Testing the Gaussian copula hypothesis for financial assets dependences

Yannick Malevergne, Didier Sornette

Testing the Gaussian Copula Hypothesis for Financial Assets Dependencies*

Y. Malevergne[1,2] and D. Sornette[1,3]

[1] Laboratoire de Physique de la Matière Condensée CNRS UMR 6622
Université de Nice-Sophia Antipolis, 06108 Nice Cedex 2, France
[2] Institut de Science Financière et d'Assurances - Université Lyon I
43, Bd du 11 Novembre 1918, 69622 Villeurbanne Cedex
[3] Institute of Geophysics and Planetary Physics and Department of Earth and Space Science
University of California, Los Angeles, California 90095, USA

JEL Classification: C12, C15, F31, G19

Keywords: Copulas, Dependence Modelisation, Risk Management, Tail Dependence.

*We acknowledge helpful discussions and exchanges with J. Andersen, P. Embrechts, J.P. Laurent, F. Lindskog, V. Pisarenko and R. Valkanov. This work was partially supported by the James S. Mc Donnell Foundation 21st century scientist award/studying complex system.

Abstract

Using one of the key property of copulas that they remain invariant under an arbitrary monotonous change of variable, we investigate the null hypothesis that the dependence between financial assets can be modeled by the Gaussian copula. We find that most pairs of currencies and pairs of major stocks are compatible with the Gaussian copula hypothesis, while this hypothesis can be rejected for the dependence between pairs of commodities (metals). Notwithstanding the apparent qualification of the Gaussian copula hypothesis for most of the currencies and the stocks, a non-Gaussian copula, such as the Student's copula, cannot be rejected if it has sufficiently many "degrees of freedom". As a consequence, it may be very dangerous to embrace blindly the Gaussian copula hypothesis, especially when the correlation coefficient between the pair of asset is too high, so that the tail dependence neglected by the Gaussian copula can became large, leading to ignore extreme events which may occur in unison.

1 Introduction

The determination of the dependence between assets underlies many financial activities, such as risk assessment and portfolio management, as well as option pricing and hedging. Following (Markovitz 1959), the covariance and correlation matrices have, for a long time, been considered as the main tools for quantifying the dependence between assets. But the dimension of risk captured by the correlation matrices is only satisfying for elliptic distributions and for moderate risk amplitudes (Sornette *et al.* 2000b). In all other cases, this measure of risk is severely incomplete and can lead to a very strong underestimation of the real incurred risks (Embrechts *et al.* 1999).

Although the unidimensional (marginal) distributions of asset returns are reasonably constrained by empirical data and their tails are more or less satisfactorily described by a power law with tail index ranging between 2 and 4 (De Vries 1994, Lux 1996, Pagan 1996, Guillaume *et al.* 1997, Gopikrishnan *et al.* 1998, McNiel and Frey 2000), by stretched exponentials (Laherrère and Sornette 1998, Gouriéroux and Jasiak 1999, Sornette *et al.* 2000a, Sornette *et al.* 2000b) or by log-Weibull distributions (Malevergne *et al.* 2003), no equivalent results have been obtained for *multivariate* distributions of asset returns. Indeed, a brute force determination of multivariate distributions is unreliable due to the limited data set (the curse of dimensionality), while the sole knowledge of marginals (one-point statistics) of each asset is not sufficient to obtain information on the multivariate distribution of these assets which involves all the n-points statistics.

Some progress may be expected from the concept of copulas, recently proposed to be useful for financial applications (Embrechts *et al.* 2001, Frees and Valdez 1998, Haas 1999, Klugman and Parsa 1999). This concept has the desirable property of decoupling the study of the marginal distribution of each asset from the study of their collective behavior or dependence. Indeed, the dependence between assets is entirely embedded in the copula, so that a copula allows for a simple description of the dependence structure between assets independently of the marginals. For instance, assets can have power law marginals and a Gaussian copula or alternatively Gaussian marginals and a non-Gaussian copula, and any possible combination thereof. Therefore, the determination of the multivariate distribution of assets can be performed in two steps : (i) an independent determination of the marginal distributions using standard techniques for distributions of a single variable ; (ii) a study of the nature of the copula characterizing completely the dependence between the assets. This exact separation between the marginal distributions and the dependence is potentially very useful for risk management or option pricing and sensitivity analysis since it allows for testing several scenarios with different kind of dependences between assets while the marginals can be set to their well-calibrated empirical estimates. Such an approach has been used by Embrechts *et al.* (2001) to provide various bounds for the Value-at-Risk of a portfolio made of depend risks, and by Rosenberg (1999) or Cherubini and Luciano (2000) to price and to analyze the pricing sensitivity of binary digital options or options on the minimum of a basket of assets.

A fundamental limitation of the copula approach is that there is in principle an infinite number of possible copulas (Genest and MacKay 1986, Genest 1987, Genest and Rivest 1993, Joe 1993, Nelsen 1998) and, up to now, no general empirical study has determined the classes of copulas that are acceptable for financial problems. In general, the choice of a given copula is guided both by the empirical evidences and the technical constraints, i.e., the number of parameters necessary to describe the copula, the possibility to obtain efficient estimators of these parameters and also the possibility offered by the chosen parameterization to allow for tractable analytical calculation. It is indeed sometimes more advantageous to prefer a simplest copula to one that fit better the data, provided that we can clearly quantify the effects of this substitution.

In this vein, the first goal of the present article is to wonder whether the Gaussian copula provides a sufficiently good approximation of the unknown true copula, on a statistical basis. Such an investigation is really anchored at the heart of many financial problems since the Gaussian copula sustains almost all current financial theories. Obviously, the Gaussian copula is rooted into the traditional theories relying on the multivariate Gaussian description, but it is also widespread in the most recent financial applications such as the modeling the dependent defaults[1] as exemplified by the model of CreditMetrics or KMV, for instance or such as the pricing of credit derivatives (). Thus, there is a real need for a test of the ability of the Gaussian copula to model financial dependencies. Our second goal is to draw the consequences of the parameterization involved in the Gaussian copula in term of potential over/underestimation of the risks, in particular for large and extreme events.

The paper is organized as follows.

In section 2, we first recall some important general definitions and theorems about copulas that will be useful in the sequel. We then introduce the concept of tail dependence that will allow us to quantify the probability that two extreme events might occur simultaneously. We define and describe the two copulas that will be at the core of our study : the Gaussian copula and the Student's copula and compare their properties particularly in the tails.

In section 3, we present our statistical testing procedure which is applied to pairs of financial time series. First of all, we determine a test statistics which leads us to compare the empirical distribution of the data with a χ^2-distribution using a bootstrap method. We also test the sensitivity of our procedure by applying it to synthetic multivariate Student's time series. This allows us to determine the minimum statistical test value needed to be able to distinguish between a Gaussian and a Student's copula, as a function of the number of degrees of freedom and of the correlation strength.

Section 4 presents the empirical results obtained for the following assets which are combined pairwise in the test statistics:

- 6 currencies,
- 6 metals traded on the London Metal Exchange,
- 22 stocks chosen among the largest companies quoted on the New York Stocks Exchange.

We show that the Gaussian copula hypothesis is very reasonable for most stocks and currencies, while it is hardly compatible with the description of multivariate behavior for metals.

Section 5 summarizes our results and concludes.

2 Generalities about copulas

2.1 Definitions and important results about copulas

This section does not pretend to provide a rigorous mathematical exposition of the concept of copula. We only recall a few basic definitions and theorems that will be useful in the following (for more information about the concept of copula, see for instance Nelsen (1998).

[1]Following the recommendations of the Basle committee on supervision banking (2001), the Gaussian copula must be chosen to model the dependence between defaults.

We first give the definition of a copula of n random variables.

DEFINITION 1 (COPULA)
A function $C : [0,1]^n \longrightarrow [0,1]$ is a n-copula if it enjoys the following properties :

- $\forall u \in [0,1]$, $C(1, \cdots, 1, u, 1 \cdots, 1) = u$,
- $\forall u_i \in [0,1]$, $C(u_1, \cdots, u_n) = 0$ if at least one of the u_i equals zero ,
- C is grounded and n-increasing, i.e., the C-volume of every box whose vertices lie in $[0,1]^n$ is positive.

It is clear from this definition that a copula is nothing but a multivariate distribution with support in [0,1]^n and with uniform marginals. The fact that such copulas can be very useful for representing multivariate distributions with arbitrary marginals is seen from the following result.

THEOREM 1 (SKLAR'S THEOREM)
Given an n-dimensional distribution function F with *continuous* marginal (cumulative) distributions $F_1, \cdots, F_n$, there exists a *unique* n-copula $C : [0,1]^n \longrightarrow [0,1]$ such that :

$$F(x_1, \cdots, x_n) = C(F_1(x_1), \cdots, F_n(x_n)) \ . \tag{1}$$

This theorem provides both a parameterization of multivariate distributions and a construction scheme for copulas. Indeed, given a multivariate distribution F with marginals $F_1, \cdots, F_n$, the function

$$C(u_1, \cdots, u_n) = F\left(F_1^{-1}(u_1), \cdots, F_n^{-1}(u_n)\right) \tag{2}$$

is automatically a n-copula[2]. This copula is the copula of the multivariate distribution F. We will use this method in the sequel to derive the expressions of standard copulas such as the Gaussian copula or the Student's copula.

A very powerful property of copulas is their invariance under arbitrary strictly increasing mapping of the random variables :

THEOREM 2 (INVARIANCE THEOREM)
Consider n *continuous* random variables $X_1, \cdots, X_n$ with copula C. Then, if $g_1(X_1), \cdots, g_n(X_n)$ are strictly increasing on the ranges of $X_1, \cdots, X_n$, the random variables $Y_1 = g_1(X_1), \cdots, Y_n = g_n(X_n)$ have exactly the same copula C.

It is this result that shows us that the full dependence between the n random variables is completely captured by the copula, independently of the shape of the marginal distributions. This result is at the basis of our statistical study presented in section 3.

[2]The quantile function F_i^{-1} of the distribution F_i can be defined by:

$$F_i^{-1}(u) = \inf\{x \mid F_i(x) \geq u\}, \quad \forall u \in (0,1).$$

When the distribution function F_i is strictly increasing, F_i^{-1} denotes the usual inverse of F_i. In fact, any quantile function can be chosen. But, for non-continuous margins, the copula (2) will depend upon the precise quantile function which will be selected.

2.2 Dependence between random variables

The dependence between two time series is usually described by their correlation coefficient. This measure is fully satisfactory only for elliptic distributions (Embrechts *et al.* 1999), which are functions of a quadratic form of the random variables, when one is interested in moderately size events. However, an important issue for risk management concerns the determination of the dependence of the distributions in the tails. Practically, the question is whether it is more probable that large or extreme events occur simultaneously or on the contrary more or less independently. This is refered to as the presence or absence of "tail dependence".

The tail dependence is also an interesting concept in studying the *contagion* of crises between markets or countries. These questions have recently been addressed by (Ang and Cheng 2001, Longin and Solnik 2001, Starica 1999) among several others. Large negative moves in a country or market are often found to imply large negative moves in others.

Technically, we need to determine the probability that a random variable X is large, knowing that the random variable Y is large.

DEFINITION 2 (TAIL DEPENDENCE 1)
Let X and Y be random variables with *continuous* marginals F_X and F_Y. The (upper) tail dependence coefficient of X and Y is, if it exists,

$$\lim_{u\to 1} \Pr\{X > F_X^{-1}(u) | Y > F_Y^{-1}(u)\} = \lambda \in [0,1] \, . \tag{3}$$

In words, given that Y is very large (which occurs with probability $1-u$), the probability that X is very large at the same probability level u defines asymptotically the tail dependence coefficient λ.

It turns out that this tail dependence is a pure copula property which is independent of the marginals. Let C be the copula of the (*assumed continuous*) variables X and Y, then

THEOREM 3
if the bivariate copula C is such that

$$\lim_{u\to 1} \frac{\bar{C}(u,u)}{1-u} = \lambda \tag{4}$$

exists (where $\bar{C}(u,u) = 1 - 2u + C(u,u)$), then C has an upper tail dependence coefficient λ.

If $\lambda > 0$, the copula presents tail dependence and large events tend to occur simultaneously, with the probability λ. On the contrary, when $\lambda = 0$, the copula has no tail dependence in this sense and large events appear to occur essentially independently. There is however a subtlety in this definition of tail dependence. To make it clear, first consider the case where for large X and Y the distribution function $F(x,y)$ factorizes such that

$$\lim_{x,y\to\infty} \frac{F(x,y)}{F_X(x)F_Y(y)} = 1 \, . \tag{5}$$

This means that, for X and Y sufficiently large, these two variables can be considered as independent. It is then easy to show that

$$\begin{aligned} \lim_{u\to 1} \Pr\{X > F_X^{-1}(u) | Y > F_Y^{-1}(u)\} &= \lim_{u\to 1} 1 - F_X(F_X^{-1}(u)) & (6) \\ &= \lim_{u\to 1} 1 - u = 0, & (7) \end{aligned}$$

so that independent variables really have no tail dependence, as one can expect.

Unfortunately, the converse does not holds : a value $\lambda = 0$ does not automatically imply true independence, namely that $F(x, y)$ satisfies equation (5). Indeed, the tail independence criterion $\lambda = 0$ may still be associated with an absence of factorization of the multivariate distribution for large X and Y. In a weaker sense, there may still be a dependence in the tail even when $\lambda = 0$. Such behavior is for instance exhibited by the Gaussian copula, which has zero tail dependence according to the definition 2 but nevertheless does not have a factorizable multivariate distribution, since the non-diagonal term of the quadratic form in the exponential function does not become negligible in general as X and Y go to infinity. To summarize, the *tail independence*, according to definition 2, is not equivalent to the *independence in the tail* as defined in equation (5).

After this brief review of the main concepts underlying copulas, we now present two special families of copulas : the Gaussian copula and the Student's copula.

2.3 The Gaussian copula

The Gaussian copula is the copula derived from the multivariate Gaussian distribution. Let Φ denote the standard Normal (cumulative) distribution and $\Phi_{\rho,n}$ the n-dimensional Gaussian distribution with correlation matrix ρ. Then, the Gaussian n-copula with correlation matrix ρ is

$$C_\rho(u_1, \cdots, u_n) = \Phi_{\rho,n}\left(\Phi^{-1}(u_1), \cdots, \Phi^{-1}(u_n)\right) , \tag{8}$$

whose density

$$c_\rho(u_1, \cdots, u_n) = \frac{\partial C_\rho(u_1, \cdots, u_n)}{\partial u_1 \cdots \partial u_n} \tag{9}$$

reads

$$c_\rho(u_1, \cdots, u_n) = \frac{1}{\sqrt{\det \rho}} \exp\left(-\frac{1}{2} y^t_{(u)} (\rho^{-1} - \mathrm{Id}) y_{(u)}\right) \tag{10}$$

with $y_k(u) = \Phi^{-1}(u_k)$. Note that theorem 1 and equation (2) ensure that $C_\rho(u_1, \cdots, u_n)$ in equation (8) is a copula.

As we said before, the Gaussian copula does not have a tail dependence :

$$\lim_{u \to 1} \frac{\bar{C}_\rho(u, u)}{1 - u} = 0, \ \ \forall \rho \in (-1, 1). \tag{11}$$

This result is derived for example in (Embrechts *et al.* 2001). But this does not mean that the Gaussian copula goes to the independent (or product) copula $\Pi(u_1, u_2) = u_1 \cdot u_2$ when (u_1, u_2) goes to one. Indeed, consider a distribution $F(x, y)$ with Gaussian copula :

$$F(x, y) = C_\rho(F_X(x), F_Y(y)). \tag{12}$$

Its density is

$$f(x, y) = c_\rho(F_X(x), F_Y(y)) \cdot f_X(x) \cdot f_Y(y), \tag{13}$$

where f_X and f_Y are the densities of X and Y. Thus,

$$\lim_{(x,y)\to\infty} \frac{f(x, y)}{f_X(x) \cdot f_Y(y)} = \lim_{(x,y)\to\infty} c_\rho(F_X(x), F_Y(y)), \tag{14}$$

which should equal 1 if the variables X and Y were independent in the tail. Reasoning in the quantile space, we set $x = F_X^{-1}(u)$ and $y = F_Y^{-1}(u)$, which yield

$$\lim_{(x,y)\to\infty} \frac{f(x,y)}{f_X(x)\cdot f_Y(y)} = \lim_{u\to 1} c_\rho(u,u). \tag{15}$$

Using equation (10), it is now obvious to show that $c_\rho(u,u)$ goes to one when u goes to one, if and only if $\rho = 0$ which is equivalent to $C_{\rho=0}(u_1,u_2) = \Pi(u_1,u_2)$ for every (u_1,u_2). When $\rho > 0$, $c_\rho(u,u)$ goes to infinity, while for ρ negative, $c_\rho(u,u)$ goes to zero as $u \to 1$. Thus, the dependence structure described by the Gaussian copula is very different from the dependence structure of the independent copula, except for $\rho = 0$.

The Gaussian copula is completely determined by the knowledge of the correlation matrix ρ. The parameters involved in the description of the Gaussian copula are very simple to estimate, as we shall see in the following.

In our tests presented below, we focus on pairs of assets, i.e., on Gaussian copulas involving only two random variables. Obviously, for risk management purposes, baskets or portfolios of $n > 2$ assets must be considered. Our restriction is not crucial since the testing procedure exposed in section 3 can be applied to any number of assets and it is only for the simplicity of the exposition that we will present the case where only two assets are considered. Moreover, testing the Gaussian copula hypothesis for two random variables gives useful information for a larger number of dependent variables constituting a large basket or portfolio. Indeed, let us assume that each pair $(a,b),(b,c)$ and (c,a) have a Gaussian copula, and in addition that the copula of the triplet (a,b,c) is elliptical, which is a reasonable assumption. Then, the triplet (a,b,c) has also a Gaussian copula. This result generalizes to an arbitrary number of random variables[3].

2.4 The Student's copula

The Student's copula is derived from the Student's multivariate distribution. Given a multivariate Student's distribution $T_{\rho,\nu}$ with ν degrees of freedom and a shape[4] matrix ρ

$$T_{\rho,\nu}(\mathbf{x}) = \frac{1}{\sqrt{\det\rho}} \frac{\Gamma\left(\frac{\nu+n}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)(\pi\nu)^{N/2}} \int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_N} \frac{d\mathbf{x}}{\left(1+\frac{x^t\rho^{-1}x}{\nu}\right)^{\frac{\nu+n}{2}}}, \tag{16}$$

the corresponding Student's copula reads :

$$C_{\rho,\nu}(u_1,\cdots,u_n) = T_{\rho,\nu}\left(t_\nu^{-1}(u_1),\cdots,t_\nu^{-1}(u_n)\right), \tag{17}$$

where t_ν is the univariate Student's distribution with ν degrees of freedom. The density of the Student's copula is thus

$$c_{\rho,\nu}(u_1,\cdots,u_n) = \frac{1}{\sqrt{\det\rho}} \frac{\Gamma\left(\frac{\nu+n}{2}\right)\left[\Gamma\left(\frac{\nu}{2}\right)\right]^{n-1}}{\left[\Gamma\left(\frac{\nu+1}{2}\right)\right]^n} \frac{\prod_{k=1}^n\left(1+\frac{y_k^2}{\nu}\right)^{\frac{\nu+1}{2}}}{\left(1+\frac{y^t\rho^{-1}y}{\nu}\right)^{\frac{\nu+n}{2}}}, \tag{18}$$

[3] An elliptical distribution, and then an elliptical copula, is fully determined by the knowledge of its mean, shape (or covariance) matrix and the generator of its type. Once the distributions of every pairs of random variables (X_i,X_j), $i,j \in \{1,\cdots,N\}$ are known, the type of the generator is fixed and the mean and the shape matrix of the joined distribution of $(X_1,X_2,\cdots,X_N)$ can be reconstructed.

[4] The shape matrix ρ is equal to the correlation matrix when ν is larger than two, namely when the second moments of the variables X_i's exist.

where $y_k = t_\nu^{-1}(u_k)$.

Since the Student's distribution tends to the normal distribution when ν goes to infinity, the Student's copula tends to the Gaussian copula as $\nu \to +\infty$. In contrast to the Gaussian copula, the Student's copula for ν finite presents a tail dependence given by :

$$\lambda_\nu(\rho) = \lim_{u \to 1} \frac{\bar{C}_{\rho,\nu}(u,u)}{1-u} = 2\bar{t}_{\nu+1}\left(\frac{\sqrt{\nu+1}\sqrt{1-\rho}}{\sqrt{1+\rho}}\right) , \tag{19}$$

where $\bar{t}_{\nu+1}$ is the complementary cumulative univariate Student's distribution with $\nu+1$ degrees of freedom (see (Embrechts *et al.* 2001) for the proof). Figure 1 shows the upper tail dependence coefficient as a function of the correlation coefficient ρ for different values of the number ν of degrees of freedom. As expected from the fact that the Student's copula becomes identical to the Gaussian copula for $\nu \to +\infty$ for all $\rho \neq 1$, $\lambda_\nu(\rho)$ exhibits a regular decay to zero as ν increases. Moreover, for ν sufficiently large, the tail dependence is significantly different from 0 only when the correlation coefficient is sufficiently close to 1. This suggests that, for moderate values of the correlation coefficient, a Student's copula with a large number of degrees of freedom may be difficult to distinguish from the Gaussian copula from a statistical point of view. This statement will be made quantitative in the following.

Figure 2 presents the same information in a different way by showing the maximum value of the correlation coefficient ρ as a function of ν, below which the tail dependence $\lambda_\nu(\rho)$ of a Student's copula is smaller than a given small value, here taken equal to $1\%, 2.5\%, 5\%$ and 10%. The choice $\lambda_\nu(\rho) = 5\%$ for instance corresponds to 1 event in 20 for which the pair of variables are asymptotically coupled. At the 95% probability level, values of $\lambda_\nu(\rho) \leq 5\%$ are undistinguishable from 0, which means that the Student's copula can be approximated by a Gaussian copula.

The description of a Student's copula relies on two parameters : the correlation matrix ρ, as in the Gaussian case, and in addition the number of degrees of freedom ν. The estimation of the parameter ν is rather difficult and this has an important impact on the estimated value of the correlation matrix. As a consequence, the Student's copula is more difficult to calibrate and use than the Gaussian copula.

3 Testing the Gaussian copula hypothesis

In view of the central role that the Gaussian paradigm has played and still plays in particular in finance, it is natural to start with the simplest choice of dependence between different random variables, namely the Gaussian copula. It is also a natural first step as the Gaussian copula imposes itself in an approach which consists in (1) performing a nonlinear transformation on the random variables into Normal random variables (for the marginals) which is always possible and (2) invoking a maximum entropy principle (which amounts to add the least additional information in the Shannon sense) to construct the multivariable distribution of these Gaussianized random variables (Sornette *et al.* 2000b, Sornette *et al.* 2000a, Andersen and Sornette 2001).

In the sequel, we will denote by H_0 the null hypothesis according to which the dependence between two (or more) random variables X and Y can be described by the Gaussian copula.

3.1 Test Statistics

We now derive the test statistics which will allow us to reject or not our null hypothesis H_0 and state the following proposition:

PROPOSITION 1
Assuming that the N-dimensional random vector $\mathbf{x} = (x_1, \cdots, x_N)$ with distribution function F and marginals F_i, satisfies the null hypothesis H_0, then, the variable

$$z^2 = \sum_{j,i=1}^{N} \Phi^{-1}(F_i(x_i))\ (\rho^{-1})_{ij}\ \Phi^{-1}(F_j(x_j)), \tag{20}$$

where the matrix ρ is

$$\rho_{ij} = \mathrm{Cov}[\Phi^{-1}(F_i(x_i)), \Phi^{-1}(F_j(x_j))], \tag{21}$$

follows a χ^2-distribution with N degrees of freedom.

To prove the proposition above, first consider an N-dimensional random vector $\mathbf{x} = (x_1, \cdots, x_N)$. Let us denote by F its distribution function and by F_i the marginal distribution of each x_i. Let us now assume that the distribution function F satisfies H_0, so that F has a Gaussian copula with correlation matrix ρ while the F_i's can be any distribution function. According to theorem 1, the distribution F can be represented as :

$$F(x_1, \cdots, x_N) = \Phi_{\rho,N}(\Phi^{-1}(F_1(x_1)), \cdots, \Phi^{-1}(F_N(x_N)))\ . \tag{22}$$

Let us now transform the x_i's into Normal random variables y_i's :

$$y_i = \Phi^{-1}(F_i(x_i))\ . \tag{23}$$

Since the mapping $\Phi^{-1}(F_i(\cdot))$ is obviously increasing, theorem 2 allows us to conclude that the copula of the variables y_i's is identical to the copula of the variables x_i's. Therefore, the variables y_i's have Normal marginal distributions and a Gaussian copula with correlation matrix ρ. Thus, by definition, the multivariate distribution of the y_i's is the multivariate Gaussian distribution with correlation matrix ρ :

$$\begin{aligned} G(\mathbf{y}) &= \Phi_{\rho,N}(\Phi^{-1}(F_1(x_1)), \cdots, \Phi^{-1}(F_N(x_N))) && (24)\\ &= \Phi_{\rho,N}(y_1, \cdots, y_N), && (25) \end{aligned}$$

and $\mathbf{y}$ is a Gaussian random vector. From equations (24-25), we obviously have

$$\rho_{ij} = \mathrm{Cov}[\Phi^{-1}(F_i(x_i)), \Phi^{-1}(F_j(x_j))]. \tag{26}$$

Consider now the random variable

$$z^2 = \mathbf{y^t}\rho^{-\mathbf{1}}\mathbf{y} = \sum_{i,j=1}^{N} y_i\ (\rho^{-1})_{ij}\ y_j\ , \tag{27}$$

where $\cdot^t$ denotes the transpose operator. It is well-known that the variable z^2 follows a χ^2-distribution with N degrees of freedom. Indeed, since $\mathbf{y}$ is a Gaussian random vector with covariance matrix[5] ρ, it follows that the components of the vector

$$\tilde{\mathbf{y}} = \mathbf{A}\mathbf{y}, \tag{28}$$

[5]Up to now, the matrix ρ was named *correlation matrix*. But in fact, since the variables y_i's have unit variance, their correlation matrix is also their *covariance matrix*.

are *independent* Normal random variables. Here, $\mathbf{A}$ denotes the square root of the matrix ρ^{-1}, obtained by the Cholevsky decomposition, so that $\mathbf{A}^t\mathbf{A} = \rho^{-1}$. Thus, the sum $\tilde{\mathbf{y}}^\mathbf{t}\tilde{\mathbf{y}} = z^2$ is the sum of the squares of N independent Normal random variables, which follows a χ^2-distribution with N degrees of freedom.

3.2 Testing procedure

The testing procedure used in the sequel is now described. We consider two[6] financial series ($N = 2$) of size T: $\{x_1(1), \cdots, x_1(t), \cdots, x_1(T)\}$ and $\{x_2(1), \cdots, x_2(t), \cdots, x_2(T)\}$. We assume that the vectors $\mathbf{x}(t) = (x_1(t), x_2(t))$, $t \in \{1, \cdots, T\}$ are independent and identically distributed with distribution F, which implies that the variables $x_1(t)$ (respectively $x_2(t)$), $t \in \{1, \cdots, T\}$, are also independent and identically distributed, with distributions F_1 (respectively F_2)[7].

The cumulative distribution $\hat{F}_i$ of each variable x_i, which is estimated empirically, is given by

$$\hat{F}_i(x_i) = \frac{1}{T}\sum_{k=1}^{T} \mathbf{1}_{\{x_i(k) \leq x_i\}}, \tag{29}$$

where $\mathbf{1}_{\{.\}}$ is the indicator function, which equals one if its argument is true and zero otherwise. We use these estimated cumulative distributions to obtain the Gaussian variables $\hat{y}_i$ as :

$$\hat{y}_i(k) = \Phi^{-1}\left(\hat{F}_i(x_i(k))\right) \quad k \in \{1, \cdots, T\}\ . \tag{30}$$

The sample covariance matrix $\hat{\rho}$ is estimated by the expression :

$$\hat{\rho} = \frac{1}{T}\sum_{i=1}^{T} \hat{\mathbf{y}}(i) \cdot \hat{\mathbf{y}}(i)^t \tag{31}$$

which allows us to calculate the variable

$$\hat{z}^2(k) = \sum_{i,j=1}^{2} \hat{y}_i(k)\ (\rho^{-1})_{ij}\ \hat{y}_j(k)\ , \tag{32}$$

as defined in (27) for $k \in \{1, \cdots, T\}$, which should be distributed according to a χ^2-distribution if the Gaussian copula hypothesis is correct.

The usual way for comparing an empirical with a theoretical distribution is to measure the distance between these two distributions and to perform the Kolmogorov test or the Anderson-Darling (Anderson and Darling 1952) test (for a better accuracy in the tails of the distribution). The

[6] As explained in section 2.3, the case $N = 2$ is not restrictive at all, even if it could, *a priori*, appear of limited interest. Indeed, for portfolio analysis and risk management purposes, larger basket of assets should be considered. However, the testing procedure exposed here can be applied to any number of assets, and it is only for the sake simplicity of the exposition that we have restricted our investigation to the bivariate case.

[7] The assumption of independently distributed data is not very realistic. Indeed, it is well-know that daily returns are uncorrelated but that their volatility exhibits long range dependence. One can then wonder why we have not filtered the data by an ARCH or GARCH process (as in Patton (2001)) and then apply our testing procedure to the residuals. The main limitation of this approach is the following. The filtering of the data does not let the dependence structure, i.e., the copula, unchanged. Thus, the copula of the residuals is not the same as the copula of the raw returns. Moreover, the copula of the residual changes with the chosen filter. Residuals are not the same when one filters the data with an ARCH, a GARCH or a Multifractal Random Walk (Muzy *et al.* 2000, Muzy *et al.* 2001). Therefore, our standpoint has been to perform a model-free analysis, and thus not to filter the data. Obviously, the price to pay for such a model-free approach is a weakening of the power of the statistical test due to the presence of (temporal) dependence between data.

Kolmogorov distance is the maximum local distance along the quantile which most often occur in the bulk of the distribution, while the Anderson-Darling distance puts the emphasis on the tails of the two distributions by a suitable normalization. We propose to complement these two distances by two additional measures which are defined as averages of the Kolmogorov distance and of the Anderson-Darling distance respectively:

$$\text{Kolmogorov}: \quad d_1 = \max_z |F_{z^2}(z^2) - F_{\chi^2}(z^2)| \tag{33}$$

$$\text{average Kolmogorov}: \quad d_2 = \int |F_{z^2}(z^2) - F_{\chi^2}(z^2)| \, dF_{\chi^2}(z^2) \tag{34}$$

$$\text{Anderson} - \text{Darling}: \quad d_3 = \max_z \frac{|F_{z^2}(z^2) - F_{\chi^2}(z^2)|}{\sqrt{F_{\chi^2}(z^2)[1 - F_{\chi^2}(z^2)]}} \tag{35}$$

$$\text{average Anderson} - \text{Darling}: \quad d_4 = \int \frac{|F_{z^2}(z^2) - F_{\chi^2}(z^2)|}{\sqrt{F_{\chi^2}(z^2)[1 - F_{\chi^2}(z^2)]}} \, dF_{\chi^2}(z^2) \tag{36}$$

The Kolmogorov distance d_1 and its average d_2 are more sensitive to the deviations occurring in the bulk of the distributions. In contrast, the Anderson-Darling distance d_3 and its average d_4 are more accurate in the tails of the distributions. We present our statistical tests for these four distances in order to be as complete as possible with respect to the different sensitivity of the tests.

The distances d_2 and d_4 are not of common use in statistics, so let us justify our choice. One usually uses distances similar to d_2 and d_4 but which differ by the square instead of the modulus of $F_{z^2}(z^2) - F_{\chi^2}(z^2)$ and lead respectively to the ω-test and the Ω-test, whose statistics are theoretically known. The main advantage of the distances d_2 and d_4 with respect to the more usual distances ω and Ω is that they are simply equal to the average of d_1 and d_3. This averaging is very interesting and provides important information. Indeed, the distances d_1 and d_3 are mainly controlled by the point that maximizes the argument within the $\max(\cdot)$ function. They are thus sensitive to the presence of an outlier. By averaging, d_2 and d_4 become less sensitive to outliers, since the weight of such points is only of order $1/T$ (where T is the size of the sample) while it equals one for d_1 and d_3. Of course, the distances ω and Ω also perform a smoothing since they are averaged quantities too. But they are the average of the square of d_1 and d_3 which can lead to an undesired overweighing of the largest events. Of course, such an overweighing of large events can be interesting when one want to particularly focus on tail events. In fact, a trade-off between the sensitivity to (desired) tail events and to (undesired) outliers must be found. That is why we have prefered d_2 and d_4, which have seemed to us more convenient, in this respect, than the omega's distances. Moreover, the square function is chosen as a convenient analytical form that allows one to derive explicitely the theoretical asymptotic statistics for the ω and Ω-tests. In contrast, using the modulus of $F_{z^2}(z^2) - F_{\chi^2}(z^2)$ instead of its square in the expression of d_2 and d_4, no theoretical test statistics can be derived analytically. In sum, the sole advantage of the standard distances ω and Ω with respect to the distances d_2 and d_4 introduced here is the theoretical knowledge of their distributions. However, this advantage disappears in our present case in which the covariance matrix is not known *a priori* and needs to be estimated from the empirical data: indeed, the exact knowledge of all the parameters is necessary in the derivation of the theoretical statistics of the ω and Ω-tests (as well as the Kolmogorov test). Therefore, we cannot directly use the results of these standard statistical tests. As a remedy, we propose a bootstrap method (Efron and Tibshirani 1986), whose accuracy is proved by (Chen and Lo 1997) to be at least as good as that given by asymptotic methods used to derive the theoretical distributions. For the present work, we have determined that the generation of 10,000 synthetic time series was sufficient to obtain a good approximation of the distribution of distances described above. Since a bootstrap method is needed to determine the tests

statistics in every case, it is convenient to choose functional forms different from the usual ones in the ω and Ω-tests as they provide an improvement with respect to statistical reliability, as obtained with the d_2 and d_4 distances introduced here.

To summarize, our test procedure is as follows.

1. Given the original time series $\mathbf{x}(t)$, $t \in \{1, \cdots, T\}$, we generate the Gaussian variables $\hat{\mathbf{y}}(t)$, $t \in \{1, \cdots, T\}$.
2. We then estimate the covariance matrix $\hat{\rho}$ of the Gaussian variables $\hat{\mathbf{y}}$, which allows us to compute the variables $\hat{z}^2$ and then measure the distance of its estimated distribution to the χ^2-distribution.
3. Given this covariance matrix $\hat{\rho}$, we generate numerically a time series of T Gaussian random vectors with the same covariance matrix $\hat{\rho}$.
4. For the time series of Gaussian vectors synthetically generated with covariance matrix $\hat{\rho}$, we estimate its sample covariance matrix $\tilde{\rho}$.
5. To each of the T vectors of the synthetic Gaussian time series, we associate the corresponding realization of the random variable z^2, called $\tilde{z}^2(t)$.
6. We can then construct the empirical distribution for the variable $\tilde{z}^2$ and measure the distance between this distribution and the χ^2-distribution.
7. Repeating 10,000 times the steps 3 to 6, we obtain an accurate estimate of the cumulative distribution of distances between the distribution of the synthetic Gaussian variables and the theoretical χ^2-distribution. This cumulative distribution represents the test statistic, which will allow us to reject or not the null hypothesis H_0 at a given significance level.
8. The significance of the distance we got at step 2 for the true variables - i.e., the probability to observe, at random and under H_0, a distance larger than the empirically estimated distance - is finally obtained by a simple reading on the complementary cumulative distribution estimated at step 7.

3.3 Sensitivity of the method

Before presenting the statistical tests, it is important to investigate the sensitivity of our testing procedure. More precisely, can we distinguish for instance between a Gaussian copula and a Student's copula with a large number of degrees of freedom, for a given value of the correlation coefficient? Formally, denoting by H_ν the hypothesis according to which the true copula of the data is the Student's copula with ν degrees of freedom, we want to determine the minimum significance level allowing us to distinguish between H_0 and H_ν.

3.3.1 Importance of the distinction between Gaussian and Student's copulas

This question has important practical implications because, as discussed in section 2.4, the Student's copula presents a significant tail dependence while the Gaussian copula has no asymptotic tail dependence. Therefore, if our tests are unable to distinguish between a Student's and a Gaussian copula, we may be led to choose the later for the sake of simplicity and parsimony and, as a

consequence, we may underestimate severely the dependence between extreme events if the correct description turns out to be the Student's copula. This may have catastrophic consequences in risk assessment and portfolio management.

Figure 1 provides a quantification of the dangers incurred by mistaking a Student's copula for a Gaussian one. Consider the case of a Student's copula with $\nu = 20$ degrees of freedom with a correlation coefficient ρ lower than $0.3 \sim 0.4$; its tail dependence $\lambda_\nu(\rho)$ turns out to be less than 0.7%, i.e., the probability that one variable becomes extreme knowing that the other one is extreme is less than 0.7%. In this case. the Gaussian copula with zero probability of simultaneous extreme events is not a bad approximation of the Student's copula. In contrast, let us take a correlation ρ larger than $0.7 - 0.8$ for which the tail dependence becomes larger than 10%, corresponding to a non-negligible probability of simultaneous extreme events. The effect of tail dependence becomes of course much stronger as the number ν of degrees of freedom decreases.

These examples stress the importance of knowing whether our testing procedure allows us to distinguish between a Student's copula with $\nu = 20$ (or less) degrees of freedom and a given correlation coefficient $\rho = 0.5$, for instance, and a Gaussian copula with an appropriate correlation coefficient ρ'.

3.3.2 Statistical test on the distinction between Gaussian and Student's copulas

To address this question, we have generated 1,000 pairs of time series of size $T = 1250$, each pair of random variables following a Student's bivariate distribution with ν degrees of freedom and a correlation coefficient ρ between the two simultaneous variables of the same pair, while the variables along the time axis are all independent. We have then applied the previous testing procedure to each of the pairs of time series.

Specifically, for each pair of Student's time series, we construct the marginal distributions and transform the Student's variables $x_i(k)$ into their Gaussian counterparts $y_i(k)$ via the transformation (23). For each pair $(y_1(k), y_2(k))$, $k \in \{1, \cdots, T\}$, we estimate its correlation matrix, then construct the time series with T realizations of the random variable $z^2(k)$ defined in (27). The set of T variables z^2 then allows us to construct the distribution of z^2 (with $N = 2$) and to compare it with the χ^2-distribution with two degrees of freedom. We then measure the distances d_1, d_2, d_3 and d_4 defined by (33-36) between the distribution of z^2 and the χ^2-distribution. The significance p_i of these distances d_i is calculated by generating 1,000 Gaussian time series with a correlation matrix equal to the correlation matrix estimated from the original Student's time series, according to the steps 3 to 8 of the testing procedure described in section 3.2. Given a Student's time series with distance d_i, the significance of this distance is

$$p_i = \frac{1}{1000} \sum_{k=1}^{1000} 1_{\{d_i(z^2(y^{(k)}),\chi^2) > d_i\}}, \tag{37}$$

where $y^{(k)} = (y_t^{(k)})_{1 \leq t \leq T}$ denotes the k^{th} replication of a bivariate Gaussian time series of length T and correlation coefficient equal to the correlation coefficient estimated from the original Student's time series.

Repeating this protocol 1,000 times for Student's time series with the same ν and ρ, we then construct the cumulative distribution function $D_i(p)$, $i \in \{1, 2, 3, 4\}$ of the significance p obtained for each of the four distances $d_1,\ d_2,\ d_3$ and d_4. It thus allows us to get the minimum significance level p such that we can discriminate a Student's copula with ν degrees of freedom and correlation

coefficient ρ from a Gaussian copula with the same correlation coefficient, at the confidence level $D_i(p)$, according to the test based upon distance d_i. For instance, the minimum significance level such that we can discriminate a Student's copula with ν degrees of freedom from a Gaussian copula with the same correlation coefficient, at the α-confidence level, according to distance d_i, is given by $D_i(p_\alpha) = \alpha$. A small value of p_α corresponds to a clear distinction between Student's and Gaussian vectors, at the α-confidence level, as it is improbable that Gaussian vectors exhibit a distance larger than found for the Student's vectors.

The cumulative distributions $D_i(p)$ for each of the four distances d_i, $i \in \{1, 2, 3, 4\}$ are shown in figure 3 for $\nu = 4$ degrees of freedom and in figure 4 for $\nu = 20$ degrees of freedom, for 5 different values of the correlation coefficient $\rho = 0.1, 0.3, 0.5, 0.7$ and 0.9. The very steep increase observed for almost all cases in figure 3 reflects the fact that most of the 1,000 Student's vectors with $\nu = 4$ degrees of freedom have a small p, i.e., their copula is easily distinguishable from the Gaussian copula. The same cannot be stated for Student's vectors with $\nu = 20$ degrees of freedom. Note also that the distances d_1, d_2 and d_4 give essentially the same result while the Anderson-Darling distance d_3 is more sensitive to ρ, especially for small ν.

Fixing for instance the confidence level at $\alpha = 95\%$, we can read from each of these curves in figures 3 and 4 the minimum $p_{95\%}$-value necessary to distinguish a Student's copula with a given ν from a Gaussian copula. This $p_{95\%}$ is the abscissa corresponding to the ordinate $D(p_{95\%}) = 0.95$. These values $p_{95\%}$ are reported in table 1, for different values of the number ν of degrees of freedom ranging from $\nu = 3$ to $\nu = 50$ and correlation coefficients $\rho = 0.1$ to 0.9. The values of $p_{95\%}(\nu, \rho)$ reported in table 1 are the minimum values that the statistical significance p should take in order to be able to reject the hypothesis that a Student's copula with ν degrees and correlation ρ can be mistaken with a Gaussian copula at the 95% confidence level.

The results of the table 1 are depicted in figures 5-6 and represent the "power" of the test. The statistical power is usually defined as the probability of rejection of null hypothesis when false. Here, we have not exactly depicted the conventional statistical power of the test, but, more precisely, the minimum significance level allowing for the discrimination between H_0 (the Gaussian copula) and the alternative hypothesis $H_{(\nu^{-1})}$ (Student's copula with ν degrees of freedom).

In the abscissa of figures 5-6 is ploted the inverse ν^{-1} of the number ν of degrees of freedom, which provides a natural "distance" between the Gaussian copula hypothesis $H_0 = H_{(\nu^{-1}=0)}$ and the Student's copula hypothesis $H_{(\nu^{-1})}$. The typical shape of these curves is a sigmoid, starting from a value very close to one for $\nu^{-1} \to 0$, decreasing as ν^{-1} increases and going to 0 as ν^{-1} becomes large enough. This typical shape simply expresses the fact that it is easy to separate a Gaussian copula from a Student's copula with a small number of degrees of freedom, while it is difficult and even impossible for too large a number of degrees of freedom.

The figure 5 shows us that the distances d_1, d_2 and d_3 are not sensitive to the value of the correlation coefficient ρ, while the discriminating power of d_3 increases with ρ. On figure 6, we note that d_2 and d_4 have the same discriminating power for all ρ's (which makes them somewhat redundant) and that they are the most efficient to differentiate H_ν from H_0 for small ρ. When ρ is about 0.5, d_2, d_3 and d_4 (and maybe d_1) are equivalent with respect to the differential power, while for large ρ, d_3 becomes the most discriminating one with high significance.

This study of the test sensitivity involves a non-parametric approach and the question may arise why it should be prefered to a direct parametric test involving for instance the calibration of the Student copula. First, a parametric test of copulas would face the "curse of dimensionality", i.e., the estimation of functions of several variables. With the limited data set available, this does not seem a reasonable approach. Second, we have taken the Student copula as an example of an alternative

to the Gaussian copula. However, our tests are independent of this choice and aim mainly at testing the rejection of the Gaussian copula hypothesis. They are thus of a more general nature than would be a parametric test which would be forced to choose one family of copulas with the problem of excluding others. The parametric test would then be exposed to the criticism that the rejection of a given choice might not be of a general nature.

In the sequel, we will choose the level of 95% as the level of rejection, which leads us to neglect one extreme event out of twenty. This is not unreasonable in view of the other significant sources of errors resulting in particular from the empirical determination of the marginals and from the presence of outliers for instance.

4 Empirical results

We investigate the following assets :

- foreign exchange rates,
- metals traded on the London Metal Exchange,
- stocks traded on the New York Stocks Exchange.

4.1 Currencies

The sample we have considered is made of the daily returns for the spot foreign exchanges for 6 currencies[8] : the Swiss Franc (CHF), the German Mark (DEM), the Japanese Yen (JPY), the Malaysian Ringgit (MYR), the Thai Baht (THA) and the British Pound (UKP). All the exchange rates are expressed against the US dollar. The time interval runs over ten years, from January 25, 1989 to December 31, 1998, so that each sample contains 2500 data points.

We apply our test procedure to the entire sample and to two sub-samples of 1250 data points so that the first one covers the time interval from January 25, 1989 to January 11, 1994 and the second one from January 12, 1994 to December 31, 1998. The results are presented in tables 2 to 4 and depicted in figures 7 to 9.

Tables 2-4 give, for the total time interval and for each of the two sub-intervals, the probability $p(d)$ to obtain from the Gaussian hypothesis a deviation between the distribution of the z^2 and the χ^2-distribution with two degrees of freedom larger than the observed one for each of the 15 pairs of currencies according to the distances d_1-d_4 defined by (33)-(36).

The figures 7-9 organize the information shown in the tables 2-4 by representing, for each distance d_1 to d_4, the number of currency pairs that give a test-value p within a bin interval of width 0.05. A clustering close to the origin signals a significant rejection of the Gaussian copula hypothesis.

At the 95% significance level, table 2 and figure 7 show that only 40% (according to d_1 and d_3) but 60% (according to d_2 and d_4) of the tested pairs of currencies are compatible with the Gaussian copula hypothesis over the entire time interval. During the first half-period from January 25, 1989 to January 11, 1994 (table 3 and figure 8), 47% (according to d_3) and up to about 75 % (according

[8]The data come from the historical database of the Federal Reserve Board.

to d_2 and d_4) of the tested currency pairs are compatible with the assumption of Gaussian copula, while during the second sub-period from January 12, 1994 to December 31, 1998 (table 4 and figure 9), between 66% (according to d_1) and about 75% (according to d_2, d_3 and d_4) of the currency pairs remain compatible with the Gaussian copula hypothesis. These results raise several comments both on a statistical and an economic point of view.

We first note that the most significant rejection of the Gaussian copula hypothesis is obtained for the distance d_3, which is indeed the most sensitive to the events in the tail of the distributions. The test statistics given by this distance can indeed be very sensitive to the presence of a single large event in the sample, so much so that the Gaussian copula hypothesis can be rejected only because of the presence of this single event (outlier). The difference between the results given by d_3 and d_4 (the averaged d_3) are very significant in this respect. Consider for instance the case of the German Mark and the Swiss Franc. During the time interval from January 12, 1994 to December 31, 1998, we check on table 4 that the non-rejection probability $p(d)$ is very significant according to d_1, d_2 and d_4 ($p(d) \geq 31\%$) while it is very low according to d_3: $p(d) = 0.05\%$, and should lead to the rejection of the Gaussian copula hypothesis. This suggests the presence of an outlier in the sample.

To check this hypothesis, we show in the upper panel of figure 10 the function

$$f_3(t) = \frac{|F_{z^2}(z^2(t) - F_{\chi^2}(\chi^2(t))|}{\sqrt{F_{\chi^2}(\chi^2)[1 - F_{\chi^2}(\chi^2)]}} \, , \tag{38}$$

used in the definition of the Anderson-Darling distance $d_3 = \max_z f_3(z)$ (see definition (35)), expressed in terms of time t rather than z^2. The function have been computed over the two time sub-intervals separately.

Apart from three extreme peaks occurring on June 20, 1989, August 19, 1991 and September 16, 1992 during the first time sub-interval and one extreme peak on September 10, 1997 during the second time sub-interval, the statistical fluctuations measured by $f_3(t)$ remain small and of the same order. Excluding the contribution of these outlier events to d_3, the new statistical significance derived according to d_3 becomes similar to that obtained with d_1, d_2 and d_4 on each sub-interval. From the upper panel of figure 10, it is clear that the Anderson-Darling distance d_3 is equal to the height of the largest peak corresponding to the event on August 19, 1991 for the the first period and to the event on September 10, 1997 for the second period. These events are depicted by a circled dot in the two lower panels of figure 10, which represent the return of the German Mark versus the return of the Swiss Franc over the two considered time periods.

The event on August 19, 1991 is associated with the coup against Gorbachev in Moscow: the German mark (respectively the Swiss franc) lost 3.37% (respectively 0.74%) in daily annualized value against the US dollar. The 3.37% drop of the German Mark is the largest daily move of this currency against the US dollar over the whole first period. On September 10, 1997, the German Mark appreciated by 0.60% against the US dollar while the Swiss Franc lost 0.79% which represents a moderate move for each currency, but a large joint move. This event is related to the contradictory announcements of the Swiss National Bank about the monetary policy, which put an end to a rally of the Swiss Franc along with the German mark against the US dollar.

Thus, neglecting the large moves associated with major historical events or events associated with unexpected incoming information[9], which cannot be taken into account by a statistical study, we obtain, for d_3, significance levels compatible with those obtained with the other distances. We can thus conclude that, according to the four distances, during the time interval from January 12,

[9]The outlier nature of the event on August 19, 1991 has been clearly demonstrated by Sornette *et al.* (2003).

1994 to December 31, 1998 the Gaussian copula hypothesis cannot be rejected for the couple German Mark / Swiss Franc.

However, the non-rejection of the Gaussian copula hypothesis does not always have minor consequences and may even lead to serious problem in stress scenarios. As shown in section 3.3, the non-rejection of the Gaussian copula hypothesis does not exclude, at the 95% significance level, that the dependence of the currency pairs may be accounted for by a Student's copula with adequate values of ν and ρ. Still considering the pair German Mark / Swiss Franc, we see in table 1 that, according to d_1, d_2 and d_4, a Student's copula with about five degrees of freedom allows to reach the test values given in table 4. But, with the correlation coefficient $\rho = 0.92$ for the German Mark/Swiss Franc couple, the Gaussian copula assumption could lead to neglect a tail dependence coefficient $\lambda_5(0.92) = 63\%$ according to the Student's copula prediction. Such a large value of $\lambda_5(0.92)$ means that when an extreme event occurs for the German Mark it also occurs for the Swiss Franc with a probability equals to 0.63. Therefore, a stress scenario based on a Gaussian copula assumption would fail to account for such coupled extreme events, which may represent as many as two third of all the extreme events, if it would turn out that the true copula would be the Student's copula with five degrees of freedom. In fact, with such a value of the correlation coefficient, the tail dependence remains high even if the number of degrees of freedom reach twenty or more (see figure 1).

The case of the Swiss Franc and the Malaysian Ringgit offers a striking difference. For instance, in the second half-period, the test statistics $p(d)$ are greater than 70% and even reach 91% while the correlation coefficient is only $\rho = 0.16$, so that a Student's copula with 7-10 degrees of freedom can be mistaken with the Gaussian copula (see table 1). Even in the most pessimistic situation $\nu = 7$, the choice of the Gaussian copula amounts to neglecting a tail dependence coefficient $\lambda_5(0.16) = 4\%$ predicted by the Student's copula. In this case, stress scenarios based on the Gaussian copula would predict uncoupled extreme events, which would be shown wrong only once out of twenty five times.

These two examples show that, more than the number of degrees of freedom of the Student's copula necessary to describe the data, the key parameter is the correlation coefficient.

From an economic point of view, the impact of regulatory mechanisms between currencies or monetary crisis can be well identified by the rejection or absence of rejection of our null hypothesis. Indeed, consider the couple German Mark / British Pound. During the first half period, their correlation coefficient is very high ($\rho = 0.82$) and the Gaussian copula hypothesis is strongly rejected according to the four distances. On the contrary, during the second half period, the correlation coefficient significantly decreases ($\rho = 0.56$) and none of the four distances allows us to reject our null hypothesis. Such a non-stationarity can be easily explained. Indeed, on January 1, 1990, the British Pound entered the European Monetary System (EMS), so that the exchange rate between the German Mark and the British Pound was not allowed to fluctuate beyond a margin of 2.25%. However, due to a strong speculative attack, the British Pound was devaluated on September 1992 and had to leave the EMS. Thus, between January 1990 and September 1992, the exchange rate of the German Mark and the British Pound was confined within a narrow spread, incompatible with the Gaussian copula description. After 1992, the British Pound exchange rate floated with respect to German Mark, the dependence between the two currencies decreased, as shown by their correlation coefficient. In this regime, we can no more reject the Gaussian copula hypothesis.

The impact of major crisis on the copula can be also clearly identified. Such a case is exhibited by the couple Malaysian Ringgit/Thai Baht. Indeed, during the period from January 1989 to January 1994, these two currencies have only undergone moderate and weakly correlated ($\rho = 0.29$) fluctuations, so that our null hypothesis cannot be rejected at the 95% significance level. On the

contrary, during the period from January 1994 to October 1998, the Gaussian copula hypothesis is strongly rejected. This rejection is obviously due to the persistent and dependent ($\rho = 0.44$) shocks incured by the Asian financial and monetary markets during the seven months of the Asian Crisis from July 1997 to January 1998 (Baig and Goldfajn 1998, Kaminsky and Schlmukler 1999).

These two cases show that the Gaussian copula hypothesis can be considered reasonable for currencies in absence of regulatory mechanisms and of strong and persistent crises. They also allows us to understand why the results of the test over the entire sample are so much weaker than the results obtained for the two sub-intervals: the time series are strongly non-stationary.

4.2 Commodities: metals

We consider a set of 6 metals traded on the London Metal Exchange: aluminum, copper, lead, nickel, tin and zinc. Each sample contains 2270 data points and covers the time interval from January 4, 1989 to December 30, 1997. The results are synthesized in table 5 and in figure 11.

Table 5 gives, for each of the 15 pairs of commodities, the probability $p(d)$ to obtain from the Gaussian hypothesis a deviation between the distribution of the z^2 and the χ^2-distribution with two degrees of freedom larger than the observed one for the commodity pair according to the distances d_1-d_4 defined by (33)-(36).

The figure 11 organizes the information shown in table 5 by representing, for each distance, the number of commodity pairs that give a test-value p within a bin interval of width 0.05. A clustering close to the origin signals a significant rejection of the Gaussian copula hypothesis.

According to the three distances d_1, d_2 and d_4, at least two third and up to 93% of the set of 15 pairs of commodities are inconsistent with the Gaussian copula hypothesis. Surprisingly, according to the distance d_3, at the 95% significance level, two third of the set of 15 pairs of commodities remain compatible with the Gaussian copula hypothesis. This is the reverse to the previous situation found for currencies. These test values lead to globally reject the Gaussian copula hypothesis.

Moreover, the largest value obtained for the distance d_3 is $p = 65\%$ for the pair copper-tin, which is significantly smaller than the 80% or 90% reached for some currencies over a similar time interval. Thus, even in the few cases where the Gaussian copula assumption is not rejected, the test values obtained are not really sufficient to distinguish between the Gaussian copula and a Student's copula with $\nu = 5 \sim 6$ degrees of freedom. In such a case, with correlation coefficients ranging between 0.31 and 0.46, the tail dependence neglected by keeping the Gaussian copula is no less than 10% and can reach 15%. One extreme event out of seven or ten might occur simultaneously on both marginals, which would be missed by the Gaussian copula.

To summarize, the Gaussian copula does not seem a reasonable assumption for metals, and it has not appeared necessary to test these data over smaller time interval.

4.3 Stocks

We now study the daily returns distributions for 22 stocks among the largest companies quoted on the New York Stock Exchange[10]: Appl. Materials (AMAT), AT&T (T), Citigroup (C), Coca Cola (KO), EMC, Exxon-Mobil (XOM), Ford (F), General Electric (GE), General Motors (GM), Hewlett Packard (HPW), IBM, Intel (INTC), MCI WorldCom (WCOM), Medtronic (MDT), Merck

[10]The data come from the Center for Research in Security Prices (CRSP) database.

(MRK), Microsoft (MSFT), Pfizer (PFE), Procter&Gamble (PG), SBC Communication (SBC), Sun Microsystem (SUNW), Texas Instruments (TXN), Wal Mart (WMT).

Each sample contains 2500 data points and covers the time interval from February 8, 1991 to December 29, 2000 and have been divided into two sub-samples of 1250 data points, so that the first one covers the time interval from February 8, 1991 to January 18, 1996 and the second one from January 19, 1996 to December 20, 2000. The results of fifteen randomly chosen pairs of assets are presented in tables 6 to 8 while the results obtain for the entire set are represented in figures 12 to 14.

At the 95% significance level, figure 12 shows that 75% of the pairs of stocks are compatible with the Gaussian copula hypothesis. Figure 13 shows that over the time interval from February 1991 to January 1996, this percentage becomes larger than 99% for d_1, d_2 and d_4 while it equals 94% according to d_3. It is striking to note that, during this period, according to d_1, d_2 and d_4, more than a quarter of the stocks obtain a test-value p larger than 90%, so that we can assert that they are completely inconsistent with the Student's copula hypothesis for Student's copulas with less than 10 degrees of freedom. Among this set of stocks, not a single one has a correlation coefficient larger than 0.4, so that a scenario based on the Gaussian copula hypothesis leads to neglecting a tail dependence of less than 5% as would be predicted by the Student's copula with 10 degrees of freedom. In addition, about 80% of the pairs of stocks lead to a test-value p larger than 50% according to the distances d_1, d_2 and d_4, so that as much as 80% of the pairs of stocks are incompatible with a Student's copula with a number of degrees of freedom less than or equal to 5. Thus, for correlation coefficients smaller than 0.3, the Gaussian copula hypothesis leads to neglecting a tail dependence less than 10%. For correlation coefficients smaller than 0.1 which corresponds to 13% of the total number of pairs, the Gaussian copula hypothesis leads to neglecting a tail dependence less than 5%.

Figure 14 shows that, over the time interval from January 1996 to December 2000, 92% of the pairs of stocks are compatible with the Gaussian copula hypothesis according to d_1, d_2 and d_4 and more than 79% according to d_3. About a quarter of the pair of stocks have a test-value p larger than 50% according to the four measures and thus are inconsistent with a Student's copula with less than five degrees of freedom.

For completeness, we present in table 9 the results of the tests performed for five stocks belonging to the computer area : Hewlett Packard, IBM, Intel, Microsoft and Sun Microsystem. We observe that, during the first half period, all the pairs of stocks qualify the Gaussian copula Hypothesis at the 95% significance level. The results are rather different for the second half period since about 40% of the pairs of stocks reject the Gaussian copula hypothesis according to d_1, d_2 and d_3. This is probably due to the existence of a few shocks, notably associated with the crash of the "new economy" in March-April 2000.

On the whole, it appears however that there is no systematic rejection of the Gaussian copula hypothesis for stocks within the same industrial area, notwithstanding the fact that one can expect stronger correlations between such stocks than for currencies for instance.

5 Conclusion and comparison with other studies

We have studied the null hypothesis that the dependence between pairs of financial assets can be modeled by the Gaussian copula.

Our test procedure is based on the following simple idea. Assuming that the copula of two assets X and Y is Gaussian, then the multivariate distribution of (X, Y) can be mapped into a Gaussian multivariate distribution, by a transformation of each marginal into a normal distribution, which leaves the copula of X and Y unchanged. Testing the Gaussian copula hypothesis is therefore equivalent to the more standard problem of testing a two-dimensional multivariate Gaussian distribution. We have used a bootstrap method to determine and calibrate the test statistics. Four different measures of distances between distributions, more or less sensitive to the departure in the bulk or in the tail of distributions, have been proposed to quantify the probability of rejection of our null hypothesis.

Our tests have been performed over three types of assets: currencies, commodities (metals) and stocks. In most cases, for currencies and stocks, the Gaussian copula hypothesis can not be rejected at the 95% confidence level. For currencies, according to three of the four distances at least,

- 40% of the pairs of currencies, over a 10 years time interval (due to non-stationary data),
- 67% of the pairs of currencies, over the first 5 years time interval,
- 73% of the pairs of currencies, over the second 5 years time interval,

are compatible with the Gaussian copula hypothesis. For stocks, we have shown that

- 75% of the pairs of stocks, over a 10 years time interval,
- 93% of the pairs of stocks, over the first 5 years time interval,
- 92% of the pairs of stocks, over the second 5 years time interval,

are compatible with the Gaussian copula hypothesis. In contrast, the Gaussian copula hypothesis cannot be considered as reasonable for metals : between 66% and 93% of the pairs of metals reject the null hypothesis at the 95% confidence level.

Notwithstanding the apparent qualification of the Gaussian copula hypothesis for most of the currencies and the stocks we have analyzed, we must bear in mind the fact that a non-Gaussian copula cannot be rejected. In particular, we have shown that a Student's copula can always be mistaken for a Gaussian copula if its number of degrees of freedom is sufficiently large. Then, depending on the correlation coefficient, the Student's copula can predict a non-negligible tail dependence which is completely missed by the Gaussian copula assumption. In other words, the Gaussian copula predicts no tail dependences and therefore does not account for extreme events that may occur simultaneously but nevertheless too rarely to modify the test statistics. To quantify the probability for neglecting such events, we have investigated the situations when one is unable to distinguish between the Gaussian and Student's copulas for a given number of degrees of freedom. Our study leads to the conclusion that it may be very dangerous to embrace blindly the Gaussian copula hypothesis when the correlation coefficient between the pair of asset is too high as the tail dependence neglected by the Gaussian copula can be as large as 0.6. In this respect, the case of the Swiss Franc and the German Mark is striking. The test values p obtained are very significant (about 33%), so that we cannot mistake the Gaussian copula for a Student's copula with less than 5-7 degrees of freedom. However, their correlation coefficient is so high ($\rho = 0.9$) that a Student's copula with, say $\nu = 30$ degrees of freedom, still has a large tail dependence.

This remark shows that it is highly desirable to test for other non-Gaussian copulas, such as the Student's copula. Breymann *et al.* (2003) have recently shown that the dependence structure of

the couple German Mark / Japanese Yen is (slightly?) better described by a Student's copula with about six degrees of freedom (for daily returns) than a Gaussian copula, according to the Akaike information criterion. This result is compatible with and precises ours, since in table 2 - where investigated period covers a time interval comparable with that used by Breymann *et al.* (2003) - we see that our test cannot reject a Student's copula with more than 3-4 degrees of freedom. However, the stationarity of the data over such a long period is not well ascertained, as proved by the results in tables 3-4, where we observe an important increase of the significance of the non-rejection of the Gaussian copula hypothesis during the second time interval with respect to the first one. In both cases, however, significance levels remain consistent with the non-rejection of a Student's copula with about 6 degrees of freedom.

In the study by Mashal and Zeevi (2002), it is claimed that the dependence between stocks is significantly better accounted for by a Student's copula with 11-12 degrees of freedom than by a Gaussian copula. Again, our results are compatible with those ones. However, contrarily to the case of currencies, the real improvement brought by the description of the dependence between stocks in terms of a Student's copula is questionable. Indeed, as underlined in section 4.3, correlation coefficients between two stocks are hardly greater than 0.4, so that the tail dependence of the Student's copula with 11-12 degrees of freedom is about 2.5% or less (see figures 1-2). In view of all the different sources of uncertainty during the estimation process in addition to the non-stationarity of the data, we doubt that such a description eventually leads to concrete improvement for practical purposes.

Finally, we want to stress that the question of the assessment of the coefficient of tail dependence must be studied for its own. Indeed, as we have seen, copulas estimation only yields poor estimate of this quantity, and are mainly based on the *a priori* assumption of the existence or not of such a tail dependence. Therefore, we think that it is necessary to develop tests that are specific to the detection of a possible tail dependence between two time series. Some results concerning stocks have been obtained by Malevergne and Sornette (2002a, b) and indicate the existence of a tail dependence ranging between about five and fifteen percent during the time period considered in the present study (see Malevergne and Sornette (2002a)). Such estimates of the coefficient of tail dependence are barely compatible with estimates performed under the Student's copula hypothesis and more generally under the elliptical copula assumption. Thus, as a conservative conclusion and in view of the different studies concerning this problem, we think that the Gaussian copula provides the most parsimonious description of the dependence between stock returns, apart from crisis periods. In such periods, the Student's copula does not bring a really better practical model since it turns out that it still underestimate the dependence of tail events.

To our knowledge, no direct investigation of the tail dependence between currencies have yet been performed. Thus, we cannot raise the same conclusion as for stocks, and assert that the Student's copula still underestimate the tail dependence. As a consequence, for such assets, the prudence leads to recommend the choice of the Student's copula with respect to the Gaussian copula for risk management purposes.

References

Andersen, J.V. and D. Sornette, 2001, Have your cake and eat it too: increasing returns while lowering large risks! Journal of Risk Finance 2, 70-82.

Anderson T. W. and D.A. Darling, 1952, Asymptotic theory of certain "goodness of fit" criteria, Annals of Mathematical Statistics 23, 193-212.

Ang, A. and J. Chen, 2001, Asymmetric correlations of equity portfolio, Working Paper.

Baig, T. and I. Goldfjn, 1998, Financial Markets Contagion in the Asian crisis, Mimeo, International Monetary Fund.

Basle commitee on supervision banking, 2001, *The New Basle Accord*.

Breymann, W., A. Dias and P. Embrechts, 2003, Dependence structures for multivariate high-frequency data in finance, *Quantitative Finance* **3**, - .

Chen, K. and S.-H. Lo, 1997, On a Mapping Approach to Investigating the Bootstrap Accuracy, Probability Theory and Related Fields 107, 197-217.

Cherubini, U. and E. Luciano, 2000, Multivariate option pricing with copulas, Working Paper.

Efron, B., and R.Tibshirani, 1986, Bootstrap method for standard errors, confidence intervals and other measures of statistical accuracy, Statistical Science 1, 54-77.

Embrechs P., A.J. McNeil and D. Straumann, 1999, Correlation : Pitfalls and Alternatives. Risk, 69-71.

2001. Embrechts, P., A. Hoeing, A. Juri, 2001, Using Copulae to bound the Value-at-Risk for functions of dependent risk, ETH Zurick Working Paper.

Embrechts P., A.J. McNeil and D. Straumann, 2001, Correlation and Dependency in Risk Management : Properties and Pitfalls, in : Dempster, M., ed., Value at Risk and Beyond (Cambridge University Press).

Frees, E. and E. Valdez, 1998, Understanding Relationships using copula,. North American Actuarial Journal 2, 1-25.

Genest, C. and R. MacKay, 1986, The Joy of Copulas, American Statistician 40, 280-283.

Genest, C., 1987, Frank's Family of Bivariate Distributions, Biometrika 74, 549-555.

Genest, C. and J.P. Rivest, 1993, Statistical Inference Procedures for Bivariate Archimedean Copulas, Journal of American Statistical Association 88, 1034-1043.

Gopikrishnan, P., M. Meyer, L.A.N. Amaral and H.E. Stanley, 1998, Inverse Cubic Law for the Distribution of Stock Price Variation, European Physical Journal B 3, 139-140.

Gouriéroux, C. and J. Jasiak, 1999, Truncated Local Likelihood and Non-parametric tail analysis, DP 99 CREST.

Guillaume, D.M., M.M Dacorogna, R.R. Davé, J.A Muller, R.B. Olsen and O.V. Pictet, 1997, From the Bird Eye to the Microscope : A Survey of the New Stylized Facts of the Intra-Daily Foreign Exchange Markets, Finance and Stochastics 1, 95-129.

Haas, C.N., 1999, On Modeling Correlated Random Variables in Risk Assessment, Risk Analysis 19, 1205-1214.

Joe, H., 1993, Parametric Families of Multivariate Distributions with Given Marginals, Journal of Multivariate Analysis 46, 262-282.

Kaminsky, G.L. and S.L. Schmukler, 1999, What Triggers Market Jitters? A Chronicle of the Asian Crisis, Journal of International Money and Finance 18, 537-560.

Klugman, S.A. and R. Parsa, 1999, Fitting Bivariate Loss Distribution with Copulas, Insurance : Mathematics and Economics 24, 139-148.

KMV-Corporation, 1997, Modelling Default Risk, *Technical Document*.

Laherrère, J. and D. Sornette, 1999, Stretched exponential distributions in Nature and Economy: "Fat tails" with characteristic scales, European Physical Journal B 2, 525-539.

Longin F. and B. Solnik, Extreme Correlation of International Equity Markets, The Journal of Finance LVI, 649-676 (2001).

Lindskog, F., 1999, Modelling Dependence with Copulas,
$http://www.risklab.ch/Papers.html\#MTLindskog$

Lux, L., 1996, The Stable Paretian Hypothesis and the Frequency of Large Returns : an Examination of Major German Stocks, Applied Financial Economics 6, 463-475.

Malevergne, Y., V. Pisarenko and D. Sornette, Empirical distributions of log-returns: Between the stretched-exponential and the power law?, Working Paper.

Malevergne, Y. and D. Sornette, 2002a, Minimizing extremes, *Risk* **15**(11), 129-132.

Malevergne, Y. and D. Sornette, 2002b, Tail dependence of factor models, *Working paper*.

Markovitz, H., 1959, Portfolio selection : Efficient diversification of investments (John Wiley and Sons, New York).

Mashal and Zeevi, 2002, Beyond correlation: *Working Paper*.

McNeil, A. and R. Frey, 2000, Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach, *Journal of Empirical Finance* **7**, 271-300.

Nelsen, R.B., 1998, An Introduction to Copulas. Lectures Notes in statistic 139 (Springer Verlag, New York).

Muzy, J.F., J. Delour and E. Bacry, 2000, Modelling fluctuations of financial time series: From cascade process to stochastic volatility model, *European Physical Journal B* **17**, 537-548.

Muzy, J.F., D. Sornette, J. Delour and A. Arnéodo, 200A, Multifractal returns and hierarchical portfolio theory, *Quantitative Finance* **1**, 131-148.

Pagan, A., 1996, The Econometrics of Financial Markets, Journal of Empirical Finance 3, 15-102.

Patton, A., 2001, Estimation of copula models for time series of possibly different lengths, *Working Paper*.

RiskMetrics-Group, 1997, CreditMetrics, *Technical Document*.

Rosenberg, J.V., 1999, Semiparametric pricing of multivariate contingent claims, NYU-Stern School of Business Working Paper.

Sornette, D., J. V. Andersen and P. Simonetti, 2000a, Portfolio Theory for "Fat Tails", International Journal of Theoretical and Applied Finance 3 , 523-535.

Sornette, D., P. Simonetti and J. V. Andersen, 2000b, ϕ^q-field theory for Portfolio optimization: "fat tails" and non-linear correlations, Physics Report 335, 19-92.

Sornette, D., Y. Malevergne and J.F. Muzy, 2003, , *Risk* **16**(2),.

Starica, C., 1999, Multivariate extremes for models with constant conditional correlations. Journal of Empirical Finance, 6, 515-553.

Vries, C.G. de, 1994, Stylized Facts of Nominal Exchange Rate Returns, in : Van der Ploeg, F., ed., The Handbook of international Macroeconomics (Blackwell, Oxford) 348-389.

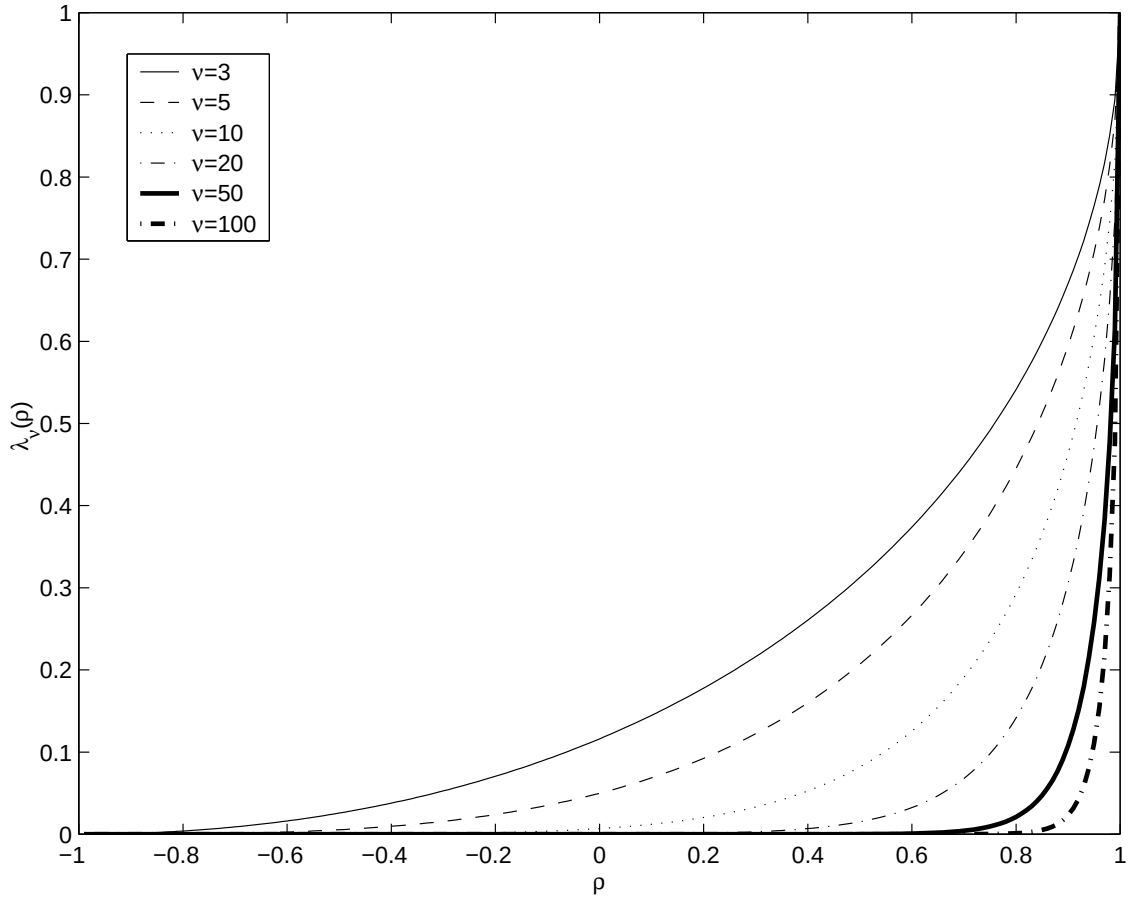

Figure 1: Upper tail dependence coefficient $\lambda_\nu(\rho)$ for the Student's copula with ν degrees of freedom as a function of the correlation coefficient ρ, for different values of ν.

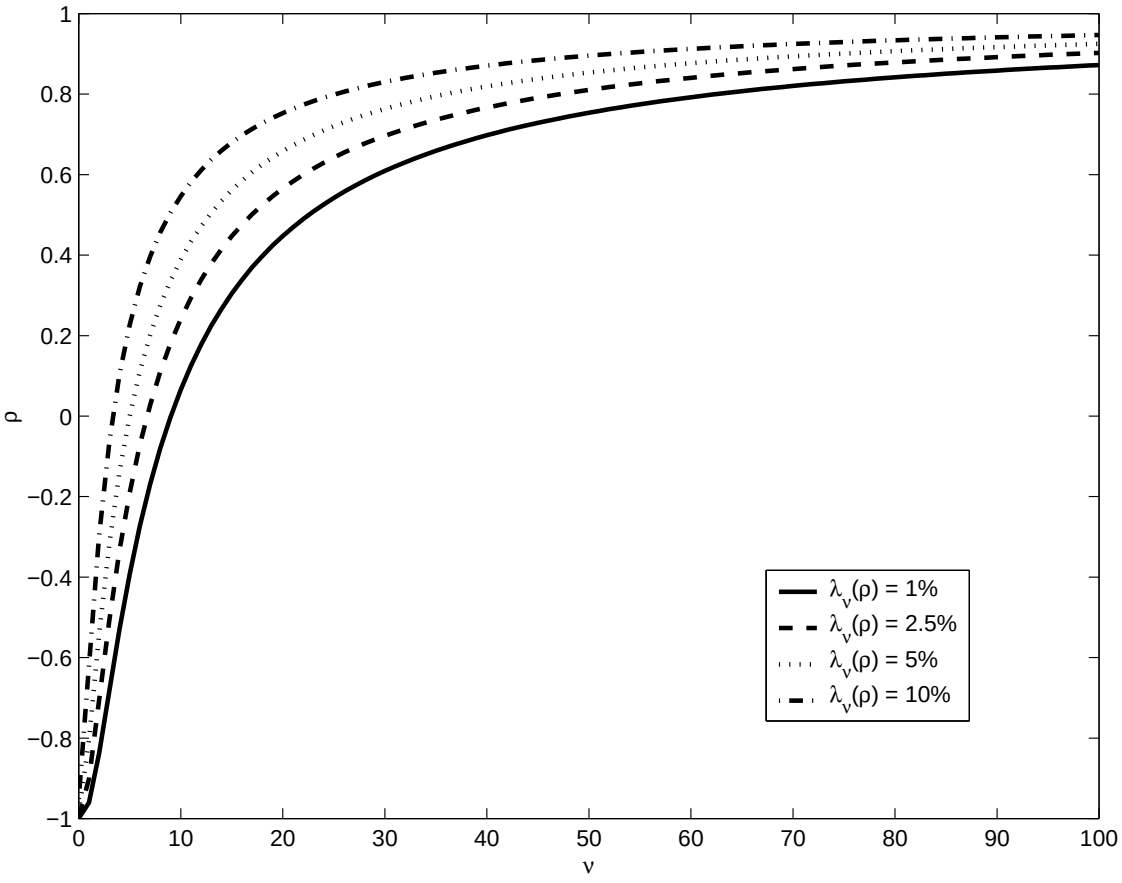

Figure 2: Maximum value of the correlation coefficient ρ as a function of ν, below which the tail dependence $\lambda_\nu(\rho)$ of a Student' copula is smaller than a given small value, here taken equal to $\lambda_\nu(\rho) = 1\%, 2.5\%, 5\%$ and 10%. The choice $\lambda_\nu(\rho) = 5\%$ for instance corresponds to 1 event in 20 for which the pair of variables are asymptotically coupled. At the $1 - \lambda_\nu(\rho)$ probability level, values of $\lambda \leq \lambda_\nu(\rho)$ are undistinguishable from 0, which means that the Student's copula can be approximated by a Gaussian copula.

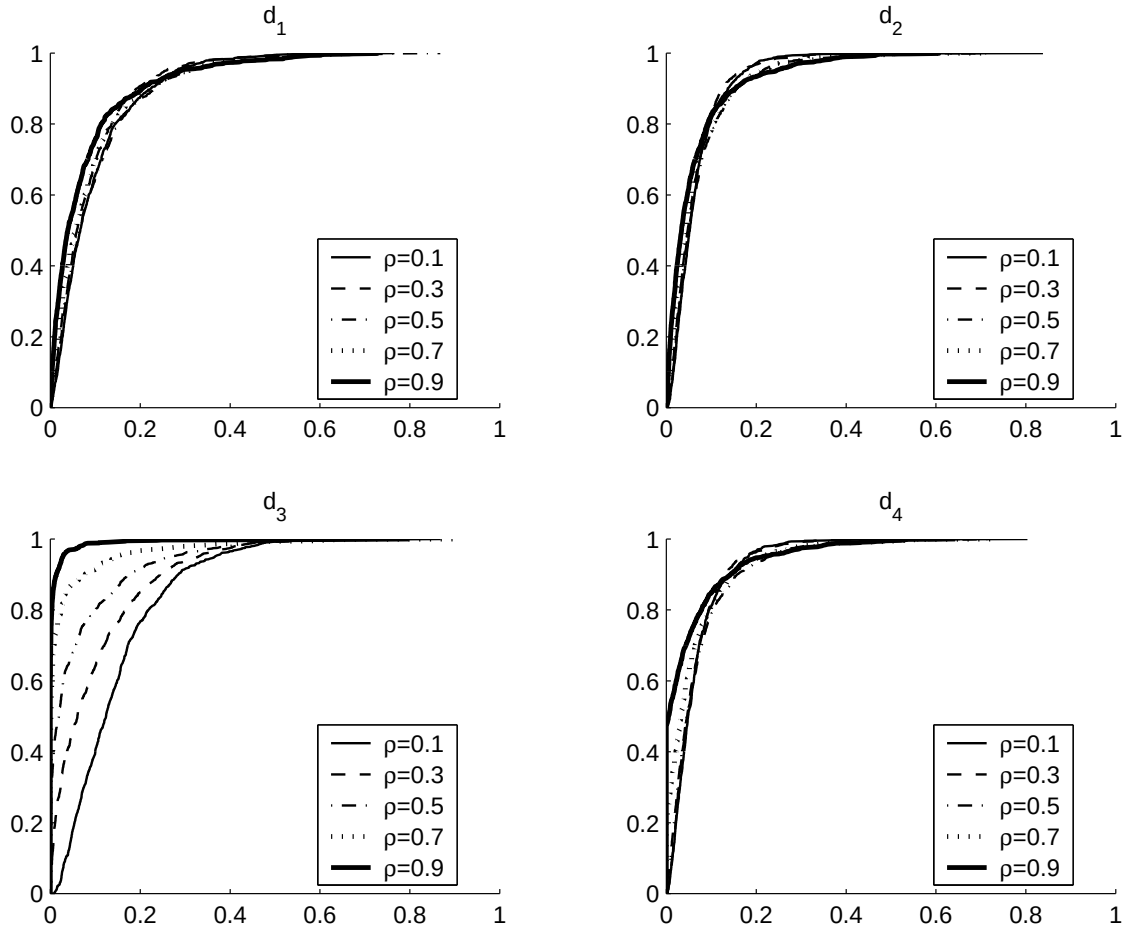

Figure 3: Cumulative distribution function $D(p)$ obtained as the fraction of Student's pairs with $\nu = 4$ degrees of freedom that exhibit a value of at lost p for the probability that Gaussian vectors can have a similar or larger distance. See the text for a detailed description of how $D(p)$ is defined and constructed. Each panel corresponds to one of the four distances d_i, $i \in \{1, 2, 3, 4\}$,defined in the text by equations (33-36). In each panel, we construct the cumulative distribution function $D(p)$ for 5 different values of the correlation coefficient $\rho = 0.1, 0.3, 0.5, 0.7$ and 0.9 of the Student's copula.

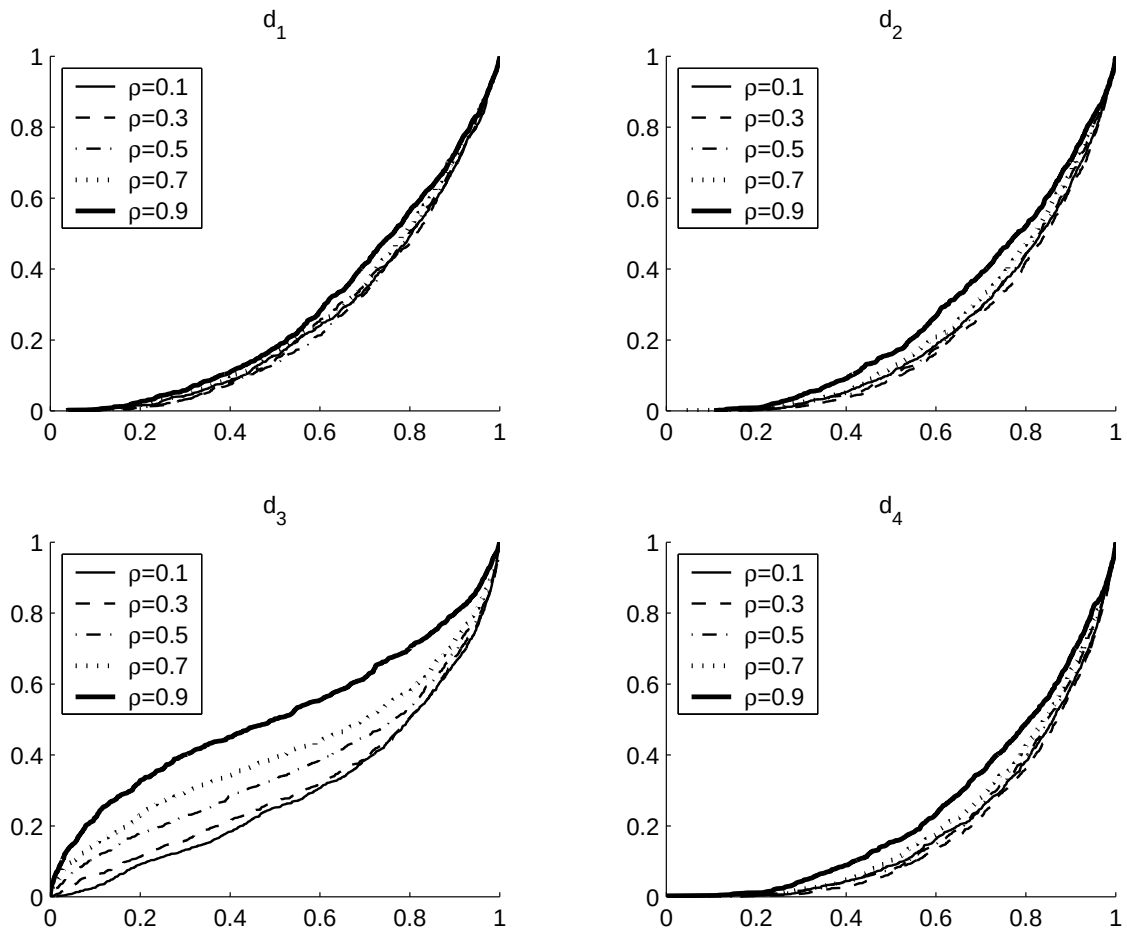

Figure 4: Same as figure 3 for Student's distributions with $\nu = 20$ degrees of freedom.

	ρ	0.1	0.3	0.5	0.7	0.9		ρ	0.1	0.3	0.5	0.7	0.9
	d_1	0.07	0.08	0.07	0.04	0.07		d_1	0.28	0.26	0.32	0.30	0.29
$\nu = 3$	d_2	0.03	0.03	0.07	0.04	0.06	$\nu = 4$	d_2	0.18	0.17	0.21	0.21	0.24
	d_3	0.22	0.17	0.08	0.03	0.01		d_3	0.36	0.33	0.26	0.15	0.03
	d_4	0.03	0.03	0.08	0.03	0.04		d_4	0.18	0.17	0.23	0.21	0.21
	ρ	0.1	0.3	0.5	0.7	0.9		ρ	0.1	0.3	0.5	0.7	0.9
	d_1	0.46	0.47	0.46	0.52	0.52		d_1	0.78	0.81	0.81	0.81	0.86
$\nu = 5$	d_2	0.36	0.34	0.39	0.44	0.43	$\nu = 7$	d_2	0.71	0.78	0.76	0.77	0.82
	d_3	0.52	0.54	0.47	0.30	0.14		d_3	0.80	0.81	0.82	0.73	0.52
	d_4	0.37	0.36	0.43	0.45	0.45		d_4	0.75	0.81	0.79	0.80	0.83
	ρ	0.1	0.3	0.5	0.7	0.9		ρ	0.1	0.3	0.5	0.7	0.9
	d_1	0.85	0.86	0.87	0.88	0.89		d_1	0.92	0.93	0.96	0.95	0.94
$\nu = 8$	d_2	0.85	0.84	0.86	0.87	0.88	$\nu = 10$	d_2	0.93	0.92	0.95	0.96	0.94
	d_3	0.91	0.91	0.91	0.81	0.70		d_3	0.96	0.96	0.96	0.95	0.88
	d_4	0.86	0.85	0.90	0.89	0.90		d_4	0.94	0.94	0.96	0.97	0.95
	ρ	0.1	0.3	0.5	0.7	0.9		ρ	0.1	0.3	0.5	0.7	0.9
	d_1	0.97	0.99	0.97	0.99	0.99		d_1	0.99	0.99	0.99	0.99	0.99
$\nu = 20$	d_2	0.99	0.99	0.97	0.99	0.99	$\nu = 50$	d_2	0.99	0.99	0.99	0.99	0.99
	d_3	0.99	0.99	0.98	0.99	0.97		d_3	0.99	0.99	0.99	0.99	0.99
	d_4	0.99	0.99	0.98	0.99	0.99		d_4	0.99	0.99	0.99	0.99	0.99

Table 1: The values $p_{95\%}(\nu, \rho)$ shown in this table give the minimum values that the significane p should take in order to be able to reject the hypothesis that a Student's copula with ν degrees and correlation ρ is undistinguishable from a Gaussian copula at the 95% confidence level. $p_{95\%}$ is the abscissa corresponding to the ordinate $D(p_{95\%}) = 0.95$ shown in figures 3 and 4. p is the probability that pairs of Gaussian random variables with the correlation coefficient ρ have a distance (between the distribution of z^2 and the theoretical χ^2 distribution) equal to or larger than the corresponding distance obtained for the Student's vector time series. A small p corresponds to a clear distinction between Student's and Gaussian vectors, as it is improbable that Gaussian vectors exhibit a distance larger than found for the Student's vectors. Different values of the number ν of degrees of freedom ranging from $\nu = 3$ to $\nu = 50$ and of the correlation coefficient $\rho = 0.1$ to 0.9 are shown. Let us take for instance the example with $\nu = 4$ and $\rho = 0.3$. The table indicates that p should be less than about 0.3 (resp. 0.2) according to the distances d_1 and d_3 (resp. d_2 and d_4) for being able to distinguish this Student's copula from the Gaussian copula at the 95% confidence level. This means that less than $20 - 30\%$ of Gaussian vectors should have a distance for their z^2 larger than the one found for the Student's. See text for further explanations.

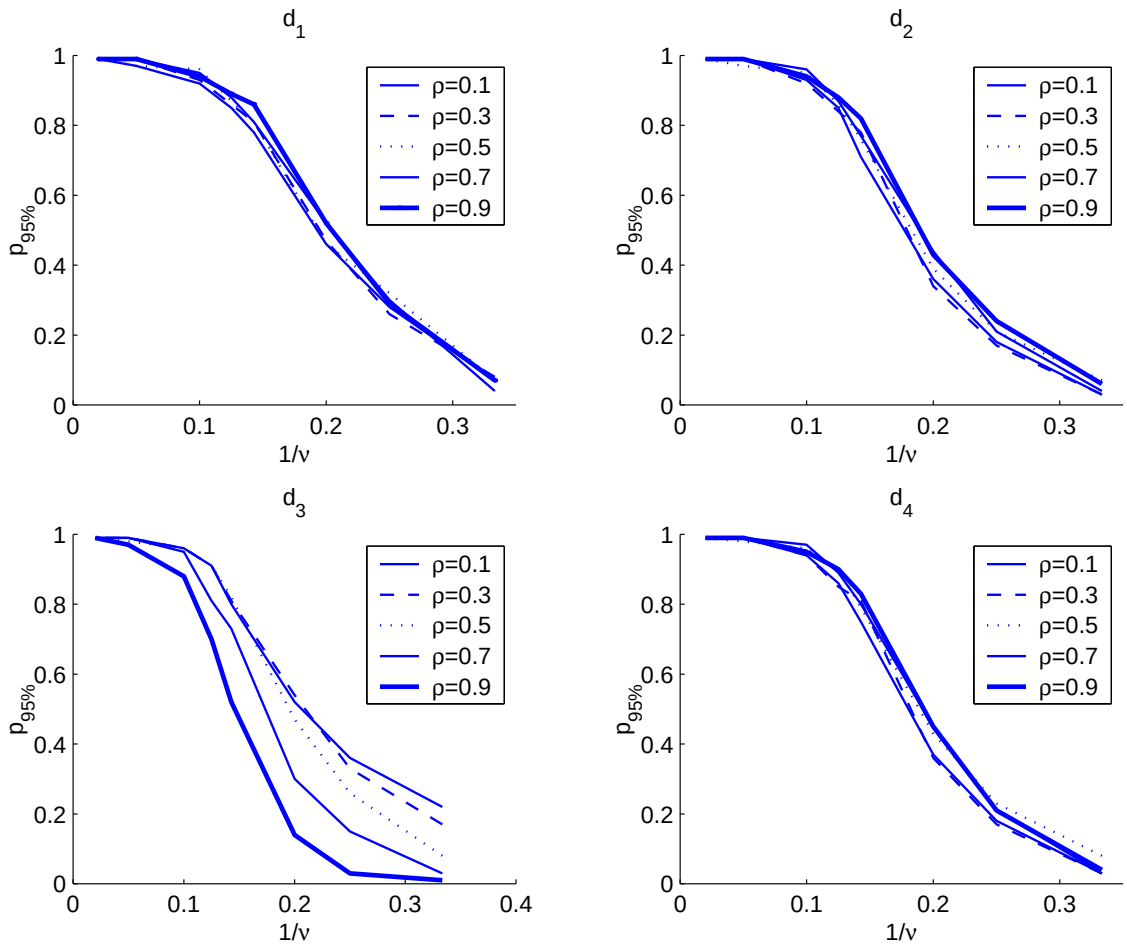

Figure 5: Graph of the minimum significance level $p_{95\%}$ necessary to distinguish the Gaussian copula hypothesis H_0 from the hypothesis of a student copula with ν degrees of freedom, as a function of $1/\nu$, for a given distance d_i and various correlation coefficients $\rho = 0.1, 0.3, 0.5, 0.7$ and 0.9.

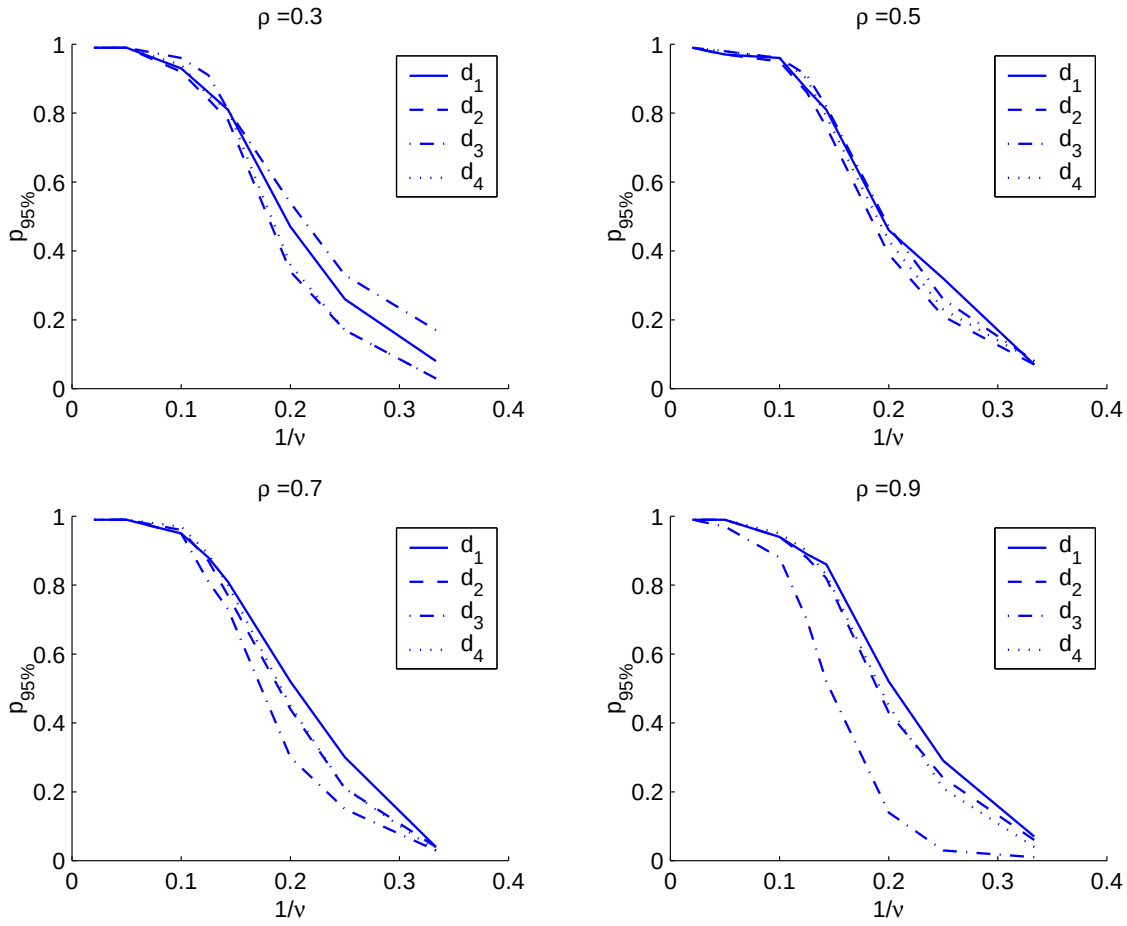

Figure 6: Same as figure 5 but comparing different distances for the same correlation coefficient ρ.

		$\hat{\rho}$	d_1	d_2	d_3	d_4
CHF	DEM	0.92	1.01%	0.67%	0.00%	0.72%
CHF	JPY	0.53	34.40%	27.10%	2.32%	28.30%
CHF	MYR	0.23	72.70%	87.10%	57.70%	92.60%
CHF	THA	0.21	3.08%	9.47%	3.31%	9.52%
CHF	UKP	0.69	0.28%	0.18%	0.06%	0.13%
DEM	JPY	0.54	2.26%	13.30%	10.00%	15.10%
DEM	MYR	0.26	42.50%	67.70%	62.20%	73.50%
DEM	THA	0.24	6.53%	13.50%	3.26%	13.20%
DEM	UKP	0.72	0.17%	0.04%	0.00%	0.04%
JPY	MYR	0.31	2.45%	6.34%	22.60%	6.86%
JPY	THA	0.34	0.00%	0.00%	3.24%	0.00%
JPY	UKP	0.41	2.85%	3.72%	5.22%	3.09%
MYR	THA	0.40	0.00%	0.00%	2.22%	0.00%
MYR	UKP	0.21	69.40%	79.40%	62.30%	83.10%
THA	UKP	0.15	52.20%	62.30%	3.21%	70.50%

Table 2: Each row gives the statistics of our test for each of the 15 pairs of currencies over a 10 years time interval from January 25, 1989 to December 31, 1998. The column $\hat{\rho}$ gives the empirical correlation coefficient for each pair determined as in section 3.1 and defined in (31). The columns d_1, d_2, d_3 and d_4 gives the probability to obtain, from the Gaussian hypothesis, a deviation between the distribution of the z^2 and the χ^2-distribution with two degrees of freedom larger than the observed one for the currency pair according to the distances d_1-d_4 defined by (33)-(36).

		$\hat{\rho}$	d_1	d_2	d_3	d_4
CHF	DEM	0.92	1.73%	1.33%	0.00%	1.31%
CHF	JPY	0.55	13.40%	14.90%	38.30%	14.10%
CHF	MYR	0.32	84.70%	70.00%	35.60%	74.00%
CHF	THA	0.17	44.00%	71.00%	3.53%	71.10%
CHF	UKP	0.79	0.31%	0.10%	0.00%	0.05%
DEM	JPY	0.56	2.46%	9.43%	16.30%	9.26%
DEM	MYR	0.35	93.20%	79.50%	35.10%	79.50%
DEM	THA	0.21	43.60%	87.70%	3.47%	87.40%
DEM	UKP	0.82	0.00%	0.00%	0.00%	0.00%
JPY	MYR	0.34	49.00%	54.90%	36.60%	59.40%
JPY	THA	0.27	38.90%	30.60%	3.37%	35.90%
JPY	UKP	0.53	0.09%	1.66%	6.72%	1.67%
MYR	THA	0.29	10.80%	8.71%	3.42%	9.30%
MYR	UKP	0.33	11.20%	28.60%	35.40%	34.50%
THA	UKP	0.21	43.40%	86.20%	3.13%	86.70%

Table 3: Same as table 2 for currencies over a 5 years time interval from January 25, 1989 to January 11, 1994.

		$\hat{\rho}$	d_1	d_2	d_3	d_4
CHF	DEM	0.92	31.50%	31.10%	0.05%	34.10%
CHF	JPY	0.52	58.40%	64.40%	1.98%	67.40%
CHF	MYR	0.16	71.10%	91.50%	88.30%	92.20%
CHF	THA	0.25	1.10%	3.87%	10.50%	3.34%
CHF	UKP	0.53	9.75%	10.30%	23.30%	9.29%
DEM	JPY	0.53	36.30%	54.00%	1.77%	65.40%
DEM	MYR	0.18	35.50%	50.00%	58.40%	56.70%
DEM	THA	0.28	1.28%	2.18%	10.80%	1.51%
DEM	UKP	0.56	11.50%	11.00%	30.20%	10.60%
JPY	MYR	0.29	7.63%	21.40%	6.67%	22.30%
JPY	THA	0.38	0.00%	0.02%	3.09%	0.02%
JPY	UKP	0.28	46.20%	23.00%	12.30%	20.70%
MYR	THA	0.44	0.05%	0.12%	5.34%	0.12%
MYR	UKP	0.11	59.40%	74.40%	69.50%	78.20%
THA	UKP	0.12	1.26%	7.66%	11.90%	6.51%

Table 4: Same as table 2 for currencies over a 5 years time interval from January 12, 1994 to December 31, 1998.

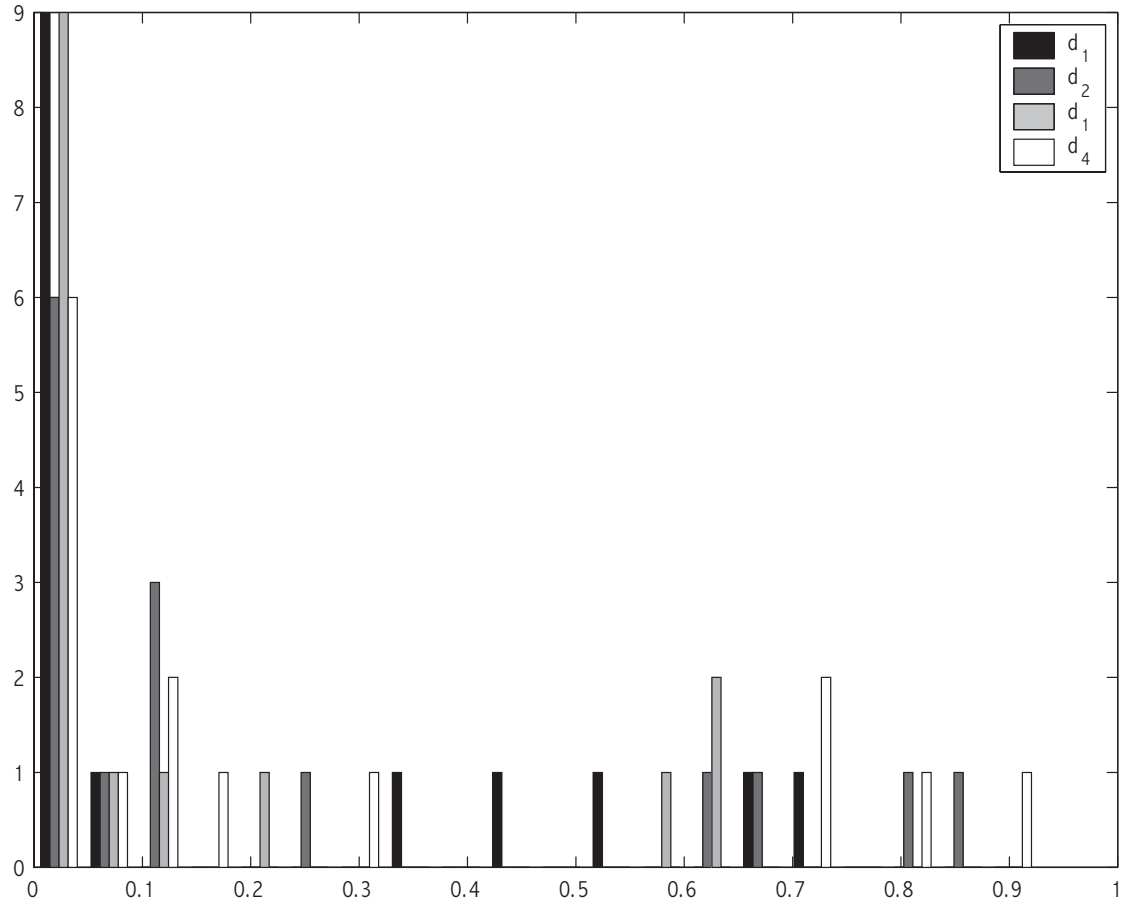

Figure 7: For each distance d_1-d_4 defined in equations (33)-(36), this figure shows the number of currency pairs that give a given p (shown on the abscissa) within a bin interval of width 0.05 for different currencies over a 10 years time interval from January 25, 1989 to December 31, 1998. p is the probability that pairs of Gaussian random variables with the same correlation coefficient ρ have a distance (between the distribution of z^2 and the theoretical chi^2 distribution) equal to or larger than the corresponding distance obtained for each currency pair. A clustering close to the origin signals a significant rejection of the Gaussian copula hypothesis.

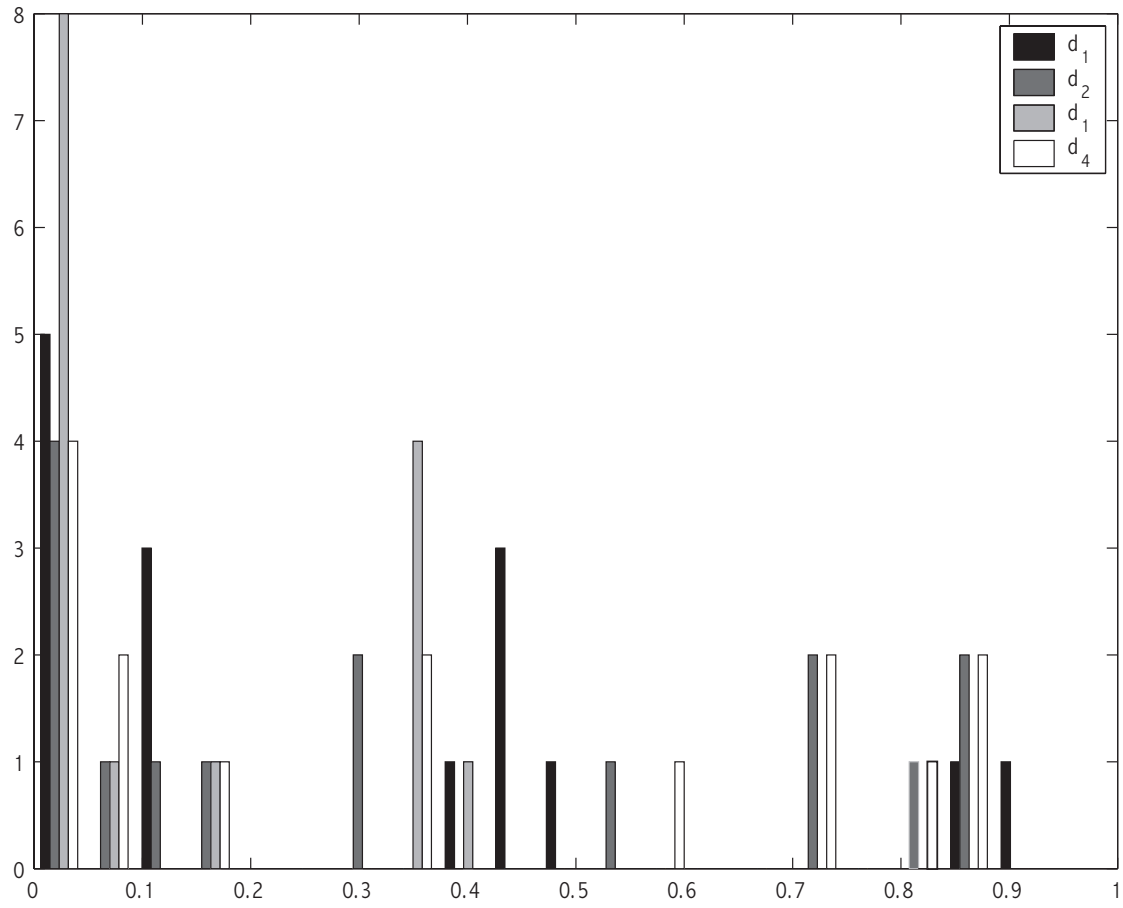

Figure 8: Same as figure 7 for currencies over a 5 years time interval from January 25, 1989 to January 11, 1994.

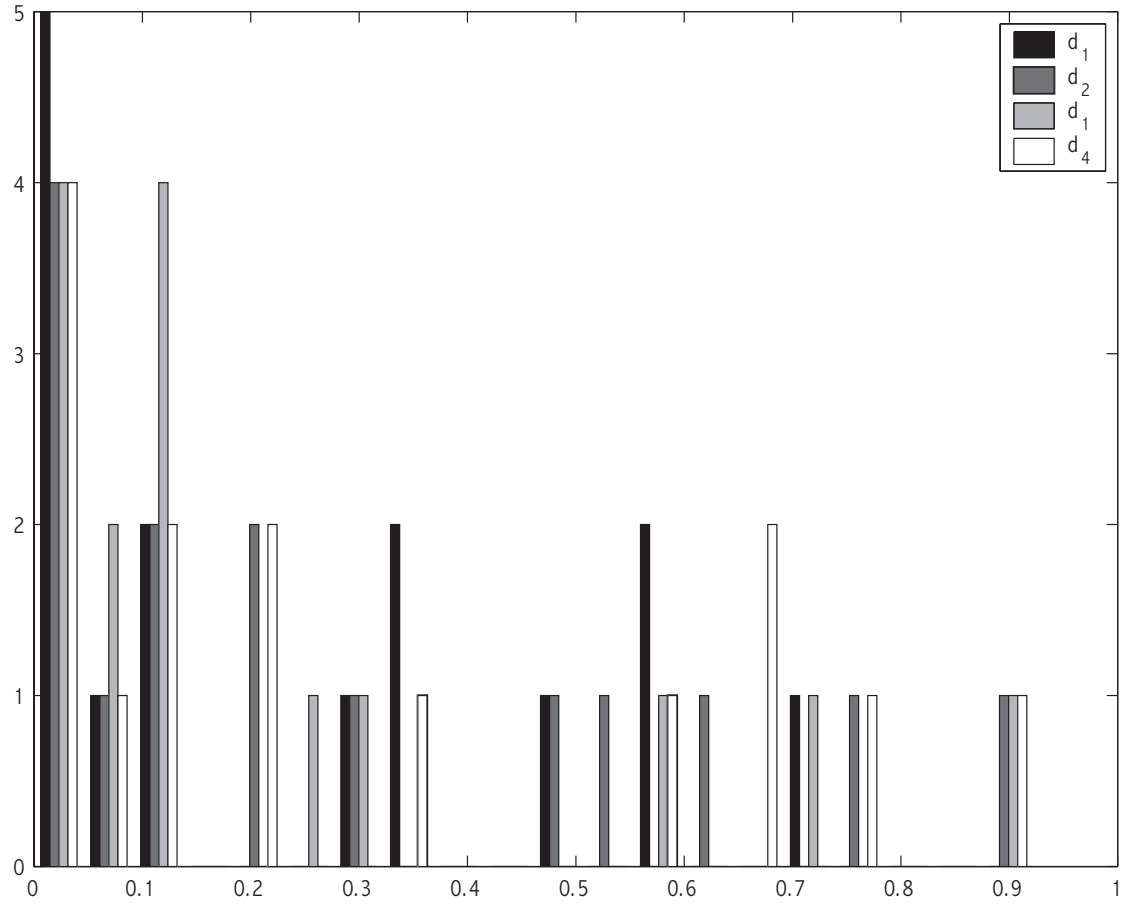

Figure 9: Same as figure 7 for currencies over a 5 years time interval from January 12, 1994 to December 1998.

		$\hat{\rho}$	d_1	d_2	d_3	d_4
aluminum	copper	0.46	6.46%	4.48%	1.45%	4.00%
aluminum	lead	0.35	11.40%	5.01%	17.00%	4.59%
aluminum	nickel	0.36	0.33%	0.51%	3.41%	0.62%
aluminum	tin	0.34	13.40%	13.80%	1.25%	15.90%
aluminum	zinc	0.36	0.23%	0.22%	6.21%	0.23%
copper	lead	0.35	4.71%	1.74%	17.90%	1.34%
copper	nickel	0.38	4.91%	4.60%	14.80%	3.80%
copper	tin	0.32	19.40%	13.50%	65.30%	14.70%
copper	zinc	0.40	3.24%	2.05%	17.50%	1.94%
lead	nickel	0.32	6.71%	3.78%	27.40%	3.62%
lead	tin	0.33	7.86%	4.04%	4.91%	3.31%
lead	zinc	0.42	0.02%	0.01%	4.59%	0.03%
nickel	tin	0.35	0.91%	0.92%	8.70%	0.76%
nickel	zinc	0.33	0.08%	0.34%	8.91%	0.35%
tin	zinc	0.31	0.53%	2.02%	10.30%	1.75%

Table 5: Same as table 2 for metals over a 9 years time interval from January 4, 1989 to December 30, 1997.

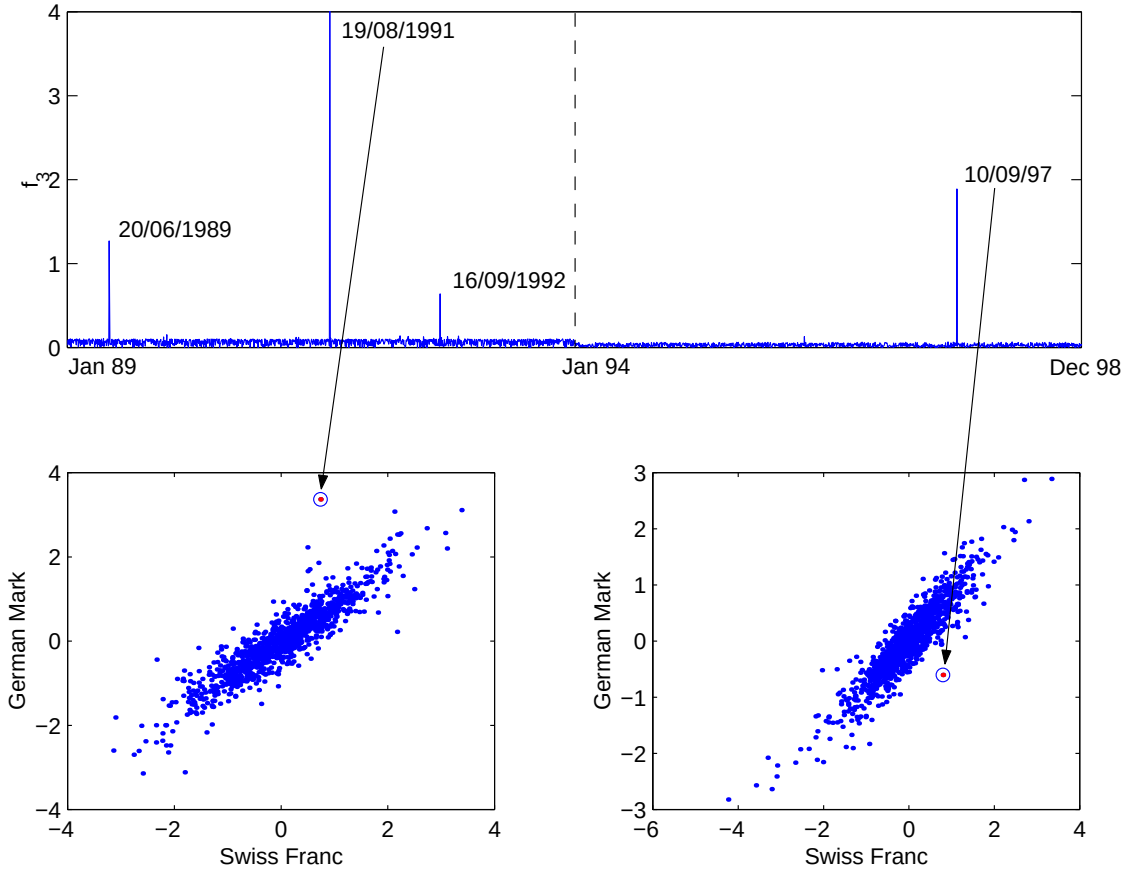

Figure 10: The upper panel represents the graph of the function $f_3(t)$ defined in (38) used in the definition of the distance d_3 for the couple Swiss Franc/German Mark as a function of time t, over the time intervals from January 25, 1989 to January 11, 1994 and from January 12, 1994 to December 31, 1998. The two lower panels represent the scatter plot of the return of the German Mark versus the return of the Swiss Franc during the two previous time periods. The circled dot, in each figure, shows the pair of returns responsible for the largest deviation of f_3 during the considered time interval.

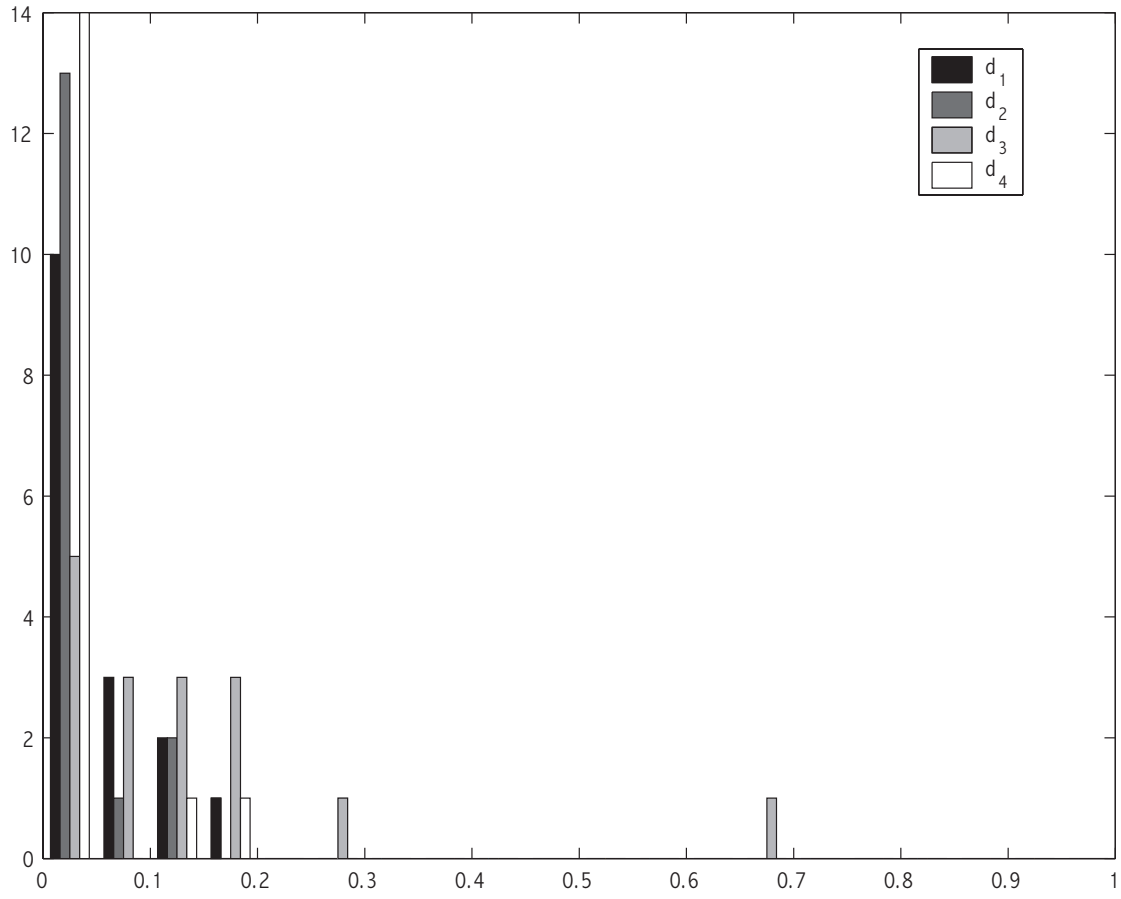

Figure 11: Same as figure 7 for metals over a 9 years time interval from January 4, 1989 to December 30, 1997.

		$\hat{\rho}$	d_1	d_2	d_3	d_4
amat	pfe	0.15	7.41%	11.20%	0.84%	11.40%
c	sunw	0.28	25.60%	48.70%	10.90%	53.90%
f	ge	0.33	25.20%	27.40%	11.50%	29.00%
gm	ibm	0.21	14.90%	38.50%	16.20%	41.80%
hwp	sbc	0.12	42.30%	16.90%	25.20%	17.20%
intc	mrk	0.17	24.80%	10.90%	64.60%	10.40%
ko	sunw	0.14	14.10%	10.10%	21.20%	9.35%
mdt	t	0.16	12.10%	28.10%	8.41%	29.80%
mrk	xom	0.19	15.40%	15.00%	11.20%	14.50%
msft	sunw	0.44	3.40%	1.85%	0.26%	1.74%
pfe	wmt	0.27	4.24%	4.12%	15.40%	3.74%
t	wcom	0.27	5.67%	8.02%	5.44%	9.07%
txn	wcom	0.28	47.90%	37.70%	15.20%	37.50%
wmt	xom	0.20	0.32%	0.00%	6.02%	0.00%

Table 6: Same as table 2 for stocks over a 10 years time interval from February 8, 1991 to December 29, 2000.

		$\hat{\rho}$	d_1	d_2	d_3	d_4
amat	pfe	0.10	58.30%	58.10%	11.80%	63.80%
c	sunw	0.23	46.60%	59.40%	43.40%	61.60%
f	ge	0.31	87.30%	78.70%	15.40%	84.80%
gm	ibm	0.21	60.00%	65.30%	10.30%	52.70%
hwp	sbc	0.11	87.30%	80.60%	28.40%	85.90%
intc	mrk	0.13	85.90%	82.10%	5.48%	86.50%
ko	sunw	0.20	35.30%	59.80%	45.10%	67.90%
mdt	t	0.14	90.90%	89.80%	16.80%	91.50%
mrk	xom	0.12	53.60%	62.10%	12.00%	61.80%
msft	sunw	0.40	26.80%	13.80%	16.00%	13.90%
pfe	wmt	0.23	29.40%	46.60%	14.10%	52.30%
t	wcom	0.19	79.20%	93.60%	4.95%	94.90%
txn	wcom	0.23	91.00%	98.30%	10.00%	99.30%
wmt	xom	0.22	71.60%	67.10%	7.35%	68.90%

Table 7: Same as table 2 for stocks over a 5 years time interval from February 8, 1991 to January 18, 1996.

		$\hat{\rho}$	d_1	d_2	d_3	d_4
amat	pfe	0.19	29.60%	33.90%	3.10%	39.50%
c	sunw	0.31	71.20%	65.80%	94.70%	70.80%
f	ge	0.34	38.00%	23.60%	32.20%	21.80%
gm	ibm	0.21	3.05%	17.90%	23.70%	21.90%
hwp	sbc	0.11	34.70%	61.30%	71.70%	64.00%
intc	mrk	0.20	13.10%	20.60%	55.70%	20.50%
ko	sunw	0.10	68.90%	34.40%	85.90%	35.20%
mdt	t	0.19	42.80%	61.10%	50.10%	57.90%
mrk	xom	0.23	35.70%	66.40%	11.30%	73.80%
msft	sunw	0.46	5.79%	7.60%	0.08%	8.07%
pfe	wmt	0.30	23.10%	21.20%	55.90%	19.80%
t	wcom	0.33	12.00%	13.70%	17.30%	14.00%
txn	wcom	0.31	56.30%	40.60%	46.40%	41.70%
wmt	xom	0.19	16.10%	5.38%	3.78%	4.94%

Table 8: Same as table 2 for stocks over a 5 years time interval from January 19, 1996 to December 29, 2000.

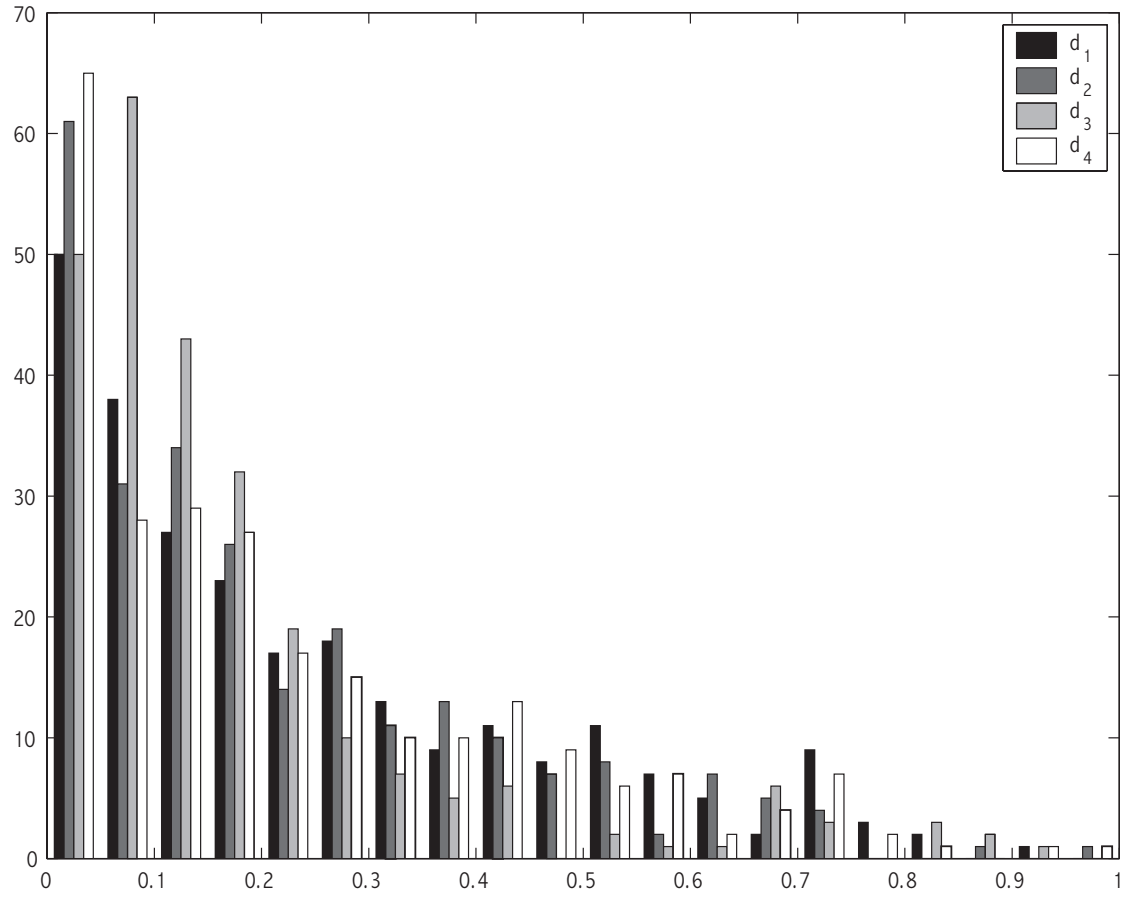

Figure 12: Same as figure 7 for stocks over a 10 years time interval from February 8, 1991 to December 29, 2000.

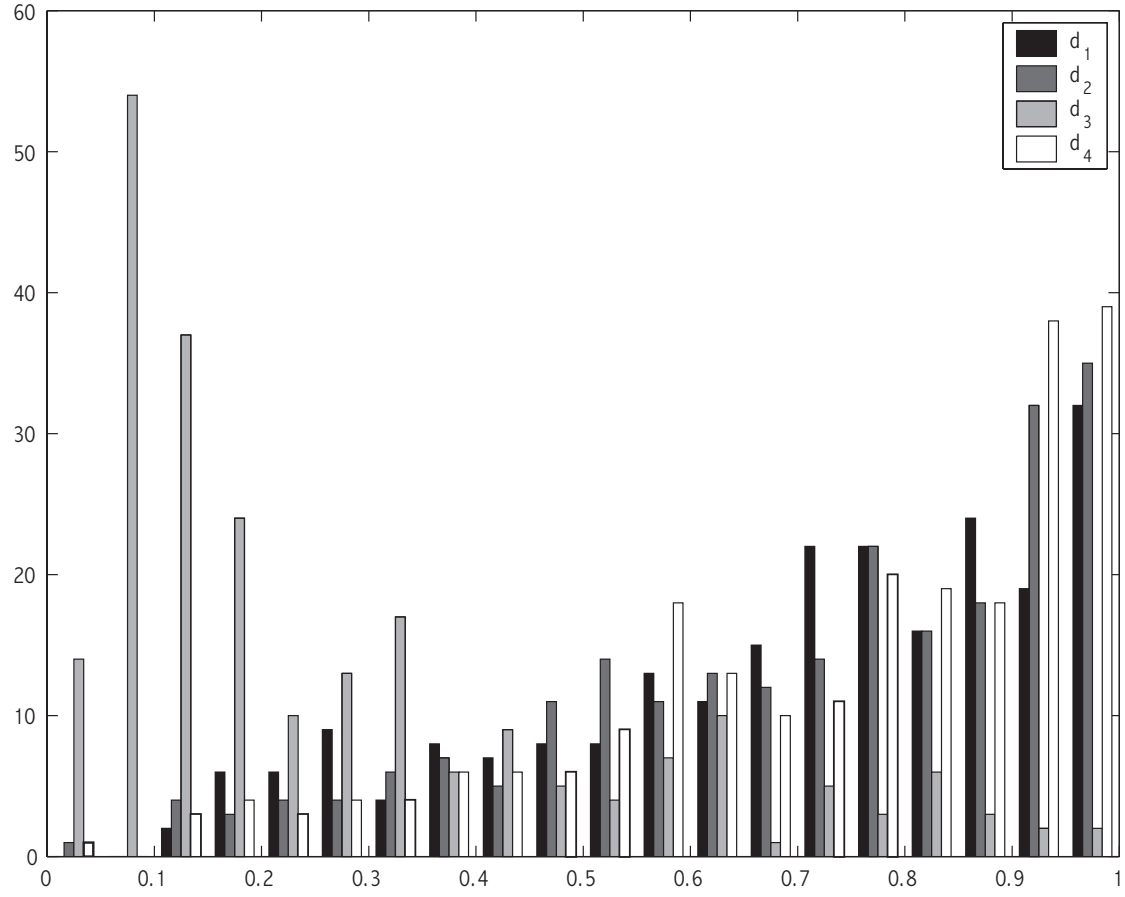

Figure 13: Same as figure 7 for stocks over a 5 years time interval from February 8, 1991 to January 18, 1996.

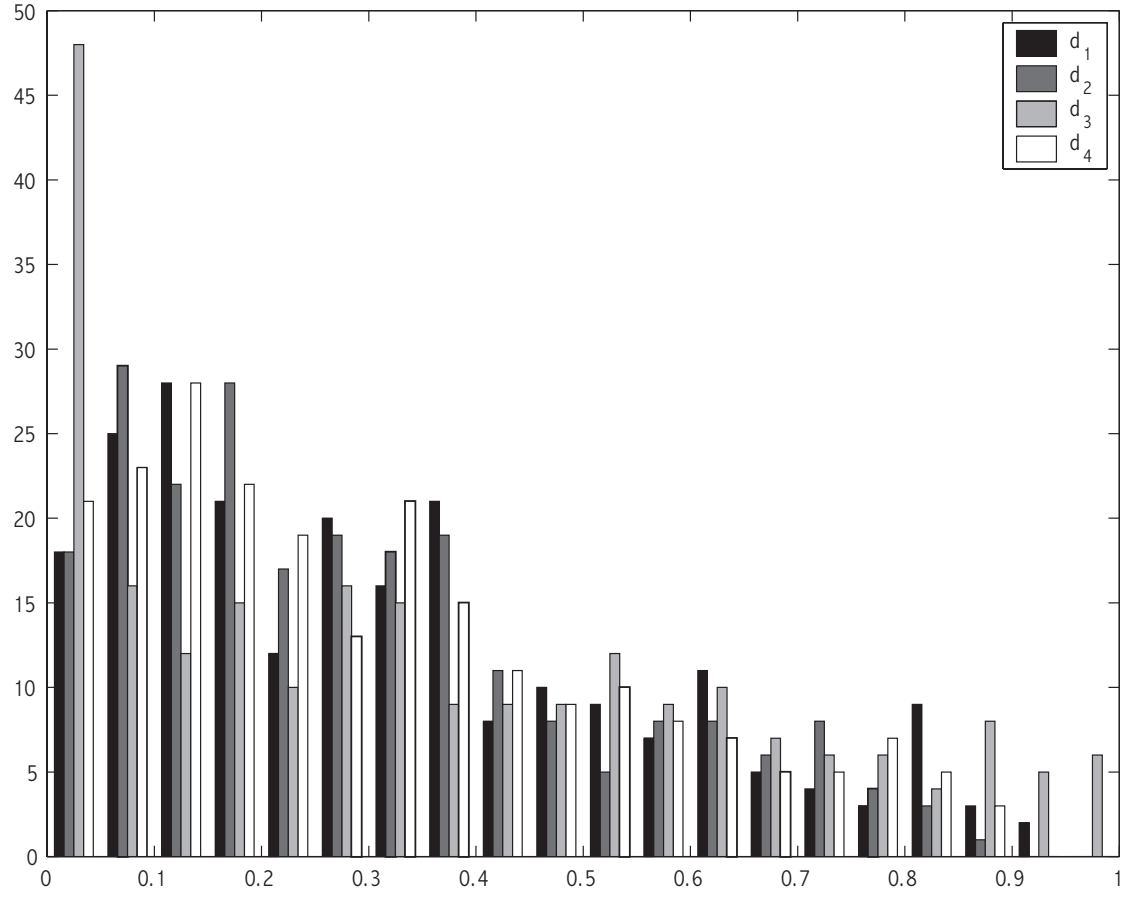

Figure 14: Same as figure 7 for stocks over a 5 years time interval from January 19, 1996 to December 30, 2000.

			$\hat{\rho}$	d_1	d_2	d_3	d_4
	hwp	ibm	0.34	33.60%	22.60%	33.30%	23.50%
	hwp	intc	0.46	30.10%	47.30%	51.20%	52.10%
	hwp	msft	0.41	76.30%	47.20%	32.30%	45.30%
	hwp	sunw	0.40	29.60%	29.80%	76.60%	35.40%
Time interval from	ibm	intc	0.30	48.10%	35.40%	4.18%	33.40%
February 8, 1991 to	ibm	msft	0.24	39.30%	66.10%	58.80%	70.70%
January 18, 1996	ibm	sunw	0.29	96.50%	97.10%	34.60%	98.60%
	intc	msft	0.47	25.90%	14.50%	4.50%	15.30%
	intc	sunw	0.40	48.10%	38.60%	4.47%	39.50%
	msft	sunw	0.40	26.80%	13.80%	16.60%	13.90%

			$\hat{\rho}$	d_1	d_2	d_3	d_4
	hwp	ibm	0.46	2.02%	3.21%	0.96%	3.96%
	hwp	intc	0.44	2.88%	4.89%	0.06%	5.80%
	hwp	msft	0.37	5.23%	9.88%	33.60%	11.80%
	hwp	sunw	0.45	56.60%	56.50%	10.80%	62.30%
Time interval from	ibm	intc	0.43	5.34%	3.31%	1.68%	2.44%
January 19, 1996 to	ibm	msft	0.39	1.00%	0.95%	2.28%	0.88%
December 29, 2000	ibm	sunw	0.46	23.50%	15.60%	33.80%	14.90%
	intc	msft	0.57	31.80%	16.10%	11.50%	17.10%
	intc	sunw	0.50	6.68%	3.55%	0.01%	4.37%
	msft	sunw	0.46	5.79%	7.60%	0.08%	8.07%

Table 9: Same as table 2 for stocks belonging to the informatic sector, over two time intervals of 5 years.