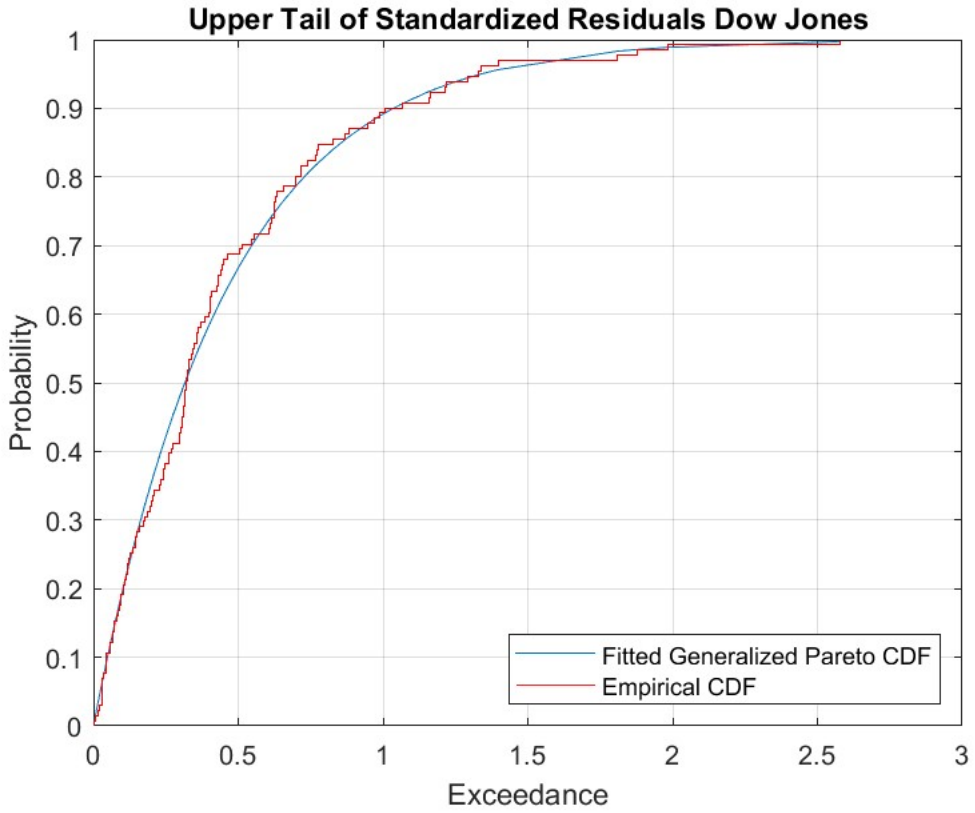

Figure 3.5: Upper tail of standardized residuals Dow Jones

a relatively large fall. Modeling indices with a Gaussian copula does not capture this behavior, because the most extreme events of individual indices in a Gaussian copula model would be independent of each other. The t copula, however, includes a parameter given by the degrees of freedom that can be used to model the tendency for extreme events to occur jointly. In this application the t copula has been calibrated by estimating the scalar parameter indicating the degrees of freedom and its linear correlation matrix through the maximum likelihood: the calibration method used is known as canonical maximum likelihood (CML). This step does what previously corresponded to the estimate of the Cumulative Distribution Function (CDF) for a single index, that is, it finds a model for the interaction between the indices , in the behavior models of the single indices. We chose to build the model also with the skew-t copula (Azzalini Capitanio 2003) because the skew-t copula is particularly appealing in finance where the data are characterized by heavy tails and skewness, and where interest is in analysing conditional distributions. However, the skew-t model is intuitively appealing in that it extends the Nor-

mal distribution by permitting tails that are heavy and asymmetric. In our model we used the dependency structure implied by the skew-t of Azzalini and Capitanio (2003). It's important to consider that for the analysis of large multivariate data sets, flexible multivariate statistical models are required that can adequately describe also the multivariate tail behavior. Standard distributions, such as the multivariate normal or Student's t distribution, are too inflexible in their marginal and joint behavior. They often require that all univariate and multivariate marginal distributions are of the same type and only allow for highly symmetric dependence structures. These characteristics are rarely satisfied in applications. For this reason we have also used in our model the Vine copula (C.Czado and T.Nagler, 2021). The class of multivariate copulas was limited for a long time to elliptical (including the Gaussian and t-copula) and Archimedean families (such as Clayton and Gumbel copulas). Both classes are rather restrictive with regard to symmetry and tail dependence properties. The class of vine copulas overcomes these limitations by building a multivariate model using only bivariate building blocks. This gives rise to highly flexible models that still allow for computationally tractable estimation and model selection procedures. We use the R package VineCopula to model the marginal series we obtain by the process explained above.

3.6 Simulation of portfolio returns

You could use the model created by the previous steps for a wide range of applications, such as calculating the expected deficit or performing dynamic portfolio analysis. Here we have chosen to derive a measure of Value at Risk, as well as the maximum loss and maximum gain simulated. Since we now have a probability model that describes the observed data quite well, we are able to generate random daily returns that will be statistically equivalent to historical data. The Monte Carlo simulation is used to analyze not just one historical study but thousands, in fact, two thousand independent random tests of employee index returns over a period of one month (or 22 trading days) are simulated. Here, the Matlab GARCH Toolbox simulation engine is used to reintroduce autocorrelation and heteroschedasticity observed in the original index returns. Then, after simulating the returns of each index and forming a comparable global portfolio, it was decided to extract from the data

the maximum economic gain, as well as the var at various confidence levels 90 % 95 % 99 % over the risk horizon of one month. We then obtain the empirical CDF of the simulated returns of the global portfolio over a time horizon that goes from 12 September 2022 to 11 October 2022 starting from the sample of data in question that ends on 9 September 2022.

3.7 Calculation of Value at Risk

In the table are shown the obtained values of Value at Risk at 90 %, 95% and 99 %, in the second column starting from the left are presented the values obtained following the application of the modeling presented in our treatment, respectively for the t copula, skew t copula and Vine copula, while in the right column there are values obtained by calculating the Value at Risk on real data. As it was possible to find on the stock market the closing price data of the nine most capitalized indices analyzed, for the period from 12 September 2022 to 11 October 2022 it was considered appropriate to perform a verification of the goodness of the proposed model. It is important to explain what is meant by saying that the var 90 %, which corresponds to 10 % cumulative probability, is about -4,10 %: it means that, during the chosen time interval, with a 90% probability the portfolio under consideration will lose no more than 4,10 %. From our analysis we can notice that the application of the model with choice of the copula Skew t has the values of the Value at Risk closer to the real ones, differing for the 0,97 % in the case of the 90 % var.

Value at Risk	EVT & t copula	EVT & Skew t	EVT & Vine	Real Data
90% VaR	-4,10%	-2,86%	-3,89%	-1,89%
95% VaR	-6,17%	-4,13%	-6,28%	-2,98%
99% VaR	-10,46%	-7,10%	-11,57%	-4,02%

3.8 Conclusions

The main objective of this paper was to study a different approach from traditional portfolio risk estimation models that assume conditional normality for logarithmic returns of financial assets or risk factors. Empirical evidence shows that distributions of logarithmic returns on financial assets are characterised by leptocurtic tails. Consistent with this statistical phenomenon, which is typical of financial data in general, a new technique of generation of

the scenarios of the returns of the nine indices analyzed has been chosen to use, going to emulate and to describe through the use of the extreme values theory the empirical evidence of the data in examination. The two-step approach described allows us to simulate and model the stock returns, in line with historical data. In the first step a probability distribution for each series is estimated. When this first phase is over you will have as many univariate models as there are variables under consideration, but it is necessary to tie these separate models together and for this purpose we use a copula. By separating the univariate description of individual variables from the multivariate description of the dependency structure, copulas offer significant theoretical and computational advantages over traditional risk management techniques. Having common movements in extremes is a widespread phenomenon in the real world, for this reason, in the second phase, it was chosen to use the t copula and the skew-t copula that include a parameter given by the degrees of freedom that can be used to model the tendency for extreme events to occur jointly. Vine copula has also been chosen because it is a flexible multivariate statistical model able to adequately describe also the behavior of the multivariate tail, and it was considered useful to compare the output returned by the EVT model and Vine copula with the two previous more inflexible in their marginal and joint behavior. In this way the nine separate models are linked together, each for each stock index under consideration, with the copulas to finally get a portfolio view, which will allow to analyze the data better. After these two steps, it was decided to use the Monte Carlo simulation to calculate the Value at Risk of the equally weighted portfolio over a period of one month. However it should be noted that you could use the template created for a wide range of applications, such as calculating the expected deficit or performing dynamic portfolio analysis. Specifically, we estimated for our three different model ,implementing the three different copulas ,the Value at Risk at 90%, 95% and 99% of the portfolio consisting of the nine most capitalized equity indices and then we verified the goodness of our model by performing a backtest on real data, such as the closing prices of the nine indices analysed over the period from 12 September 2022 to 11 October 2022. The comparison between the VaR calculated on the sample of simulated data from the three different models implemented and the VaR calculated on the real data shows that among the three models implemented the EVT and Skew-t copula model has values of Value at Risk closer to that calculated on real data.

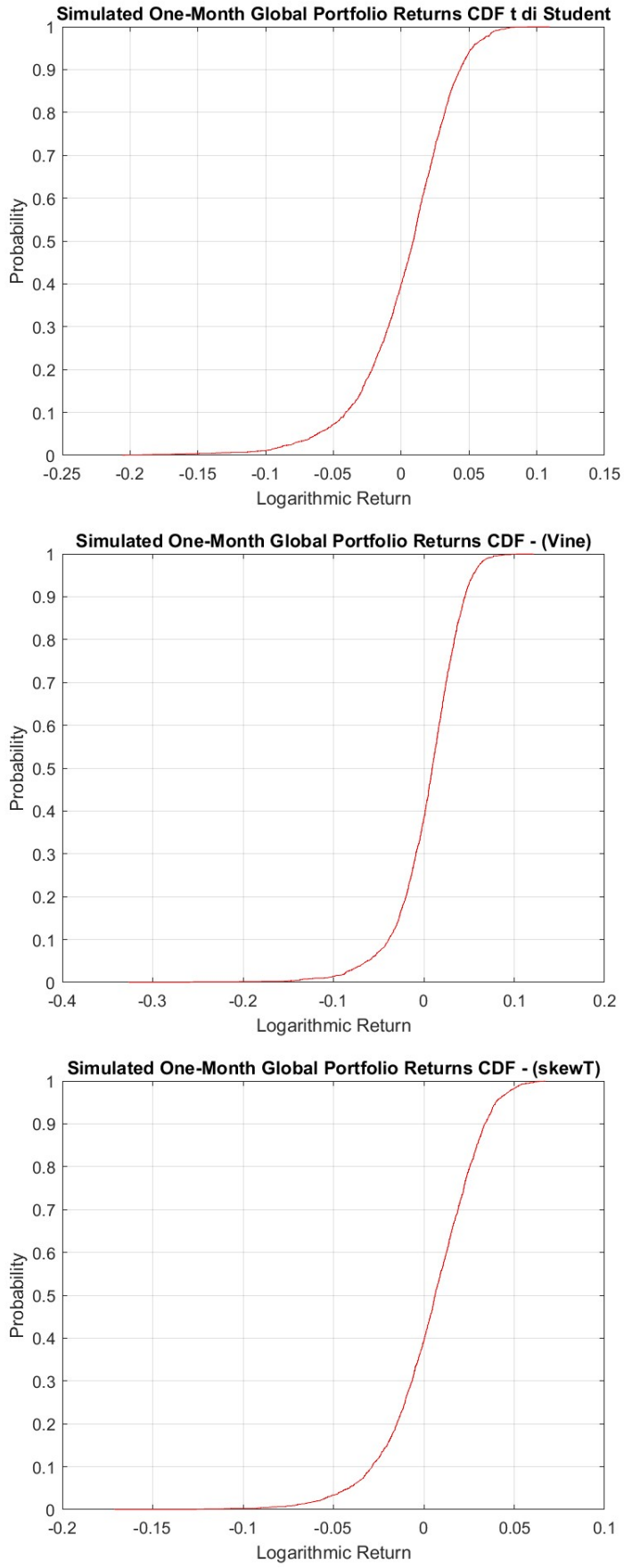

Figure 3.6: Simulated one-month global portfolio returns CDF

Appendix

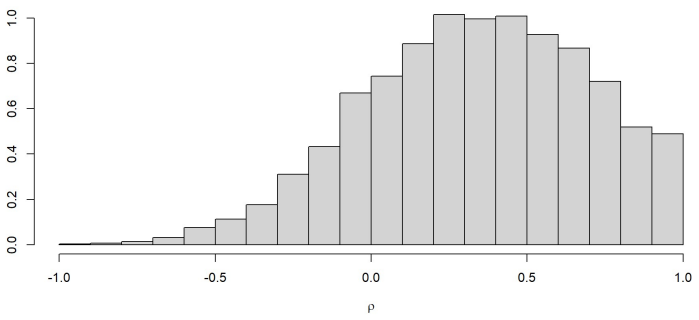

Figure 3.7: posterior distribution of ρ for Dow Jones and Nasdaq100 (α= 0,01)

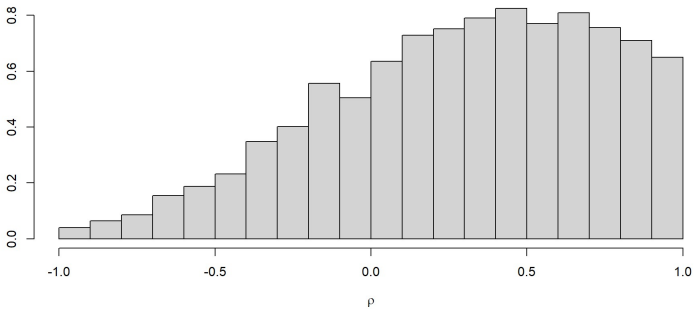

Figure 3.8: posterior distribution of ρ for Dow Jones and FTSE100 (α= 0,01)

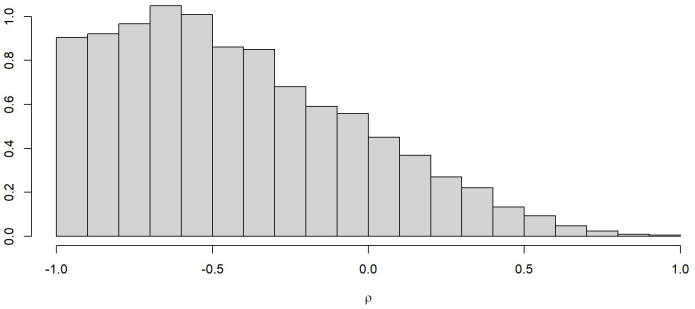

Figure 3.9: posterior distribution of ρ for Dow Jones and Nikkei225 (α= 0,01)

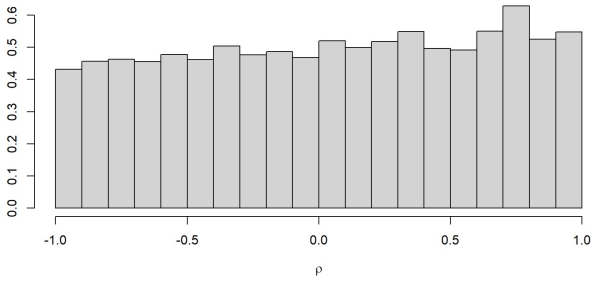

Figure 3.10: posterior distribution of ρ for Dow Jones and SSEComposite (α= 0,01)

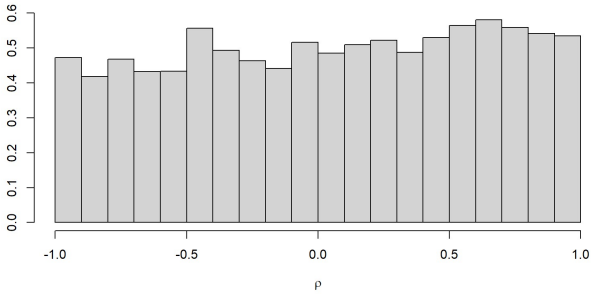

Figure 3.11: posterior distribution of ρ for Dow Jones and SZSEComponent (α= 0,01)

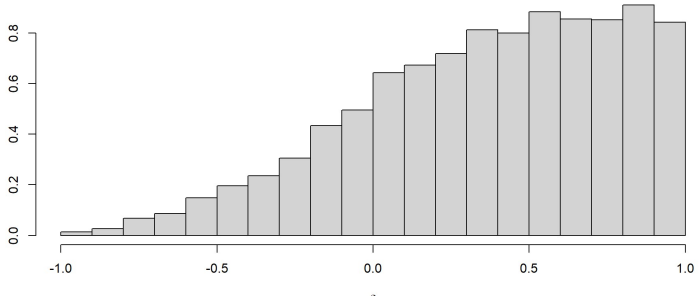

Figure 3.12: posterior distribution of ρ for Dow Jones and Euronext100 (α= 0,01)

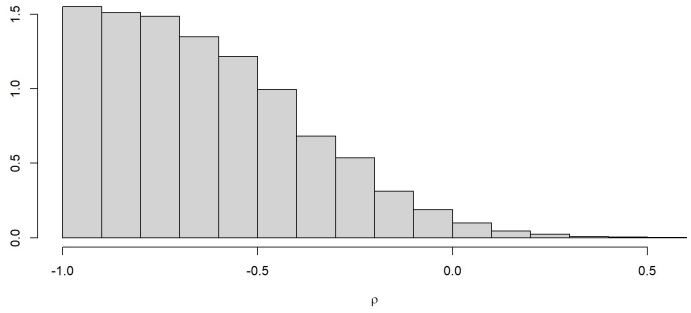

Figure 3.13: posterior distribution of ρ for Dow Jones and HANGSENG (α= 0,01)

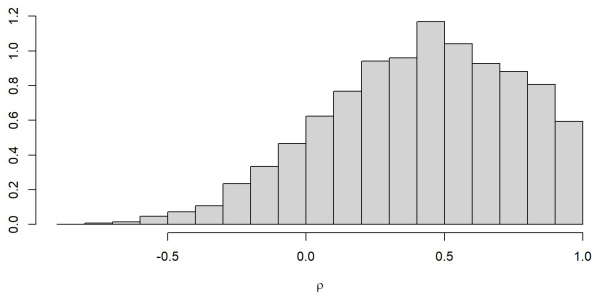

Figure 3.14: posterior distribution of ρ for S&P500 and Nasdaq100 (α= 0,01)

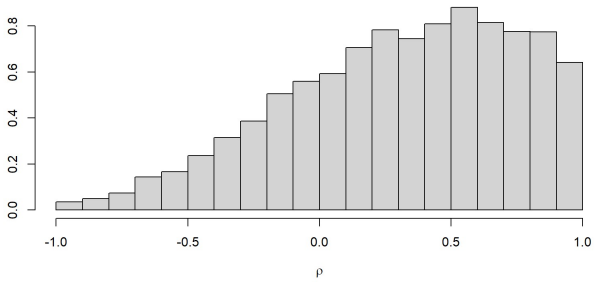

Figure 3.15: posterior distribution of ρ for S&P500 and FTSE100 (α= 0,01)

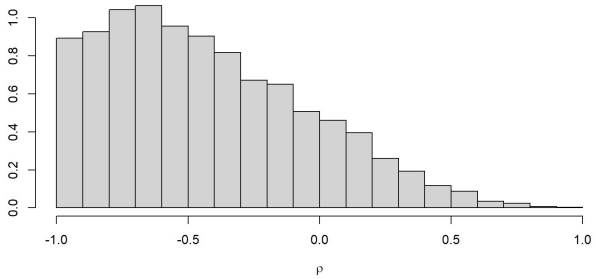

Figure 3.16: posterior distribution of ρ for S&P500 and Nikkei225 (α= 0,01)

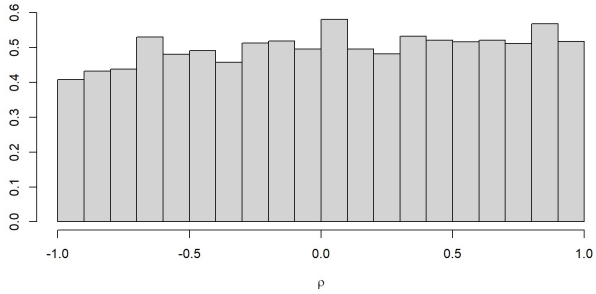

Figure 3.17: posterior distribution of ρ for S&P500 and SSEComposite (α= 0,01)

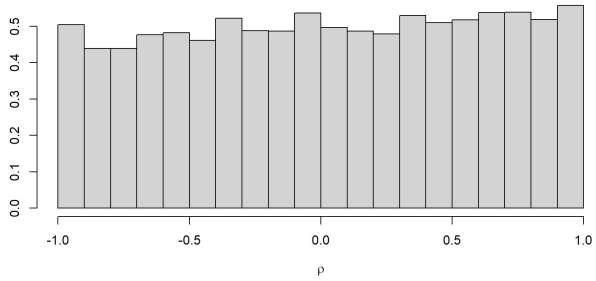

Figure 3.18: posterior distribution of ρ for S&P500 and SZSEComponent (α= 0,01)

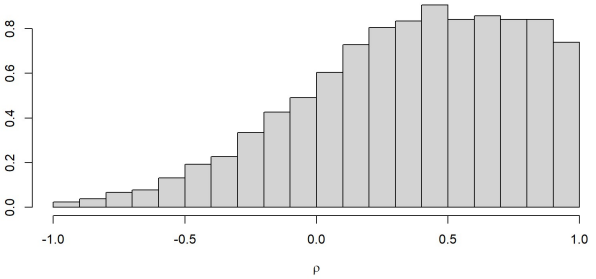

Figure 3.19: posterior distribution of ρ for S&P500 and Euronext100 (α= 0,01)

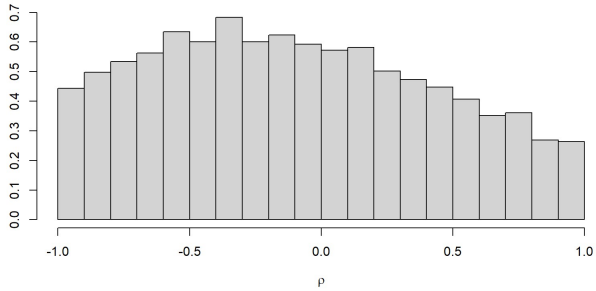

Figure 3.20: posterior distribution of ρ for S&P500 and HANGSENG (α= 0,01)

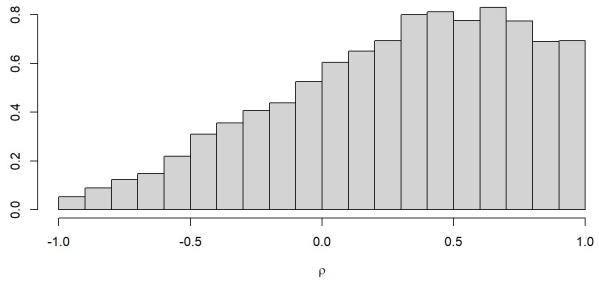

Figure 3.21: posterior distribution of ρ for Nasdaq100 and FTSE100(α= 0,01)

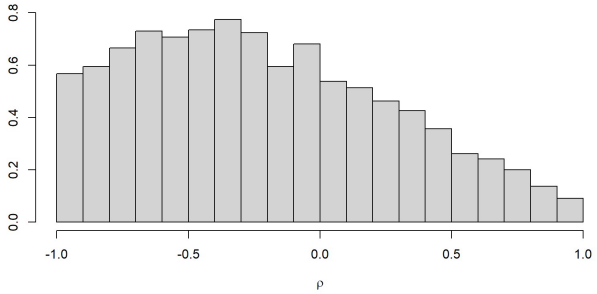

Figure 3.22: posterior distribution of ρ for Nasdaq100 and Nikkkei225 (α= 0,01)

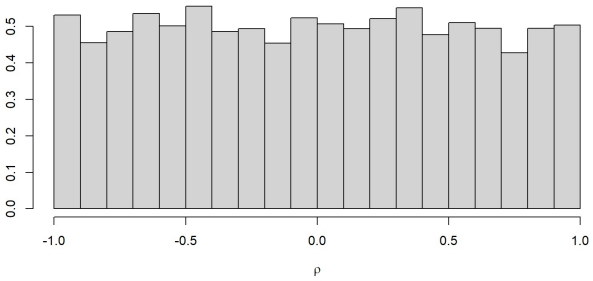

Figure 3.23: posterior distribution of ρ for Nasdaq100 and SSE Composite (α= 0,01)

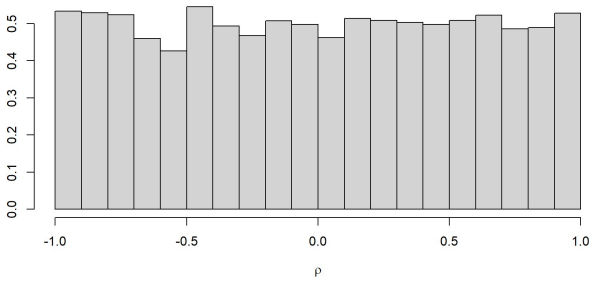

Figure 3.24: posterior distribution of ρ for Nasdaq100 and SZSEComponent (α= 0,01)

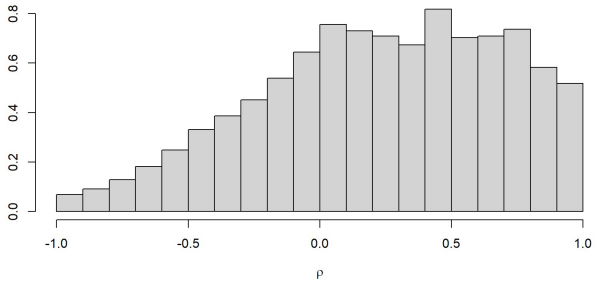

Figure 3.25: posterior distribution of ρ for Nasdaq100 and Euronext100 (α= 0,01)

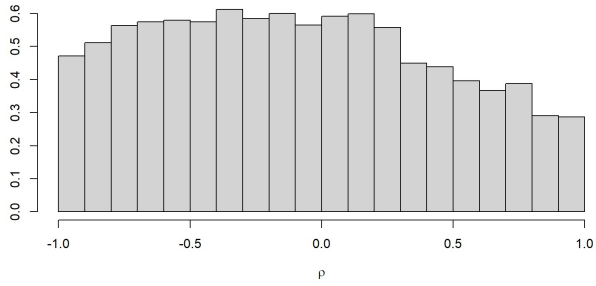

Figure 3.26: posterior distribution of ρ for Nasdaq100 and HANG SENG (α= 0,01)

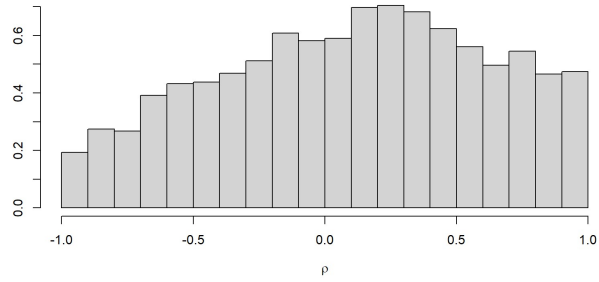

Figure 3.27: posterior distribution of ρ for FTSE100 and Nikkei225 (α= 0,01)

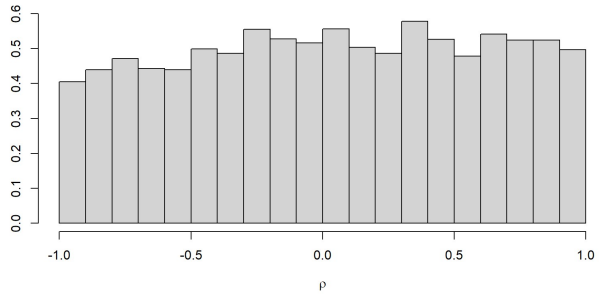

Figure 3.28: posterior distribution of ρ for FTSE100 and SSE Composite (α= 0,01)

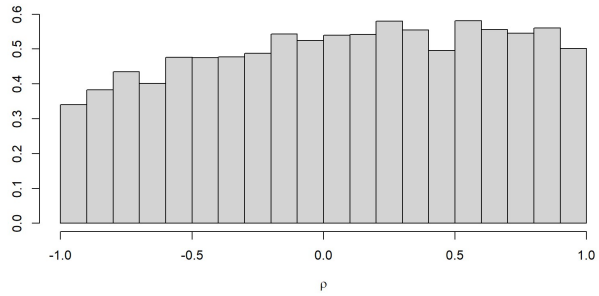

Figure 3.29: posterior distribution of ρ for FTSE100 and SZSE Component (α= 0,01)

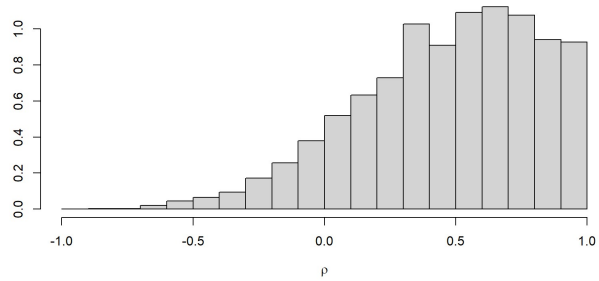

Figure 3.30: posterior distribution of ρ for FTSE100 and Euronext 100 (α= 0,01)

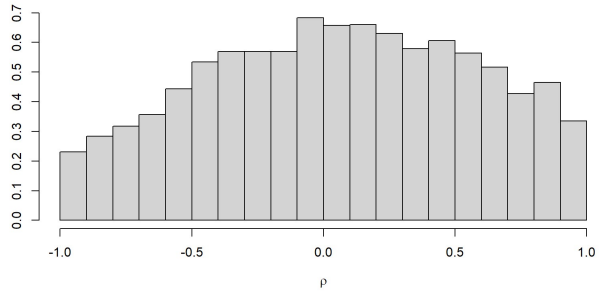

Figure 3.31: posterior distribution of ρ for FTSE100 and HANG SENG (α= 0,01)

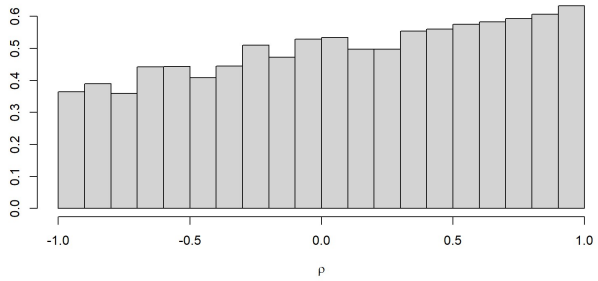

Figure 3.32: posterior distribution of ρ for Nikkei 225 and SSE Composite (α= 0,01)

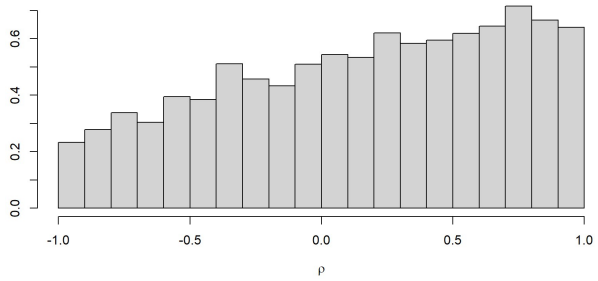

Figure 3.33: posterior distribution of ρ for Nikkei 225 and SZSE Component (α= 0,01)

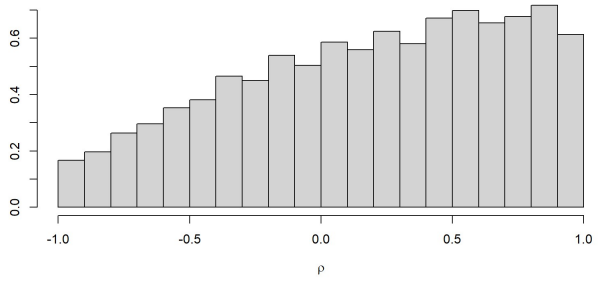

Figure 3.34: posterior distribution of ρ for Nikkei 225 and Euronext100 (α= 0,01)

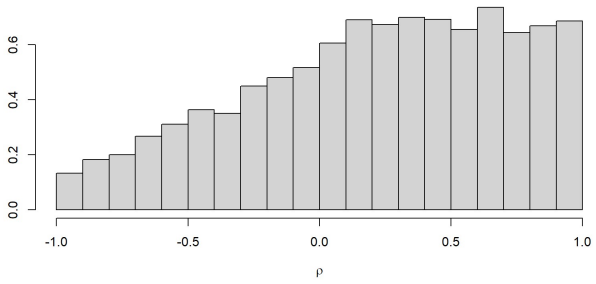

Figure 3.35: posterior distribution of ρ for Nikkei 225 and HANG SENG (α= 0,01)

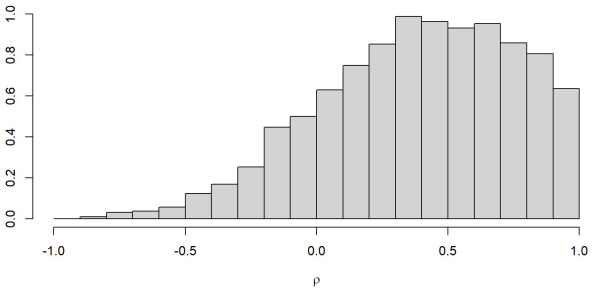

Figure 3.36: posterior distribution of ρ for SSE Composite and SZSE Component (α= 0,01)

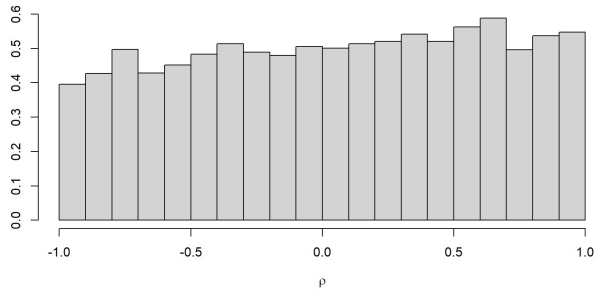

Figure 3.37: posterior distribution of ρ for SSE Composite and Euronext 100 (α= 0,01)

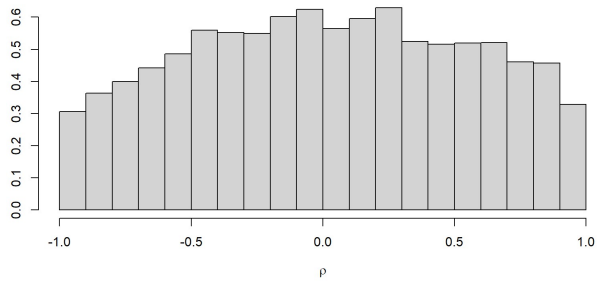

Figure 3.38: posterior distribution of ρ for SSE Composite and HANG SENG (α= 0,01)

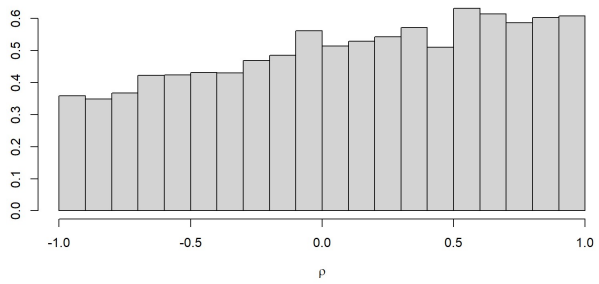

Figure 3.39: posterior distribution of ρ for SZSE Component and Euronext 100 (α= 0,01)

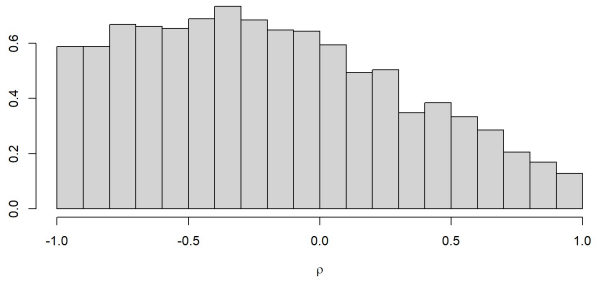

Figure 3.40: posterior distribution of ρ for SZSE Component and HANG SENG (α= 0,01)

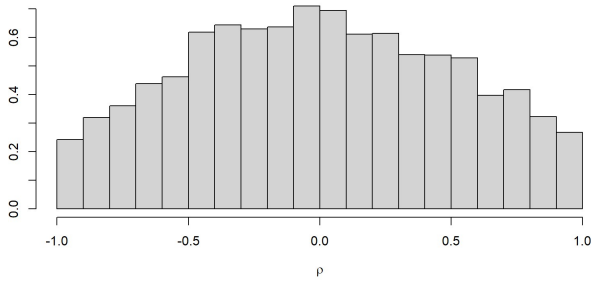

Figure 3.41: posterior distribution of ρ for Euronext100 and HANG SENG (α= 0,01)

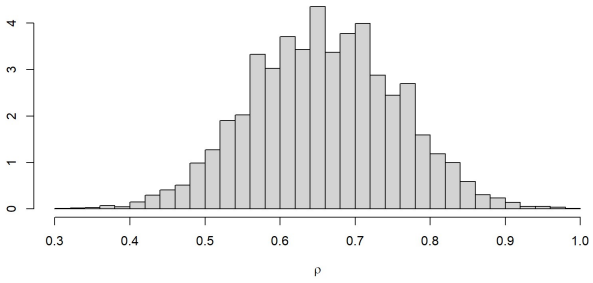

Figure 3.42: posterior distribution of ρ for Dow Jones and Nasdaq100 (α= 0,25)

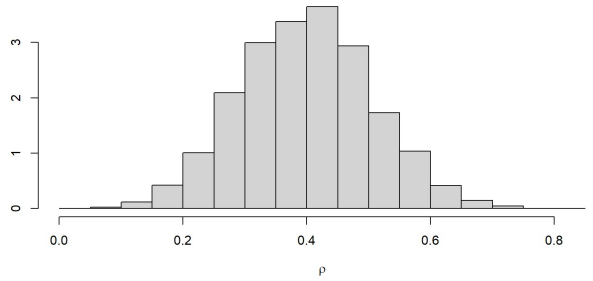

Figure 3.43: posterior distribution of ρ for Dow Jones and FTSE100 (α= 0,25)

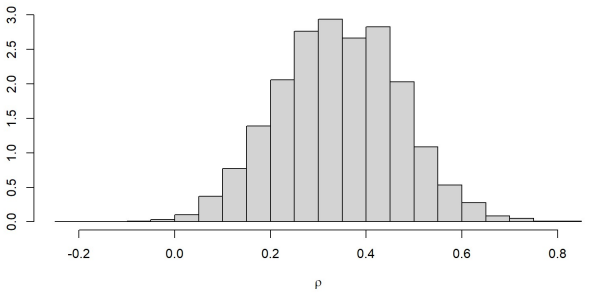

Figure 3.44: posterior distribution of ρ for Dow Jones and Nikkei225 (α= 0,25)

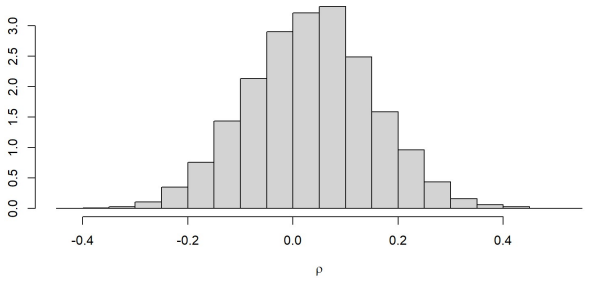

Figure 3.45: posterior distribution of ρ for Dow Jones and SSEComposite (α= 0,25)

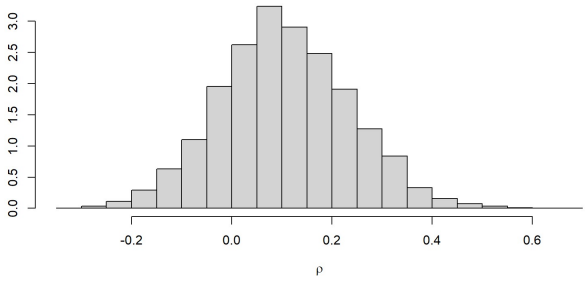

Figure 3.46: posterior distribution of ρ for Dow Jones and SZSEComponent (α= 0,25)

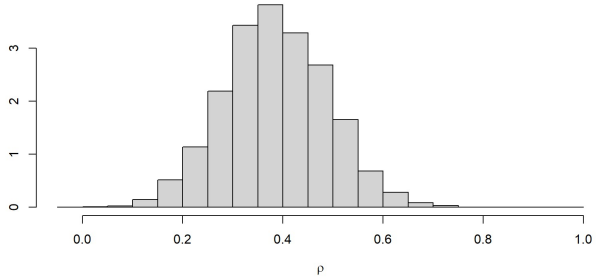

Figure 3.47: posterior distribution of ρ for Dow Jones and Euronext100 (α= 0,25)

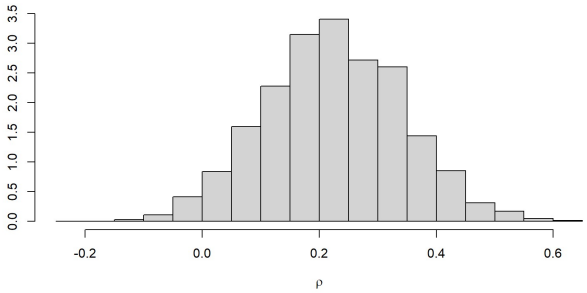

Figure 3.48: posterior distribution of ρ for Dow Jones and HANGSENG (α= 0,25)

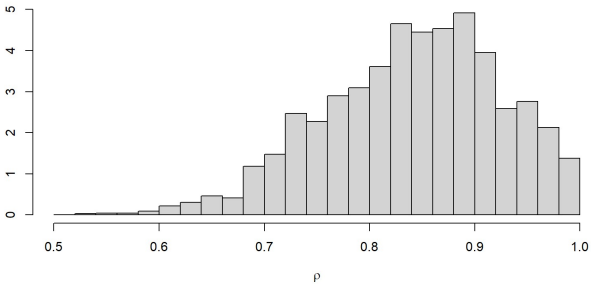

Figure 3.49: posterior distribution of ρ for Dow Jones and S&P500 (α= 0,25)

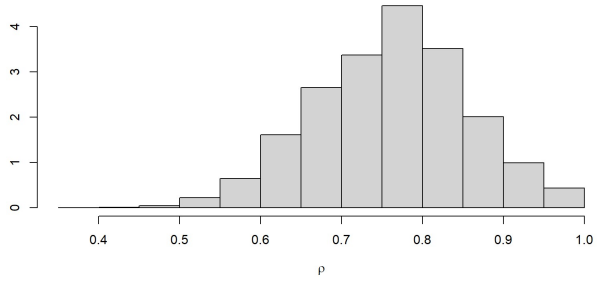

Figure 3.50: posterior distribution of ρ for S&P500 and Nasdaq100 (α= 0,25)

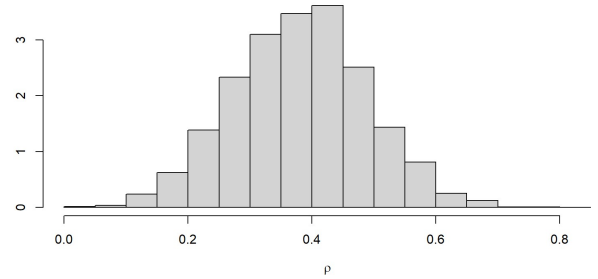

Figure 3.51: posterior distribution of ρ for S&P500 and FTSE100 (α= 0,25)

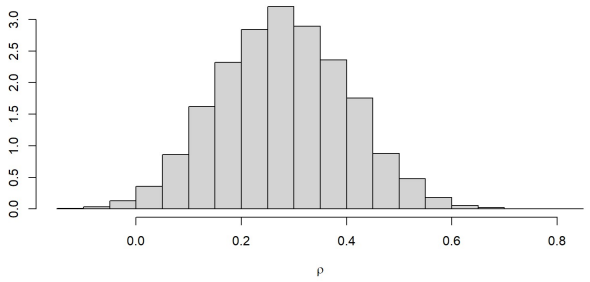

Figure 3.52: posterior distribution of ρ for S&P500 and Nikkei225 (α= 0,25)

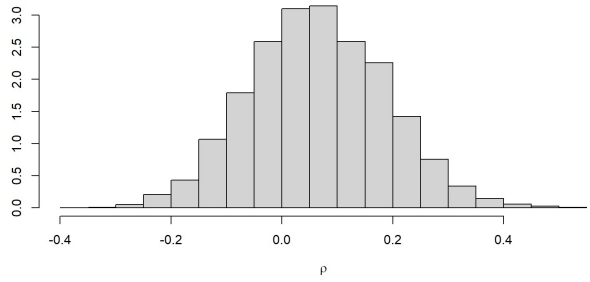

Figure 3.53: posterior distribution of ρ for S&P500 and SSEComposite (α= 0,25)

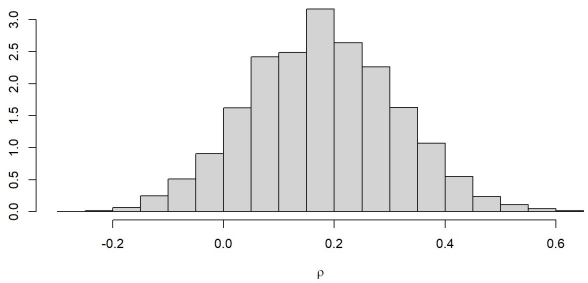

Figure 3.54: posterior distribution of ρ for S&P500 and SZSEComponent (α= 0,25)

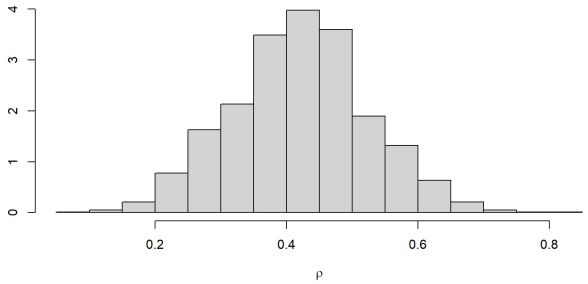

Figure 3.55: posterior distribution of ρ for S&P500 and Euronext100 (α= 0,25)

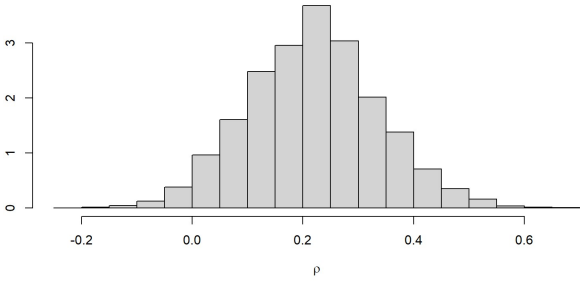

Figure 3.56: posterior distribution of ρ for S&P500 and HANGSENG (α= 0,25)

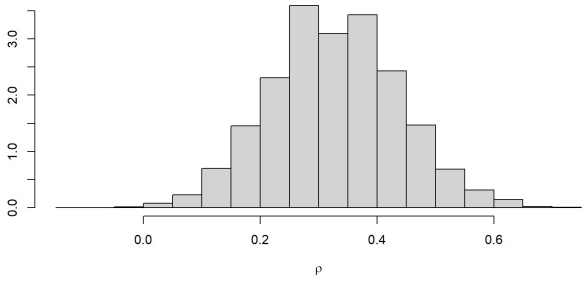

Figure 3.57: posterior distribution of ρ for Nasdaq100 and FTSE100(α= 0,25)

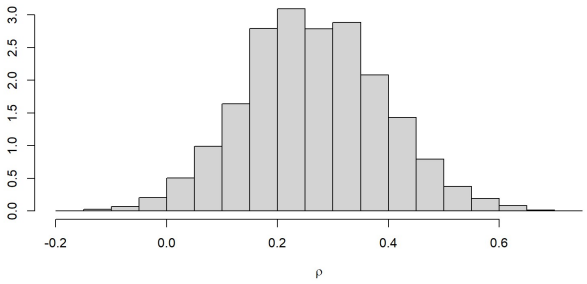

Figure 3.58: posterior distribution of ρ for Nasdaq100 and Nikkkei225 (α= 0,25)

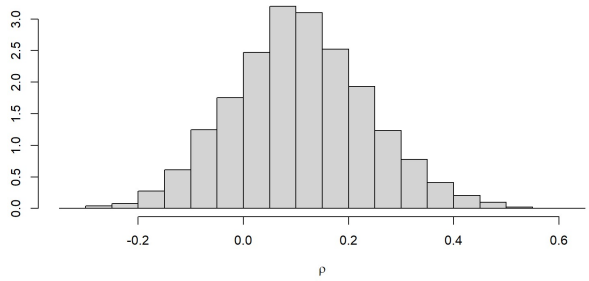

Figure 3.59: posterior distribution of ρ for Nasdaq100 and SSE Composite (α= 0,25)

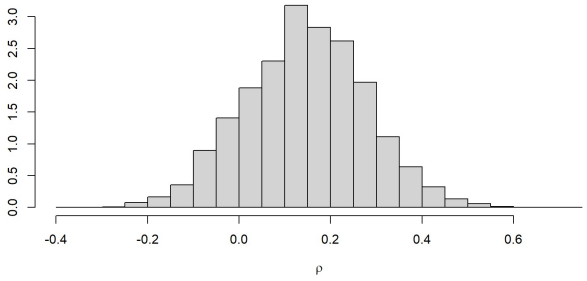

Figure 3.60: posterior distribution of ρ for Nasdaq100 and SZSEComponent (α= 0,25)

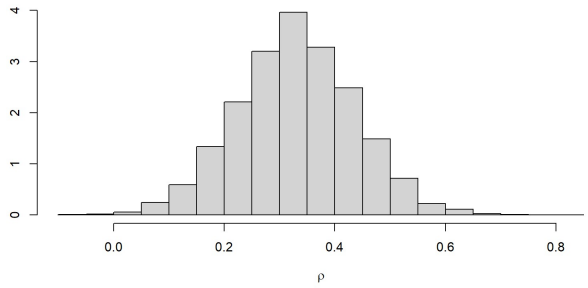

Figure 3.61: posterior distribution of ρ for Nasdaq100 and Euronext100 (α= 0,25)

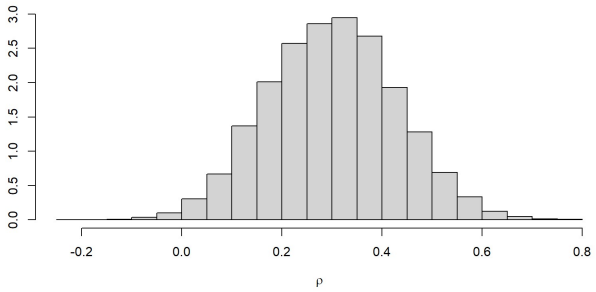

Figure 3.62: posterior distribution of ρ for Nasdaq100 and HANG SENG (α= 0,25)

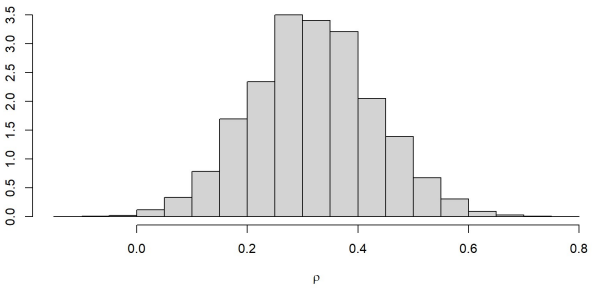

Figure 3.63: posterior distribution of ρ for FTSE100 and Nikkei225 (α= 0,25)

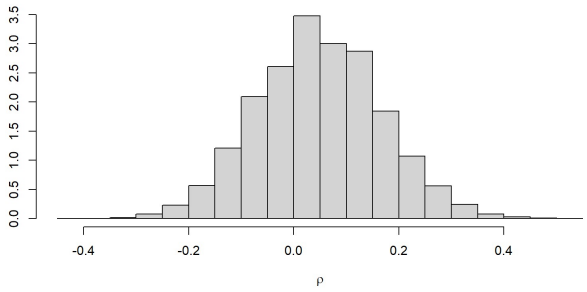

Figure 3.64: posterior distribution of ρ for FTSE100 and SSE Composite (α= 0,25)

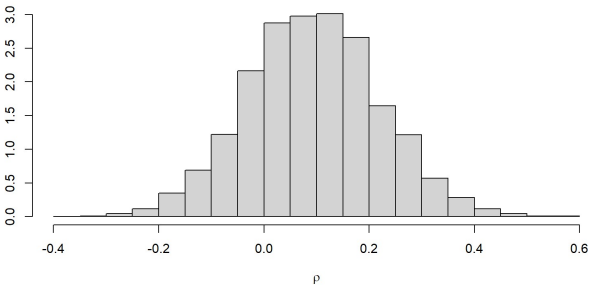

Figure 3.65: posterior distribution of ρ for FTSE100 and SZSE Component (α= 0,25)

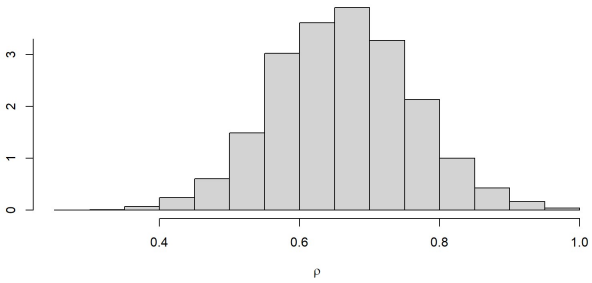

Figure 3.66: posterior distribution of ρ for FTSE100 and Euronext 100 (α= 0,25)

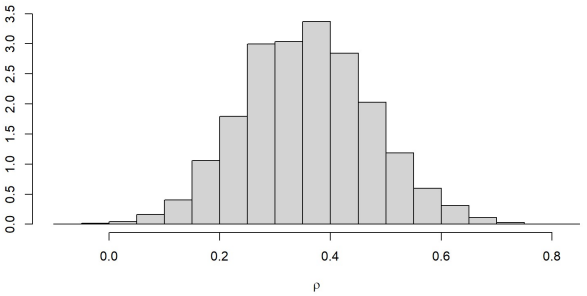

Figure 3.67: posterior distribution of ρ for FTSE100 and HANG SENG (α= 0,25)

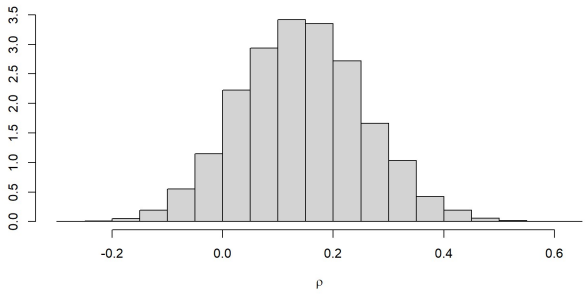

Figure 3.68: posterior distribution of ρ for Nikkei 225 and SSE Composite (α= 0,25)

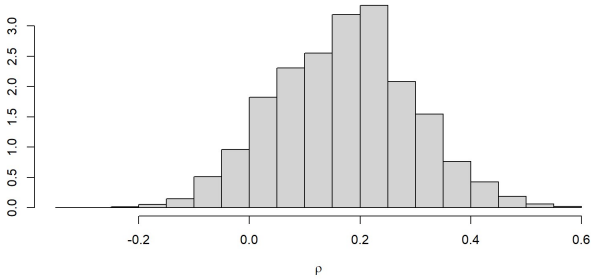

Figure 3.69: posterior distribution of ρ for Nikkei 225 and SZSE Component (α= 0,25)

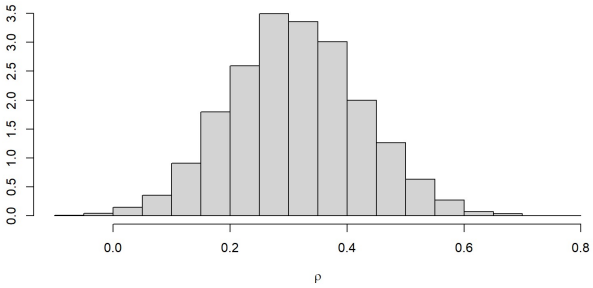

Figure 3.70: posterior distribution of ρ for Nikkei 225 and Euronext100 (α= 0,25)

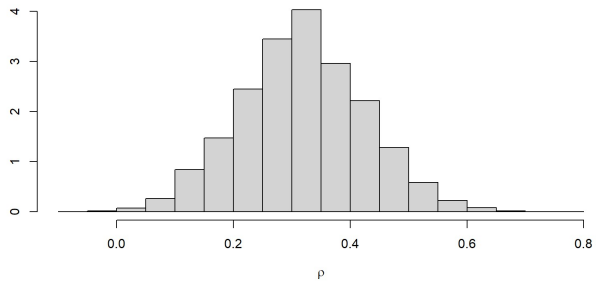

Figure 3.71: posterior distribution of ρ for Nikkei 225 and HANG SENG (α= 0,25)

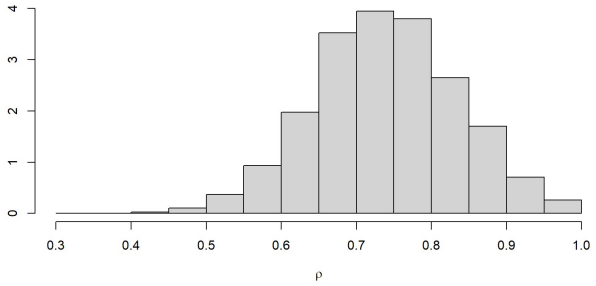

Figure 3.72: posterior distribution of ρ for SSE Composite and SZSE Component (α= 0,25)

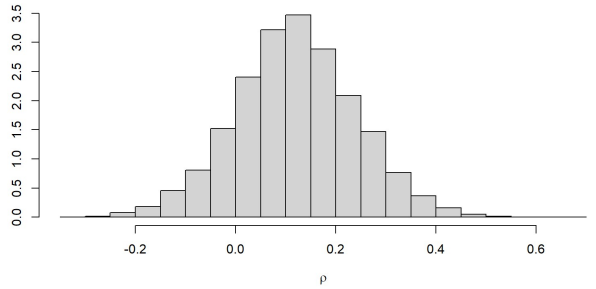

Figure 3.73: posterior distribution of ρ for SSE Composite and Euronext 100 (α= 0,25)

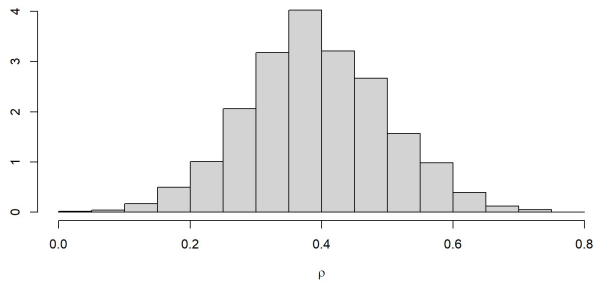

Figure 3.74: posterior distribution of ρ for SSE Composite and HANG SENG (α= 0,25)

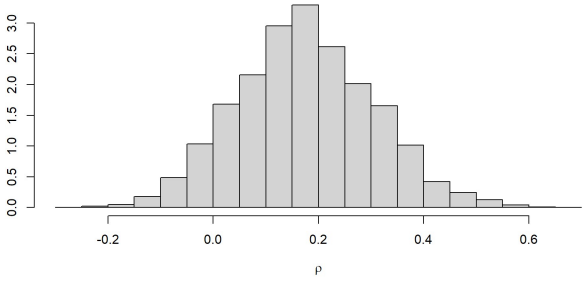

Figure 3.75: posterior distribution of ρ for SZSE Component and Euronext 100 (α= 0,25)

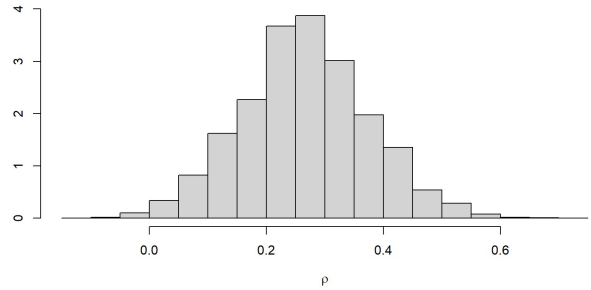

Figure 3.76: posterior distribution of ρ for SZSE Component and HANG SENG (α= 0,25)

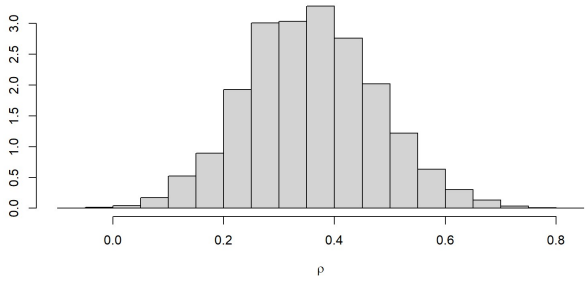

Figure 3.77: posterior distribution of ρ for Euronext100 and HANG SENG (α= 0,25)

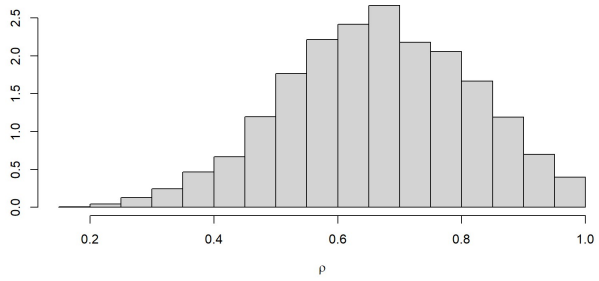

Figure 3.78: posterior distribution of ρ for Dow Jones and Nasdaq100 (α= 0,1)

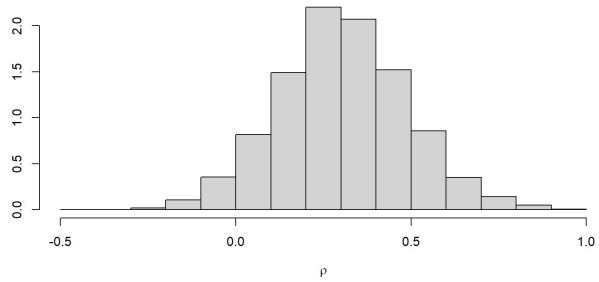

Figure 3.79: posterior distribution of ρ for Dow Jones and FTSE100 (α= 0,1)

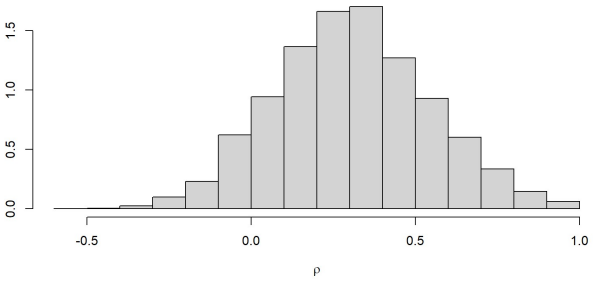

Figure 3.80: posterior distribution of ρ for Dow Jones and Nikkei225 (α= 0,1)

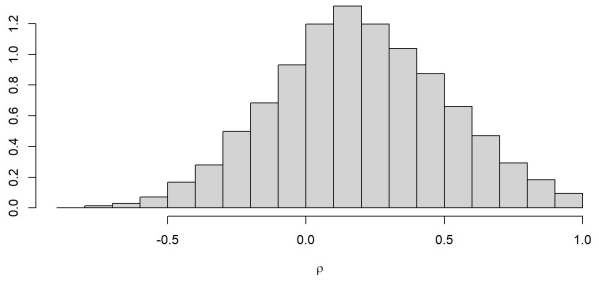

Figure 3.81: posterior distribution of ρ for Dow Jones and SSEComposite (α= 0,1)

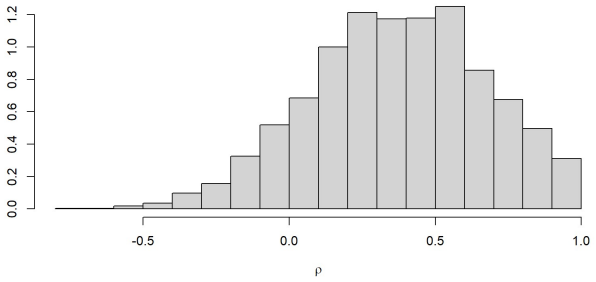

Figure 3.82: posterior distribution of ρ for Dow Jones and SZSEComponent (α= 0,1)

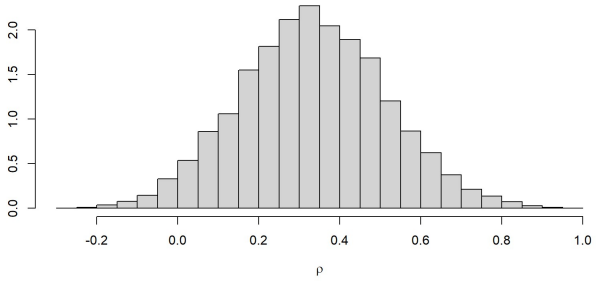

Figure 3.83: posterior distribution of ρ for Dow Jones and Euronext100 (α= 0,1)

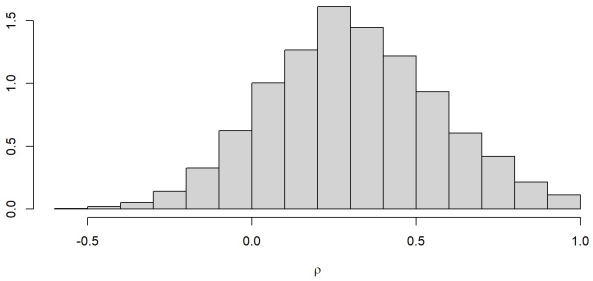

Figure 3.84: posterior distribution of ρ for Dow Jones and HANGSENG (α= 0,1)

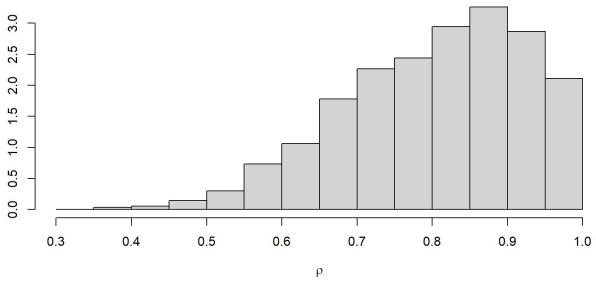

Figure 3.85: posterior distribution of ρ for Dow Jones and S&P500 (α= 0,1)

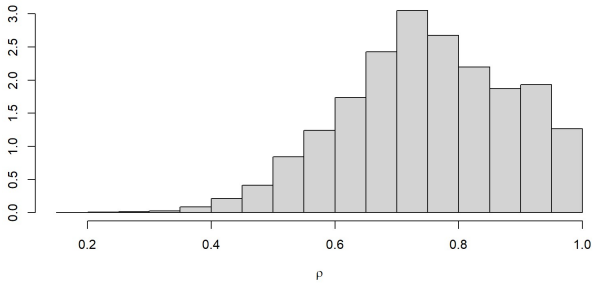

Figure 3.86: posterior distribution of ρ for S&P500 and Nasdaq100 (α= 0,1)

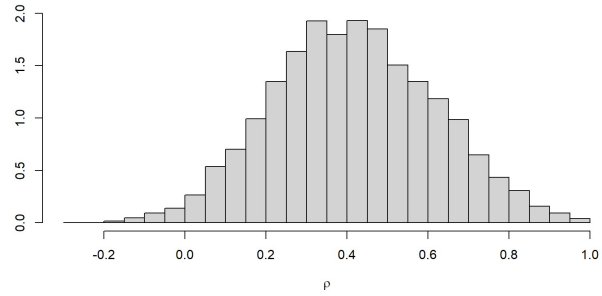

Figure 3.87: posterior distribution of ρ for S&P500 and FTSE100 (α= 0,1)

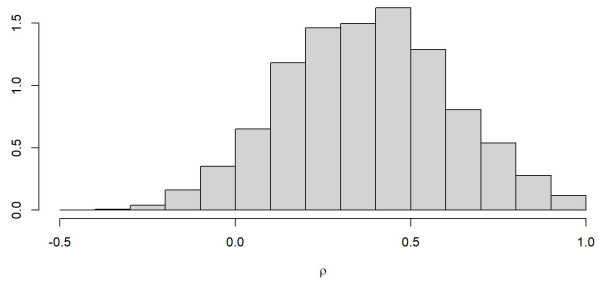

Figure 3.88: posterior distribution of ρ for S&P500 and Nikkei225 (α= 0,1)

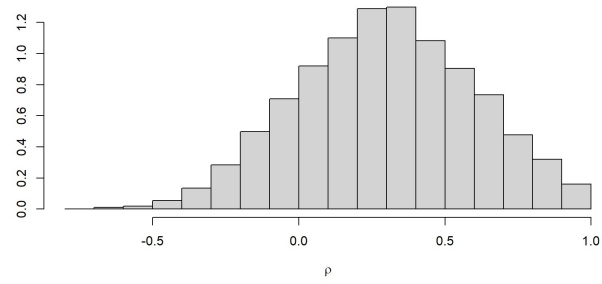

Figure 3.89: posterior distribution of ρ for S&P500 and SSEComposite (α= 0,1)

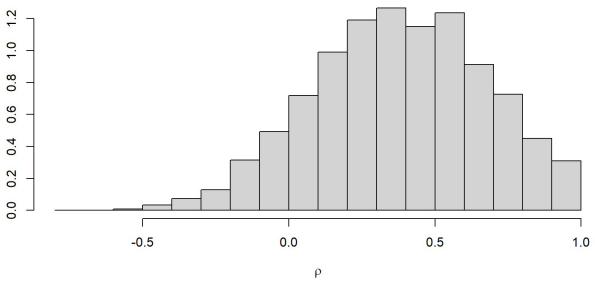

Figure 3.90: posterior distribution of ρ for S&P500 and SZSEComponent (α= 0,1)

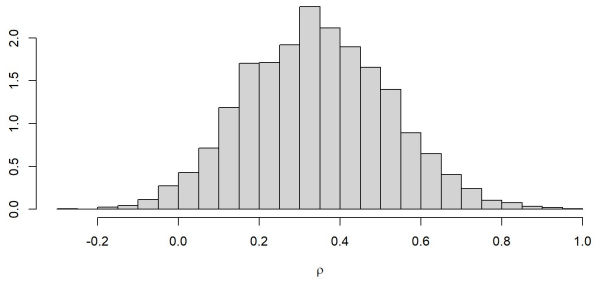

Figure 3.91: posterior distribution of ρ for S&P500 and Euronext100 (α= 0,1)

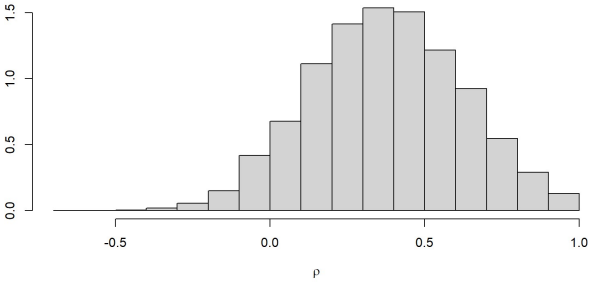

Figure 3.92: posterior distribution of ρ for S&P500 and HANGSENG (α= 0,1)

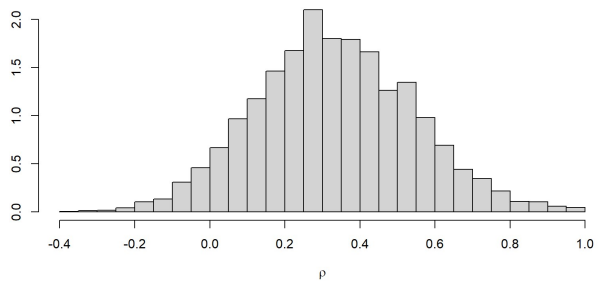

Figure 3.93: posterior distribution of ρ for Nasdaq100 and FTSE100(α= 0,1)

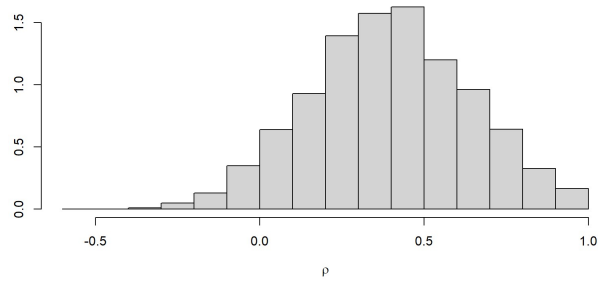

Figure 3.94: posterior distribution of ρ for Nasdaq100 and Nikkkei225 (α= 0,1)

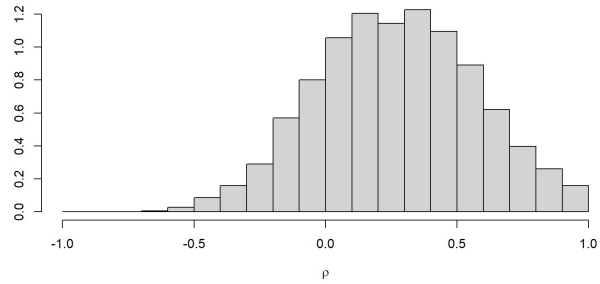

Figure 3.95: posterior distribution of ρ for Nasdaq100 and SSE Composite (α= 0,1)

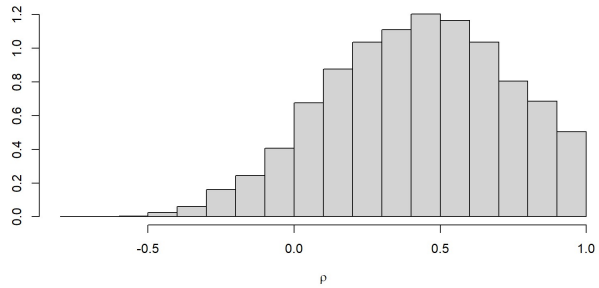

Figure 3.96: posterior distribution of ρ for Nasdaq100 and SZSEComponent (α= 0,1)

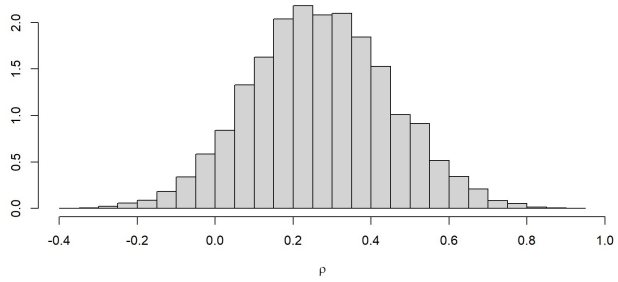

Figure 3.97: posterior distribution of ρ for Nasdaq100 and Euronext100 (α= 0,1)

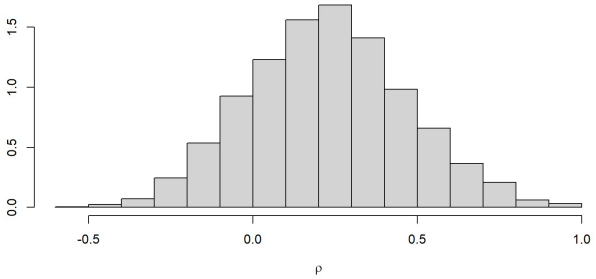

Figure 3.98: posterior distribution of ρ for Nasdaq100 and HANG SENG (α= 0,1)

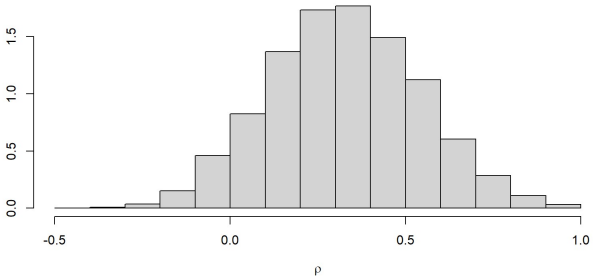

Figure 3.99: posterior distribution of ρ for FTSE100 and Nikkei225 (α= 0,1)

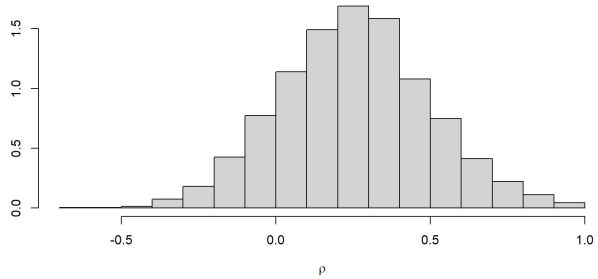

Figure 3.100: posterior distribution of ρ for FTSE100 and SSE Composite (α= 0,1)

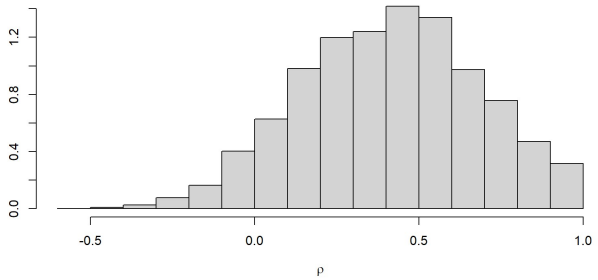

Figure 3.101: posterior distribution of ρ for FTSE100 and SZSE Component (α= 0,1)

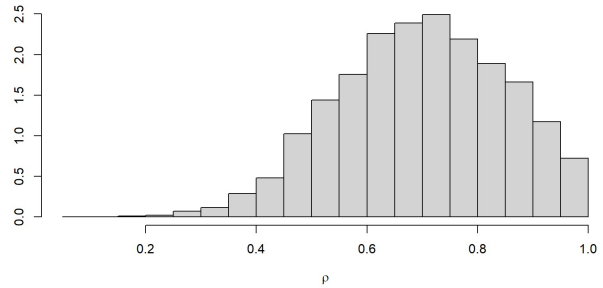

Figure 3.102: posterior distribution of ρ for FTSE100 and Euronext 100 (α= 0,1)

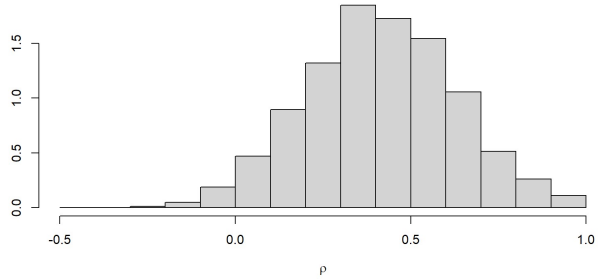

Figure 3.103: posterior distribution of ρ for FTSE100 and HANG SENG (α= 0,1)

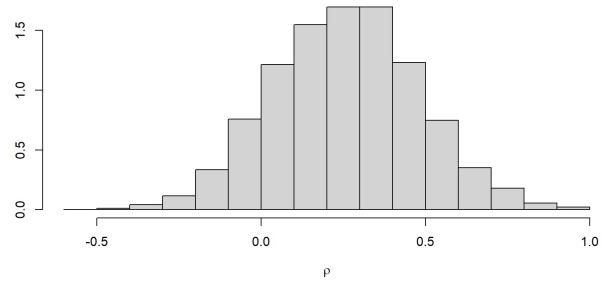

Figure 3.104: posterior distribution of ρ for Nikkei 225 and SSE Composite (α= 0,1)

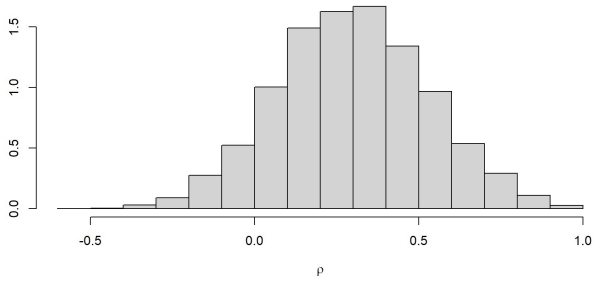

Figure 3.105: posterior distribution of ρ for Nikkei 225 and SZSE Component (α= 0,1)

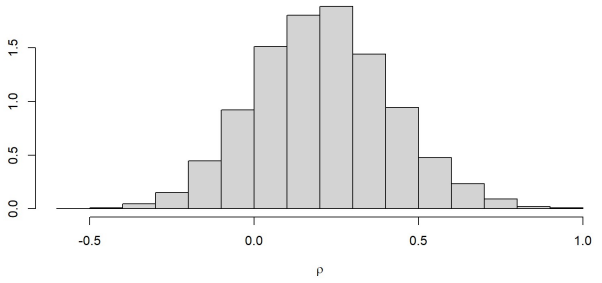

Figure 3.106: posterior distribution of ρ for Nikkei 225 and Euronext100 (α= 0,1)

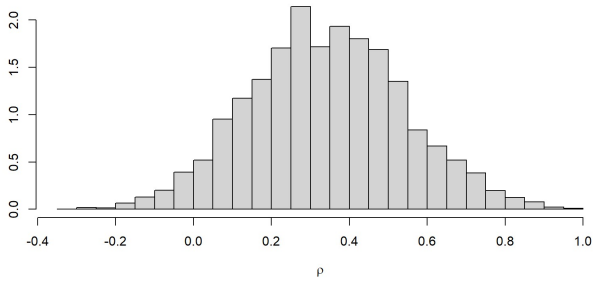

Figure 3.107: posterior distribution of ρ for Nikkei 225 and HANG SENG (α= 0,1)

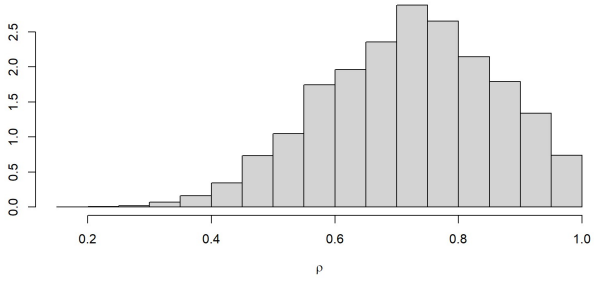

Figure 3.108: posterior distribution of ρ for SSE Composite and SZSE Component (α= 0,1)

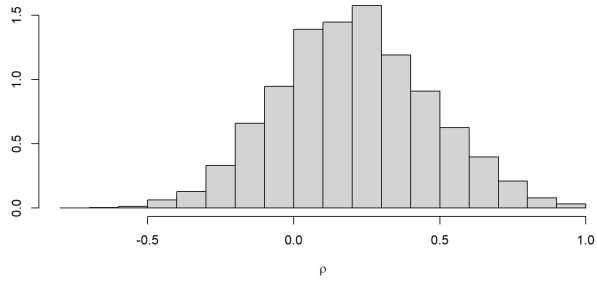

Figure 3.109: posterior distribution of ρ for SSE Composite and Euronext 100 (α= 0,1)

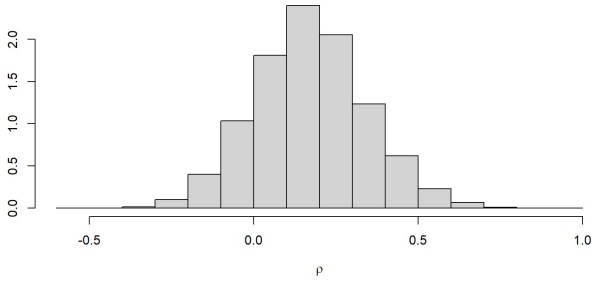

Figure 3.110: posterior distribution of ρ for SSE Composite and HANG SENG (α= 0,1)

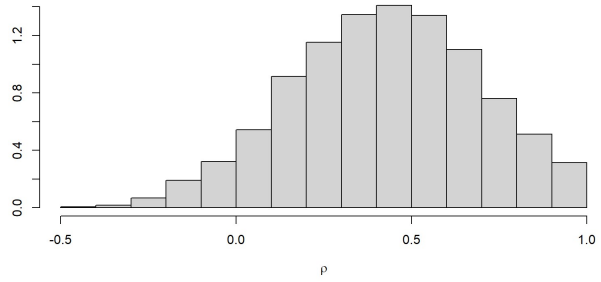

Figure 3.111: posterior distribution of ρ for SZSE Component and Euronext 100 (α= 0,1)

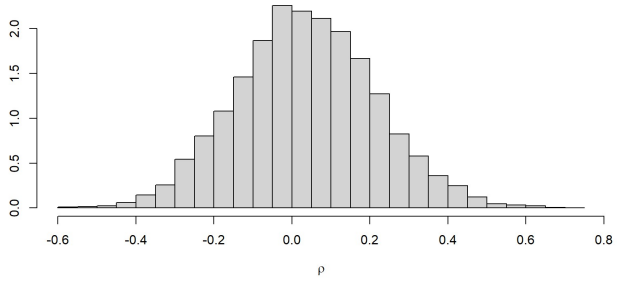

Figure 3.112: posterior distribution of ρ for SZSE Component and HANG SENG (α= 0,1)

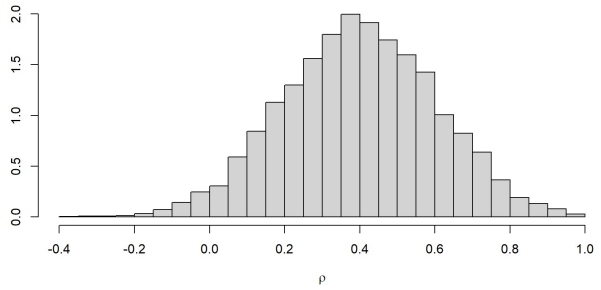

Figure 3.113: posterior distribution of ρ for Euronext100 and HANG SENG (α= 0,1)

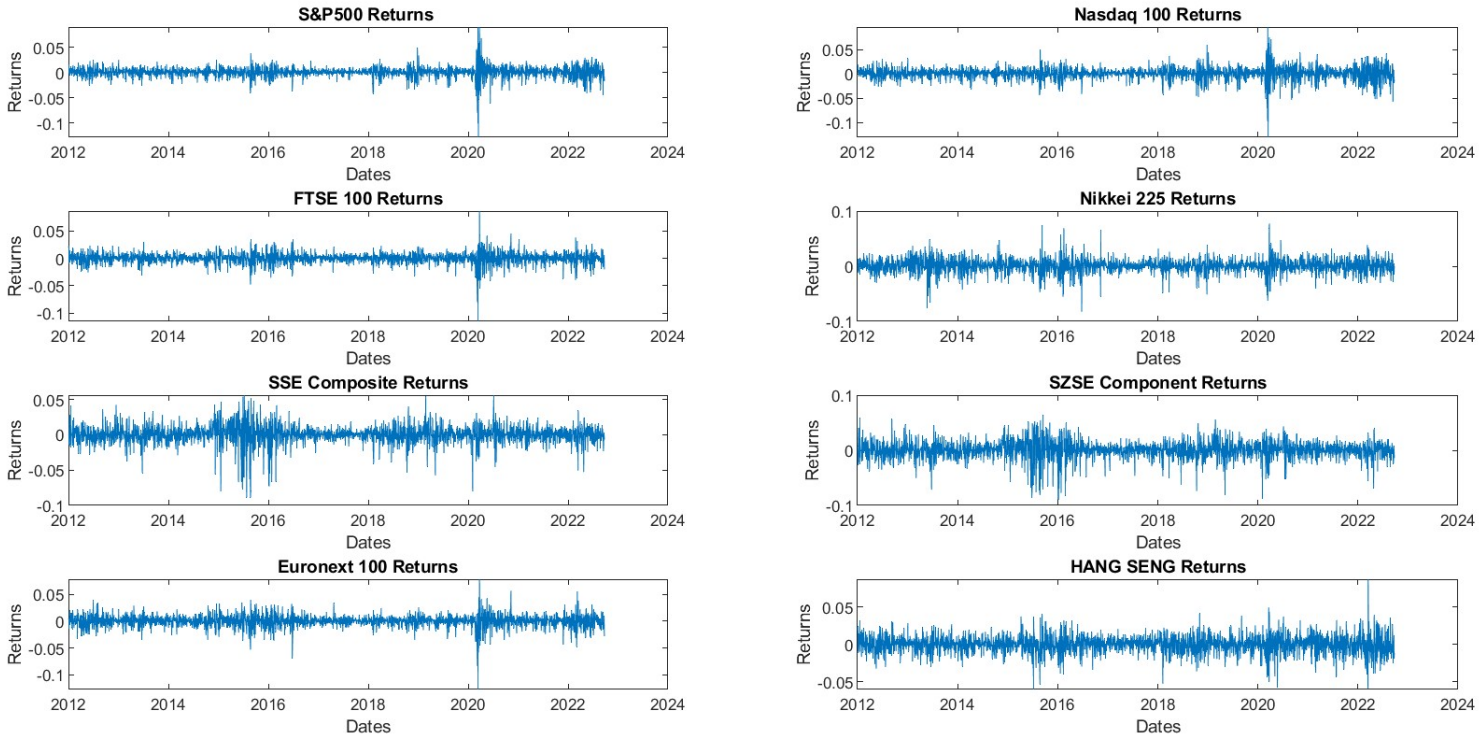

Figure 3.114: Indices Returns

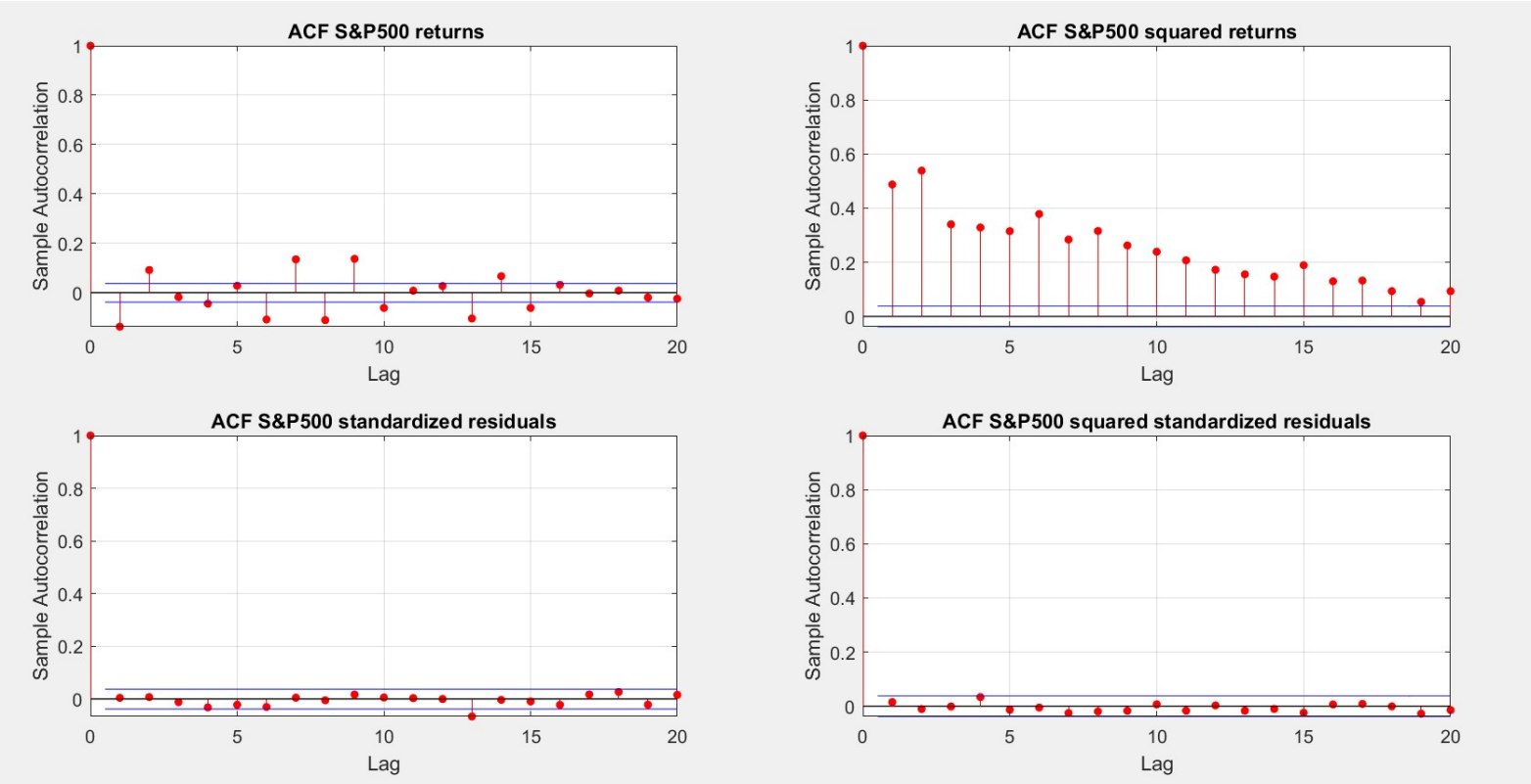

Figure 3.115: Autocorrelation Function

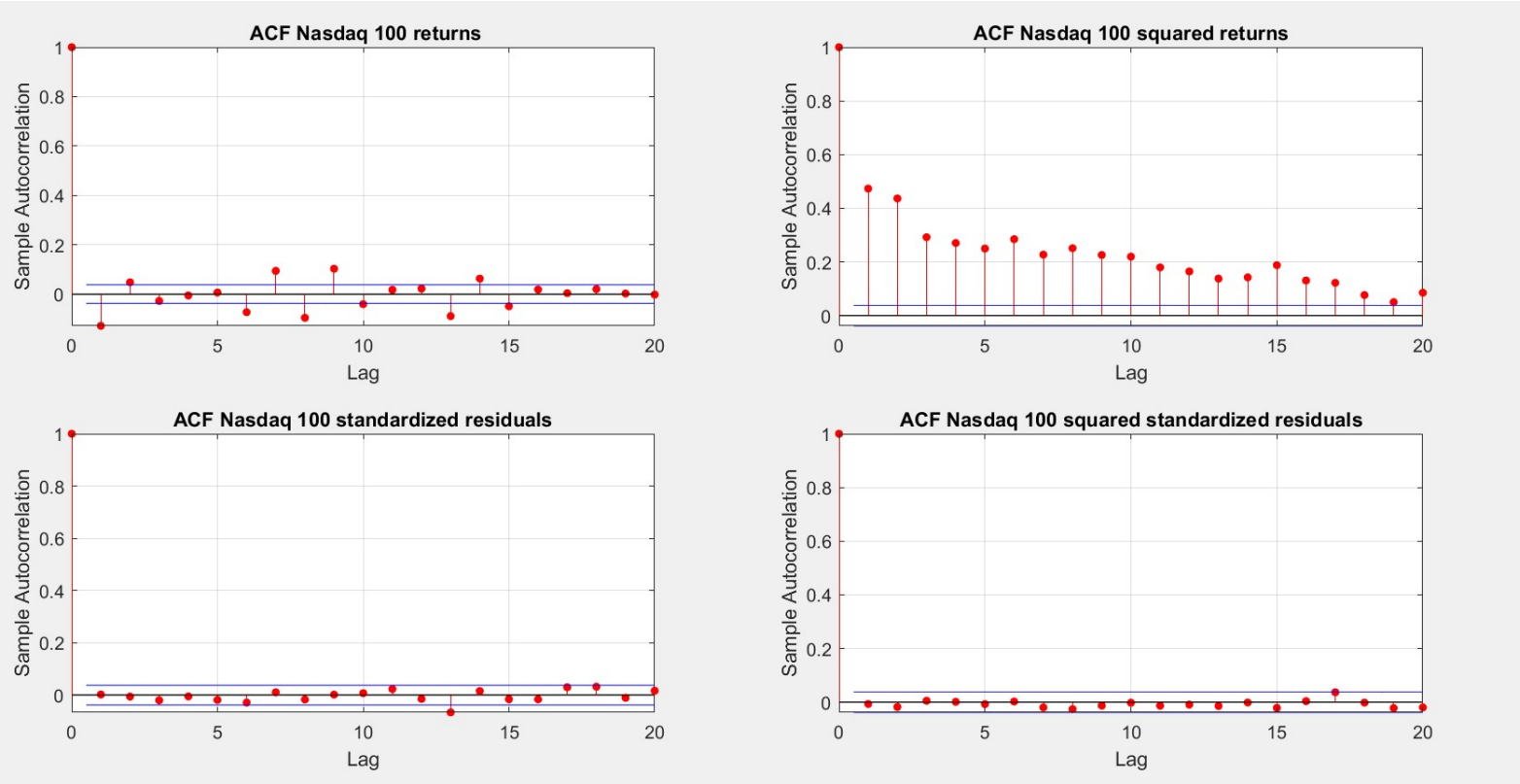

Figure 3.116: Autocorrelation Function

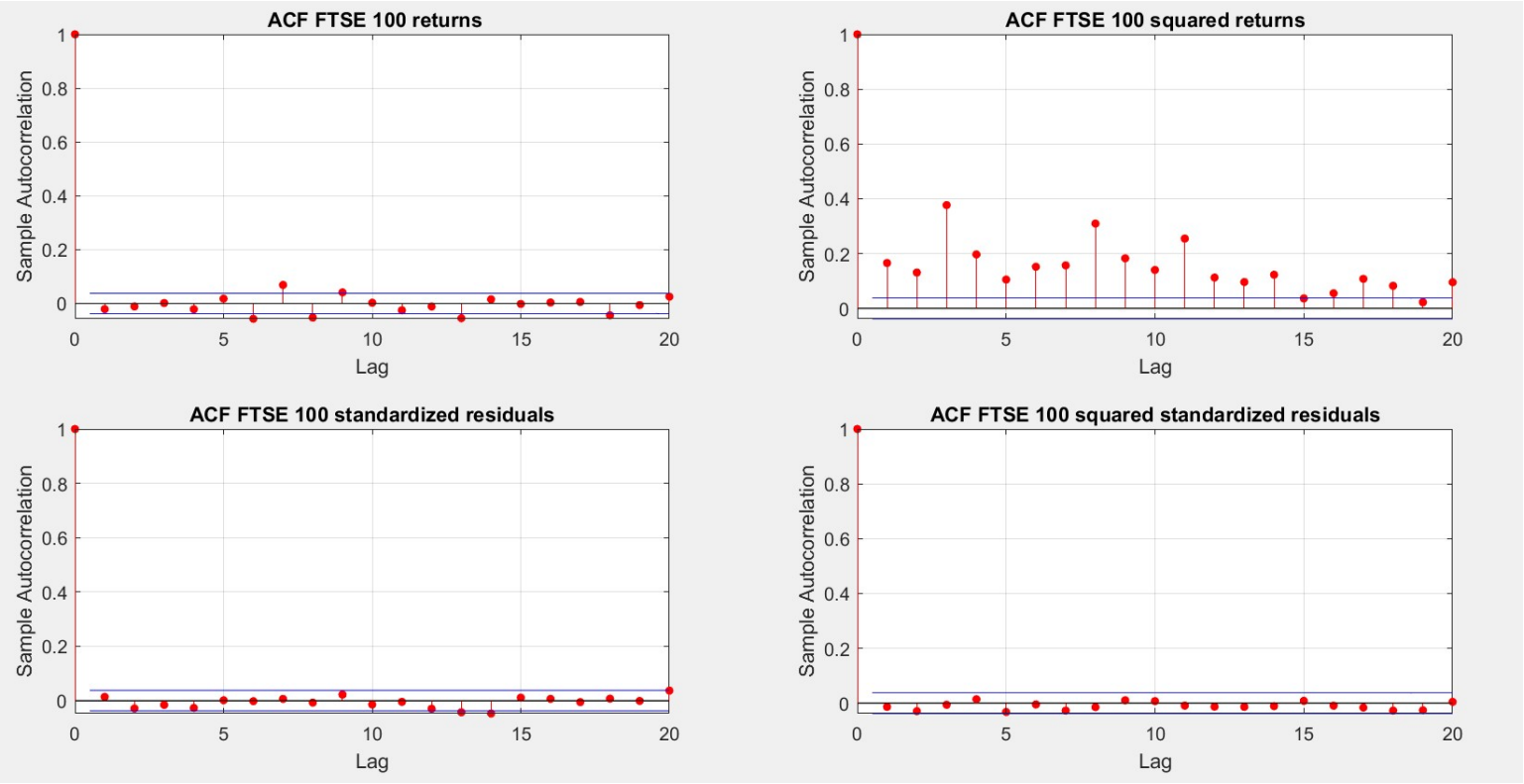

Figure 3.117: Autocorrelation Function

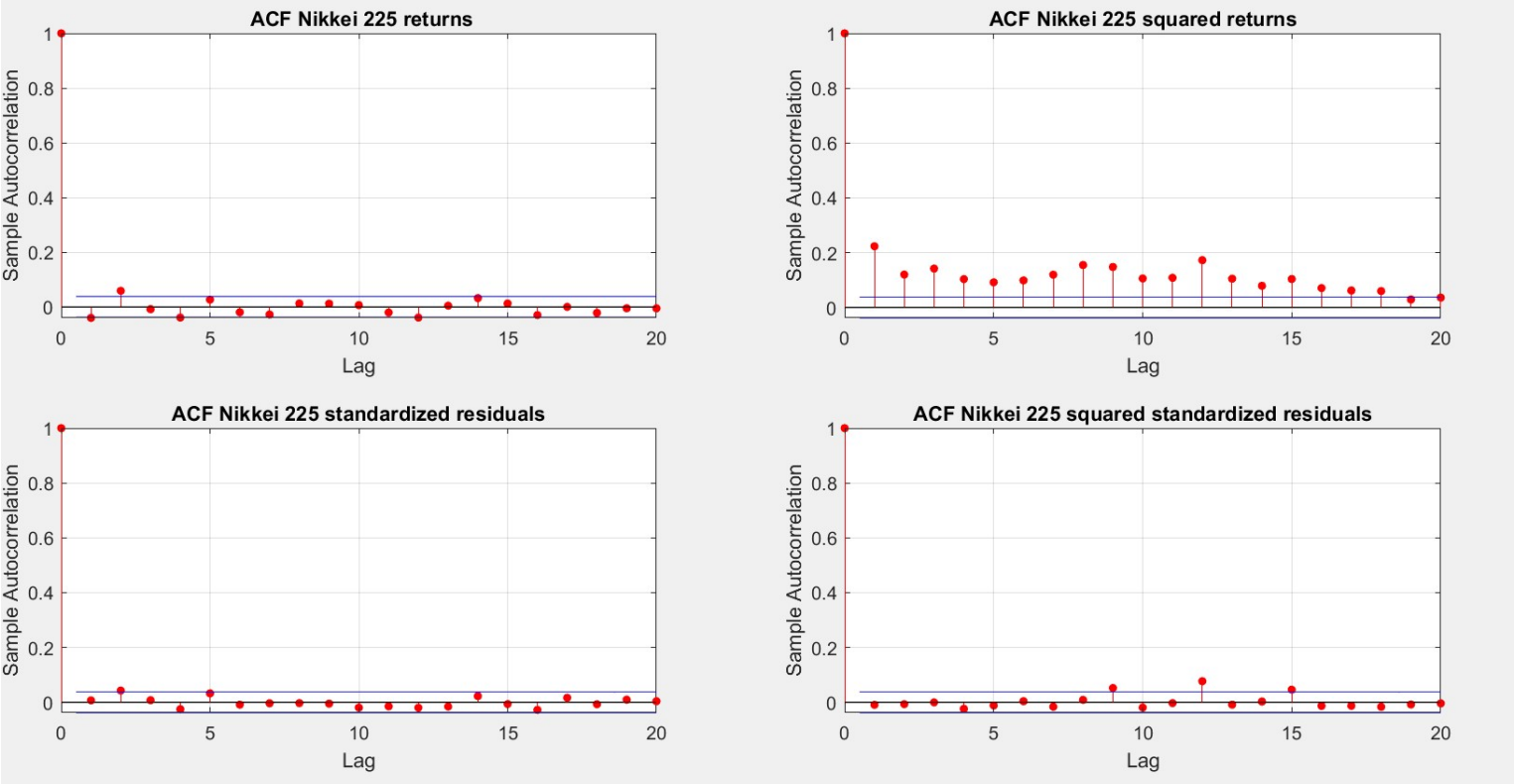

Figure 3.118: Autocorrelation Function

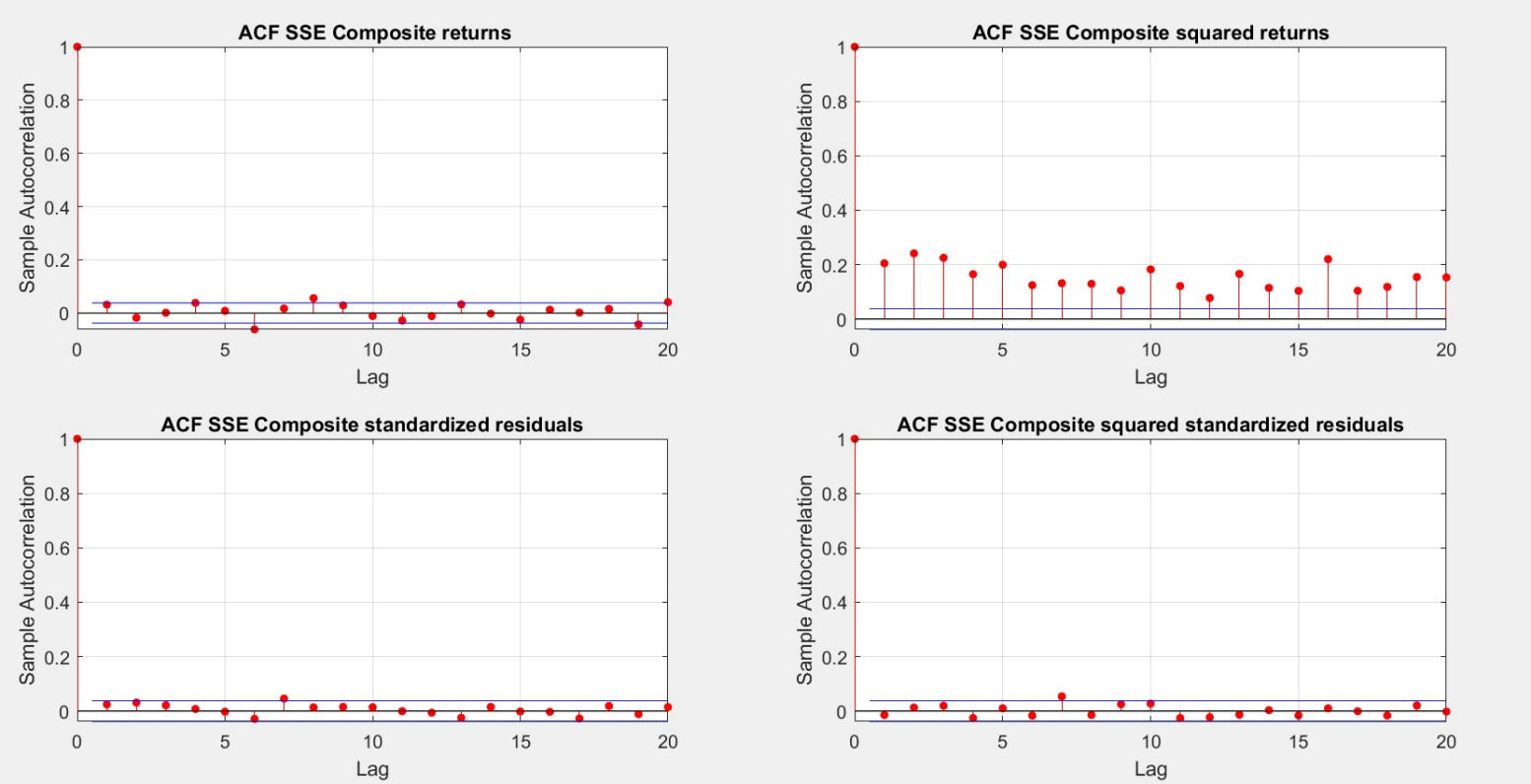

Figure 3.119: Autocorrelation Function

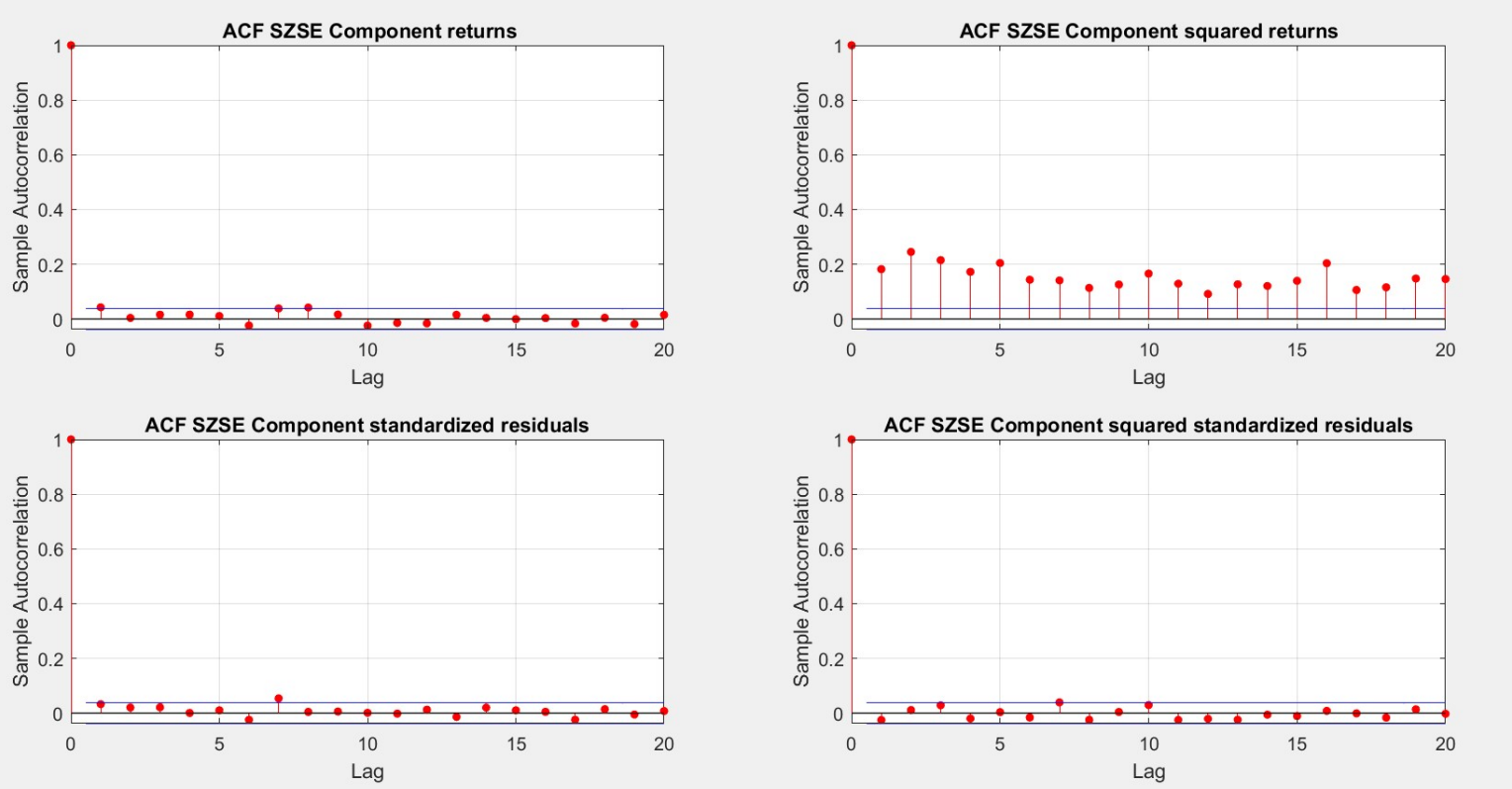

Figure 3.120: Autocorrelation Function

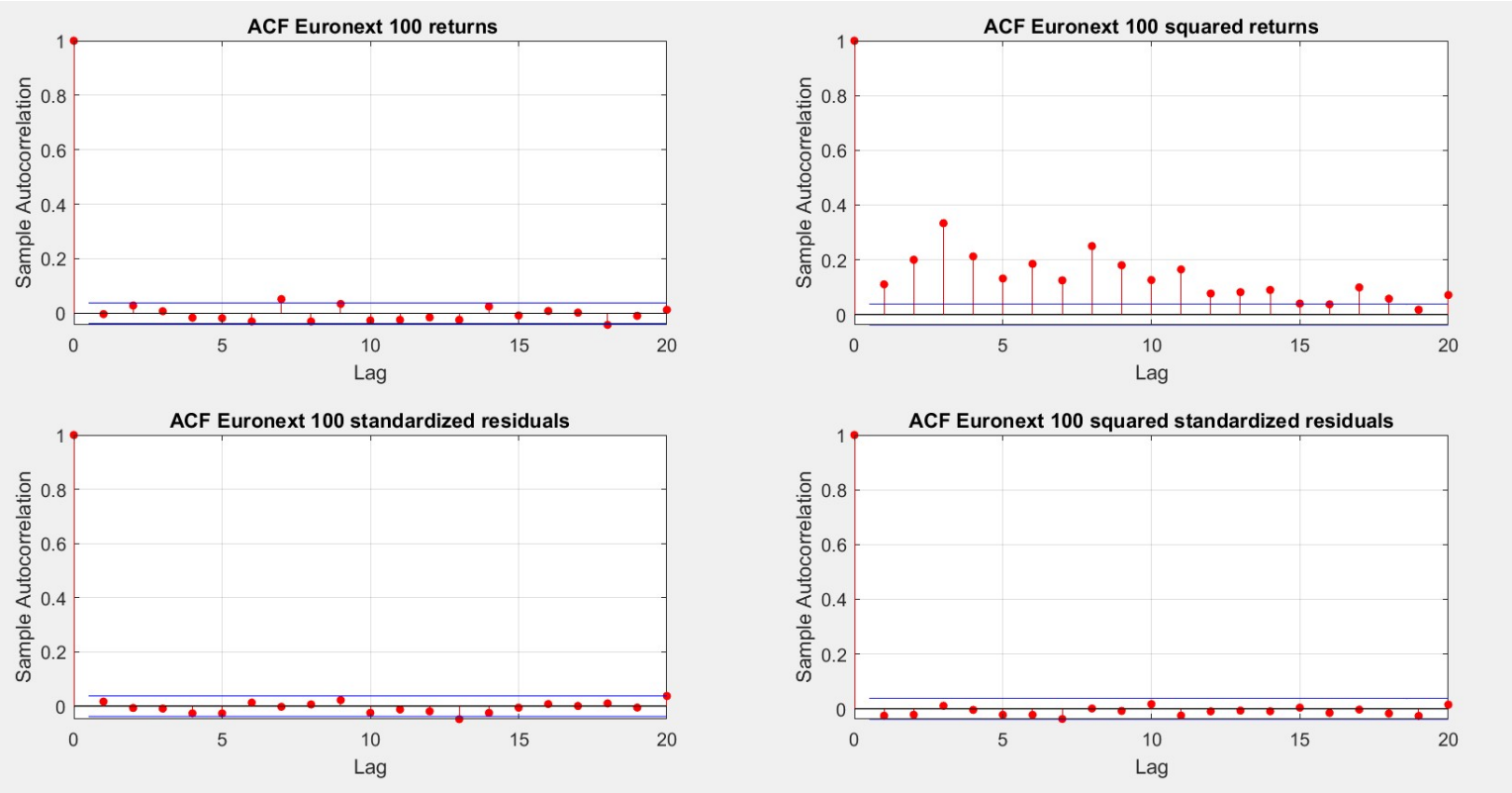

Figure 3.121: Autocorrelation Function

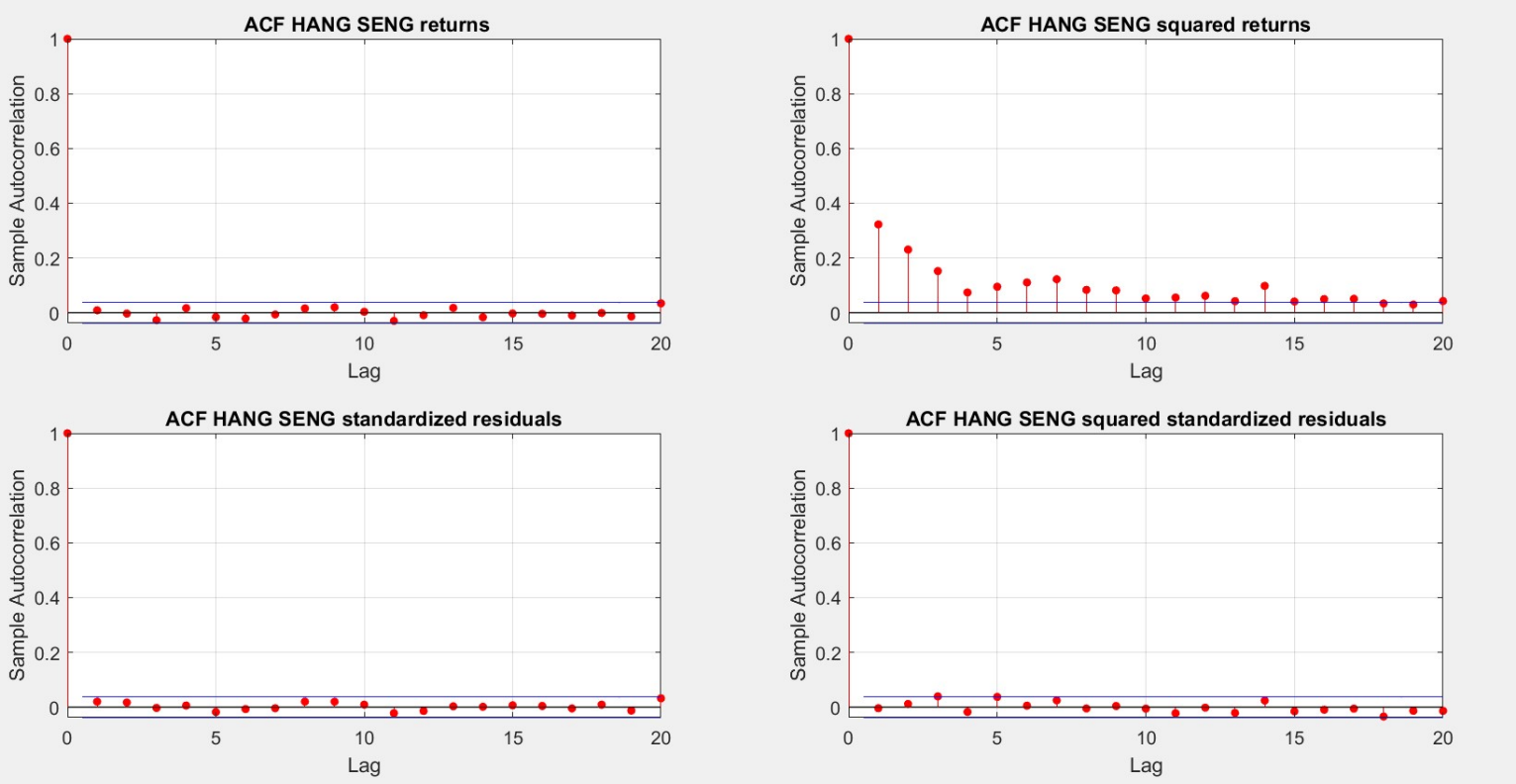

Figure 3.122: Autocorrelation Function

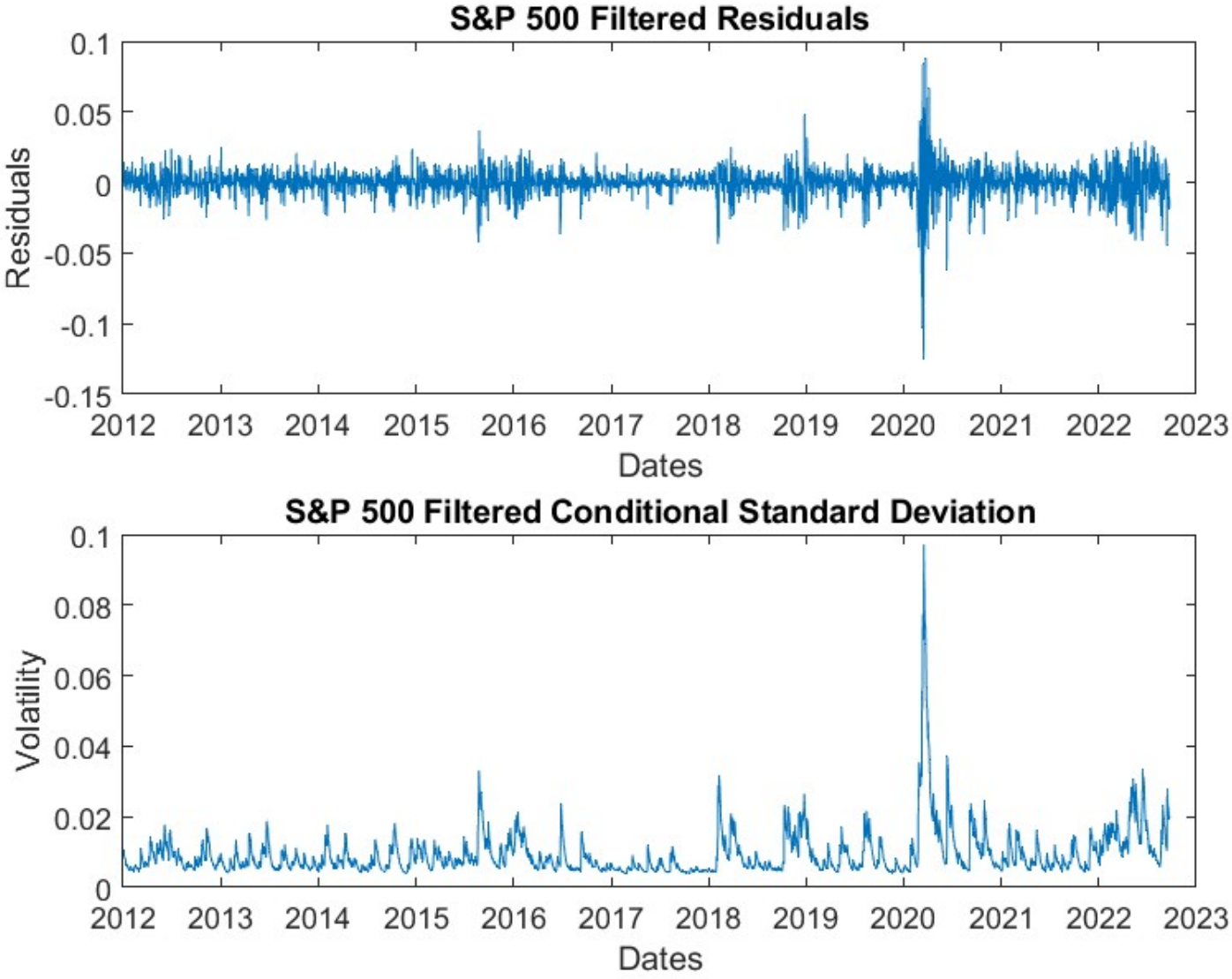

Figure 3.123: Residuals and Volatility

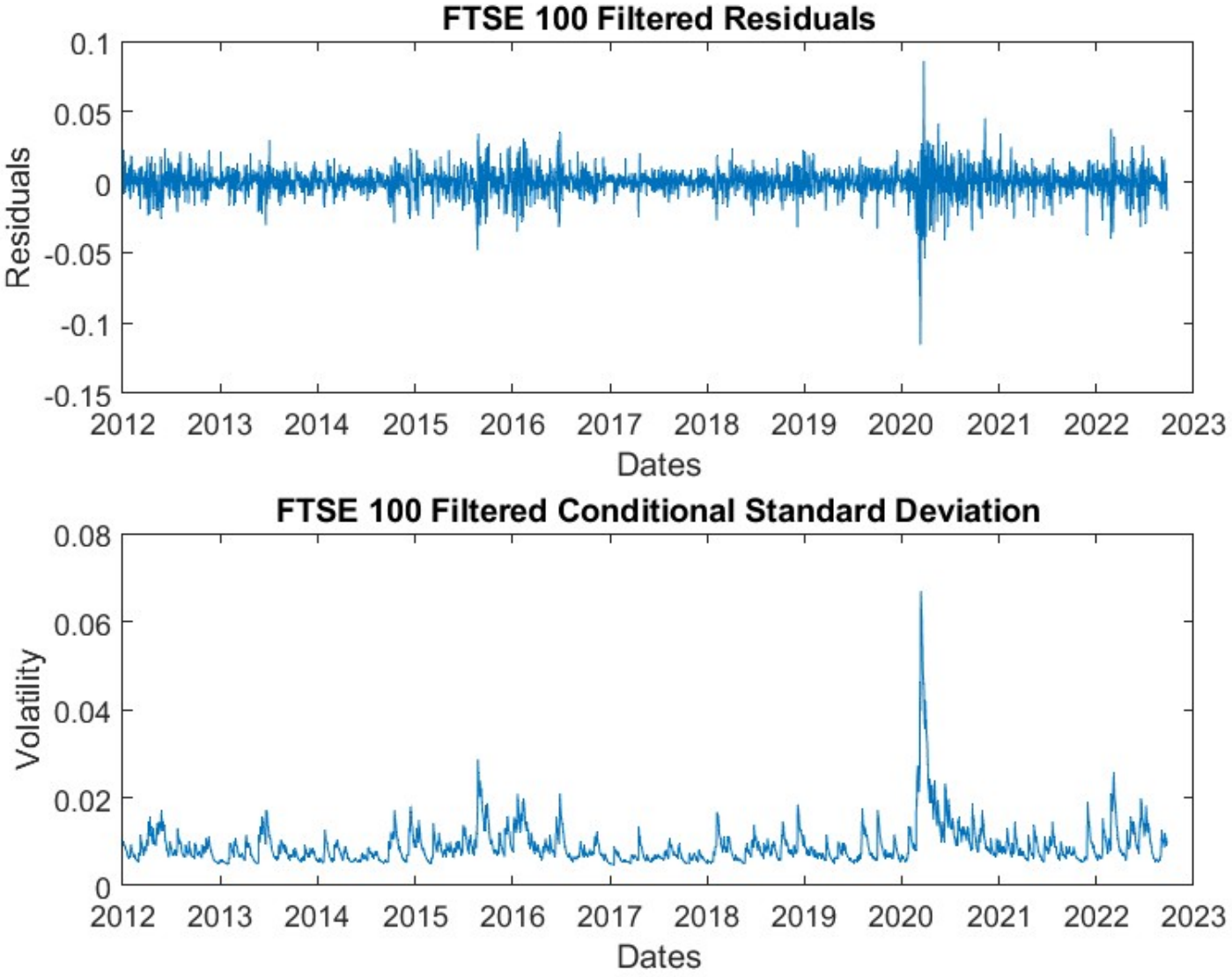

Figure 3.124: Residuals and Volatility

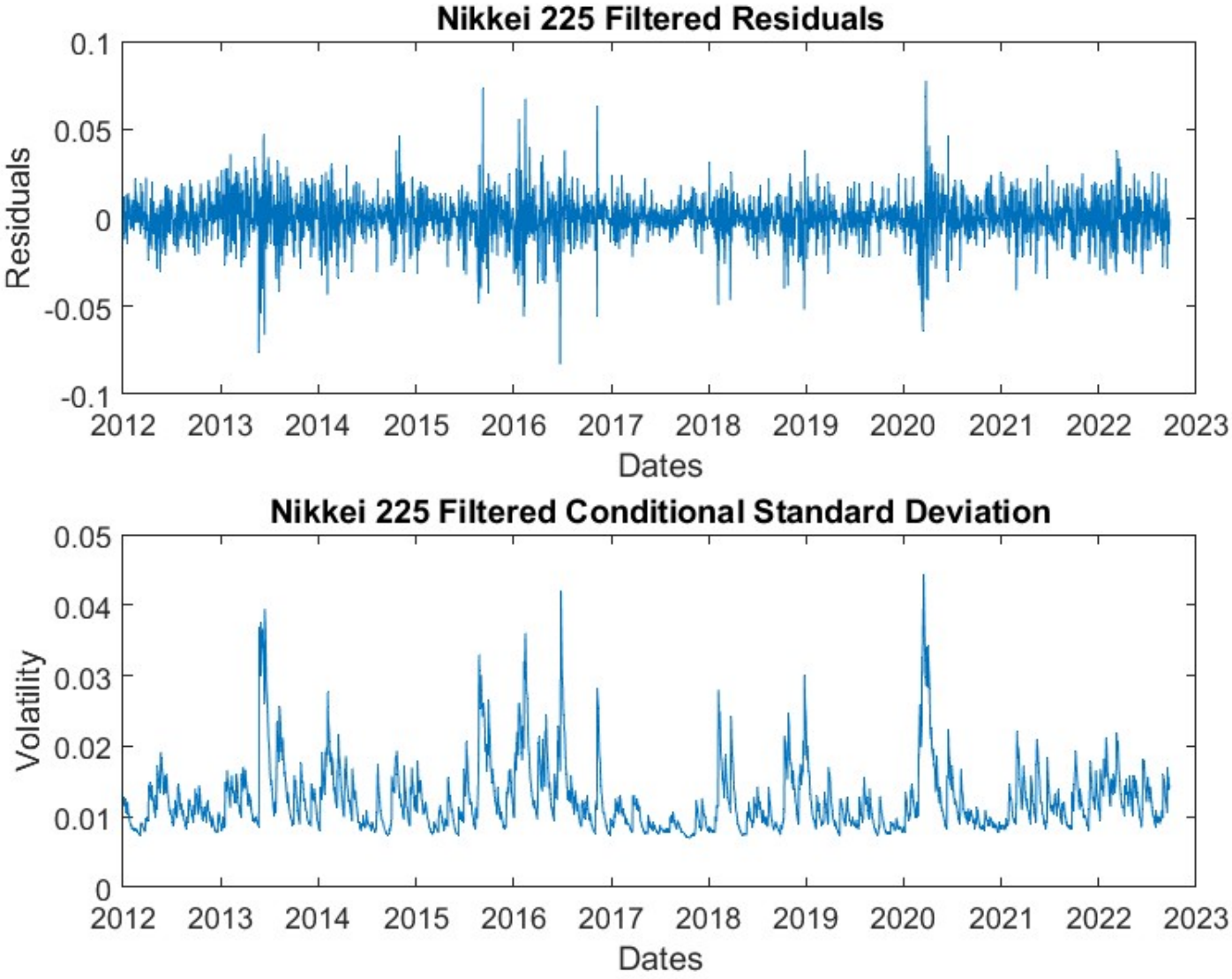

Figure 3.125: Residuals and Volatility

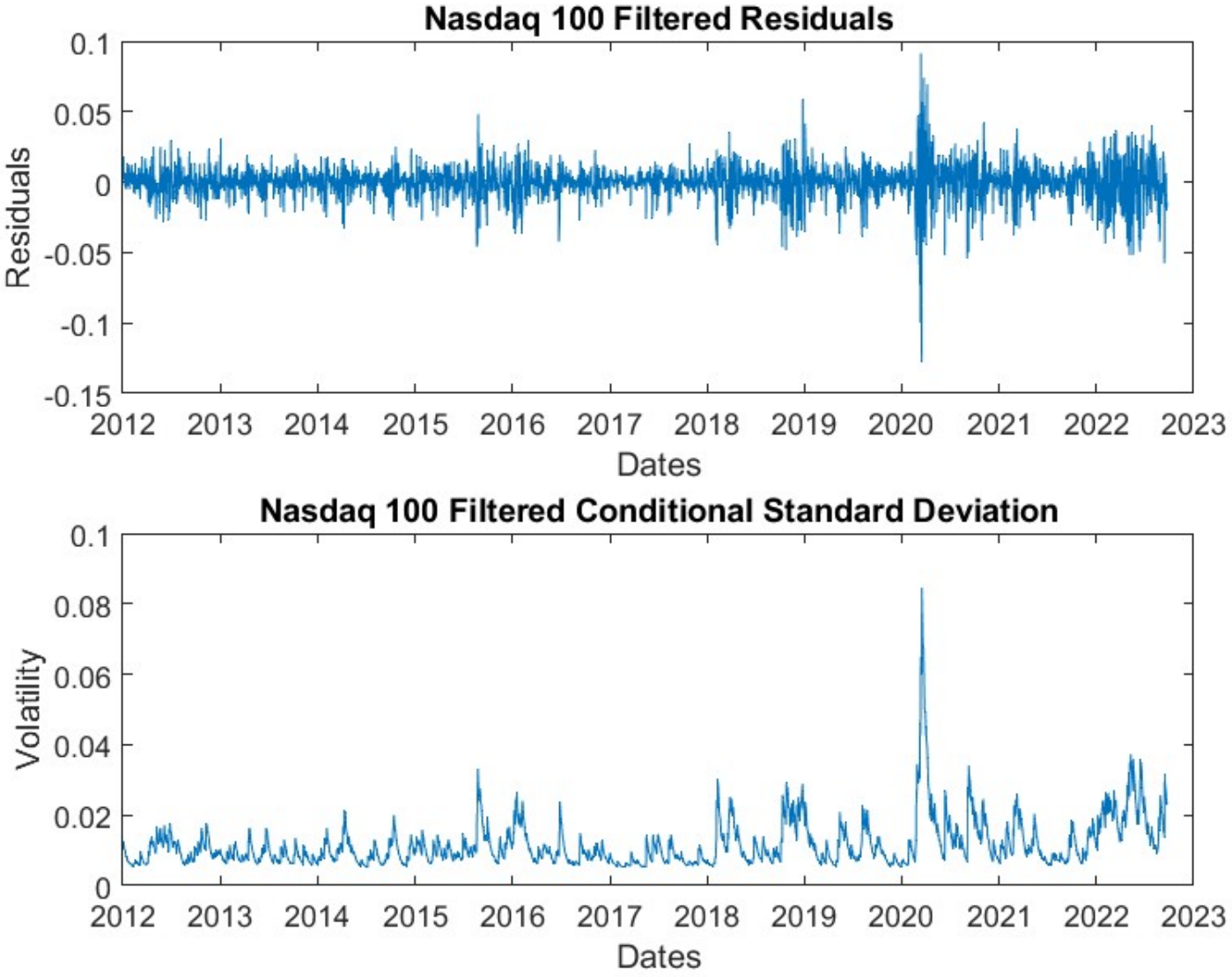

Figure 3.126: Residuals and Volatility

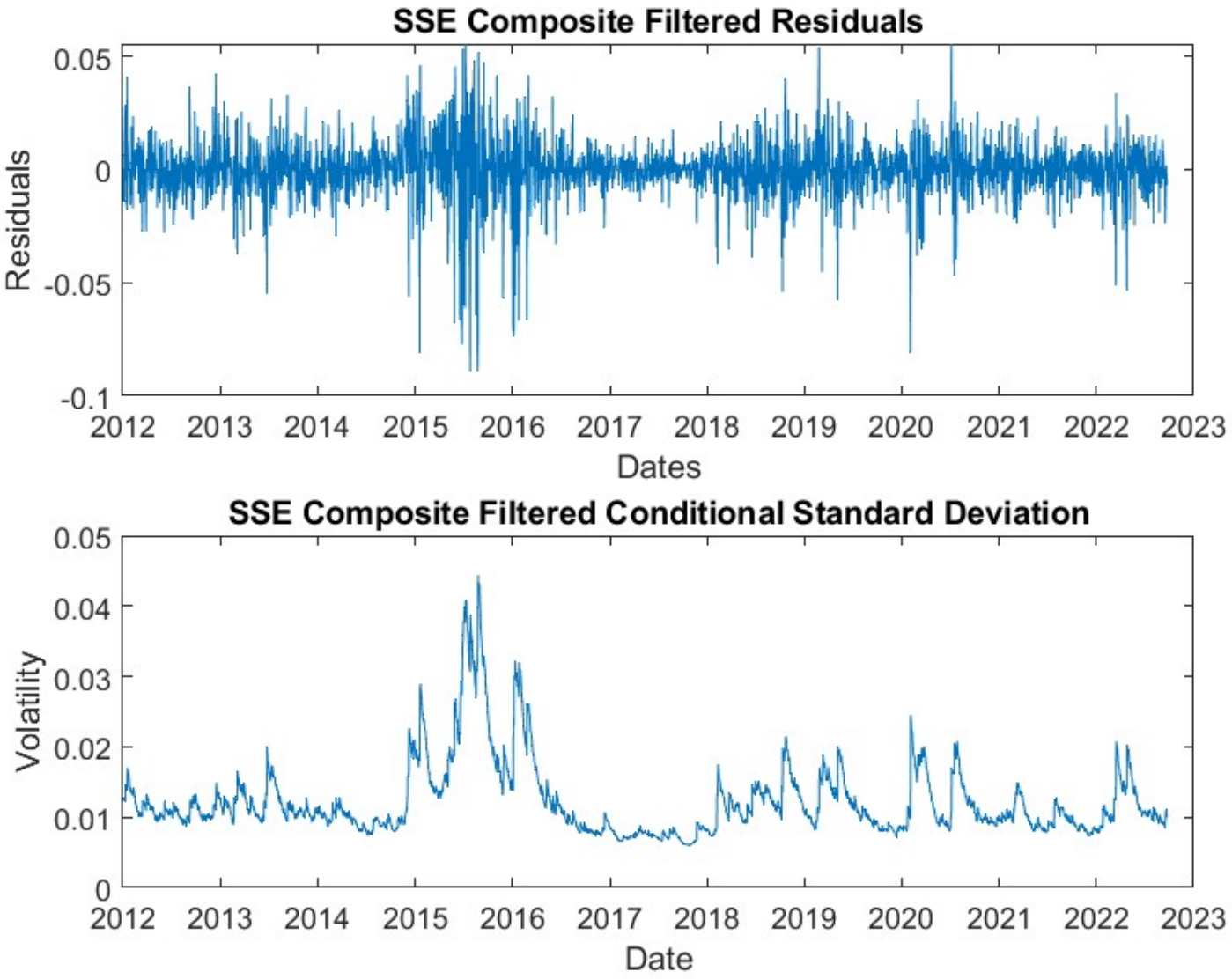

Figure 3.127: Residuals and Volatility

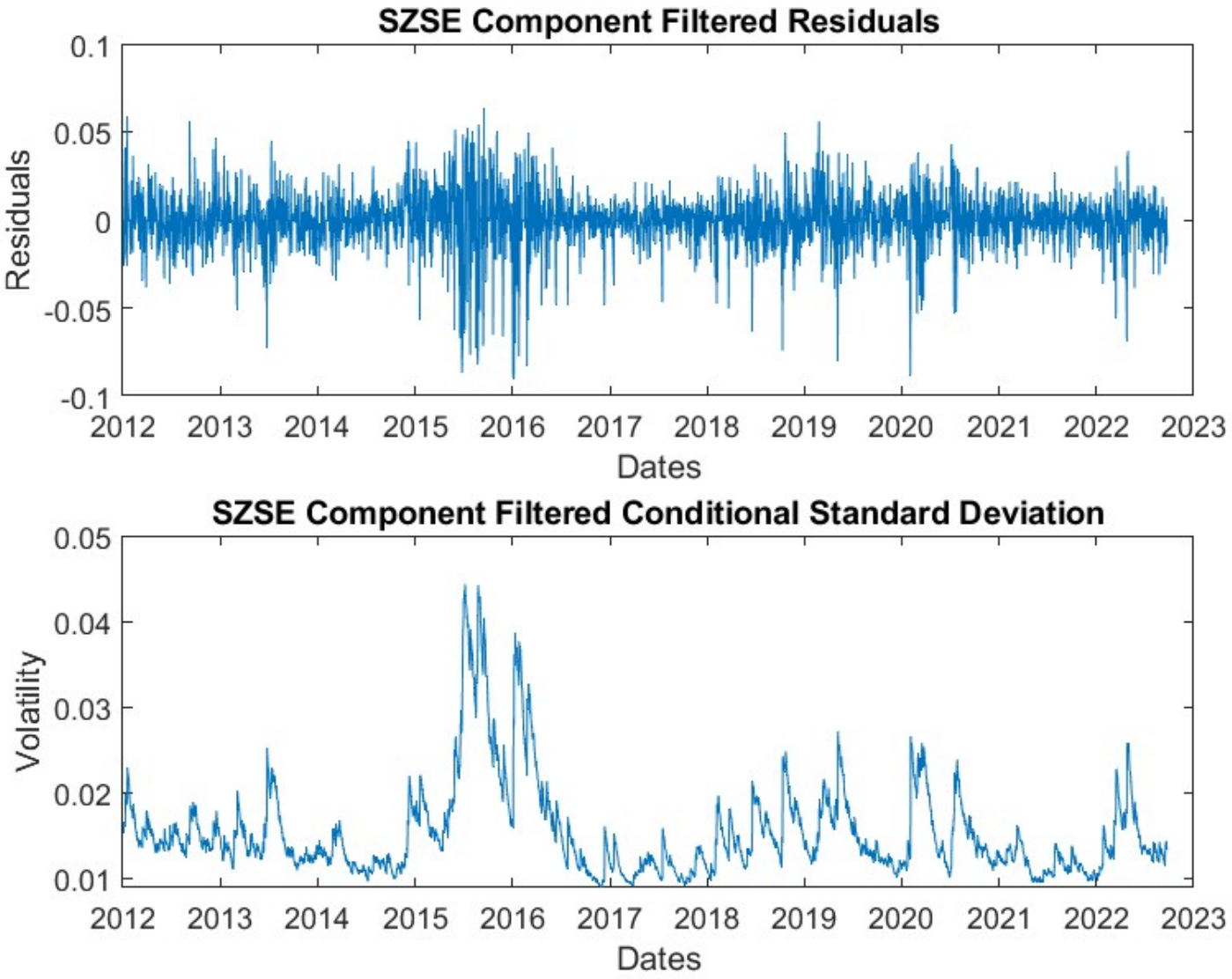

Figure 3.128: Residuals and Volatility

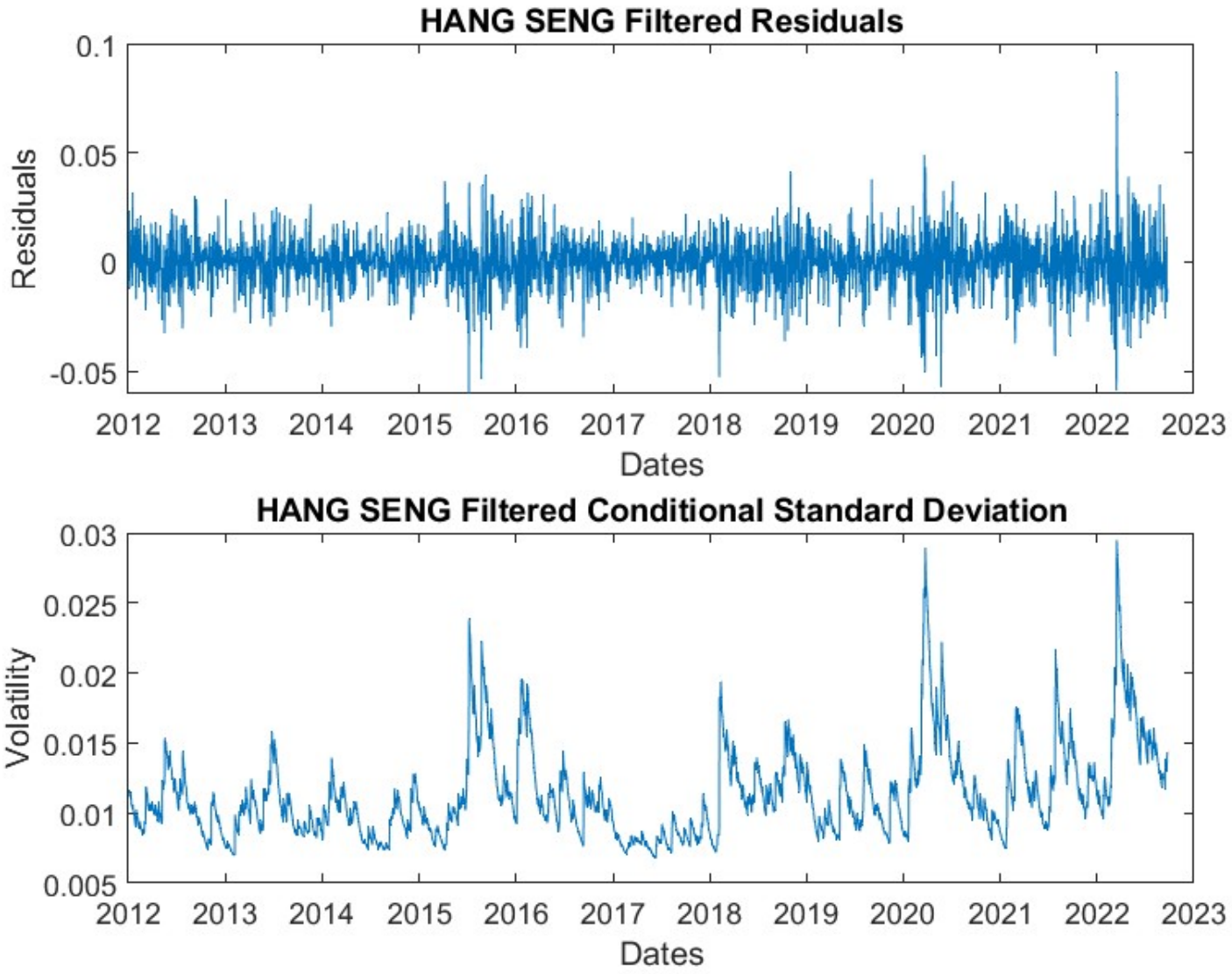

Figure 3.129: Residuals and Volatility

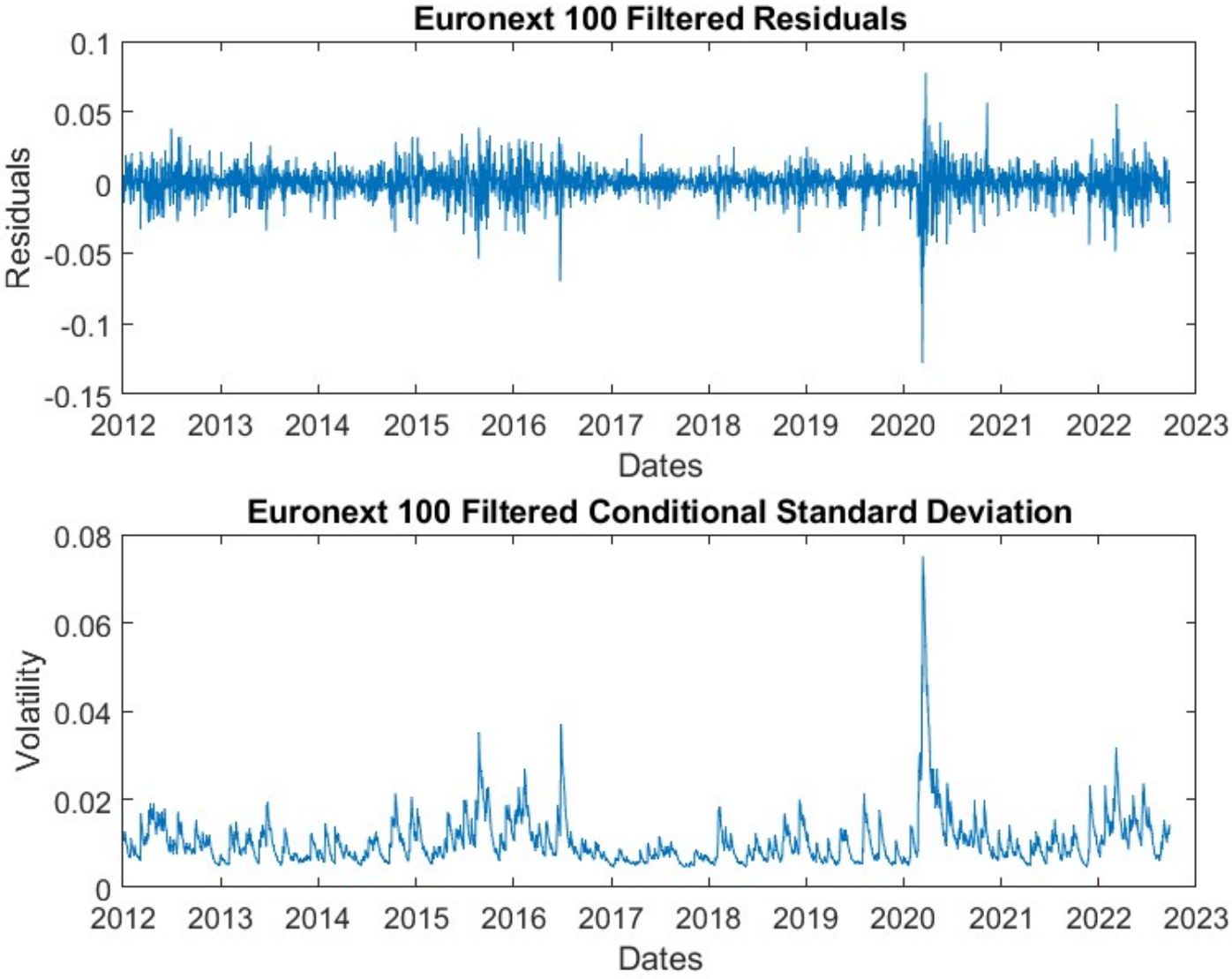

Figure 3.130: Residuals and Volatility

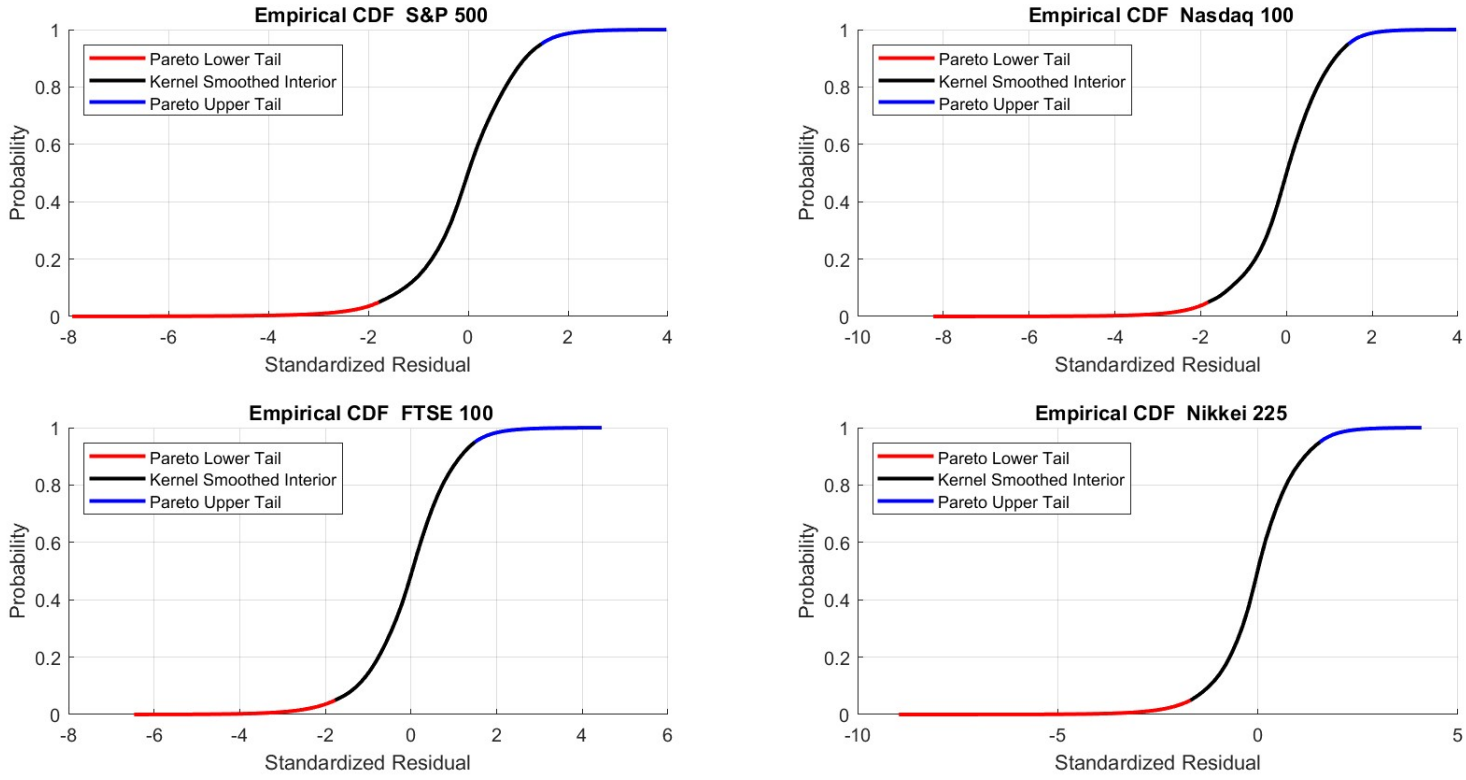

Figure 3.131: Empirical CDF

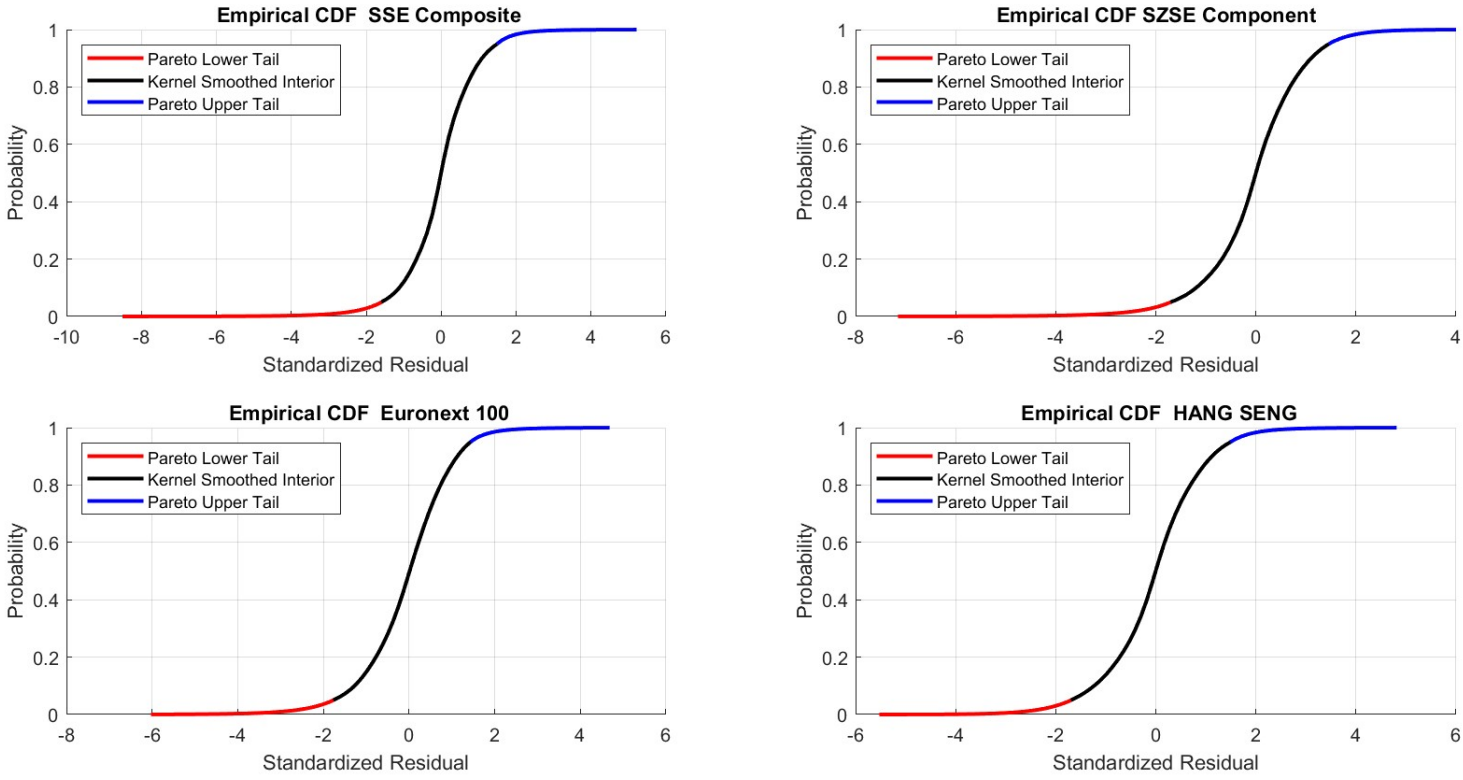

Figure 3.132: Empirical CDF

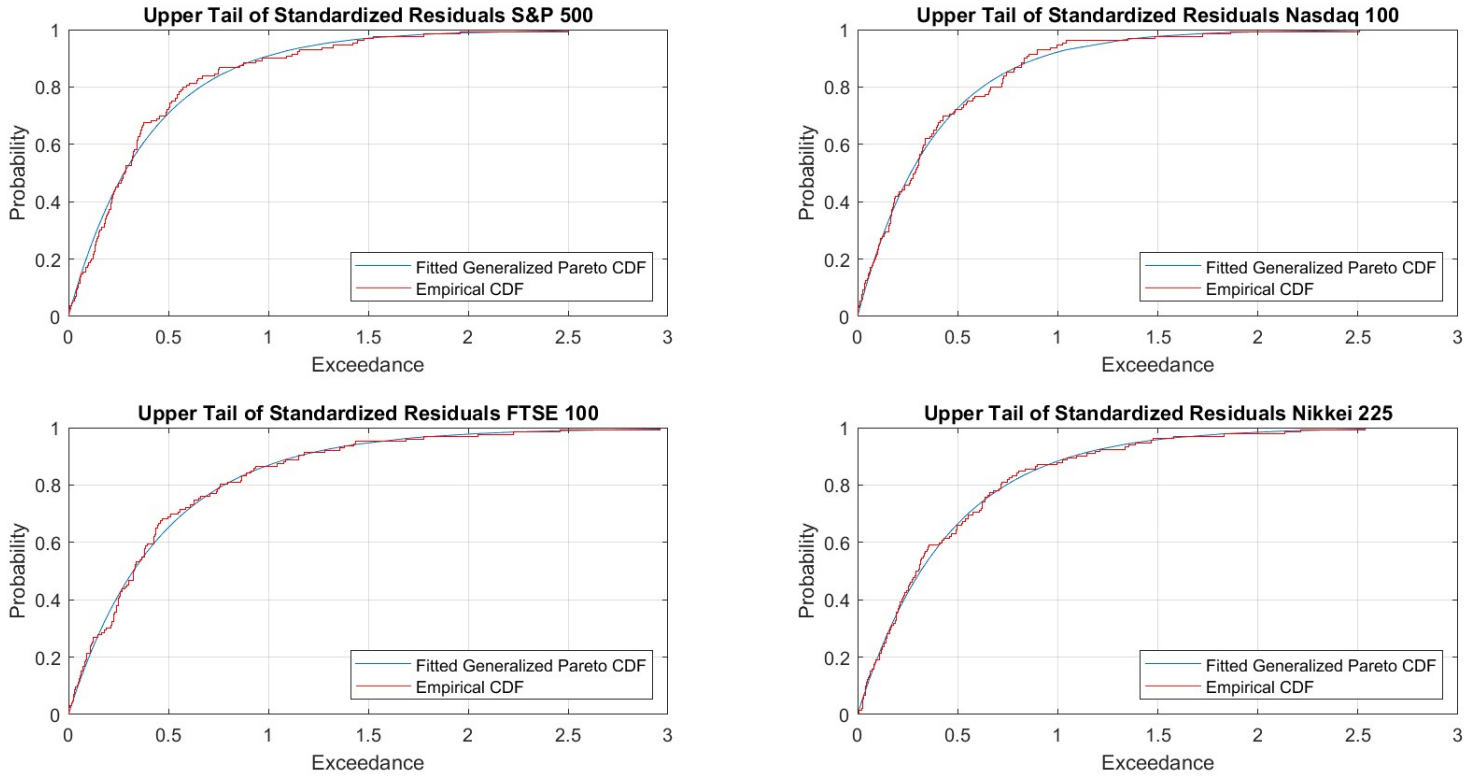

Figure 3.133: Upper Tail of Standardized Residuals

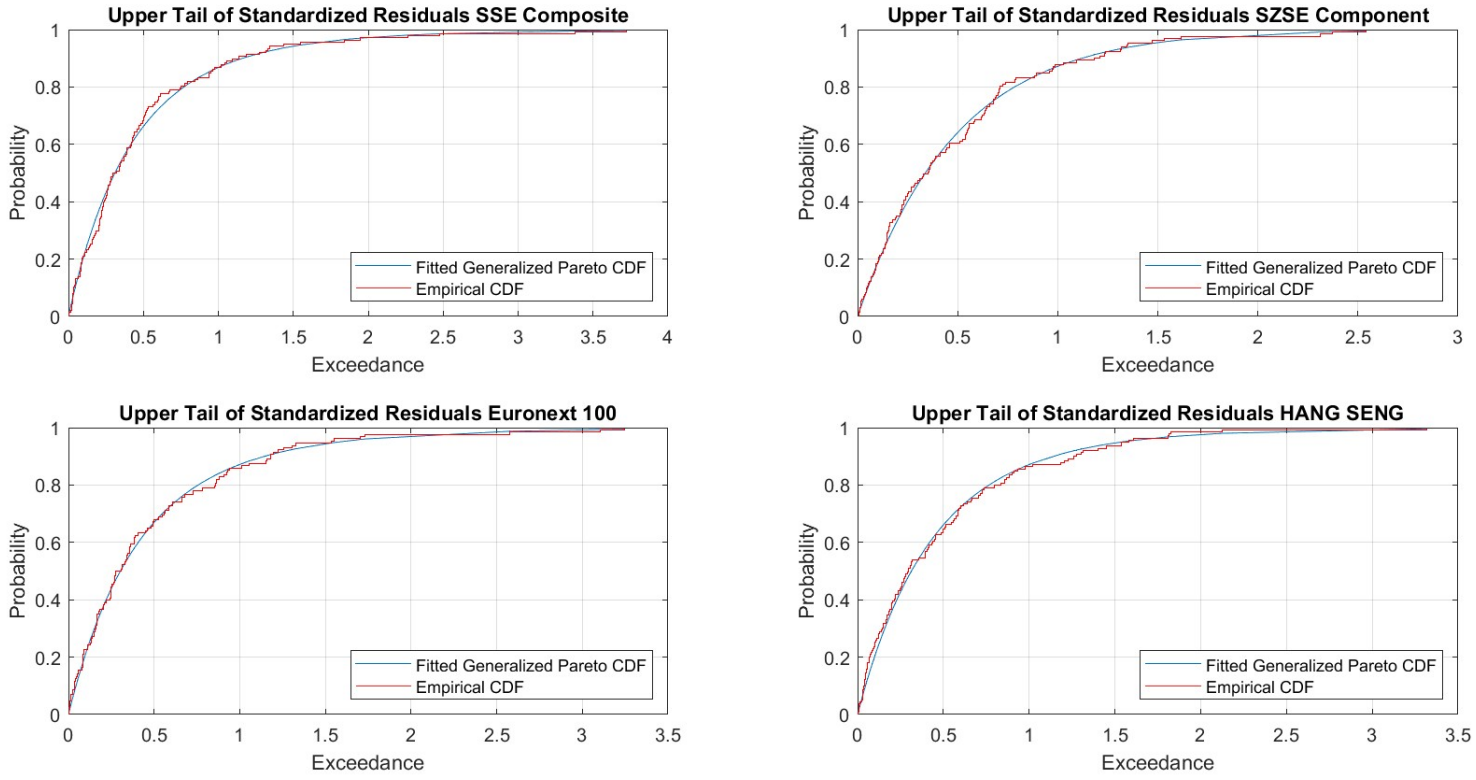

Figure 3.134: Upper Tail of Standardized Residuals

Bibliography

A. Azzalini, A. Capitanio.The Skew-Normal and Related Families. Cambridge University Press. Multivariate Normal Variance Mixtures in R.

M. Billio, L. Pelizzon(2014). Misure econometriche di connettività e rischio sistemico nei settori finanziario e assicurativo europei, SYRTO Working Paper Series ,Working paper No 10.

E. Brodin, C. Klüppelberg(2006). Modeling, estimation and visualization of multi-variate dependence for high-frequency data. Munich University of Technology.

V. Chavez-Demoulin, P. Embrechts(2009). An EVT primer for credit risk, The Oxford Handbook of Credit Derivatives 73 (1), 500-532.

A. Di Clemente (2015). Modellizzazione e Gestione dei Rischi Finanziari attraverso un Approccio di Portafoglio. McGraw-Hill Education.

C.Czado,T.Nagler 2021. Vine copula based modeling. Annual Review of Statistics and Its Applications.

J.Danielsson(2011). Financial Risk Forecasting, The Theory and Practice of Forecasting Market Risk, with Implementation in R and Matlab, pp 97-207. John Wiley & Sons Ltd, West Sussex, United Kingdom.

G. De Luca, P.Zuccolotto (2021). Hierarchical time series clustering on tail dependence with linkage based on a multivariate copula approach. International Journal of Approximate Reasoning.
www.elsevier.com/locate/ijar

G. De Luca, P.Zuccolotto(2011). A tail dependence-based dissimilarity measure for financial time series clustering pp323-340. Springer-Verlag.

F.Durante, R.Pappadà, N.Torelli (2013). Clustering of financial time series

in risky scenarios, pp359-376. Springer Verlag Berlin Heidelberg.

F.Durante, E.Foscolo, P.Jaworski, H.Wang (2013).A spatial contagion measure for financial time series.Expert Systems with Applications, Volume 41, Pages 4023-4034.

S. Demarta, A. J. McNeil(2004). The t Copula and Related Copulas, Department of Mathematics, Federal Institute of Technology, ETH Zentrum CH-8092, Zurich.

V. Durrleman, A. Nikeghbali and T. Roncalli (2000).Copulas for Finance A Reading Guide and Some Applications.
Electronic copy available at: https://ssrn.com/abstract=1032533

P. Embrechts, A. McNeil, D. Straumann(1998). Correlation and Dependency in Risk Management: Properties and Pitfalls

P. Embrechts et al.(2002). Correlation and dependence in risk management: Properties and pitfalls in Risk Management: Value at Risk and Beyond (M. Dempster, Ed.) Cambridge University Press, Cambridge, 176–223

R.Gencay, F. Selcuk, A. Ulugülyağci(2001). EVIM: A Software Package for Extreme Value Analysis in MATLAB, Studies in Nonlinear Dynamics and Econometrics,213–239.

C. Genest, A. C. Favre(2007). Everything You Always Wanted to Know about Copula Modeling but Were Afraid to Ask, Journal of Hydrologic Engineering,368. Asce.

C. Genest and L.P. Rivest (1993). Statistical Inference Procedures for Bivariate Archimedean Copulas, Journal of the American Statistical Association , Sep., 1993, Vol. 88, No. 423 , pp. 1034-1043. Taylor & Francis, Ltd. on behalf of the American Statistical Association.

I. Gijbels, V. Kika, M. Omelka, (2020). Multivariate Tail Coefficients: Properties and Estimation. Entropy 2020, 22, 728.
www.mdpi.com/journal/entropy

C.Grazian, B.Liseo (2017). Approximate Bayesian Inference in Semiparametric Copula Models pp991-1016. International Society of Bayesian Analysis.

C.Grazian, B.Liseo. Approximate Bayesian Computation for Copula Es-

timation pp 1-9. Objective Bayesian analysis for the multivariate skew -t model. Statistical Methods and Applications manuscript.

F. F. Heinz, D. Rusinova(2015). An alternative view of exchange market pressure episodes in emerging Europe: an analysis using Extreme Value Theory (EVT), Working Paper Series N 1818, Europian Central Bank, Eurosystem.

E. Hintz, M. Hofert, C. Lemieux (2022). The R Package nvmix. Journal of Statistical Software.

P. Jaworski, F. Durante, W. Härdle, T. Rychlik (2010). Copula Theory and Its Applications, Proceedings of the Workshop held in Warsaw, 25–26 September 2009,Lecture Notes in Statistics–Proceedings, Springer-Verlag Berlin Heidelberg.

H. Joe (1997) Multivariate Models and Dependence Concepts. Monographs in Statistics and Probability. Chapman and Hall, London.

H. Joe, H. Li, A.K. Nikoloulopoulos(2010). Tail dependence functions and vine copulas, Journal of Multivariate Analysis 101, pp. 252-270. Academic Press, Inc. Orlando, FL, USA.

Joe, Harry (2015). Dependence Modeling with Copulas, Monographs on Statistics and Applied Probability, 134. CRC Press, Boca Raton, FL.

E. Këllezi, M. Gilli (2017). An Application of Extreme Value Theory for Measuring Financial Risk, Revised version of a paper appeared in Computational Economics 27(2-3) (2006).

D.O. Ledenyov, V. O. Ledenyov(2013). To the problem of evaluation of market risk of global equity index portfolio in global capital markets, MPRA Paper No. 47708. James Cook University, Townsville, Australia.

T.Leung Lai, H. Xing(2008). Statistical Models and Methods for Financial Markets, Springer Texts in Statistics, pp 175-334.

J.F. Mai, M. Scherer(2012). Simulating Copulas, Stochastic Models, Sampling Algorithms and Applications,pp 161-256. Imperial College Press, 57 Shelton Street Covent Garden, London.

A. J. McNeil, R. Frey and P. Embrechts(2015). Quantitative Risk Management: Concepts, Techniques and Tools, Princeton Series in Finance. Prince-

ton University Press, Princeton,NJ.

Y. Malevergne and D. Sornette (2006).Extreme Financial Risks, From Dependence to Risk Management. Springer Berlin, Heidelberg.

R. Nekhili(2006). Extreme contagion in emerging stock markets using extreme value copulas, The VII workshop on Quantitative Finance, University of Wollongong, Australia.

M. Rocco(2014). Extreme Value Theory in Finance: a survey, Financial Stability Unit, Bank of Italy, Journal of Economic Surveys (2014) Vol. 28, No. 1, pp. 82–108.2012 John Wiley & Sons Ltd.

F.Schmid and R. Schmidt (2007). Multivariate extensions of Spearman's ρ and related statistics. Statistics & Probability Letters,2007, vol. 77, issue 4, 407-416.

R. Smith, (1985) Statistics of Extreme Values. International Statistical Institute, Amsterdam

M. S. Smith (2011). Bayesian Approaches to Copula Modelling, Hierarchical Models and MCMC: A Tribute to Adrian Smith, Melbourne Business School, University of Malbourne.

M. S. Smith, Q. Gan, R. J. Kohn (2012). Modelling dependence using skew t copulas: Bayesian inference and applications, Journal of Applied Econometrics, 27(3), 500–522.

T. Yoshiba(2015) Maximum likelihood estimation of skew-t copulas with its applications to stock returns, Bank of Japan, Chuo-ku, Tokyo 103-8660,Japan, The Institute of Statistical Mathematics, Tachikawa, Tokyo 190-8562, Japan.

T. Yoshiba(2018). Maximum likelihood estimation of skew-t copulas with its applications to stock returns. Journal of Statistical Computation and Simulation.