Dynamic Copula Modelling for Value at Risk

Dean Fantazzini•

Abstract

This paper proposes dynamic copula and marginals functions to model the joint distribution of risk factor returns affecting portfolios profit and loss distribution over a specified holding period. By using copulas, we can separate the marginal distributions from the dependence structure and estimate portfolio Value-at-Risk, assuming for the risk factors a multivariate distribution that can be different from the conditional normal one. Moreover, we consider marginal functions able to model higher moments than the second, as in the normal. This enables us to better understand why VaR estimates are too aggressive or too conservative. We apply this methodology to estimate the 95%, 99% VaR by using Monte-Carlo simulation, for portfolios made of the SP500 stock index, the Dax Index and the Nikkei225 Index. We use the initial part of the sample to estimate the models, and the the remaining part to compare the out-of-sample performances of the different approaches, using various back-testing techniques.

Keywords: Copulae, Value at Risk, Dynamic Modelling
JEL classification: G11, G12, G32

• Moscow School of Economics – Moscow State University. E-mail: fantazzini@mse-msu.ru Phone ++74955105267, Fax + 74955105256.
An earlier version of this paper has been presented at conferences in Konstanz, Venice, Dublin. For helpful comments and suggestions I'd like to thank, Luc Bauwens, Elena M. DeGiuli, Mario Maggi, Winfried Pohlmeier, Eduardo Rossi, Giovanni Urga, and Bas Werker. Financial support is gratefully acknowledged by the Research and Training Network (RTN) "Microstructure of Financial Markets in Europe".

1 - Introduction

Following the increase in financial uncertainty in the 90s, there has been intensive research from financial institutions, regulators and academics to better develop sophisticated models for market risk estimation (see Jorion (1997), Crouhy, Galai, and Mark (2001), Dowd (1998)). The most well known risk measure is the Value-at-Risk ("VaR"), which is defined as the maximum loss which may be incurred by a portfolio, at a given time horizon and at a given level of confidence. Jorion (2000) provides an introduction to Value at Risk as well as discussing its estimation, while the www.gloriamundi.org website comprehensively cites the Value at Risk literature as well as providing other VaR resources.

VaR is now recognized by both official bodies and private sector groups as an important market risk measurement tool: official bodies that have embraced the concept of VaR include the Basle Committee on Banking supervision (Basle Committee on Banking Supervision (1998)) and the Bank for International Settlements (Fisher-Report (1994)). Private sector organizations, such as the Group of-Thirty (1993), which consists of bankers and other derivatives market participants, have advocated the use of Var for setting risk management standards.

However, we remark that Value-at-Risk (VaR) as a risk measure is criticized for not being sub-additive (see Embrechts (2000) for an overview of the criticism). This means that the risk of a portfolio can be larger than the sum of the stand-alone risks of its components when measured by VaR (Artzner, Delbaen, Eber, and Heath (1997) and Artzner, Delbaen, Eber, and Heath (1999)). Hence, managing risk by VaR may fail to stimulate diversification. Still, the final Basle Capital Accord that will come in force since 2007 focus on VaR only (see the Basle Committee on Banking Supervision (2005)). Therefore, we explore the features of VaR in a multivariate setting.

In order to estimate VaR, we can use either non-parametric or parametric methods. The historical simulation methodology belongs to the former family and, instead of making distributional assumptions on returns, it uses realized past returns to estimate the given percentile. The parametric methods make specific distributional assumptions on returns and, therefore computes (analytically or by simulation) the corresponding percentile.

Crucial for the VaR is the distribution function of the returns on market value. Usually, as allowed by the Basle Committee, a normal distribution is assumed for the market return. However it is a known stylized fact that the normal distribution underestimates the probability in the tail and hence the VaR (see McNeil, Frey, and Embrechts (2005) for a general review).

Nowadays, alternative distributions and methods are proposed that focus more on the tail behavior of the returns. See, for example, Embrechts, Kluppelberg, and Mikosch (1997), Lucas and Klaassen (1998), Tsay (2002) for a discussion. Popular alternatives in the financial literature include GARCH-type models to allow for time-varying volatility and the Student-t distribution, since it allows for more probability mass in the tail than the normal distribution. For a review on (G)ARCH models we refer to Bollerslev, Engle, and Nelson (1994).

While univariate VaR estimation has been widely investigated, the multivariate case has been dealt only by a smaller and more recent literature, regarding the forecasting of correlations between assets, too: empirical works which deal with this issue are those of Engle and Sheppard (2001), Giot and Laurent (2003), Bauwens and Laurent (2005) and Chong (2004).

When we use parametric methods, VaR estimation for a portfolio of assets can become very difficult due to the complexity of joint multivariate modelling. Moreover, computational problems arise when increasing the number of assets.

In order to overcome these problems, we can resort to Copula theory: copula functions allow to construct flexible multivariate distribution with different margins and different dependence structure, without the constraints of the traditional joint normal distribution. Hull and White (1998) were among the first to consider this kind of modelling, even though they did not refer explicitly to copulas, but rather to mapping observations from the assumed distribution of daily changes into a standard normal distribution on a “fractile-to-fractile” basis. More recently, copulae have been explicitly used for measuring portfolio Value at Risk by Bouye’, Durrleman, Nikeghbali, Riboulet, and Roncalli (2001), Embrechts, Lindskog, and McNeil (2003) and Cherubini, Vecchiato, and Luciano (2004).Glasserman, Heidelberger, and Shahabuddin (2002) extended this framework to a portfolio of derivatives by considering a particular extension of the multivariate t distribution. Yet, the applications made so far dealt with *unconditional* distributions, only.

What we do in this work is to propose a Monte Carlo methodology for measuring portfolio Value at Risk by performing simulations from different *conditional* multivariate distributions, in order to evaluate what are the main determinants when doing VaR forecasts for a portfolio of assets. We compare different distribution assumptions for the marginals, as well as different dynamic specifications for their moments, to understand whether the proper modelling of the latter is more important than the type of distribution.

In a similar fashion, we consider different type of copulas and different conditional specifications for their parameters.

We analyze three portfolios composed of the Standard and Poor 500 stock index, the Dax Index and the Nikkei225 index, with daily data taking into consideration the very volatile period between January 1th 1994 till August 1th 2000. This period of time witnessed the fast raise of world financial markets followed by the the burst of the high-tech bubble. Thus, this sample is perfectly suited to highlight the differences among the considered models. We use the initial part of the sample to estimate the 95%, 99% VaR, and the remaining part to compare the out-of-sample performances of the different models, using various backtesting techniques.

The rest of the paper is organized as follows. In Section 2 we provide an outline of multivariate modelling, focusing the attention to copula theory and skewed marginals. Section 3 presents the necessary steps to perform Monte-Carlo simulation for VaR forecasting within a flexible copula framework, while Section 4 presents the empirical results. We conclude in Section 5.

2 - Conditional multivariate modelling

Copula theory provide an easy way to deal with (otherwise) complex multivariate modelling. The essential idea of the copula approach is that a joint distribution can be factored into the marginals and a dependence function called a copula. The term copula comes the Latin language and means “link”: the copula couples the marginal distributions together to form a joint distribution. The dependence relationship is entirely determined by the copula, while scaling and shape (mean, standard deviation, skewness, and kurtosis) are entirely determined by the marginals.

Copulas can be useful for combining risks when the marginal distributions are estimated individually: marginals are initially estimated separately and can then be combined in a joint density by a copula that preserves the characteristics of the marginals. Copulas can also be used to obtain more realistic multivariate densities than the traditional joint normal. For example, the normal dependence relation can be preserved using a normal copula, but marginals can be entirely general (for example, a normal copula with one T-student marginal and one logistic marginal). In this case we have the so-called meta joint distribution functions, whose theoretical background has been studied recently by Embrechts, Hoing, and Juri (2003) and Fang and Fang (2002).

In the following, we briefly describe the conditional marginal and copula functions used to construct multivariate distributions.

2.1 Copula modelling

In what follows, the definition of a copula function and some of its basic properties are given, while the reader interested in a more detailed treatment is referred to Nelsen (1999) and Joe (1997). As time series analysis is often interested in random variables conditioned on some variables, we allow for conditioning variables represented by F_{t-1} (for example, lagged returns). Besides, we consider the bivariate case since it will be later used in the empirical analysis.

Definition 2.1 (Conditional Copula): *The conditional copula of* $(x,y)|\ \Phi_{t-1}$, *where* $x|\ \Phi_{t-1} \sim F_t$ *and* $y|\ \Phi_{t-1} \sim G_t$, *is the conditional joint distribution function of* $U_t \equiv F_t\ (x|\ \Phi_{t-1})$ *and* $V_t \equiv G_t\ (y|\ \Phi_{t-1})$, *given* Φ_{t-1}.

It is easy to observe that a two-dimensional conditional copula is the conditional joint distribution of the probability integral transforms of each of the variables X_t and Y_t with respect to their marginal distributions, F_t and G_t.

In order to understand why the univariate margins and the multivariate dependence structure can be separated, we have to consider the following theorem:

Theorem 2.1 (Sklar's theorem for conditional distributions): *Let* F_t *be the conditional distribution of* $x|\Phi_{t-1}$, *given the conditioning set* Φ_{t-1}, G_t *be the conditional distribution of* $y|\Phi_{t-1}$, *and* H_t *be the joint conditional bivariate distribution of* $(x, y|\Phi_{t-1})$. *Assume that* F_t *and* G_t *are continuous in x and y. Then there exists a unique conditional copula* C_t *such that*

$$H_t\ (x, y|\Phi_{t-1}) = C_t\ (F_t\ (x|\Phi_{t-1}), G_t\ (y|\Phi_{t-1})|\ \Phi_{t-1}) \qquad (2.1)$$

Conversely, if we let F_t *and* G_t *be the conditional distributions of the two random variables* X_t *and* Y_t , *and* C_t *be a conditional copula, then the function* H_t *defined as (2.1) is a conditional bivariate distribution function with conditional marginal distributions* F_t *and* G_t (*see Sklar (1959), Patton (2006), for a proof*).

A copula is thus a function that, when applied to univariate marginals, results in a proper multivariate probability density function (p.d.f.): since this p.d.f. embodies all the information about the random vector, it contains all the information about the dependence structure of its components. Using copulas in this way splits the distribution of a random vector into individual components (the marginals), with a dependence structure among them given by the copula, without losing any information.

The Sklar's theorem for conditional distributions implies that the conditioning variable(s), Φ_{t-1}, must be the same for both marginal distributions and the copula: if we do not use the same conditioning variable for F_t, G_t and C_t, the function H_t will not be, in general, a joint conditional distribution function. The only case when H_t will be the joint distribution of $(x, y|\Phi_{t-1}) = (x, y|w_1,w_2)$, is when $F_t(x|\ w_1) = F_t(x|\ w_1,\ w_2)$ and $G_t(y|\ w_2) = G_t\ (y|\ w_1,\ w_2)$, that is when some variables affect the conditional distribution of one variable but not the other. It is straightforward to see that it is possible to extract the implied conditional copula from any bivariate conditional distribution. By applying Sklar's theorem and using the relation between the distribution and the density function, we can derive the bivariate copula density $c_t(F_t(x|\Phi_{t-1}),\ G_t(y|\Phi_{t-1})|\ \Phi_{t-1})$, associated to a copula function $C_t(F_t\ (x|\Phi_{t-1}),\ G_t(y|\Phi_{t-1})|\ \Phi_{t-1})$:

$$h_t(x,\ y\,|\Phi_{t-1}) = \frac{\partial^2[C_t(F_t(x\,|\Phi_{t-1}),G_t(y\,|\Phi_{t-1})\,|\Phi_{t-1})]}{\partial F_t(x\,|\Phi_{t-1}),\partial G_t(y\,|\Phi_{t-1})}\cdot\frac{\partial F_t(x\,|\Phi_{t-1})}{\partial x}\cdot\frac{\partial G_t(y\,|\Phi_{t-1})}{\partial y} =$$

$$= c_t(F_t(x\,|\Phi_{t-1}),G_t(y\,|\Phi_{t-1})\,|\Phi_{t-1})\cdot f_t(x\,|\Phi_{t-1})\cdot g_t(y\,|\Phi_{t-1}) \ \rightarrow$$

$$c_t(u,\ v\,|\Phi_{t-1}) = \frac{h_t(x,\ y\,|\Phi_{t-1})}{f_t(x\,|\Phi_{t-1})\cdot g_t(y\,|\Phi_{t-1})} \tag{2.2}$$

where $u \equiv F_t\ (x|\Phi_{t-1})$ and $v \equiv G_t\ (y|\Phi_{t-1})$. By using this procedure, we can derive the Normal and the T-copula:

1 The copula of the bivariate Normal distribution is the *Normal-copula*, whose probability density function in the bivariate case is the following

$$c(u_t,\ v_t;\rho_t\,|\mathcal{F}_{t-1}) = \frac{1}{\sqrt{1-\rho_t^2}}\cdot$$
$$\cdot\exp\left(-\frac{1}{2}\cdot\frac{[(\Phi^{-1}(u_t))^2\ +\ (\Phi^{-1}(v_t))^2\ -\ 2\rho_t\cdot\Phi^{-1}(u_t)\cdot\Phi^{-1}(v_t)]}{1-\rho_t^2}\right)$$
$$\cdot\exp\left(\frac{1}{2}[(\Phi^{-1}(u_t))^2\ +\ (\Phi^{-1}(v_t))^2]\right) \qquad (2.3)$$

where ρ_t is the conditional linear correlation, given the conditioning set Φ_{t-1}, and Φ^{-1} is the inverse of the standard univariate Normal distribution.

2 On the other hand, the copula of the bivariate Student's t-distribution is the Student's *T-copula*, whose density function is

$$c(u_t,\ v_t;\rho_t,\upsilon_t\,|\mathcal{F}_{t-1}) = \frac{\Gamma\left(\frac{\upsilon_t+2}{2}\right)\Gamma\left(\frac{\upsilon_t}{2}\right)}{\Gamma\left(\frac{\upsilon_t+1}{2}\right)^2\sqrt{1-\rho_t^2}}\cdot$$
$$\cdot\exp\left(-\frac{1}{2}\cdot\frac{[(t_{\upsilon_t}^{-1}(u_t))^2\ +\ (t_{\upsilon_t}^{-1}(v_t))^2\ -\ 2\rho_t\cdot t_{\upsilon_t}^{-1}(u_t)\cdot t_{\upsilon_t}^{-1}(v_t)]}{(1-\rho_t^2)\cdot\upsilon_t}\right)^{-\frac{\upsilon_t+2}{2}}$$
$$\cdot\exp\left[\left(1+\frac{(t_{\upsilon_t}^{-1}(u_t))^2}{\upsilon_t}\right)\left(1+\frac{(t_{\upsilon_t}^{-1}(v_t))^2}{\upsilon_t}\right)\right]^{\frac{\upsilon_t+1}{2}} \qquad (2.4)$$

where ρ_t is the conditional linear correlation, υ_t are the conditional degrees of freedom, while $t_{\upsilon_t}^{-1}$ denotes the inverse of the Student's t cumulatived istribution function.

Both these copulae belong to the class of Elliptical copulae[1]. An alternative to Elliptical copulae is given by Archimedean copulae: however, they present the serious limitation to model only positive dependence (or only partial negative dependence), while their multivariate extension involve strict restrictions on bivariate dependence parameters. This is why we do not consider them here.

[1] See Cherubini, Vecchiato, and Luciano (2004) for more details.

2.2 Marginal modelling

The marginal distributions that we used to build a joint multivariate distribution are the following four:

1. Normal;
2. Skew-Normal;
3. T-student;
4. Skew-T student.

Two well known deviations from normality are fat tails and asymmetry. One distribution that is used to allow for excess kurtosis is the Students-T and it has been generalized toallow for skewness by Hansen (1994). Despite other generalizations have been proposed, we chose this one due to its simplicity and its past success in modelling economic variables (Patton (2004), Patton (2006), Jondeau and Rockinger (2003)).

Definition 2.2 (Skew-T distribution): *Let y_t be a random variable which follows a conditional Skewed-T distribution with density function $f(\upsilon_t, \lambda_t)$ and mean zero and variance one by construction, in order to be a suitable model for the standardized residuals of a conditional mean and variance model. The conditional parameters υ_t, λ_t control the kurtosis and skewness of the variable, respectively, while the density function is reported below (for more details, see Hansen (1994)):*

$$\text{Skewed-T } f(y_t; v_t, \lambda_t \,|\mathcal{F}_{t-1}) = \begin{cases} bc\left(1+\dfrac{1}{\upsilon_t+2}\left(\dfrac{by_t+a}{1-\lambda_t}\right)^2\right)^{-\frac{\upsilon_t+1}{2}} & for\ y_t \le -\dfrac{a}{b} \\ bc\left(1+\dfrac{1}{\upsilon_t+2}\left(\dfrac{by_t+a}{1+\lambda_t}\right)^2\right)^{-\frac{\upsilon_t+1}{2}} & for\ y_t > -\dfrac{a}{b} \end{cases}$$

where

$$c = \frac{\Gamma\left(\dfrac{\upsilon_t+1}{2}\right)}{\Gamma\left(\dfrac{\upsilon_t}{2}\right)\sqrt{\pi(\upsilon_t-2)}}$$

$$b = \sqrt{1+3\lambda_t^2-a^2}$$

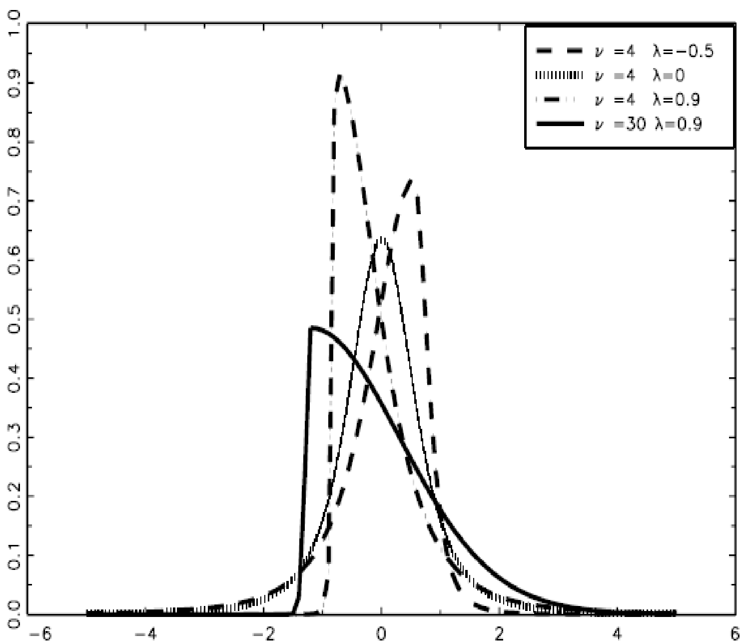
$$a = 4\lambda_t\, c\left(\frac{\upsilon_t-2}{\upsilon_t-1}\right)$$

This density is defined for $2<\upsilon_t<\infty$ and $-1 < \lambda_t <+1$. Moreover, this density encompasses a large set of conventional densities:

1. if λ_t =0, the Skew-t reduces to the traditional Student-T distribution.
2. If $\upsilon_t \to \infty$ we have the skew-normal density.
3. If $\lambda_t = 0$ and $\upsilon_t \to \infty$ we have the normal density.

Similarly to Student's T, given the restriction $\upsilon_t > 2$, this distribution is well defined and its second moment exists, while skewness exists if $\upsilon_t > 3$ and kurtosis is defined if $\upsilon_t > 4$. The parameter λ_t controls for skewness: If it is bigger than zero, we have positive skewness, while if it is smaller than zero the distribution is negative skewed. Fig.1 displays different densities obtained for various values of degrees of freedom υ and skewness parameter λ. The cumulative distribution function (c.d.f.), the inverse-c.d.f. and relative proofs are reported in Appendix A.

Figure 1: Density of Hansen's Skewed-T distribution



2.3 Estimation: the inference for margins method (IFM)

As we have seen, the joint density function in the bivariate conditional case is:

$$h_t(x, y \mid \mathcal{F}_{t-1};\theta_h) = f_t(x \mid \mathcal{F}_{t-1};\theta_f)\cdot g_t(y \mid \mathcal{F}_{t-1};\theta_g)\cdot c_t(u,v \mid \mathcal{F}_{t-1};\theta_c) \quad (2.5)$$

where $u \equiv F_t\ (x|\Phi_{t-1};\ \theta_f)$ and $v \equiv G_t\ (y|\Phi_{t-1};\ \theta_g)$, and θ_h, θ_f, θ_g, θ_c are the joint density, marginals and copula parameters' vectors, respectively, with θ_h, ≡ [θ_f', θ_g', θ_c']. Maximum likelihood analysis implies,

$$Lxy(\theta_h) = Lx(\theta_f) + Ly(\theta_g) + Lc(\theta_f,\ \theta_g,\ \theta_c),$$

where $Lxy(\theta_h) \equiv \log h_t\ (x, y|\Phi_{t-1};\ \theta_h)$, $Lx(\theta_f) \equiv \log f_t\ (x|\Phi_{t-1};\ \theta_f)$, $Ly(\theta_g) \equiv \log g_t(y|\Phi_{t-1};\ \theta_g)$, and $Lc(\theta_f,\ \theta_g,\ \theta_c) \equiv \log c(u,\ v|\Phi_{t-1};\ \theta_c)$.

It will not always be the case that the parameter vector of the joint distribution θ_h decomposes so neatly into n+1 components, associated with the n margins (in this case n = 2) and the copula. Examples of common models where such a decomposition is possible are vector autoregressions and multivariate autoregressive conditional heteroscedasticity (ARCH) models: in these cases, the n margins' parameter vectors are variation free w.r.t each other, and can be estimated separately.

Common models where the previous decomposition is not possible are multivariate ARMA models and multivariate generalized ARCH models. When this decomposition fail to hold, all the n marginal models must be estimated jointly.

According to the IFM method, the parameters of the marginal distributions are estimated separately from the parameters of the copula. In other words, the estimation process is divided into the following two steps:

1.Estimating the parameters θ_f and θ_g of the marginal distribution F_t and G_t using the ML method:

$$\hat{\theta}_f = \arg\max L(\theta_f) \arg\max \sum_{t=1}^{T} \log f_t(x_t;\theta_f)$$
$$\hat{\theta}_g = \arg\max L(\theta_g) \arg\max \sum_{t=1}^{T} \log g_t(y_t;\theta_g) \quad (2.6)$$

if the parameters' vectors are variation free; otherwise,

$$\hat{\theta}_f, \hat{\theta}_g = \arg\max[L(\theta_f) + L(\theta_g)] = \arg\max \sum_{t=1}^{T} [\log f_t(x_t;\theta_f) + \log g_t(x_t;\theta_g)] \quad (2.7)$$

2.Estimating the copula parameters θ_c , given step 1):

$$\hat{\theta}_c = \arg\max[L(\theta_c)] = \arg\max \sum_{t=1}^{T} \log[c_t(F_t(x_t;\hat{\theta}_f), G_t(x_t;\hat{\theta}_g);\theta_c)] \quad (2.8)$$

Like the ML estimator it verifies the properties of asymptotic normality, but the covariance matrix must be modified (Joe and Xu (1996), Joe (1997), Patton (2006)):

$$\sqrt{T}(\hat{\theta}_h - \theta_0) \rightarrow N(0, V^1(\theta_0))$$

where $V(\theta_0) = \mathbf{D}^{-1}\mathbf{M}(\mathbf{D}^{-1})^T$ is the Godambe Information Matrix, while $\mathbf{D} = E[\partial g(\theta)^T/\partial\,\theta]$, $\mathbf{M} = E\,[g(\theta)^T\, g(\theta)]$, and $g(\theta) = (\partial L_x/\partial\theta_f\,, \partial L_y/\partial\theta_g, \partial Lc/\partial\theta_c)$ is the score function.

2.3.1 In-sample results: the dataset

We consider three important risk factors, such as the Standard and Poor 500 index, the Dax Index and the Nikkei225 Index, with daily data taking into consideration the very volatile period between January 1th 1994 till August 1th 2000. This period of time witnessed the fast raise of world financial markets followed by the the burst of the high-tech bubble. The descriptive statistics of the (raw) log-returns are presented in Table 1.

The three indexes generally exhibited negative skewness (the Nikkei slightly positive, instead) and excess kurtosis, while the Jarque-Bera statistic indicates that neither series is unconditionally normal. Besides, the means and volatilities are very similar, as expected.

Table 1: **SP500, DAX, NIKKEI225 Log-Returns Descriptive statistics.**

	SP500	*DAX*	*NIKKEI225*
# of obs.	1700	1700	1700
Mean	0.00068	0.00067	0.00001
Median	0.00041	0.00070	0.00000
Maximum	0.04989	0.05894	0.07655
Minimum	-0.07113	-0.06450	-0.07234
Std. Dev.	0.00994	0.01291	0.01376
Skewness	-0.44087	-0.36694	0.06449
Kurtosis	8.62053	5.48054	6.26267
Jarque-Bera	2292.72	473.99	755.20

2.3.2 In sample results: marginals models

We started modelling each marginal time series by a general AR(1)-Threshold GARCH(1,1) model[2] for the continuously compounded returns $y_t = 100[\log(P_t) - \log(P_{t-1})]$, given by:

$$y_t = \mu + \phi_1 y_{t-1} + \varepsilon_t \tag{2.9}$$

$$\varepsilon_t = \eta_t \sqrt{h_t}\,,\ \eta_t \overset{i.i.d}{\sim} f(0.1) \tag{2.10}$$

$$h_t = \omega + \alpha\varepsilon_{t-1}^2 + \gamma\varepsilon_{t-1}^2 D_{t-1} + \beta h_{t-1} \tag{2.11}$$

where $D_{t-1} = 1$ if $\varepsilon_{t-1} < 0$, and 0 otherwise. Good news $\varepsilon_{t-1} > 0$ and bad news $\varepsilon_{t-1} < 0$, have different effects on the conditional variance in this model: good news has an impact of α, while bad news has an impact of $\alpha + \gamma$. If $\gamma > 0$ we say that the leverage effect exists, while if $\gamma \neq 0$ the news impact is asymmetric.

We tried an EGARCH specification Nelson (1991), too: however, since the diagnostic tests and the Schwartz criterion were similar to that of the TGARCH, but the numerical maximization of the log-likelihood function resulted much more difficult when working with not normal distributions, we resort to the TGARCH specification (this evidence confirms similar results by Angelidis, Benos, and Degiannakis (2003)).

We estimate the AR(1)-TGARCH(1,1) model assuming four different density functions $f(0, 1)$ for η_t: the Normal, the Skew-Normal, the Student's-T and the Skew-T as presented in section 2.2. When working with the latter three distributions, we have to specify a dynamic model for the conditional skewness parameter and/or the conditional degrees of freedom, as well. We propose here a specification similar to Hansen (1994) and Rockinger and Jondeau (2001):

$$\lambda_t = \Lambda(\zeta + \delta\,\varepsilon_{t-1}) \tag{2.12}$$

$$\upsilon_t = \Gamma(\zeta + \tau\,\varepsilon_{t-1}) \tag{2.13}$$

[2] For the SP500 we used an AR(3)-Threshold GARCH(1,1) model, instead, since Likelihood Ratio tests and diagnostic tests indicate this as the best model.

where $\Lambda(\cdot)$ is a modified logistic transformation designed to keep the conditional skewness parameter λ_t in (-1,1) at all times, while $\Gamma(\cdot)$ is a logistic transformation designed to keep the conditional degrees of freedom in (2,30) at all times (see Hansen 1994).

We avoid an autoregressive specification, in so far as it may lead to spuriously significant parameters (see Rockinger and Jondeau (2001), for a proof). Moreover, we tried different specifications, but with similar results and increased computational time. This is why we resort to this simple modelling.

The models estimated using the entire dataset available are reported below.

Table 2: **AR(1)-TGARCH(1,1) models with different distributions assumptions.**

		Mean		*Volatility*				*Skewness*		*D.o.F*		*Log-Lik.*
Asset	*Distribution*	μ	ϕ	ω	α	γ	β	ζ	δ	θ	τ	
SP500 (*)	Normal	0.065 (0.019)	−0.051 (0.025)	0.015 (0.002)	0.008 (0.011)	0.131 (0.013)	0.913 (0.008)	/	/	/	/	−2167
	Skew-Normal	0.060 (0.019)	−0.063 (0.024)	0.015 (0.002)	0.008 (0.011)	0.135 (0.014)	0.908 (0.010)	−0.276 (0.062)	0.196 (0.045)	/	/	−2142
	Student's T	0.085 (0.017)	−0.084 (0.024)	0.008 (0.002)	0.012 (0.016)	0.096 (0.021)	0.934 (0.012)	/	/	−2.051 (0.229)	−0.303 (0.183)	−2114
	Skew-T	0.075 (0.018)	−0.084 (0.024)	0.010 (0.003)	0.012 (0.016)	0.113 (0.022)	0.924 (0.012)	−0.136 (0.075)	0.162 (0.052)	−1.892 (0.256)	−0.394 (0.186)	−2106
DAX	Normal	0.064 (0.026)	0.007 (0.028)	0.022 (0.004)	0.050 (0.011)	0.039 (0.011)	0.916 (0.011)	/	/	/	/	−2655
	Skew-Normal	0.057 (0.026)	0.011 (0.026)	0.019 (0.004)	0.051 (0.011)	0.033 (0.012)	0.920 (0.011)	−0.233 (0.059)	0.120 (0.051)	/	/	−2641
	Student's T	0.084 (0.025)	−0.007 (0.025)	0.014 (0.006)	0.053 (0.017)	0.035 (0.018)	0.923 (0.014)	/	/	−1.247 (0.329)	−0.454 (0.206)	−2630
	Skew-T	0.066 (0.026)	0.012 (0.026)	0.014 (0.006)	0.054 (0.017)	0.036 (0.018)	0.922 (0.014)	−0.191 (0.075)	0.124 (0.060)	−1.035 (0.394)	−0.543 (0.230)	−2621
NIKKEI	Normal	−0.005 (0.029)	−0.042 (0.026)	0.047 (0.007)	0.027 (0.008)	0.096 (0.014)	0.901 (0.011)	/	/	/	/	−2823
	Skew-Normal	−0.009 (0.030)	−0.034 (0.026)	0.047 (0.008)	0.029 (0.009)	0.100 (0.015)	0.898 (0.015)	−0.059 (0.042)	0.095 (0.032)	/	/	−2820
	Student's T	−0.003 (0.026)	−0.056 (0.023)	0.023 (0.009)	0.020 (0.011)	0.074 (0.019)	0.932 (0.013)	/	/	−1.938 (0.225)	−0.212 (0.128)	−2767
	Skew-T	−0.012 (0.028)	−0.045 (0.025)	0.024 (0.010)	0.023 (0.012)	0.077 (0.020)	0.928 (0.014)	−0.061 (0.070)	0.058 (0.046)	−1.922 (0.230)	−0.216 (0.132)	−2765

(*) The - parameter correspond to an AR(3) term

The best specification is given by the Skew-T distribution, as expected, even though all four distributions passed Box-Pierce tests on the standardized residuals (not reported). The estimated parameters are rather similar across different distributions: however, Table 2 shows that when skewness and kurtosis are not taken into account, both the parameters of the mean and of the variance equations appear to be slightly overestimated; this happens for the Normal distribution, too. These results confirm Newey and Steigerwald (1997), who showed that Quasi-ML estimators can give not consistent estimates under certain conditions.

The leverage effect is statistically significant in all the considered series, as well as the conditional skewness and degrees of freedom. However, the skewness parameters are not significant for the Nikkei255, showing that the skewness for this series is not statistically different from zero, as we've previously seen in the descriptive statistics in Table 1.

2.3.3 In-sample results: copula models

After having estimated the parameters of the marginal distributions $\{F_t,G_t\}$ in the first stage, we proceeded to estimate the copula parameters in the second stage, as previously explained in paragraph 2.3 .

Since we'll consider three portfolios for VaR evaluation, composed of the SP500 and the Dax, the SP500 and the NIkkei225, and the Nikkei225 and the Dax, respectively, we fitted the Normal (2.3) and T-copula (2.4) to these asset pairs, by using the cumulative distribution function of the standardized residuals $\hat{\eta}_t$ estimated from the marginal models:

$$\left(\frac{X_t-\mu_t^x}{\sqrt{h_t^x}},\frac{X_t-\mu_t^x}{\sqrt{h_t^y}}\right) \sim C_{Normal-copula}\left(F_t(\eta_t^x),\cdot G_t(\eta_t^y);\rho_t \,|\mathcal{F}_{t-1}\right)$$

$$\left(\frac{X_t-\mu_t^x}{\sqrt{h_t^x}},\frac{X_t-\mu_t^x}{\sqrt{h_t^y}}\right) \sim C_{T-copula}\left(F_t(\eta_t^x),\cdot G_t(\eta_t^y);\rho_t,\upsilon_t \,|\mathcal{F}_{t-1}\right) \qquad (2.14)$$

where $[\rho_t,\ \upsilon_t]$ are the conditional correlation and conditional degrees of freedom, respectively, $[\mu_t,\ h_t]$ the conditional means and variances, while $\{F_t,G_t\}$ can be Normal / Skew-Normal / Student's T / Skew T.

When we consider a dynamic specification for a copula, it is important to remember that the Sklar's theorem for conditional distributions implies that the conditioning variable(s) must be the same for both marginal distributions and the copula. Following Patton (2006) and Embrechts and Dias (2003), we suggest this general evolution equation as starting point for our dynamic copula specifications:

$$\lambda_t = \Lambda(\xi+\vartheta\,|\,u_{t-1}-v_{t-1}\,|) \qquad (2.15)$$

$$\upsilon_t = \Gamma(\varsigma+\varphi\,|\,u_{t-1}-v_{t-1}\,|) \qquad (2.16)$$

where $\Lambda(x) \equiv (1-e^{-x})/(1+e^{-x})$ is the modified logistic transformation, designed to keep the conditional correlation ρ_t in $(-1,1)$ at all times, while $\Gamma(\cdot)$ is a logistic transformation designed to keep the conditional degrees of freedom in (2, 100) at all times.

The models estimated using the entire dataset available are presented in Table 3, while the conditional correlation for the three asset pairs, as well as the conditional degrees of freedom for the T-copula are reported in Figures 2-3 (for sake of space, we report the case of Normal and Skew-T marginals, only).

Table 3: **Different copula models with different marginals**

		Correlation		*D.o.F*		*Log-Lik.*
Assets pair	*Copula (Marginals)*	ξ	ϱ	ς	φ	
SP500 DAX	Normal (Normal m.)	0.620 (0.078)	−0.055 (0.246)	/	/	76.5
	Normal (Skew-T m.)	0.615 (0.075)	−0.021 (0.223)	/	/	76.8
	T-copula (Normal m.)	0.620 (0.080)	−0.059 (0.252)	2.869 (6.265)	1.712 (8.917)	76.6
	T-copula (Skew-T m.)	0.615 (0.079)	−0.027 (0.231)	−0.003 (6.696)	0.005 (2.489)	77.0
SP500 NIKKEI	Normal (Normal m.)	−0.005 (0.087)	0.390 (0.217)	/	/	4.3
	Normal (Skew-T m.)	−0.018 (0.084)	0.420 (0.197)	/	/	5.3
	T-copula (Normal m.)	−0.005 (0.089)	0.392 (0.223)	0.008 (0.931)	0.002 (2.236)	4.6
	T-copula Skew-T	−0.018 (0.085)	0.423 (0.201)	0.012 (2.191)	0.007 (2.506)	5.4
NIKKEI DAX	Normal (Normal m.)	0.461 (0.065)	−0.114 (0.210)	/	/	48.2
	Normal (Skew-T m.)	0.516 (0.074)	−0.117 (0.217)	/	/	48.8
	T-copula (Normal m.)	0.506 (0.087)	−0.121 (0.248)	−4.044 (0.896)	11.21 (12.94)	52.0
	T-copula (Skew-T m.)	0.503 (0.090)	−0.088 (0.241)	−5.583 (1.755)	11.82 (11.42)	53.1

Figure 2: **Conditional. Correlation T-copula: SP500-DAX, SP500-NIKKEI, NIKKEI-DAX.**

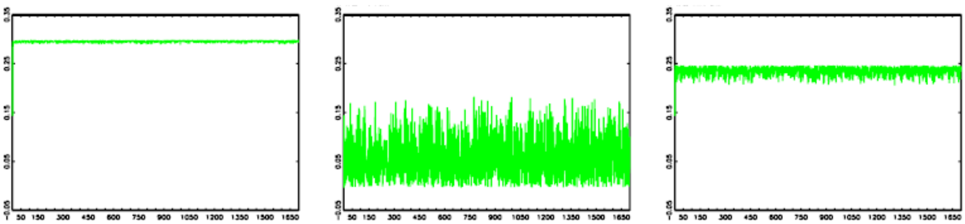
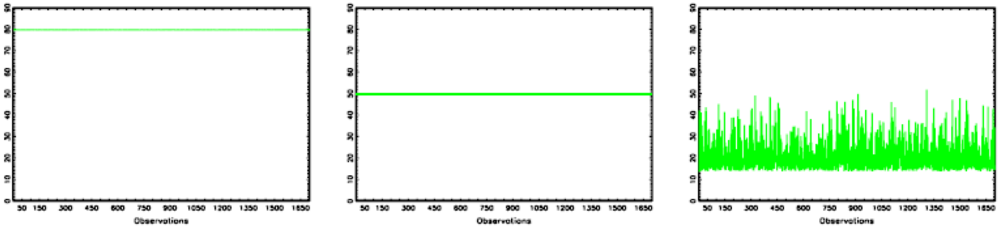


Figure 3: **Conditional. D.o.F. T-copula: SP500-DAX, SP500-NIKKEI, NIKKEI-DAX.**



As Table 3 and figures 2-3 clearly show, the simple normal copula with constant correlation is the best solution, since the dynamic parameters are not significant and the degrees of freedom of the T-copula are always above 20: that is the T-copula and the Normal copula are no more distinguishable.

3 - Value at risk applications

3.1 Introduction

Value at Risk (or VaR) is a concept developed in the field of risk management that is defined as the maximum amount of money that one could expect to lose with a given probability over a specific period of time. While VaR is widely used, it is, nonetheless, a controversial concept, primarily due to the diverse methods used in obtaining VaR, the widely divergent values so obtained and the fear that management will rely too heavily on VaR with little regard for other kinds of risks. The VaR concept embodies three factors:

1. A given time horizon. A risk manager might be concerned about possible losses over one day, one week, etc.
2. VaR is associated with a probability. The stated VaR represents the possible loss over a given period of time with a given probability.
3. The actual amount of money invested.

If we call $\Delta V(l)$ the change in the value of the assets in the financial position from t to t+l and $F_l(x)$ the cumulative distribution function ("cdf") of $\Delta V(l)$, the VaR is formally defined as follows:

Definition 3.1 (Value at Risk) *We define the VaR of a long position over time horizon l with probability p as the real $VaR_t(p, l)$ such that*

$$p = \Pr[\Delta V(l) \le VaR_t(p, l)] = F_l(VaR_t(p, l)).$$

where VaR is defined as a negative value (loss). If the cdf is known, then VaR is simply its p-th quantile times the value of the financial position: however, the cdf is not known in practice and must be estimated.

We remark that for any univariate cdf $F_l(x)$ and probability p, the quantity $x_p = \inf\{x | F_l(x) \geq p\}$ is called the p-th quantile of $F_l(x)$, i.e. the smallest real number x satisfying $F_l(x) \geq p$. Besides, the definition of VaR for a long position continuous to apply to a short position if one uses the distribution of $-\Delta V(l)$. Therefore is suffices to discuss methods of VaR calculations using a long position (for more details see Jorion (2000), Tsay (2002)).

3.2 VaR estimation: problem setup

Our goal is to estimate the daily Value at Risk for a portfolio of two assets (or "risk factors", in a general framework), by using the time series of daily returns for each risk factor. At this stage, we have to find the most appropriate joint distribution function which describes these data: This is done by choosing specific marginals for the single risk returns and a copula to link them together into a joint distribution function.

Let x_t and y_t denote the assets log-returns at time t and be $\beta \in (0, 1)$ the allocation weight, so that the portfolio return is given by $z_t = \beta x_t + (1-\beta) y_t$. The conditional joint distribution function estimated at time $t-1$ is given as in (2.1):

$$H_t(x, y|\Phi_{t-1}) = C_t(F_t(x|\Phi_{t-1}), G_t(y|\Phi_{t-1})| \Phi_{t-1}), \tag{3.1}$$

while the relative density function is given in (2.2). Hence, the cumulative distribution function for the portfolio return Z is given by:

$$\zeta(z) = \Pr(Z \leq z_t) = \Pr(\beta X + (1-\beta) Y \leq z_t) =$$
$$= \int_{-\infty}^{+\infty} \left\{ \int_{-\infty}^{\frac{1}{\beta} z - \frac{1-\beta}{\beta} y_t} c_t(F_t(x|\mathsf{F}_{t-1}), G_t(y|\mathsf{F}_{t-1})|\mathsf{F}_{t-1}) \cdot f_t(x|\mathsf{F}_{t-1}) dx \right\} \cdot g_t(y|\mathsf{F}_{t-1}) dy \tag{3.2}$$

The one-step-ahead VaR computed in $t-1$ for the portfolio at a confidence level p is the solution z^* of the equation $\zeta(z^*)=p$, times the value of the financial position at t-1 .

Apart for the multivariate normal and multivariate T case, when using copulas with different marginals the computation of the one-step-ahead VaR is not straightforward, since there are no analytic and easy-to-use formulae to switch from the conditional means and volatilities to the long and short VaR of the portfolio. However, the VaR can be computed using a simple Monte Carlo simulation as widely used in quantitative finance and VaR applications, see Jorion (2000), Giot and Laurent (2003) and Bauwens and Laurent (2005). Besides, Giot and Laurent (2003) show that a choice of 100,000 simulations provides accurate estimates of the quantile.

Following this solution, we generate a large number of one-day-ahead returns $\{x_t, y_t\}$ for the two assets, by simulating 100,000 random returns with the conditional distribution function (3.1) and we revaluate the portfolio at time t. We then determine the Value at Risk at a given confidence level p, by simply taking the empirical quantile at p of the simulated portfolio profit and loss distribution.

The detailed steps of the procedure for estimating the 95%, 99% VaR over a one-day holding period are the following:

1. Let consider the portfolio z which contains one position for each of the two risk factors (or assets), whose value at time $t-1$ is:

$$P_{z,t-1} = P_{x,t-1} + P_{y,t-1}$$

where $P_{x,t-1}$, $P_{y,t-1}$ are the market prices of the two assets at time $t-1$.

2. We simulate j = 100,000 MC scenarios for each asset log-returns, $\{x_{j,t}\ , y_{j,t}\}$, over the time horizon $[t-1,t]$, using the conditional joint distributional function (3.1).

(a) First, we have to simulate a j random variate $(u_{j,x},\ u_{j,y})'$ from the copula $C_t(\cdot)$. See Cherubini, Vecchiato, and Luciano (2004), for a discussion about copula simulation.

i. → If we are using a 2-dimensional conditional *Normal copula* (2.3), we have to run the following algorithm:

· Find the Cholesky decomposition **A** of the 2×2 correlation matrix Σ, i.e. [1 $\hat{\rho}_t$, $\hat{\rho}_t$ 1], where $\hat{\rho}_t$ is the forecasted conditional correlation, estimated with the dynamic model described in section 2.3.3;
· Simulate 2 independent standard normal random variates $\mathbf{z_j} = (z_{j,1}, z_{j,2})'$;
· Set the vector $b_j = \mathbf{A}z_j$;
· Determine the components $(u_{j,x}, u_{j,y})' = (\phi(b_{j,1}), \phi(b_{j,2}))'$, where $\phi(\cdot)$ is the standard normal c.d.f.
The vector $(u_{j,x}, u_{j,y})'$ is a *j* random variate from the 2-dimensional Normal copula $C_{Normal}(\cdot ; \hat{\rho}_t | \Phi_{t-1})$

ii. → If we are using a 2-dimensional conditional *T-copula* (2.4), we have to run the following algorithm:
· Find the Cholesky decomposition **A** of the 2×2 correlation matrix Σ, i.e. [1 $\hat{\rho}_t$, $\hat{\rho}_t$ 1], where $\hat{\rho}_t$ is the forecasted conditional correlation, estimated with the dynamic model described in section 2.3.3;
· Simulate 2 independent standard normal random variates $\mathbf{z_j} = (z_{j,1}, z_{j,2})'$;
· Simulate a random variate, s_j , from a $\chi^2_{\hat{\upsilon}_t}$ distribution, independent of z_j, where $\hat{\upsilon}_t$ is the forecasted degrees of freedom parameter, estimated with the dynamic model described in section 2.3.3;
· Determine the vector $\mathbf{b_j = Az_j}$;
· Set $\mathbf{c_j} = \frac{\sqrt{\hat{\upsilon}_t}}{\sqrt{s_j}}\mathbf{b_j}$
· Determine the components $(u_{j,x}, u_{j,y})' = (t_{\hat{\upsilon}_t}(c_{j,1}), t_{\hat{\upsilon}_t}(c_{j,2}))'$, where $t_{\hat{\upsilon}_t}(\cdot)$ is the Student's T c.d.f. with degrees of freedom equal to υ_t.
The vector $(u_{j,x}, u_{j,y})'$ is a j random variate from the 2-dimensional T-copula $C_{T-cop}(\cdot ; \hat{\rho}_t , \hat{\upsilon}_t | \Phi_{t-1})$.

(b) Second, we get the standardized asset log-returns by using the inverse functions of the estimated marginals, which can be Normal / Skew-Normal / T-Student / Skew-T, as described in section 2.2 and Appendix A:

$$\mathbf{Q_j} = (q_{j,x}, q_{j,x})' = (\hat{F}_t^{-1}(u_{j,x}); \hat{G}_t^{-1}(u_{j,x}))$$

(c) Third, we rescale the standardized assets log-returns by using the forecasted means and variances, estimated with AR-GARCH models as described in section 2.3.2:

$$\{x_{j,t},\, y_{j,t}\} \;=\; \left(\hat{\mu}_{x,t} + q_{j,x}\sqrt{\hat{h}_{x,t}}\,,\; \hat{\mu}_{y,t} + q_{j,y}\sqrt{\hat{h}_{y,t}}\right)$$

(d) Finally, we repeat this procedure for j = 100, 000 times.

3. By using these 100,000 scenarios, the portfolio is being revaluated at time t, that is:

$$P_{z,t}^{\;j} = P_{x,t-1}\exp(x_{j,t}) + P_{y,t-1}\exp(y_{j,t}), \quad \text{j} = 1,\ldots,100.000$$

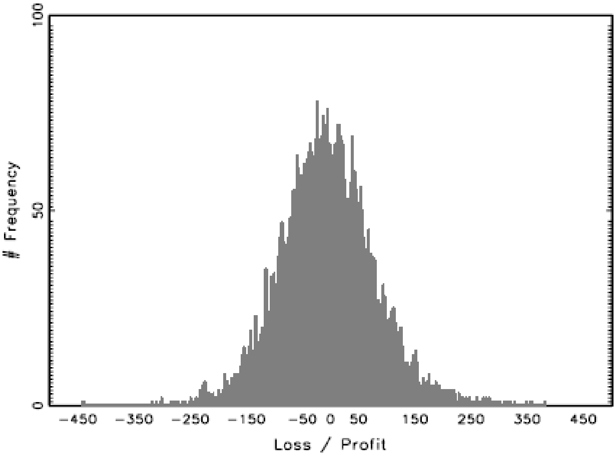
4. Portfolio Losses in each scenario j are then computed:

$$\text{Loss}_j = P_{z,t}^{\;j} - P_{z,t\text{-}1}, \qquad j = 1, \ldots, 100, 000$$

5. The calculus of 95%, 99% VaR is very simple:
   a) We order the 100,000 Loss_j in increasing order;
   b) 99% VaR is the 1000th ordered scenario;
   c) 95% VaR is the 5000th ordered scenario.

As an example, Figure 4 shows a simulated profit and loss distribution for a given portfolio.

Figure 4: **Profit & Loss distribution for a given portfolio z.**



We want to remark that the traditional joint normal distribution is a special case nested in the more general copula-marginals framework: looking at equation (2.2), we can easily see that a random vector (X,Y) is multivariate normal if the univariate margins F_t, G_t are Normal, and the dependence structure among the margins is described by the Normal copula (2.3).

4 - VaR evaluation and results

The goal of this work is to evaluate what are the main determinants when doing VaR forecasts for a portfolio of assets. We saw in Section 2 that four elements have to be considered in a conditional multivariate analysis:

1. The choice of the marginals distribution;
2. The specification of the conditional moments of the marginals, ranging from the mean till the kurtosis;
3. The choice of the copula;
4. The specification of the conditional copula parameters

The in-sample results showed that an AR(1)-TGARCH(1,1) model with a dynamic Skew-T distribution resulted to be the best choice for marginals specifications, while a simple normal copula with constant correlation was the best option for the bivariate asset pairs.

In order to evaluate how important are the cited four elements, we generate out-of-sample VaR forecasts for the considered three portfolios (SP500-DAX, SP500-NIKKEI225, NIKKEI225-DAX), by using these different conditional multivariate distributions:

• *Dynamic Normal copula + Normal marginals*, with the following specifications:

1. Constant mean and Constant Variance;
2. AR(1) specification for the mean and Constant Variance;
3. AR(1) specification for the mean and GARCH(1,1) for the Variance;
4. AR(1) specification for the mean and T-GARCH(1,1) for the Variance;

• *Dynamic Normal copula + Skew-Normal marginals*, with the following specifications:

1. Constant mean, Constant Variance, Constant Skewness parameter;
2. AR(1) specification for the mean, Constant Variance, and Constant Skewness parameter;
3. AR(1) specification for the mean and GARCH(1,1) for the Variance, and Constant Skewness parameter;
4. AR(1) specification for the mean and T-GARCH(1,1) for the Variance, and Constant Skewness parameter;
5. AR(1) specification for the mean and T-GARCH(1,1) for the Variance, and Dynamic Skewness parameter given by (2.12)

• *Dynamic Normal copula + Student's T marginals*, with the following specifications:

1. Constant mean, Constant Variance and Constant Degrees of Freedom;
2. AR(1) specification for the mean, Constant Variance, and and Constant Degrees of Freedom;
3. AR(1) specification for the mean and GARCH(1,1) for the Variance, and Constant Degrees of Freedom;
4. AR(1) specification for the mean and T-GARCH(1,1) for the Variance, and Constant Degrees of Freedom;
5. AR(1) specification for the mean and T-GARCH(1,1) for the Variance, and Dynamic Degrees of Freedom given by (2.13);

• *Dynamic Normal copula + Skew-T marginals*, with the following specifications:

1. Constant mean, Constant Variance, Constant Skewness parameter, and Constant Degrees of Freedom;
2. AR(1) specification for the mean, Constant Variance, Constant Skewness parameter, and Constant Degrees of Freedom;
3. AR(1) specification for the mean and GARCH(1,1) for the Variance, Constant Skewness parameter, and Constant Degrees of Freedom;
4. AR(1) specification for the mean and T-GARCH(1,1) for the Variance, Constant Skewness parameter, and Constant Degrees of Freedom;
5. AR(1) specification for the mean and T-GARCH(1,1) for the Variance, Dynamic Skewness parameter given by (2.12), and Constant Degrees of Freedom;
6. AR(1) specification for the mean and T-GARCH(1,1) for the Variance, Dynamic Skewness parameter given by (2.12), and Dynamic Degrees of Freedom given by (2.13);

• *Constant Normal copula + Normal / Skew-Normal / Student's T /Skew-T marginals*, with the previous specifications;

• *Dynamic T-copula + Normal / Skew-Normal / Student's T /Skew-T marginals*, with the previous specifications;

• *Constant T-copula + Normal / Skew-Normal / Student's T /Skew-T marginals*, with the previous specifications.

We generate portfolio Value at Risk forecasts at the 95 % - 99 % confidence levels, following the procedure described in section 3.2. The predicted one-step-ahead VaR forecasts are then compared with the observed portfolio losses and both results are recorded for later assessment using a back-testing procedure. The first estimation sample is given by the first 700 observation: at the j-th iteration where j goes from 700 to 1699 (for a total of 1000 observations), the estimation sample is augmented to include one more observation, and this procedure is iterated until all days have been included in the estimation sample.

We finally assess the performance of the competing models over these 1000 observations using the following back-testing techniques:

• Kupiec's unconditional coverage test;

• Christoffersen's conditional coverage test;

• Loss functions to evaluate VaR forecasts accuracy.

4.1 Unconditional model evaluation: Kupiec's test

This test is based on binomial theory and tests the difference between the observed and expected number of VaR exceedances of the effective portfolio losses. Since VaR is based on a confidence level p, when we observe N losses in excess of VaR out of T observations, hence we observe N/T proportion of excessive losses: the Kupiec's test answers the question whether N/T is statistically significantly different from p.

Following binomial theory, the probability of observing N failures out of T observations is $(1-p)^{T-N} p^N$, so that the test of the null hypothesis that the expected exception frequency $N/T = p$ is given by a likelihood ratio test statistic:

$$LR_{UC} = -2 \cdot \ln[(1-p)^{T-N} p^N] + 2 \cdot \ln[(1 - N/T)^{T-N} (N/T)^N] \qquad (4.1)$$

which is distributed as $\chi^2(1)$ under the H_0 . This test can reject a model for both high and low failures but, as stated in Kupiec (1995), its power is generally poor.

4.2 Conditional model evaluation: Christoffersen's test

A more complete test was made by Christoffersen (1998), who developed a likelihood ratio statistic to test the joint assumption of unconditional coverage and independence of failures. Its main advantage over the previous statistic is that it takes account of any conditionality in our forecast: for example, if volatilities are low in some period and high in others, the VaR forecast should respond to this clustering event. The Christoffersen procedure enables us to separate clustering effects from distributional assumption effects. His statistic is computed as:

$$LR_{CC} = -2 \ln[(1-p)^{T-N} p^N] + 2 \ln[(1 - \pi_{01})^{n00} \pi^{n01}{}_{01} (1 - \pi_{11})^{n10} \pi^{n11}{}_{11}] \qquad (4.2)$$

where n_{ij} is the number of observations with value i followed by j for $i, j = 0,1$ and

$$\pi_{ij} = n_{ij} \,/\, (\Sigma_j \;\; n_{ij}) \tag{4.3}$$

are the corresponding probabilities. Under the H_0, this test is distributed as a $\chi^2(2)$. If the sequence of N losses is independent, then the probabilities to observe or not observe a VaR violation in the next period must be equal, which can be written more formally as $\pi_{01} = \pi_{11} = p$. The main advantage of this test is that it can reject a VaR model that generates either too many or too few clustered violations, although it needs several hundred observations in order to be accurate.

4.3 Measures of accuracy: loss functions

The previous tests do not show any power in distinguishing among different, but close, alternatives. Moreover, as noted by the Basle Committee on Banking Supervision (1998), the magnitude as well as the number of exceptions are a matter of regulatory concern. This concern can be readily incorporated into a so called "loss function" by introducing a magnitude term. This was first accomplished by Lopez (1998). The general form of these loss functions is:

$$C_{t+1} = \begin{cases} f\;(L_{t+1},\; VaR_{t+1}) & \textbf{if} \quad L_{t+1} \;>\; VaR_{t+1} \\ g\;(L_{t+1},\; VaR_{t+1}) & \textbf{if} \quad L_{t+1} \;\leq\; VaR_{t+1} \end{cases}$$

where $f(x, y)$ and $g(x, y)$ are arbitrary functions such that $f(x, y) \geq g(x, y)$ for a given y, while L is the portfolio loss. Under very general conditions, accurate VaR estimates will generate the lowest possibile score. Many loss functions can be constructed. Lopez (1998) has proposed the following quadratic loss function:

$$C_{t+1} = \begin{cases} 1 + (L_{t+1} - VaR_{t+1})^2 & \textbf{if} \quad L_{t+1} \;>\; VaR_{t+1} \\ 0 & \textbf{if} \quad L_{t+1} \;\leq\; VaR_{t+1} \end{cases} \tag{4.4}$$

Thus, as before, a score of one is imposed when an exception occurs, but now, an additional term based on its magnitude is included. The numerical score increases with the magnitudo of the exception and can provide additional information on how the underlying VaR model forecasts the lower tail of the underlying L_{t+1} distribution.

Blanco and Ihle (1999) suggest an alternative way to deal with the size of exceptions by focusing on the average size of the exceptions:

$$C_{t+1} = \begin{cases} \dfrac{L_{t+1} - VaR_{t+1}}{VaR_{t+1}} & \textbf{if} \quad L_{t+1} > VaR_{t+1} \\ 0 & \textbf{if} \quad L_{t+1} \leq VaR_{t+1} \end{cases} \tag{4.5}$$

A strategy that can be implemented with the previous approaches is the following:

1. Apply Kupiec's and Christoffersen tests at a first stage to choose the best models;
2. Then use the loss functions to compare the costs of different admissible choices.

4.3 VaR out-of-sample results

Tables 4-6 report the real VaR exceedances *N/T*, the *p*-value p_{UC} of Kupiec's Unconditional Coverage test, and the *p*-value p_{CC} of Christoffersen's Conditional Coverage test, for the the VaR forecasts at 95% - 99% confidence levels, relative to the three different considered portfolios (SP500-DAX, SP500-NIKKEI225, NIKKEI225-DAX).

Table 4: **VaR results: SP500-DAX portfolio.**

	DYNAMIC NORMAL COPULA	**SP500 - DAX**					
		95%			**99%**		
Marg. Distr.	*Moment specification*	N/T	p_{UC}	p_{CC}	N/T	p_{UC}	p_{CC}
NORMAL	Constant: Mean, Variance	11,59%	0,000	0,000	5,39%	0,000	0,000
	AR(1) Constant, Variance	11,49%	0,000	0,000	5,19%	0,000	0,000
	AR(1) GARCH(1,1)	6,69%	0,019	0,058	1,40%	0,232	0,396
	AR(1) T-GARCH(1,1)	5,89%	0,206	0,303	0,70%	0,312	0,567
SKEW-NORMAL	Constant: Mean, Variance, Skewness	10,49%	0,000	0,000	4,60%	0,000	0,000
	AR(1) Constant: Variance, Skewness	10,99%	0,000	0,000	4,70%	0,000	0,000
	AR(1) GARCH(1,1) Const. Skewness	5,79%	0,260	0,469	1,10%	0,757	0,834
	AR(1) T-GARCH(1,1) Const. Skewness	5,09%	0,891	0,910	0,60%	0,169	0,372
	AR(1) T-GARCH(1,1) Dyn. Skewness	4,30%	0,295	0,550	0,20%	0,002	0,008
STUDENT'S T	Constant: Mean, Variance, D.o.F.	12,59%	0,000	0,000	4,10%	0,000	0,000
	AR(1) Constant: Variance, D.o.F.	13,89%	0,000	0,000	4,30%	0,000	0,000
	AR(1) GARCH(1,1) Constant D.o.F.	8,19%	0,000	0,000	1,10%	0,757	0,834
	AR(1) T-GARCH(1,1) Constant D.o.F.	7,69%	0,000	0,001	0,70%	0,312	0,567
	AR(1) T-GARCH(1,1) Dynamic D.o.F.	6,99%	0,006	0,019	1,00%	0,997	0,895
SKEW-T	Const: Mean, Variance, Skew., D.o.F.	11,99%	0,000	0,000	3,50%	0,000	0,000
	AR(1) Const: Variance, Skew, D.o.F.	12,89%	0,000	0,000	4,10%	0,000	0,000
	AR(1) GARCH(1,1) Const: Skew, D.o.F.	6,89%	0,009	0,026	0,80%	0,508	0,747
	AR(1) T-GARCH(1,1) Const: Skew, D.o.F.	6,19%	0,094	0,193	0,50%	0,078	0,206
	AR(1)T-GARCH(1,1) Dyn Skew,Const D.o.F.	5,09%	0,891	0,653	0,40%	0,030	0,093
	AR(1) T-GARCH(1,1) Dyn Skew,Dyn D.o.F.	4,50%	0,457	0,312	0,30%	0,009	0,032

Table 5: **VaR results: SP500-NIKKEI portfolio.**

	DYNAMIC NORMAL COPULA	SP500 - NIKKEI					
		95%			99%		
Marg. Distr.	*Moment specification*	N/T	p_{UC}	p_{CC}	N/T	p_{UC}	p_{CC}
NORMAL	Constant: Mean, Variance	7,29%	0,002	0,007	2,90%	0,000	0,000
	AR(1) Constant Variance	8,09%	0,000	0,000	3,70%	0,000	0,000
	AR(1) GARCH(1,1)	5,79%	0,260	0,489	0,90%	0,744	0,866
	AR(1) T-GARCH(1,1)	4,30%	0,295	0,430	0,20%	0,002	0,008
SKEW-NORMAL	Constant: Mean, Variance, Skewness	7,29%	0,002	0,007	3,00%	0,000	0,000
	AR(1) Constant: Variance, Skewness	8,09%	0,000	0,000	3,90%	0,000	0,000
	AR(1) GARCH(1,1) Constant Skewness	5,69%	0,324	0,573	1,00%	0,997	0,895
	AR(1) T-GARCH(1,1) Constant Skewness	4,40%	0,371	0,640	0,20%	0,002	0,008
	AR(1) T-GARCH(1,1) Dynamic Skewness	3,10%	0,003	0,004	0,10%	0,000	0,001
STUDENT'S T	Constant: Mean Variance D.o.F.	8,09%	0,000	0,000	1,50%	0,140	0,155
	AR(1) Constant: Variance D.o.F.	9,69%	0,000	0,000	2,40%	0,000	0,000
	AR(1) GARCH(1,1) Constant D.o.F.	6,79%	0,013	0,041	0,90%	0,744	0,173
	AR(1) T-GARCH(1,1) Constant D.o.F.	6,09%	0,124	0,285	0,30%	0,009	0,032
	AR(1) T-GARCH(1,1) Dynamic D.o.F.	5,49%	0,480	0,736	0,30%	0,009	0,032
SKEW-T	Constant: Mean, Variance,Skewness, D.o.F.	7,99%	0,000	0,000	1,50%	0,140	0,155
	AR(1) Constant: Variance, Skewness, D.o.F.	9,49%	0,000	0,000	2,10%	0,002	0,002
	AR(1) GARCH(1,1) Const: Skew D.o.F.	6,79%	0,013	0,041	0,70%	0,312	0,567
	AR(1) T-GARCH(1,1) Const: Skew, D.o.F.	5,79%	0,260	0,469	0,50%	0,078	0,206
	AR(1)T-GARCH(1,1) Dyn Skew,Const D.o.F.	4,40%	0,371	0,485	0,20%	0,002	0,008
	AR(1) T-GARCH(1,1) Dyn Skew, Dyn D.o.F.	4,20%	0,230	0,314	0,10%	0,000	0,001

Table 6: **VaR results: NIKKEI-DAX portfolio**

	DYNAMIC NORMAL COPULA	NIKKEI - DAX					
		95%			99%		
Marg. Distr.	*Moment specification*	N/T	p_{UC}	p_{CC}	N/T	p_{UC}	p_{CC}
NORMAL	Constant: Mean, Variance	6,99%	0,006	0,004	3,00%	0,000	0,000
	AR(1) Constant Variance	7,99%	0,000	0,000	3,90%	0,000	0,000
	AR(1) GARCH(1,1)	5,09%	0,891	0,865	1,00%	0,997	0,895
	AR(1)T-GARCH(1,1)	4,00%	0,132	0,058	0,30%	0,009	0,032
SKEW-NORMAL	Constant: Mean, Variance, Skewness	6,99%	0,006	0,008	3,20%	0,000	0,000
	AR(1) Constant: Variance Skewness	7,99%	0,000	0,000	3,90%	0,000	0,000
	AR(1) GARCH(1,1) Constant Skewness	5,00%	0,994	0,895	0,90%	0,744	0,866
	AR(1) T-GARCH(1,1) Constant Skewness	4,10%	0,176	0,067	0,50%	0,078	0,206
	AR(1) T-GARCH(1,1) Dynamic Skewness	3,00%	0,002	0,003	0,20%	0,002	0,008
STUDENT'S T	Constant: Mean, Variance, D.o.F.	7,79%	0,000	0,000	1,30%	0,364	0,551
	AR(1) Constant: Variance, D.o.F.	9,29%	0,000	0,000	2,10%	0,002	0,002
	AR(1) GARCH(1,1) Constant D.o.F.	6,79%	0,013	0,035	0,80%	0,508	0,747
	AR(1) T-GARCH(1,1) Constant D.o.F.	5,59%	0,397	0,582	0,50%	0,078	0,206
	AR(1) T-GARCH(1,1) Dynamic D.o.F.	5,00%	0,994	0,902	0,50%	0,078	0,206
SKEW-T	Constant: Mean, Variance, Skewness, D.o.F.	7,79%	0,000	0,000	1,40%	0,232	0,201
	AR(1) Constant: Variance, Skewness, D.o.F.	8,89%	0,000	0,000	1,70%	0,043	0,012
	AR(1) GARCH(1,1) Const: Skewness D.o.F.	6,39%	0,052	0,142	0,80%	0,508	0,747
	AR(1) T-GARCH(1,1) Const: Skew D.o.F.	5,00%	0,994	0,902	0,50%	0,078	0,206
	AR(1)T-GARCH(1,1) Dyn Skew,Const D.o.F.	4,00%	0,132	0,293	0,20%	0,002	0,008
	AR(1) T-GARCH(1,1) Dyn Skew, Dyn D.o.F.	3,60%	0,032	0,094	0,10%	0,000	0,001

The two loss functions discussed in Section 4.3 are presented in Appendix B, Tables 10 -12, instead.

The results using the Normal copula with constant correlation, as well as the T-copula with dynamic and constant parameters, delivered almost identical results, thus confirming the in-sample results: a simple normal copula with constant correlation is sufficient to describe the dependence structure of the considered risk-factors. For sake of space, we report only the VaR results relative to the AR(1)-TGARCH(1,1) specification with a dynamic Skew-T distribution. The complete set of results is available from the author.

Table 7: **VaR results: SP500-DAX portfolio**
(Dynamic Skew-T + other copulas)

		95%			99%		
Marg. Distr.	*Copula specification*	N/T	p_{UC}	p_{CC}	N/T	p_{UC}	p_{CC}
Dynamic SKEW-T	CONSTANT NORMAL COPULA	4,40%	0,371	0,249	0,10%	0,000	0,001
Dynamic SKEW-T	DYNAMIC T-COPULA	4,30%	0,295	0,193	0,20%	0,002	0,008
Dynamic SKEW-T	CONSTANT T-COPULA	4,40%	0,371	0,193	0,20%	0,002	0,008

Table 8: **VaR results: SP500-NIKKEI portfolio**
(Dynamic Skew-T + other copulas)

		95%			99%		
Marg. Distr.	*Copula specification*	N/T	p_{UC}	p_{CC}	N/T	p_{UC}	p_{CC}
Dynamic SKEW-T	CONSTANT NORMAL COPULA	4,10%	0,176	0,241	0,10%	0,000	0,001
Dynamic SKEW-T	DYNAMIC T-COPULA	4,20%	0,230	0,314	0,10%	0,000	0,001
Dynamic SKEW-T	CONSTANT T-COPULA	4,10%	0,176	0,241	0,10%	0,000	0,001

Table 9: **VaR results: NIKKEI-DAX portfolio**
(Dynamic Skew-T + other copulas)

		95%			99%		
Marg. Distr.	*Copula specification*	N/T	p_{UC}	p_{CC}	N/T	p_{UC}	p_{CC}
Dynamic SKEW-T	CONSTANT NORMAL COPULA	3,50%	0,021	0,067	0,20%	0,002	0,008
Dynamic SKEW-T	DYNAMIC T-COPULA	3,60%	0,032	0,082	0,10%	0,000	0,001
Dynamic SKEW-T	CONSTANT T-COPULA	3,50%	0,021	0,067	0,10%	0,000	0,001

The previous tables highlights some interesting results:

1. The GARCH specification for the variance is absolutely fundamental to have good VaR forecasts, whatever the marginal distribution is;

2. The asymmetric GARCH specification is important to get precise VaR estimates at the 95% confidence level. However, it can produce conservative estimates at the 99% level when dealing with strongly leptokurtic assets and the normal (or skew-normal) distribution is used;

3. The AR specification of the mean is not relevant in all cases;

4. The GARCH specification seems to model most of the leptokurtosis present in the data. However, when the assets are strongly leptokurtic, a Skew-T distribution is the best choice (similarly to point 2.);

5. Using a Student's T distribution with strongly skewed assets can produce very aggressive VaR forecasts at the 95% level;

6. The Skew-T and Skew-Normal distributions present the most precise VaR forecasts, according to the tests and Loss functions used;

7. Allowing for dynamics in the skewness and degrees of freedom parameters produces more conservative VaR forecasts in almost all cases;

8. The type of copula as well as the dynamics in its parameters are not relevant: a simple normal copula with constant correlation resulted to be sufficient in all cases.

These results seem to point out that a Skew-T distribution with a T-GARCH(1,1) specification and constant skewness and degrees of freedom parameters for the marginals, together with a constant normal copula, should be a good compromise for precise VaR estimates when dealing with a portfolio of different assets.

This evidence is consistent with other results reported in Marshal and Zeevi (2002) and Chen, Fan, and Patton (2004) that found the normal copula to be a plausible choice when dealing with bivariate portfolios, while this is not the case for larger collections of financial assets. Particularly, Chen, Fan, and Patton (2004) point out that the Student's t copula is a better approximation to the true copula when portfolio dimensionality increases, but the possibility of misspecification still exists.

5 – Conclusions

We have described in this work a model for estimating portfolio VaR by Monte Carlo simulation and conditional multivariate distributions. Traditional risk measurement models assume that the risk factors distribution is multivariate normal: unfortunately empirical evidence shows that the true marginal distributions of financial assets usually present fat tails and strong skewness. We therefore introduced a model to generate scenarios for portfolio risk factors from different conditional multivariate distributions, to find out what are the main determinants when doing VaR forecasts for a portfolio of assets: we used dynamic Gaussian or T-copulas to model assets' dependence structure, together with different dynamic marginal specifications.

We estimated the 95 %, 99% VaR for three different portfolios composed of the SP500 and the DAX, the SP500 and the NIKKEI225, and the NIKKEI225 and the DAX, respectively, using 100,000MC scenarios. We then compared the accuracy of the different VaR estimates with a backtesting procedure over a time window of 1000 daily observations: we apply Kupiec's and Christoffersen's tests in the first stage, and loss functions in the second stage to compare the costs of different admissible choices.

The empirical analysis showed that correct marginals specification is absolutely crucial for VaR forecasting, while copula specification plays a minor role. Particularly, a T-GARCH(1,1) specification together with a Skew-T or Skew-normal distribution gave the best results. The use of symmetric distributions, such as the Student's T, with skewed assets, produced aggressive VaR estimates at the 95 % confidence level, instead.

The analysis showed that the choice of copula was not relevant for correct VaR estimates, and a simple constant normal copula was sufficient in all cases. This evidence confirms previous results for bivariate portfolios but with a different approach, since we used a VaR backtesting procedure instead of goodness-of-fit tests for copulas as in Chen, Fan, and Patton (2004).

Possible extensions that can be considered for future research include the passage from conditional bivariate to conditional multivariate copula, following the suggestion of Engle (2004) who recently proposed the Rank-DCC model. Secondly, a solution to the trade-off between parsimonious modelling and unbiased estimates may be given by a copula factor model. The direct consideration of transaction costs can surely help in detecting which factors are useful and which are not. Thirdly, explore the features of VaR with portfolios composed of different financial assets, such as bonds or exchange rates.

References

Angelidis, T., A. Benos, and S. Degiannakis, 2003. The Use of GARCH Models in VaR Estimation. Discussion paper, University of Piraeus.

Artzner, P., F. Delbaen, J. Eber, and D. Heath, 1997. Thinking Coherently. *Risk*, 10(11), 68–71.

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

Basle Committee on Banking Supervision,1998, 2005. Amendment to the Capital Accord to Incorporate Market Risks.

Bauwens, L., and S. Laurent , 2005. A New Class of Multivariate Skew Densities, with Application to GARCH Models. *Journal of Business and Economic Statistics*, 23(3), 46–354.

Blanco, C., and G. Ihle, 1999. How good is your VaR? Using backtesting to assess system performance. *Financial Engineering News*, August, 1-2.

Bollerslev, T., R. Engle, and D. Nelson, 1994. ARCH Models, vol. 4 of *Handbook of Econometrics*. (Elsevier, Amsterdam).

Bouye', E., V. Durrleman, A. Nikeghbali, G. Riboulet, and T. Roncalli, 2001. Copulas: an open field for risk management, Discussion paper, Groupe de Recherche Operationnelle, Credit Lyonnais.

Chen, X., Y. Fan, and A. Patton, 2004. Simple Tests for Models of Dependence Between Multiple Financial Time Series with Applications to U.S. Equity Returns and Exchange Rates. *Discussion Paper n. 483, Financial Markets Group, London School of Economics*.

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

Chong, J., 2004. Value at Risk from Econometric Models and Implied from Currency Options. *Journal of Forecasting*, 23, 603–620.

Christoffersen, P. F. 1998. *Evaluating Interval Forecasts*, *International Economic Review*, 39, 841– 862.

Crouhy, M., D. Galai, and R. Mark, 2001. *Risk Management*. (McGraw Hill, New York).

Dowd, K., 1998. *Beyond Value at Risk: The New Science of Risk Management*. (John Wiley and Sons, New York).

Embrechts, P. 2000. Extreme Value Theory: Potential and Limitations as an Integrated Risk Management Tool. *Derivatives Use, Trading and Regulation*, 6(1), 449–456.

Embrechts, P., and A. Dias, 2003. Dynamic copula models for multivariate high-frequency data in finance. Discussion paper, Department of Mathematics ETH Zurich.

Embrechts, P., A. Hoing, and A. Juri, 2003. Using copulae to bound the Value at Risk for functions of dependent risks. *Finance and Stochastics*, 7(2), 145–167.

Embrechts, P., C. Kluppelberg, and T. Mikosch 1997. *Modelling Extreme Events*. (Springer Verlag, Berlin).

Embrechts, P., F. Lindskog, and A. J. McNeil, 2003. Modelling dependence with copulas and applications to risk management, in *Handbook of heavy tailed distributions in finance*, (Elsevier/North-Holland, Amsterdam).

Engle, R., 2004. Dynamic Conditional Correlation: Econometric Results and Financial Applications, Seminar Universidad Carlos III Madrid, May 24 2004.

Engle, R., and K. Sheppard, 2001. Theorical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH, Mimeo, UCSD.

Fang, H., and K. Fang, 2002. The metaelliptical distributions with given marginals. *Journal of Multivariate Analysis*, 82, 1–16.

Fisher-Report (1994): *Basle Euro-currency Standing Committe of the Central Banks of the Group of Ten Countries. Public disclosure of market and credit risks by financial intermediaries*.

Giot, P., and S. Laurent, 2003. Value-at-Risk for Long and Short Positions. *Journal of AppliedEconometrics*, 18, 641–663.

Glasserman, P., P. Heidelberger, and P. Shahabuddin, 2002. Portfolio Value-at-Risk with Heavily-Tailed Risk Factors. *Mathematical Finance*, 12(3), 239–269.

Group of-Thirty, 1993. *Derivatives: practices and principles. Global Derivatives Study Group* (New York).

Hansen, B. E., 1994. Autoregressive Conditional Density Estimation, *International Economic Review*, 35(3), 705–730.

Hull, J., and A. White, 1998. Value at risk when daily changes in market variables are not normally distributed. *Journal of derivatives*, 5, 9–19.

Joe, H., and J. Xu, 1996. The estimation method of inference functions for margins for multivariate models, Discussion Paper, Department of Statistics, University of British Columbia.

Joe, U. 1997. *Multivariate Models and Dependence Concepts*. (Chapman Hall, London).

Jondeau, E., and M. Rockinger, 2003. Conditional volatility, skewness, and kurtosis: existence, persistence, and comovements. *Journal of Economic Dynamics and Control*, 27, 1699 – 1737.

Jorion, P. 1997. *Value at Risk. The New Benchmark for Controlling Derivatives Risk.* (McGraw Hill, New York).

Jorion, P. 2000. *Value at Risk: The New Benchmark for Managing Financial Risk.* (2nd Edition, McGraw Hill, New York).

Kupiec, P., 1995. Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives*, 2, 173–184.

Lopez, J. 1998. Methods for evaluating value-at-risk estimates. *Federal Reserve Bank of New York research paper*, n. 9802.

Lucas, A., and P. Klaassen, 1998. Extreme returns, downside risk and optimal asset allocation. *The Journal of Portfolio Management*, 25, 71–79.

Marshal, R., and A. Zeevi, 2002. Beyond Correlation: Extreme Co-movements Between Financial Assets, Discussion paper, Columbia University.

McNeil, A., R. Frey, and P. Embrechts, 2005. *Quantitative Risk Management: Concepts, Techniques and Tools*. (Princeton University Press).

Nelsen, R., 1999. *An Introduction to Copulas*. Lecture Notes in Statistics 139, (Springer, N.Y.).

Nelson, D., 1991. Conditional Heteroskedasticity in Asset Returns: A New Approach, *Journal of Econometrics*, 43, 227–251.

Newey, W., and D. Steigerwald, 1997. Asymptotic Bias for Quasi Maximum Likelihood Estimators in Conditional Heteroskedasticity Models, *Econometrica*, 3, 587–599.

Patton, A., 2004. On the Out-of-Sample Importance of Skewness and Asymmetric Dependence for Asset Allocation, *Journal of Financial Econometrics*, 2(1), 130 – 168.

Patton, A., 2006. Modelling Asymmetric Exchange Rate Dependence, International Economic Review, 47(2).

Rockinger, M., and E. Jondeau, 2001. Conditional dependency of financial series: an application of copulas. *Discussion Paper NER 82, Banque de France*.

Sklar, A., 1959. Fonctions de repartition `a n dimensions et leurs marges. *Publ. Inst. Statis. Univ. Paris*, 8, 229–231.

Tsay, R. S., 2002. *Analysis of Financial Time Series*. (Wiley, New York).

Appendices

A. Skew-T Cdf and Inverse Cdf

A.1 CDF

Proposition 3: *Let X be a standard Student's T random variable, with mean zero and variance (υ/(υ-2)), and define the c.d.f. of X as F(t; υ). The c.d.f. of a Skewed-t random variable y with mean zero and variance 1, in terms of F(t;υ) is the following:*

$$G(y;\upsilon,\lambda)\begin{cases}(1-\lambda)\cdot F\left(\sqrt{\dfrac{\upsilon}{\upsilon-2}}\left(\dfrac{by+a}{1-\lambda}\right);\ \upsilon\right), & for\ \ y<-\dfrac{a}{b}\\ (1-\lambda)/2 & for\ \ y=-\dfrac{a}{b}\\ (1+\lambda)\cdot F\left(\sqrt{\dfrac{\upsilon}{\upsilon-2}}\left(\dfrac{by+a}{1+\lambda}\right);\ \upsilon\right), & for\ \ y>-\dfrac{a}{b}\end{cases}$$

where a, b, c where defined before for the density.

Proof:
a) Consider the case where $y < a/b$. Given the skewed-t density we have:

$$G(y;\upsilon,\lambda) = \int_{-\infty}^{t} bc\left(1+\frac{1}{\upsilon-2}\left(\frac{by+a}{1-\lambda}\right)^2\right)^{-(\upsilon+1)/2} dy =$$

We now change variable, substituting $u = (by + a)/(1 - \lambda)$ and the relative Jacobian:

$$= \int_{-\infty}^{(by+a)/(1-\lambda)} \frac{\Gamma((\upsilon+1)/2)}{\Gamma(\upsilon/2)} \frac{1}{\sqrt{\pi(\upsilon-2)}}\left(1+\frac{u^2}{\upsilon-2}\right)^{-(\upsilon+1)/2} du =$$

We change the variable again, substituting $z = \sqrt{\frac{\upsilon}{\upsilon-2}} \cdot u$ and its Jacobian. By recalling that the standard Student's T Cdf is:

$$F(t) = \int_{-\infty}^{t} \frac{\Gamma((\upsilon+1)/2)}{\Gamma(\upsilon/2)} \frac{1}{\sqrt{\pi\upsilon}}\left(1+\frac{x^2}{\upsilon}\right)^{-(\upsilon+1)/2} dx =$$

We finally get the Cdf:

$$G(y;\upsilon,\lambda) = (1-\lambda)F\left(\left(\frac{by+a}{1-\lambda}\right)\sqrt{\frac{\upsilon}{\upsilon-2}};\upsilon\right)$$

b) When $y = a/b$

$$G(y;\upsilon,\lambda) = \frac{(1-\lambda)}{2}$$

c) When $y > a/b$

$$G(y;\upsilon,\lambda) = G(-a/b;\upsilon,\lambda) + \int_{-a/b}^{t} bc\left(1+\frac{1}{\upsilon-2}\left(\frac{by+a}{1-\lambda}\right)^2\right)^{-(\upsilon+1)/2} dy =$$

By using analogous computations we have...

$$G(y;\upsilon,\lambda) = \frac{(1-\lambda)}{2} + (1+\lambda)F\left(\left(\frac{by+a}{1-\lambda}\right)\sqrt{\frac{\upsilon}{\upsilon-2}};\upsilon\right) - \frac{(1+\lambda)}{2} =$$

$$= \quad (1+\lambda)F\left(\left(\frac{by+a}{1-\lambda}\right)\sqrt{\frac{\upsilon}{\upsilon-2}};\upsilon\right) - \lambda$$

A.2 Inverse CDF

We express the inverse c.d.f. of the Skewed-t distribution in terms of the inverse distribution of a Student's T random variable, denoted as $F^{-1}(u;\ \upsilon)$. If we indicate with u = $G(y;\ \upsilon,\ \lambda)$ an uniform variate with support [0; 1], and we use the previous CDF, we get

$$G^{-1}(y;\upsilon,\lambda)\begin{cases} \dfrac{1}{b}\left[(1-\lambda)\sqrt{\dfrac{\upsilon-2}{\upsilon}}\cdot F^{-1}\left(\dfrac{u}{1-\lambda};\ \upsilon\right)-a\right], & for\ 0<u<\dfrac{1-\lambda}{2} \\ \dfrac{1}{b}\left[(1+\lambda)\sqrt{\dfrac{\upsilon-2}{\upsilon}}\cdot F^{-1}\left(\dfrac{u+\lambda}{1+\lambda};\ \upsilon\right)-a\right], & for\ \dfrac{1-\lambda}{2}\le u<1 \end{cases}$$

The inverse Cdf can be used to generate random draws from the skewed t distribution with the following algorithm:

1. Simulate n uniform draws from the *Uniform*(0, 1) distribution;
2. Then use the previous formula to compute $y_t = G^{-1}(u_t;\ \upsilon,\ \lambda)$

B. Loss functions tables

Table 10: **Loss functions: SP500-DAX portfolio.**

	DYNAMIC NORMAL COPULA	SP500 - DAX			
		95%		99%	
Marg. Distr.	*Moment specification*	Lopez	Blanco-Ihle	Lopez	Blanco-Ihle
NORMAL	Constant: Mean, Variance	574195,3	67,1	255879,0	22,9
	AR(1) Constant Variance	564928,7	69,5	255875,8	23,9
	AR(1) GARCH(1,1)	130001,5	18,5	8026,1	1,9
	AR(1)T-GARCH(1,1)	82808,6	13,3	1316,8	0,6
SKEW-NORMAL	Constant: Mean, Variance, Skewness	534959,3	60,5	224011,0	19,3
	AR(1) Constant: Variance Skewness	544099,3	64,6	233545,9	20,8
	AR(1) GARCH(1,1) Constant Skewness	111031,0	15,0	4274,1	1,1
	AR(1) T-GARCH(1,1) Constant Skewness	77321,6	11,5	633,7	0,4
	AR(1) T-GARCH(1,1) Dynamic Skewness	50334,3	8,7	211,4	0,1
STUDENT'S T	Constant: Mean, Variance, D.o.F.	661811,0	80,6	190942,5	16,3
	AR(1) Constant: Variance, D.o.F.	771999,6	102,0	236067,5	20,7
	AR(1) GARCH(1,1) Constant D.o.F.	172566,8	23,9	7146,9	1,5
	AR(1) T-GARCH(1,1) Constant D.o.F.	136215,7	20,0	1631,7	0,6
	AR(1) T-GARCH(1,1) Dynamic D.o.F.	117027,2	17,8	2803,1	0,9
SKEW-T	Constant: Mean, Variance, Skewness, D.o.F.	615342,0	72,5	170421,3	14,1
	AR(1) Constant: Variance, Skewness, D.o.F.	724984,1	92,2	212342,5	18,1
	AR(1) GARCH(1,1) Const: Skewness D.o.F.	143596,1	19,2	4622,0	1,0
	AR(1) T-GARCH(1,1) Const: Skew D.o.F.	113709,0	16,2	800,4	0,4
	AR(1)T-GARCH(1,1) Dyn Skew,Const D.o.F.	63746,9	10,7	81,0	0,1
	AR(1) T-GARCH(1,1) Dyn Skew, Dyn D.o.F.	52449,9	9,3	436,6	0,2

Table 11: **Loss functions: SP500-NIKKEI portfolio.**

	DYNAMIC NORMAL COPULA	SP500 - NIKKEI			
		95%		99%	
Marg. Distr.	*Moment specification*	Lopez	Blanco-Ihle	Lopez	Blanco-Ihle
NORMAL	Constant: Mean, Variance	3851367,0	32,9	1571774,8	9,4
	AR(1) Constant Variance	5038879,0	46,1	2184673,2	13,6
	AR(1) GARCH(1,1)	1477611,9	16,0	162099,1	0,9
	AR(1)T-GARCH(1,1)	756699,7	8,6	22643,9	0,1
SKEW-NORMAL	Constant: Mean, Variance, Skewness	3918815,2	33,8	1610904,9	9,7
	AR(1) Constant: Variance Skewness	5072825,9	46,6	2204347,8	13,8
	AR(1) GARCH(1,1) Constant Skewness	1459672,3	15,7	153811,7	0,8
	AR(1) T-GARCH(1,1) Constant Skewness	746088,8	8,5	17698,4	0,1
	AR(1) T-GARCH(1,1) Dynamic Skewness	338531,4	4,4	1,7	0,0
STUDENT'S T	Constant: Mean, Variance, D.o.F.	4394741,9	39,9	931425,0	4,9
	AR(1) Constant: Variance, D.o.F.	6318888,4	64,4	1538091,5	8,2
	AR(1) GARCH(1,1) Constant D.o.F.	2689904,7	27,1	219203,5	1,2
	AR(1) T-GARCH(1,1) Constant D.o.F.	1662747,4	17,1	41109,0	0,3
	AR(1) T-GARCH(1,1) Dynamic D.o.F.	1376633,1	14,1	67619,3	0,4
SKEW-T	Constant: Mean, Variance, Skewness, D.o.F.	4360381,1	39,4	907213,8	4,8
	AR(1) Constant: Variance, Skewness, D.o.F.	6201146,7	62,3	1448283,1	7,7
	AR(1) GARCH(1,1) Const: Skewness D.o.F.	2574438,8	25,7	191700,8	1,0
	AR(1) T-GARCH(1,1) Const: Skew D.o.F.	1607348,2	16,2	35800,8	0,3
	AR(1)T-GARCH(1,1) Dyn Skew,Const D.o.F.	878962,7	9,2	135,0	0,0
	AR(1) T-GARCH(1,1) Dyn Skew, Dyn D.o.F.	769501,0	8,1	5291,1	0,1

Table 12: **Loss functions: NIKKEI-DAX portfolio**

	DYNAMIC NORMAL COPULA	NIKKEI - DAX			
		95%		99%	
Marg. Distr.	*Moment specification*	**Lopez**	**Blanco-Ihle**	**Lopez**	**Blanco-Ihle**
NORMAL	Constant: Mean, Variance	4510996,4	33,8	1724578,6	9,1
	AR(1) Constant Variance	5802481,2	45,3	2385543,3	13,2
	AR(1) GARCH(1,1)	1730165,7	13,8	192633,7	1,2
	AR(1)T-GARCH(1,1)	956943,6	8,0	21718,4	0,3
SKEW-NORMAL	Constant: Mean, Variance, Skewness	4626696,7	35,0	1788903,1	9,7
	AR(1) Constant: Variance Skewness	5833715,4	45,9	2397754,6	13,3
	AR(1) GARCH(1,1) Constant Skewness	1679017,6	13,6	164767,8	1,0
	AR(1) T-GARCH(1,1) Constant Skewness	1001726,8	8,3	22041,1	0,3
	AR(1) T-GARCH(1,1) Dynamic Skewness	517597,2	4,8	2746,4	0,1
STUDENT'S T	Constant: Mean, Variance, D.o.F.	5070840,7	39,6	967386,1	4,6
	AR(1) Constant: Variance, D.o.F.	7267202,7	62,3	1564577,0	7,3
	AR(1) GARCH(1,1) Constant D.o.F.	2906143,9	23,1	230168,3	1,4
	AR(1) T-GARCH(1,1) Constant D.o.F.	1925515,0	15,3	47020,3	0,5
	AR(1) T-GARCH(1,1) Dynamic D.o.F.	1625858,0	12,8	70050,5	0,6
SKEW-T	Constant: Mean, Variance, Skewness, D.o.F.	4991445,4	38,8	914465,5	4,4
	AR(1) Constant: Variance, Skewness, D.o.F.	7072329,9	59,7	1516976,0	7,0
	AR(1) GARCH(1,1) Const: Skewness D.o.F.	2758012,6	21,7	192470,9	1,2
	AR(1) T-GARCH(1,1) Const: Skew D.o.F.	1798403,9	14,0	33183,9	0,4
	AR(1)T-GARCH(1,1) Dyn Skew,Const D.o.F.	1100270,5	8,4	85,6	0,0
	AR(1) T-GARCH(1,1) Dyn Skew, Dyn D.o.F.	960117,5	7,4	300,8	0,0