Conditional Dependency of Financial Series: An Application of Copulas

Michael Rockinger[a,*] and Eric Jondeau[b]

February 2001

Abstract

We develop a new methodology that measures conditional dependency. We achieve this by using copula functions that link marginal distributions, here chosen to obey a GARCH-type model with time-varying skewness and kurtosis. We apply this model to daily returns of stock-market indices. We find strong evidence of persistence in dependency both for local currency and $ US denominated series. For European stock markets, we also find evidence that large simultaneous returns of either sign lead to higher subsequent dependency.

We show that dependency changes through time, as well. For stock markets within Europe, dependency increased whereas it decreased since the mid 90s when involving the S&P 500 or the Nikkei.

We also suggest extensions for conditional asset pricing models involving time variation of co-skewness and co-kurtosis.

Keywords: International correlation, Market integration, ARCH, Stock indices, Exchange rates

JEL classification: C51, F37, G11

[a] HEC - School of Management, Department of Finance, 1, rue de la Libération, 78351 Jouy-en-Josas

[b] Banque de France, DEER, 41-1391 Centre de recherche, 31, rue Croix des Petits Champs, 75049 Paris

*Corresponding author. Tel: +33/(0)1 39 67 72 59, e-mail: rockinger@hec.fr. Michael Rockinger is also CEPR, and scientific consultant at the Banque de France. He acknowledges help from the HEC Foundation and the European Community TMR Grant: "Financial Market Efficiency and Economic Efficiency." We are grateful to Sanvi Avouyi-Dovi, Christian Gourieroux, and Christian Pfister for precious comments. The numerical implementation used the Fortan optimization program SNOPT which was kindly provided by Philip Gill. This paper was started during a sabbatical of the first author at UCSD. The usual disclaimer applies. The Banque de France does not necessarily endorse the views expressed in this paper.

Conditional Dependency of Financial Series: An Application of Copulas

Abstract

We develop a new methodology that measures conditional dependency. We achieve this by using copula functions that link marginal distributions, here chosen to obey a GARCH-type model with time-varying skewness and kurtosis. We apply this model to daily returns of stock-market indices. We find strong evidence of persistence in dependency both for local currency and $ US denominated series. For European stock markets, we also find evidence that large simultaneous returns of either sign lead to higher subsequent dependency.

We show that dependency changes through time, as well. For stock markets within Europe, dependency increased whereas it decreased since the mid 90s when involving the S&P 500 or the Nikkei.

We also suggest extensions for conditional asset pricing models involving time variation of co-skewness and co-kurtosis.

1 Introduction

The aim of this paper is to provide a new methodology to describe the multivariate conditional distribution of returns in the presence of non-normal innovations. On the one hand, our methodology builds on so-called "copulas," i.e. functions connecting marginal distributions. On the other hand, we draw on recent advances in the modelization of conditional returns that allow for time-varying second, third, and fourth moments.

Copula functions are well-studied objects in the statistical literature. At textbook level, one may mention the contributions of Joe (1997) or Nelsen (1999). These functions have been introduced to model a joint (multivariate) distribution when only marginal distributions are known. This type of link between marginal distributions is particularly useful in situations where multivariate normality does not hold. In financial applications, it is well established that univariate distributions are fat-tailed. In such situations, the statistics literature provides only little guidance about what type of multivariate distribution to chose, i.e. Kotz, Balakishnan, and Johnson (2000). In a recent application to finance, Embrechts, McNeil, and Strautman (1999) have used copulas to model extreme values.

Whereas the copula is a rather recent concept, much progress has been made in the modeling of univariate series with conditional higher moments. For instance, in the spirit of Engle's (1982) ARCH and Bollerslev's (1986) GARCH model, Hansen (1994) as well as Harvey and Siddique (1999) developed models where the fat-tailed conditional density has parameters which vary deterministically over time. Whereas Hansen's model builds on a generalized student-t distribution, Harvey and Siddique's relies on a non-central gamma density. As shown by Jondeau and Rockinger (2000), some difficulties related to the numerical estimation of Hansen's model may be overcome with the use of a sophisticated numerical estimation technique. Premaratne and Bera (2000) achieve time variation in skewness and kurtosis by using the Pearsonian densities. Rockinger and Jondeau (2001) use an entropy density. An other recent contribution in this field is by Engle and Manganelli (1999), who focus on the modelization of large realizations using quantile regressions. The modelization of conditional marginal distributions appears, therefore, as being rather well understood.

In this work, we build on Hansen's (1994) generalized student-t univariate model to obtain conditional marginal distributions. Then, we use Plackett's copula function to link these marginal distributions. Given that the copula function introduces an explicit parameter that may be interpreted intuitively as a correlation, it is easy to render this parameter conditional. In other words, our model allows for conditionally dependent marginal distributions. According to the way in which the dependency parameter arises, we can also introduce asymmetries in the model, in the spirit of Engle and Ng (1993), Glosten, Jagannathan, and Runkle (1993), Gourieroux and Monfort (1992), as well as Zakoïan (1994). Our model thus provides an alternative to multivariate GARCH models. It presents the advantage of allowing for fat-tailed innovations.

We apply our model to daily returns of stock indices. Our aim is to contribute to the debate on the relationship between international correlation and stock-market turbulence. There are several studies of this phenomenon, which is of great importance for portfolio choice applications. Empirical evidence is often contradictory. For instance, Kaplanis (1988) uses monthly data of various markets and cannot reject the assumption of the constancy of the correlation matrix. Ratner (1992) confirms this result. Koch and Koch (1991) use daily data and find that correlation increases through time. King, Sentana, and Wadhwani (1994) object that this result is due to the 1987 stock-market crash. Another strand of the literature is based on the ARCH framework.[1] For instance, using multivariate GARCH models, Hamao, Masulis, and Ng (1990) and Susmel and Engle (1994) measure the interdependence of returns and volatilities across major stock markets. Longin and Solnik (1995) use monthly data and a multivariate GARCH model, in which they condition correlation on the S&P return exceeding a given threshold. They obtain an increase in correlation in S&P turbulent periods. However, most of these results are subject to the Boyer, Gibson, and Loretan (1997) critique.[2] These authors show that the correlation between two series appears much higher if the latter are conditioned on one of the series exceeding a threshold. As a consequence, studies, in which correlation is computed on a subsample where one of the series exceeds a given level, will find an artificially high correlation.

More recently, tests of a constant correlation in a GARCH context have been proposed by Bera and Kim (1996) and Tse (2000). Ramchand and Susmel (1998) and Ang and Bekaert (1999) estimate a multivariate Markov-switching model and test the hypothesis of a constant international conditional correlation between stock markets. Ball and Torous (2000), using a model with stochastic correlation, find that the correlation does not significantly increase as markets become more turbulent. Rigobon (1999) develops a model that tests whether dependency changed over a given time gap. In Longin and Solnik (2000), dependency for monthly data is investigated with extreme value theory. This theory, however, focuses only on the tails and neglects the central part of the distribution. A model capturing the tail behavior as well as the more central part seems, therefore, to be of value.

[1]See Hamao, Masulis, and Ng (1990), Susmel and Engle (1994), Longin and Solnik (1995), Bekaert and Harvey (1995), Karolyi (1995), Karolyi and Stutz (1996), Bera and Kim (1996), Tse (2000).

[2]See also Forbes and Rigobon (1999).

In the empirical part of this paper, we show that dependency is conditional and that there is persistence in dependency for daily stock index returns of five important markets. As a consequence of large joint stock market movements, dependency increases.

We also provide evidence that dependency evolves through time. For European countries we find a symmetric increase of dependency whereas dependencies involving the S&P500 or the Nikkei are decreasing since the mid 90s.

Our model may also be used for multivariate Value-at-Risk considerations and the testing of asset pricing models involving conditional higher moments such as conditional co-skewness and co-kurtosis. The computation of co-skewness and co-kurtosis is possible in our framework since we focus not only on the tails of the distribution but also on the entire distribution.

In the next section, we introduce copula functions and derive one particularly useful copula: Plackett's. In section 3, we first introduce our univariate model allowing for time-varying volatility, skewness, and kurtosis, and then show how to link the univariate models with copulas. In section 4, we describe the data and discuss our results. Section 5 contains a conclusion and ideas for further research.

2 Copula distribution functions

2.1 Generalities

As mentioned by Nelsen (1999, p. 1), the study of copulas is a recent phenomenon in statistics. Hence, it is not astonishing that copulas have not yet found their way into empirical finance. In order to understand the usefulness of copulas, consider two random variables X and Y with marginal distributions $F(x) = \Pr[X < x]$ and $G(y) = \Pr[Y < y]$. In this paper we only consider situations where all cumulative distribution functions (cdf) are continuous. The random variables may also have joint distribution function $H(x, y) = \Pr[X < x, Y < y]$. All the distribution functions, $F(\cdot)$, $G(\cdot)$ and $H(\cdot, \cdot)$ belong to the interval $[0, 1]$. In some cases, a multivariate distribution exists, so that the function H has an explicit expression. One such case is the multivariate normal distribution. In many cases, however, a description of $F(\cdot)$ and $G(\cdot)$ is relatively easy to obtain, whereas an explicit expression of the joint distribution $H(\cdot, \cdot)$ may be difficult to obtain. This is where copulas come in handy.

We now will define copulas more formally. In the next section we construct a useful copula for finance applications. We would like to emphasize from the onset that many results developed in this paper extend to a tri-variate or even higher dimensional framework. Some of the results, however, only hold for the bi-variate framework.

Definition 1 A two-dimensional copula is a function $C : [0,1]^2 \rightarrow [0,1]$ having three properties:

1. $C(u,v)$ is increasing in u and v.
2. $C(0,v) = C(u,0) = 0$, $C(1,v) = v$, $C(u,1) = u$.
3. $\forall u_1, u_2, v_1, v_2$ in $[0,1]$ such that $u_1 < u_2$ and $v_1 < v_2$ we have $C(u_2,v_2) - C(u_2,v_1) - C(u_1,v_2) + C(u_1,v_1) \geq 0$.

Property 1 states that when one marginal distribution is constant, the joint probability will increase provided that the other marginal distribution increases.

Property 2 reveals conditions one would expect for a joint distribution. That is, if one margin has zero probability the joint occurrence also has zero probability to occur. Consequently, if on the contrary one margin is certain to occur, then the probability of a joint occurrence is determined by the remaining margin probability.

Property 3 indicates that if u and v both increase then the joint probability also increases. This property is therefore a multivariate extension of the condition that a cdf is increasing.

Furthermore, if we set $u = F(x)$ and $v = G(y)$, then $C(F(x), G(y))$ yields a description of the joint distribution of x and y. Having obtained this intuitive definition, further properties may be obtained.

Proposition 1 *If u and v are independent, then $C(u,v) = uv$.*

Proof. The proof of this property follows immediately from the definition of independent random variables. ■

Proposition 2 *(Sklar's Theorem). Let H be a joint distribution function with margins F and G. Then, there exists a copula C such that for all real numbers x, y*

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

Furthermore, if F and G are continuous, then C is unique. Conversely, if F and G are distributions, then the function H defined by equation (1) is a joint distribution function with margins F and G.

Proof. The proof of the theorem first appeared in Sklar (1959). A relatively simple proof may be found in Schweizer and Sklar (1974). ■

This theorem justifies the importance of copulas for empirical research. We now show how to obtain a copula that is relevant for finance.

2.2 Construction of a useful Copula

In finance applications, it is necessary to express a positive and negative dependence between variables. We are now going to construct a copula allowing marginal distributions to be either positively or negatively dependent. This construction will yield the so-called Plackett's copula. Plackett's copula is a way of joining two marginal distributions. For higher dimensional problems, other types of copula or combinations thereof could be used.[3]

Consider Figure 1, where we assume for the moment that we have two random variables X and Y. Both variables may take two discrete states, say high and low. As indicated in the figure, we associate probabilities a, b, c, d to the various simultaneous realizations. Intuitively, if along the 45^o diagonal the probabilities are high, then we would have a positive dependence situation. In fact if one state is high, the other state will be high as well. If along the a, b diagonal there are as many observations as along the c, d diagonal, then the random variables may be considered independent.

These observations suggest

$$\theta = \frac{ab}{cd}$$

as a natural measure of dependency. If $\theta = 1$ there will be independence, if $\theta < 1$ dependence will be negative, and if $\theta > 1$ dependence will be positive. Plackett (1965) had the idea to associate with the states L the marginal cdf $F(x)$ and $G(y)$ in $[0, 1]$. Also, the probabilities a, b, c, d may be expressed in terms of the copula function. If F and G still represent the marginal distributions and H the joint distribution, we obtain

$$\theta = \frac{H(x,y)[1 - F(x) - G(y) + H(x,y)]}{[F(x) - H(x,y)][G(y) - H(x,y)]}.$$

In general, one may expect that, for a given joint distribution, θ will be a function of x and y. There may be situations where H does not exist. In such cases one wishes to create a function C playing the role of H. In other words, one may ask if there exists a function C from $[0, 1]^2$ into $[0, 1]$, having as arguments $u = F(x)$ and $v = G(y)$, for which θ does not depend on x and y.

Tedious but straightforward computations show that the object

$$C_\theta(u,v) = \begin{cases} \frac{1}{2(\theta-1)} \left[1 + (\theta-1)(u+v) - \sqrt{[1 + (\theta-1)(u+v)]^2 - 4uv\theta(\theta-1)}\right] & \text{if } \theta \neq 1, \\ uv & \text{if } \theta = 1, \end{cases}$$

defined for $\theta > 0$, satisfies the three conditions that define a copula function. In this case the function $C_\theta(F(x), G(y))$ is the joint cdf of x and y.

[3]Many other copulas exist. We decided to focus on Plackett's because of its intuitive appeal and because many of its properties have already been established.

2.3 Link with traditional dependence measures

In economic applications, and in mean-variance portfolio analysis in particular, a widely used measure of dependence is given by Pearson's correlation coefficient, R.

We recall that Pearson's R is defined for a set of random variables X and Y with joint distribution $H(x,y)$ and marginal ones, $F(x)$ and $G(y)$, as

$$R = \frac{\mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y]}{\sqrt{\mathrm{Var}[X]\mathrm{Var}[Y]}} = \frac{\int xydH(x,y) - \int xdF(x)\int ydG(y)}{\sqrt{\mathrm{Var}[X]\mathrm{Var}[Y]}}$$

where

$$\mathrm{Var}[X] = \int (x - \mathrm{E}[x])^2 dF(x)$$

and similarly for Y. Another measure of association is given by Spearman's ρ. It is defined as the correlation between the marginal distributions. Setting $u = F(x)$ and $v = G(y)$, we obtain

$$\rho_\theta = \frac{\mathrm{E}[uv] - \mathrm{E}[u]\mathrm{E}[v]}{\sqrt{\mathrm{Var}[u]\mathrm{Var}[v]}}.$$

It may be shown that

$$\rho_\theta = 12\int_{[0,1]^2} uvdC(u,v) - 3 = 12\int_{[0,1]^2} C(u,v)dudv - 3.$$

This observation is very useful, because it allows the verification of the estimation procedure. Indeed, the dependency parameter θ must be close to Pearson's correlation of the margins (u,v). For Plackett's copula, an explicit formula for the correlation is available, presented in the following proposition.

Proposition 3 *For Plackett's copula, C_θ, Spearman's ρ_θ is given by*

$$\rho_\theta = \begin{cases} \frac{\theta+1}{\theta-1} - \frac{2\theta}{(\theta-1)^2}\ln(\theta), & \text{if } \theta \neq 1 \\ 0 & \text{if } \theta = 1. \end{cases} \tag{2}$$

Proof. See Nelsen (1997), p. 138. ■

Proposition 4 $\rho_{1/\theta} = -\rho_\theta$.

Proof. The proof follows by replacing θ in (2) by $1/\theta$ and rearranging the expression. The meaning of this result is that one may easily convert a given positive dependency into a negative dependency. ■

Last but not least, it is to be observed that θ is only defined for positive values. In numerical applications, this restriction may be easily implemented by using a logarithmic transform of θ. In this case, independency corresponds to a value of $\ln(\theta) = 0$. If $\ln(\theta)$ is positive, then we have positive dependency and symmetrically for a negative value.

2.4 Simulation

In many finance applications, such as Value-at-Risk, where models get tested using simulations, it is necessary to simulate data following a multivariate distribution, given here by the copula function. One of the techniques used to simulate data distributed as a Plackett's copula is based on Johnson (1986). It consists in simulating an observation from one of the margins and then simulating the other margin using the copula. This yields the following algorithm:

1. simulate two independent realizations, i.e. u and t, distributed uniformly on the interval [0,1].
2. define $a = t(1-t)$, $b = \sqrt{\theta}\sqrt{\theta + 4au(1-u)(1-\theta)^2}$
3. compute $v = [2a(u\theta^2 + 1 - u) + \theta(1-2a) - (1-2t)b]/[2\theta + a(\theta-1)^2]$
4. u and v will be distributed according to a Plackett's distribution.

Once margin realizations have been simulated, it is also possible to construct trajectories for x_t and y_t.

2.5 Estimation of the model

In practical bi-variate situations, one observes a sample (x_t, y_t), $t = 1, \cdots, T$. It is assumed that x_t gets generated by a continuous marginal distribution $F(\cdot, w_x)$, where w_x represents a vector of parameters. Similarly, y_t is generated by a continuous distribution, $G(\cdot, w_y)$, where w_y is a parameter vector. For instance, F could represent a GARCH model describing the marginal distribution of x_t.

For convenience, we now define $u_t \equiv F(x_t, w_x)$ and $v_t \equiv G(y_t, w_y)$.[4] We notice that the dependency parameter θ appears explicitly in the copula function. This means that θ can be easily conditioned. We define w_θ as the vector of parameters involved in the modeling of the dependency parameter. In this case, we may write

$$\theta_{t|t-1} = \Theta(u_{t-1}, v_{t-1}; w_\theta).$$

It is easy to establish the density of a Plackett's copula as

$$c_\theta(u,v) \equiv \frac{\partial^2 C_\theta(u,v)}{\partial u \partial v} = \frac{\theta[1 + (u - 2uv + v)(\theta - 1)]}{\left([1 + (\theta-1)(u+v)]^2 - 4uv\theta(\theta-1)\right)^{\frac{3}{2}}}.$$

In Figures 2 and 3, we display copula functions for the case of positive dependency, e.g. $\theta = 5$, and of corresponding negative dependency, that is $\theta = 1/5$. If $\theta = 1$, then the density turns out to be a flat surface corresponding to independence.

[4]In this section, we tried to keep the notation as simple as possible.

Also, writing f and g as the marginal densities, the joint density of an observation (x_t, y_t) is

$$l(x_t, y_t; w_x, w_y, w_\theta) = c_{\Theta(F(x_{t-1},w_x),G(y_{t-1},w_y);w_\theta)}(F(x_t, w_x), G(y_t, w_y))f(x_t, w_x)g(y_t, w_y).$$

As a consequence, the log-likelihood of a sample becomes

$$\begin{aligned}\mathcal{L}((x_t, y_t), t = 1, \cdots, T; w_x, w_y, w_\theta) = \quad & \sum_{t=1}^{T} \left(\ln \left[c_{\Theta(F(x_{t-1},w_x),G(y_{t-1},w_y);w_\theta)}(F(x_t, w_x), G(y_t, w_y))\right]\right. \\ & \left. + \ln\left[f(x_t, w_x)\right] + \ln\left[g(y_t, w_y)\right]\right). \qquad (3)\end{aligned}$$

Ideally, one would like to maximize the likelihood simultaneously over all the parameters, yielding the parameter estimates written as $\hat{w}_x, \hat{w}_y, \hat{w}_\theta$. In practical applications, this estimation may be difficult.

First, the dimension of the problem can be large. In such a case, it may be necessary to help the estimation by providing starting values obtained from the marginal estimations

$$\tilde{w}_x \in \operatorname{argmax} \sum_{t=1}^{T} \ln[f(x_t, w_x)], \qquad (4)$$

$$\tilde{w}_y \in \operatorname{argmax} \sum_{t=1}^{T} \ln[g(y_t, w_y)]. \qquad (5)$$

Secondly, the dependency parameter of the copula function may be a convoluted expression of the parameters. In such a case, an analytical expression of the gradient of the likelihood might not exist.[5] Therefore, only numerical gradients may be computable with the associated slowing down of the numerical procedure.

For complicated situations, it is therefore recommended to use the set $(\tilde{w}_x, \tilde{w}_y, \tilde{w}_\theta)$ obtained by estimating in a first step (4) and (5) before solving for

$$\tilde{w}_\theta \in \operatorname{argmax} \mathcal{L}((x_t, y_t), t = 1, \cdots, T; \tilde{w}_x, \tilde{w}_y, w_\theta).$$

2.6 Computing co-moments

In several applications such as conditional asset pricing models, it is necessary to compute expressions such as co-skewness or co-kurtosis. Such expressions will typically involve moments

[5] Or it may, at least, be a very complicated expression. One such example is provided by our model where, as shown later and in an appendix, the computation of the gradient involves the computation of the derivative of a student-t cdf with respect to its parameter. This means that one would need to know the integral of the derivative of the student-t distribution.

of the form

$$m_{i,j} = \int_{x\in\mathcal{R}}\int_{y\in\mathcal{R}} x^i y^j\, c_\theta(F(x), G(y)) f(x) g(y) dx\, dy.$$

Such integrals may be efficiently evaluated using a change in variables $u = F(x)$, $v = G(y)$. With this change we get

$$m_{i,j} = \int_{u\in[0,1]}\int_{y\in[0,1]} (F^{-1}(u))^i (G^{-1}(v))^j\, c_\theta(u, v) du\, dv.$$

For numerical purposes, under the assumption that both F and G are generated by a generalized student-t, F^{-1} and G^{-1} are elementary expressions involving standard functions.[6] We leave the application of this technique to further research.

In VaR applications, it is necessary to compute the probability that a portfolio exceeds a certain threshold. Again, once the marginal models are known, the exceedance probability may be numerically computed as a simple integration using the fact that if the pair (X, Y) has some joint distribution function $C(F(x), G(y))$ then

$$\Pr[\delta X + (1-\delta)Y > \gamma] = \int_{\delta x+(1-\delta)y>\gamma} dC(F(x), G(y)).$$

Again, this expression is rather easy to implement numerically.

3 A model for the marginal distributions

Our margin model builds on the GARCH model of Engle (1982) and Bollerslev (1986).[7] It is well known that the residuals obtained for a GARCH model are non-normal. This observation has led to the introduction of fat-tailed distributions for innovations. Nelson (1991) considers the generalized error distribution. Bollerslev and Wooldridge (1992) consider the case of a student-t distribution.[8] Engle and Gonzalez-Rivera (1991) model residuals non-parametrically. Even though these contributions recognize the fact that errors have fat tails, they do not render them time-varying, i.e., the parameters of the error distribution are assumed to be constant over time.

[6] For instance, under Fortran, the IMSL routine `TIN` may be used to compute the inverse of the student-t distribution.

[7] The literature concerning GARCH models is huge. Several reviews of the literature are available, e.g., Bera and Higgins (1993), Bollerslev, Chou, and Kroner (1992), as well as Bollerslev, Engle, and Nelson (1994).

[8] For a definition of the traditional student-t distribution, see, for instance, Mood, Graybill, and Boes (1982).

3.1 Hansen's generalized student-t density

Hansen (1994) is the first to propose a model where the first four moments are conditional and, therefore, time varying. He achieves this by introducing a generalization of the student-t distribution. This generalization allows for asymmetries of large return realizations while maintaining the assumption of a zero mean and unit variance. The conditioning is obtained by defining parameters as functions of past realizations.

Hansen's generalized student-t distribution is defined by

$$d(z|\eta,\lambda) = \begin{cases} bc\left(1+\frac{1}{\eta-2}\left(\frac{bz+a}{1-\lambda}\right)^2\right)^{-\frac{\eta+1}{2}} & \text{if } z < -a/b, \\ bc\left(1+\frac{1}{\eta-2}\left(\frac{bz+a}{1+\lambda}\right)^2\right)^{-\frac{\eta+1}{2}} & \text{if } z \geq -a/b \end{cases} \tag{6}$$

where

$$a \equiv 4\lambda c\frac{\eta-2}{\eta-1}, \qquad b^2 \equiv 1+3\lambda^2-a^2, \qquad c \equiv \frac{\Gamma\left(\frac{\eta+1}{2}\right)}{\sqrt{\pi(\eta-2)}\Gamma\left(\frac{\eta}{2}\right)}.$$

If a random variable Z has the density $d(z|\eta,\lambda)$, we will write $Z \sim D(z|\eta,\lambda)$. Inspection of the various formulas reveals that this density is defined for $2 < \eta < \infty$ and $-1 < \lambda < 1$. Furthermore, it encompasses a large set of conventional densities. For instance, if $\lambda = 0$, Hansen's distribution is reduced to the traditional student-t distribution. We recall that the traditional student-t distribution is not skewed. If in addition $\eta = \infty$, the student-t distribution collapses to a normal density.

It is well known that a traditional student-t with η degrees of freedom allows for the existence of all moments up to the ηth. Therefore, given the restriction $\eta > 2$, Hansen's skewed t distribution is well defined and its second moment exists. The higher moments are not directly given by the parameter η, although formulas exist for these moments.

Proposition 5 *If $Z \sim D(z|\eta,\lambda)$, then Z has zero mean and unit variance.*

Proof. See Hansen (1994). ■

Proposition 6 *Define $X = bZ + a$, and write the jth moment of X as $m_j \equiv E[X^j]$. Then X has mean a and standard deviation b. Furthermore*

$$E[Z^3] = [m_3 - 3a\,m_2 + 2a^3]/b^3, \tag{7}$$

$$E[Z^4] = [m_4 - 4a\,m_3 + 6a^2m_2 - 3a^4]/b^4, \tag{8}$$

where

$$
\begin{aligned}
m_2 &= 1+3\lambda^2, \\
m_3 &= 16c\,\lambda(1+\lambda^2)\frac{(\eta-2)^2}{(\eta-1)(\eta-3)} && \textit{if } \eta>3, \\
m_4 &= 3\frac{\eta-2}{\eta-4}(1+10\lambda^2+5\lambda^4) && \textit{if } \eta>4.
\end{aligned}
$$

Proof. Straightforward but tedious computations. ■

Since Z has zero mean and unit variance, we obtain that skewness (Sk) and kurtosis (Ku) are directly related to the third and fourth moments: $\mathrm{Sk}[Z] = \mathrm{E}[Z^3]$ and $\mathrm{Ku}[Z] = \mathrm{E}[Z^4]$. Excess kurtosis is defined as $\mathrm{XKu} = \mathrm{Ku} - 3$.

We emphasize that the density and the various moments do not exist for all parameters. Given the way asymmetry is introduced, we must have $-1 < \lambda < 1$. As already mentioned, the density d is meaningful only if $\eta > 2$. Furthermore, careful scrutiny of the algebra yielding equation (7) shows that skewness exists if $\eta > 3$. Last, kurtosis in equation (8) is well defined if $\eta > 4$.[9]

In the continuous time finance literature, asset prices are often assumed to follow a Brownian motion combined with jumps. This translates into returns data with occasional very large realizations. Our model captures such instances since, if η is small, e. g. close to 2, not even skewness exists.

3.2 The distribution of the generalized student-t

The copula involves marginal distributions rather than densities. For this reason, we now derive the cumulative distribution function (cdf) of Hansen's density. To do so, we recall that the conventional student-t distribution is defined by

$$
f(x) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})}\frac{1}{\sqrt{\pi n}}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}
$$

where n is the degree of freedom parameter. Numerical evaluation of the cdf of the conventional student-t is well known and procedures are provided in most software packages, and in particular in Fortran, which is the language retained in our study. We write the cdf of a student-t with n degrees of freedom as

$$
A(t;n) = \int_{-\infty}^{t} f(x)dx.
$$

[9]In empirical applications, we will only impose that $\eta > 2$ and let the data decide for itself if, for a given time period, a specific moment exists or not.

The following proposition presents the cdf of Hansen's generalized student-t.

Proposition 7 *Defining $D(t) = Pr[Z < t]$, where Z follows the density (6), yields*

$$D(t) = \begin{cases} (1-\lambda)A\left(\frac{bt+a}{1-\lambda}\sqrt{\frac{\eta-2}{\eta}},\eta\right) & \text{if } t < -a/b, \\ \frac{(1-\lambda)}{2} + \left(A\left(\frac{bt+a}{1+\lambda}\sqrt{\frac{\eta-2}{\eta}},\eta\right) - \frac{1}{2}\right)(1+\lambda) & \text{if } t \geq -a/b. \end{cases}$$

Proof. Suppose that $t < -a/b$. Given the definition of $D(t)$, we have

$$D(t) = \int_{-\infty}^{t} bc\left(1 + \frac{1}{\eta-2}\left(\frac{bz+a}{1-\lambda}\right)^2\right)^{-\frac{\eta+1}{2}} dz$$

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

$$= (1-\lambda)A\left(\frac{bt+a}{1-\lambda}\sqrt{\frac{\eta-2}{\eta}},\eta\right).$$

The second equation follows from a change of variable involving $u = (bz+a)/(1-\lambda)$. The last equation follows from a trivial change of variable. In the case where $t = -a/b$, we obtain that

$$D(t) = \frac{(1-\lambda)}{2}.$$

For $t > -a/b$, we have

$$D(t) = D(-a/b) + \int_{-a/b}^{t} bc\left(1 + \frac{1}{\eta-2}\left(\frac{bz+a}{1-\lambda}\right)^2\right)^{-\frac{\eta+1}{2}} dz.$$

The result now follows from a computation analogous to the case $t < -a/b$. ■

It is easy to verify that $D(-\infty) = 0$, and $D(\infty) = 1$.

3.3 A GARCH model allowing for conditional skewness and kurtosis

Let r_t, for $t = 1, \cdots, T$, be returns of a given series. The margin model that allows for time-varying volatility and for conditional skewness and kurtosis is defined by

$$r_t = \mu_t + z_t, \quad (9)$$

$$z_t = \sigma_t \epsilon_t, \quad (10)$$

$$\sigma_t^2 = a_0 + a_1^+(z_{t-1}^+)^2 + a_1^-(z_{t-1}^-)^2 + a_2\sigma_{t-1}^2, \quad (11)$$

$$\epsilon_t \sim D(\epsilon_t|\eta_t, \lambda_t). \quad (12)$$

Equation (9) decomposes the return of time t into a conditional mean, μ_t, and an innovation, z_t. Equation (10) defines this innovation as the product between conditional volatility, σ_t, and a residual, ϵ_t. The next equation (11) determines the dynamics of volatility. We use the notation $z^+ = \max(z, 0)$ and $z^- = \max(-z, 0)$. Such a specification has been suggested by Glosten, Jagannathan, and Runkle (1993), and by Zakoïan (1994). In a similar spirit, one may mention Campbell and Hentschel (1992) or Gourieroux and Monfort (1992). Equation (12), specifies that residuals follow a generalized student-t with time-varying parameters (η_t, λ_t). We defined this density above.

It is tempting to use for η_t and λ_t a specification similar to an ARMA(1,1), thus resembling equation (11). Such a specification is, however, hazardous. Indeed, for financial data, there exist outliers (such as the October 1987 crash). This in turn may lead to spuriously significant parameters. To see how such spurious parameters may arise, let us proceed with a thought experiment. We assume an ARMA(1,1) type specification for the parameter λ such as

$$\lambda_t = a + bz_{t-1} + c\lambda_{t-1}.$$

Furthermore, we consider that the data results from i.i.d. normal data. Now, the estimates of b and c will be small and statistically non-significant. Because of random variation, b and c will not be equal to zero. Assume that they take positive values. For our thought experiment, we consider now the replacement of r_{t-1}, the return at time $t-1$, by a large positive perturbation. Had such an event existed in reality, it would have created heavy tailedness. Because, at time t, λ_t needs to be bounded above by 1, the program will converge to a solution where the impact is undone at time t. This is achieved by the choice of a large negative c that may appear statistically significant even if robust estimates of the standard error are used (as suggested by White, 1980). For this reason, we will use a specification without lagged parameter:

$$\eta_t = b_0 + b_1 z_{t-1} + b_2 z_{t-2}, \quad (13)$$

$$\lambda_t = c_0 + c_1 z_{t-1} + c_2 z_{t-2}. \quad (14)$$

We also have the following constraints on the parameters $a_1^+ + a_2 < 1$, $a_1^- + a_2 < 1$, $2 < \eta_t$, $1 < \lambda_t < 1$. The first two equations guarantee stationarity for the volatility process given by equation (11). The following $T-2$ constraints are necessary to guarantee that the density is well defined (the first two values η_1 and η_2 are not needed). The last restrictions involve $2(T-2)$ inequality constraints that guarantee that skewness will be well defined.

We also experimented with a logistic map and found that the resulting dynamics of η_t and λ_t differed from ours for extreme returns. One explanation of this finding is that the logistic map becomes flat, even for relative small deviations from 0, and as a consequence it is unable to distinguish large deviations from only relatively large ones.[10]

A consequence of the way in which we implement the model is that, for simulation purposes or forecasting experiments, η_t and λ_t may not obey the restrictions. For such situations, we recommend to truncate the parameters.[11]

3.4 Numerical estimation of the marginal model

The estimation of model (9) to (14), under the constraints, represents a formidable task. Given that the likelihood is defined only if the constraints are not violated, it is necessary to use an optimization algorithm in which the constraints are always satisfied. This implies using an interior optimization algorithm. Furthermore, the sample is of a rather large size and hence the number of constraints becomes very large, say several thousands. For this reason, speed becomes a very important factor. Given the structure of our problem, we use a program especially developed by Gill, Murray, and Saunders (1997, 1999) for large optimizations involving many constraints, called SNOPT.

For a given set of initial values, this program first verifies that all initial values satisfy the restrictions. If this is not the case, it searches initial values verifying the restrictions and that are closest to the proposed initial values with respect to the Euclidean norm. Next, it uses a sequential quadratic programming algorithm whereby it is guaranteed that the linear constraints are always satisfied. In appendix A and B, we provide the gradients required for the implementation.

3.5 Specification of conditional dependency

Many different specifications of the dependency parameter are possible. In this paper we follow Gourieroux and Monfort (1992) and adapt a modelization where θ_t is conditional on the position of past joint realizations in the unit square. This means that we decompose the unit square of joint past realizations into a grid. The parameter θ_t will be constant for each element of the grid. More precisely, the basic model considered in this paper is

$$\ln(\theta_t) = \sum_{j=1}^{16} d_j I[(u_{t-1}, v_{t-1}) \in \mathcal{A}_j]$$

[10] This observation also holds with the arc-tangent map.

[11] This implies for $\eta_t < 2$ a value of 2.001. If $\lambda_t < -1$ we set it to -0.999 and similarly if $\lambda_t > 1$ we restrict it to 0.999. The restrictions were only rather seldom binding.

where $\mathcal{A}_j$ is the j-th element of the unit-square grid. To each parameter d_j an area $\mathcal{A}_j$ is associated.[12] For instance $\mathcal{A}_1 = [0, 1/4[\times[0, 1/4[$ and $\mathcal{A}_2 = [1/4, 1/2[\times[0, 1/4[$. The choice of 16 subintervals is purely arbitrary. This choice of parameterization has the advantage to allow the easy testing of various conjectures concerning the impact of past joint returns on subsequent dependency while still allowing for a large number of observations per area. For instance one may test if large joint realizations have a greater impact than small joint realizations on subsequent dependency. This could be investigated in a test of the null $d_1 = d_{16}$ versus $d_1 > d_{16}$.

In other applications, one may question whether dependency evolved linearly through time or followed some other complex patterns through time. This may be achieved by setting

$$\ln(\theta_t) = d_0 + d_1 t + d_2 t^2, \text{ for } t = 1, \cdots, T.$$

This parameterization may be viewed as a second-order approximation of a more general function. It obviously nests the case of a linear trend, $d_2 = 0$, and allows for several evolutions of dependency. Essentially, it represents an arc of a parabola. This parameterization allows for a decreasing dependency followed by an increasing one. This is the case if d_2 is positive and if there is a minimum, given by $t^* = -d_1/(2d_2)$ in the sample, i.e. $t \in [0, T]$. If d_2 is negative, then if $-d_1/(2d_2)$ belongs to $[0, T]$ a maximum has occurred at t^*.[13] In this case dependency would have increased and then decreased. If the extremum occurs before $t = 1$ then if d_2 is positive, as t gets larger, dependency augments quadratically and inversely if d_2 is negative.

Also, there may be situations in which dependency relies on additional variables. In such a situation one may chose a model such as

$$\ln(\theta_t) = z_t w_\theta$$

where z_t is a $(1 \times K)$ vector of explanatory variables and w_θ is a corresponding $(K \times 1)$ vector of parameters.

4 Empirical Results

4.1 The data

The choice of data is guided by the fact that we want to investigate the dependency of stock indices in local currency as well as in the \$ US referential. Given that exchange rates changed

[12] Figure 6 illustrates the position of the areas $\mathcal{A}_j$. How the figure is constructed is discussed in detail below.

[13] In numerical implementations, it may be necessary to rescale t for instance by using $\ln(\theta_t) = d_0 + d_1(t/\kappa) + d_2(t/\kappa)^2, \forall t$ where κ is a scaling factor such as 10^3.

from fixed to floating in 1972, we decided to start our study on January 1st, 1975. This eliminates the issue that there may have been some learning by market participants in the early period. The sample ends with January 24th, 2001. We downloaded all the data from Datastream. The chosen frequency is daily. To eliminate spurious correlation generated by holidays, we eliminated from the database those observations when a holiday occurred at least for one country. This reduced the sample from 6799 observations to 6669. Preliminary estimations revealed that the crash of October 1987 was of such importance that the dynamics of our model would be very much influenced by this event. For the S&P500, on that date, the index dropped by -22%. The second largest drop was only -8.64%. For this reason, we eliminated the data between October 17th and 24th. This reduces the sample by 6 observations to a total of 6663 observations.

The mnemonics are SP500 for the S&P 500, NIK for the NIKKEI, FTSE for the Financial Times stock index, DAX for the Deutsche Aktien Index, and CAC for the French Cotation Automatique Continue index. We also have Japanese, UK, German, and French exchange rates with respect to the dollar.

Descriptive statistics[14] show that stock-index returns are more volatile than currency returns.[15] For all series, the assumption of normality may be rejected using the Jarque-Bera test. The Engle test statistic reveals strong evidence of heteroskedasticity for all series.

4.2 Estimation of the marginal model

Table 1 presents the results of the general model with possible asymmetries in the impact of past good and bad news on conditional volatility and where skewness and kurtosis are time varying. The last row of that table presents a likelihood-ratio test statistic of the restriction of our general model to Bollerslev's (1986) GARCH (1,1) model where we assume that innovations follow a generalized student-t distribution with constant parameters. This test corresponds to the null of $a_1^+ = a_1^-$ and $b_1 = b_2 = c_1 = c_2 = 0$.[16] The likelihood-ratio test statistic, distributed as a χ^2 with 5 degrees of freedom, rejects the restrictions for all series. It is possible to separate the test of the impact of asymmetry, i.e. of $a_1^+ = a_1^-$ from the impact of past realizations on the tails of the distribution. Since both tests turned out statistically significant, we only report the results of the joint test.

[14] Available on request.

[15] This observation suggests that the results will not be very much affected by the choice of numeraire. We report in this paper mostly results for local currency returns and only sketch those in the $ US referential. We also checked the stability of our results using a pound sterling referential.

[16] We do not present the results of the estimation. The estimates are available on request.

A discussion of the actual impact on skewness and kurtosis of the various signs of the b_j and c_j, $j = 1, 2$ parameters is not straightforward since both the asymmetry parameter, λ_t, and the tail-fatness parameter, η_t, are closely entangled. Nonetheless, inspection of the parameters involved in Table 1 reveals that for most estimations, the parameters c_1 and c_2 are positive and in particular so for $c_1 + c_2$. Another regularity is that b_1, b_2, and the sum $b_1 + b_2$ tend to be negative. These regularities lead to the following dynamics: Subsequent to a large negative shock, the asymmetry parameter λ_t, also takes a negative value implying that the density of the innovations will have a fatter left tail. This impact is reinforced by the positive impact on η_t that generally controls the left and right fat tailedness. Subsequent to a large positive shock, the impact of the directional parameter λ_t and the tail-fatness parameter η_t offset each other and the global impact cannot be inferred directly.

We, thus, confirm the results of Harvey and Siddique (1999) that a negative return has an impact at different levels. First, subsequent to a negative return, as compared to a positive return of the same magnitude, volatility increases. This is the well-known impact documented by Campbell and Hentschel (1992), Glosten, Jagannathan, and Runkle (1993), as well as Zakoïan (1994). The impact in the extremes is expressed through the density via λ_t and η_t. The impact of this negative return is also to increase the left tail of the distribution, increasing the probability of a subsequent large negative event.

4.3 Estimation of the multivariate model

4.3.1 Parameter estimates

For all stock-market pairs available, we estimate the multivariate model. We present in Figures 4 and 5 scatterplots of the marginal cumulative density functions u and v. We notice that many points are distributed in the shape of a cross. These are associated with days when returns are very small or zero. We notice, furthermore, that except for the region where one margin is large and the other small, the unit square is rather uniformly filled with realizations. For the DAX-CAC pair, there is, in addition, a higher concentration along the diagonal. This is suggestive of the fact that the dependency of the French and German stock markets is higher than for the American and Japanese markets. From a modeling point of view, these scatterplots suggest that even if we introduce 16 areas for the conditioning, each one will contain enough observations so as to yield good estimates.

It should be emphasized that these scatterplots are unable to tell anything about temporal dependency. To establish if a temporal dependency exists or not, it is necessary to estimate the dynamic model. We now turn to the discussion of the parameter estimates. Figures 6 and

7 present the estimates and the associated standard errors of the various d_j. Whereas Figure 6 presents the estimates for the SP500-NIK pair, Figure 7 presents those of the DAX-CAC pair. Inspection of the parameter estimates of Figure 6 reveals that the extreme elements along the diagonal, ($d_1 = 1.7579$ and $d_{16} = 1.6108$), take rather similar values and that these values are larger than those at the center of the diagonal, ($d_6 = 1.4621$ and $d_{11} = 1.4102$). This observation suggests that subsequent to a joint large variation, i.e. both u and v simultaneously very large or very small, the dependency is higher. In terms of returns, this implies that simultaneous large positive or large negative returns yield subsequent increased dependency.

Inspection of the off-diagonal elements, and especially those where one return is large and positive and the other is large but negative, i.e. d_4 and d_{13} reveals that dependency is much smaller subsequent to such an event. The point estimate $d_{13} = 0.5503$ is nearly a third of $d_1 = 1.7579$. As may be expected from the scatterplot presented in Figure 4, the associated standard error is also larger.

The rather loose statements made so far will be formally tested below. Before doing so, we present in Figure 7 the same parameter estimates for the DAX-CAC pair. We first notice that the value taken by the parameters is larger that for the previous pair. This may be explained by the fact that the dependency of these series is higher, as we noticed already in Figure 5. For instance, the diagonal elements now take values ranging from 2.0331 for d_6 up to 2.3325 for d_{16}. The standard errors of these estimates is of a similar magnitude as for the SP500-NIK pair. This translates into higher t-ratios, meaning that the temporal dependency will be more pronounced for the DAX-CAC pair. We also notice that the pattern of magnitude of the estimates in this figure is of similar magnitude that the one of Figure 6, i.e. diagonal extreme estimates, (d_1, d_{16}), is larger than the central elements, (d_6, d_{11}), and on-diagonal elements are larger than off-diagonal ones.

4.3.2 Formal tests of conditional dependency

We now turn to the formal tests of the conditional dependency relations presented in Table 2. All the statistics shown in this table are based on Wald tests.[17] A first test investigates if the piecewise constant grid of the u-v unit square yields truly different values for the d_j or not. This test corresponds to the null hypothesis if $d_1 = d_2 = \cdots = d_{16}$ versus inequality for at least one pair of elements. The p-values presented in Table 2 reveal a rather uniform message. Except for the SP500-FTSE and NIK-FTSE pairs, dependency is conditional on past joint realizations. For certain stock-market pairs such as the FTSE-DAX, FTSE-CAC, and

[17] For certain tests such as test 1, a likelihood-ratio test may also be performed. We found that the Wald test yielded similar results to a likelihood-ratio test.

DAX-CAC, this relation is even very strongly significant.

The next test checks if joint bad news (u and v in the lower left area) induce greater dependency than joint good news (u and v in the upper right area). Formally this corresponds to a test of $d_1 = d_{16}$ versus $d_1 > d_{16}$. Inspection of the test 2 statistic and its associated p-value shows that this test is rejected only for the NIK-FTSE and the FTSE-DAX pairs. Unlike the univariate finding by Campbell and Hentschel (1992), i.e. that bad news create greater volatility, at the multivariate level we are unable to detect a similar pattern for correlation.

We may now pursue the tests with investigating if large joint return realizations, be they of positive or negative sign, yield to higher dependency than if returns are small. We formulate this question as a test of the null hypothesis $d_1 = d_6 = d_{11} = d_{16}$ against $d_1 = d_{16} > d_6 = d_{11}$. Inspection the test 3 statistics shows that for most cases the null cannot be rejected. Exceptions are the SP500-NIK and especially the stock-market pairs situated within Europe: the FTSE-DAX, FTSE-CAC and DAX-CAC pairs. This means that within Europe, stock markets tend to be more dependent subsequent to joint return realizations. One may view this finding as evidence that there is persistence of dependency within Europe.

Given that we found under test 1 evidence of conditional dependency for most countries, there remains the question whether some regularity may be established. Building on previous observations that there is some evidence that realizations of returns of the same sign yield further dependency, one is tempted to formulate a weaker test. This test investigates if dependency is higher subsequent to a joint variation than after opposite variations, i.e. the null is $d_1 = d_6 = d_{11} = d_{16} = d_{13} = d_9 = d_{14} = d_3 = d_4 = d_8$ against $d_1 = d_6 = d_{11} = d_{16} > d_{13} = d_9 = d_{14} = d_3 = d_4 = d_8$. We find, as the last two lines of Table 2 reveal, that this test is indeed rejected for nearly all countries. Again the three European stock markets distinguish themselves with particularly high values of the test statistics (5.42, 7.32, and 7.01) all having zero p-values, since these tests are distributed as a normal under the null and the test is unilateral.

4.3.3 Tests for \$ US denominated series

We now present some results in order to investigate how the results change when the stock indices are converted into \$ US. As a first stage, we study if the dependency parameter $\ln(\theta)$ is affected by a change of numeraire. In Table 3, Panel A presents those estimates in local currency denomination. We find that dependency is the rule between all stock-market pairs. Panel A also presents the test if the results are dependent on the choice of the margin model. We construct a likelihood-ratio test based on a restriction of the margin model to a GARCH(1,1) without time variation in skewness and kurtosis but still with generalized student-t distributed errors.

For all pairs, we strongly reject the simple GARCH(1,1), model. Thus, this test corroborates our using a model with time-varying skewness and kurtosis.

In Panel B, we present the same tests as in Panel A yet in a $ US referential. We find that the restriction to a simple GARCH(1,1) is still rejected. Turning to the dependency parameter, we notice that, for those country pairs where one of the series is the SP500, the unconditional dependency is smaller. For instance for the SP500-NIK, $\ln(\theta)$ drops from 1.42 to 1.21 and similarly for the SP500-CAC where the estimate falls from 1.59 to 1.26. We also notice that for country pairs not involving the SP500, dependency rises. This observation may be explained by the fact that dollar variations introduce additional dependency.

We now turn to Panel C that replicates the tests of Table 2. Test 1 is slightly weakened in that, now, for many cases the test of no conditional dependency cannot be rejected. This test is of course rather weak and if one tests for more specific patterns these may turn out statistically significant. Test 2 and Test 3 show that elements along the diagonal have no privileged role to play. In most cases, neither we find that joint large negative returns have a greater impact on dependency than joint positive returns, nor that joint large returns of any sign affect dependency more than joint small realizations. Presently, we turn to test 4 to investigate if the diagonal elements play a more important role than elements far from the diagonal. We find that all test statistics, except the one for the SP500-NIK and the SP500-FTSE, are significant. This corroborates the previous finding that dependency will be higher subsequent to joint stock market movements. We can reformulate this by stating that dependency is persistent even if one changes numeraire.

4.4 Dependency as a function of time

Previous research, see Longin and Solnik (1995), also investigated if dependency varied as a function of time. Whereas they used monthly data and a linear time trend, we decided to focus on daily returns and a quadratic trend. We report the results for local currency denominated series. This implies that there will be no contamination due to possible Dollar movements. Table 4 presents the results of this estimation.

The first question to which we turn is if there has been a change of dependency through time. This is tantamount to a test of the null hypothesis $d_1 = d_2 = 0$ against $d_1 \neq 0$ and $d_2 \neq 0$. We perform this test with a likelihood-ratio test. For all series, except the SP500-NIK and the NIK-CAC, we find that dependency changed with time. If d_2 is different from zero, then the specification represents a parabola, thus, there will be a minimum if d_2 is negative, or a maximum if d_2 is positive. Also when an extreme occurs, it may take place before, in or after our sample. The last line displays the occurrence of this minimum. A dash indicates that the extremum occurred before our sample started.

For the FTSE-DAX, FTSE-CAC, and the DAX-CAC, the minimum took place before our sample started. However, among these pairs, only for the DAX-CAC pair do we find a significant parameter d_2. This implies that for the three European pairs dependency increased. Only for the DAX-CAC there is a faster than linear trend.

Turning to the other pairs, we notice that the d_1 estimate is always positive and that the d_2 estimate is always negative. This implies a pattern where dependency increased up to a certain period of time and then decreased. Focusing only on those pairs for which d_2 is significant, we obtain for the SP500-DAX an increase of dependency till July 1991 followed by a decrease. Similar dates are June 1982 for the SP500-CAC, October 1993 for the NIK-FTSE and March 1990 for the NIK-DAX. It should be noticed, however, that the parameter is badly estimated for the SP500-CAC pair. As a consequence, lots of uncertainty surrounds the date of this latter pair. Considered together, for the first seven pairs except for the SP500-NIK and the NIK-CAC, there appears a decrease in dependency since the mid 90s. This observation is of importance for portfolio allocation purposes.

5 Conclusion

For many years, the traditional mean-variance framework developed by Sharpe, Lintner, and Markowitz has been the reference for portfolio management. This framework, involving only first and second moments, seems overly simplistic as testified by the many contributions to Value-at-Risk which focus on large deviations. Early theoretical contributions as by Kraus and Litzenberger (1976), Friend and Westerfield (1980), Barone-Adesi (1985), and Ingersoll (1990) extend the traditional framework by introducing higher moments such as skewness and kurtosis. In this light, it is astonishing that no more effort has been made to test portfolio models involving higher moments. An exception being Harvey and Siddique (2000), who test a portfolio model involving higher moments. One possible reason is that presently there exists no technique to express these higher moments in a conditional multivariate framework which also allows for non-normality of innovations. Our framework appears to provide the necessary technology within which this type of model may be tested.

We apply this model to daily returns of stock-market indices over the period from 1975 till the beginning of 2001. We find strong evidence of persistence in dependency both for local currency and \$ US denominated series. For European stock markets, we also find evidence that large simultaneous returns of either sign lead to higher subsequent dependency. We also find evidence that dependency changed through time. For stock markets within Europe, dependency increased whereas dependency involving the SP500 or the NIKKEI decreased over the recent period.

Our model may also be used to investigate the dependency properties of other markets than considered here. For instance, it may be of interest to consider emerging markets. The volatility spillovers among such markets have been investigated for instance in Beakert and Harvey (1995) and in Rockinger and Urga (2001). This framework may also be used to investigate the spillover of large realizations.

Furthermore, a straightforward extension of our framework could yield a model for the joint distribution of returns, volume, and duration between transactions. For instance, Marsh and Wagner (2000) investigate the return-volume dependence when extreme events occur. In this case, one could use a tri-variate copula or one could proceed in successive steps, modeling first the dependency between volume and duration using a first copula. Then, in a second step, one could link this copula to the returns series through another copula. Hence, our model may be adapted to settings, where the data of each margin is not of the same nature.

Appendix A

In the following appendix, we present the computations of the gradient of the log-likelihood for the margin model. To simplify notations, we focus on the gradient of a single observation. Summation of these gradients yields the sample gradients. We define $d = (br/\sigma + a)/(1 - \lambda s)$ where s is a sign dummy taking the value of 1 if $br/\sigma + a < 0$ and $s = -1$ otherwise. We also define $v_1 = 1 + d^2/(\eta - 2)$. We recall that the likelihood of an observation is

$$l = \ln(b) + \ln\left(\Gamma\left(\frac{\eta+1}{2}\right)\right) - \frac{1}{2}\ln(\pi) - \frac{1}{2}\ln(\eta-2) - \ln\left(\frac{\eta}{2}\right) - \ln(\sigma) - \frac{\eta+1}{2}\ln(v_1).$$

To obtain the gradients with respect to the various parameters $a_{0,}b_0^+, b_0^-, c_0, a_{1,}b_{11}, b_{12}, a_{1,}b_{21}, b_{22}$, we decompose the problem and make frequent use of the chain rule of differentiation. The necessary ingredients to obtain the gradients are:

$$\frac{\partial l_t}{\partial \sigma} = -\frac{1}{\sigma} + \frac{\eta+1}{2}\frac{1}{v_1}\frac{2d}{\eta-2}\frac{br}{(1-\lambda s)\sigma^2}.$$

Next we have

$$\begin{aligned}
\frac{\partial a}{\partial \lambda} &= 4c(\eta-2)(\eta-1)^{-1}, \\
\frac{\partial b}{\partial \lambda} &= (3\lambda - a\,\partial a/\partial\lambda)/b, \\
\frac{\partial d}{\partial \lambda} &= \left(\frac{\partial b}{\partial \lambda}\frac{r}{\sigma} + \frac{\partial a}{\partial \lambda}\right)(1-\lambda s)^{-1} + sz(1-\lambda s)^{-2}, \\
\frac{\partial v_1}{\partial \lambda} &= \frac{2}{\eta-2}d\frac{\partial d}{\partial \lambda},
\end{aligned}$$

so that

$$\frac{\partial l_t}{\partial \lambda} = \frac{1}{b}\frac{\partial b}{\partial \lambda} - \frac{\eta+1}{2}\frac{1}{v_1}\frac{\partial v_1}{\partial \lambda}.$$

To obtain $\partial l_t/\partial\eta$, we proceed similarly. First, we notice that $\partial c/\partial\eta = c\,\partial\ln(c)/\partial\eta$ and

$$\begin{aligned}
\frac{\partial \ln(c)}{\partial \eta} &= \frac{1}{2}\Psi(\frac{\eta+1}{2}) - \frac{1}{2}\frac{1}{\eta-2} - \frac{1}{2}\Psi\left(\frac{\eta}{2}\right), \\
\frac{\partial a}{\partial \eta} &= 4\lambda(\eta-2)(\eta-1)^{-1}\frac{\partial c}{\partial \eta} + 4\lambda c[(\eta-1)^{-1} - (\eta-2)(\eta-1)^{-2}], \\
\frac{\partial b}{\partial \eta} &= -\frac{a}{b}\frac{\partial a}{\partial \eta}, \qquad \frac{\partial d}{\partial \eta} = \left(\frac{\partial b}{\partial \eta}\frac{r}{\sigma} + \frac{\partial a}{\partial \eta}\right)/(1-\lambda s), \\
\frac{\partial v_1}{\partial \eta} &= -(\eta-2)^{-2}d^2 + 2(\eta-2)^{-1}d\frac{\partial d}{\partial \eta}, \\
\frac{\partial l_t}{\partial \eta} &= \frac{1}{b}\frac{\partial b}{\partial \lambda} + \frac{\partial \ln(c)}{\partial \eta} - \frac{1}{2}\ln(v_1)6\frac{\eta+1}{2}\frac{1}{v_1}\frac{\partial v_1}{\partial \eta},
\end{aligned}$$

where $\Psi(\cdot)$ is the derivative of the log of the gamma function. This derivative is known as the di-gamma or psi function, which may be implemented with desired accuracy. The Fortran library IMSL implements this function.

Now, we can compute the partials with respect to the actual parameters by using:

$$
\begin{aligned}
\frac{\partial l_t}{\partial a_0} &= \frac{\partial l_t}{\partial \sigma}\left(1 + c_0 \frac{\partial h_{t-1}}{\partial a_0}\right), \\
\frac{\partial l_t}{\partial b_0} &= \frac{\partial l_t}{\partial \sigma}\frac{1}{2\sigma}\left(r_{t-1}^2 + c_0 \frac{\partial h_{t-1}}{\partial b_0}\right), \\
\frac{\partial l_t}{\partial c_0} &= \frac{\partial l_t}{\partial \sigma}\frac{1}{2\sigma}\left(h_{t-1} + c_0 \frac{\partial h_{t-1}}{\partial c_0}\right), \\
\frac{\partial h_1}{\partial a_0} &= 1 + c_0 \frac{\partial h_0}{\partial a_0} = 1, \quad \frac{\partial h_2}{\partial a_0} = 1 + c_0, \\
\frac{\partial l_t}{\partial a_1} &= \frac{\partial l_t}{\partial \eta}, \quad \frac{\partial l_t}{\partial b_{11}} = \frac{\partial l_t}{\partial \eta} r_{t-1}, \quad \frac{\partial l_t}{\partial b_{12}} = \frac{\partial l_t}{\partial \eta} r_{t-2}, \\
\frac{\partial l_t}{\partial a_2} &= \frac{\partial l_t}{\partial \lambda}, \quad \frac{\partial l_t}{\partial b_{21}} = \frac{\partial l_t}{\partial \lambda} r_{t-1}, \quad \frac{\partial l_t}{\partial b_{22}} = \frac{\partial l_t}{\partial \lambda} r_{t-2}.
\end{aligned}
$$

Appendix B

In the following appendix, we provide elements useful to maximize the copula density and will point out a limitation of this technique. We have

$$\begin{aligned}
\frac{\partial \ln c}{\partial \theta} &= \frac{1}{\theta} + \frac{u - 2uv + v}{1 + (u - 2uv + v)(\theta - 1)} - \frac{3}{2}\frac{2(u+v) + 2(\theta - 1)(u+v)^2 - 4uv(2\theta - 1)}{[1 + (\theta - 1)(u+v)]^2 - 4uv\theta(\theta - 1)} \\
\frac{\partial \ln c}{\partial u} &= \frac{(1 - 2v)(\theta - 1)}{1 + (u - 2uv + v)(\theta - 1)} - \frac{3}{2}\frac{2[1 + (\theta - 1)(u+v)](\theta - 1) - 4v(\theta - 1)}{[1 + (\theta - 1)(u+v)]^2 - 4uv\theta(\theta - 1)} \\
\frac{\partial \ln c}{\partial v} &= \frac{(1 - 2u)(\theta - 1)}{1 + (u - 2uv + v)(\theta - 1)} - \frac{3}{2}\frac{2[1 + (\theta - 1)(u+v)](\theta - 1) - 4u(\theta - 1)}{[1 + (\theta - 1)(u+v)]^2 - 4uv\theta(\theta - 1)}.
\end{aligned}$$

We notice now that u and v being cdfs will involve parameters. This implies that if derivatives with respect to those parameters are required it will be necessary to use the chain rule. For instance, if $u = F_X(x, w_X)$, then

$$\frac{\partial l_t}{\partial w_X} = \frac{\partial l_t}{\partial u}\frac{\partial F_X(x, w_X)}{\partial w_X}. \tag{B.1}$$

This computation reveals that the explicit computation of the score is far from being trivial. In our model the marginal cdf is given by the D function. Inspection of the required derivatives to compute (15) shows that one would need a technique to compute the integral of a derivative of a student-t with respect to the degree of freedom parameter. To our best knowledge such a technique does not exist though it may be obtained using continuous fractions or similar numerical approaches. We leave such developments for further research.

References

[1] Ang, A., and G. Bekaert (1999), International Asset Allocation with Time-Varying Correlations, NBER Working Paper 7056.

[2] Ball, C. A., and W. N. Torous (2000), Stochastic Correlation across International Stock Markets, *Journal of Econometrics*, 7(3-4), 373–388.

[3] Barone-Adesi, G. (1985), Arbitrage Equilibrium with Skewed Asset Returns, *Journal of Financial and Quantitative Analysis*, 20(3), 299–313.

[4] Bekaert, G., and C. L. Harvey (1995), Time-Varying World Market Integration, *Journal of Finance*, 50, 403–444.

[5] Bera, A. K., and M. L. Higgins (1993), ARCH Models: Properties, Estimation and Testing, *Journal of Economic Surveys,* 7(4), 305–362.

[6] Bera, A. K., and S. Kim (1996), Testing Constancy of Correlation with an Application to International Equity Returns, CIBER Working Paper 96–107.

[7] Bollerslev, T. (1986), Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, 31(3), 307–327.

[8] Bollerslev, T., R. Y. Chou, and K. F. Kroner (1992), ARCH Modelling in Finance: A Review of the Theory and Empirical Evidence, *Journal of Econometrics*, 52(1-2), 5–59.

[9] Bollerslev, T., R. F. Engle, and D. B. Nelson (1994), ARCH Models, in R. F. Engle and D. L. McFadden (eds.), *Handbook of Econometrics*, Chapter 49, Vol. IV, 2959–3038, Elsevier Science, Amsterdam.

[10] Bollerslev, T., and J. M. Wooldridge (1992), Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances, *Econometric Reviews*, 11(2), 143–172.

[11] Boyer, B. H., M. S. Gibson, and M. Loretan (1997), Pitfalls in Tests for Changes in Correlations, International Finance Discussion Paper 597, Board of Governors of the Federal Reserve System.

[12] Campbell, J. Y., and L. Hentschel (1992), No News Is Goos News: An Asymmetric Model of Changing Volatility, *Journal of Financial Economics*, 31(3), 281–318.

[13] Embrechts, P., A. J. McNeil, and A. J. Strautman (1999), Correlation and Dependency in Risk Management: Properties and Pitfalls, Working Paper, ETH Zurich.

[14] Engle, R. F. (1982), Auto-Regressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica*, 50(4), 987–1007.

[15] Engle, R. F., and G. Gonzalez-Rivera (1991), Semi-Parametric ARCH Models, *Journal of Business and Economic Statistics*, 9(4), 345–359.

[16] Engle, R. F., and S. Manganelli (1999), CAViaR: Conditional Value at Risk by Quantile Regression, NBER, Working Paper 7341.

[17] Engle, R. F., and V. K. Ng (1993), Measuring and Testing the Impact of News on Volatility, *Journal of Finance*, 48, 1749–1778.

[18] Forbes, K., and R. Rigobon (1999), No Contagion, Only Interdependence: Measuring Stock Market Co-movements, NBER, Working Paper 7267.

[19] Friend, I., and R. Westerfield (1980), Co-Skewness and Capital Asset Pricing, *Journal of Finance*, 35(4), 897–913.

[20] Gill, Ph. E., W. Murray, and M. A. Saunders (1997), SNOPT: An SQP Algorithm of Large-Scale Constrained Optimization, Report NA 97-2, Department of Mathematics, UCSD, California.

[21] Gill, Ph. E., W. Murray, and M. A. Saunders (1999), User's Guide for SNOPT 5.3: A Fortran Package for Large-Scale Nonlinear Programming, Working Paper, Department of Mathematics, UCSD, California.

[22] Glosten, R. T., R. Jagannathan, and D. Runkle (1993), On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks, *Journal of Finance*, 48(5), 1779–1801.

[23] Gourieroux, C., and A. Monfort (1992), Qualitative Threshold ARCH Models, *Journal of Econometrics*, 52, 159–199.

[24] Hamao, Y., R. W. Masulis, and V. K. Ng (1990), Correlations in Price Changes and Volatility Across International Stock Markets, *Review of Financial Studies*, 3(2), 281–307.

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

[26] Harvey, C. R., and A. Siddique (1999), Autoregressive Conditional Skewness, *Journal of Financial and Quantitative Analysis*, 34(4), 465–487.

[27] Harvey, C. R., and A. Siddique (2000), Conditional Skewness in Asset Pricing Tests, *Journal of Finance*, 55(3), 1263–1295.

[28] Ingersoll, J. Jr. (1990), *Theory of Financial Decision Making*, Rowman and Littlefield, Totowa, New Jersey.

[29] Joe, H. (1997), *Multivariate Models and Dependence Concepts*, Chapman and Hall, London.

[30] Johnson, M. E. (1987), *Multivariate Statistical Simulation*, John Wiley and Sons, New York.

[31] Jondeau, E., and M. Rockinger (2000), Conditional Volatility, Skewness, and Kurtosis: Existence and Persistence, HEC working paper.

[32] Kaplanis, E. C. (1988), Stability and Forecasting of the Co-Movement Measures of International Stock Market Returns, *Journal of International Money and Finance*, 8(1), 63–95.

[33] Karolyi, G. A. (1995), A Multivariate GARCH Model of International Transmissions of Stock Returns and Volatility: The Case of the United States and Canada, *Journal of Business and Economic Statistics*, 13(1), 11–25.

[34] Karolyi, G. A., and R. M. Stutz (1996), Why Do Markets Move Together? An Investigation of U.S.-Japan Stock Return Comovements, *Journal of Finance*, 51(3), 951–986.

[35] King, M., E. Sentana, and S. Wadhawini (1994), Volatility and Links between National Stock Markets, *Econometrica*, 62(4), 901–933.

[36] Koch, P. D., and T. W. Koch (1991), Evolution in Dynamic Linkages across National Stock Indices, *Journal of International Money and Finance*, 10(2), 231–251.

[37] Kotz, S., N. Balakishnan, and N. L. Johnson (2000), *Continuous Multivariate Distributions, Volume 1: Models and Applications*, Wiley Interscience, New York.

[38] Kraus, A., and R. H. Litzenberger (1976), Skewness Preference and the Valuation of Risk Assets, *Journal of Finance*, 31(4), 1085–1100.

[39] Longin, F., and B. Solnik (1995), Is the Correlation in International Equity Returns Constant: 1960–1990?, *Journal of International Money and Finance*, 14(1), 3–26.

[40] Longin, F., and B. Solnik (2000), Extreme Correlation of International Equity Markets, forthcoming, *Journal of Finance*.

[41] Marsh, T. A., and N. Wagner (2000) Return-Volume Dependence and Extremes in International Equity Markets, Haas School of Business, working paper.

[42] Mood, A. M., F. A. Graybill, and D. C. Boes (1982), *Introduction to the Theory of Statistics*, McGraw-Hill, New York.

[43] Nelsen, R. B. (1999), *An Introduction to Copulas*, Springer Verlag, New York.

[44] Nelson, D. B. (1991), Conditional Heteroskedasticity in Asset Returns: A New Approach, *Econometrica*, 59(2), 397–370.

[45] Plackett, R. L. (1965), A Class of Bivariate Distributions, *Journal of the American Statistical Association*, 60, 516–522.

[46] Premaratne, G., and A. K. Bera (2000), Modeling Asymmetry and Excess Kurtosis in Stock Return Data, University of Illino, Working Paper.

[47] Ramchand, L., and R. Susmel (1998), Volatility and Cross Correlation across Major Stock Markets, *Journal of Empirical Finance*, 5(4), 397–416.

[48] Ratner, M. (1992), Portfolio Diversification and the Inter-Temporal Stability of International Indices, *Global Finance Journal*, 3, 67–78.

[49] Rigobon, R. (1999), On the Measurement of the International Propagation of Shocks, NBER, Working paper 7354.

[50] Rockinger, M., and E. Jondeau (2001), Entropy Densities with an Application to Autoregressive Conditional Skewness and Kurtosis, forthcoming in *Journal of Econometrics*.

[51] Rockinger, M., and G. Urga (2001), A Time-Varying Parameter Model to Test for Predictability and Integration in the Stock Markets of Transtion Economies, *Journal of Business and Economic Statistics*, 19 (1), 73–84.

[52] Schweizer, B., and A. Sklar (1974), Operations on Distribution Functions Not Derivable from Operations on Random Variables, *Studia Mathematica,* 52, 43–52.

[53] Sklar, A. (1959), Fonctions de répartition à n dimensions et leurs marges, *Publications de l'Institut Statistique de l'Université de Paris*, 8, 229–231.

[54] Susmel, R., and R. F. Engle (1994), Hourly Volatility Spillovers between International Equity Markets, *Journal of International Money and Finance*, 13(1), 3–25.

[55] Tse, Y. K. (2000), A Test for Constant Correlations in a Multivariate GARCH Model, *Journal of Econometrics*, 98(1), 107–127.

[56] White, H. (1980), A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity, *Econometrica*, 48(4), 817–838.

[57] Zakoïan, J. M. (1994), Threshold Heteroskedastic Models, *Journal of Economic Dynamics and Control*, 18(5), 931–955.

Table 1: Parameters of the margin model with asymmetric GARCH and time varying skewness and kurtosis.

Each individual series of returns $\{r_t\}_{t=1}^T$ is modeled as a GARCH with time-varying skewness and kurtosis. Formally: $r_t = \mu + y_t$, $y_t = \sigma_t \epsilon_t$, $\sigma_t^2 = a_0 + a_1^+ (y_{t-1}^+)^2 + a_1^- (y_{t-1}^-)^2 + a_2 \sigma_{t-1}^2$, where ϵ_t follows a generalized student t with parameters λ_t and η_t whose dynamics is $\eta_t = b_0 + b_1 y_{t-1} + b_2 y_{t-2}$ and $\lambda_t = c_0 + c_1 y_{t-1} + c_2 y_{t-2}$.

The first line, time, presents the required time for the estimation. Lik is the value of the log-likelihood. LRT is the likelihood ratio test statistic of the restriction $a_1^+ = a_1^-$ and $b_1 = b_2 = c_1 = c_2 = 0$. Bold numbers are statistically significant.

	SP500	NIK	FTSE	DAX	CAC
time	39.16	32.73	60.37	40.75	57.34
a_0	**0.0065**	**0.0070**	**0.0156**	**0.0142**	**0.0150**
	(0.003)	(0.002)	(0.004)	(0.003)	(0.004)
a_1^+	**0.0250**	**0.0406**	**0.0623**	**0.0641**	**0.0629**
	(0.013)	(0.008)	(0.008)	(0.009)	(0.010)
a_1^-	**0.0641**	**0.1378**	**0.0770**	**0.1028**	**0.0966**
	(0.026)	(0.021)	(0.009)	(0.015)	(0.016)
a_2	**0.9474**	**0.9108**	**0.9128**	**0.9053**	**0.9124**
	(0.0156)	(0.0132)	(0.0101)	(0.0124)	(0.0127)
b_0	**6.8378**	**5.0170**	**20.2651**	**9.0528**	**5.7035**
	(3.160)	(0.328)	(4.570)	(1.223)	(0.470)
b_1	-0.2773	**-0.2965**	-1.5395	**-0.4133**	**-0.4075**
	(4.337)	(0.034)	(2.361)	(0.133)	(0.064)
b_2	0.4703	**-0.2051**	0.0035	**-0.9013**	**-0.3315**
	(1.532)	(0.033)	(3.663)	(0.266)	(0.051)
c_0	-0.0067	**-0.0467**	**-0.0565**	**-0.0399**	**-0.0226**
	(0.014)	(0.013)	(0.019)	(0.018)	(0.014)
c_1	0.0433	**0.0317**	**0.0397**	**0.0210**	-0.0135
	(0.028)	(0.013)	(0.017)	(0.015)	(0.012)
c_2	**0.0587**	0.0046	**0.0485**	**0.0382**	**0.0401**
	(0.005)	(0.012)	(0.015)	(0.017)	(0.013)
Lik.	-8150.94	-8238.39	-8778.08	-8979.88	-9196.59
LRT	46.13	109.69	19.55	35.22	32.59
p-value	0.00	0.00	0.00	0.00	0.00

Table 2: Test of restrictions on the conditional dependency parameter.
Estimates for local-currency denominated stock indices.

All test statistics of this table are Wald tests. Test 1 tests the null hypothesis that all the parameters describing the dependency parameter θ are constant, i. e. $d_1 = d_2 = \cdots = d_{16}$. This statistic is distributed as a χ^2 with 15 degrees of freedom.

Test 2 tests the null if $d_1 = d_{16}$ versus $d_1 > d_{16}$, that is if large joint variations lead to higher subsequent dependency than if joint variations are small. This test is unilateral and normally distributed.

Test 3 tests if $d_1 = d_6 = d_{11} = d_{16}$ versus $d_1 = d_{16} > d_6 = d_{11}$, that is if large joint variations lead to higher subsequent dependency than if joint variations are small.

Test 4 tests if dependency is higher subsequent to a joint variation than after opposite variations, i.e. the null is $d_1 = d_6 = d_{11} = d_{16} = d_{13} = d_9 = d_{14} = d_3 = d_4 = d_8$ against $d_1 = d_6 = d_{11} = d_{16} > d_{13} = d_9 = d_{14} = d_3 = d_4 = d_8$.

	SP-NIK	SP-FT	SP-DAX	SP-CAC	NIK-FT	NIK-DAX	NIK-CAC	FT-DAX	FT-CAC	DAX-CAC
Test 1	32.61	14.04	32.71	26.30	14.49	29.30	26.04	61.49	102.80	64.10
p-value	0.01	0.52	0.01	0.03	0.49	0.01	0.04	0.00	0.00	0.00
Test 2	0.53	-0.18	-0.06	-0.08	1.51	0.96	-1.16	2.01	1.01	-1.14
p-value	0.30	0.57	0.52	0.53	0.07	0.17	0.88	0.02	0.16	0.87
Test 3	1.65	-0.95	-1.38	-0.65	0.74	1.38	0.59	1.88	2.85	1.88
p-value	0.05	0.83	0.92	0.74	0.23	0.08	0.28	0.03	0.00	0.03
Test 4	1.16	2.28	3.94	3.60	1.25	1.43	1.62	5.42	7.32	7.01
p-value	0.12	0.01	0.00	0.00	0.11	0.08	0.05	0.00	0.00	0.00

Table 3: Comparison of dependency for local currency and for \$US denominated stock indices.

Panel A presents the estimates of the unconditional dependency parameter under the assumption that the marginal distributions are GARCH with time varying skewness and kurtosis. Residuals follow a generalized student t. If $\ln(\theta) = 0$ then margins would not be dependent. LRT tests the restriction to non-time varying marginal models ($a_1^+ = a_1^-$ and $b_1 = b_2 = c_1 = c_2 = 0$).
Panel B presents the same estimates and tests as Panel A but using \$US denominated stock indices.
Panel C presents Wald tests of similar nature than presented in Table 2.

Panel A

	SP-NIK	SP-FT	SP-DAX	SP-CAC	NIK-FT	NIK-DAX	NIK-CAC	FT-DAX	FT-CAC	DAX-CAC
$\ln(\theta)$	1.42	1.00	1.42	1.59	0.80	1.25	1.47	1.26	1.64	1.85
se	0.04	0.04	0.04	0.04	0.05	0.05	0.04	0.04	0.04	0.04
Lik.	-16065.66	-16689.22	-16684.22	-16875.54	-16902.39	-16969.45	-17135.13	-17341.98	-17310.74	-17409.90
LRT	70.81	61.26	63.21	58.42	122.37	140.37	133.61	52.08	47.85	95.44

Panel B

	SP-NIK	SP-FT	SP-DAX	SP-CAC	NIK-FT	NIK-DAX	NIK-CAC	FT-DAX	FT-CAC	DAX-CAC
$\ln(\theta)$	1.21	0.93	1.27	1.26	1.12	1.54	1.53	1.58	1.82	2.15
se	0.05	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.03	0.03
Lik	-17988.49	-17939.12	-17902.40	-17975.65	-19771.25	-19648.29	-19742.56	-19385.06	-19198.72	-18978.34
LRT	68.14	47.84	58.35	28.85	82.21	109.49	61.45	39.53	6.58	37.17

Panel C

	SP-NIK	SP-FT	SP-DAX	SP-CAC	NIK-FT	NIK-DAX	NIK-CAC	FT-DAX	FT-CAC	DAX-CAC
Test 1	17.08	5.76	29.81	15.05	19.73	17.11	29.60	31.17	53.50	37.52
p-value	0.31	0.98	0.01	0.45	0.18	0.31	0.01	0.01	0.00	0.00
Test 2	-0.23	-0.46	0.27	0.46	0.27	0.53	0.01	0.28	0.77	-0.96
p-value	0.59	0.68	0.40	0.32	0.39	0.30	0.50	0.39	0.22	0.83
Test 3	2.45	-0.04	-0.64	-0.92	2.53	0.75	-0.66	1.72	0.11	0.71
p-value	0.01	0.51	0.74	0.82	0.01	0.23	0.74	0.04	0.46	0.24
Test 4	-0.58	-0.40	1.95	2.29	1.22	2.88	3.34	3.88	6.52	5.30
p-value	0.72	0.65	0.03	0.01	0.11	0.00	0.00	0.00	0.00	0.00

Table 4: Test for a time variation in dependency

This table presents the results of the estimation of $\ln(\theta) = d_0 + d_1(t/1000) + d_2(t/1000)^2$ that is if dependency varied through time. The likelihood ratio test LRT tests if there is a dynamics, i. e. if $d_1 = d_2 = 0$ versus $d_1 \neq 0$ and $d_2 \neq 0$. Standard errors are presented between parenthesis. The given dynamics represents a parabola. If d_2 is negatif (positif), then for large values the parabola presents a(n) downward (upward) trend.

The last line indicates when the parabola reached an extremum. A dash indicates that the extremum took place before the sample starts.

	SP-NIK	SP-FT	SP-DAX	SP-CAC	NIK-FT	NIK-DAX	NIK-CAC	FT-DAX	FT-CAC	DAX-CAC
time	0.71	0.94	0.61	1.21	0.87	0.88	0.72	0.66	1.09	1.04
d0	1.2643	0.7507	1.0027	1.6402	0.0954	0.7948	1.5593	0.2642	0.3084	1.0965
	(0.1360)	(0.1145)	(0.1235)	(0.1138)	(0.1501)	(0.1812)	(0.1328)	(0.1239)	(0.1421)	(0.1433)
d1	0.0768	0.1249	0.2624	0.0855	0.3891	0.3099	0.0125	0.2350	0.3309	0.0392
	(0.0903)	(0.0809)	(0.0799)	(0.0795)	(0.0985)	(0.1130)	(0.0915)	(0.0785)	(0.0832)	(0.0866)
d2	-0.0070	-0.0110	-0.0311	-0.0224	-0.0404	-0.0400	-0.0084	0.0117	0.0108	0.0389
	(0.0127)	(0.0117)	(0.0112)	(0.0116)	(0.0140)	(0.0155)	(0.0134)	(0.0108)	(0.0110)	(0.0116)
LRT	2.26	7.05	13.97	13.95	31.14	11.00	4.43	264.10	455.17	276.43
p-value	0.32	0.03	0.00	0.00	0.00	0.00	0.11	0.00	0.00	0.00
extr. Date	96/07/19	97/02/24	91/07/11	82/06/16	93/10/25	90/03/02	77/11/21	-	-	-

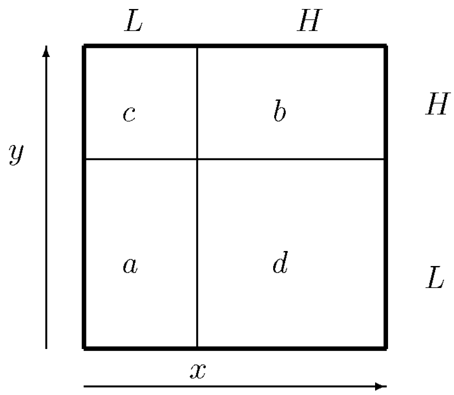
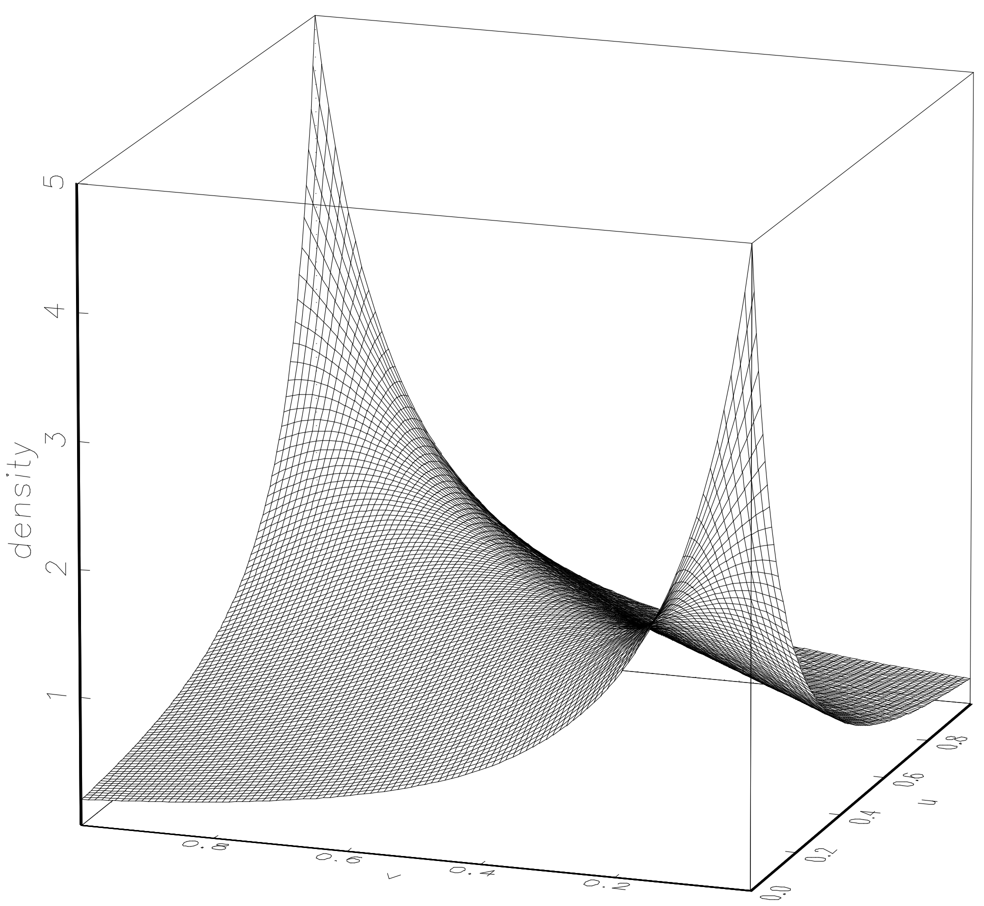
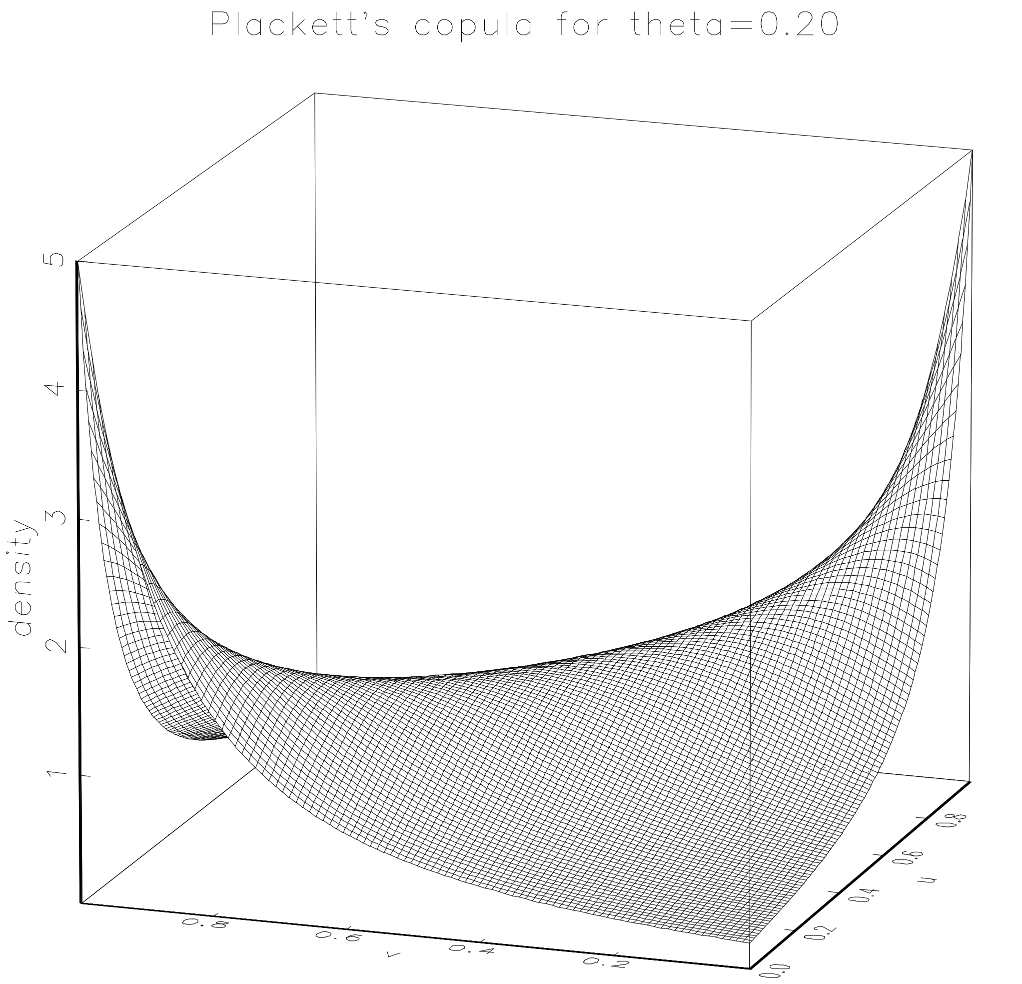
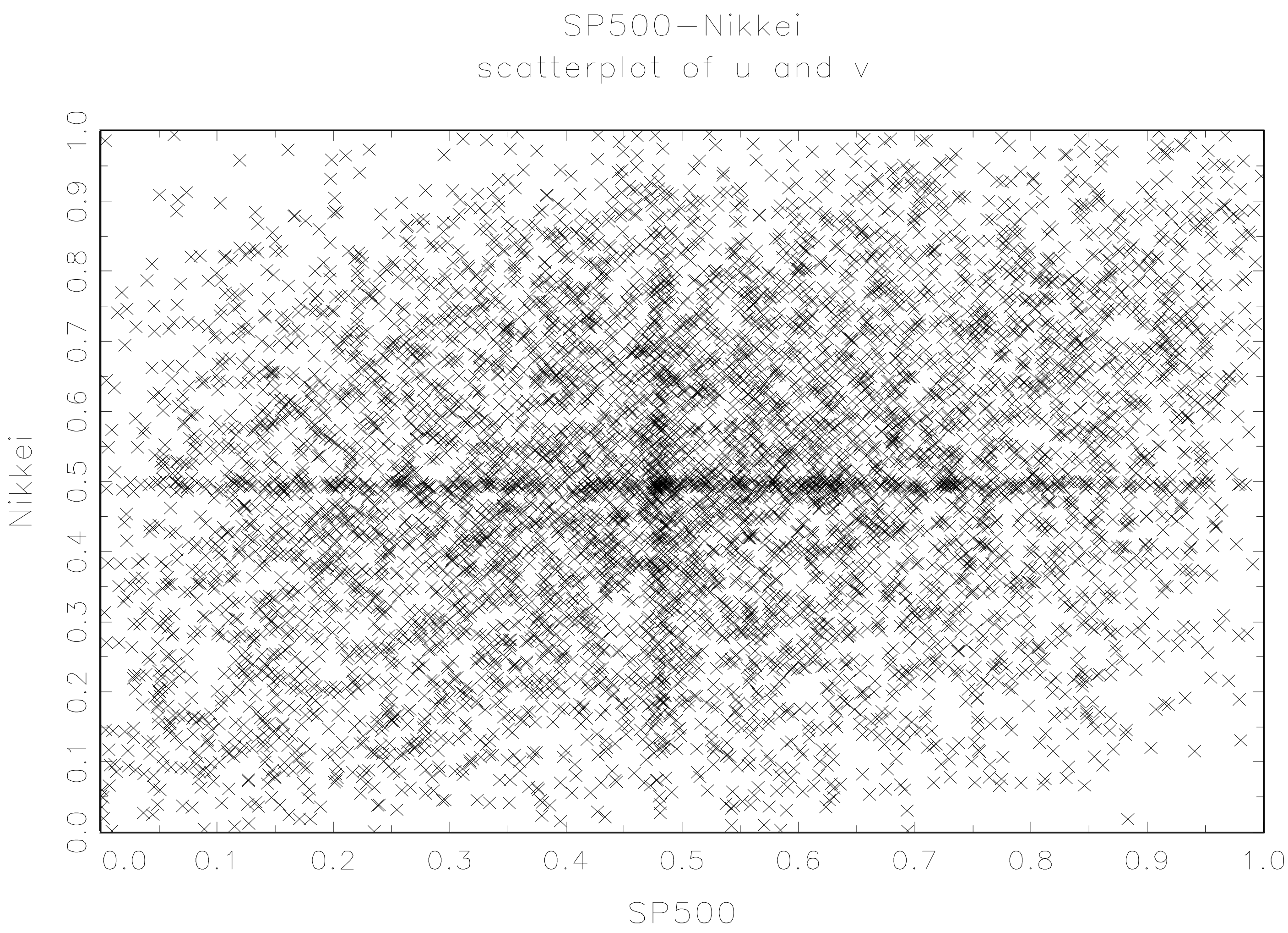


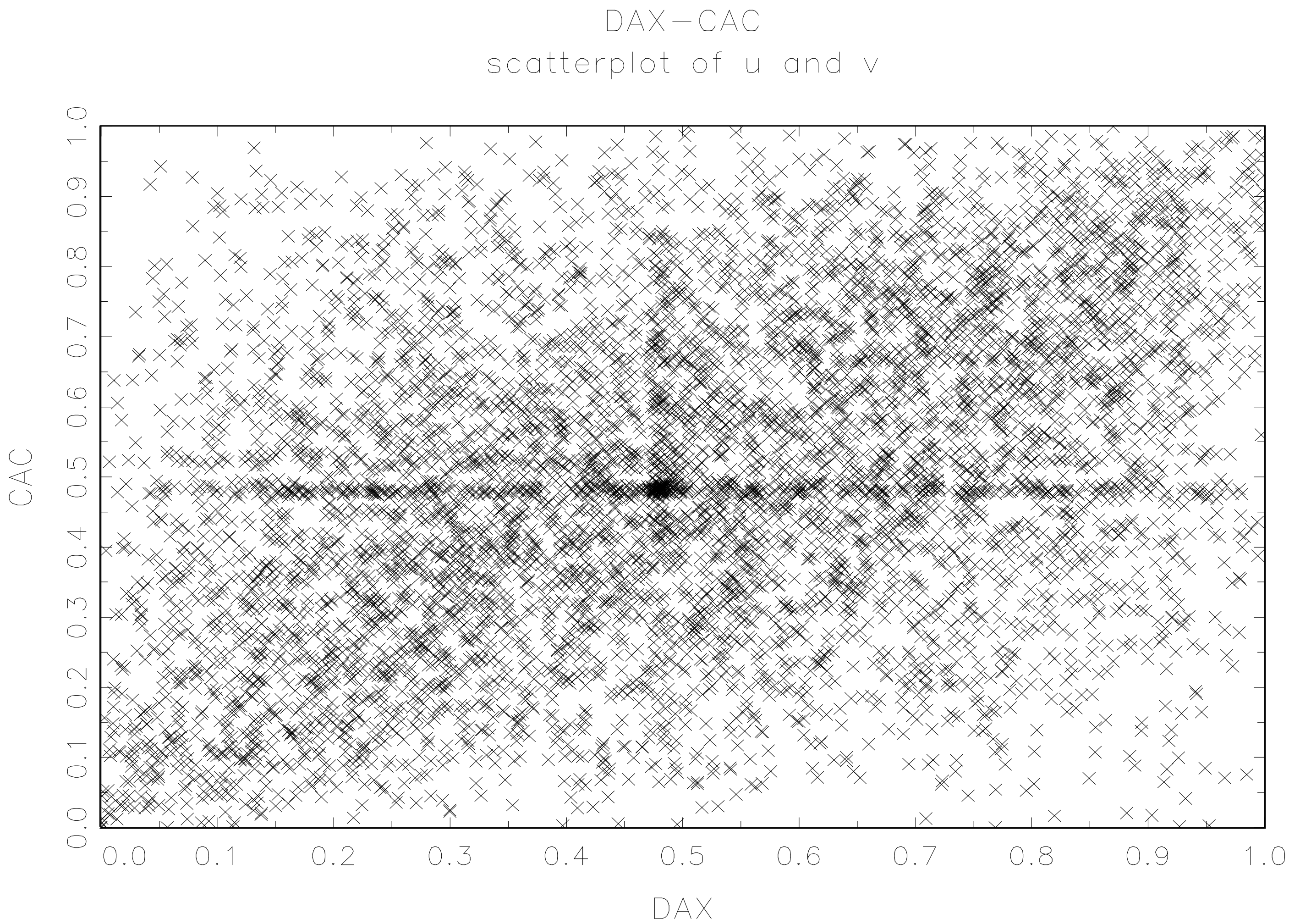
Figure 1.

Plackett's copula for theta=5.00




Plackett's copula for theta=0.20
density
5
4
3
2
1
0.8
0.6
v
0.4
0.2
0.0
0.2
0.4
0.6
u
0.8


SP500−Nikkei
scatterplot of u and v
Nikkei
SP500
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0


DAX−CAC
scatterplot of u and v
CAC
DAX
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0

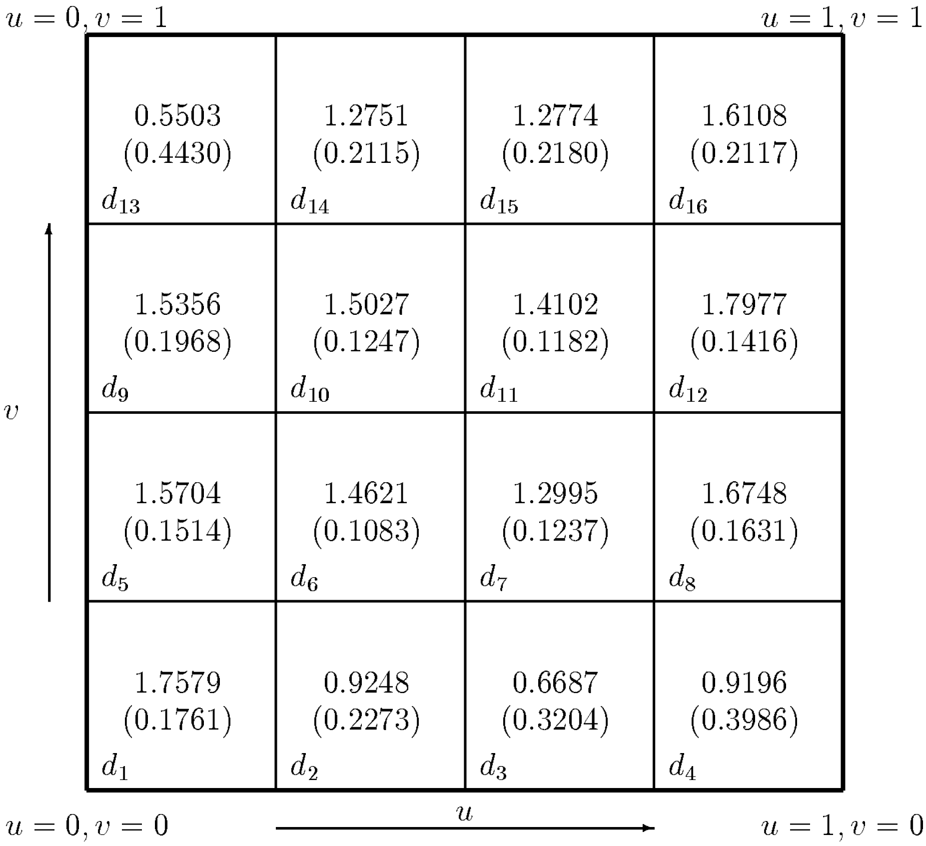

$u=0, v=1$
$u=1, v=1$
0.5503 (0.4430) d_{13}
1.2751 (0.2115) d_{14}
1.2774 (0.2180) d_{15}
1.6108 (0.2117) d_{16}
1.5356 (0.1968) d_9
1.5027 (0.1247) d_{10}
1.4102 (0.1182) d_{11}
1.7977 (0.1416) d_{12}
v
1.5704 (0.1514) d_5
1.4621 (0.1083) d_6
1.2995 (0.1237) d_7
1.6748 (0.1631) d_8
1.7579 (0.1761) d_1
0.9248 (0.2273) d_2
0.6687 (0.3204) d_3
0.9196 (0.3986) d_4
$u=0, v=0$
u
$u=1, v=0$

Figure 6.

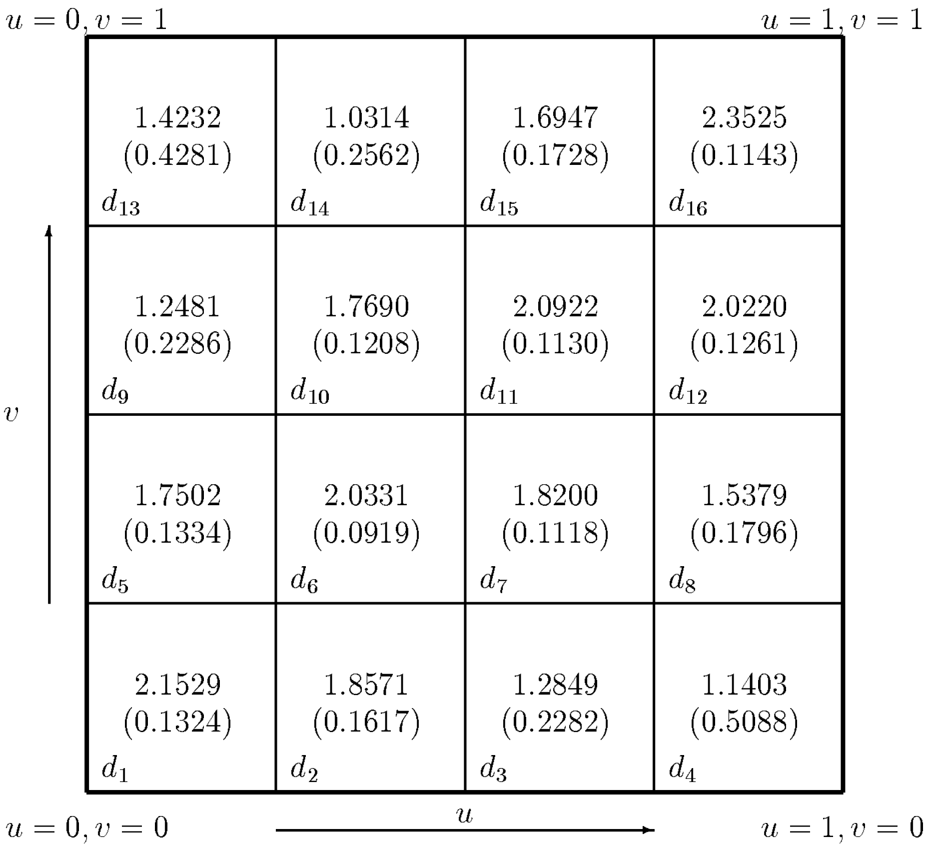

u = 0, v = 1
u = 1, v = 1
1.4232 (0.4281) d13
1.0314 (0.2562) d14
1.6947 (0.1728) d15
2.3525 (0.1143) d16
1.2481 (0.2286) d9
1.7690 (0.1208) d10
2.0922 (0.1130) d11
2.0220 (0.1261) d12
v
1.7502 (0.1334) d5
2.0331 (0.0919) d6
1.8200 (0.1118) d7
1.5379 (0.1796) d8
2.1529 (0.1324) d1
1.8571 (0.1617) d2
1.2849 (0.2282) d3
1.1403 (0.5088) d4
u
u = 0, v = 0
u = 1, v = 0

Figure 7.