Università degli Studi di Torino
Politecnico di Torino

PhD in Pure and Applied Mathematics

# MODELING DEPENDENCE IN ESTIMATION OF PORTFOLIO'S VALUE AT RISK

Supervisors:
Prof. Franco Pellerey
Prof.ssa Laura Sacerdote

Candidate:
Ottavia Telve

October $4^{th}$, 2018

# Contents

# Preface

For financial institutions (shortly, FI) facing continuing market volatility and upheaval, risk management plays a double role.
Defensive ex–post risk management, focused on regulatory compliance, controls and, basically, on staying out of trouble, remains vitally important (Stress Test, IFRS9, Risk Appetite Framework, introduction of Risk Policy, Regulatory rules).
But, more and more often, risk management is playing an active role (ex–ante) in helping FI to identify and achieve growth and avoid losses from market volatility and crisis also supporting Investment Team in the use of quantitative asset allocation models, see for example Cowell (2002) or Faber (2013).
Modern Risk Management is evolving in this direction using more and more sophisticated quantitative models for forecasting the distribution of portfolios returns, in order to intercept sudden changes in market equilibrium and in diversification of asset allocation.
These information are translated in risk indicators that allow to summarize the complex dynamics of portfolios (Artzner, Delbaen, Eber, & Heath, 1999).
Based on these indicators the FI can create 'Risk Policies' that translates the risk appetite of the group in 'Risk Indicators limits'. It is essential to monitor these limits on real–time bases in order to promptly adjust portfolio allocations in case of sudden changes in market condition or in group risk appetite.
It is, therefore, important for FI to have frequently updated risk indicators, in order to monitor the situation and to have time to act in one way or another. Unfortunately this is difficult because of the complexity of models and high dimensional portfolios (for example a medium Investment Management Company has at least $[200 - 300]$ portfolios invested in $[5000 - 10000]$ assets). Hence the trade-off is between reliability of forecasting and small computing time and simple application to complex allocation.

Two are the targets to focus on: volatility and diversification. The first is both specific of each asset, then, in a mathematical view, concerns the univariate evolution of a financial time series and on the portfolio. The second tries to describe the dependence dynamic of assets in a portfolio, hence deals with multivariate models/distributions.
In last years most of the financial literature investigate efficient ways to describe and forecast asset returns and volatility in univariate or at least bivariate case, see,

e.g. (Dias & Embrechts, 2010), (Manner & Reznikova, 2012), (Patton, 2006), (Okimoto, 2008). There is, on the other side, lack of studies on multivariate dependence forecasting, for example Embrechts and Hofert (2014) or Stoeber and Czado (2012).
The reason is the complexity of structures and the time-consuming calculations. Hence in FI the most used models are still based on multivariate Gaussian distributions, even if it is known that in this way the forecasting looses the specific non-linear dependence. Another approach involves the historical simulation with some adjustment like EWMA(Equally Weighted Moving Average) like J.P.Morgan (1985), with approximation of dependence via PCA (Principal Component Analysis) or parallel historical bootstrap, as in Zenti and Pallotta (2002).

The aim of this thesis is try to give a contribution to the trade off between applicability and reliability of forecasting models. We presents two different approaches:

- Operating approach: uses well known models focusing on dependence to end up with an improved model that is also fast in calculation, can deal with huge portfolio dimensions and with good results in back-testing.

- Innovative approach: investigate the use of an instrument similar to Copula that can be useful for specific portfolio analysis.

We focus on dependence dynamic of asset returns. It changes over time, especially in Crisis periods, when there are fast-moving dependence structures that can compromise some allocation actions finalized to preserve the portfolio returns looking at diversification.
In Chapter 1 we review the known univariate models most used in finance, and also in this thesis. We return the outcome in modeling dependence. In forecasting financial returns there are evidences of tail dependence, comonotonicity in turbulent periods and non-linear dependent between assets. We consider Copula function as the proper instrument to deal with the characteristics described above, hence we recall the theoretical back-ground of Copula, Nelsen (2006). Looking at different Copula families we underlying useful properties and limitations. In the next chapter we present different approaches in forecasting time–series, with some examples. In each approach we try to enlighten key strengths and limits.
Last chapter presents the original contributions of of this thesis and is organized in two main sections introducing the application of two different models. We choose a portfolio composed by International stock indexes and apply forecasting models on this portfolio in 3 different turbulent financial periods long enough to perform also some out-of-sample back–test and compared with analog method existing.

# Chapter 1

# Mathematical background for financial modeling

The theory and practice of asset price evolution over time involves financial time series analysis and determines a key role of statistical theory and methods for many financial studies. Concepts such as leptokurticity of the marginal distributions, asymmetries, stationarity, autocorrelation, white noise, innovation are central in this context. Primary models to refer to are, for example, the autoregressive integrated moving average (ARIMA) models and vector autoregressive models (VAR) or Autoregressive conditional heteroskedasticity (ARCH).
Another important topic in financial time series analysis is represented by comovements and dependence between assets and risk factors. To manage the risk of a portfolio it is important to understand the diversification effect of asset allocation, the higher contributors to risk and prevent comonotonicity effects.
In this chapter we recall mathematical background, basic definitions and main properties of financial time series, stochastic processes. Furthermore, we describe the copula framework as an instrument to deal with dependence separately from the marginal behavior of assets.

## 1.1 Time series for financial modeling

Financial data include different time series. However, most financial studies involve returns, instead of prices, of assets. (J. Campbell, Lo, & A.C., 1997) give two main motivations for using returns. First, for average investors, return of an asset is a complete and scale-free summary of the investment opportunity. Second, return series are easier to handle than price series because **the former have more attractive statistical properties** (Asset prices are generally non stationary. Returns are usually stationary).
There are, however, several definitions of an asset return.

**Definition 1.1.1.** *Let $T \subseteq \mathbb{Z}$ and consider the sequence $\mathcal{P} = \{P_{i,t}, t \in T\}$ of the daily observations of asset i price. Let $\mathcal{X} = \{X_t, t \in T\}$ be the corresponding sequence of asset return; we define*
***Compound return** (Groppelli & Nikbakht, 2000) as:*

$$X_{i,[t-1,t]} = X_{i,t} = \frac{P_{i,t} - P_{i,t-1}}{P_{i,t-1}} \tag{1.1.1}$$

*While, we define **continuous return** (Groppelli & Nikbakht, 2000) as:*

$$X_{i,[t-1,t]} = X_{i,t} = ln\left(\frac{P_{i,t}}{P_{i,t-1}}\right) = ln\left(P_{i,t}\right) - ln\left(P_{i,t-1}\right) \tag{1.1.2}$$

In this thesis we use continuous asset return (1.1.2). This is commonly used choice because it simplifies the passage from daily to weekly or monthly returns. Indeed, let $X_{i,t}$ be the return of asset $i$ between $t-1$ and $t$ with daily interval and let $X_{i,[t,t+s]}$ be the return between $t$ and $t+s$ with continuous returns, it holds:

$$X_{i,[t,s]} = ln\left(\frac{P_{i,s}}{P_{i,t}}\right) = ln\left(P_{i,s}\right) - ln\left(P_{i,t}\right) = \sum_{v=t}^{s} X_{i,v}.$$

Typically $s = 5$ indicates weekly returns and $s = 21$ means monthly returns. The historical sample used in models is the last price of the day, hence the analyzed financial time series are discrete in time (daily frequency) with continuous domain.

### 1.1.1 Conditional mean: ARMA models

A class of processes often considered to describe asset returns includes the ARMA (AutoRegressive Moving Average) models:

**Definition 1.1.2.** *The process $\mathbf{X} = \{X_t, t \in T\}$ follows an ARMA(p,q) process if for every t the r.v. $X_t$ satisfies:*

$$X_t = \sum_{j=1}^{p} \phi_j X_{t-j} + \sum_{i=1}^{q} \theta_i Z_{t-i} + Z_t. \tag{1.1.3}$$

*where $Z_t$ is a white noise and $\phi_j$ and $\theta_i$ are the ARMA parameters.*

**Definition 1.1.3.** *$\mathcal{Z} = \{Z_t, t \in T\}$ is a white-noise process if it is a sequence of uncorrelated random variables: $\forall t \neq s : cov(Z_t Z_s) = 0$.*

Before detailing of this kind of processes, we recall some definitions.
There exist different type of stationarity:

**Definition 1.1.4.** *$\mathcal{X} = \{X_t, t \in T\}$ is said **strong stationary** process if $\forall\, t_1, t_2, \ldots, t_n$ and $\forall k$ the joint distribution of $\{X_{t_1}, X_{t_2}, \ldots, X_{t_n}\}$ is the same of the joint distribution of $\left\{X_{t_{1+k}}, X_{t_{2+k}}, \ldots, X_{t_{n+k}}\right\}$.*

**Definition 1.1.5.** *$\mathcal{X} = \{X_t, t \in T\}$ is a process with **stationarity of order** $m$ if $\forall t_1, t_2, \ldots, t_n$ and $\forall k$ the joint moments to order $m$ of $\{X_{t_1}, X_{t_2}, \ldots, X_{t_n}\}$ are the same of $\{X_{t_{1+k}}, X_{t_{2+k}}, \ldots, X_{t_{n+k}}\}$.*

**Definition 1.1.6.** *$\mathcal{X} = \{X_t, t \in T\}$ is a process with **stationarity of order** $2$ if:*

1. *$E[X_t] = \mu$, with $\mu$ not time dependent;*

2. *$E[X_t^2] = \mu_2$, with $\mu_2$ not time dependent;*

3. *$E[X_t X_s] = \phi(t-s)$, $\forall t, s$.*

From now on, we consider only stationary processes of order 2.
An important aspect of these kind of processes is the presence of dependence between different times of the same process:

**Definition 1.1.7.** *The autocovariance function of $\mathcal{X} = \{X_t, t \in T\}$ is:*

$$\gamma_\tau = \gamma(X_t, X_{t+\tau}) = cov(X_t X_{t+\tau}) = E[(X_t - \mu)(X_{t+\tau} - \mu)] \tag{1.1.4}$$

*where $\tau$ is the time lag and $\gamma_0$ is the variance.*

**Definition 1.1.8.** *The autocorrelation function of $\mathcal{X} = \{X_t, t \in T\}$ is:*

$$\rho(\tau) = \rho_{X_t, X_{t+\tau}} = \frac{cov(X_t X_{t+\tau})}{\sqrt{var(X_t) var(X_{t+\tau})}} \tag{1.1.5}$$

**Definition 1.1.9.** *Partial autocorrelation coefficient function (PACF) between $X_t$ and $X_{t-r}$, given $X_{t-1}$, is:*

$$\rho_{X_t, X_{t-r}|X_{t-1}} = \frac{\rho_{X_t, X_{t-r}} - \rho_{X_t, X_{t-1}} \rho_{X_{t-1}, X_{t-2}}}{\sqrt{\left(1 - \rho^2_{X_t, X_{t-1}}\right)\left(1 - \rho^2_{X_{t-1}, X_{t-2}}\right)}} \tag{1.1.6}$$

A specific class of processes belonging to ARMA class is:

**Definition 1.1.10.** *$\mathcal{X} = \{X_t, t \in T\}$ is an autoregressive process of order $p$ $[AR(p)]$ if it satisfies:*

$$X_t = c + \sum_{k=1}^{p} \phi_k X_{t-k} + Z_t \tag{1.1.7}$$

*with $Z_t$ a white noise process and $\phi_k$ the AR parameters.*

The PACF of an $AR(p)$ process is zero at lag $\tau \geq p+1$, so the appropriate maximum lag $p$ is the one beyond which the partial autocorrelations (1.1.6) are all zero.
In particular $AR(1)$ is the most used process in financial time series and it is stationary of order 1 and if $|\phi| < 1$ is asymptotically stationary of order 2.

Now we focus on the estimation of correct lag $p$ and related AR parameters. We recall here some methods and we introduce the necessary notations while we refer to the literature for a more detailed presentation. Main references for this paragraph are Box, Jenkins, and Reinsel (1994), Greene (1997), Enders (1995) and Hamilton (1994).

There are many techniques to estimate the coefficients of an AR model, such as the ordinary least squares (OLS) procedure or method of moments (MM).
The first step requests to estimate the $p$ order that better fits the data.
Then, for the $AR(p)$ model we should estimate the parameters $\phi_k, k = 1, \cdots, p$.
Since there is a direct correspondence between the $\phi_k$ parameters and the covariance function of the process, we can invert this correspondence inverted to determine the parameters from the autocorrelation function (1.1.8). This is done using the Yule–Walker equations.

The Yule–Walker equations, named for Yule (1927) and Walker (1931), are:

$$\gamma_m = \sum_{k=1}^{p} \phi_k \gamma_{m-k} + \sigma_\varepsilon^2 \delta_{m,0} \tag{1.1.8}$$

where $m = 0, \ldots, p$, yielding $p+1$ equations. Here $\gamma_m$ is the autocovariance function of $X_t$, $\sigma_\varepsilon$ is the standard deviation of the input noise process, and $\delta_{m,0}$ is the Kronecker delta function.
For $m = 0$, (1.1.8) becomes:

$$\gamma_0 = \sum_{k=1}^{p} \phi_k \gamma_{-k} + \sigma_\varepsilon^2, \tag{1.1.9}$$

which can be solved to determine $\sigma_\varepsilon^2$, once $\phi_m,\ m = 1, 2, \cdots, p$ are known.
The remaining set of equations can be rewritten in matrix form, thus getting

$$\begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \vdots \\ \gamma_p \end{bmatrix} = \begin{bmatrix} \gamma_0 & \gamma_{-1} & \gamma_{-2} & \cdots \\ \gamma_1 & \gamma_0 & \gamma_{-1} & \cdots \\ \gamma_2 & \gamma_1 & \gamma_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \\ \gamma_{p-1} & \gamma_{p-2} & \gamma_{p-3} & \cdots \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \vdots \\ \phi_p \end{bmatrix}$$

which can be solved for all $\phi_m,\ m = 1, 2, \cdots, p$.

An alternative formulation makes use of the autocorrelation function. The AR parameters are determined by the first $p+1$ elements $\rho(\tau)$ of the autocorrelation function. The full autocorrelation function can be expressed through

$$\rho(\tau) = \sum_{k=1}^{p} \phi_k \rho(k-\tau). \tag{1.1.10}$$

The above equations (the Yule–Walker equations) provide several methods to estimating the parameters of an $AR(p)$ model, by replacing the theoretical covariances with estimated values.

Alternatively, we can reformulate the problem as an extended form of OLS prediction problem (Von Storch and Zwiers (2001)). Here two sets of prediction equations are combined into a single estimation scheme and a single set of normal equations. One set is the set of forward-prediction equations and the other is a corresponding set of backward prediction equations, relating to the backward representation of the AR model:

$$X_t = c + \sum_{i=1}^{p} \phi_i X_{t-i} + Z_t$$

.
Here, predicted values of $X_t$ are based on the $p$ last observed values of the series. This way of estimating the AR parameters is due to Burg (1968) and is called the Burg method. Burg and later authors (Bos, De Waele, and Broersen (2002) and Brockwell, Dahlhaus, and Trindade (2005)) called these particular estimates "maximum entropy estimates", but the reasoning behind this technique can be applied to any set of estimated AR parameters. Compared to the estimation scheme using only the forward prediction equations, different estimates of the auto-covariances are produced, and the estimates have different stability properties. Burg estimates are associated with maximum entropy spectral estimation.

Other possible approaches to estimation include maximum likelihood estimation (Fisher (1925) , Self and Liang (1987) and Kiefer and Wolfowitz (1956)). Two distinct variants of maximum likelihood are available: in one (broadly equivalent to the forward prediction least squares scheme) the likelihood function considered corresponds to the conditional distribution of later values in the series given the initial $p$ values in the series; in the second, the used likelihood function corresponds to the unconditional joint distribution of all the values in the observed series. Substantial differences in the results of these approaches can occur if the observed series is short, or if the process is close to non-stationarity. For further details see Enders (1995), Greene (1997) and Hamilton (1994).
In this thesis we use the conditional maximum likelihood estimation.

### 1.1.2 Conditional Volatility: ARCH models

A last important phenomenon characterizing financial time series is the volatility clustering, usually referred to as conditional heteroscedasticity. It is often observed that periods of high price volatility follow periods of low volatility and vice versa, in contrast to the often-assumed log-normal distribution of asset price returns.
Volatility clustering is usually approached by modeling the price process with an ARCH-type model. The two most widely-used models are the autoregressive conditional heteroskedasticity (ARCH) introduced by Engle (1982) and later extended to GARCH (Generalized AutoRegressive Conditional Heteroskedasticity) models by Bollerslev (1986).
In this subsection we describe these models used for modeling the volatility of financial time series.

**Definition 1.1.11.** *The process $\{\xi_t, t \in T\}$ follows an ARCH model if for every $t$ the r.v. $\xi_t$ satisfies:*

$$E_{t-1}[\xi_t] = 0. \tag{1.1.11}$$

*and the conditional variance is*

$$\sigma_t^2 = Var_{t-1}[\xi_t] = E_{t-1}[\xi_t^2] = \alpha_0 + \sum_{j=1}^{p} \alpha_j \xi_{t-j}^2 \tag{1.1.12}$$

Note that the variance depends non trivially on the $\sigma$-field generated by the past observations: $\xi_{t-1}, \xi_{t-2}, \ldots$, hence large observations in recent past strongly influence the volatility.
The Generalized version of this process is called GARCH:

**Definition 1.1.12.** *The process $\{\xi_t, t \in T\}$ follows a GARCH model if for every $t$ the r.v. $\xi_t$ satisfies:*

$$E_{t-1}[\xi_t] = 0. \tag{1.1.13}$$

*and the conditional variance*

$$\sigma_t^2 = \alpha_0 + \sum_{j=1}^{p} \alpha_j \xi_{t-j}^2 + \sum_{k=1}^{q} \beta_k \sigma_{t-k}^2 \tag{1.1.14}$$

Note that the variance depends again non trivially on the $\sigma$-field generated by the past observations: $\xi_{t-1}, \xi_{t-2}, \ldots$. Here, $\alpha_j > 0, \beta_k > 0$, so that more persistence with respect to the ARCH model is built into the volatility.

Let $X_t$ denote the stochastic process of interest with conditional mean $\mu_t = E_{t-1}[X_t]$ and GARCH conditional volatility; the equation that describes this process is:

$$\begin{cases} X_t = \mu_t + \xi_t, \\ \xi_t = \sigma_t Z_t \end{cases} \tag{1.1.15}$$

for $t \in T$. Here the conditional mean $\mu_t = E[X_t | \mathfrak{F}_{t-1}]$ and the volatility $\sigma_t = E[(X_t - \mu_t)^2 | \mathfrak{F}_{t-1}]^{1/2}$ are measurable functions of the filtration $\mathfrak{F}_{t-1}$ generated by $\mathbf{X}_{t-1} = \{X_s, s \leq t-1, s \in T\}$, i.e., of the history of the process up to time $t-1$, while $Z_t$ is assumed to be independent on $\mathfrak{F}_{t-1}$. The terms $\xi_t$ are usually referred as the *innovation terms* of the process.

Implementation of these models through non-Gaussian distributed residuals have been proposed (see, e.g., (McNeil, Frey, & Embrechts, 2005)) to account for the time-varying effect. Moreover, a variety of alternative definitions of ARCH and GARCH models have been proposed and applied (see, e.g., Palm (1996) for a large list of related models).
In this thesis the estimation of GARCH(1,1) parameters is done using maximum-likelihood estimation.

Concerning the multivariate setting, many generalizations of the previous models have been defined and applied in financial contexts. Engle and Kraft (1983) introduced the basic M-GARCH (Multivariate Generalized AutoRegressive Conditional Heteroskedasticity) model for a vector $\mathbf{X}_t = (X_{1,t}, X_{2,t}, \ldots, X_{n,t})'$ of log-returns as

$$\begin{cases} \mathbf{X}_t = \mu_t + \xi_t, \\ \xi_t = \Sigma_t^{1/2} \mathbf{z}_t, \end{cases} \tag{1.1.16}$$

for $t \in T$. Here $\mu_t$ denotes the $(n, 1)$ vector of conditional means, $\Sigma_t$ denotes the $(n, n)$ conditional covariance matrix of the innovation term $\xi_t$, and the standardized vector $\mathbf{z}_t = (z_{1,t}, \ldots, z_{n,t})$ of residual term comes from a sequence of iid vectors with mean $E[\mathbf{z}_t] = 0$ and variance-covariance matrix $V[\mathbf{z}_t] = I_n$ (commonly, $\mathbf{z}_t$ is assumed to be normally distributed). Here, $\Sigma_t^{1/2}$ is any $n \times n$ positive definite matrix such that $\Sigma_t$ is the conditional covariance matrix of $\mathbf{X}_t$, e.g., it may be obtained by the Cholesky decomposition of $\Sigma_t$. Starting from this model, a long list of multivariate ARCH or GARCH models have been proposed, having the specification of $\Sigma_t$ as the main difference. These models allow for inclusion of asymmetric and time-varying effects. For a comprehensive survey on multivariate GARCH models see, e.g., Bauwens, Laurent, and Rombouts (2006), or Jondeau, Poon, and Rockinger (2007). In particular, non-Gaussian assumptions for $\mathbf{z}_t$ can be considered, in order to capture asymmetry or dependency in the tails of returns distributions.

## 1.2 Dependence in finance

Copulae are the typical tool to model dependence between random variables. There are many copulae families, in this subsection we focus on the most used in finance: elliptical and Archimedean copulae. For details please refer to Nelsen (2006) and Genest, Quesada Molina, and Rodriguez Lallena (1995). In both cases the parameters are strictly connected to the dependence, we present the definition, the link with dependence and the conditional distribution.

Cherubini, Luciano, and Vecchiato (2004) focus on a main fact: linear correlation, which represents the standard tool used in risk management units to measure the co-movement of markets, may turn out to be a flawed instrument in the presence of a non-normal return.
Linear correlation between the returns $X_{i,t}$ and $X_{j,t}$ of assets $i$ and $j$ may be written as

$$
\begin{aligned}
corr(X_{i,t}, X_{j,t}) = \frac{cov(X_{i,t}, X_{j,t})}{\sigma_{i,t}\sigma_{j,t}} = \frac{1}{\sigma_{i,t}\sigma_{j,t}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [F_{i,j}(x,y) \\
- F_i(x)F_j(y)\; dx\; dy
\end{aligned}
\tag{1.2.1}
$$

where $\sigma_{i,t}$ and $\sigma_{j,t}$ represent volatility. Notice that the correlation depends on the marginal distributions of the returns. Its maximum value can be computed by substituting the upper Fréchet bound (Fréchet, 1935 and Fréchet, 1951) in the formula

$$
\begin{aligned}
corr_{max}(X_{i,t}, X_{j,t}) = \frac{1}{\sigma_{i,t}\sigma_{j,t}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [min\,\{F_i(x), F_j(y)\} \\
- F_i(x)F_j(y)]\; dx\; dy.
\end{aligned}
\tag{1.2.2}
$$

Furthermore, the value corresponding to perfect negative correlation can be obtained by substituting the Fréchet lower bound

$$
\begin{aligned}
corr_{min}(X_{i,t}, X_{j,t}) = \frac{1}{\sigma_{i,t}\sigma_{j,t}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [max\,\{F_i(x) + F_j(y) - 1, 0\} \\
- F_i(x)F_j(y)]\; dx\; dy.
\end{aligned}
\tag{1.2.3}
$$

Note that the intuition on values +1 and - 1 of the correlation determined by these bounds is false or at least is not true in general.
Of course, that is what happens when returns are normal and the result also holds in the more general case of elliptic distributions. However, this does not hold for other arbitrary choices for the dependencies. Looking at the problem from a different viewpoint, correlation is an effective way to represent co-movements between variables if they are linked by linear relationships, but it may be severely flawed in the presence of nonlinear links. Readers may check this in the simple case of a variable z normally distributed and $z^2$ which is obviously perfectly correlated with the first one, but has a chi-squared distribution.
Hence, using linear correlation to measure the co–movements of markets in the presence of non–linear relationships may become misleading since it does not

cover the whole range from $-1$ to +1 while two markets can be moved by the same factor, becoming perfectly dependent.
Hence, in the Section 1.2.2, we introduce other measures of dependence.

### 1.2.1 Copulae - Definition and basic properties

It is possible to have assets with different distributions, eventually non-parametric, and with specific tail dependence characteristic or positive dependence. In these cases it is difficult to build the multivariate model of the overall portfolio without split marginal specific characteristics from dependence structure.
Copula functions may be of great help to address these problems. As we will see, copula functions provide a way to represent the dependence structure between different random variables, ignoring the marginal distribution of each of them. For this reason they represent a useful instrument to describe market co-movements in a world in which the marginal distributions of returns are not normal and change with the assets.
This subsection firstly focus on two-dimensional copulas and their characteristics and properties, in next sections we present multivariate cases.

In order to define the copula function, we need some preliminary definitions. We use the notation of $\overline{\mathbb{R}} = [-\infty, +\infty]$ and of $\overline{\mathbb{R}}^2 = \overline{\mathbb{R}} \times \overline{\mathbb{R}}$ and $\mathbf{I}^2 = \mathbf{I} \times \mathbf{I}$ with $\mathbf{I} = [0, 1]$.

**Definition 1.2.1.** *Let $A_1$ and $A_2$ be two non-empty subsets of $\overline{\mathbb{R}}$ and let $H : A_1 \times A_2 \to \mathbb{R}$ be a function. Given $B = [x_1, x_2] \times [y_1, y_2]$ a rectangle whose vertices are in DomH. Then the $H - volume$ of $B$ is*

$$V_H(B) = H(x_2, y_2) - H(x_2, y_1) - H(x_1, y_2) + H(x_1, y_1)$$

**Definition 1.2.2.** *A 2-place real function $H$ is 2-increasing if $V_H(B) \geq 0$ for all $B = [x_1, x_2] \times [y_1, y_2]$ whose vertices lie in DomH.*

Note that the 2-increasing definition does not imply nor is implied by the fact that $H$ is non-decreasing in each argument. In order to have a similar relation, we need to introduce a new definition.

**Definition 1.2.3.** *Given $A_1$ and $A_2$ two non-empty subsets of $\overline{\mathbb{R}}$, with a least an element $s_1$ of $A_1$ and $s_2$ of $A_2$, the function $H : A_1 \times A_2 \to \mathbb{R}$ is grounded if $H(x, s_2) = 0 = H(s_1, y)$ for all $(x, y)$ in $A_1 \times A_2$.*

**Definition 1.2.4.** *Let $A_1$ and $A_2$ be two non-empty subsets of $\overline{\mathbb{R}}$, let $b_1$ be the bigger element of $A_1$ and let $b_2$ be the bigger element of $A_2$. The function $H : A_1 \times A_2 \to \mathbb{R}$ admits marginals $F$ and $G$ defined:*

$$\begin{aligned} DomF = A_1 \quad &and \quad F(x) = H(x, b_2) \qquad \forall\, x \in A_1 \\ DomG = A_2 \quad &and \quad G(y) = H(b_1, y) \qquad \forall\, y \in A_2. \end{aligned}$$

We have, now, all the instruments to define the bi-variate copula function.

**Definition 1.2.5.** *A two-dimensional copula $C : \mathbf{I}^2 \to \mathbf{I}$ is a function that has the following properties:*

- *for every $u, v \in \mathbf{I}$*

$$C \textit{ is grounded}\colon C(u,0) = 0 = C(0,v) \tag{1.2.4a}$$
$$C(u,1) = u \quad \textit{and} \quad C(1,v) = v; \tag{1.2.4b}$$

- *$C$ is 2-increasing: for every $u_1, u_2, v_1, v_2 \in \mathbf{I}$ such that $u_1 \leq u_2$ and $v_1 \leq v_2$*

$$C(u_2,v_2) - C(u_2,v_1) - C(u_1,v_2) + C(u_1,v_1) \geq 0 \tag{1.2.5}$$

We now introduce two important functions. They represent, in $\mathbf{I}^2$, the copula boundaries:

**Theorem 1.2.1.** *Given a copula C, for every $(u, v) \in \mathbf{I}^2$*

$$max(u+v-1,0) \leq C(u,v) \leq min(u,v) \tag{1.2.6}$$

The functions $min(u, v)$ and $max(u + v - 1, 0)$ are two copulas. $M = min(u, v)$ and $W = max(u + v - 1, 0)$ are called *upper bound* and *lower bound of Fréchet Hoeffding*.

From the definition of copula and Theorem 1.2.1, we have that any copula is a continuous surface in the unit cube $\mathbf{I}^3$, whose boundary is the skew quadrilateral with vertices (0,0,0), (1,0,0), (1,1,1), (0,1,0) and that lies between the Fréchet-Hoeffding bounds. Another important copula is the copula product $\Pi(u, v) = uv$. In Figure 1.2.1 we can see these 3 different copulas:

**Corollary 1.2.2.** *Let $C$ be a copula and $a$ any number in $\mathbf{I}$. The horizontal section of C at a is the function from $\mathbf{I}$ to $\mathbf{I}$ given by $t \to C(t,a)$; the vertical section of C at a is the function given by $t \to C(a,t)$; and the diagonal section of C is defined as $\delta_C(t) = C(t,t)$. All of these functions are non-decreasing and uniformly continuous on $\mathbf{I}$.*

As far as the the level set $\{(u, v) \in \mathbf{I}^2 : C(u, v) = K\}$ is concerned, we observe that for $K$ fixed, the level set lies in the triangle described by the Fréchet-Hoeffding bounds: $\{(u, v) : max(u + v - 1, 0) = K\}$ and $\{(u, v) : min(u, v) = K\}$.

This type of representation is called *contour diagram.*

We have introduced copulas as "simple" functions without reference to probability theory. But it is this the field where they assume a very important role: they determine the relationship between multivariate distribution functions and their univariate marginals.

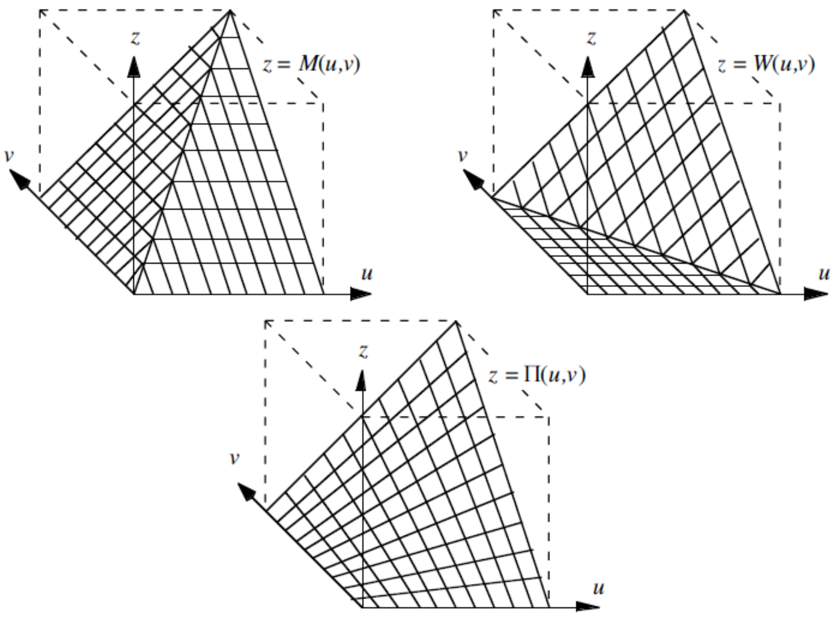


Figure 1.1: Graphs of the copulas $M$, $\Pi$ and $W$

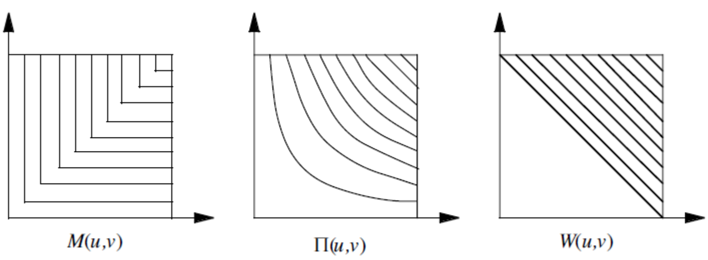


Figure 1.2: Contourn diagrams of the copulas $M$, $\Pi$ and $W$

Let $X$ and $Y$ be two random variables with marginal distributions $F(x) = P(X \leq x)$ and $G(y) = P(Y \leq y)$ respectively and joint distribution $H(x, y) = P(X \leq x, Y \leq y)$. We can associate to every $(x, y)$ three numbers: $F(x), G(y)$ and $H(x, y)$. Each of these numbers lies in $\mathbf{I}$. In other words, each $(x, y)$ is associated with $(F(x), G(y))$, and this pair corresponds to a number $H(x, y)$. We will show that this correspondence, which assigns the value of the joint distribution function to each ordered pair of values of the individual distribution functions, is indeed a function, and this function corresponds to a copula.

We recall the principal properties of distribution functions.

**Proposition 1.2.1.** *Let $X$ be a random variable and $F(x) = P(X \leq x)$ its distribution function. $F$ is such that:*

- *$F$ is non-decreasing;*
- $lim_{x \to +\infty} F(x) = 1$*;*
- $lim_{x \to -\infty} F(x) = 0$*;*
- *$F$ is right continuous.*

**Proposition 1.2.2.** *Let $X$ and $Y$ be two random variables with joint distribution function $H(x, y) = P(X \leq x, Y \leq y)$. $H$ is such that:*

- *$H$ is two-increasing;*
- $\lim_{x \to -\infty} H(x, y) = \lim_{y \to -\infty} H(x, y) = 0$*;*
- $\lim_{(x,y) \to (+\infty, +\infty)} H(x, y) = 1$*;*
- *$H$ is right continuous.*

**Remark 1.2.1.** *From Proposition 1.2.2 we get that $H$ is grounded. Since its domain is $\overline{\mathbb{R}}^2$, it has marginals $F(x) = H(x, +\infty)$ and $G(y) = H(+\infty, y)$.*

The following theorem, proved in Sklar (1959), is essential to define the relationship between copulas and distribution functions. Sklar also was the first one to use the name copula.

**Theorem 1.2.3** (Sklar's theorem)**.** *Let $H$ be a joint distribution function with margins $F$ and $G$. Then there exists a copula $C$ such that for all $x, y \in \overline{\mathbb{R}}^2$,*

$$H(x, y) = C(F(x), G(y)). \tag{1.2.7}$$

*If $F$ and $G$ are continuous, then $C$ is unique; otherwise, $C$ is uniquely determined on $RanF \times RanG$. Conversely, if $C$ is a copula and $F$ and $G$ are distribution functions, then the function $H$ defined by (1.2.7) is a joint distribution function with margins $F$ and $G$.*

Equation 1.2.7 gives an expression for joint distribution functions in terms of a copula and two univariate distribution functions. That is why copula is called *dependence function*. This equation can be inverted to express copulas in terms of a joint distribution function and the inverses of the two margins. Since we cannot we invert a margin if it is not strictly increasing, we introduce the definition of "quasi-inverses".

**Definition 1.2.6.** *Let $F$ be a distribution function. The quasi-inverse of $F$ is any function $F^{(-1)}$ with domain $\mathbf{I}$ such that*

- *if $t \in RanF$, then $F^{(-1)}(t)$ corresponds to any number $x \in \overline{\mathbb{R}}$ such that $F(F^{(-1)}(t)) = F(x) = t$;*
- *if $t \notin RanF$, then $F^{(-1)}(t) = \inf\{x : F(x) \geq t\} = \sup\{x : F(x) \leq t\}$.*

*If $F$ is strictly increasing, then it has a single quasi-inverse, which is of course the ordinary inverse, for which we use the customary notation $F^{-1}$.*

**Corollary 1.2.4.** *Let $H$ be a joint distribution function with $F$ and $G$ continuous marginals. Then there exists a unique copula $C$ such that for every $(u, v) \in \mathbf{I}^2$*

$$H(x, y) = C(F^{-1}(x), G^{-1}(y)) \tag{1.2.8}$$

**Proposition 1.2.3.** *Let $X$ and $Y$ be random variables with joint distribution function $H$ and marginals $F$ and $G$. Then the copula associated with $H$ is the distribution function $C$ of $(F(X), G(Y))$.*

**Remark 1.2.2.** *$F(X)$ and $G(Y)$ are standard uniform random variables. The proof is straightforward: if we call $U = F(X)$, then*

$$P(U \leq u) = P(F(X) \leq u) = P(X \leq F^{-1}(u)) = F(F^{-1}(u)) = u. \tag{1.2.9}$$

*Hence the copula can be considered a joint distribution function of standard uniform random variables.*

**Theorem 1.2.5.** *Let $X$ and $Y$ be continuous random variables. Then $X$ and $Y$ are independent if and only if $C_{X,Y} = uv$.*

*Proof.* The proof follows from the fact that $X$ and $Y$ are independent if and only if $H(x, y) = F(x)G(y)$. □

Having understood the usefulness of the copulas, we can rethink to the nonlinearity problem of assets dependence in terms of copulas. However their use is often complex and statistical indexes can be a good alternative method to investigate the presence of dependence. Hence we consider some non-parametric dependence measures, the Concordance measures.

### 1.2.2 Concordance measures

Generally speaking, two random variables $X$ and $Y$ are "concordant" if for high values of one correspond high values of the other and, viceversa, if for small values of one correspond small values of the other.
Formally we identify the concordance measure between $X$ and $Y$ with $M_{X,Y}$ or, if $C$ is the copula describing their dependence, with $M_C$.
The concordance measure has the following properties:

**Definition 1.2.7.** *$M_{X,Y} = M_C$ is the concordance measure between random variables $X$ and $Y$ with copula $C$ if and only if:*

- *is defined for each couple of random variables;*
- *is a standardized measure, $M_X, Y \in [-1, 1]$;*
- *is symmetric, $M_{X,Y} = M_{Y,X}$;*
- *if $X$ and $Y$ are independent, then $M_{X,Y} = 0$;*
- *$M_{-X,Y} = M_{X,-Y} = -M_{X,Y}$;*
- *$M_{X,Y}$ converges when the copula $C$ converges point-wise, if $\{(X_n, Y_n)\}$ is a series of continuous random variables with copula $C_n$ such that $\forall(u, v) \in \boldsymbol{I}^2$*

$$lim_{n\to\infty} C_n(u, v) = C(u, v)$$

*then*

$$lim_{n\to\infty} M_{X_n,Y_n} = M_{X,Y}.$$

This definition implies invariance among increase transformation and the existence of limits for $M$ in case of comonotonicity or contromonotonicity.

**Theorem 1.2.6.** *Let $\pi$ and $\psi$ two strictly increasing functions respectively on $RanF$ and $RanG$, then $M_{X,Y} = M_{\pi(X),\psi(Y)}$.*

**Theorem 1.2.7.** *If $X$ and $Y$ are comonotone, then $M_{X,Y} = 1$; if they are contromonotone, then $M_{X,Y} = -1$.*

It is important to notice that independence is a sufficient, but not necessary, condition of $M = 0$.

The most important concordance measures are $\tau$ of Kendall and $\rho$ of Spearman.

**Definition 1.2.8.** ***Kendall's*** *$\tau$ of a couple of random variables $(X, Y)$ is defined as:*

$$\tau(X,Y) = P\left[(X_1 - X_2)(Y_1 - Y_2) > 0\right] - P\left[(X_1 - X_2)(Y_1 - Y_2) < 0\right] \tag{1.2.10}$$

*where $(X_1, Y_1)$ and $(X_2, Y_2)$ are couples of independent and identically distributed random variables with the same joint distribution $H$ of $(X, Y)$.*

The following theorem underlines the link between $\tau(X, Y)$ and copula function $C$ of $(X, Y)$.

**Theorem 1.2.8.** *Let $X$ and $Y$ be two continuous random variables with copula $C$. Then Kendall's $\tau$ is:*

$$\tau(X,Y) = 4\int\int_{[0,1]^2} C(u,v)dC(u,v) - 1. \tag{1.2.11}$$

The integral of Equation 1.2.11 can be seen as the expectation of $C(U, V)$, where $U$ and $V$ are standard uniform with joint distribution $C(U, V)$. Then we can rewrite (1.2.11) as:

$$\tau(X,Y) = 4E[C(U,V)] - 1. \tag{1.2.12}$$

**Definition 1.2.9.** *Spearman's $\rho$ of a couple of random variables $(X, Y)$ is defined as:*

$$\rho_S(X,Y) = 3\ \left(P\left[(X_1 - X_2)(Y_1 - Y_3) > 0\right] - P\left[(X_1 - X_2)(Y_1 - Y_3) < 0\right]\right) \tag{1.2.13}$$

*where $(X_1, Y_1)$, $(X_2, Y_2)$ and $(X_3, Y_3)$ are couples of independent and identically distributed random variables with the same joint distribution $H$ of $(X, Y)$.*

In line with Theorem 1.2.8, it follows the theorem that describe the link between Spearman's $\rho_S$ and $(X, Y)$ copula $C$:

**Theorem 1.2.9.** *Let $X$ and $Y$ two continuous random variables with copula $C$. Then Spearman's $\rho_S$ is:*

$$\rho_S(X,Y) = 12\int\int_{[0,1]^2} uv dC(u,v) - 3 = 12\int\int_{[0,1]^2} C(u,v) du\, dv - 3. \tag{1.2.14}$$

Let $F$ and $G$ the marginal distribution respectively of $X$ and $Y$, defining $U = F(X)$ and $V = G(Y)$ the (1.2.14) becomes:

$$\begin{aligned} \rho_S(X,Y) &= 12\int\int_{[0,1]^2} uv dC(u,v) - 3 = 12E[UV] - 3 = \\ &= \frac{E[UV] - 1/4}{1/12} = \frac{E[UV] - E[U]E[V]}{\sqrt{Var(U)}\sqrt{Var(V)}} = \\ &= \frac{Cov(U,V)}{\sqrt{Var(U)}\sqrt{Var(V)}} = \rho(F(X), G(Y)). \end{aligned} \tag{1.2.15}$$

This means that Spearman's $\rho_S$ corresponds to the linear correlation coefficient $\rho$ of $F(X)$ and $G(Y)$: Spearman's $\rho_S$ is called *Rank correlation coefficient* because is the correlation coefficient of integral transformation of $X$ and $Y$. Also Kendall's $\tau$ is a *Rank correlation coefficient.*

Having understood the usefulness of the copulas, we can rethink to the non-linearity problem of assets dependence in terms of copulas. However their use is often complex and statistical indexes can be a good alternative method to investigate the presence of dependence. Hence we consider here non-parametric dependence measures, such as Spearman's $\rho_S$ and Kendall's $\tau$ . The non-parametric feature of these measures implies that they do not depend on the marginal probability distributions.

These measures are directly linked to the copula function. We recall the following relationships:

$$\rho_S = 12 \int_0^1 C(u,v)du\, dv - 3 \tag{1.2.16}$$

$$\tau = 4 \int_0^1 C(u,v)dC(u,v) - 1. \tag{1.2.17}$$

Notice that the specific shape of the marginal probability distributions does not enter in these relationships. Furthermore, it may be proved that substituting the maximum and minimum copulas in these equations gives values of $-1$ and $+1$ respectively.

Differently from the linear correlation measure, then, if the two variables are perfectly dependent we get $+1$ for Spearman's $\rho_S$ and Kendall's $\tau$ , while a score -1 corresponds to perfect negative dependence.

The relationship between non–parametric dependence measures and copula functions can also be applied to recover a first calibration technique of the copula function itself. In some cases the relationship between these non-parametric statistics and the parameters of the copula function may become particularly easy.

### 1.2.3 Tail dependence

The departure from normality in a multivariate system and the need to represent the co-movement of markets as closely as possible raises a second dimension of the problem. We know that non-normality at the univariate level is associated with skewness and leptokurtosis phenomena. This is known as the fat-tail problem. In a multivariate setting, the fat tail problem can be referred both to the marginal univariate distributions or to the joint probability of movements in a large market. This last instance is called tail dependence. Intuitively, we may conceive markets in which the marginal distributions are endowed with fat tails, but extreme market movements are orthogonal, or cases in which the returns on each market are normally distributed, but large market movements are likely to occur together. The use of copula functions enables us to model these two features, fat tails and tail

dependence, separately.
To represent tail dependence we consider the likelihood that one event with probability lower than $q$ occurs in first variable $X_i$, given that an event with probability lower than $q$ occurs in the second one ($X_j$). Concretely, we compute the probability to observe, for example, a crash with probability lower than $q = 1\%$ in the Asset $i$, given that a crash with probability lower than $1\%$ has occurred in the Asset $j$:

$$\lambda(q) = P(F_i(X_i) \leq q | F_j(X_j) \leq q) = \frac{P(F_i(X_i) \leq q, F_j(X_j) \leq q)}{P(F_j(X_j) \leq q)} = \frac{C(q,q)}{q} \tag{1.2.18}$$

where $F_i$ and $F_j$ are cdf respectively of $X_i$ and $X_j$.

If we compute this dependence measure far in the lower tail, that is, for very small values of $q$, we obtain the so-called tail index, in particular the lower tail index:

$$\lambda_L = lim_{q \to 0} P(F_i(X_i) \leq q | F_j(X_j) \leq q) \tag{1.2.19}$$

and the upper tail index:

$$\lambda_U = lim_{q \to 1} P(F_i(X_i) > q | F_j(X_j) > q) \tag{1.2.20}$$

### 1.2.4 Copulae in Finance

There exist many copulae families but in this subsection we focus only on the most used in finance: elliptical and Archimedean copulae.
For each family, we give the copula definition, its density and its conditional distribution.
In both cases the parameters are strictly connected to the dependence properties and, whenever possible, we focus on the relationship between these parameters and the measures of concordance or tail dependence defined above.
For further details please refer to Nelsen (2006).
Elliptical copulas are simpler to deal with, they are defined through (1.2.7) for an elliptical distribution function $F$, here we refers to the most used.

**Gaussian copula**

**Definition 1.2.10.** *The Gaussian copula is defined as follows:*

$$C_\rho^{Ga}(u,v) = \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v)) \tag{1.2.21}$$

*where $\Phi_\rho$ is the bivariate joint distribution of a standard normal vector of dimension 2, with linear correlation $\rho$ and $\Phi$ is the standard normal distribution function.*

Therefore:

$$C_\rho^{Ga}(u,v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{\frac{s^2 - 2\rho st + t^2}{2(1-\rho^2)}\right\} ds\, dt. \tag{1.2.22}$$

An alternative representation of (1.2.22) (demonstrated by (Roncalli, 2002)) is:

$$C_\rho^{Ga}(u,v) = \int_0^u \Phi\left(\frac{\Phi^{-1}(v) - \rho\,\Phi^{-1}(t)}{\sqrt{1-\rho^2}}\right) dt. \tag{1.2.23}$$

This representation is usefull to calculate the conditional distribution function via copula. The $C_{v|u}$ corresponds to $\frac{\partial C(u,v)}{\partial u}$, hence we have:

$$C_\rho^{Ga}(v|u) = \Phi\left(\frac{\Phi^{-1}(v) - \rho\,\Phi^{-1}(u)}{\sqrt{1-\rho^2}}\right). \tag{1.2.24}$$

This copula may generate the Gaussian bivariate joint distribution function. Specifically, we have the following:

**Proposition 1.2.4.** *The Gaussian copula generates the joint normal standard distribution function (via Sklar's theorem) if and only if the marginals are standard normal.*

*Proof.* Observe that

$$C_\rho^{Ga}(F_1(x), F_2(y)) = \int_{-\infty}^{x} \int_{-\infty}^{y} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{\frac{s^2 - 2\rho st + t^2}{2(1-\rho^2)}\right\} ds\, dt. \tag{1.2.25}$$

if and only if $\Phi^{-1}(F_1(x)) = x$ and $\Phi^{-1}(F_2(y)) = y$, equivalent to $F_1 = F_2 = \Phi$. □

For any other choice of the marginals, the Gaussian copula is not the distribution of a standard jointly normal vector. Both in the positive and in the negative correlation cases, the same copula, together with different marginals, determines a different joint distribution.
An example is Gaussian copula with T-Student marginals.
In order to have a visual representation of the phenomenon, and more generally of the effect of "coupling" the same copula with different marginals, let us consider the example reported in Cherubini et al. (2004). The joint density functions in the following figures is obtained coupling the Gaussian copula with standard Gaussian margins (above) and with three Student's t degree of freedom (dof); we consider both the case $\rho = 0.2$, in Figure 1.3 , and $\rho = 0.9$, in Figure 1.4.

Note that marginal Student distributions increase the tail probabilities. In general, Figures 1.3 and 1.4 provide examples of the modeling flexibility obtained

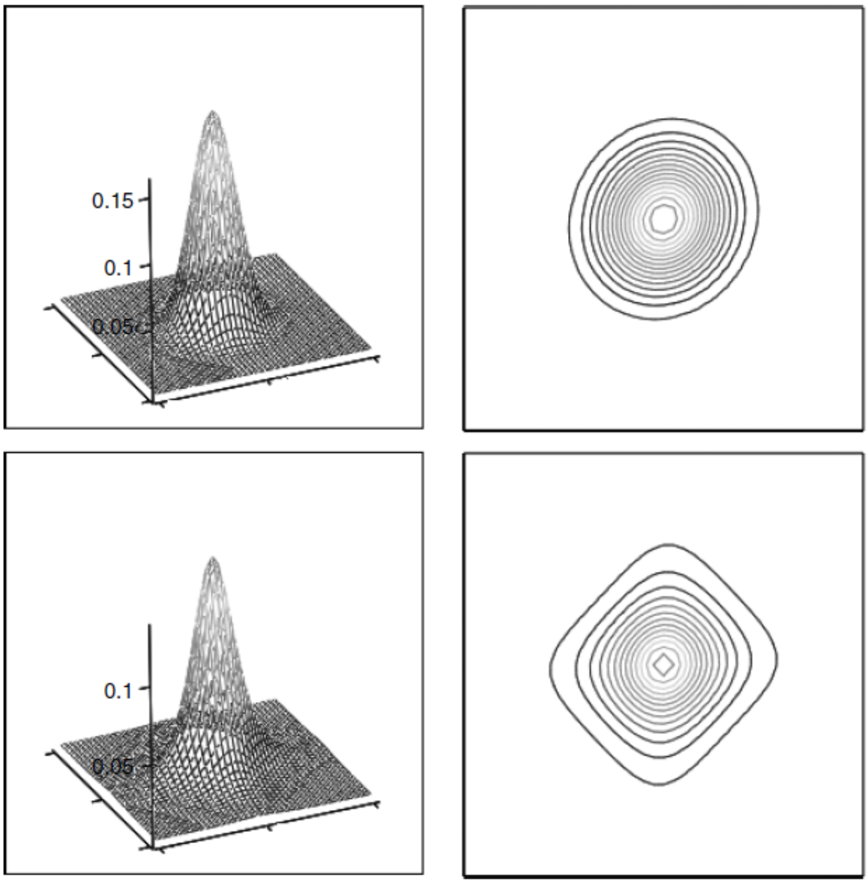


Figure 1.3: Density and level curves of the distribution obtained coupling the Gaussian copula with standard normal marginals (top) and 3-d.o.f. Student ones (bottom), $\rho = 0.2$

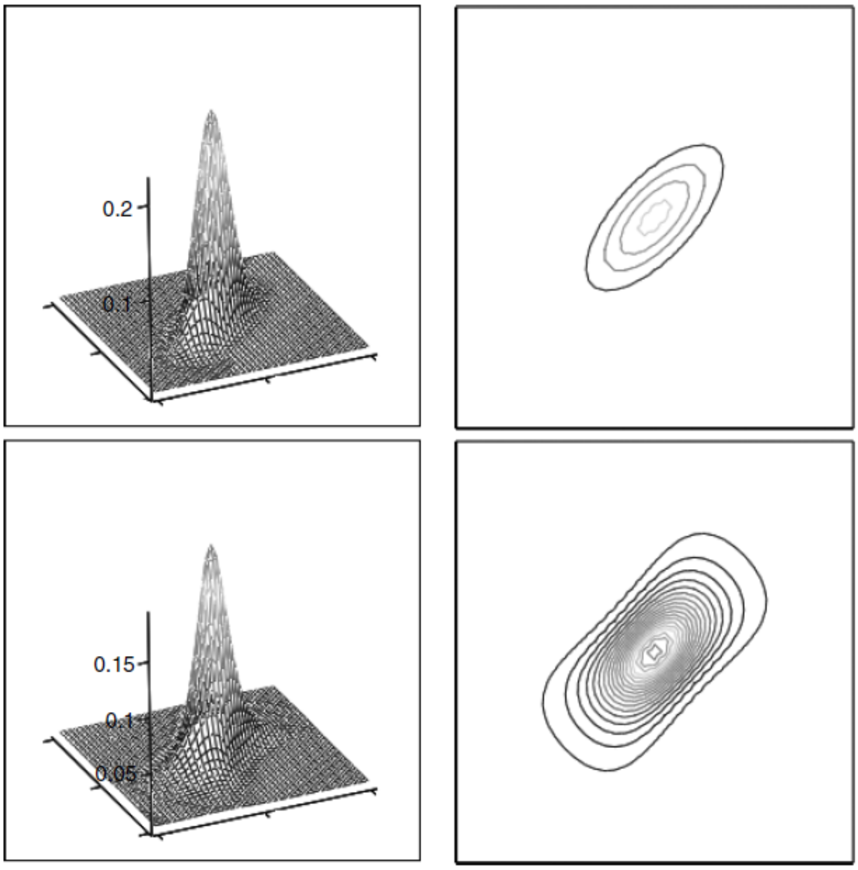


Figure 1.4: Density and level curves of the distribution obtained coupling the Gaussian copula with standard normal marginals (top) and 3-d.o.f. Student ones (bottom), $\rho = 0.9$

using copula functions instead of joint distribution functions.

Using the definition of Kendall's $\tau$ and Spearman's $\rho_S$, one can show that

$$\tau = \frac{2}{\pi} arcsin\rho \tag{1.2.26}$$

and

$$\tau = \frac{6}{\pi} arcsin\frac{1}{2}\rho \tag{1.2.27}$$

Furthermore, it can be shown that Gaussian copulas have neither upper nor lower tail dependence, unless $\rho = 1$:

$$\lambda_U = \lambda_L = \begin{cases} 0 & iff \rho < 1 \\ 1 & iff \rho = 1 \end{cases}$$

Finally, they present PQD if $\rho \geq 0$.

**T-Student copula**

**Definition 1.2.11.** *The T-Student copula is defined as follows:*

$$C^T_{\rho,\upsilon}(u, v) = t_{\rho,\upsilon}(t_\upsilon^{-1}(u), t_\upsilon^{-1}(v)) \tag{1.2.28}$$

*where $t_{\rho,\upsilon}$ is the bivariate joint distribution of a T-Student vector of dimension $2$ and dof $\upsilon$, with linear correlation $\rho$. Furthermore, $t_\upsilon$ is the univariate T-Student (with $\upsilon$ dof) distribution function.*

Therefore:

$$C^T_{\rho,\upsilon}(u, v) = \int_{-\infty}^{t_\upsilon^{-1}(u)} \int_{-\infty}^{t_\upsilon^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \left(1 + \frac{s^2 - 2\rho st + t^2}{\upsilon(1-\rho^2)}\right)^{-\frac{\upsilon+2}{2}} ds\ dt. \tag{1.2.29}$$

When the number of degrees of freedom diverges, the copula converges to the Gaussian one.
If $\upsilon > 2$, each margin admits a (finite) variance $\frac{\upsilon}{(\upsilon 2)}$, and  can be interpreted as a linear correlation coefficient.

T-Student copula has no close form link between $\rho$ and Kendall $\tau$ or Spearman $\rho_S$.

The conditional distribution function via copula is ((Roncalli, 2002)):

$$C^T_{\rho,\upsilon}(v|u) = t_{\upsilon+1}\left(\sqrt{\frac{\upsilon+1}{\upsilon + t_\upsilon^{-1}(u)^2}} \frac{t_\upsilon^{-1}(v) - \rho\, t_\upsilon^{-1}(u)}{\sqrt{1-\rho^2}}\right). \tag{1.2.30}$$

T-Student copula has tail dependence for finite $v$:

$$\lambda_U = \lambda_L = \begin{cases} > 0 & \text{if f}\rho > -1 \\ 0 & \text{if f}\rho = 1 \end{cases}$$

They present PQD if $\rho \geq 0$.

**Archimedean copulae**
Differently from Gaussian and T-Student copulas, the Archimedean copulas are not derived from bivariate distributions through the Sklar theorem.
The construction of such copulas is based on the possibility of separating the joint distribution H from the two marginals F and G through a function $\lambda(t)$, positive in the interval **I**: $\lambda(H(x,y)) = \lambda(F(x))\lambda(G(y))$. If we replace with $\varphi(t) = -ln(\lambda(t))$ the relation becomes $\varphi(H(x,y)) = \varphi(F(x)) + \varphi(G(y))$ in terms of copulae we get

$$\varphi(C(u,v)) = \varphi(u) + \varphi(v). \tag{1.2.31}$$

In order to proceed further, we need to define the pseudo-inverse function of $\varphi$.

**Definition 1.2.12.** *Let $\varphi$ a function $\varphi : \boldsymbol{I} \to [0, +\infty]$ continuous and strictly decreasing with $\varphi(1) = 0$ and $\varphi(0) \leq +\infty$. The $\varphi$ pseudo-inverse is the function $\varphi^{[-1]}$ with Dom$\varphi^{[-1]} = [0, +\infty]$ and Ran$\varphi^{[-1]} = \boldsymbol{I}$ given by:*

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1} & 0 \leq t \leq \varphi(0) \\ 0 & \varphi(0) \leq t \leq +\infty \end{cases} \tag{1.2.32}$$

**Observation 1.** *Note that the function $\varphi^{[-1]}$ is continuous, non-decreasing on $[0, +\infty]$ and strictly increasing on $[0, \varphi(0)]$. Furthermore $\varphi^{[-1]}(\varphi(u)) = u$ on $\boldsymbol{I}$, and*

$$\varphi(\varphi^{[-1]}(t)) = \begin{cases} t & 0 \leq t \leq \varphi(0) \\ \varphi(0) & \varphi(0) \leq t \leq +\infty \end{cases}$$

*Finally, if $\varphi(0) = +\infty$, then $\varphi^{[-1]} = \varphi^{-1}$.*

Having defined the pseudo-inverse of function $\varphi$, we can reverse the previous Equation 1.2.31 to get:

$$C(u,v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)).$$

**Theorem 1.2.10.** *Let $\varphi : \boldsymbol{I} \to [0, +\infty]$ a continuous, strictly decreasing function,such that $\varphi(1) = 0$ and let $\varphi^{[-1]}$ the pseudo-inverse of $\varphi$ defined in (1.2.12). Then the function $C : \boldsymbol{I}^2 \to \boldsymbol{I}$ defined as:*

$$C(u,v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)) \tag{1.2.33}$$

*is a copula if and only if $\varphi$ is convex.*

**Definition 1.2.13.** *The bivariate Archimedean copulae are defined as:*

$$C^A(u,v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)) \tag{1.2.34}$$

*where $\varphi$ is called generator. If $\varphi(0) = +\infty$ then $\varphi^{[-1]} = \varphi^{-1}$ and $\varphi$ is called strict generator, generating a strict Archimedean copula.*

**Theorem 1.2.11.** *Let $C^A$ an Archimedean copula with generator $\varphi$. Then:*

- *$C^A$ is simmetric: $C^A(u,v) = C^A(v,u) \forall u, v \in \boldsymbol{I}$;*
- *$C^A$ is associative: $C^A(C^A(u,v), w) = C^A(u, C^A(v,w)) \forall u, v, w \in \boldsymbol{I}$;*
- *if $\varphi$ is a generator of $C^A$, then $\forall c > 0$ constant c $\varphi$ is still a generator of $C^A$.*

The Archimedean copulas are linked to the association measures introduced in Section 1.2.2. If $X$ and $Y$ are continuous random variables with unique Archimedean copula $C^A$ and with generator $\varphi$, then the relationship between $\varphi$ and Kendall $\tau$ is:

$$\tau = 1 + \int_0^1 \frac{\varphi(t)}{\varphi'(t)} dt. \tag{1.2.35}$$

In this thesis we consider the Archimedean copulas with 1 parameter, defined using generator $\varphi_\theta$ depending on real parameter $\theta$. Here a table of the most relevant Archimedean copulas in finance and their properties.

Table 1.1: Generator function of some 1-parameter Archimedean copulas

| **Name** | **Generator function** |
|---|---|
| Gumbel | $\varphi_\theta(t) = (ln(t))^\theta, \theta \in [1, +\infty)$ |
| Clayton | $\varphi_\theta(t) = \frac{(t^{-\theta}-1)}{\theta}, \theta \in [-1, 0) \cup (0, +\infty)$ |
| Frank | $\varphi_\theta(t) = -ln\frac{exp(-\theta t)-1}{exp(-\theta)-1}, \theta \in (-\infty, 0), \cup(0, +\infty)$ |

Table 1.2: Copula function of some 1-parameter Archimedean copulas

| **Name** | **Copula** |
|---|---|
| Gumbel | $C_\theta^{Gu}(u,v) = e^{-[(-ln\ u)^\theta + (-ln\ v)^\theta]^{\frac{1}{\theta}}}$ |
| Clayton | $C_\theta^{Cl}(u,v) = max[(u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}}, 0]$ |
| Frank | $C_\theta^{Fr}(u,v) = -\frac{1}{\theta} ln(1 + \frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}$ |

### 1.2.5 Multivariate Copulas

We extend here to higher dimensions, the results presented in previous paragraphs. In particular, for $n \geq 2$ the definitions of $\overline{\mathbb{R}}^n$, $\mathbf{I}^n$, groundness and $n$-increasing

Table 1.3: Conditional distribution via copula of some 1-parameter Archimedean copulas

| **Name** | **Conditional distribution via copula** |
|---|---|
| Gumbel | $C_\theta^{Gu}(v\|u) = C_\theta^{Gu}(u,v)\frac{1}{u}(-ln\ u)^{\theta-1}[(-ln\ u)^\theta(-ln\ v)^\theta]^{\frac{1}{\theta}-1}$ |
| Clayton | $C_\theta^{Cl}(v\|u) = u^{-\theta-1}[(u^{-\theta}+v^{-\theta}-1)^{-\frac{1}{\theta}-1}]$ |
| Frank | $C_\theta^{Fr}(v\|u) = \frac{(e^{-\theta u}-1)(e^{-\theta v}-1)+e^{-\theta v}-1}{(e^{-\theta u}-1)(e^{-\theta v}-1)e^{-\theta}-1}$ |

Table 1.4: Link between parameter and Kendall $\tau$ of most relevant 1-parameter Archimedean copulas

| **Name** | **$\tau$ of Kendall** |
|---|---|
| Gumbel | $\tau = 1 - \frac{1}{\theta}$ |
| Clayton | $\tau = \frac{\theta}{\theta+2}$ |
| Frank | $\tau = 1 - \frac{4}{\theta}(1 - \frac{1}{\theta}\int_0^\theta \frac{x}{e^x-1}dx)$ |

function are the straightforward extention of the previously given ones with $n = 2$. Let $H$ be a function whose domain Dom$H$ is a subset of $\overline{\mathbb{R}}^n$ and whose range Ran$H$ is a subset of $\overline{\mathbb{R}}^n$.

**Definition 1.2.14.** *Let $A_1, \dots, A_n$ be $n$ non empty subsets of $\overline{\mathbb{R}}^n$ and $H$ be a real valued function*
*$H\colon A_1 \times \cdots \times A_n \to \mathbb{R}$. $H$ is n-increasing if*

$$V_H(B) = \sum(sgn(\mathbf{c})H(\mathbf{c})) \geq 0$$

*for every $B = [\mathbf{a}, \mathbf{b}]$ rectangle $\subset A_1 \times \cdots \times A_n$. The sum is defined on every vertex $\mathbf{c}$ of $B$ and*

$$sgn(\mathbf{c}) = \begin{cases} +1 & \textit{if } c_k = a_k \textit{ an even number of times} \\ -1 & \textit{se } c_k = a_k \textit{ an odd number of times} \end{cases}$$

**Definition 1.2.15.** *Let $A_1, \dots, A_n$ be $n$ non empty subsets of $\overline{\mathbb{R}}^n$ and $H$ be a real valued function $H\colon A_1 \times \cdots \times A_n \to \mathbb{R}$. Let $a_k$ be the smaller element of every $A_k$, $k = 1, \dots, n$. $H$ is grounded if $H(\mathbf{t}) = 0$ for every $\mathbf{t} \in DomH$ such that $t_k = a_k$ for at least one $k$.*

**Definition 1.2.16.** *Let $A_1, \dots, A_n$ be $n$ non empty subsets of $\overline{\mathbb{R}}^n$ and $H$ a real valued function $H\colon A_1 \times \cdots \times A_n \to \mathbb{R}$. Let $b_k$ be the smaller elemen of every $A_k$, $k = 1, \dots, n$. Then $H$ has marginals and the univariate marginals are the functions $H_k$, with domain Dom$H_k = A_k$, defined as:*

$$H_k(x) = H(b_1, \dots, b_{k-1}, x, b_{k+1}, \dots, b_n) \qquad \forall\, x \in A_k.$$

Now, we extend the notion of copula to the multivariate case.

**Definition 1.2.17.** *A $n$-dimensional copula is a function $C\colon \boldsymbol{I}^n \to \boldsymbol{I}$ such that:*

- *for every $\boldsymbol{u} \in \boldsymbol{I}^n$*

$$C(\boldsymbol{u}) = 0 \tag{1.2.36a}$$

*if at least one of the coordinates of $\boldsymbol{u}$ is equal to 0, for every $k = 1, \ldots, n$*

$$C(1, \ldots, 1, u_k, 1, \ldots, 1) = u_k; \tag{1.2.36b}$$

- *for every $\boldsymbol{a}$, $\boldsymbol{b} \in \boldsymbol{I}$ such that $\boldsymbol{a} \leq \boldsymbol{b}$ we have that*

$$\sum(sgn(\boldsymbol{c})C(\boldsymbol{c})) \geq 0. \tag{1.2.37}$$

**Theorem 1.2.12.** *Let $C$ be a multidimensional copula. Then $C$ is uniformly continuous in its domain.*

Now, we extend the properties of the joint distribution function of $n$ variables, $n > 2$.

**Proposition 1.2.5.** *Let $\mathbf{X} = (X_1, \ldots, X_n)$ be a random vector with joint distribution function $H(\mathbf{x}) = P(X_1 \leq x_1, \ldots, X_n \leq x_n)$. Then $H$ satisfies the following properties:*

- *$H$ is $n$-increasing;*
- *$H(\boldsymbol{x}) = 0 \qquad \forall \boldsymbol{x} \in \overline{\mathbb{R}}^n$ such that $x_k = -\infty$ for at least one $k$;*
- *$\lim_{\boldsymbol{x} \to +\infty} H(\boldsymbol{x}) = 1$;*
- *$H$ is right continuous.*

We can enunciate the extension of Sklar's Theorem in $n$-dimensions.

**Theorem 1.2.13** (Sklar's theorem: $n$ dimensional version, $n > 2$)**.** *Let $\mathbf{X} = (X_1, \ldots, X_n)$ be a vector of random variables with joint distribution $H(\mathbf{x})$ and marginals $F_1(x_1), \ldots,$
$F_n(x_n)$. Then there exists a $n$-dimensional copula such that for every $\mathbf{x} \in \overline{\mathbb{R}}^n$ we have:*

$$H(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)). \tag{1.2.38}$$

*If the marginals are continuous, then $C$ is unique. Otherwise, it is uniquely defined on $RanF_1 \times \cdots \times RanF_n$. Conversely, if $C$ is a $n$-copula and $F_1(x_1), \ldots, F_n(x_n)$ are the distribution functions of $(X_1, \ldots, X_n)$, then the function $H$ defined in (1.2.38) is a joint distribution function with marginals $F_1(x_1), \ldots, F_n(x_n)$.*

**Corollary 1.2.14.** *Let $H$ be a joint distribution with univariate marginals $F_1(x_1), \ldots, F_n(x_n)$. For every $\mathbf{u} \in \mathbf{I}^n$ we get:*

$$C(u_1, \ldots, u_n) = H(F^{(-1)}(u_1), \ldots, F^{(-1)}(u_n)). \tag{1.2.39}$$

*If the marginals are strictly increasing, we have:*

$$C(u_1, \ldots, u_n) = H({F_1}^{-1}(u_1), \ldots, {F_n}^{-1}(u_n)). \tag{1.2.40}$$

**Observation 2.** *If the marginals are continuous, $C$ is unique and we can contruct it.*

**Theorem 1.2.15.** *Let $C$ be a $n$-dimensional copula. For $n > 2$, $1 < k < n$, the $k$-dimensional marginal of $C$ are $k$-dimensional copulas.*

We can also extend $M, W, \Pi$ to the n-dimensional case. There exist the extensions $M^n, W^n, \Pi^n$ defined by $M^n = \min(u_1, \ldots, u_n)$, $W^n = \max(u_1 + \cdots + u_n - n + 1, 0)$, $\Pi^n = u_1 \ldots u_n$. We observe a relevant difference between $W$ and $W^n$: $W^n$ is no more a copula for $n > 2$. However, it is again the best lower bound for a copula. We get, then, an order:

**Theorem 1.2.16.** *For every $C$ $n$-dimensional copula:*

$$W^n(\boldsymbol{u}) \leq C(\boldsymbol{u}) \leq M^n(\boldsymbol{u}) \qquad \forall \boldsymbol{u} \in \mathbf{I}^n.$$

There exists the analogous theorem concerning the role of $\Pi^n$ with independent random variables.

**Theorem 1.2.17.** *Let $X_1, \ldots, X_n$ be continuous and random variables. They are independent if and only if the copula associated is $\Pi^n$.*

**Proposition 1.2.6.** *Let $X_1, \ldots, X_n$ be continuous random variables with joint distribution function $H$ and marginals $F_1, \ldots, F_n$. The the copula $C$ associated is the joint distribution function of $(F_1(X_1), \ldots, F_n(X_n))$.*

Since $(F_1(X_1), \ldots, F_n(X_n))$ are standard uniforms, the copula associated to them corresponds to the joint distribution function of $n$ standard uniform random variables.

### 1.2.6 Impossibility Theorem

In this last short section we introduce a crucial property that determines the necessity to introduce another class of dependence functions: the Linkages.

From Sklar's theorem, we know that if $C$ is a bivariate copula and $F$ and $G$ are *univariate* distribution functions, then $C(F(x), G(y))$ is a distribution function. But what if $F$ and $G$ are *multivariate* functions? On this, an important result was proved by Genest et al. (1995) with the so called *theorem of impossibility*.

**Theorem 1.2.18** (Theorem of impossibility)**.** *Let $m$ and $n$ be two positive integer such that $m + n \geq 3$ and suppose $C$ is a bivariate copula such that $H(\mathbf{x}, \mathbf{y}) = C(F(\mathbf{x}), G(\mathbf{y}))$ is a distribution function of dimension $m + n$ with marginals $H(\mathbf{x}, +\infty) = F(\mathbf{x})$ and $H(+\infty, \mathbf{y}) = G(\mathbf{y})$ for every $m$-dimensional distribution functions $F$ and $n$-dimensional $G$. Then, $C = \Pi$.*

Namely, the only possible copula which works with multidimensional marginals is the independent one. In the next section we present an alternative instrument that can deal with this limitation.

### 1.2.7 Linkages - The standard construction

In their paper dated 1996, Li, Scarsini and Shaked (Li, Scarsini, & Shaked, 1996) define a new tool, the linkage function, that can be used for the study of multivariate distributions with given multivariate marginals. The linkage emphasizes the roles of dependence structure between given marginals, and the dependence structure within each of them, like copula, but overcomes the limitation that the copula function can't handle with multivariate marginals with rare exemptions (see Theorem 1.2.18).

Given $X_1, X_2, \dots, X_n$ univariate random variables with marginals $F_i$ and joint distribution $F$, let us consider $F_{i+1|1,2,\dots,i}(\cdot|x_1, x_2, \dots, x_i)$ the conditioned distribution of $X_{i+1}$ given $X_1 = x_1, \dots, X_i = x_i$. We should also introduce the inverse of $F_i$ and of $F_{i+1|1,2,\dots,i}(\cdot|x_1, x_2, \dots, x_i)$ for every $(x_1, \dots, x_i)$ in the support of $X_1, \dots, X_i$. The first one is defined (as we did in Definition 1.2.6) as $F_i^{-1}(u) = \sup(x : F(x) \leq u)$, $u \in [0, 1]$, and similarly for the others.

**Definition 1.2.18.** *For $i = 1, 2, \dots, k$, the transformation $\Psi : \mathbb{R}^n \to [0, 1]^n$ is defined as*

$$\Psi_F(x_1, \dots, x_n) = (F_1(x_1), F_{2|1}(x_2|x_1), \dots, F_{n|1,\dots,n-1}(x_n|x_1, \dots, x_{n-1})) \tag{1.2.41}$$

*for any $(x_1, \dots x_n)$ in the support of $(X_1, \dots, X_n)$ .*

**Lemma 1.2.19.** *Let $X_1, X_2, \dots, X_n$ be $n$ random variables as described above with an absolutely continuous joint distribution function $F$. Consider*

$$(U_1, \dots, U_n) = \Psi_F(X_1, \dots, X_n) \tag{1.2.42}$$

*Then $U_1, \dots, U_n$ are independent uniform $[0, 1]$ random variables.*

For the proof please refer to Li et al. (1996).

It is possible to invert $\Psi_F$ by induction:

$$x_1 = F^{-1}(u_1) \tag{1.2.43a}$$

$$x_i = F^{-1}_{i+1|1,2,\dots,i}(u_{i+1}|x_1, x_2, \dots, x_i), \quad i = 2, \dots, n \tag{1.2.43b}$$

The function $\Psi_F^* : [0,1]^n \to \mathbb{R}^n$ is defined as

$$\Psi_F^*(u_1, \dots, u_n) = (x_1, \dots, x_n), \quad (u_1, \dots, u_n) \in [0,1]^n \tag{1.2.44}$$

where $x_i$ are the ones defined in 1.2.43b.

Let

$$(\hat{X}_1, \hat{X}_2, \dots, \hat{X}_n) \equiv \Psi_F^*(U_1, U_2, \dots, U_n). \tag{1.2.45}$$

Then

$$(\hat{X}_1, \hat{X}_2, \dots, \hat{X}_n) =_{st} (X_1, X_2, \dots, X_n), \tag{1.2.46}$$

where $=_{st}$ means equality in law. In fact, if $F$ is absolutely continuous, then

$$\Psi_F^* \circ \Psi_F(X_1, X_2, \dots, X_n) =_{a.s.} (X_1, X_2, \dots, X_n), \tag{1.2.47}$$

where $=_{a.s.}$ means equality almost surely under the probability measure associated with $F$.

Let now $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_k$ be k random vectors of dimensions $m_1, m_2, \dots, m_k$, respectively, that can be independent or not. Let them have marginal distributions $F_1, F_2, ..., F_k$, of dimension $m_1, m_2, ..., m_k$, and let $F$ be the corresponding $\sum_{i=1}^k m_i$-dimensional joint distribution function. It is possible to associate to the random vector $(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_k)$ the above defined linkage function as follows.

Let $\Psi_{F_i} : \mathbb{R}^{m_i} \to [0,1]^{m_i}$ be defined as in (1.2.41). By using it, we can define the vectors $\mathbf{U}_i = \Psi_{F_i}(\mathbf{X}_i)$, $i = 1, 2, \dots, k$, where each $\mathbf{U}_i$ is made by $m_i$ independent uniformly distributed on $[0,1]$ variables. Notice that, since $\mathbf{X}_i$ are not necessarily independent, then the $\mathbf{U}_i$ could be not independent.
The joint distribution $L$ of

$$(\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_k) = (\Psi_{F_1}(\mathbf{X}_1), \Psi_{F_2}(\mathbf{X}_2), \dots, \Psi_{F_k}(\mathbf{X}_k)) \tag{1.2.48}$$

is called the *linkage* corresponding to $(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_k)$.

It is important to note that different distributions $F$ can have the same linkage. The linkage emphasizes the dependence between the $\mathbf{X}_i$, but does not contain any information about the dependence properties within the $\mathbf{X}_i$. These types of info are lost when we compute $\mathbf{U}_i$.

**Theorem 1.2.20.** *If $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_k$ have joint distribution $F$ and $\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_n$ have joint distribution the linkage L, then*

$$(\hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2, \dots, \hat{\mathbf{X}}_n) \equiv (\Psi_{F_1}^*(\mathbf{U}_1), \Psi_{F_1}^*(\mathbf{U}_2), \dots, \Psi_{F_1}^*(\mathbf{U}_n)) \tag{1.2.49}$$

*is such that*

$$(\hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2, \dots, \hat{\mathbf{X}}_n) =_{st} (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_k). \tag{1.2.50}$$

# Chapter 2

# Financial modeling

Forecasting financial stock returns or modeling stochastic volatilities is a very active area of research in recent decades ((Brooks, 2014), (Taylor, 1994) and (Clements, Hendry, & Mills, 2012)).
Traders, academic and risk managers have proposed several financial, statistical and econometric theories attempting to explain the features, the patterns and the dynamics of stock prices.
Modeling or forecasting financial time series remains a very difficult task due to the complexity of the phenomena driving the dynamics of market prices.
However, the composition of optimal portfolio requests the formulation of reliable and efficient forecasting models. The presence of non-linearity, skewness, fat tails, volatility clustering, leverages effects, co-movements in volatility of the existing financial returns models, from the Integrated Autoregressive Moving Average models (ARIMA) to the General Autoregressive and Conditionally Heteroskedastic (GARCH) and others stochastic volatility models including those by Bera and Higgins (1995), Bollerslev, Chou, and Kroner (1992), Bollerslev, Engle, and Nelson (1994) determines the necessity of elaborating alternative forecasting statistical methods.
Furthermore the diversification effect between assets in a portfolio of hundred of assets is often understimated in existing models. When the market is turbolent the the traders can obtain gains only with a well diversified portfolio, However, in these periods comonotonic behaviors of assets increase their role. Hence, both traders and risk managers need to forecast the rapidity of changes in the right direction of the asset allocation of a portfolio.

The purpose of this chapter is to illustrate some well known forecasting models used in Risk Management industry and to introduce other sophisticated models recently published in academic papers not yet regularly used for industry application. A preliminary part of this Chapter presents some used Risk indicators and the Back-test used both by Regulators and Risk Managers.

## 2.1 Risk Management

In financial institutions Risk Management identifies, computes, understands and monitors risks. This task is performed to maintain the risk under control.
There are several causes of risks, but one particularly relevant is the *market risk*. Market risk is the possibility for an investor to experience losses due to factors that affect the overall performance of the financial markets. It is also called "systematic risk" and cannot be eliminated through diversification, though it can be hedged against. The risk that a major natural disaster will cause a decrease in the market as a whole is an example of market risk. Other sources of market risk include recessions, political turmoil, changes in interest rates. Financial institutes created statistical models that try to compute and control market risk.

### 2.1.1 Measures of risk

To compute and monitoring risk it is useful to identify some indicators that summarize in a number the whole risk, a Risk Measure. Artzner et al. (1999) analyzed some risk measures. In the following we introduce some financial definitions that will be used in the next sections.

**Definition 2.1.1.** *A measure of risk is a mapping from a set of random variables $X_i$ (portfolio returns) to the real numbers*

$$\rho\colon L \to \overline{\mathbb{R}}$$

*that satisfies these properties:*

- *translative: if $a \in \mathbb{R}$ and $Z \in L$, then $\rho(Z + a) = \rho(Z) + a$;*
- *monotone: if $Z_1, Z_2 \in L$ and $Z_1 \leq Z_2$, then $\rho(Z_2) \leq \rho(Z_1)$;*
- *homogeneous: if $a \in \mathbb{R}$ and $Z \in L$, then $\rho(\lambda Z) = \lambda\rho(Z)$;*
- *subadditivity: $\rho(Z_1 + Z_2) \leq \rho(Z_1) + \rho(Z_2), \quad Z_1, Z_2 \in L$.*

A measure of risk that satisfies all these properties is called *coherent*.
*Modern portfolio theory* (MPT) is a theory of finance that attempts to maximize portfolio expected returns for a given amount of portfolio risk, or equivalently to minimize the risk for a given level of expected return, by carefully choosing the proportions of various assets.

*Value at risk* (VaR) is the most used measure of risk. Folklore attributes the introduction of this indicator to Dennis Weatherstone at J. P. Morgan in 1990. He was looking for a way to convey meaningful information on risk exposure to the financial institution's board but he wanted to avoiding the request of significant technical expertise to understand the presentation. However, there are evidences that it was already internally used, at least at Federal Reserve Bank of New York, from 1985.

**Definition 2.1.2.** *For a given portfolio at time t, a time horizon $T$ and a probability $\alpha$, the $VaR_\alpha$ is defined as a threshold loss value, such that the probability that the loss on the portfolio over the given time horizon exceeds this value is $\alpha$.*

Or more formally:

**Definition 2.1.3.** *Let $X_{[t,T]}$ be the variable whose fulfillment are the possible performances of the portfolio (or asset) at time t with time horizon $T$ and let $f_{X_{[t,T]}}(x)$ its density at time $T$, the VaR at time $T$ at significance level $\alpha \in (0,1)$ is the quantile of $X_{[t,T]}$:*

$$VaR_\alpha(X_{[t,T]}) = -q_\alpha(X_{[t,T]})$$

*where q indicates the $\alpha$-quantile of $X_{[t,T]}$.*
*That is the value $VaR_\alpha$ such that:*

$$P(X_{[t,T]} \leq VaR_\alpha) = \alpha$$

.

Note that VaR is not coherent, because it does not satisfy subadditivity. Furthermore, VaR is not a good measure to limit risks assumed by traders since it answers to the question “How bad can things go?”, but it is useless to deal with other questions such as “If things go bad, how much do we loss?”.
In order to consider this type of instances, it was introduced another measure, the Expected Shortfall.

**Definition 2.1.4.** *Espected Shorfall of level $\alpha \in (0,1)$ of a portfolio at time horizon $T$ is*

$$ES_\alpha(X_{[t,T]}) = -\frac{1}{\alpha}\int_0^\alpha q_u(X_{[t,T]})\,du, = \frac{1}{\alpha}\int_0^\alpha VaR_u(X_{[t,T]})\,du,$$

*that is the expected value of VaR of level $\leq \alpha$ computed at level $\alpha$. $X_{[t,T]}$ is defined as in 2.1.3.*

If $X_{[t,T]}$ is absolutely continuous, it holds also this definition:

$$ES_\alpha(X_{[t,T]}) = -E[x|X_{[t,T]} \leq q_\alpha(X_{[t,T]})] = E[X_{[t,T]}|X_{[t,T]} \leq VaR_\alpha(X_{[t,T]})]$$

A key advantage of VaR over most other measures of risk such as Expected Shortfall is the availability of several back-testing procedures for validating a set of VaR forecasts, Kupiec method (Kupiec (1995)) is the most used. Early examples of back-tests can be found in Christoffersen (1998).

### 2.1.2 VaR Model Back-Test

The shortcomings of methods we will present in this chapter and VaR in general are the most significant reason for a double check of the accuracy of the risk estimates. In fact, in order to evaluate the quality of the estimates, the model should be verified with different back-testing tools. Back-testing is a statistical procedure where actual profits and losses are systematically compared to corresponding VaR estimates, a violation occurs when the daily loss is larger than the VaR calculated the previous day (Out-of-Sample Back-Test). For example, if the confidence level used for daily computation of VaR is 99%, we expect exceptions to occur once in every 100 days on average. In the back-testing process we statistically examine whether the frequency of exceptions over some specified time interval is in line with the selected confidence level. These kinds of tests are known as tests of *unconditional coverage* (S. D. Campbell (2007)). Their implementation is straightforward, since they do not take into account the time when the exceptions occur. However, a good VaR model not only produces the correct amount of exceptions, but also exceptions that are evenly spread over time, i.e., are independent of each other. Clustering of exceptions indicates that the model does not accurately capture the changes in market volatility and correlations. Suitable tests of conditional coverage should be performed to check the independence hypothesis for the times between exceptions.
Among these tests we consider here those mentioned in Basel Committee's traffic light approach (on Banking Supervision (1996)), i.e., the *Kupiec's proportion of failures-test* (Kupiec (1995)) and the *Christoffersen's interval forecast test* (Christoffersen (1998)).

The Basel Regulation back-testing process is based on the fact that if $V$ counts the number of successes in a number $T$ of independent experiments then $V$ is a binomial random variable, whose $p$ parameter is estimated by the number of exceptions. Thus, the Basel Regulation back-testing process test requests of comparing the last 250 daily 99% VaR estimates with corresponding daily trading outcomes. The accuracy of the model (Basel Committee, 1996) is then evaluated by counting the number of exceptions during this period, and accepted or rejected accordingly to the values shown in Figure 2.1. Assuming that the model is correct, the expected number of exceptions is 2.5. If there are zero to four exceptions observed, the model falls into green zone and is defined to be accurate as the probability of accepting an inaccurate model is quite low.

Some references (see, e.g., Haas (2001)) remind that the Basel traffic light approach cannot be used to evaluate the goodness of a VaR model. An example of reason for this criticism is that it ignores the independence of exceptions. Due to the severe drawbacks of the Basel framework, this method is mainly used as a preliminary test for VaR accuracy.
In any kind credible model validation process, the traffic light approach is simply inadequate, and more advanced tests should be performed. Kupiec (1995) sug-

| Zone | Number of exceptions | Increase in scaling factor | Cumulative probability |
|---|---|---|---|
| | 0 | 0.00 | 8.11 % |
| | 1 | 0.00 | 28.58 % |
| Green Zone | 2 | 0.00 | 54.32 % |
| | 3 | 0.00 | 75.81 % |
| | 4 | 0.00 | 89.22 % |
| | 5 | 0.40 | 95.88 % |
| | 6 | 0.50 | 98.63 % |
| Yellow Zone | 7 | 0.65 | 99.60 % |
| | 8 | 0.75 | 99.89 % |
| | 9 | 0.85 | 99.97 % |
| Red Zone | 10 or more | 1.00 | 99.99 % |

Figure 2.1: Traffic light approach: Cumulative probability is the probability of obtaining a given number or fewer exceptions when the model is correct (i.e. true coverage is 99%). The boundaries are based on a sample of 250 observations. For other sample sizes, the yellow zone begins at the point where cumulative probability exceeds 95%, and the red zone begins at cumulative probability of 99.99%

gested the most widely used unconditional coverage test based on failure rates.
Kupiec's test, also known as the POF-test (*Proportion Of Failures*), measures whether the number of exceptions is consistent with the confidence level. The number $V$ of exceptions follows the binomial distribution as discussed previously. Hence, the only information required to implement a POF-test are the total number of observations ($T$) and the observed number of exceptions ($v$). The null hypothesis is that the frequency of violations is consistent with the VaR significance level $\alpha$, i.e., $\alpha = \hat{\alpha} = \frac{v}{T}$, where $\alpha$ is the pre-specified level of VaR. As well-known, under the null hypothesis the ratio

$$\frac{v - \alpha T}{\sqrt{\alpha(1-\alpha)T}}$$

converges to a standard normal distribution as $T$ increases, and the likelihood-ratio

$$LR_k = -2ln\left(\frac{(1-\alpha)^{T-v}\alpha^v}{\left[1-\left(\frac{v}{T}\right)\right]^{T-v}}\left(\frac{v}{T}\right)^v\right)$$

converges to a $\chi_1^2$ distribution (see Kupiec (1995) for details). Hence, the null hypothesis is validated when the $p$-value of the test is lower than 95% or 99%. Resulting $p$-values for Kupiec's test under different confidence levels and sample sizes have been computed.

The Basel framework and unconditional coverage tests, such as the POF-test, focus only on the number of exceptions. In theory, however, we would expect these exceptions to be evenly spread over time. Good VaR models are capable of reacting to changing volatility and correlations in a way that exceptions occur independently of each other, whereas bad models tend to produce a sequence of consecutive exceptions.

Clustering of exceptions is something that VaR users want to be able to detect since large losses occurring in rapid succession are more likely to lead to disastrous events than individual exceptions taking place every now and then. Tests of conditional coverage try to deal with this problem by examining the frequency of VaR violations together with their occurrence times. For this aim, the Christoffersen's Interval Forecast Test ((Christoffersen, 1998)) is suggested. It uses the same log-likelihood testing framework as in Kupiec's test, but extends it to include also a separate statistic for checking the independence of exceptions. With respect to the correct rate of coverage, this test examines whether the probability of an exception on any day depends on the outcome of the previous day. To perform this check, in addiction to Kupiec indicator $v$, this test defines $n_{i,j}$ as the number of days when condition $j$ occurred assuming that condition $i$ occurred on the previous day. We denote with $\pi_i$ the probability of observing an exception conditional on state $i$ on the previous day, i.e., $\pi_0 = \frac{n_{0,1}}{n_{0,0}+n_{0,1}}$, $\pi_1 = \frac{n_{1,1}}{n_{1,0}+n_{1,1}}$. Then the term $\pi = \frac{n_{0,1}+n_{1,1}}{n_{1,0}+n_{1,1}+n_{0,0}+n_{0,1}}$ plays the same role as $\hat{p} = \frac{v}{T}$ in Kupiec's test. If the model is accurate, then an exception today should not depend on whether or not an exception occurred on the previous day. In other words, under the null hypothesis the probabilities $\pi_1$ and $\pi_0$ should be the equal. The relevant test statistic for independence of exceptions is the likelihood-ratio

$$LR_{ind} = -2ln\left(\frac{(1-\pi)^{n_{0,0}+n_{1,0}}\ \pi^{n_{0,1}+n_{1,1}}}{(1-\pi_0)^{n_{0,0}}\ \pi_0^{n_{0,1}}\ (1-\pi_1)^{n_{1,0}}\ \pi_1^{n_{1,1}}}\right)$$

By combining this independence statistic with Kupiec's POF-test we obtain a joint test that examines both properties of a good VaR model, the correct failure rate and independence of exceptions, i.e. conditional coverage: $LR = LR_k + LR_{ind}$. Under the null hypothesis the quantity $LR$ is sum of two $\chi^2_1$ quantities, hence it is distributed as a $\chi^2$ with 2 degrees of freedom.
Christoffersen's framework allows examining the reason for not passing the test deciding if it is caused by inaccurate coverage by clustered exceptions or by both of them. This evaluation can be done simply by calculating each statistic, $LR_k$ and $LR_{ind}$,separately and using $\chi^2$ distribution with one degree of freedom as the critical value for both statistics. S. D. Campbell (2007) reminds that in some cases it is possible that the model passes the joint test while it fails either the independence test or the coverage test. Therefore it is advisable to run the separate tests even when the joint test yields a positive result.

## 2.2 Forecasting Models

In this section we present forecasting models used in industry and those recently proposed in academic world. The list of the considered models can be divided according with three main types of models: parametric models, historical simulation and Monte Carlo simulation

### 2.2.1 Historical Simulation

The historical simulation is based on the hypothesis that the past observed returns are a good and complete representation of future returns probability density function (pdf). This is sometimes called the empirical distribution of returns.
It has a huge use in industries because of its simplicity and lack of distributional assumption about underlying process of returns. it only assumes that the historically observed returns used in the simulation are taken from independent and identical distributions (iid) and that the same distributions can be applied for forecasting.

For the *standard* VaR, historical simulation (empirical distribution) method suppose to observe data from day 1 to day t. Indicating with $X_t$ the return of asset on day $t$, then the method is based on a series of observed returns $x_1, x_2, \ldots, x_t$. The Value at Risk with coverage rate $\alpha$ is calculated as:

$$VaR^{\alpha}_{t+1} = percentile[x_1, x_2, \ldots, x_t, \alpha]$$

Bootstrap historical simulation approach is an extension of traditional historical simulation. It is a simple and intuitive estimation procedure. The bootstrap technique draws a sample from the data set, records the VaR from that particular sample and returns the data. This procedure is repeated over and over and records multiple sample VaRs. Since the data is always returned to the data set, this procedure is like sampling with replacement.The best VaR estimate from the full data set is the average of all sample VaRs. (See Chernick, Gonzalez-Manteiga, Crujeiras, and Barrios (2011))

An alternative to bootstrapping consists in the use of kernel estimation of historical distribution. The kernel density estimator ((Silverman, 1986), (Sheather & Marron, 1990)) is a way of determining the probability density function from the data. While a histogram results in a density that is piecewise constant, a kernel estimator results in a smooth density. Any continuous shape spread around each data point allows to smooth the data to obtain a continuous density. As the sample size grows, the net sum of all the smoothed points approaches the true pdf, whatever that may be, irrespective of the method of smoothing the data. The most used is the Gaussian Kernel density. In the first step we estimate the pdf and cumulative distribution function (cdf) of asset returns. Define the pdf of the asset returns as $f$ and the cdf of the returns as $F$. The kernel estimator of the pdf , using a fixed Gaussian kernel, is given by

$$\hat{f}(x) = \frac{1}{n} \frac{1}{x\,h} \sum_{i=1}^{n} \frac{1}{\sqrt{2\pi} e^{-\frac{1}{2}\left(\frac{x-x_i}{h}\right)^2}}$$

where $h$ is the bandwidth, which in this case can be interpreted as a standard deviation and $x_i$ is the return of day $i$and $n$ is the sample size.

In the second step we estimate the distribution of a percentile of the order statistic.
Using pdfs estimated with the kernel density estimator above, it can be derived the pdf of the $j-th$ order statistic and calculate its mean and variance. The pdf is not known in closed form but we can compute its moments through numerical methods. The mean implied by that pdf is the estimate of VaR. From the standard error of the estimate one can calculate confidence intervals.
The distribution of the $j$-th order statistic is derived as follows (Stuart and Ord (1987)).
Suppose our observations $x_1, x_2, \ldots, x_n$ come from some known distribution (or cumulative density) function $F(x)$, with $j-th$ order statistic $x_{(j)}$. Hence, $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$. The probability that $j$ of our $n$ observations do not exceed a fixed value $x$ must then obey the following binomial distribution:

$$g_j(x) = \frac{n!}{j!(n-j)!} F(x)^j (1 - F(x))^{n-j}. \tag{2.2.1}$$

It follows that the probability that at least $j$ observations in the sample do not exceed $x$ is also a binomial

$$G_j(x) = \sum_{k=j}^{n} \frac{n!}{k!(n-k)!} F(x)^k (1 - F(x))^{n-k}, \tag{2.2.2}$$

$G_j(x)$ is therefore the distribution function of our order statistic. It follows, in turn, that $G_j(x)$ also gives the distribution function of our VaRs. For further details we refer to Butler and Schachter (1996) and Hendricks (1996).

As already mentioned, the main strength of the historical simulation approach is that it is non-parametric (no specific distributional assumptions about the data are made ex ante, and no distributional parameters need to be estimated). Therefore, the data determine the shape of the return distribution. Hendricks (1996) showed that the historical simulation approach provided good estimates of the first percentile of the distribution using simulated spot foreign exchange portfolios with departures from normality in the return distribution. Mahoney (1995) obtained a similar result studying simulated spot currency and equity portfolios.
When the historical sampling period is too short arises one of the main shortcoming of historical simulation. In this case arises the risk of unreliable estimation of VaR ; ? (?) found that longer historical sample periods result in less variability in the VaR estimate. In applying this approach a trade-off must be made between long sample periods, which potentially violate the assumption of i.i.d. observations, or a parametric model such as GARCH with short sample periods, which reduce the precision of the estimate.
A related problem in the historical simulation approach is that the only possible

changes in returns in the forecast distribution are those that are observed in the historical sample period. This problem may be especially significant in the estimation of tail probabilities or to intercept a 'black swan' event.

### 2.2.2 Filtered Historical Simulation

Filtered historical simulation (FHS) is a generalized historical simulation method proposed by Barone-Adesi, Giannopoulos, and Vosper (1999). They try to render returns i.i.d. removing any serial correlation and volatility clusters present in the dataset, after filtering ARMA-GARCH residuals result i.i.d. then easier to simulate. They propose to capture volatility clusters by modeling returns as GARCH processes ((Bollerslev, 1986)) and the serial dependency with ARMA models.
The resulting filter is a combination of ARMA and GARCH models described in Chapter 1.

**Definition 2.2.1.** *The process* $\mathbf{X} = \{X_t, t \in T\}$ *follows an ARMA(1,1)-GARCH(1,1) process if for every* $t$ *the r.v.* $X_t$ *satisfies:*

$$X_t = \phi X_{t-1} + \theta \xi_{t-1} + \xi_t. \tag{2.2.3}$$

*where* $\theta$ *is the MA parameter,* $\phi$ *is the AR parameter and* $\{\xi_t, t \in T\}$ *follows a GARCH model where for every* $t$ *the r.v.* $\xi_t$ *satisfies:*

$$E_{t-1}[\xi_t] = 0 \tag{2.2.4}$$

*and the conditional variance*

$$\sigma_t^2 = k + \alpha \xi_{t-1}^2 + \beta \sigma_{t-1}^2 \tag{2.2.5}$$

*depends non trivially on the* $\sigma$*-field generated by the past observations:* $\xi_{t-1}, \xi_{t-2}, \ldots$ *and where* $k$ *is a constant,* $\alpha$ *and* $\beta$ *are the GARCH(1,1) parameters.*

To standardize residual returns we need to divide the estimated residual $\hat{\xi}$ by the corresponding volatility $\hat{\sigma}_t$. Thus the standardized residual return is given as $z_t = \frac{\hat{\xi}_t}{\hat{\sigma}_t}$.
Under the GARCH hypothesis the set of standardized residuals are independent and identically distributed (i.i.d.) and therefore suitable for historical simulation.
It is clear that this method is similar to a Monte Carlo simulation. However, traditional approaches based on Monte Carlo simulation typically use a set of stochastic differential equations for generating returns over the time horizon. So traditional Monte-Carlo simulation uses arbitrary assumptions about the distribution of returns, which define "a priori" the structure of risk that is supposed to investigate.
The results of FHS are sensitive to changing market conditions and can predict losses outside the historical range. From a computational standpoint, this method is very appealing for large portfolios and empirical evidence supports its predictive ability. Various authors extensively tested it, see for example Zenti and Pallotta

Table 2.1: Historical standardized residuals of portfolio assets

| | | | | |
|---|---|---|---|---|
| $z_{1,-T}$ | $z_{2,-T}$ | $\dots$ | $z_{n-1,-T}$ | $z_{n,-T}$ |
| $z_{1,-T+1}$ | $z_{2,-T+1}$ | $\dots$ | $z_{n-1,-T+1}$ | $z_{n,-T+1}$ |
| $\dots$ | $\dots$ | $\dots$ | $\dots$ | $\dots$ |
| $z_{1,t-1}$ | $z_{2,t-1}$ | $\dots$ | $z_{n-1,t-1}$ | $z_{n,t-1}$ |
| $z_{1,t}$ | $z_{2,t}$ | $\dots$ | $z_{n-1,t}$ | $z_{n,t}$ |
| $z_{1,t+1}$ | $z_{2,t+1}$ | $\dots$ | $z_{n-1,t+1}$ | $z_{n,t+1}$ |
| $\dots$ | $\dots$ | $\dots$ | $\dots$ | $\dots$ |
| $z_{1,0}$ | $z_{2,0}$ | $\dots$ | $z_{n-1,0}$ | $z_{n,0}$ |

(2002).

This model is a univariate model, it focus only on an asset or a portfolio returns directly, without taking in consideration the diversification effects of asset allocation in portfolio or the dependence between assets.
A possible evolution of this model is the Parallel Filtered Bootstrapping (Marsala, Pallotta, and Zenti (2004)). The main idea is to use implicit correlation of asset returns focusing on implicit historical correlation of standardized residuals $z_t$.

Given a portfolio of $n$ assets, let $X_{i,t}$ be the return at time $t$ of asset $i$, with $i = 1, \dots, n$, we estimate the ARMA-GARCH filter of each asset returns (as in 2.2.1) and calculate the standardized residuals $z_{i,t} = \frac{\hat{\xi}_{i,t}}{\hat{\sigma_{i,t}}}$ to get:

Instead of bootstrapping the single standardized return $z_{i,t}$ they bootstrap the vector of standardized returns of all portfolio assets at time $t$, hence $[z_{1,t}, z_{2,t}, \dots, z_{n-1,t}, z_{n,t}]$. In this way they reproduce the historical correlation of past time $t$ in forecasting simulation, reflecting actual time-varying correlations.

The parallel Filtered Bootstrap approach to the simultaneous assessment of market risk of assets returns relies on an empirical multivariate probability distribution inferred from financial data in a semi-parametric way. The model captures most of the features observed in financial time series, such as conditional heteroskedasticity, autocorrelation and leptkurtosis. As bootstrap is applied in a parallel way to all the risk factors (asset returns), their comovements are captured in a semi-parametric way, without any restrictive hypothesis about the correlation structure, that in general is not linear and it is variable over time.
These features characterize a suitable and flexible way to generate future scenarios both on short and medium term horizons. Therefore, this model is particularly appropriate for asset management companies with a broad investment universe, usually managing several balanced portfolios, with multiple horizons to keep under control.
A possible improvement requests the use of an explicit correlation (may be time-varying) in parallel bootstrapping in absence of crisis period in historical sample

and no comonotonic phenomena in historical data is observed. In these conditions, the observed correlations are the only possible. The existence of an explicit expression for the correlation allows to stress correlation between assets and gives a more intuitive view of diversification effect in portfolio. It clarifies not only the most important contributors to risk, but also the dependence of these contributors to the rest of assets in the portfolio, allowing more efficient asset allocation.
In Chapter 3 we propose an alternative to this model trying to catch the above mentioned improvements through the use of a GARCH-Copula approach.

### 2.2.3 Parametric Model - EVT

The basic idea of parametric models is to find a specific well known distribution of returns (e.g. Normal, Weibull, t-student, ...). The main strength of these models is that they use distributions that are known in close form (hence the quantile, the volatility, the expected returns, ... are known too).
This fact avoids the use of simulations and only the parameter estimations become necessary.
The classical example is the multivariate normal distribution. The model hypothesize that every asset returns in a portfolio fits the Gaussian distribution. Hence, known mean vector and variance-covariance matrix it is possible to get all risk indicators without necessity of any simulation. Even after 2008 financial crisis, that showed how this approach is insufficient to measure the risk in stressed condition, there are companies that still provide risk indicators that use this model for their calculation (Riskmetrics®, Barra®).
It is possible that some asset returns can be described with parametric distribution, however the fact that the theoretical distribution fits returns for one asset does not mean that it works well for every asset in the portfolio. Hence, it is not possible to use a multivariate parametric distribution but, in this case, it is possible to model separately the marginal distributions of each assets and then model separately the dependence between them with multidimensional copulas.
Building multidimensional copulas has always been considered a difficult problem. As we saw in Chapter 1, there are large number of families of bivariate copulas, but the set of the multidimensional ones is very limited. Attempts have been made to build multivariate extensions of Archimedean bivariate copulas, however constructions of this type have limitations: one of the strongest is that it can be applied only on variables that exhibit the same dependence structure. To solve the problem, it is introduced a new multivariate model construction methodology: pair-copula construction (PCC).

### 2.2.4 Pair Copula Construction and Vine Copulas

The pair-copula construction (PCC) theory concerns the construction of multivariate structures through the use of bivariate copulas, the so-called *pair-copulas*. It is

introduced by Joe (1996) and then developed in Bedford and Cooke (2001),Bedford and Cooke (2002). With Aas, Czado, Frigessi, and Bakken (2009), Czado (2010) and Dissman, Brechmann, Czado, and Kurowicka (2013) we get industry application examples.
Nowadays, this method represents the most flexible and intuitive way to construct multivariate copulas.

**Definition 2.2.2.** *A pair copula is a bivariate copula with conditioned copulas as argument.*

To better understand this definition let us consider $(X_1, \dots, X_n)$ random variables with joint distribution function $F$. We can write their density $f$ uniquely as:

$$f(x_1, \dots, x_n) = f_1(x_1)\, f_{2|1}(x_2|x_1)\, f_{3|12}(x_3|x_1, x_2)\ \dots\ f_{n|1\dots n-1}(x_n|x1, \dots, x_{n-1}). \tag{2.2.6}$$

If $F$ is absolutely continuous with continuous and strictly increasing marginals $F_i(x_i)$, using Sklar theorem (1.2.7), we can rewrite (2.2.6) as

$$f(x_1, \dots, x_n) = c_{1,2,\dots,n}(F_1(x_1), F_2(x_2), \dots, F_n(x_n))\, f_1(x_1)\ \dots\ f_n(x_n). \tag{2.2.7}$$

where $c_{1,2,\dots,n}$ is the density of copula $C_{1,2,\dots,n}$ that is unique.
The objective is rewrite (2.2.6) using (2.2.7).
The second factor $f_{2|1}(x_2|x_1)$ can be rewritten as:

$$\begin{aligned} f_{2|1}(x_2|x_1) &= \frac{f(x_1, x_2)}{f_1(x_1)} \\ &= \frac{c_{12}(F_1(x_1), F_2(x_2)) f_1(x_1)\, f_2(x_2)}{f_1(x_1)} \\ &= c_{12}(F_1(x_1), F_2(x_2)) f_2(x_2) \end{aligned} \tag{2.2.8}$$

The third factor $f_{3|12}(x_3|x_1, x_2)$ can be rewritten as:

$$\begin{aligned} f_{3|12}(x_3|x_1, x_2) &= \frac{f_{23|1}(x_2, x_3|x_1)}{f_{2|1}(x_2|x_1)} \\ &= \frac{c_{23|1}(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1)) f_{2|1}(x_2|x_1)\, f_{3|1}(x_3|x_1)}{f_{2|1}(x_2|x_1)} \\ &= c_{23|1}(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1))\, f_{3|1}(x_3|x_1) \\ &= c_{23|1}(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1))\, c_{13}(F_1(x_1), F_3(x_3)) f_3(x_3). \end{aligned} \tag{2.2.9}$$

We can see that this third factor of (2.2.6) is the product of two pair-copulas and one univariate density.
Generalizing we get the n-dimensional case. We omit functions arguments to get

a simpler notation and we consider $c_{i,j|i_1,\ldots,i_k}$ with $i < j$ and $i_1 < \ldots < i_k, k \in [1, n-1]$ conditional copula densities and we get:

$$\begin{aligned} f_{n|1,\ldots,n-1} &= c_{1,n|2\ldots n-1}(F_{1|2\ldots n-1}, F_{n|2\ldots n-1}) f_{n|2\ldots n-1} \\ &= \left[\prod_{s=1}^{n-2} c_{s\ n|s+1\ldots n-1}(F_{s|s+1\ldots n-1}, F_{n|s+1\ldots n-1})\right] c_{n-1\ n}(F_{n-1}, F_n)\ f_n. \end{aligned} \tag{2.2.10}$$

Substituting (2.2.10) in (2.2.6) we get the pair copula construction of multivariate density:

$$f = \left[\prod_{j=1}^{n-1}\prod_{i=1}^{n-j} c_{i\ i+j|v_{ij}}(F_{i|v_{ij}}, F_{j|v_{ij}})\right] \prod_{k=1}^{n} f_k \tag{2.2.11}$$

where $v_{ij}$ is the set of indexes $i+1, \ldots, i+j-1$.
If $v_{ij}$ is 1-dimensional:

$$F(x|v) = \frac{\partial C_{xv}(F_x(x), F_v(v))}{\partial F_v(v)}. \tag{2.2.12}$$

When $x$ and $v$ are uniform variables (f(x)=f(v)=1), we define $F(x, v) = h(x, v, \Theta)$ and obtain

$$h(x, v, \Theta) = F(x|v) = \frac{\partial C_{xv}(F_x(x), F_v(v))}{\partial F_v(v)} \tag{2.2.13}$$

where $h$ is the conditional variable and $\Theta$ is the set of parameters of the copula $C_{xv}$.

**Vines**
For high dimension distributions, there is a large number of possible PCCs. Bedford and Cooke ((Bedford & Cooke, 2001),(Bedford & Cooke, 2002)) have introduced graphical models that help to organize possible pair-copulas decompositions. These models have been called *regular vines*. To define the vines, first we remember the definition of tree:

**Definition 2.2.3.** *A tree $T = N, E$ is an acyclic connected graph, where $N$ is the set of nodes and $E$ set of edges, non-ordered nodes couples.*

A tree with $n$ nodes has $n-1$ edges. A tree is completely determined if all its edges are specified (see an example in Figure 2.2).

**Definition 2.2.4.** *V is a vine of $n$ elements if:*

- *$V = (T_1, \ldots, T_n - 1)$;*
- *$T_1$ is a tree with $N_1 = 1, \ldots, n$ set of nodes and $E_1$ set of edges;*

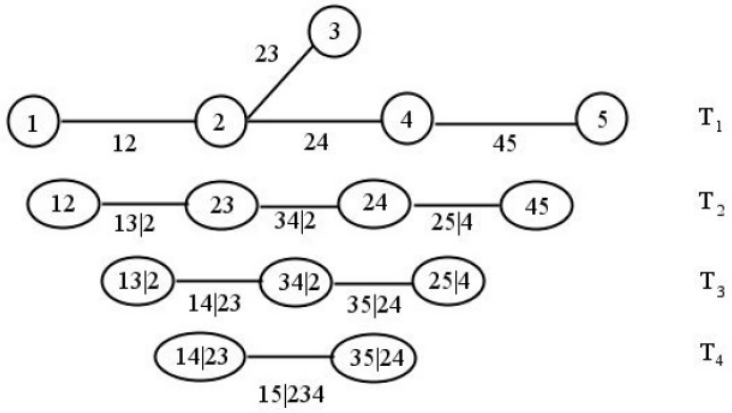


Figure 2.2: *Regular vine* trees representation for $d = 5$.

- *$T_i$ is a tree with $N_i = E_{i-1}$ set of nodes and with card$(N_i) = n - (i-1)$ for $i = 2, \ldots, n$;*

*V is a regular vine if it holds also:*

- *proximity condition: if $a = \{a_1, a_2\}$ and $b = \{b_1, b_2\}$ are two nodes of $T_{i+1}$ such that $\{a, b\} \in E_{i+1}$, then one of $a_j$ is equal to one of $b_j$ for $i = 2, \ldots, n-1$ and $j = 1, 2$.*

**Proposition 2.2.1.** *Let V $= (T_1, \ldots, T_{n-1})$ be a R-Vine, then:*

- *the V edges number is $\frac{n(n-1)}{2}$;*
- *Each conditioning set corresponds to a couple and each couple of indexes appears only one time as a conditioning set;*
- *If two edges have the same conditioning set then they are the same edge.*

The R-vines is a wide class and it is possible to identify large number of structures but we focus on two specific families of vines: Canonical-Vines (C-vines) and Drawable-Vines (D-vines).

**Definition 2.2.5.** *A R-vines is:*

- *a C-vines if each $T_i$ has only one node of degree $n-i$, with $i = 1, \ldots, n-1$;*
- *a D-vines if each node in $T_1$ has at most degree 2.*

Working with these two families of vines is often convenient: their structure is completely determined by the first level tree $T - 1$, in D-vine case, or by the hierarchy of conditioning of the variables, in C-vine case. These properties also allow to calculate the number of different C-vines and D-vines that can be obtained with

$n$ variables.
In the C-vines the conditioning set is the same for all the edges of a tree. In general, for the choice of the only element of the conditioning set in the $T_2$ tree there are $n$ possible choices, for the tree $T_3$ the number of choices decreases to $n-1$ and so on up to the tree $T_{n-2}$ for which there are three choices. So in total there are $\frac{n!}{2}$ different C-vines with $n$ variables.
For a n-dimensional D-vine there are $n!$ different permutations for the indexes in $T_1$; however, the edges of a generic vine are not oriented: the order of the nodes in the first tree of a D-vine can be inverted without changing the type of vine. As a consequence, the number of permutations that leads to actually different D-vines is halved, reducing again to $\frac{n!}{2}$.

#### Link between R-Vines and PCC

By introducing the R-vines, we obtain models whose purpose is to provide help in organizing the various decompositions in the PCC. In particular, the link between these graphic structures and pair-copulas occurs through the combination of R-vines edges to pair-copulas. To each R-vine corresponds therefore a R-vine copula specification, formally defined in this way:

**Definition 2.2.6.** $(\boldsymbol{F}, V, B)$ *is a R-vine copula specification (or vine-copula) if:*

- $\boldsymbol{F} = F_1, \ldots, F_n$ *is a continuous and invertible distribution function vector;*
- *V is a R-vine of $n$ elements;*
- $B = B_e | e \in E_i, i = 1, \ldots, n-1$ *is a set of bivariate copulas, in this case pair-copulas.*

**Definition 2.2.7.** *The joint distribution $F$ of a random vector $(X_1, \ldots, X_n)$ realizes a R-vine copula specification $(\boldsymbol{F}, V, B)$ if for $i = 1, \ldots, n-1$ and for each edge $e = a, b \in E_i$ then $B_e \in B$ is the bivariate copula of $X_{C_{e,a}}$ and $X_{C_{e,b}}$ conditioned by $X_{D_e} = X_i | i \in D_e$.*
*$c_{C_{e,a}, C_{e,b}|D_e} = c_{a,b|D_e}$ is the density function of bivariate copula $B_e$ for edge $e = a, b$.*
*$F_j$ is the marginal distribution of $X_j$ for $j = 1, \ldots, n$.*

The density of a R-vine can be factored into the product of the conditioned and unconditioned bivariate copulas associated with each of its edges.

**Theorem 2.2.1.** *Let $(\boldsymbol{F}, V, B)$ be a vine-copula on $n$ elements, then there is one and only one distribution function of R-vine, and the density function $f$ is given by:*

$$f = \prod_{k=1}^{n} f_k \prod_{i=1}^{n-1} \prod_{e \in E_i} c_{a,b|D_e}(F_{a|D_e}, F_{b|D_e}) \tag{2.2.14}$$

*where $e = a, b$ and $f_i$ is the marginal density corresponding to marginal distribution $F_i$, $i = 1, \ldots, n$. The density arguments are omitted for simpler writing.*

Equation (2.2.14) shows that the vine-copula has closed form density when $F_1, \dots, F_n$ and bivariate copula $B$ are differentiable.
For a C-vine the set of bivariate copula is $B = C_{i_1,i_2|1,\dots,i_1-1} : 1 \le i_1 < i_2 \le n$, while for a D-vine is $B = C_{i_1,i_2|i_1+1,\dots,i_2-1} : 1 \le i_1 < i_2 \le n$.
Using these notions we get C-vine and D-vine density functions.
The $n$-dimensional density $f$ corresponding to a canonical vine is given by

$$f = \prod_{k=1}^{n} f_k \prod_{j=1}^{n-1} \prod_{i=1}^{n-j} c_{j,j+i|1,,j-1}(F_{j|1,,j-1}, F_{j+i|1,,j-1}). \tag{2.2.15}$$

The $n$-dimensional density $f$ corresponding to a drawable vine is given by

$$f = \prod_{k=1}^{n} f_k \prod_{j=1}^{n-1} \prod_{i=1}^{n-j} c_{i,i+j|i+1,,i+j-1}(F_{i|i+1,,i+j-1}, F_{i+j|i+1,,i+j-1}) \tag{2.2.16}$$

We will present an empirical application of these techniques in Chapter 3, Section 3.3.2.

### 2.2.5 Conclusion

In this chapter we described different techniques for forecasting financial time series, each of them has some drawback.
Filtered Historical Simulation presented in Section 2.2.2 use an implicit correlation, hence it is not possible use this model for a "Stress Test" on increasing correlation and we cannot monitor the correlation evolution of the assets or do some analysis on dependence.
The vine approach is effective only if we estimate the best fitting bivariate copula at each step of vine estimation and it is onerous for portfolio with large number of assets.
An alternative is define "a priori" the bivariate copula to use, but also in this case we lose the dependence focus we need to describe correctly our portfolio dynamics.
In the next chapter we try to improve these models and analyze the difference between "old" and "new" also using the back–test techniques presented at the beginning of this chapter.

# Chapter 3

# Modern approaches and new results

We have seen in Chapter 1 and Chapter 2 different tools and models to forecast assets and portfolios returns. The focus is on both marginal distribution of each asset and dependence structure of all assets in portfolio allocation. An important task that is not analyzed in the previous chapter is the stability of the chosen model (hence parameters) over time; in fact, all presented models have not time-varying parameters and describe a static portfolio. This is a good approximation if the forecasting time-horizon is one day, but it can be misleading if the time-horizon is higher because the market condition evolves over time and the same should happen to the parameters.
There exists a huge number of references where authors try to construct a mathematical tools to reproduce time–varying dependence between assets. All these models seem to be appropriate in specific contexts, but they require an initial effort in the model construction that financial industry cannot allow. An example is Dias and Embrechts (2010), where a new flexible time–varying copula approach with new correlation specifications to describe the dependence between exchange rates is proposed. In their paper they present a new approach that allows the conditional correlation between exchange rates to be both time–varying and modeled independently from the marginal distributions. They introduce a dynamic specification for the correlation using the Fisher transformation. Applied to Euro/US dollar and Japanese Yen/US dollar, their results reveal a significantly time–varying correlation, dependent on the past return realizations. The dynamic copula model outperforms at six different time horizons, ranging from hourly to daily, confirming the model specification. In this case the model performs better than other well-known models, but under a bivariate specification. In our opinion, it can be extremely hard to create a similar model to describe time–varying correlation for financial portfolios involving many assets.
In this chapter we try to give an alternative formulation or an improvement of the existing models focusing on stability of parameters and dependence, also having

in mind their impact and flexibility for applicative purposes. In particular, we describe the proposed model, related inference procedures and simulations applied to a specific financial portfolio.
In the first section we present and describe the portfolio used in analysis, with some additional statistical analysis.
In the next section we present a filtered GARCH-Copula model which is based, on historical simulations, on a Montecarlo simulation and on a parametric assumption for the dependence structure. We focus on the analysis of dependence structure of this model and on the stability and evolution of model parameters. We present the back-testing and a comparison with FHS.
Last section is devoted to propose a new tool, similar to copula, that overcomes the limitation highlighted by the *'Impossibility Theorem'* and that can be used in forecasting model as alternative to Vine Copulae.

## 3.1 Financial application

### 3.1.1 Data Description

In our analysis, we considered an equally weighted portfolio of seven MSCI (Morgan Stanley Capital International) regional indexes of financial markets. Being the understanding of the global financial situation in the last decade among our goals, with reference to the evolution of dependence in financial markets across the world, we focused on the macro regions World, North America, Asia-except-Japan, Japan, Europe, Italy and Switzerland. Particular attention on market behaviors in Europe and Italy compared with Switzerland and Japan was given. In details, the following indexes have been considered.

- The *MSCI World Index*, which captures large and mid cap representation across 23 Developed Markets (Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Hong Kong, Ireland, Israel, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, Switzerland, the UK and the US).

- The *MSCI North America Index*, which is designed to measure the performance of the large and mid cap segments of the US and Canada markets.

- The *MSCI AC Asia except Japan Index*, which captures large and mid cap representation across Developed Markets countries in Asia (Hong Kong and Singapore, excluding Japan) and 8 Emerging Markets countries (China, India, Indonesia, Korea, Malaysia, the Philippines, Taiwan and Thailand).

- The *MSCI Europe Index* captures large and mid cap representation across 15 Developed Markets (Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Italy, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and the UK) countries in Europe.

- The *MSCI Japan Index* captures large and mid cap segments of the Japanese market. With 319 constituents, the index covers approximately 85% of the free float-adjusted market capitalization in Japan.
- The *MSCI Italy Index* captures large and mid cap segments of the Italian market. With 23 constituents, the index covers about 85% of the equity universe in Italy.
- The *MSCI Switzerland Index* captures large and mid cap segments of the Swiss market. With 38 constituents, the index covers approximately 85% of the free float-adjusted market capitalization in Switzerland.

It should be pointed out that the top constituents of both Europe and Switzerland are the same, with different weight, and this cause a positive correlation between the two corresponding indexes.The same situation occurs in the case of MSCI World Index and MSCI North America Index. This fact is confirmed in Figure 3.1, where the evolution of these historical correlations is always above 0.75.

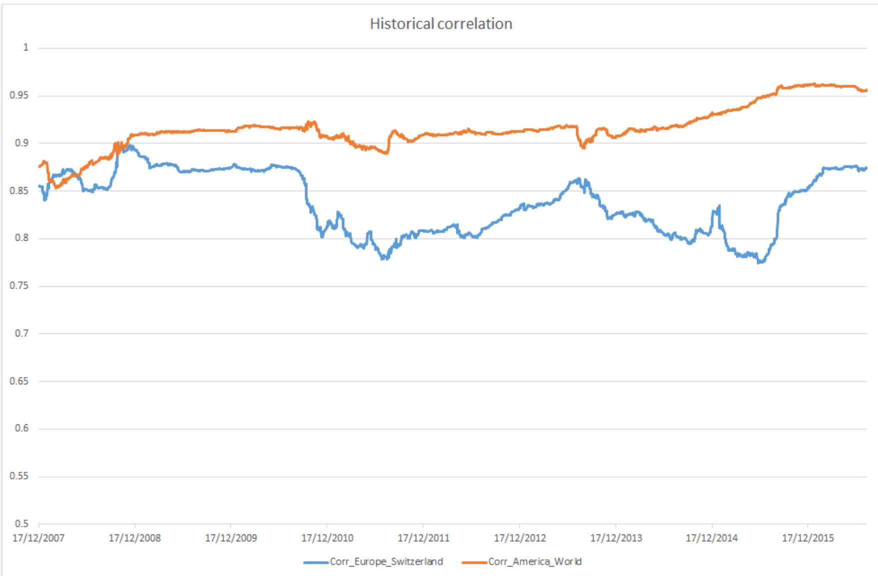


Figure 3.1: Evolution of historical correlation (based on calculation of 2 years of daily returns correlation) between MSCI Switzerland and MSCI Europe (Blu line) and MSCI North America and MSCI World. Period 12/2005 to 07/2016

The data is composed of daily observations (last prices), comprising the period from 29-th December 2005 to 29-th July 2016, with a total of 2762 points. During this period all the indexes were affected by the *financial crisis* in the United States that started in 2008 (in our analysis from end of 2007 to end of 2009), the European debt crisis of 2011 (in our analysis from begin of 2011 to end of 2012), the Chinese stock market turbulence in summer 2015 and at the beginning of 2016, and the Brexit event for the European market. Temporal evolutions of these indexes are shown in Figure 3.2.

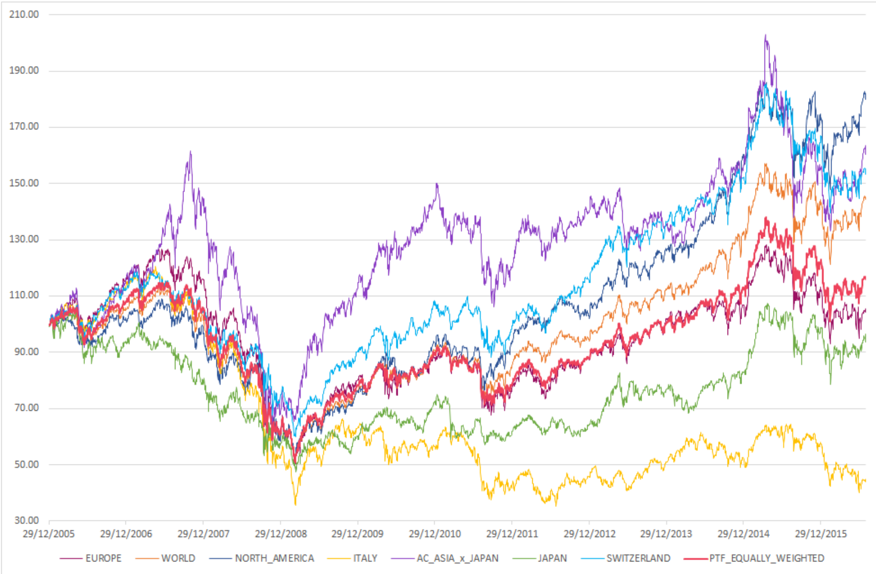


Figure 3.2: Price evolution of indexes and the Equally-Weighted analyzed portfolio. Period 12/2005 to 07/2016

Our analyses are daily, hence we prefer to focus on a different market period to see if our model fitting is good enough in every condition. To this aim, we choose three different crisis periods of last decade described above. To account for the last Crisis we analyze the 2 years from September 2014 to August 2016 that are characterized also by fast mini-crisis. In all the cases the modeling is performed with previous daily returns of last 2 years (504 daily observations). This means that, for example for *2008 Crisis*, the data used for the first step are from December 2005 to December 2007 and for the last step are from October 2009 to October 2011.
The VaR simulated with data return observed from $t$ to $t + 503$ is compared with the observation of the return at the end of the working-day $t + 504$.

### 3.1.2 Statistical analysis

In this section we verify that, as we expected, the Gaussian Hypothesis is not correct and that Financial Crisis can be identify with high-volatility periods.
For each rolling window used for model estimation (504 daily observations) we verify the rejection of Jarque-Bera Null Hypothesis (JB $p$-value $< 0.01$ reject the null hypothesis of Gaussian distribution, (Jarque & Bera, 1987)), and calculate some statistical indicators.

Figures 3.3 and 3.4 show an example of the descriptive statistics done at each step of estimation for the continuous returns of such data . We can see that this data usually exhibits both negative asymmetry and kurtosis excess. Also, according to Jarque-Bera test statistics, there is evidence that the log-returns are not normally distributed, i.e., the marginal of returns are not in an elliptical world.

**Descriptive statistics for the period 30/12/2005 - 05/12/2007**

| | EUROPE | WORLD | NORTH AMERICA | ITALY | ASIA ex JAPAN | JAPAN | SWITZERLAND |
|---|---|---|---|---|---|---|---|
| Mean | 0.03% | 0.01% | -0.01% | 0.01% | 0.08% | -0.04% | 0.02% |
| St.d. | 0.89% | 0.71% | 0.89% | 0.86% | 1.19% | 1.23% | 0.80% |
| Min | -3.54% | -3.11% | -4.08% | -3.43% | -5.12% | -8.60% | -3.06% |
| Max | 2.69% | 2.23% | 3.04% | 2.18% | 5.94% | 3.31% | 2.24% |
| Skewness | -0.51 | -0.52 | -0.47 | -0.58 | -0.38 | -0.72 | -0.55 |
| Kurtosis | 4.63 | 4.26 | 4.87 | 4.10 | 5.67 | 7.56 | 4.28 |
| JB p value | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |

Figure 3.3: Descriptive statistics of Indexes log–returns. 12/2005 - 12/2007

**Descriptive statistics for the period 29/10/2009 - 04/10/2011**

| | EUROPE | WORLD | NORTH AMERICA | ITALY | ASIA ex JAPAN | JAPAN | SWITZERLAND |
|---|---|---|---|---|---|---|---|
| Mean | -0.02% | 0.01% | 0.04% | -0.09% | 0.00% | 0.01% | 0.01% |
| St.d. | 1.22% | 0.90% | 1.08% | 1.66% | 1.19% | 1.36% | 1.01% |
| Min | -4.82% | -5.06% | -6.56% | -6.62% | -4.90% | -8.55% | -4.01% |
| Max | 6.82% | 3.96% | 4.51% | 10.43% | 3.85% | 8.58% | 3.87% |
| Skewness | -0.23 | -0.68 | -0.67 | -0.05 | -0.49 | -0.29 | -0.53 |
| Kurtosis | 6.42 | 6.80 | 7.51 | 7.08 | 4.02 | 10.68 | 5.44 |
| JB p value | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |

Figure 3.4: Descriptive statistics of Indexes log–returns. 10/2009 - 10/2011

In following graphs one can observe the evolution of Statistical Indicators. As expected the skewness and kurtosis values confirm the non Gaussianity of historical samples at each step, with few exceptions for some intervals of *MSCI AC Asia except Japan Index* and *MSCI Japan Index* (Kurtosis near 3 in Figure 3.8).

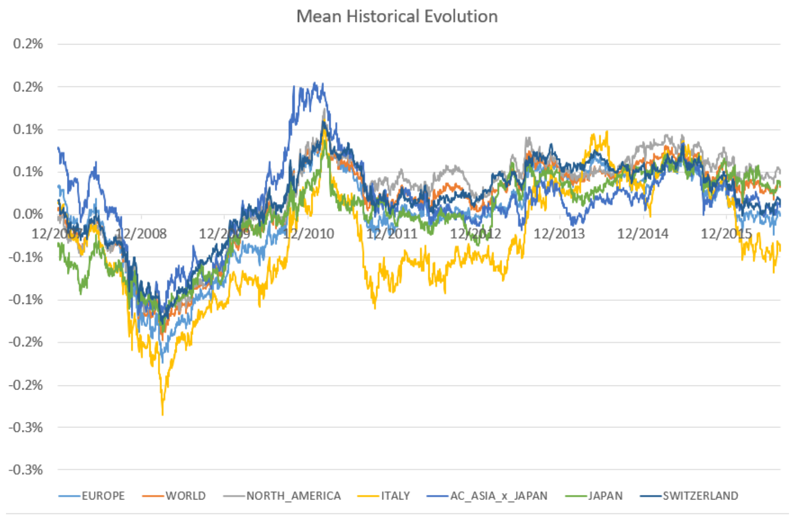


Figure 3.5: Mean evolution of Indexes log–returns.

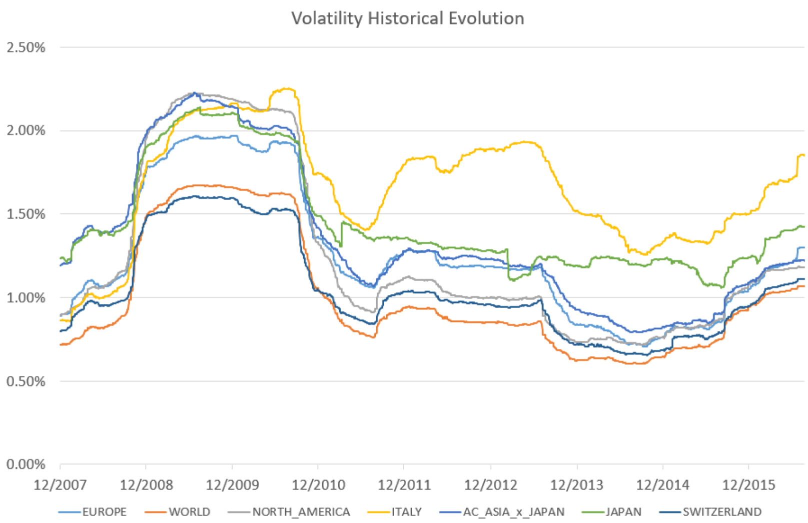


Figure 3.6: Daily Volatility evolution of Indexes log–returns.

## 3.2 GARCH-Copula

For financial industry the main goal is end up with a model which is easy to understand and maintain, that could help in the interpretation of the final risk indicators, and that could quickly enough react to the changes in market conditions without onerous model updating. The idea of this section is to investigate if a model based on standard assumptions and constant parameters, where the time–varying depen-

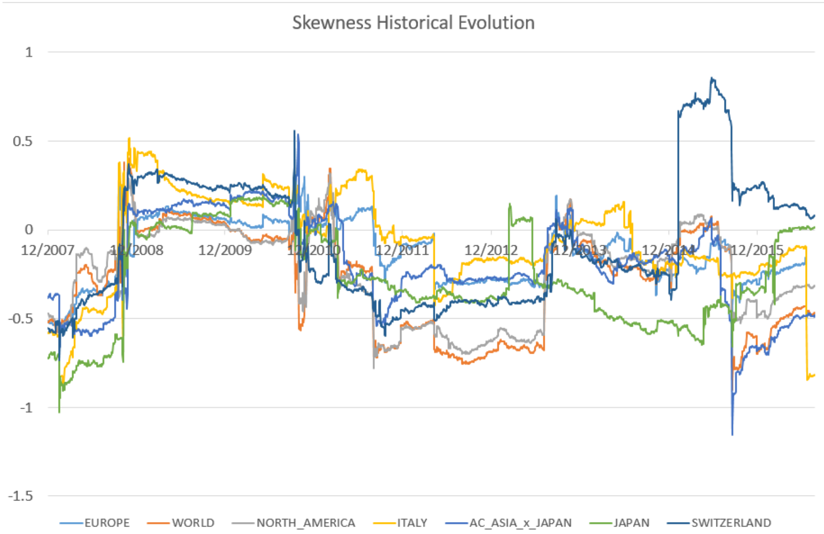


Figure 3.7: Skewness evolution of Indexes log–returns.

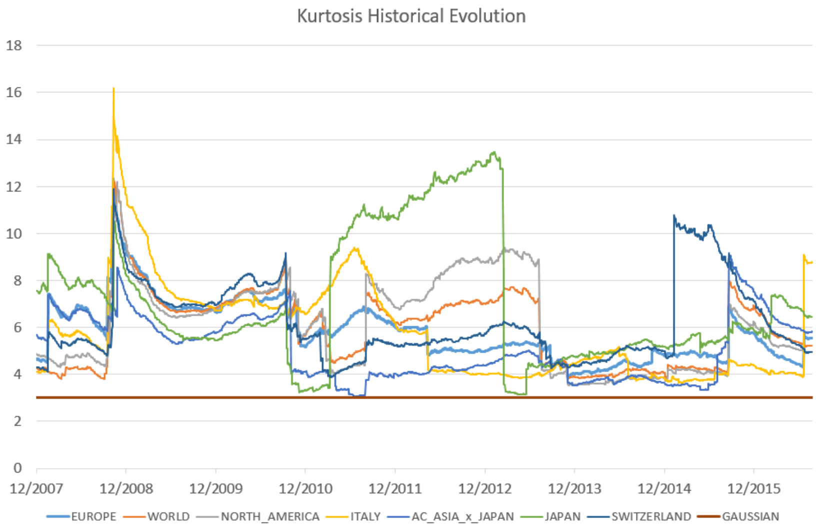


Figure 3.8: Kurtosis evolution of Indexes log–returns.

dence is modeled through frequent updating of the parameters, instead of through a fixed initial estimation of the time-varying evolution law of the parameters, can be a suitable alternative model for forecasting the risk of a multivariate portfolio in different financial situation.

We use the Student-$t$ copula to describe the dependence between residuals of our filtered marginal distributions. We assumed the correlations of residuals to be constant in the model, but since they are affected by evolution in time, their estimation is daily updated by rolling the observations used in estimation. In details, denoted with $t = 0$ the first day of the rolling set of observations, and with $t = T$ the last day, for any pair of assets $i$ and $j$, with $i \neq j$, the correlation between innovations $\xi_{i,T+1}$ and $\xi_{j,T+1}$ given the algebra $\mathfrak{F}_T$ is defined as

$$\rho(\xi_{i,T+1}, \xi_{j,T+1}) = \frac{Cov(\xi_{i,T+1}, \xi_{j,T+1})}{\sqrt{Var(\xi_{i,T+1})\, Var(\xi_{j,T+1})}},$$

where

$$\begin{aligned} Cov(\xi_{i,T+1}, \xi_{j,T+1}) &= E[\xi_{i,T+1}\ \xi_{j,T+1}] - E[\xi_{i,T+1}]\ E[\xi_{j,T+1}] \\ &= E[\sigma_{i,T+1}\ z_{i,T+1}\ \sigma_{j,T+1}\ z_{j,T+1}] \\ &= E[\sigma_{i,T+1}\ \sigma_{j,T+1}]\ E[z_{i,T+1}\ z_{j,T+1}] \\ &= E[\sigma_{i,T+1}\ \sigma_{j,T+1}]\ \rho_{i,j} \end{aligned}$$

and where $\rho_{i,j}$ denotes the constant correlation coefficient between the residuals $z_i$ and $z_j$.
Hence, observing that in terms of 1-day forecasting volatilities are constant given $\mathfrak{F}_T$, it holds:

$$\begin{aligned} \rho(\xi_{i,T+1}, \xi_{j,T+1}) &= \frac{\sigma_{i,T+1}\ \sigma_{j,T+1}\ \rho_{i,j}}{\sigma_{i,T+1}\ \sqrt{Var(z_{i,T+1})}\ \sigma_{j,T+1}\ \sqrt{Var(z_{j,T+1})}} \\ &= \frac{\rho_{i,j}}{\sqrt{Var(z_{i,T+1})\ Var(z_{j,T+1})}}. \end{aligned} \tag{3.2.1}$$

Thus, the dependence between innovations $\xi_{i,T}$ and $\xi_{j,T}$ depends on $\rho_{i,j}$ and on the volatility of residuals, i.e., there is an adjustment of correlation between innovation and residuals when the volatility of residuals is different from one.

It can be seen that correlations among residuals evolve along time, i.e., are not constant, while, on the contrary, correlations of historical returns have low-reacts to the changes in market dependencies. This fact can be clearly observed in the correlation evolution the first financial periods under study: after the crisis, the dependence among assets remains high and stable for almost two years before reaching again normal levels. This fact is related with the equally-weighed information that we get for the last two years of data, while with the residual correlation we have a sort of adjust effect in correlation through the update of the parameters of marginal models.

An empirical evidence is the focus on the evolution of bivariate correlation between the MSCI European and the MSCI Japanese indexes in the 2008-2010 period. Figure 3.9shows the evolution of estimated bivariate correlations between recorded returns and estimated residuals (2-years rolling, hence 504 working days observed in the past). It also shows the correlations between simulated forecasting (based on historical sample of 2-years, 504 working days) and returns (all based on a rolling of 1000 days). Note that it is immediately observable the jump from positive to negative correlation in late 2010. Such instability arises because in late 2010 the historical sample used for the estimation does no more contain the observations of 2008 crisis, and hence the correlation decreases significantly. On the contrary, looking at the evolution of the residuals' correlation, we observe a smoother decrease, with some increase. Hence the one day forecasting simulation in 2009 has smaller correlation than the observed one, but is more representative of what was going on.

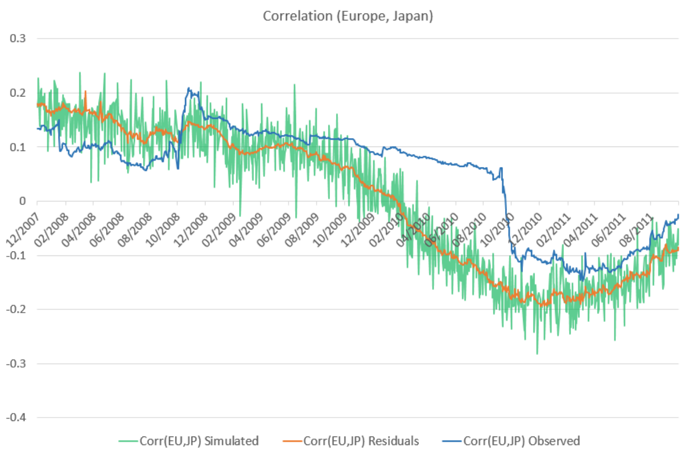


Figure 3.9: Evolution of 2-year correlation between MSCI Europe and MSCI Japan. Historical vs Simulated one, period 2007-2009

This observation is of fundamental importance in the study of the time varying dynamic of dependence, since it is a mitigation of the jumps that occur in historical dynamics of correlation caused by the rolling data-set used. The idea of GARCH models is to focus on more recent data, in order to have a fast reaction of the model to the sudden changes in volatility, and modeling the dependence of GARCH through residuals we recreate a more stable and realistic description of correlation among assets, where the correlation becomes more reactive to the market also for the $X_{i,t}$. This is not the case of Vine-Copulas approach, as we will see in Section 3.3.2.

### 3.2.1 Inference and Simulation

We first make inferences on the marginal models for the log-returns, then we observe the residuals according to the selected univariate model, and we dynamically estimate the parameters of the copula connecting the observed residuals. The resulting marginal models and the estimated copula for the residuals are finally used to evaluate daily portfolio's returns through different measures of risk. For the copula describing the dependence among residuals, we use the Student-$t$ copula (having a number of degrees of freedom and correlation matrix $P$ dynamically estimated along time), since, as we will see in next sections, these parameters are not constants over time, but are stable enough to understand the dynamic of the model.

At the end of the market-day we collect the last 2 years of observed daily log-return (504 observations) for every asset in our portfolio and we use it has historical sample to estimate the 1-day forecasting model. This approach, whose goal is to obtain a tractable tool for forecasting purposes, is based on the following steps.

1. **Model estimation**, which consists in:

   (a) Estimation of the parameters of the univariate models, using the last 2 years of observed daily log-return (504 observations), using conditional maximum likelihood estimation;

   (b) Computation of the empirical correlated residuals according to the marginal univariate models (details are provided next in this section).

   (c) Estimation of the correlation matrix $P$ and the degrees of freedom $\nu$ of the Student-$t$ copula starting from empirical correlated residuals.

2. **Simulation**. Once the marginal model is fixed, a simulation of 1000 possible 1-day forecasting portfolio returns to estimate their distribution is performed. At any step, the simulation consists in:

   (a) generation of a multivariate vector having marginal uniformly distributed variables and the previously estimated Student-$t$ copula as the joint distribution (whose dimension is the number of assets in the portfolio);

   (b) generation of the correlated vector of residual through the *inverse empirical distribution function* (by Kernel function) of the past residuals created during the estimation phase (so that no assumptions on the marginal distribution of residuals is needed);

   (c) generation of new future correlated stock returns using marginal models with estimated parameters, starting from last observed residuals and returns;

   (d) generation of the empirical distribution for the portfolio subject of the study.

3. **Risk valuation**. Common risk factors as the *Value at Risk* (VaR), the *Expected Shortfall* (ES) and the volatility can now be estimated through the empirical joint distribution of the portfolio's returns at 1-day time horizon forecast.

4. **Back-testing**. For this final phase the VaR at 1-day horizon and the realized returns of the next day (out-of-sample back-test), for at least 1 year (250

data), are considered. All the previous phases are repeated for 250/500/750 working days, and back-testing analysis are performed.

For the estimation of the parameters of the copula, historical residuals recreated from assets' marginal models are used. Given the estimated parameters $\mu_i$ and $\varphi_i$ and the observed returns $X_{i,t}$ with $t = 0, ..., T$, one can first recreate the historical innovations as

$$\hat{\xi}_{i,t} = X_{i,t} - \mu_i - \varphi_i \, X_{i,t-1}.$$

Then, given the initial values of variance model $\xi_{i,0}$ and $\sigma_{i,0}$, we can recreate the historical conditional variance time series through

$$\hat{\sigma}^2_{i,t} = K_i + \alpha_i \, \hat{\xi}^2_{i,t-1} + \beta_i \, \hat{\sigma}^2_{i,t-1}, \;\; t = 1, \ldots, T,$$

where the coefficients $K_i$, $\alpha_i$ and $\beta_i$ are those estimated for the marginal models. The resulting historical time series for the residuals is then given by $\hat{z}_{i,t} = \hat{\xi}_{i,t} \;/\; \hat{\sigma}_{i,t}$. Finally, the correlations $\rho_{i,j,T+1} = \rho(z_{i,T+1}, z_{i,T+1})$ between residuals of assets $i$ and $j$ at time $T + 1$ are estimated through the sequences of pairs $\{(\hat{z}_{i,t}, \hat{z}_{j,t}), t = 0, \ldots, T\}$. Also the degree of freedom of the Student-$t$ copula is estimated from this multivariate sample, hence the copula is parametrized.

In order to catch the fat tail and anomalous skewness in different market period we do not impose a parametric distribution for the simulation of residuals, but we used the empirical distribution with a kernel smoothness function. In this way the combined effect between estimated and non constant tail dependence using the Student-$t$ copula and the empirical distribution of residual allow us to forecast a 1-day distribution of future log-returns of a portfolio that quickly catch the changes in market.

### 3.2.2 Back-testing and Comparison with FHS

Once the marginal model for the univariate evolutions of the indexes in the portfolio is fixed, and the copula of the residuals is dynamically estimated, it is possible to provide daily forecast computing the empirical distribution of the portfolio, based on the marginal empirical distributions of the log-returns and the related estimated connecting copula.

In our analysis, the importance of VaR is linked to the possibility of performing back-testing of the model through generation of samples of portfolio's value in the next day, and then to provide the validation of negative tail of future returns. In fact, daily estimation of log-returns by means of the previously described procedure can easily allows for generation of 1000 samples of portfolio's returns, and thus of estimates of risks measures.

In all three cases the forecast passed the back testing according the Basel Regulation, as shown in Figure 3.10.

| Copula-GARCH Model | | | |
|---|---|---|---|
| **Crisis** | **Period** | # Violation VaR 1% | Traffic Light |
| Financial crisis | 2008-2009 | 4 | GREEN |
| European debt crisis | 2010-2012 | 6 | GREEN |
| Chinese stock market turbulence | 2015-2016 | 8 | GREEN |

Figure 3.10: Results of back-testing of GARCH-Copula model according to Market Risk Capital Requirements (Basel) test.

As we saw in Section 2.1.2, this is the Regulatory back-test but it has statistical limits. Hence we focus on Kupiec (Kupiec (1995)) and Christoffersen (Christoffersen (1998)) back-tests.
The Kupiec results are summarized in Figure 3.11, and clearly show the admissibility of the estimated VaR.

| Copula-GARCH Model | | | | | |
|---|---|---|---|---|---|
| **Crisis** | **Period** | #% VaR 1% | #% VaR 5% | P value VaR 1% | P value VaR 5% |
| Financial crisis | 2008-2009 | 0.74% | 4.81% | 47.70% | 16.51% |
| European debt crisis | 2010-2012 | 1.15% | 4.80% | 26.59% | 16.82% |
| Chinese stock market turbulence | 2015-2016 | 1.54% | 6.72% | 74.52% | 91.34% |

Figure 3.11: Results of Kupiec test on GARCH-Copula model for 2008 Financial Crisis (12/2007-12/2009), European Debt Crisis (01/2011-12/2012) and Chinese Turbulence (08/2014-07/2016)

Also the resulting $p$-values for Christoffersen's test under different confidence levels and sample sizes have been computed. The results are summarized in Figure 3.12, and also in this case show the admissibility of the estimated VaR, except for few cases.

| Copula-GARCH Model | | | | | |
|---|---|---|---|---|---|
| **Crisis** | **Period** | #% VaR 1% | #% VaR 5% | P value VaR 1% | P value VaR 5% |
| Financial crisis | 2007-2009 | 0.74% | 4.81% | 35.94% | 75.55% |
| European debt crisis | 2010-2012 | 1.15% | 4.80% | 86.76% | 91.43% |
| Chinese stock market turbulence | 2015-2016 | 1.54% | 6.72% | 92.81% | 97.17% |

Figure 3.12: Results of Christoffersen test on GARCH-Copula model for 2008 Financial Crisis (12/2007-12/2009), European Debt Crisis (01/2011-12/2012) and Chinese Turbulence (08/2014-07/2016)

In fact, in $2014-2016$ the Christoffersen test fail for confidence level of $95\%$

(pointed out in red color). Problems also raised for the European and Italian indexes probably due to the Brexit effect; here the market volatility is increased because of political/exogenous reasons and it is not possible to head off the phenomenon from past observed data, but a risk manager can integrate the qualitative view (increasing of market uncertainty) with the model estimation and VaR. It can prevent some draw-down with coverage instrument or low the European exposure of the portfolio for that period. Of course in our equally-weighted portfolio the Brexit effect affects also the VaR estimation of the overall portfolio.

An interesting comparison can be done with the Parallel Filtered Bootstrapping Approach (PFB).
One relevant difference between the two models is that the GARCH-Copula has a Montecarlo simulation while The PFB model is based in historical simulation (see Section 2.2.2 for details).
Another significant difference is the dependence modeling, in GARCH-Copula we use a semi-parametric approach with the estimation of explicit correlation structure of our portfolio, while in PFB model the correlation is not estimated and is implicitly considered using the parallel bootstrapping.
These two differences have a great impact on the back-testing results, as we can see in Figures 3.13,3.14 and 3.15, this model does not perform well in Crisis periods.

| PFB Model | | | |
|---|---|---|---|
| **Crisis** | **Period** | # Violation VaR 1% | Traffic Light |
| Financial crisis | 2008-2009 | 12 | YELLOW |
| European debt crisis | 2010-2012 | 12 | YELLOW |
| Chinese stock market turbulence | 2015-2016 | 17 | YELLOW |

Figure 3.13: Results of back-testing of PFB model according to Market Risk Capital Requirements (Basel) test.

| PFB Model | | | | | |
|---|---|---|---|---|---|
| **Crisis** | **Period** | #% VaR 1% | #% VaR 5% | P value VaR 1% | P value VaR 5% |
| Financial crisis | 2007-2009 | 2.22% | 7.58% | 98.59% | 98.97% |
| European debt crisis | 2010-2012 | 2.30% | 4.61% | 98.94% | 32.30% |
| Chinese stock market turbulence | 2015-2016 | 3.26% | 8.06% | 99.99% | 99.69% |

Figure 3.14: Results of Kupiec test on PFB model for 2008 Financial Crisis (12/2007-12/2009), European Debt Crisis (01/2011-12/2012) and Chinese Turbulence (08/2014-07/2016)

Another usefull term of comparison between these two models is look at the evolution of estimated VaRs linked to effective returns of the portfolio.
Figures 3.16, 3.18 and 3.20 show the estimated VaR at percentiles 1% and 5%, and the true value of the asset in the previous day for Copula-GARCH model, while

| PFB Model | | | | | |
|---|---|---|---|---|---|
| **Crisis** | **Period** | #% VaR 1% | #% VaR 5% | P value VaR 1% | P value VaR 5% |
| Financial crisis | 2007-2009 | 2.22% | 7.58% | 99.82% | 99.96% |
| European debt crisis | 2010-2012 | 2.30% | 4.61% | 99.99% | 99.99% |
| Chinese stock market turbulence | 2015-2016 | 3.26% | 8.06% | 99.99% | 99.99% |

Figure 3.15: Results of Christoffersen test on PFB model for 2008 Financial Crisis (12/2007-12/2009), European Debt Crisis (01/2011-12/2012) and Chinese Turbulence (08/2014-07/2016)

Figures 3.17, 3.19 and 3.21 show the same information for Parallel Filtered Bootstrap model. The good performance of the estimated VaR is graphically confirmed, even if with quite different behavior; both the risk indicators react quickly to financial crisis, and in 2008 big one in particular, but also in the beginning of 2011 and in late 2015, or in the beginning of 2016, proving that both the models give a reliable estimations of risk also in the *after-crisis* periods.

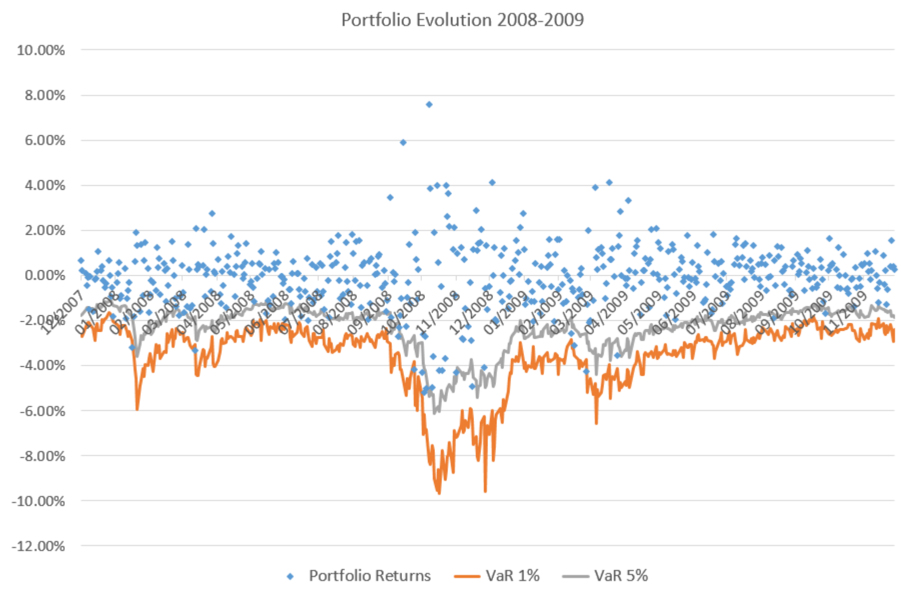


Figure 3.16: Observed log-returns of the portfolio, and estimated (via Copula-GARCH) 1% and 5% corresponding VaR, period from 12-2007 to 12-2009.

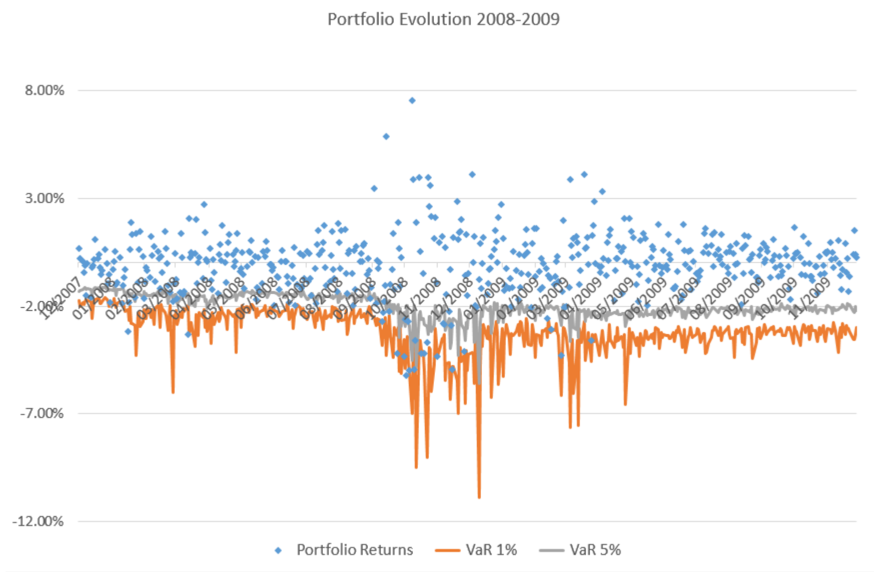


Figure 3.17: Observed log-returns of the portfolio, and estimated (via Parallel Filtered Bootstrap) 1% and 5% corresponding VaR, period from 12-2007 to 12-2009.

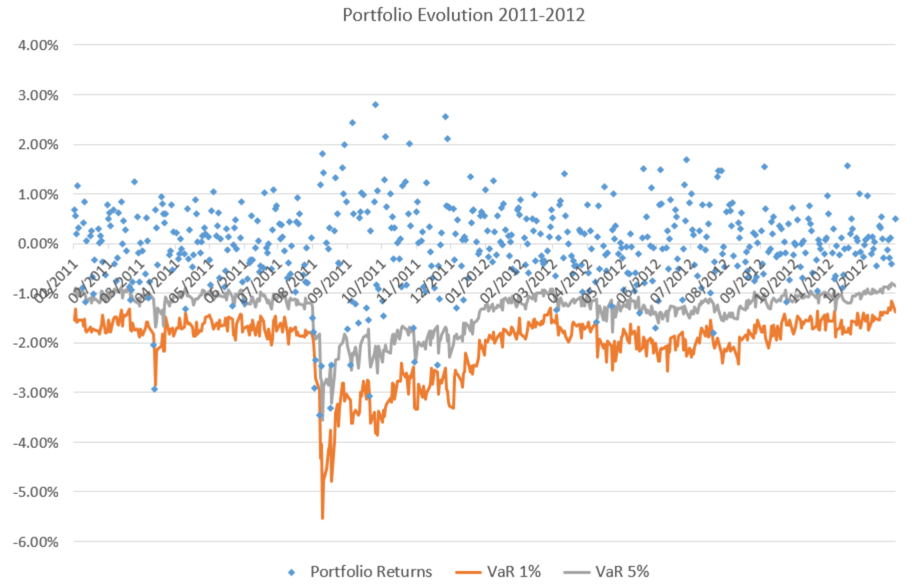


Figure 3.18: Observed log-returns of the portfolio, and estimated (via Copula-GARCH) 1% and 5% corresponding VaR, period from 01-2011 to 12-2012.

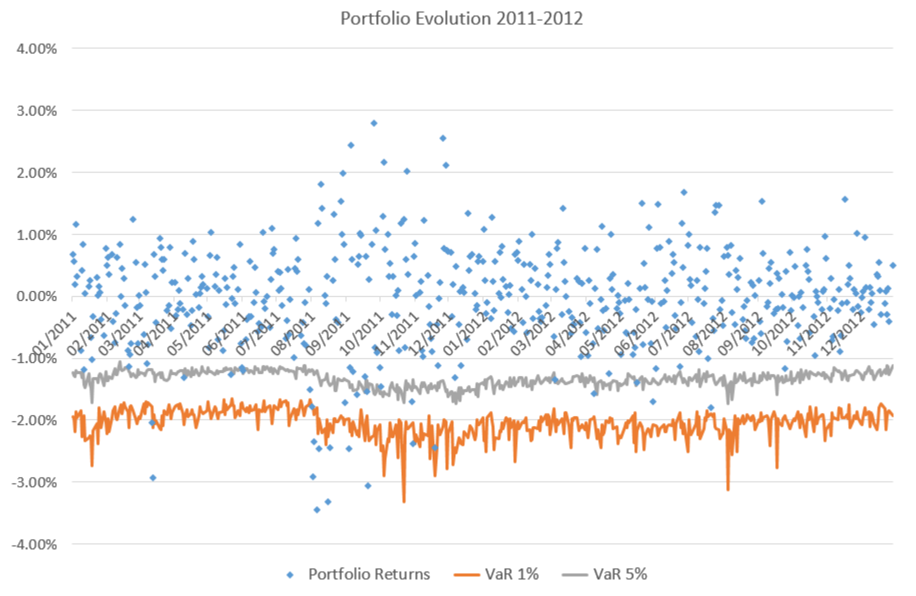


Figure 3.19: Observed log-returns of the portfolio, and estimated (via Parallel Filtered Bootstrap) 1% and 5% corresponding VaR, period 01-2011 to 12-2012.

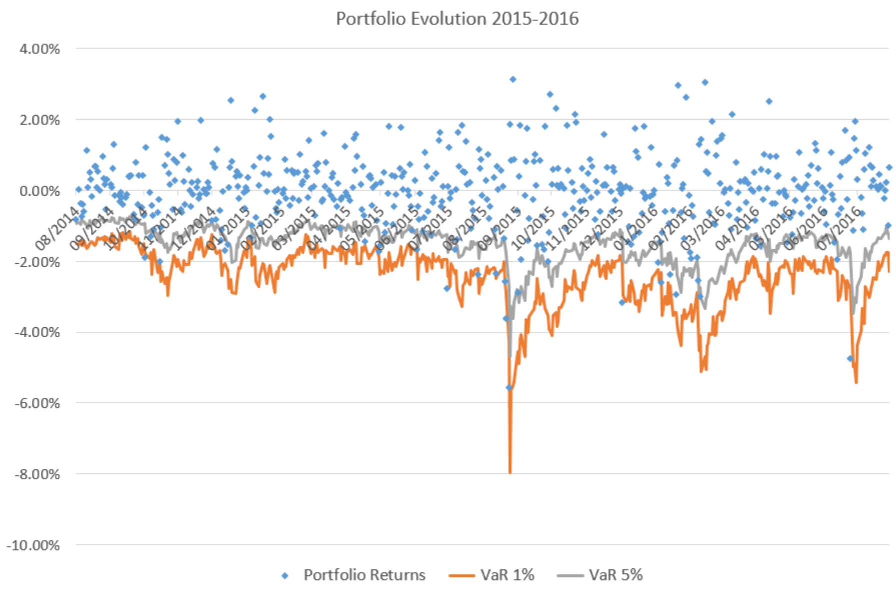


Figure 3.20: Observed log-returns of the portfolio, and estimated (via Copula-GARCH) 1% and 5% corresponding VaR, period 08-2014 to 07-2016.

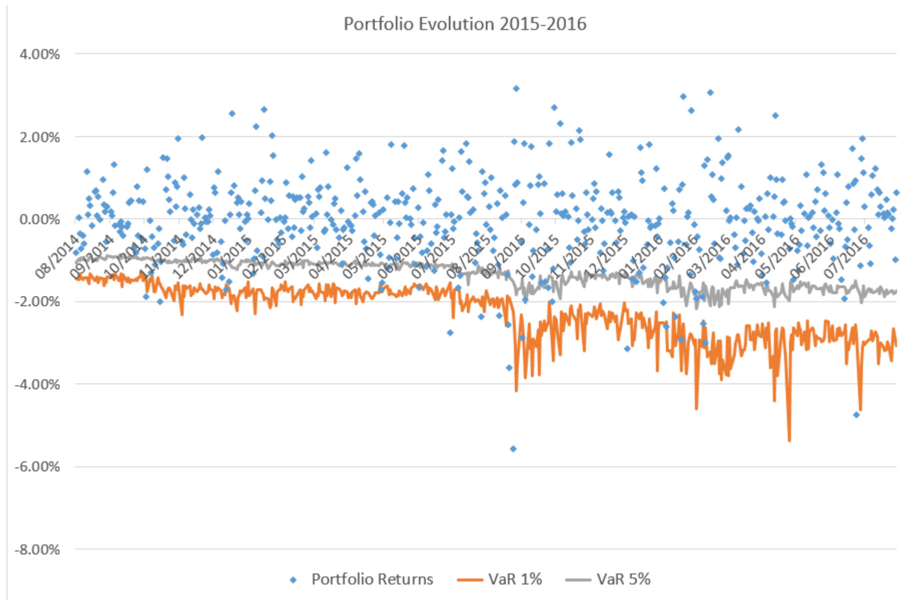


Figure 3.21: Observed log-returns of the portfolio, and estimated (via Parallel Filtered Bootstrap) 1% and 5% corresponding VaR, period 08-2014 to 07-2016.

A relevant difference between the two models is their stability despite the fast-reactivity to market conditions. The Copula-GARCH model even in daily forecasting has a more stable adjustment with respect to the Parallel Filtered Bootstrap one. It is undeniable if we compare the different VaR evolution in all time series considered:

- In Global Financial Crisis of 2008 the Copula-GARCH model reacts almost instantaneously (with a delay maybe of 1 day) to the Crisis Beginning and identifies the first huge negative returns as Crisis (VaR remain high, in absolute term, for week). Furthermore, it recognizes fast the intermediate recovery in April 2009 and the end of the Crisis in summer of 2009 (Figure 3.16) with less conservative VaR. In Figure 3.17 we see that the reaction of PFB model is quite precise but the rebound of the market returns determines the presence of fluctuations on its values. This means an excessive sensibility of the model to new positive observations and a lack of stability.

- In European debt Crisis the situation is quite different, in this case only Copula-GARCH model reacts at appropriate time, while Parallel Filtered Bootstrap is too slow, the reason is that the data sample used for September 2011 start at September 2009, hence the data used in historical forecasting came from a positive market period, with not enough 'tail event' to predict a Crisis. Indeed the VaR becomes more conservative (higher in absolute value) after some large negative returns occur and becomes part of the data sample used for the forecasting.
  In this the Parallel Filtered Bootstrap Model is also slow in understanding the change in correlation between assets. Using the historical implicit correlation it takes a while to capture from the data the comonotonicity typical of financial Crisis, while in previous analysis we showed that Copula-GARCH model reacts faster in describing correlation between assets and identifying comonotonicity events.

- The analysis of the last 2 years confirms what anticipated in the first item. Copula-GARCH is quick to understand the end of high-volatility financial period. In these years there were some fast 'mini-crisis', the beginning of 2015, the Summer of 2015, the beginning of 2016 and the Brexit event, but in Figure 3.20 the evolution of VaR identifies all these periods. In Figure 3.21 , instead, the PFB model does not recognized that there is a period between 'mini-crisis' of normal financial situation. This difficulty is determined by the presence in historical data of negative returns related with high volatility periods. The results is that with Parallel Filtered Bootstrap we predict too conservative VaRs.

## 3.3 Linkages

In this section we introduce an application of Linkages, instruments similar to Copulas presented in Section 1.2.7. As already underlined, Linkages emphasize roles of dependence structure between given marginals, and dependence structure within each of them, like copula, but overcome some limitation that the copula function cannot handle. We use empirical distribution function aiming to comparing the evolution and back-testing results with the similar approach of D-Vines Copulas, but different multivariate parametric distributions could be used for each group of random variables.

### 3.3.1 Inference and Simulation of Linkages model

As for previous analysis we call $X_i$ the vector of 504 returns of title *i*, $i = 1, \ldots, 7$. Their joint distribution will be denoted as $F$ and the marginals $F_i$, $i = 1, \ldots, 7$. With some specific tests (Anderson-Darling test or Jacque-Bera or Kolmogorov-Smirnov) it's possible to study if $X_i$ is distributed as a normal random variable. In section 3.1.2 we see that our data in general are not Gaussian distributed but in this case, only for the construction of the Linkages function and the linear correlation structure, we assume the Gaussianity of all variables. However in the simulation phase we will proceed as for GARCH-Copula with a Montecarlo simulation of empirical distribution of all assets.

We divided the $X_i$ in three groups, the improvement of this method is that we can choose how to group the variables if we have financial constrains. In this case we decide to group Europe Area Indexes (MSCI Europe, MSCI Italy and MSCI Switzerland), the rest of the world (MSCI World, MSCI North America and MSCI Asia ex Japan) and MSCI Japan in the last group, alone because the Japan financial evolution is pretty different from others.
Each variable has mean $\mu_i$, $i = 1, \ldots, 7$ and their covariance matrix is $\Sigma_X$.
For an easier computation we standardize the $X_i$, $i = 1, \ldots, 7$, variables and we rename them in $Z_i$:

$$Z_i = \frac{X_i - \mu_i}{\sigma_i}. \tag{3.3.1}$$

The covariance matrix of $Z_i$, $i = 1, \ldots, 7$ is the correlation matrix of $X_i$, $i = 1, \ldots, 7$, hence $\Sigma_Z = \rho_X$.
For example Table 3.1 shows the correlation matrix of portfolio calculated with returns observed in $25/09/2012 - 29/08/2014$.

For each group we calculate the conditional variables $Y_i$, $i = 1, \ldots, 7$, in this way the $Y_i$ of the same group are independent from each other:

Table 3.1: Historical Spearmann correlation of X, equal to the covariance matrix of Z estimated with data sample $25/09/2012 - 29/08/2014$

| | | | | | | |
|---|---|---|---|---|---|---|
| 1.0000 | 0.7814 | 0.8012 | 0.7641 | 0.5509 | 0.3519 | 0.0996 |
| 0.7814 | 1.0000 | 0.5095 | 0.5484 | 0.3920 | 0.1665 | -0.0198 |
| 0.8012 | 0.5095 | 1.0000 | 0.5821 | 0.3639 | 0.3117 | 0.1872 |
| 0.7641 | 0.5484 | 0.5821 | 1.0000 | 0.9191 | 0.4845 | 0.3056 |
| 0.5509 | 0.3920 | 0.3639 | 0.9191 | 1.0000 | 0.3267 | 0.0794 |
| 0.3519 | 0.1665 | 0.3117 | 0.4845 | 0.3267 | 1.0000 | 0.4427 |
| 0.0996 | -0.0198 | 0.1872 | 0.3056 | 0.0794 | 0.4427 | 1.0000 |

$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \\ Y_5 \\ Y_6 \\ Y_7 \end{pmatrix} = \begin{pmatrix} Z_1 \\ Z_2|Z_1 \\ Z_3|Z_1, Z_2 \\ Z_4 \\ Z_5|Z_4 \\ Z_6|Z_4, Z_5 \\ Z_7 \end{pmatrix} \tag{3.3.2}$$

To simplify our computations we re-standardize these conditional variables and rename them according with their belonging to the three groups in $M = (M_1, M_2, M_3)$ with joint distribution function $F_M$, $K = (K_4, K_5, K_6)$ with joint distribution $F_K$ and $Y_7$.

$$\begin{pmatrix} M_1 \\ M_2 \\ M_3 \\ K_4 \\ K_5 \\ K_6 \\ Y_7 \end{pmatrix} = \begin{pmatrix} Y_1 \\ \frac{Y_2 - \mu_{2|1}}{\sigma_{2|1}} \\ \frac{Y_3 - \mu_{3|1,2}}{\sigma_{3|1,2}} \\ Y_4 \\ \frac{Y_5 - \mu_{5|4}}{\sigma_{5|4}} \\ \frac{Y_6 - \mu_{6|4,5}}{\sigma_{6|4,5}} \\ Y_7 \end{pmatrix} \tag{3.3.3}$$

The linkage function, then, is the joint distribution function $L$ of

$$\begin{pmatrix} U_1 \\ U_2 \\ U_3 \\ V_4 \\ V_5 \\ V_6 \\ W_7 \end{pmatrix} = \begin{pmatrix} \Phi(M_1) \\ \Phi(M_2) \\ \Phi(M_3) \\ \Phi(K_4) \\ \Phi(K_5) \\ \Phi(K_6) \\ \Phi(Y_7) \end{pmatrix} \tag{3.3.4}$$

with $U_1, U_2, U_3$ independent uniform $[0, 1]$ random variables and $V_4, V_5, V_6$ too. Note that the two groups are not independent within each other and are not independent from $W_7$. In order to get some insight about the distribution $L$, we recall that it is the linkage of $[M_1, M_2, M_3, K_4, K_5, K_6, Y_7]$ and according with our hypothesis it is a 7-variate normal random variable with parameters:

$$\mu_L = (0, 0, 0, 0, 0, 0, 0)$$

$$\Sigma_L = \begin{pmatrix} 1 & 0 & 0 & \mathrm{C}(M_1, K_4) & \mathrm{C}(M_1, K_5) & \mathrm{C}(M_1, K_6) & \mathrm{C}(M_1, K_7) \\ 0 & 1 & 0 & \mathrm{C}(M_2, K_4) & \mathrm{C}(M_2, K_5) & \mathrm{C}(M_2, K_6) & \mathrm{C}(M_2, K_7) \\ 0 & 0 & 1 & \mathrm{C}(M_3, K_4) & \mathrm{C}(M_3, K_5) & \mathrm{C}(M_3, K_6) & \mathrm{C}(M_3, K_7) \\ \mathrm{C}(K_4, M_1) & \mathrm{C}(K_4, M_2) & \mathrm{C}(K_4, M_3) & 1 & 0 & 0 & 0 \\ \mathrm{C}(K_5, M_1) & \mathrm{C}(K_5, M_2) & \mathrm{C}(K_5, M_3) & 0 & 1 & 0 & 0 \\ \mathrm{C}(K_6, M_1) & \mathrm{C}(K_6, M_2) & \mathrm{C}(K_6, M_3) & 0 & 0 & 1 & 0 \\ \mathrm{C}(Y_7, M_1) & \mathrm{C}(Y_7, M_2) & \mathrm{C}(Y_7, M_3) & \mathrm{C}(Y_7, K_4) & \mathrm{C}(Y_7, K_5) & \mathrm{C}(Y_7, K_6) & 1 \end{pmatrix}$$

The variance/covariance matrix calculation is not difficult, but it is time consuming because of the presence of conditional variables. For example

$$\mathrm{C}(M_1, K_4) = \mathrm{C}(Y_1, Y_4) = \mathrm{C}(Z_1, Z_4) = \rho_X(1, 4)$$

$$\begin{aligned} \mathrm{C}(M_1, K_5) &= E[M_1\ K_5] - E[M_1]\ E[K_5] \\ &= E[Z_1\ \frac{Z_5 - \rho_{X_{(4,5)}} * Z_4}{\sqrt{1-\rho^2_{X_{(4,5)}}}}] \\ &= E[\frac{Z_1 * Z_5 - \rho_{X_{(4,5)}} * Z_4 * Z_1}{\sqrt{1-\rho^2_{X_{(4,5)}}}}] \\ &= \frac{E[Z_1 * Z_5] - E[Z_4 * Z_1] * \rho_{X_{(4,5)}}}{\sqrt{1-\rho^2_{X_{(4,5)}}}} \\ &= \frac{\rho_{X_{(1,5)}} - \rho_{X_{(4,1)}} * \rho_{X_{(4,5)}}}{\sqrt{1-\rho^2_{X(4,5)}}} \end{aligned} \tag{3.3.5}$$

In analogous way we calculate the remaining conditional correlations. The corresponding dependence parameter of Table 3.1 (the correlation matrix on $25/09/2012-29/08/2014$), for example , is 3.2:

At this point, we need to generate $u_1, u_2, u_3, v_4, v_5, v_6, w_7$ random uniform. Note that we hypothesize the independence within the groups but not the independence between groups. It results easier to parametrize the copula using the correlation, this can be done using elliptical Copulas function with parameters the estimated correlation matrix (and in case of T-Copula a specific degree of freedom).

Having $u_1, \ldots, v_6, w_7$, we generate $\hat{y}_1, \ldots, \hat{y}_7$ distributed as the performances

Table 3.2: Correlation matrix of $[M_1, M_2, M_3, K_4, K_5, K_6, Y_7]$ calculated on $25/09/2012 - 29/08/2014$ period.

| | | | | | | |
|---|---|---|---|---|---|---|
| 1.0000 | 0.0000 | 0.0000 | 0.7641 | -0.3840 | -0.1629 | 0.0996 |
| 0.0000 | 1.0000 | 0.0000 | -0.0779 | 0.0253 | -0.1564 | -0.1564 |
| 0.0000 | 0.0000 | 1.0000 | -0.0784 | -0.2142 | 0.1085 | -0.2751 |
| 0.7641 | -0.0779 | -0.0784 | 1.0000 | 0.0000 | 0.0000 | 0.3056 |
| -0.3840 | 0.0253 | -0.2142 | 0.0000 | 1.0000 | 0.0000 | -0.5112 |
| -0.1629 | -0.1564 | 0.1085 | 0.0000 | 0.0000 | 1.0000 | 0.1715 |
| 0.0996 | -0.1564 | -0.2751 | 0.3056 | -0.5112 | 0.1715 | 1.0000 |

$y_1, \ldots, y_7$ of our portfolio, i.e. recalling the definition of $u_1, \ldots, v_6, w_7$:

$$
\begin{aligned}
u_1 &= F_1(y_1) \\
u_2 &= F_{2|1}(y_2|y_1) \\
u_3 &= F_{3|1,2}(y_3|y_1, y_2) \\
v_4 &= F_4(y_4) \\
v_5 &= F_{5|4}(y_5|y_4) \\
v_6 &= F_{6|4,5}(y_6|y_4, y_5) \\
w_7 &= F_7(y_7)
\end{aligned}
$$

Inverting, in analogous way used for copula function but with conditional distributions:

$$
\begin{aligned}
\hat{y}_1 &= F_1^{-1}(u_1) \\
\hat{y}_2 &= F_{2|1}^{-1}(u_2|\hat{y}_1) \\
\hat{y}_3 &= F_{3|1,2}^{-1}(u_3|\hat{y}_1, \hat{y}_2) \\
\hat{y}_4 &= F_4^{-1}(v_4) \\
\hat{y}_5 &= F_{5|4}^{-1}(v_5|\hat{y}_4) \\
\hat{y}_6 &= F_{6|4,5}^{-1}(v_6|\hat{y}_4, \hat{y}_5) \\
\hat{y}_7 &= F_7^{-1}(w_7)
\end{aligned}
$$

At this point becomes evident of the possibility to use different multivariate distribution functions. We can use $F_{1,2,3}$ multivariate distribution for the first group and $G_{4,5,6}$ multivariate distribution for the second one, without limitation.
But to compare the results with D-Vine copulas and to avoid the use of parametric marginal models we prefer to use marginal and conditional empirical distribution functions $F_1, F_{2|1}, F_{3|1,2}, F_4, F_{5|4}, F_{6|4,5}, F_7$. As for the GARCH-Copula model, to estimate them, we use the *kernel density estimation*.

Doing iteratively these passages we get the n-simulated 1-day forecast correlated asset returns of our portfolio. (The distribution of this forecasting is given in the next Section 3.3.2).

### 3.3.2 Linkages back-testing and Comparison with D-vine copulas

As in Section 3.2.2, also for this model we perform some back-testing in the three analyzed periods.
In all three cases the Linkages forecast passed the back testing according the Basel Regulation, as shown in Table 3.22.

| Linkages Model | | | |
|---|---|---|---|
| Crisis | Period | # Violation VaR 1% | Traffic Light |
| Financial crisis | 2008-2009 | 14 | YELLOW |
| European debt crisis | 2010-2012 | 8 | GREEN |
| Chinese stock market turbulence | 2015-2016 | 9 | GREEN |

Figure 3.22: Results of back-testing of Linkages model according to Market Risk Capital Requirements (Basel) test.

As we mentioned in Section 2.1.2, this back-test is the Regulatory one but it has shortcut in statistical background hence we focus on Kupiec (Kupiec (1995)) and Christoffersen (Christoffersen (1998)) back-test.
The Kupiec results are summarized in Table 3.23, and show the admissibility of the estimated VaR.

| Linkages Model | | | | | |
|---|---|---|---|---|---|
| Crisis | Period | #% VaR 1% | #% VaR 5% | P value VaR 1% | P value VaR 5% |
| Financial crisis | 2007-2009 | 2.65% | 6.99% | 99.80% | 93.76% |
| European debt crisis | 2010-2012 | 1.54% | 3.26% | 74.52% | 94.73% |
| Chinese stock market turbulence | 2015-2016 | 1.73% | 5.76% | 86.96% | 85.58% |

Figure 3.23: Results of Kupiec test on Linkages model for 2008 Financial Crisis (12/2007-12/2009), European Debt Crisis (01/2011-12/2012) and Chinese Turbulence (08/2014-07/2016)

Table 3.23 confirms the good performance of the method through back-test. The only exception is given by the values obtained in the period 2007-2009 for Var 1%. This period included financial crisis of 2008 that is not well recognized. This fact, together with other observations that we will introduce in the following, is determined by slower reaction times of this model with respect to the AR-GARCH models. A cause of this difficulty should be recognized in the use of the empirical distribution in alternative to the previously used filter.
This is confirmed by the resulting $p$-values for Christoffersen's test under different confidence levels and sample sizes. Looking at the results summarized in Table

3.24 it is clear that the model has the right number of exceeded but in the first two analyzed periods they are not independent, while they become independent in last period. This happens because in the first two cases the negative returns are gathered in few close-range days the empirical marginal distribution is adapting to the new situation without the introduction of a filter. This fact slows down in 2015-16 period, otherwise, there were a number of mini-crisis distant in time and the model works well.

| Linkages Model | | | | | |
|---|---|---|---|---|---|
| **Crisis** | **Period** | #% VaR 1% | #% VaR 5% | P value VaR 1% | P value VaR 5% |
| Financial crisis | 2007-2009 | 2.65% | 6.99% | 99.90% | 99.90% |
| European debt crisis | 2010-2012 | 1.54% | 3.26% | 99.90% | 99.90% |
| Chinese stock market turbulence | 2015-2016 | 1.73% | 5.76% | 94.68% | 85.58% |

Figure 3.24: Results of Christoffersen test on Linkages model for 2008 Financial Crisis (12/2007-12/2009), European Debt Crisis (01/2011-12/2012) and Chinese Turbulence (08/2014-07/2016)

An interesting comparison can be done with the Vine Copulas (see previous Section 2.2.4).

We simulate the D-Vine copulas instead of the C-Vines because in this approach, studying the correlation matrix of the variables, we can select the path, with nodes corresponding to the variables, that maximize the sum of the absolute values of $\tau$ of Kendall. This procedure corresponds to the resolution of the famous problem of the traveling salesman ((Lawler, 1985)), that is the problem of the choice of the minimum path (in this specific case we are looking for the maximum one).
To solve the problem we use the *nearest neighbor* algorithm. This operation is very simple: we choose a node as the beginning of the journey and at each step we consider the following node as the node at maximum distance, i.e. the node with highest correlation value. The main advantage presented by nearest neighbor is that the choice of the path is automatically processed. This path corresponds to the first level tree. The subsequent trees are determined in an analogous way.

We use this method to model the dependence structure of our portfolio and then to simulate correlated uniform random variables used to end up with the forecast of each portfolio assets using empirical distribution.

The relevant difference between the two models is the dependence construction: Vine Copulas construct overall dependence starting from bivariate copulae, selecting the best copula at each step, while with Linkages we focus on specific qualitative groups that can have dimension higher than two.
Another significant difference is that in Vine copula the composite dependence

structure is 'hidden' (we use the best-fitting bivariate copulas at each step but we loose the interpretability of the results), while, with this alternative method, we can see the interaction between groups in Linkages case.

These two differences have a great impact to the back-testing results, as illustrated in Figures 3.25, 3.26 and 3.27. Unfortunately this model does not perform well in Crisis periods.

| D-Vine Model | | | |
|---|---|---|---|
| **Crisis** | **Period** | # Violation VaR 1% | Traffic Light |
| Financial crisis | 2008-2009 | 15 | YELLOW |
| European debt crisis | 2010-2012 | 9 | GREEN |
| Chinese stock market turbulence | 2015-2016 | 17 | YELLOW |

Figure 3.25: Results of back-testing of D-Vines model according to Market Risk Capital Requirements (Basel) test.

| D-Vine Model | | | | | |
|---|---|---|---|---|---|
| **Crisis** | **Period** | #% VaR 1% | #% VaR 5% | P value VaR 1% | P value VaR 5% |
| Financial crisis | 2007-2009 | 2.77% | 8.69% | 99.93% | 99.97% |
| European debt crisis | 2010-2012 | 1.73% | 5.18% | 15.06% | 86.96% |
| Chinese stock market turbulence | 2015-2016 | 3.26% | 9.02% | 99.99% | 99.99% |

Figure 3.26: Results of Kupiec test on D-Vines model for 2008 Financial Crisis (12/2007-12/2009), European Debt Crisis (01/2011-12/2012) and Chinese Turbulence (08/2014-07/2016)

| D-Vine Model | | | | | |
|---|---|---|---|---|---|
| **Crisis** | **Period** | #% VaR 1% | #% VaR 5% | P value VaR 1% | P value VaR 5% |
| Financial crisis | 2007-2009 | 2.77% | 8.69% | 99.99% | 99.99% |
| European debt crisis | 2010-2012 | 1.73% | 5.18% | 99.99% | 99.99% |
| Chinese stock market turbulence | 2015-2016 | 3.26% | 9.02% | 99.99% | 99.99% |

Figure 3.27: Results of Christoffersen test on D-Vines model for 2008 Financial Crisis (12/2007-12/2009), European Debt Crisis (01/2011-12/2012) and Chinese Turbulence (08/2014-07/2016)

Another useful term of comparison between these two models considers the evolution of estimated VaRs linked to effective returns of the portfolio.
Figures 3.28, 3.29,3.30, 3.31,3.32 and 3.33 show the evolution of estimated VaR at percentiles $1\%$ and $5\%$ for Linkages model and D-Vine one and the true value of the asset in the previous day.
A relevant difference between these two models is stability despite the fast-reactivity to market conditions. The Linkages model has a more stable adjustment with respect to the D-Vine one.
Looking at the graphs, the intuition is also that Linkages model reacts faster than D-Vine one in case of comonotonicity events (when the 2008 and 2011 crisis started the Linkages model adjusted the estimation correctly, while D-Vine one had some reverse estimations from one day to the next one).

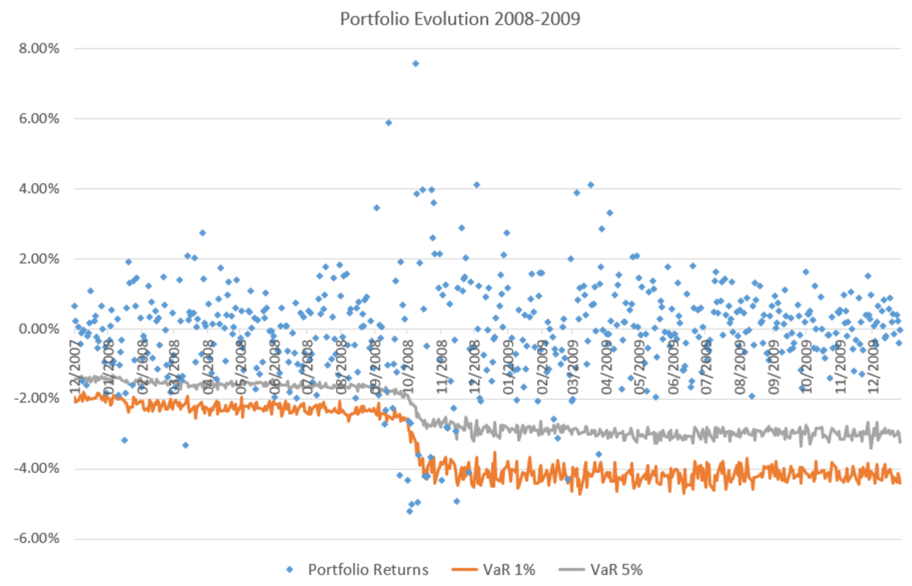


Figure 3.28: Observed log-returns of the portfolio, and estimated (via Linkages) 1% and 5% corresponding VaR, period from December 2007 to December 2009.

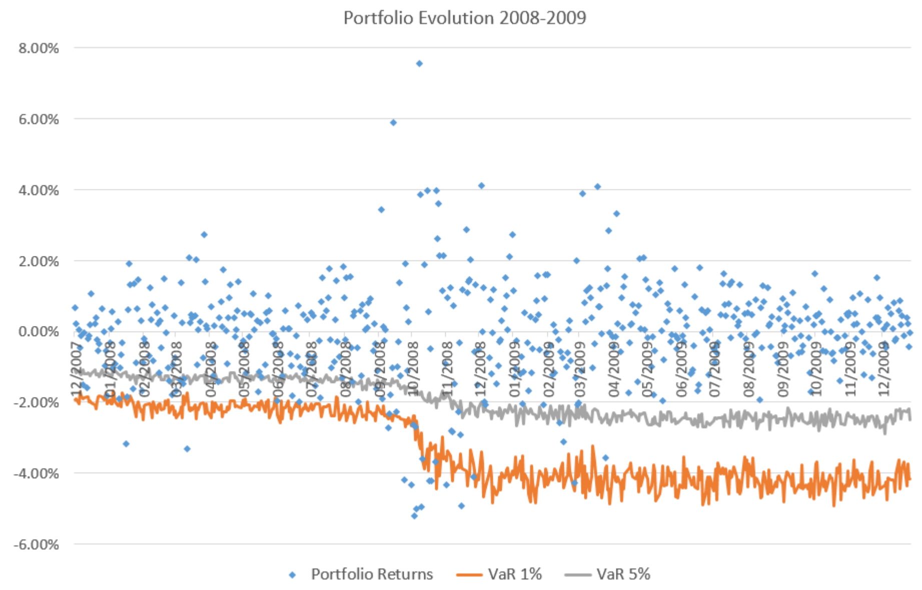


Figure 3.29: Observed log-returns of the portfolio, and estimated (via D-Vines) 1% and 5% corresponding VaR, period December 2007 to December 2009.

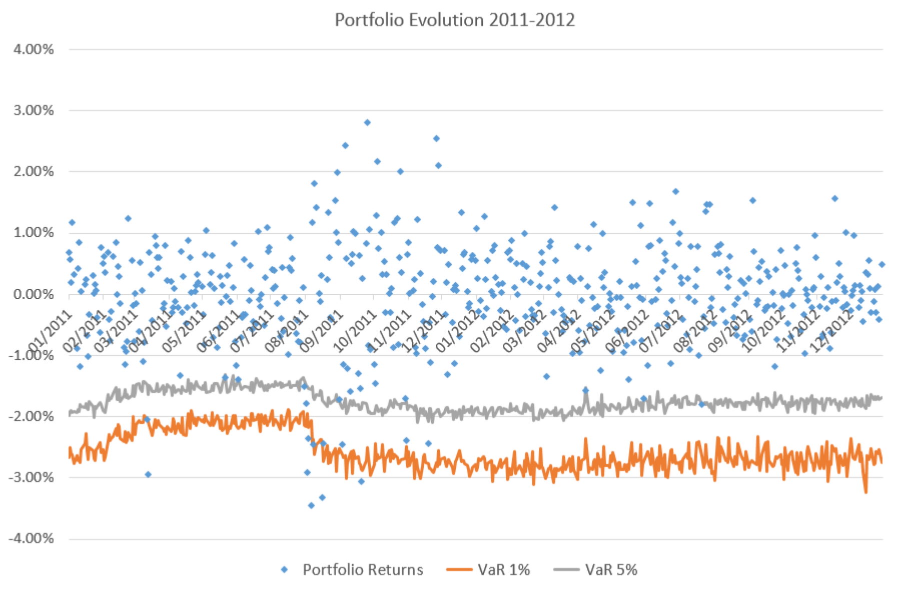


Figure 3.30: Observed log-returns of the portfolio, and estimated (Linkages) 1% and 5% corresponding VaR, period from January 2011 to December 2012.

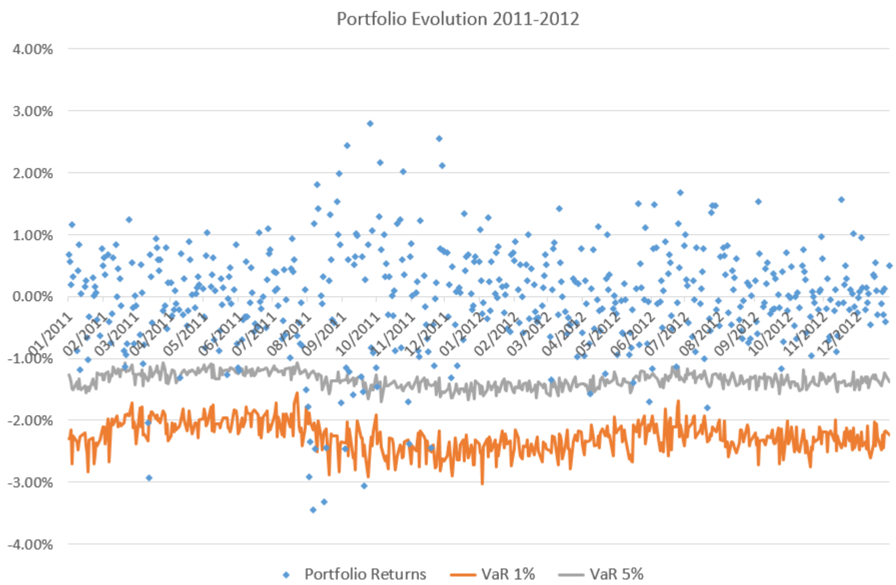


Figure 3.31: Observed log-returns of the portfolio, and estimated (D-Vines) 1% and 5% corresponding VaR, period January 2011 to December 2012.

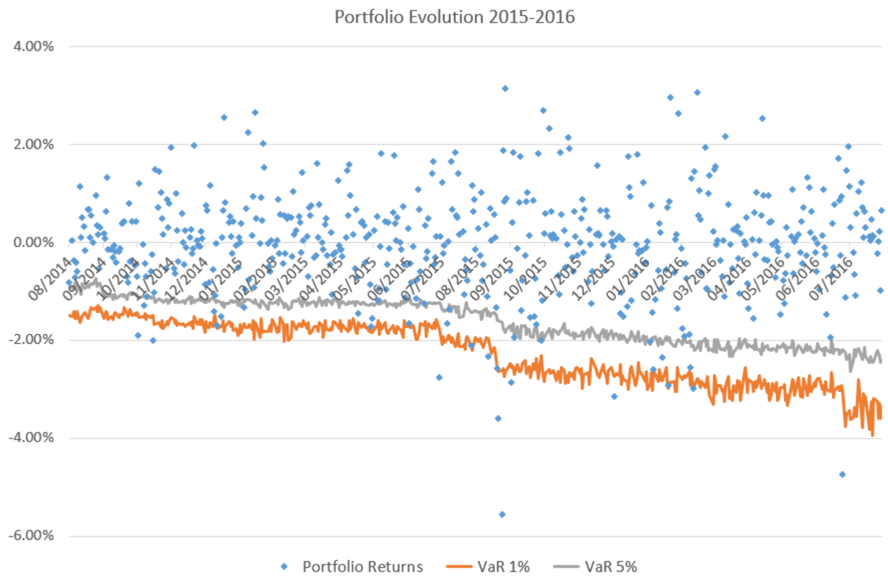


Figure 3.32: Observed log-returns of the portfolio, and estimated (Linkages) 1% and 5% corresponding VaR, period August 2014 to July 2016.

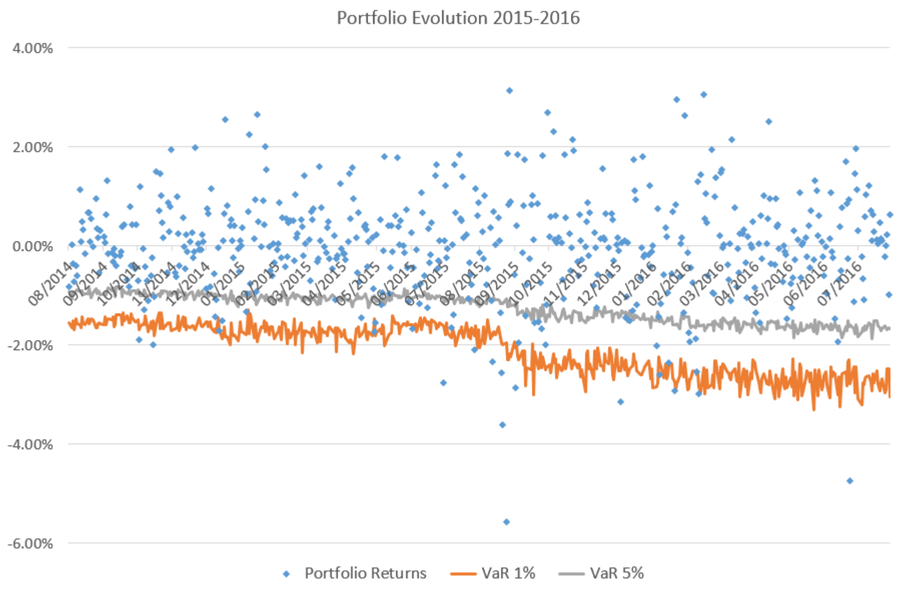


Figure 3.33: Observed log-returns of the portfolio, and estimated (D-Vines) 1% and 5% corresponding VaR, period August 2014 to July 2016.

# Conclusions

In this thesis we present two forecasting models focused on innovative tools to deal with dependence analysis of portfolio.
The performed analysis seem to suggest that the Copula-GARCH model is appropriate in periods of big financial crisis; it appears to quickly reacts at beginning of crisis but also, equally important, also intercepts in right time the end of crisis. Sometime, however, the filter does not intercept mini-crisis, as we saw in 2015-16 back-test.
In particular, the forecasting of European and Italian indexes fail probably because of the Brexit effect. In this case, the market volatility increases because of political/exogenous reasons and it is not possible to head off the phenomenon from past observed data. However, a risk manager can integrate the model and VaR estimation with qualitative views, preventing draw-downs with coverage instrument or decreasing the European exposure of portfolio. Of course, in our equally-weighted portfolio the Brexit effect also affects the VaR estimation of the overall portfolio. The problem seems to be overcome by using the Linkages model, which better performs in this period.
Linkages model has the right number of exceeded, but it should be pointed out that in the first two analyzed periods the exceeding are not independent. The reason is that in the first two cases the negative returns are gathered in few close-range days, hence the empirical marginal distribution adapts too slowly to the new situations without a filter. In 2015-16 period, otherwise, there were a number of mini-crisis not too close in time, that allows for a correct fat tail empirical distribution estimate.
The slower reaction on negative returns in the second model is caused by the use of empirical distribution function without the application of the AR-GARCH filter, hence this is a limit due to univariate projections rather than lacks in dependence modeling.
Back-tests and graphical analysis confirm that both models suitably intercept in right time the changes in dependence structure. They also confirm the superiority of ARMA-GARCH models in describing the univariate forecasting of asset returns with respect to an approach based directly on historical simulations, empirical distribution or bootstrapping.

# References


Aas, K., Czado, C., Frigessi, A., & Bakken, H. (2009). Pair-copula constructions of multiple dependence. *Insurance, mathematics and Economics*, *22*(2), 182–198.

Artzner, P., Delbaen, F., Eber, J., & Heath, D. (1999). Coherent Measures of Risk. *Mathematical Finance*, *9*(3), 203-228.

Barone-Adesi, G., Giannopoulos, K., & Vosper, L. (1999). VaR without Correlations for Non-Linear Portfolios. *Journal of Futures Markets*, *19*, 583–602.

Bauwens, L., Laurent, S., & Rombouts, J. V. K. (2006). Multivariate GARCH models: a survey. *Journal of Applied Econometrics*, *21*(1), 79–109.

Bedford, T., & Cooke, R. (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of mathematics and artificial intelligence*, *32*(1-4), 245–268.

Bedford, T., & Cooke, R. (2002). Vines - a new graphical model for dependent random variables. *Annals of statistics*, *30*(4), 1031–1068.

Bera, A. K., & Higgins, M. L. (1995). On ARCH models: properties, estimating and testing. *Surveys on Econometrics*.

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

Bollerslev, T., Chou, R., & Kroner, K. F. (1992). ARCH modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics*, *52*(1-2), 5-59.

Bollerslev, T., Engle, R. F., & Nelson, D. B. (1994). ARCH model. *Handbook of Econometrics*, *IV*, 2961 3031.

Bos, R., De Waele, S., & Broersen, P. M. T. (2002). Autoregressive spectral estimation by application of the burg algorithm to irregularly sampled data. *IEEE Transactions on Instrumentation and Measurement*(51 (6)).

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (1994). *Time Series Analysis: Forecasting and Control.* (E. Cliffs, Ed.). Prentice Hall.

Brockwell, P. J., Dahlhaus, R., & Trindade, A. A. (2005). Modified Burg Algorithms for Multivariate Subset Autoregression. *Statistica Sinica*(15), 197–213.

Brooks, C. (2014). *Introductory Econometrics for Finance*. Cambridge University Press.

Burg, J. P. (1968). A new analysis technique for time series data. *NATO Advanced Study Institute on Signal Processing with Emphasis on Underwater Acoustics*.

Butler, J. S., & Schachter, B. (1996). Improving Value-At-Risk Estimates By Combining Kernel Estimation With Historical Simulation. *EconWPA - Finance*(9605001).

Campbell, J., Lo, A., & A.C., M. (1997). *The econometrics of financial markets*. Princeton University.

Campbell, S. D. (2007). A review of backtesting and backtesting procedures. *The Journal of Risk*, 1–17.

Chernick, M. R., Gonzalez-Manteiga, W., Crujeiras, R. M., & Barrios, E. B. (2011). Bootstrap methods. *International Encyclopedia of Statistical Science*, 169–174.

Cherubini, U., Luciano, E., & Vecchiato, W. (2004). *Copula Methods in Finance.* John Wiley and Sons.

Christoffersen, P. (1998). Evaluating Interval Forecasts. *International Economic Review*, *39*, 841-862.

Clements, M. P., Hendry, D. F., & Mills, T. C. (2012). *Forecasting financial time series*. Oxford University Press.

Cowell, F. (2002). Quantitative Asset Allocation Models. *Practical Quantitative Investment Management with Derivatives*, *4*, 55–93.

Czado, C. (2010). Pair-Copula Constructions of Multivariate Copulas. *Copula Theory and its Applications, Lecture Notes in Statistics*, *198*, 93-109.

Dias, A., & Embrechts, P. (2010). Modeling exchange rate dependence dynamics at different time horizons. *Journal of International Money and Finance*, *29*(8), 1687-1705.

Dissman, J., Brechmann, E., Czado, C., & Kurowicka, D. (2013). Selecting and estimating regular vine copulae and application to financial returns. *Computational Statistics and Data Analysis*, *59*, 52-69.

Embrechts, P., & Hofert, M. (2014). Statistics and Quantitative Risk Management for Banking and Insurance. *Annual Review of Statistics and its Applications*, *1*, 493–514.

Enders, W. (1995). *Applied Econometric Time Series.* John Wiley and Sons.

Engle, R. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, *50*(4), 987-1007.

Engle, R., & Kraft, D. (1983). Multiperiod forecast error variances of inflation estimated from ARCH models, in: A. ZeUner, ed., Applied time series analysis of economic data. *Bureau of the Census Washington DC*, 293-302.

Faber, M. (2013). A Quantitative Approach to Tactical Asset Allocation. *The Journal of Wealth Management*.

Fisher, R. A. (1925). Theory of statistical estimation. *Mathematical Proceedings of the Cambridge Philosophical Society*, *22*(5), 700–725.

Genest, C., Quesada Molina, J. J., & Rodriguez Lallena, J. A. (1995). De limpossibilité de construire des lois á marges multidimensionelles donneés á partir

de copules. *C. R. Acad. Sci. Paris*, *320*, 723–726.
Greene, W. H. (1997). *Econometric Analysis.* Prentice Hall.
Groppelli, A. A., & Nikbakht, E. (2000). *Barron's Finance*. New York.
Haas, M. (2001). New methods in backtesting. *Financial Engineering*.
Hamilton, J. D. (1994). *Time Series Analysis.* Princeton University Press.
Hendricks, D. (1996). Evaluation of Value-at-Risk Models Using Historical Data. *Economic Policy Review*, *2*(1).
Jarque, C. M., & Bera, A. K. (1987). A Test for Normality of Observations and Regression Residuals. *International Statistical Review*, *51*, 123-130.
Joe, H. (1996). Families of m-variate distributions with given margins and m(m-1)/2 bivariate dependence parameters. *Distribution with fixed marginals related topics. Institute of mathematical statistics, Hayward.*, 120-141.
Jondeau, E., Poon, S., & Rockinger, M. (2007). *Financial Modeling Under Non-Gaussian Distributions.* Springer Finance, New York.
J.P.Morgan. (1985). *Risk metrics - technical document*. John Wiley Sons Inc.
Kiefer, J., & Wolfowitz, J. (1956). Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters. *The Annals of Mathematical Statistics*, 887-906.
Kupiec, P. H. (1995). Tecniques for verifying the accuracy of risk measurement models. *Journal of Derivatives*, *3*, 73-84.
Lawler, E. L. (1985). *The traveling salesman problem: A guided tour of combinatorial optimization*. John Wiley Sons Inc.
Li, H., Scarsini, M., & Shaked, M. (1996). Linkages: A Tool for the Construction of Multivariate Distributions with Given Nonoverlapping Multivariate Marginals. *Journal of Multivariate Analysis*, *56*, 20-41.
Mahoney, J. (1995). Empirical-based Versus Model-based Approaches to Value-at-Risk. *Unpublished manuscript, Federal Reserve Bank of New York*.
Manner, H., & Reznikova, O. (2012). A Survey on Time-Varying Copulas: Specification, Simulations, and Application. *Econometric Reviews*, *31*(6), 654–687.
Marsala, C., Pallotta, M., & Zenti, R. (2004). Integrated Risk Management with a Filtered Bootstrap approach. *Economic Notes*, *33*, 375–398.
McNeil, A., Frey, R., & Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques, Tools.* Princeton Series in Finance.
Nelsen, R. (2006). *An introduction to copulas.* Springer, New York.
Okimoto, T. (2008). New evidence of asymmetric dependence structures in international equity markets. *Journal of Financial and Quantitative Analysis*, *43*(3), 787–816.
on Banking Supervision, B. C. (1996). Supervisory Framework for the Use of 'Backtesting' in Conjunction with the Internal Models Approach to Market Risk Capital Requirements.
Palm, F. (1996). GARCH models of volatility. *Handbook of Statistics*, *14*, 209–240.

Patton, A. (2006). Modelling asymmetric exchange rate dependence. *International Economic Review*, *47*(2), 527–556.

Roncalli, A. (2002). *Gestion des risques multiples.* Lecture notes - Cours ENSAI de $3^{eme} anne.$

Self, S. G., & Liang, K. Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association*, *398*(82), 605–610.

Sheather, S. J., & Marron, J. S. (1990). Kernel Quantile Estimators. *Journal of the American Statistical*(85), 410–416.

Silverman, B. (1986). *Density Estimation for Statistics and Data Analysis.* London: Chapman and Hall.

Sklar, A. (1959). Fonctions de rrtition imensions et leur marges. *Publ. Inst. Statist. Univ. Paris 8*, 229-231.

Stoeber, J., & Czado, C. (2012). Detecting regime switches in the dependence structure of high dimensional financial data. *Computational Statistics Data Analysis*, *76*.

Stuart, A., & Ord, J. (1987). *Kendalls Advanced Theory of Statistics*.

Taylor, S. J. (1994). Modeling stochastic volatility: a review and comparative study. *Mathematical Finance*, *4*(2), 183-204.

Von Storch, H., & Zwiers, F. W. (2001). Statistical analysis in climate research. *Cambridge Univ Pr.*.

Walker, G. (1931). On Periodicity in Series of Related Terms. *Proceedings of the Royal Society of London*, 518–532.

Yule, G. U. (1927). On a Method of Investigating Periodicities in Disturbed Series, with Special Reference to Wolfer's Sunspot Numbers. *Philosophical Transactions of the Royal Society of London*, 267–298.

Zenti, R., & Pallotta, M. (2002). Are asset managers properly using tracking error estimates? *Journal of Asset Management*, *3*(3), 279-289.