Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Journal of International Money and Finance 25 (2006) 827–853
Journal of International Money and Finance
www.elsevier.com/locate/jimf

The Copula-GARCH model of conditional dependencies: An international stock market application

Eric Jondeau, Michael Rockinger*

Swiss Finance Institute and University of Lausanne, Lausanne, Switzerland

Abstract

Modeling the dependency between stock market returns is a difficult task when returns follow a complicated dynamics. When returns are non-normal, it is often simply impossible to specify the multivariate distribution relating two or more return series. In this context, we propose a new methodology based on copula functions, which consists in estimating first the univariate distributions and then the joining distribution. In such a context, the dependency parameter can easily be rendered conditional and time varying. We apply this methodology to the daily returns of four major stock markets. Our results suggest that conditional dependency depends on past realizations for European market pairs only. For these markets, dependency is found to be more widely affected when returns move in the same direction than when they move in opposite directions. Modeling the dynamics of the dependency parameter also suggests that dependency is higher and more persistent between European stock markets.

JEL classification: C51; F37; G11

Keywords: Stock indices; International correlation; Dependency; GARCH model; Skewed Student-*t* distribution; Copula function

* Corresponding author. University of Lausanne, Ecole des HEC, Department of Finance and Insurance, 1015 Lausanne, Switzerland. Tel.: +41 2 16 92 33 48; fax: +41 2 16 92 34 35.

E-mail addresses: eric.jondeau@unil.ch (E. Jondeau), michael.rockinger@unil.ch (M. Rockinger).

0261-5606/$ - see front matter © 2006 Published by Elsevier Ltd.
doi:10.1016/j.jimonfin.2006.04.007

1. Introduction

An abundant literature has investigated how the correlation between stock market returns varies when markets become agitated. In a multivariate GARCH framework, for instance, Hamao et al. (1990), Susmel and Engle (1994), and Bekaert and Harvey (1995) have measured the interdependence of returns and volatilities across stock markets. More specifically, Longin and Solnik (1995) have tested the hypothesis of a constant conditional correlation between a large number of stock markets. They found that correlation generally increases in periods of high-volatility of the U.S. market. In addition, in a similar context, tests of a constant correlation have been proposed by Bera and Kim (2002) and Tse (2000). Recent contributions by Kroner and Ng (1998), Engle and Sheppard (2001), Engle (2002), and Tse and Tsui (2002) have developed GARCH models with time-varying covariances or correlations. As an alternative approach, Ramchand and Susmel (1998) and Ang and Bekaert (2002) have estimated a multivariate Markov-switching model and tested the hypothesis of a constant international conditional correlation between stock markets. They obtained that correlation is generally higher in the high-volatility regime than in the low-volatility regime.

In this context, an important issue is how dependency between stock markets can be measured when returns are non-normal. In the GARCH framework, some recent papers have focused on multivariate distributions which allow for asymmetry as well as fat tails. For instance, multivariate skewed distributions, and in particular the skewed Student-*t* distribution, have been studied by Sahu et al. (2001) and Bauwens and Laurent (2002). In addition, in the Markov-switching context, Chesnay and Jondeau (2001) have tested for a constant correlation between stock returns, while allowing for Student-*t* innovations.[1] For most types of univariate distributions, however, it is simply impossible to specify a multivariate extension that would allow the dependency structure to be captured. In this paper, we present a new methodology to measure conditional dependency in a GARCH context. Our methodology builds on so-called "copula" functions. These functions provide an interesting tool to model a multivariate distribution when only marginal distributions are known. Such an approach is, thus, particularly useful in situations where multivariate normality does not hold. An additional interesting feature of copulas is the ease with which the associated dependency parameter can be conditioned and rendered time varying, even when complicated marginal dynamics are estimated.

We use this methodology to investigate the impact of certain joint stock return realizations on the subsequent dependency of international markets. Many univariate models have been proposed to specify the dynamics of returns. However, given the focus of this work, we draw on recent advances in the modeling of conditional returns that allow second, third, and fourth moments to vary over time. Our univariate model builds on Hansen's (1994) seminal paper. In that paper, a so-called skewed Student-*t* distribution is derived. This distribution allows for a control of asymmetry and fat-tailedness. By rendering these characteristics conditional, it is possible to obtain time-varying higher moments.[2] This model, therefore, extends Engle's (1982) ARCH and Bollerslev's (1986) GARCH models. In an extension to Hansen (1994),

[1] Some papers also considered how correlation varies when stock market indices are simultaneously affected by very large (positive or negative) fluctuations. Longin and Solnik (2001), using extreme value theory, found that dependency increases more during downside movements than during upside movements. Poon et al. (2004) adopted an alternative statistical framework to test conditional dependency between extreme returns and showed that such a tail dependency may have been overstated once the time-variability of volatility is accounted for.

[2] Higher moments refer to the standardized third and fourth central moments.

Jondeau and Rockinger (2003a,b) determine the expression of skewness and kurtosis of the skewed Student-*t* distribution and show how the cumulative distribution function (cdf) and its inverse can be computed. They show how to simulate data distributed as a skewed Student-*t* distribution and discuss how to parametrize time-varying higher moments.

We then consider two alternative copula functions which have different characteristics in terms of tail dependency: the Gaussian copula that does not allow any dependency in the tails of the distribution, and the Student-*t* copula that is able to capture such a tail dependency. Finally, we propose several ways to condition the dependency parameter with past realizations. It is thus possible to test several hypotheses of the way in which dependency varies during turbulent periods.

In the empirical part of the paper, we investigate the dependency structure between daily returns of major stock market indices over 20 years. As a preliminary step, we provide evidence that the skewed Student-*t* distribution with time-varying higher moments fits very well the univariate behavior of the data. Then, we check that the Student-*t* copula is able to capture the dependency structure between market returns. Further scrutiny of the data reveals that the dependency between European markets increases, subsequent to movements in the same direction, either positively or negatively. Furthermore, the strong persistence in the dynamics of the dependency structure is found to reflect a shift, over the sample period, in the dependency parameter. This parameter has increased from about 0.3 over the 1980s to about 0.6 over the next decade. Such a pattern is not found to hold for the dependency structure between the U.S. market and the European markets.

The remainder of the paper is organized as follows. In Section 2, we first introduce our univariate model which allows for volatility, skewness, and kurtosis, to vary over time. In Section 3, we introduce copula functions and describe the copulas used in the empirical application. We also describe how the dependency parameter may vary over time. In Section 4, we present the data and discuss our empirical results. Section 5 summarizes our results and concludes.

2. A model for the marginal distributions

It is well known that the residuals obtained from a GARCH model are generally non-normal. This observation has led to the introduction of fat-tailed distributions for innovations. For instance, Nelson (1991) considered the generalized error distribution, while Bollerslev and Wooldridge (1992) focused on Student-*t* innovations. Engle and Gonzalez-Rivera (1991) modeled residuals non-parametrically. Even though these contributions recognize the fact that errors have fat tails, they generally do not render higher moments time varying, i.e., parameters of the error distribution are assumed to be constant over time. Our margin model builds on Hansen (1994).

2.1. Hansen's skewed Student-t distribution

Hansen (1994) was the first to propose a GARCH model, in which the first four moments are conditional and time varying. For the conditional mean and volatility, he built on the usual GARCH model. To control higher moments, he constructed a new density with which he modeled the GARCH residuals. The new density is a generalization of the Student-*t* distribution while maintaining the assumption of a zero mean and unit variance. The conditioning is obtained by defining parameters as functions of past realizations.

Some extensions to this seminal contribution may be found in Theodossiou (1998) and Jondeau and Rockinger (2003a).[3]

Hansen's skewed Student-*t* distribution is defined by

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

where

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

and η and λ denote the degree-of-freedom parameter and the asymmetry parameter, respectively. If a random variable Z has the density $d(z;\eta,\lambda)$, we will write $Z\sim \text{ST}(\eta,\lambda)$.

Additional useful results are provided in the Appendix A. In particular, we characterize the domain of definition for the distribution parameters η and λ, give formulas relating higher moments to η and λ, and we describe how the cdf of Hansen's skewed Student-*t* distribution can be computed. This computation is necessary for the evaluation of the likelihood of the copula function.

2.2. A GARCH model with conditional skewness and kurtosis

Let the returns of a given asset be given by $\{r_t\}$, $t=1,\ldots,T$. Hansen's margin model with time-varying volatility, skewness, and kurtosis, is defined by

$$r_t=\mu_t+\varepsilon_t, \tag{2}$$

$$\varepsilon_t=\sigma_t z_t, \tag{3}$$

$$\sigma_t^2=a_0+b_0^+\left(\varepsilon_{t-1}^+\right)^2+b_0^-\left(\varepsilon_{t-1}^-\right)^2+c_0\sigma_{t-1}^2, \tag{4}$$

$$z_t\sim \text{ST}(\eta_t,\lambda_t). \tag{5}$$

Eq. (2) decomposes the return of time t into a conditional mean, μ_t, and an innovation, ε_t. The conditional mean is modeled with 10 lags of r_t and day-of-the-week dummies. Eq. (3) then defines this innovation as the product between conditional volatility, σ_t, and a residual, z_t. Eq. (4) determines the dynamics of volatility. We use the notation $\varepsilon_t^+=\max(\varepsilon_t,0)$ and $\varepsilon_t^-=\max(-\varepsilon_t,0)$. For positivity and stationarity of the volatility process to be guaranteed, parameters are assumed to satisfy the following constraints: $a_0>0$, b_0^+, b_0^-, $c_0\geq 0$, and

[3] Harvey and Siddique (1999) have proposed an alternative specification, based on a non-central Student-*t* distribution, in which higher moments also vary over time. This distribution is designed so that skewness depends on the non-centrality parameter and the degree-of-freedom parameter. Note also that the specification of the skewed Student-*t* distribution adopted by Lambert and Laurent (2002) corresponds to the distribution proposed by Hansen, but with asymmetry parametrized in a different way.

$c_0 + \left(b_0^+ + b_0^-\right)/2 < 1$. Such a specification has been suggested by Glosten et al. (1993). Eq. (5) specifies that residuals follow a skewed Student-t distribution with time-varying parameters η_t and λ_t.

Many specifications could be used to describe the dynamics of η_t and λ_t. To ensure that they remain within their authorized range, we consider an unrestricted dynamic that we constrain via a logistic map.[4] The type of functional specification that should be retained is discussed in Jondeau and Rockinger (2003a). The general unrestricted model that we estimate is given by

$$\tilde{\eta}_t = a_1 + b_1^+ \varepsilon_{t-1}^+ + b_1^- \varepsilon_{t-1}^- + c_1 \tilde{\eta}_{t-1} \tag{6}$$

$$\tilde{\lambda}_t = a_2 + b_2^+ \varepsilon_{t-1}^+ + b_2^- \varepsilon_{t-1}^- + c_2 \tilde{\lambda}_{t-1}. \tag{7}$$

We map this dynamic into the authorized domain with $\eta_t = g_{]L_\eta,U_\eta[}(\tilde{\eta}_t)$ and $\lambda_t = g_{]L_\lambda,U_\lambda[}(\tilde{\lambda}_t)$.

Several encompassing restrictions of this general specification are tested in the empirical section of the paper. In particular, we test, within the class of GARCH models of volatility, the following restrictions: a Gaussian conditional distribution, a standard Student-t distribution, and a skewed Student-t distribution with constant skewness and kurtosis. We will see that the most general model cannot be rejected for all the stock indices considered.

3. Copula distribution functions

3.1. The copula function

Consider two random variables X_1 and X_2 with marginal cdfs $F_i(x_i) = \Pr[X_i \le x_i]$, $i = 1,2$. The joint cdf is denoted $H(x_1,x_2) = \Pr[X_1 \le x_1, X_2 \le x_2]$. All cdfs $F_i(\cdot)$ and $H(\cdot,\cdot)$ range in the interval [0,1]. In some cases, a multivariate distribution exists, so that the function $H(\cdot,\cdot)$ has an explicit expression. One such case is the multivariate normal distribution. In many cases, however, the margins $F_i(\cdot)$ are relatively easy to describe, while an explicit expression of the joint distribution $H(\cdot,\cdot)$ may be difficult to obtain.[5]

In such a context, copulas can be used to link margins into a multivariate distribution function. The copula function extends the concept of multivariate distribution for random variables which are defined over [0,1]. This property allows to define a multivariate distribution in terms of margins $F_i(x_i)$ instead of realizations x_i. Then, as highlighted in Sklar's theorem (see the Appendix A), one has the equality between the cdf H defined over realizations of the random variables x_i and the copula function C defined over margins $F_i(x_i)$, so that $H(x_1,x_2) = C(F_1(x_1),F_2(x_2))$.

We now describe the two copula functions used in our empirical application. For notational convenience, set $u_i \equiv F_i(x_i)$. The Gaussian copula is defined by the cdf

$$C(u_1,u_2;\rho) = \Phi_\rho\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2)\right),$$

and the density by

[4] The logistic map, $g_{]L,U[}(x) = L + (U - L)(1 + e^{-x})^{-1}$ maps $\mathcal{R}$ into the interval $]L,U[$. In practice, we use the bounds $L_\eta = 2$, $U_\eta = 30$ for η and $L_\lambda = -1$, $U_\lambda = 1$ for λ.

[5] It may be argued that multivariate extensions of the skewed Student-t distribution exist (see, in particular, Bauwens and Laurent, 2002). In fact, the difficulty comes in this case from the joint estimation of two or more distributions, each involving a large number of unknown parameters.

$$c(u_1,u_2;\rho)=\frac{1}{\sqrt{1-\rho^2}}\exp\left(-\frac{1}{2}\psi'\left(R^{-1}-I_2\right)\psi\right),$$

where $\psi=\left(\Phi^{-1}(u_1),\Phi^{-1}(u_2)\right)'$. The matrix R is the (2,2) correlation matrix with ρ as dependency measure between X_1 and X_2. Φ_ρ is the bivariate standardized Gaussian cdf with correlation ρ, $-1<\rho<1$. The letter Φ represents the univariate standardized Gaussian cdf.

Similarly, the Student-t copula is defined by

$$C(u_1,u_2;\rho,n)=T_{n,\rho}\left(t_n^{-1}(u_1),t_n^{-1}(u_2)\right),$$

and its associated density is

$$c(u_1,u_2;\rho,n)=\frac{1}{\sqrt{1-\rho^2}}\frac{\Gamma\left(\frac{n+2}{2}\right)\Gamma\left(\frac{n}{2}\right)}{\left(\Gamma\left(\frac{n+1}{2}\right)\right)^2}\frac{\left(1+\frac{1}{n}\psi'\Omega^{-1}\psi\right)^{-[(n+2)/2]}}{\prod_{i=1}^{2}\left(1+\frac{1}{n}\psi_i^2\right)^{-[(n+1)/2]}},$$

where $\psi=\left(t_n^{-1}(u_1),t_n^{-1}(u_2)\right)'$. $T_{n,\rho}$ is the bivariate Student-t cdf with degree-of-freedom parameter n and correlation $\rho\in[-1,1]$, while t_n is the univariate Student-t cdf with degree-of-freedom parameter n.

These two copula functions have different characteristics in terms of tail dependence. The Gaussian copula does not have tail dependence, while the Student-t copula has got it (see, for instance, Embrechts et al., 2003). Such a difference is likely to have important consequences on the modeling of the dependency parameter. Indeed, Longin and Solnik (2001) have shown, using an alternative methodology, that the correlation between market returns is higher in case of extreme events.[6] Finally, the two copula functions under study are symmetric. Therefore, when the dependency parameter is assumed to be constant, large joint positive realizations have the same probability of occurrence as large joint negative realizations. In Section 3.2, we relax this assumption by allowing the dependency parameter to be conditional on past realizations.

Both Gaussian and Student-t copulas belong to the elliptical-copula family. Thus, when margins are elliptical distributions with finite variances, ρ is just the usual linear correlation coefficient and can be estimated using a linear correlation estimator (see Embrechts et al., 1999). In the following, however, we provide evidence that margins can be well approximated by the skewed Student-t distribution, which does not belong to the elliptical-distribution family. It follows that ρ is not the linear Pearson's correlation and needs to be estimated via maximum-likelihood.

3.2. Alternative specifications for conditional dependency

Let us consider a sample $\{z_{1t},z_{2t}\}$, $t=1,\ldots,T$. It is assumed that z_{it} gets generated by a continuous cdf $F_i(\cdot;\theta_i)$, where θ_i represents the vector of unknown parameters pertaining to the marginal distribution of Z_{it}, $i=1,2$. In our context, z_{it} is the residual of the univariate GARCH model presented in Section 2.2.

[6] Poon et al. (2004) have obtained that much of this increase in dependency may be explained by changes in volatility.

The key observation is that the copula depends on parameters that can be easily conditioned. We define ρ_t as the value taken by the dependency parameter at time t.

For the Student-t copula, the degree-of-freedom parameter n may be conditioned as well. Several different specifications of the dependency parameter are possible in our context. As a first approach, we follow Gouriéroux and Monfort (1992) and adopt a semi-parametric specification in which ρ_t depends 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, with parameter ρ_t held constant for each element of the grid. More precisely, we consider an unrestricted specification:

$$\rho_t = \sum_{j=1}^{16} d_j I\left[(z_{1t-1}, z_{2t-1}) \in \mathcal{A}_j\right], \tag{8}$$

where $\mathcal{A}_j$ is the jth element of the unit square grid and $d_j \in [-1,1]$. To each parameter d_j, an area $\mathcal{A}_j$ is associated. For instance, $\mathcal{A}_1 = [0, p_1[\times [0, q_1[$ and $\mathcal{A}_2 = [p_1, p_2[\times [0, q_1[$.[7] The choice of 16 subintervals is somewhat arbitrary. It should be noticed, however, that it has the advantage of providing an easy testing of several conjectures concerning the impact of past joint returns on subsequent dependency while still allowing for a large number of observations per area. In the empirical section, we test several hypotheses of interest on the parameters d_j.

It should be recognized that such a specification is not able to capture persistence in ρ_t. Therefore, we first consider a time-varying correlation (TVC) approach, as proposed by Tse and Tsui (2002) in their modeling of the Pearson's correlation in a GARCH context. The dependency parameter ρ_t is assumed to be driven by the following model:

$$\rho_t = (1 - \alpha - \beta)\rho + \alpha \xi_{t-1} + \beta \rho_{t-1}, \tag{9}$$

where

$$\xi_t = \frac{\sum_{i=0}^{m-1} z_{1t-i} z_{2t-i}}{\sqrt{\sum_{i=0}^{m-1} z_{1t-i}^2 \sum_{i=0}^{m-1} z_{2t-i}^2}}$$

represents the correlation between the residuals over the recent period, with $m \geq 2$. For stationarity to be guaranteed, the following constraints are imposed: $0 \leq \alpha$, $\beta \leq 1$, $\alpha + \beta \leq 1$, and $-1 \leq \rho \leq 1$.

Our empirical analysis, however, reveals a very large persistence in ρ_t, with $\alpha + \beta$ very close to 1 in most cases. This result suggests that the TVC model may be inappropriate and that the long-memory feature may be in fact the consequence of a model with large but infrequent breaks (see Lamoureux and Lastrapes, 1990; Diebold and Inoue, 1999; or Gouriéroux and Jasiak, 2001).[8] This approach has been followed, among others, by Ramchand and Susmel

[7] Fig. 2 illustrates the position of the areas d_js. In the figure, we have set equally spaced threshold levels, i.e., p_1, p_2, and p_3 take the values 0.25, 0.5, and 0.75, respectively. The same for q_1, q_2, and q_3, respectively. In the empirical part of the paper, we use as thresholds the values 0.15, 0.5, and 0.85. The reason for this choice is that we want to focus on rather large values. If we had used 0.25, 0.5, and 0.75, the results would have been similar, however.

[8] We are grateful to a referee for suggesting this interpretation.

(1998), Chesnay and Jondeau (2001), and Ang and Bekaert (2002). These authors have obtained, for several market pairs, evidence of the presence of a high-volatility/high-correlation regime and a low-volatility/low-correlation regime. Thus, we also consider, as an alternative approach, a model in which parameters of the Student-t copula are driven by the following model

$$\rho_t = \rho_0 S_t + \rho_1 (1 - S_t), \tag{10}$$

$$n_t = n_0 S_t + n_1 (1 - S_t), \tag{11}$$

where S_t denotes the unobserved regime of the system at time t. S_t is assumed to follow a two-state Markov process, with transition probability matrix given by

$$\begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix}$$

with

$$p = \Pr[S_t = 0 | S_{t-1} = 0],$$

$$q = \Pr[S_t = 1 | S_{t-1} = 1].$$

Note that, in this model, we do not assume that univariate characteristics of returns should also shift. Rather, only parameters pertaining to the dependence structure are driven by the Markov-switching model.

Quasi maximum-likelihood estimation of this model can be easily obtained using the approach developed by Hamilton (1989) and Gray (1995). For the degree-of-freedom parameter, we investigated several hypotheses. In particular, we tested whether it is regime independent ($n_0 = n_1$) or infinite, so that the Gaussian copula would prevail for a given regime. We also investigated time-variation in transition probabilities, along the lines of Schaller and van Norden (1997). We tested specifications in which transition probabilities are allowed to depend on past volatilities (σ_{1t-1} and σ_{2t-1}) as well as on correlation between the residuals over the recent period (ξ_{t-1}). We were unable, however, to obtain a significant time-variation in the probabilities using such variables.

3.3. Estimation

We now assume that the copula function depends on a set of unknown parameters through the function $\Theta(z_{1t-1}, z_{2t-1}; \theta_c)$. We have $\theta_c = (d_1, \ldots, d_{16}, n)'$ for the semi-parametric specification, $\theta_c = (\rho, \alpha, \beta, n)'$ for the TVC specification, and $\theta_c = (\rho_0, \rho_1, n_0, n_1, p, q)'$ for the Markov-switching model. We also denote f_i as the marginal density of z_{it}. In the context of a skewed Student-t marginal distribution, as presented in Section 2, this density is simply defined by $f_i(z_{it}; \theta_i) = d(z_{it}; \theta_i)$, $i = 1,2$. We set $\theta = \left(\theta_1', \theta_2', \theta_c'\right)'$ the vector of all parameters to be estimated. Consequently, the log-likelihood of a sample is given by

$$\ell(\theta) = \sum_{t=1}^{T} \ell_t(\theta), \tag{12}$$

where

$$\ell_t(\theta) = \ln c(F_1(z_{1t};\theta_1), F_2(z_{2t};\theta_2); \Theta(z_{1t-1}, z_{2t-1};\theta_c)) + \sum_{i=1}^{2} \ln f_i(z_{it};\theta_i).$$

Maximum-likelihood estimation involves maximizing the log-likelihood function (12) simultaneously over all parameters, yielding parameter estimates denoted $\widehat{\theta}_{\mathrm{ML}} = (\widehat{\theta}'_1, \widehat{\theta}'_2, \widehat{\theta}'_c)'$, such that

$$\widehat{\theta}_{\mathrm{ML}} = \arg\max \ell(\theta).$$

In some applications, however, the ML estimation method may be difficult to implement, because of a large number of unknown parameters or of the complexity of the model.[9] In such a case, it may be necessary to adopt a two-step ML procedure, also called inference functions for margins. This approach, which has been introduced by Shih and Louis (1995) and Joe and Xu (1996), can be viewed as the ML estimation of the dependence structure given the estimated margins. First, parameters pertaining to the marginal distributions are estimated separately:

$$\tilde{\theta}_i \in \arg\max \sum_{t=1}^{T} \ln f_i(z_{it};\theta_i) \quad i = 1, 2. \tag{13}$$

Second, parameters pertaining to the copula function are estimated by solving the following equation:

$$\tilde{\theta}_c \in \arg\max \sum_{t=1}^{T} \ln c(F_1(z_{1t};\tilde{\theta}_1),\ F_2(z_{2t};\tilde{\theta}_2);\ \Theta(z_{1t-1}, z_{2t-1};\theta_c)).$$

Patton (2006) has shown that this two-step estimation yields asymptotically efficient and normal parameter estimates. If θ_0 denotes the true value of the parameter vector, the asymptotic distribution of $\tilde{\theta}_{\mathrm{TS}} = (\tilde{\theta}'_1, \tilde{\theta}'_2, \tilde{\theta}'_c)'$ is given by

$$\sqrt{T}(\tilde{\theta}_{\mathrm{TS}} - \theta_0) \to N(0, \Omega),$$

where the asymptotic covariance matrix Ω may be estimated by the robust estimator $\widehat{\Omega} = \widehat{M}^{-1}\widehat{V}\widehat{M}^{-1}$, with

$$\widehat{M} = -\sum_{t=1}^{T} \frac{\partial^2 \ell_t(\tilde{\theta})}{\partial\theta\,\partial\theta'},$$

$$\widehat{V} = \sum_{t=1}^{T} \frac{\partial \ell_t(\tilde{\theta})}{\partial\theta} \sum_{t=1}^{T} \frac{\partial \ell_t(\tilde{\theta})}{\partial\theta}'.$$

[9] 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. Therefore, only numerical gradients may be computable, implying a dramatic slowing down of the numerical procedure.

4. Empirical results

4.1. The data

We investigate the interactions between four major stock indices. The labels are SP for the S&P 500, FTSE for the Financial Times 100 stock index, DAX for the Deutsche Aktien Index, and CAC for the French Cotation Automatique Continue index. Our sample covers the period from January 1, 1980 to December 29, 2000.

All the data are from Datastream, sampled at a daily frequency. To eliminate spurious correlation generated by holidays, we eliminated those observations when a holiday occurred at least for one country from the database. This reduced the sample from 5479 observations to 4578. Note that such an observation would not affect the dependency between stock markets during extreme events. Yet, it would affect the estimation of the return marginal distribution and, subsequently, the estimation of the distribution of the copula. In particular, the estimation of the copula would be distorted to account for the excessive occurrence of null returns in the distribution. To take into account the fact that international markets have different trading hours, we use once lagged U.S. returns, although this does not significantly affect the correlation with European markets (because trading times are partially overlapping). Preliminary estimations also revealed that the crash of October 1987 was of such importance that the standard errors of our model would be very much influenced by this event. For the S&P, on that date, the index dropped by −22%, while the second largest drop was −9% only. For this reason, we eliminated the data between October 17 and 24. This further reduces the sample to a total of 4572 observations.

Table 1 provides summary statistics on stock market returns. Returns (r_t) are defined as $100 \times \ln(P_t/P_{t-1})$, where P_t is the value of the index at time t. Statistics are computed after holidays have been removed from the time series. Therefore, the number of observations is the same one for all markets, and the series do not contain days when a market was closed. We begin with the serial dependency of returns. The LM(K) statistic tests whether the squared return is serially correlated up to lag K. This statistic clearly indicates that ARCH effects are likely to be found in all market returns. Also, when considering the Ljung–Box statistics, QW(K), after correction for heteroskedasticity, we obtain that returns are serially correlated for all the retained indices.

We also consider the unconditional moments of the various series, with standard errors computed using a GMM-based procedure. We notice that for all series skewness, Sk, is negative. Moreover, excess kurtosis, XKu, is significant for all return series. This indicates that the empirical distributions of returns display fatter tails than the Gaussian distribution. The Wald statistic of the joint test of significance of skewness and excess kurtosis corroborates this finding.[10]

Finally, the unconditional correlation matrix indicates that a rather large dependency between market returns is expected. The correlation is the smallest between the SP and the CAC, and the largest between the DAX and the CAC.

[10] When the 1987 crash is not removed, the SP distribution is characterized by a very strong asymmetry (with a skewness equal to −2.55) and fat tails (with an excess kurtosis as high as 57). Yet, due to uncertainty around higher-moment point estimates, the Wald test would not reject normality.

Table 1
Summary statistics on daily returns

	SP	FTSE	DAX	CAC
Mean	0.049[a]	0.044[a]	0.044[b]	0.041[b]
s.e.	(0.014)	(0.014)	(0.018)	(0.019)
Std	0.963[a]	0.923[a]	1.211[a]	1.190[a]
s.e.	(0.026)	(0.020)	(0.042)	(0.042)
Sk	−0.399[b]	−0.164	−0.720[b]	−0.683[b]
s.e.	(0.186)	(0.103)	(0.299)	(0.362)
XKu	5.147[a]	2.082[a]	8.604[a]	8.369[b]
s.e.	(1.389)	(0.525)	(3.132)	(3.550)
Wald stat.	17.454[a]	15.837[a]	7.581[b]	6.022[b]
p-Values	(0.000)	(0.000)	(0.023)	(0.049)
LM(1)	82.745[a]	103.585[a]	222.750[a]	113.279[a]
p-Values	(0.000)	(0.000)	(0.000)	(0.000)
LM(5)	188.927[a]	315.646[a]	318.219[a]	345.317[a]
p-Values	(0.000)	(0.000)	(0.000)	(0.000)
QW(5)	13.187[b]	17.264[a]	10.161[c]	10.404[c]
p-Values	(0.022)	(0.004)	(0.071)	(0.065)
QW(10)	22.526[b]	25.574[a]	18.515[b]	21.470[b]
p-Values	(0.013)	(0.004)	(0.047)	(0.018)
Correlation matrix				
SP	1	0.272	0.317	0.269
FTSE	0.272	1	0.475	0.524
DAX	0.317	0.475	1	0.554
CAC	0.269	0.524	0.554	1

This table reports summary statistics on daily stock market returns. Mean, Std, Sk, and XKu denote the mean, the standard deviation, the skewness, and the excess kurtosis of returns, respectively. Standard errors are computed using a GMM-based procedure. Wald stat. is the Wald statistics which tests the null hypothesis that skewness and excess kurtosis are jointly equal to zero. It is distributed, under the null, as a $\chi^2(2)$. The LM(K) statistics for heteroskedasticity is obtained by regressing squared returns on K lags. QW(K) is the Ljung–Box statistics for serial correlation, corrected for heteroskedasticity, computed with K lags. Since international markets have different trading hours, the correlation matrix is computed using once lagged U.S. returns. In this table as well as all the following ones, significance is denoted by superscripts at the 1% ([a]), 5% ([b]), and 10% ([c]) levels.

4.2. Estimation of the marginal model

In a preliminary step, we consider several restrictions of the general marginal model as possible candidates for adjusting the empirical return distribution. Table 2 reports the goodness-of-fit tests for these distributions. For this purpose, we follow Diebold et al. (1998), henceforth DGT, who suggested that, if a marginal distribution is correctly specified, the margin u_{it} should be iid Uniform(0,1). The test is performed in two steps. First, we evaluate whether u_{it} is serially correlated. Hence, we examine the serial correlation of $(u_{it} - \bar{u}_i)^k$, for $k = 1,\ldots,4$ by regressing $(u_{it} - \bar{u}_i)^k$ on 20 own lags.[11] The LM test statistic is defined as $(T-20)R^2$, where R^2 is the coefficient of determination of the regression, and is distributed as a $\chi^2(20)$ under the null. In Table 2, these tests, labeled DGT-AR$^{(k)}$, generally do not reject the null hypothesis. In particular, even when residuals are assumed to be normal, the first four moments are found to be

[11] Zero correlation is equivalent to independence, only under normality. The correlogram is, therefore, only suggestive of possible independence.

Table 2
Goodness-of-fit test statistics

	SP	FTSE	DAX	CAC
Panel A: Normal distribution				
DGT-AR$^{(1)}$	27.233	17.756	31.781[b]	21.162
DGT-AR$^{(2)}$	30.816[c]	17.882	27.983	33.140[b]
DGT-AR$^{(3)}$	23.947	23.661	28.945[c]	17.501
DGT-AR$^{(4)}$	27.856	13.532	16.536	25.317
DGT-H(20)	109.057[a]	36.615[a]	67.215[a]	63.708[a]
ℓL	−5932.938	−5834.696	−6817.531	−6817.807
AIC	2.603	2.560	2.991	2.991
SIC	2.608	2.565	2.996	2.996
Panel B: Student-t distribution				
DGT-AR$^{(1)}$	28.057	17.900	31.441[b]	21.154
DGT-AR$^{(2)}$	29.269[c]	17.464	19.211	21.498
DGT-AR$^{(3)}$	25.116	22.810	29.855[c]	17.409
DGT-AR$^{(4)}$	27.549	13.176	9.764	14.645
DGT-H(20)	26.014	38.579[a]	45.900[a]	14.125
ℓL	−5791.183	−5802.548	−6623.191	−6645.547
LRT(1)	283.511	64.297	388.680	344.519
p-Values	(0.000)	(0.000)	(0.000)	(0.000)
AIC	2.541	2.546	2.906	2.916
SIC	2.548	2.553	2.913	2.923
Panel C: Skewed Student-t distribution				
DGT-AR$^{(1)}$	27.978	18.099	31.029[c]	21.209
DGT-AR$^{(2)}$	29.351[c]	17.076	21.219	21.741
DGT-AR$^{(3)}$	24.726	22.538	27.923	17.147
DGT-AR$^{(4)}$	27.675	12.696	11.983	14.788
DGT-H(20)	26.426	24.918	28.566[c]	13.818
ℓL	−5790.403	−5796.170	−6612.153	−6641.937
LRT(1)	1.559	12.756	22.076	7.220
p-Values	(0.212)	(0.000)	(0.000)	(0.007)
AIC	2.541	2.544	2.901	2.915
SIC	2.550	2.552	2.910	2.923
Panel D: Skewed Student-t distribution with time-varying parameters				
DGT-AR$^{(1)}$	30.934[c]	19.201	32.978[b]	24.599
DGT-AR$^{(2)}$	22.487	15.327	14.356	19.027
DGT-AR$^{(3)}$	24.877	20.815	28.561[c]	17.422
DGT-AR$^{(4)}$	20.840	10.752	7.188	13.688
DGT-H(20)	18.886	24.208	30.004[c]	12.152
ℓL	−5770.644	−5776.923	−6601.303	−6634.668
LRT(6)	39.518	38.495	21.700	14.539
p-Values	(0.000)	(0.000)	(0.001)	(0.024)
AIC	2.535	2.538	2.899	2.914
SIC	2.552	2.555	2.916	2.931

This table reports goodness-of-fit statistics for several distributional restrictions of the general univariate model. The first-part of each panel contains the LM test statistics for the null of no serial correlation of the kth centered moments of the u_{it}, labeled DGT-AR$^{(k)}$. Under the null of no autocorrelation of the residuals, the statistics is distributed as a $\chi^2(20)$. The table also reports the test statistics for the null hypothesis that the cdf of residuals is Uniform(0,1), labeled DGT-H(20). Under the null, the statistics is distributed as a $\chi^2(20)$. Finally, the table presents the log-likelihood (ℓL) and the AIC and SIC information criteria. The models A, B, C, and D are encompassing each other. The statistics LRT(p) tests the null hypothesis that the restricted version of a model is not rejected as one moves from one panel to the other. The parameter p is the number of constraints under the null. In this table as well as all the following ones, significance is denoted by superscripts at the 1% ([a]), 5% ([b]), and 10% ([c]) levels.

non-serially correlated for the SP and the FTSE, while the DAX return and the CAC volatility are found to be serially correlated.

Second, we test the null hypothesis that u_{it} is Uniform(0,1). For this purpose, we cut the empirical and theoretical distributions into N bins and test whether the two distributions significantly differ on each bin. An advantage of the approach suggested by DGT is that it permits a graphical representation which can be used to identify areas where the theoretical distribution fails to fit the data. Table 2 reports the test statistics, labeled DGT-H(N), for various distributions with p-values computed with $N-1$ degrees of freedom.[12] We consider the case where $N=20$ bins. Similarly, Fig. 1 displays, for the SP and the CAC indices, the estimates of the density of u_{it} for each bin for the normal distribution and for the skewed Student-t distribution with time-varying skewness and kurtosis. While the table indicates that the normal distribution is strongly rejected for all markets, at any significance level, the figure reveals that the rejection of the normal distribution is attributable to its inability to fit the fat tails of the empirical distribution. The standard Student-t distribution is not rejected for the SP and the CAC, suggesting that asymmetry is not a major feature for these indices. The skewed Student-t distribution is found to fit the data quite well, except for a few number of bins for the DAX. Finally, when skewness and kurtosis are allowed to vary over time, we do not reject the null hypothesis that the theoretical distribution provides a good fit of the empirical distribution for any return series, at the 5% level.

Table 3 presents estimates of the general model in which volatility, skewness, and kurtosis are time varying.[13] We can summarize our empirical evidence for margins as follows. First, a negative return has a stronger effect on subsequent volatility than a positive return of the same magnitude. This result is consistent with the well-known leverage effect, documented by Campbell and Hentschel (1992), and Glosten et al. (1993).

Second, the impact of large returns on the subsequent distribution is measured via λ_t and η_t. The unrestricted dynamics of $\tilde{\lambda}_t$ and $\tilde{\eta}_t$ gets mapped into λ_t and η_t with the logistic map. The dynamics of the degree-of-freedom parameter η_t is found to be rather persistent, except for the FTSE. The significantly negative sign of b_1^+ suggests that, subsequent to large positive realizations, tails thin down. In contrast, we do not obtain significant estimates of b_1^-. This result suggests that a crash is more likely to be followed by a subsequent large return (of either sign) than by a boom.

The asymmetric impact of large returns on the subsequent distribution is measured by the dynamics of λ_t. We find that, in general, past positive returns enlarge the right tail, while past negative returns enlarge the left tail. The effect of positive returns is slightly larger than the effect of negative returns, although not always significantly. Last, the asymmetry parameter is found to be persistent, in particular in European markets.

4.3. Estimation of the multivariate model

4.3.1. Model with constant dependency parameter

Table 4 reports parameter estimates for the Gaussian and Student-t copulas, when the dependency parameter is assumed to be constant over time. For all market pairs, the estimate of the

[12] As shown by Vlaar and Palm (1993), under the null, the correct distribution of the test statistic is bounded between a $\chi^2(N-1)$ and a $\chi^2(N-K-1)$ where K is the number of estimated parameters.

[13] See Jondeau and Rockinger (2003b) for more details on the estimation method.

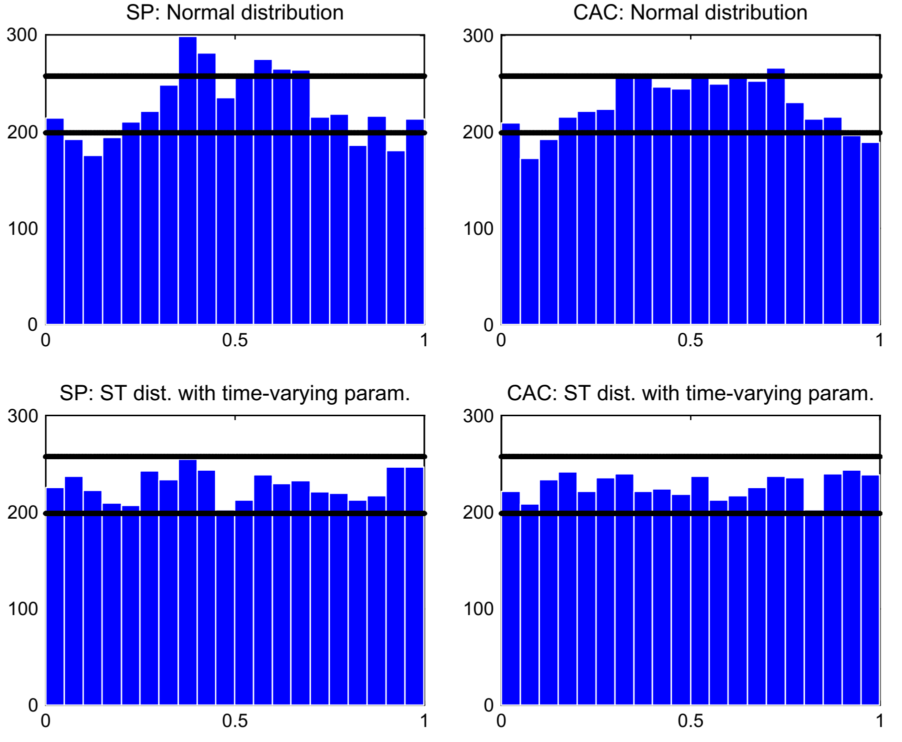

Fig. 1. The estimates of the density of u_{it}, for the normal distribution and the skewed Student-t distribution with time-varying skewness and kurtosis, for the SP and the CAC indices. Horizontal lines correspond to the 95% confidence band.

dependency parameter is found to be positive and strongly significant. As expected, it is statistically and economically much larger between European pairs than between pairs involving the SP. This result reflects the fact that European stock markets are more widely integrated. We also performed a goodness-of-fit test to investigate whether a given copula function is able to fit the dependence structure observed in the data, along the lines of DGT.[14] For all market pairs, we obtain that the Student-t copula fits the data very well, since the null hypothesis is never rejected. In contrast, the Gaussian copula is unable to adjust the dependence structure between European markets.

To provide further insight on the ability of the chosen copulas to fit the data, we also report the log-likelihood and the AIC and SIC information criteria. A LR statistic tests for the null hypothesis that the degree-of-freedom parameter of the Student-t copula is infinite, so that the Student-t copula would reduce to the Gaussian copula. On the basis of the information criteria as well as the LR test, we strongly reject the Gaussian copula specification. As will be shown later on, this result is consistent with the finding that the dependency is stronger in the tails of the distribution than in the middle of the distribution.

[14] The reported DGT-H test statistics are computed by splitting the joint distribution as a (5,5) square and by evaluating for each bin whether the empirical and theoretical distributions significantly differ.

Table 3
Parameter estimates of the model with skewed Student-*t* distribution and time-varying moments

	SP	FTSE	DAX	CAC
Volatility equation				
a_0	0.006[a] (0.002)	0.021[a] (0.005)	0.022[a] (0.006)	0.022[a] (0.006)
b_0^+	0.043[a] (0.010)	0.045[a] (0.010)	0.081[a] (0.016)	0.067[a] (0.013)
b_0^-	0.071[a] (0.012)	0.085[a] (0.012)	0.125[a] (0.018)	0.107[a] (0.014)
c_0	0.941[a] (0.009)	0.910[a] (0.013)	0.887[a] (0.014)	0.900[a] (0.012)
Degree-of-freedom parameter equation				
a_1	−0.510[a] (0.198)	0.390 (0.466)	−0.496[b] (0.242)	−0.429 (0.230)
b_1^+	−0.616[a] (0.149)	−0.994[a] (0.291)	−0.664[a] (0.136)	−0.436[a] (0.126)
b_1^-	0.060 (0.174)	−0.414 (0.358)	0.001 (0.098)	0.326 (0.369)
c_1	0.628[a] (0.130)	0.049 (0.453)	0.422[a] (0.154)	0.626[a] (0.142)
Asymmetry parameter equation				
a_2	−0.084[c] (0.049)	−0.117[a] (0.033)	−0.046 (0.051)	−0.057[c] (0.034)
b_2^+	0.239[a] (0.071)	0.271[a] (0.056)	−0.033 (0.091)	0.083[c] (0.047)
b_2^-	−0.088[c] (0.051)	−0.058 (0.042)	−0.100[b] (0.050)	−0.033 (0.031)
c_2	0.253 (0.178)	0.745[a] (0.075)	0.508[c] (0.304)	0.669[a] (0.159)
Summary statistics				
LM(1)	0.017	0.028	0.052	0.003
p-Values	(0.897)	(0.868)	(0.820)	(0.956)
LM(5)	0.868	1.181	0.717	0.614
p-Values	(0.973)	(0.947)	(0.982)	(0.987)
QW(5)	2.027	2.867	13.543	6.511
p-Values	(0.845)	(0.721)	(0.019)	(0.260)
QW(10)	7.387	4.776	15.371	10.012
p-Values	(0.689)	(0.906)	(0.119)	(0.439)
ℓL	−5770.644	−5776.923	−6601.303	−6634.668

This table reports parameter estimates and residual summary statistics for the model with a skewed Student-*t* distribution and time-varying higher moments. Parameters are defined in Eqs. (4), (6), and (7). Summary statistics include LM(K) statistics for heteroskedasticity, obtained by regressing squared returns on K lags, and the QW(K) statistic for serial correlation, corrected for heteroskedasticity, computed with K lags. ℓL is the sample log-likelihood of the model. In this table as well as all the following ones, significance is denoted by superscripts at the 1% ([a]), 5% ([b]), and 10% ([c]) levels.

Since the Student-*t* copula is preferred over the Gaussian one, we focus in the following on the Student-*t* copula only.[15]

4.3.2. Semi-parametric model for dependency

We first turn to the discussion of the estimation of the model in which the dependency parameter ρ is rendered conditional on past realizations, using the bivariate semi-parametric model (8), as suggested by Gouriéroux and Monfort (1992). Due to the large number of parameters, we do not report the estimates for all market pairs. Instead, we display in Fig. 2 the unit square with parameter estimates of the various d_js and their standard errors, for the SP−FTSE as well as the FTSE−CAC. These two pairs can be viewed as two polar cases. The first pair has

[15] We performed the same estimations with the Gaussian copula and found that the reported results were not altered by the change of copula. Results are available upon request from the authors.

Table 4
Parameter estimates of the copula with constant dependency parameter

	SP–FTSE	SP–DAX	SP–CAC	FTSE–DAX	FTSE–CAC	DAX–CAC
Panel A: Gaussian copula						
ρ	0.240	0.305	0.275	0.406	0.488	0.459
s.e.	(0.014)	(0.013)	(0.013)	(0.012)	(0.010)	(0.011)
DGT-H(25)	14.780	23.944	16.069	41.312[b]	42.424[b]	54.217[a]
ℓL	135.13	220.39	178.40	408.67	620.63	536.47
AIC	−0.059	−0.096	−0.078	−0.178	−0.271	−0.234
SIC	−0.057	−0.095	−0.076	−0.177	−0.270	−0.233
Panel B: Student-t copula						
ρ	0.239	0.304	0.275	0.406	0.494	0.465
s.e.	(0.014)	(0.014)	(0.014)	(0.013)	(0.012)	(0.012)
n	23.773	18.425	12.939	9.320	6.975	6.666
s.e.	(9.056)	(5.380)	(2.794)	(1.522)	(0.851)	(0.779)
DGT-H(25)	16.985	24.228	12.346	22.513	21.147	27.427
ℓL	139.052	227.500	191.420	433.927	667.453	592.095
LRT	7.847	14.215	26.035	50.507	93.655	111.256
p-Values	(0.005)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
AIC	−0.060	−0.099	−0.083	−0.189	−0.291	−0.258
SIC	−0.057	−0.096	−0.080	−0.186	−0.288	−0.255
Empirical ρ	0.222	0.285	0.261	0.383	0.470	0.440

This table reports parameter estimates for the copula functions when the dependency parameter is assumed to be constant over time. Parameters are ρ for the Gaussian copula, and ρ and n for the Student-t copula. Following Diebold et al. (1998), the table also reports the test statistic for the null hypothesis that the cdf of residuals is Uniform(0,1). Under the null, the statistic is distributed as a $\chi^2(25)$. We also report the log-likelihood (ℓL) as well as the AIC and SIC information criteria. For the Student-t copula, LRT is the LR test statistic for the null hypothesis that $1/n = 0$. Finally, empirical ρ is the sample correlation between the margins. In this table as well as all the following ones, significance is denoted by superscripts at the 1% ([a]), 5% ([b]) levels.

a very low dependency parameter ($\rho = 0.24$), while the second one has the largest ρ ($\rho = 0.49$). Inspection of the figures indicates that the extreme diagonal elements for the FTSE–CAC take the values 0.64 and 0.59 that compare with 0.17 and 0.28 for the SP–FTSE. Inspection of the figures, and comparison with the off-diagonal elements show that, subsequent to dissimilar events, i.e., one market goes up, and the other goes down, the likelihood to find a similar event is small. This observation holds also for most market pairs under investigation. We wish now to test the feature of the dependency parameter more formally.

In Table 5, we report the results of the tests for conditional dependency. Since we essentially focus on the parameters along the diagonal, we first present those parameters which correspond to the level of dependency whenever lagged realizations of both markets belong to the same quartile. Then, the table displays Wald test statistics associated with the different hypotheses presented in Section 3.2.

In a first test, we investigate whether the piecewise constant grid of the $u_1 - u_2$ unit square yields significantly different values for the d_js. This test corresponds to the null hypothesis H_1: $d_1 = d_2 = \cdots = d_{16}$, versus inequality for at least one pair of elements. The test statistic is distributed under the null as a $\chi^2(15)$. For market pairs involving the SP, we do not reject the null hypothesis at the 5% significance level. In contrast, the dependency of ρ on past realizations is not rejected for European market pairs. Our evidence is broadly consistent with previous empirical results

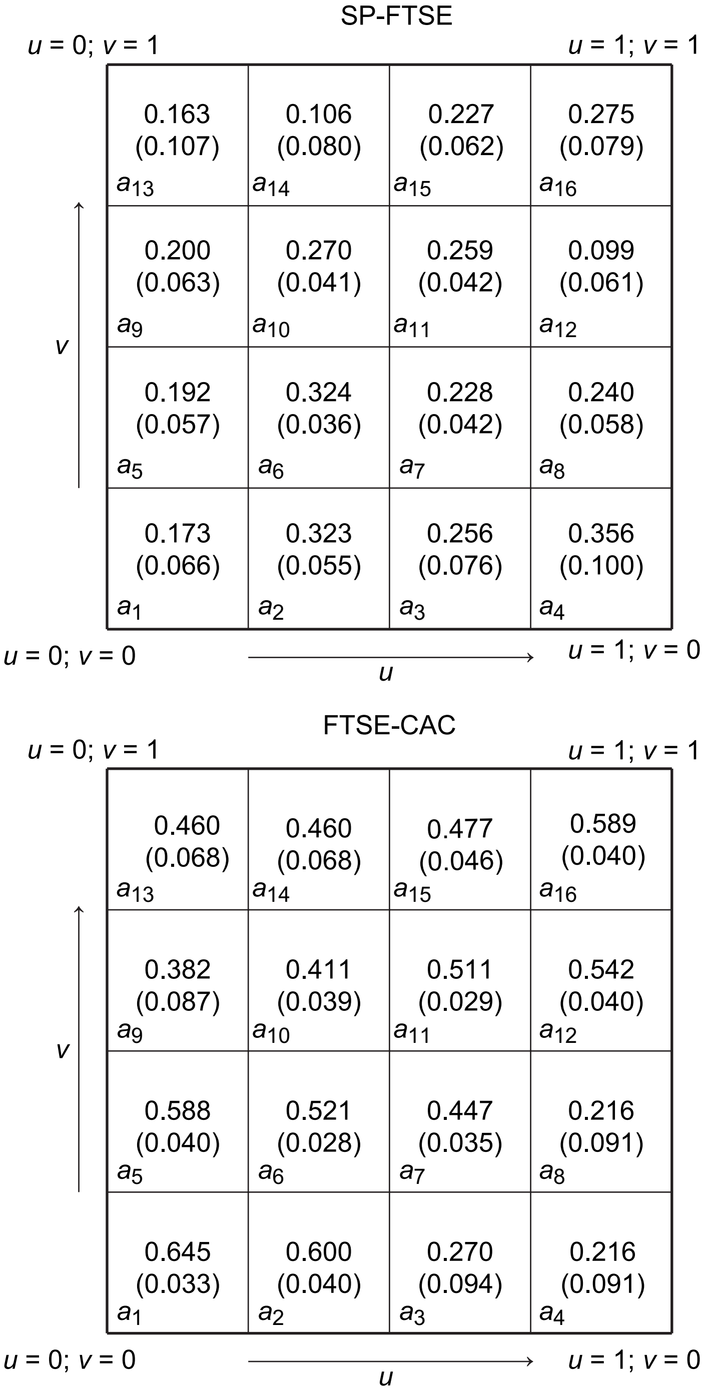

Fig. 2. The unit square with estimates of parameters d_js and their standard errors, for the SP–FTSE and the FTSE–CAC pairs, respectively.

Table 5
Parameter estimates of the Student-*t* copula with dependency parameter driven by the semi-parametric model (8)

	SP−FTSE	SP−DAX	SP−CAC	FTSE−DAX	FTSE−CAC	DAX−CAC
Parameters on the diagonal						
d_1	0.173	0.282	0.280	0.496	0.645	0.553
s.e.	(0.066)	(0.059)	(0.061)	(0.045)	(0.033)	(0.042)
d_6	0.324	0.363	0.353	0.489	0.521	0.530
s.e.	(0.036)	(0.037)	(0.035)	(0.029)	(0.028)	(0.028)
d_{11}	0.259	0.345	0.260	0.431	0.511	0.489
s.e.	(0.042)	(0.036)	(0.039)	(0.033)	(0.029)	(0.031)
d_{16}	0.275	0.206	0.286	0.455	0.589	0.555
s.e.	(0.079)	(0.078)	(0.071)	(0.057)	(0.040)	(0.045)
Composite tests						
H_1: $d_i = d$, $i = 1,\ldots,16$	20.914	12.778	16.734	30.329	72.603	39.508
p-Values	(0.140)	(0.619)	(0.335)	(0.011)	(0.000)	(0.001)
H_2: $d_1 > d_{16}$	−0.991	0.616	−0.004	0.316	1.181	−0.001
p-Values	(0.320)	(0.433)	(0.948)	(0.574)	(0.277)	(0.971)
H_3: $d_1 = d_{16} > d_{11} = d_6$	−1.418	−1.855	−0.557	0.406	3.303	1.156
p-Values	(0.146)	(0.071)	(0.342)	(0.367)	(0.002)	(0.204)
H_4: $d_1 = d_6 = d_{11} = d_{16} > d_3 = d_4 = d_8 = d_9 = d_{13} = d_{14}$	1.534	1.469	1.265	2.026	5.298	4.790
p-Values	(0.123)	(0.136)	(0.179)	(0.051)	(0.000)	(0.000)

This table reports (some) parameter estimates and test statistics for the Student-*t* copula when the dependency parameter ρ depends on the position of past joint realizations in the unit square. Parameters d_1, d_6, d_{11}, and d_{16} correspond to ρ when u_{1t} and u_{2t} belong to the same quartile along the diagonal. Composite tests are described in Section 4.3. For the null hypothesis H_1, the statistics is distributed as a $\chi^2(15)$. For the other hypotheses, the test statistic is distributed as an $N(0,1)$.

obtained, among others, by Bera and Kim (2002) and Tse (2000), who provided formal tests for a constant conditional correlation between stock markets. On one hand, Bera and Kim (2002) found that for most market pairs, the hypothesis of a constant conditional correlation should be strongly rejected. On the other hand, Tse (2000) provided more mixed evidence, suggesting that the test developed by Bera and Kim may have low power under non-normality. Our test procedure does not require any distributional assumption, since marginal models are estimated in a preliminary step. In addition, due to its semi-parametric design, this test does not assume any particular form under the alternative. Note finally that the results of this test are consistent with the evidence provided in the rest of this section, based on a particular parameterization under the alternative.

Next, we consider a test of asymmetry in the persistence of extreme events. Whereas Longin and Solnik (2001), Hartmann et al. (2001), and Poon et al. (2004) focus on the contemporaneous correlations in the tails, finding that correlation is stronger in downside markets than in upside markets, we question whether the dependency between markets is stronger subsequent to downside markets than subsequent to upside markets. Therefore, we compare the magnitude of parameters d_1 and d_{16}. The value d_1 measures the dependency subsequent to downside markets, while d_{16} measures the dependency subsequent to upside markets. We construct a formal test of the hypothesis that H_2: $d_1 = d_{16}$ versus $d_1 > d_{16}$. We notice first that, for most pairs, the estimates of $(d_1 - d_{16})$ take a positive value, meaning that joint downside movements create stronger dependence than corresponding upside movements. Yet, the difference is significant for neither of the market pairs. Therefore, unlike the evidence obtained by Longin and Solnik (2001) using the extreme value theory, we find that a crash and a boom of the same magnitude have generally a similar effect on subsequent correlation, so that it is the increase in volatility which primarily affects correlation. Such an interpretation has been recently put forward by Poon et al. (2004).

To confirm this result, we presently investigate whether large joint returns, be they of positive or negative sign, yield higher dependency than small joint returns. We formulate this hypothesis as a test of H_3: $d_1 = d_6 = d_{11} = d_{16}$ versus $d_1 = d_{16} > d_6 = d_{11}$. Inspection of the test statistics shows that, for most cases, the null cannot be rejected. Longin and Solnik (1995) performed a similar test of asymmetry in correlation between stock markets (see their Table 8). They investigated whether the U.S. market shocks affect correlation in a different extent depending on both the sign and the magnitude of the shocks. They obtained that in most cases large shocks in the U.S. market, either positive or negative, have a similar effect on subsequent correlation than small shocks. Interestingly, we find very similar patterns for pairs involving the SP, since parameters d_1 and d_{16} are smaller than parameters d_6 and d_{11}. In contrast, for European market pairs, we obtain that large shocks have a stronger effect on volatility than small shocks, although the difference is significant for the FTSE–CAC pair only. This suggests that correlation between European markets tends to increase during agitated periods, while this is not likely to be the case between the U.S. market and the European markets.

The last test investigates whether joint variations, be they large or small, have a stronger effect on conditional dependency than opposite variations. Therefore, we test the null hypothesis H_4: $d_1 = d_6 = d_{11} = d_{16} = d_{13} = d_9 = d_{14} = d_3 = d_4 = d_8$ versus $d_1 = d_6 = d_{11} = d_{16} > d_{13} = d_9 = d_{14} = d_3 = d_4 = d_8$. We find that the null hypothesis is not rejected for pairs involving the SP. In contrast, the three European market pairs display large values of the test statistics (2, 5.3, and 4.8), all having small p-values. This test indicates that joint variations (whatever the sign and magnitude) increase subsequent dependency of returns, suggesting a persistence in dependency.

4.3.3. Modeling persistence in dependency

The last issue we address in this paper is the persistence in the dependency parameter. Estimations presented above have shown that, in many circumstances, past joint realizations affect the international dependency. We now measure the extent to which the persistence in dependency is likely to attenuate this link. We thus estimate two specifications which are designed to capture such a persistence, the TVC model (Eq. (9)) and the Markov-switching model (Eqs. (10) and (11)).

QML estimates of the TVC model are reported in Table 6, with $m = 5$ lags in the computation of ξ_t, so that short-term correlation is computed over one week of data. (The results are not altered when we select $m = 10$ or 20.) This table reveals that persistence in the dynamics of dependency is very strong with a parameter, β, ranging between 0.942 and 0.995. The effect of the past short-term correlation between residuals, measured by α, is strongly significantly positive for European market pairs, but barely significant for pairs involving the SP. These results suggest that the TVC model may be inappropriate for modeling the dynamics of the dependency parameter for pairs involving the SP. Furthermore, dependency appears to be very large and persistent between European markets, a feature which may be the sign of large but infrequent breaks in the dynamics of dependency. We thus turn to the estimation of a Markov-switching model for capturing the persistence in the dependency structure.

Table 6
Parameter estimates of the Student-t copula with dependency parameter driven by the TVC model (9)

	SP–FTSE	SP–DAX	SP–CAC	FTSE–DAX	FTSE–CAC	DAX–CAC
ρ	0.242	0.320	0.311	0.504	0.608	0.601
s.e.	(0.017)	(0.025)	(0.048)	(0.053)	(0.052)	(0.072)
β	0.942	0.963	0.995	0.984	0.980	0.983
s.e.	(0.068)	(0.016)	(0.003)	(0.004)	(0.004)	(0.004)
α	0.008	0.017	0.004	0.013	0.017	0.015
s.e.	(0.008)	(0.006)	(0.002)	(0.003)	(0.003)	(0.003)
n	23.579	18.143	12.800	10.769	9.398	8.868
s.e.	(9.460)	(5.224)	(2.625)	(1.954)	(1.515)	(1.370)
ℓL	140.101	239.233	201.314	499.231	826.917	724.092
AIC	−0.060	−0.103	−0.086	−0.217	−0.360	−0.315
SIC	−0.054	−0.097	−0.081	−0.211	−0.354	−0.309

This table reports parameter estimates for the Student-t copula when the dynamics of the dependency parameter ρ is given by a TVC model. Parameters are defined in Eq. (9). We also report the log-likelihood and the AIC and SIC information criteria.

Table 7 reports QML parameter estimates of the Markov-switching model in which both correlation and degree-of-freedom parameters are regime dependent. State 0 is associated with a low dependency between stock markets, while State 1 corresponds to the high-dependency regime. First, for all market pairs, the difference between the correlations estimated for the two regimes are strongly significant. In regime 0, correlations are about 0.23 for pairs involving the SP and about 0.3 for European market pairs. In regime 1,

Table 7
Parameter estimates of the Student-t copula with parameters driven by the Markov-switching model (10) and (11)

	SP–FTSE	SP–DAX	SP–CAC	FTSE–DAX	FTSE–CAC	DAX–CAC
ρ_0	0.222	0.248	0.234	0.262	0.280	0.332
s.e.	(0.028)	(0.018)	(0.018)	(0.022)	(0.020)	(0.018)
ρ_1	0.514	0.583	0.456	0.559	0.698	0.735
s.e.	(0.229)	(0.042)	(0.035)	(0.016)	(0.011)	(0.016)
n_0	29.052	∞	16.188	14.226	22.143	13.031
s.e.	(18.297)	–	(4.659)	(4.928)	(10.977)	(3.518)
n_1	10.836	7.142	9.733	11.396	9.133	7.263
s.e.	(25.801)	(2.684)	(4.060)	(3.074)	(2.089)	(1.986)
p	0.996	0.995	0.999	0.998	0.998	0.995
s.e.	(0.009)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)
q	0.942	0.979	0.998	0.998	0.998	0.988
s.e.	(0.075)	(0.010)	(0.002)	(0.001)	(0.001)	(0.004)
ℓL	140.291	249.223	204.163	488.740	829.692	699.747
AIC	−0.060	−0.107	−0.087	−0.212	−0.361	−0.304
SIC	−0.054	−0.102	−0.080	−0.206	−0.356	−0.297
Expected duration (in days)						
State 0	278	213	1667	625	455	189
State 1	17	47	417	556	400	83

This table reports parameter estimates for the Student-t copula when the dynamics of the dependency parameter ρ and the degree-of-freedom parameter n are given by a Markov-switching model. Parameters are defined in Eqs. (10) and (11). We also report the log-likelihood and the AIC and SIC information criteria.

correlations are about 0.5 and 0.65, respectively. Another interesting result is that, in some cases, the difference between the degree-of-freedom parameters is rather large, but barely significant. For the SP–DAX, the Student-t copula in regime 0 reduces to the Gaussian copula. For other market pairs, the dependency parameters are significantly different between the two regimes, while the degree-of-freedom parameters may be insignificantly different one from the other.

Another noticeable result provided by these estimates concerns the expected duration in a given regime. Expected duration (in days) is computed as $(1-p)^{-1}$ for regime 0 and $(1-q)^{-1}$ for regime 1. Results reported in Table 7 indicate that, for pairs involving the SP, the expected duration in the low-correlation regime is very long as compared with the expected duration in the high-correlation regime. For instance, for the SP–FTSE, the corresponding expected durations are 278 days and 17 days, respectively. This result suggests that the second regime may be economically irrelevant. In contrast, for European markets, the durations expected for the two regimes are very close. Therefore, the two regimes are more balanced and are likely to be relevant in order to capture the dynamics in the dependence structure.[16]

To provide further insight on the dynamics of the dependency structure between European markets, we consider now the evolution through time of the dependency parameter. Fig. 3 displays the evolution of parameter ρ_t for the FTSE–DAX and FTSE–CAC, estimated by the TVC model as well as by the Markov-switching model. For the latter model, we plot the aggregated-over-regimes (or implied) conditional dependency at time t. It is defined as $(\rho_0 p_{0t} + \rho_1(1-p_{0t}))$, where p_{0t} denotes the ex ante probabilities $\Pr[S_t = 0 \,|\, I_{t-1}]$. Interestingly, we notice that the evolutions of dependency estimated with the TVC model and the Markov-switching model look rather similar. This corroborates the rather close values of the information criteria obtained for the two models for these market pairs. The figure reveals that the dependency between the FTSE and the DAX is characterized essentially by two long subperiods. The first one, from 1980 to 1989, corresponds to a high probability of being in the low dependency regime 0 (with $\rho_0 = 0.26$). During this period, two short-lasting increases in the dependency between the FTSE and the DAX occurred in September 1981 and October 1987, when the two markets experienced simultaneous crashes. The second period, from 1990 on, is mainly associated with the high-dependency regime 1 (with $\rho_1 = 0.56$). The initial increase in dependency corresponds to the Kuwait crisis, from Iraq's invasion on August 2, 1990 through the conclusion of the Gulf war on March 3, 1991. Then the dependency between the FTSE and the DAX experienced a short-lasting decrease around September 1992. It can be explained by the EMS crisis, which was accompanied by a sharp increase in the FTSE. In contrast, the dependency briefly decreased in September and October 1995, when the DAX suffered from a marked decrease. We may observe that such events affected the dependency negatively, because only one market was under pressure. Finally, at the end of the period under study, from 1997 on, the two markets are strongly dependent. The increase in 1997 may be linked to the South-east Asian crisis, which started around June 1997. The period from 1998 to 1999 may also be related to

[16] Since traditional asymptotics do not apply in this setting, we performed experiments, to assess the statistical significance of each regime using a Monte-Carlo simulation, along the lines of Hansen (1992) and Ang and Bekaert (2002). These experiments turned out to be very cumbersome and indicated the second regime is strongly significant for European market pairs, while it is barely (or not) significant for pairs involving the SP.

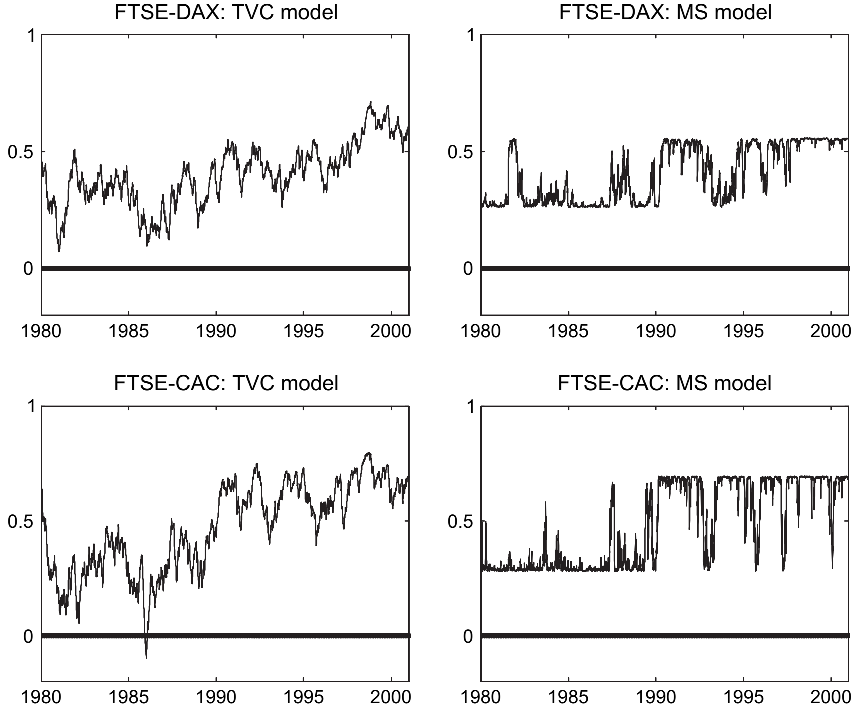

Fig. 3. The evolution of the parameter ρ_t obtained with the TVC model (9) and with the Markov-switching model (10) and (11), for the FTSE–DAX and the FTSE–CAC pairs, respectively.

the Russian crisis, which started with the collapse of the bond market at the beginning of August 1998.

The second pair, between the FTSE and the CAC returns, displays a similar pattern. One noticeable exception is the period 1981 and 1982. While the dependency between the FTSE and the DAX strongly increased over this period, the FTSE and the CAC did not experience such a pattern. This may be explained by the fact that the French market was, at this time, strongly affected by domestic political changes.

5. Conclusion

The methodology developed in this paper resorts to copula functions for modeling dependency between time series, when univariate distributions are complicated and cannot be easily extended to a multivariate setup. We use this methodology to investigate the dependency structure between daily stock market returns over the period 1980–1999.

We first provide empirical evidence that the distribution of daily returns may be well described by the skewed Student-*t* distribution, with volatility, skewness, and kurtosis varying over time. In such a context, modeling several returns simultaneously, using the multivariate extension of this distribution, would be extremely cumbersome. Subsequently, we use copula functions to join these complicated univariate distributions. This approach leads to a multivariate distribution which fits the data well, without involving time-consuming estimations. Finally, we describe how the dependency parameter can be rendered conditional and we propose alternative specifications to model the dynamics of the dependency parameter. On one hand, the dependency parameter is allowed to depend on the position of past realizations of margins in the unit square. This model provides a semi-parametric description of the dependency parameter. The dependency between European markets is found to increase significantly subsequently to movements in the same direction, either a crash or a boom. On the other hand, the dynamics of the dependency structure is captured through a time-varying correlation model and a Markov-switching model. Empirical evidence reveals that the dependency between European market pairs is time varying and has significantly increased between 1980 and 1999, contrary to the dependency between the U.S. and European markets.

This methodology may also be used in several contexts, such as conditional asset allocation or Value-at-Risk computation in a non-normal framework. To implement asset allocation in such a context, it is necessary to compute expressions involving multiple integrals of the joint distribution, while VaR applications require computing the probability that a portfolio exceeds a given threshold. Once the marginal models are known, such computations can be performed rather easily using numerical integration.

Acknowledgements

Michael Rockinger is also CEPR. The authors acknowledge support from the National Center of Competence: Financial Valuation and Risk Management (NCCR FINRISK). We are grateful to Sanvi Avouyi-Dovi, Giovanni Barone-Adesi, Patrick Gagliardini, Christian Gouriéroux, Christian Pfister, Jean-Michel Zakoïan, seminar participants at the 2001 European Investment Review conference, Eurandom, CREST, City University College, GREQAM, and ERUDITE seminars for their comments and suggestions. The authors are grateful to the two anonymous referees for their helpful comments. This paper was started when Jondeau was with the Banque de France and when Rockinger was on sabbatical at UCSD. A first version of this paper circulated as a working paper of the Banque de France (NER # 82) in 2001. The usual disclaimer applies.

Appendix A. Additional results on the copula and the skewed Student-*t* distribution

A1. Copula functions

In this Appendix, we define copulas and provide a theorem which justifies their use.[17]

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

1. $C(u_1, u_2)$ is increasing in u_1 and u_2.
2. $C(0, u_2) = C(u_1, 0) = 0$, $C(1, u_2) = u_2$, $C(u_1, 1) = u_1$.

[17] See Joe (1997) and Nelsen (1999) at textbook level. The following definition and theorem may be found in Nelsen (1999).

3. $\forall u_1, u_1', u_2, u_2'$ in [0,1] such that $u_1 < u_1'$ and $u_2 < u_2'$, we have $C(u_1', u_2') - C(u_1', u_2) - C(u_1, u_2') + C(u_1, u_2) \geq 0$.

This definition provides a multivariate extension of the definition of the cdf. An important property of the copula function is that it is defined over variables uniformly distributed over [0,1]. Now, if we set $u_i = F_i(x_i)$, then the copula function $C(F_1(x_1), F_2(x_2))$ describes the joint distribution of X_1 and X_2. We can now propose the following theorem, which first appeared in Sklar (1959).

Theorem 2. (*Sklar's theorem for continuous distributions*). *Let H be a joint distribution function with margins F_1 and F_2. Then, there exists a copula C such that, for all real numbers x_1 and x_2, one has the equality*

$$H(x_1, x_2) = C(F_1(x_1), F_2(x_2)). \quad (14)$$

Furthermore, if F_1 and F_2 are continuous, then C is unique. Conversely, if F_1 and F_2 are distributions, then the function H defined by Eq. (14) *is a joint distribution function with margins F_1 and F_2.*

In this work, we resort to the "converse" part of this theorem and construct a multivariate density from marginal ones.

A2. Skewed Student-t distribution

We now provide additional results on the skewed Student-*t* distribution, which are useful in the empirical application. First of all, Hansen (1994) shows that if $Z \sim \mathrm{ST}(\eta, \lambda)$, then Z has zero mean and unit variance. For this result to hold, it must be that $\eta > 2$. Jondeau and Rockinger (2003a,b) set $m_2 = 1 + 3\lambda^2$, $m_3 = 16c\lambda(1 + \lambda^2)(\eta - 2)^2 / [(\eta - 1)(\eta - 3)]$, defined if $\eta > 3$, and $m_4 = 3(\eta - 2)(1 + 10\lambda^2 + 5\lambda^4)/(\eta - 4)$, defined if $\eta > 4$. With these notations, they show that if $Z \sim \mathrm{ST}(\eta, \lambda)$ then the third and fourth moments of Z are given by:

$$E[Z^3] = [m_3 - 3am_2 + 2a^3]/b^3, \quad (15)$$

$$E[Z^4] = [m_4 - 4am_3 + 6a^2 m_2 - 3a^4]/b^4. \quad (16)$$

Since Z has zero mean and unit variance, skewness (Sk) and kurtosis (Ku) are directly related to the third and fourth moments: $\mathrm{Sk}[Z] = E[Z^3]$ and $\mathrm{Ku}[Z] = E[Z^4]$. The density and the various moments do not necessarily exist for all parameter values. Given the way asymmetry is introduced, we have $-1 < \lambda < 1$. As already mentioned, the distribution is meaningful only if $\eta > 2$. Furthermore, careful scrutiny of the algebra yielding Eq. (15) shows that skewness exists if $\eta > 3$. Last, kurtosis in Eq. (16) is well defined if $\eta > 4$.[18]

Hansen's skewed Student-*t* encompasses a large set of conventional densities. For instance, if $\lambda = 0$, Hansen's distribution reduces to the traditional Student-*t* distribution, which is not skewed. If, in addition, $\eta \to \infty$, it reduces to the normal density.

The copula involves marginal cdfs rather than densities. For this reason, we now derive the cdf of Hansen's skewed Student-*t* distribution. To do so, we recall that the conventional

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

Student-t distribution is defined by

$$t(x;\eta)=\frac{\Gamma\left(\frac{\eta+1}{2}\right)}{\Gamma\left(\frac{\eta}{2}\right)}\frac{1}{\sqrt{\pi\eta}}\left(1+\frac{x^2}{\eta}\right)^{-[(\eta+1)/2]},$$

where η 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. We write the cdf of a Student-t distribution with η degrees of freedom as

$$A(y;\eta)=\int_{-\infty}^{y}t(x)\mathrm{d}x.$$

The following proposition presents the cdf of the skewed Student-t distribution.

Proposition 3. *Let $D(z;\eta,\lambda)=Pr[Z<z]$, where Z has the skewed Student-t distribution given by Eq. (1). Then $D(z;\eta,\lambda)$ is defined as*

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

Proof. Let $w/\sqrt{\eta}=(bz+a)/\left[(1-\lambda)\sqrt{\eta-2}\right]$. The result follows then from the change of variable in Eq. (1) of z into w. □

References

Ang, A., Bekaert, G., 2002. International asset allocation with time-varying correlations. Review of Financial Studies 15 (4), 1137–1187.

Bauwens, L., Laurent, S., 2002. A new class of multivariate skew densities, with application to GARCH model. Working Paper, CORE, Université de Liège and Université Catholique de Louvain.

Bekaert, G., Harvey, C.R., 1995. Time-varying world market integration. Journal of Finance 50 (2), 403–444.

Bera, A.K., Kim, S., 2002. Testing constancy of correlation and other specifications of the BGARCH model with an application to international equity returns. Journal of Empirical Finance 9 (2), 171–195.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31 (3), 307–327.

Bollerslev, T., Wooldridge, J.M., 1992. Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. Econometric Reviews 11 (2), 143–172.

Campbell, J.Y., Hentschel, L., 1992. No news is good news: an asymmetric model of changing volatility. Journal of Financial Economics 31 (3), 281–318.

Chesnay, F., Jondeau, E., 2001. Does correlation between stock returns really increase during turbulent periods? Economic Notes 30 (1), 53–80.

Diebold, F.X., Inoue, A., 1999. Long memory and structural change. Stern School of Business Discussion Paper.

Diebold, F.X., Gunther, T.A., Tay, A.S., 1998. Evaluating density forecasts with applications to financial risk management. International Economic Review 39 (4), 863–883.

Embrechts, P., McNeil, A.J., Strautman, A.J., 1999. Correlation and dependency in risk management: properties and pitfalls. Working Paper, ETH Zurich.

Embrechts, P., Lindskog, F., McNeil, A., 2003. Modelling dependence with copulas and applications to risk management. In: Rachev, S.T. (Ed.), Handbook of Heavy Tailed Distributions in Finance. Elsevier/North-Holland, Amsterdam.

Engle, R.F., 1982. Auto-regressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50 (4), 987–1007.

Engle, R.F., 2002. Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. Journal of Business and Economic Statistics 20 (3), 339–350.

Engle, R.F., Gonzalez-Rivera, G., 1991. Semi-parametric ARCH models. Journal of Business and Economic Statistics 9 (4), 345–359.

Engle, R.F., Sheppard, K., 2001. Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. NBER Working Paper 8554.

Glosten, R.T., Jagannathan, R., Runkle, D., 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.

Gouriéroux, C., Jasiak, J., 2001. Memory and infrequent breaks. Economics Letters 70, 29–41.

Gouriéroux, C., Monfort, A., 1992. Qualitative threshold ARCH models. Journal of Econometrics 52 (1–2), 159–199.

Gray, S.F., 1995. An analysis of conditional regime-switching models. Working Paper, Fuqua School of Business, Duke University.

Hamao, Y., Masulis, R.W., Ng, V.K., 1990. Correlations in price changes and volatility across international stock markets. Review of Financial Studies 3 (2), 281–307.

Hamilton, J.D., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica 57 (2), 357–384.

Hansen, B.E., 1992. The likelihood ratio test under non-standard conditions: testing the Markov switching model of GNP. Journal of Applied Econometrics 7, S61–S82.

Hansen, B.E., 1994. Autoregressive conditional density estimation. International Economic Review 35 (3), 705–730.

Hartmann, P., Straetmans, S., de Vries, C.G., 2001. Asset market linkages in crisis periods. ECB Working Paper 71.

Harvey, C.R., Siddique, A., 1999. Autoregressive conditional skewness. Journal of Financial and Quantitative Analysis 34 (4), 465–487.

Joe, H., 1997. Multivariate Models and Dependence Concepts. Chapman and Hall, London.

Joe, H., Xu, J.J., 1996. The estimation method of inference functions for margins for multivariate models. Technical Report 166, Department of Statistics, University of British Columbia.

Jondeau, E., Rockinger, M., 2003a. Conditional volatility, skewness, and kurtosis: existence, persistence, and comovements. Journal of Economic Dynamics and Control 27 (10), 1699–1737.

Jondeau, E., Rockinger, M., 2003b. User's guide. Journal of Economic Dynamics and Control 27 (10), 1739–1742.

Kroner, K.F., Ng, V.K., 1998. Modeling asymmetric comovements of asset returns. Review of Financial Studies 11 (4), 817–844.

Lambert, P., Laurent, S., 2002. Modeling skewness dynamics in series of financial data using skewed location-scale distributions. Working Paper, Université Catholique de Louvain and Université de Liège.

Lamoureux, C.B., Lastrapes, W.D., 1990. Heteroskedasticity in stock return data: volume versus GARCH effects. Journal of Finance 45 (1), 221–229.

Longin, F., Solnik, B., 1995. Is the correlation in international equity returns constant: 1960–1990? Journal of International Money and Finance 14 (1), 3–26.

Longin, F., Solnik, B., 2001. Extreme correlation of international equity markets. Journal of Finance 56 (2), 649–676.

Nelsen, R.B., 1999. An Introduction to Copulas. Springer Verlag, New York.

Nelson, D.B., 1991. Conditional heteroskedasticity in asset returns: a new approach. Econometrica 59 (2), 347–370.

Patton, A.J., 2006. Estimation of copula models for time series of possibly different lengths. Journal of Applied Econometrics 21 (2), 147–173.

Poon, S.H., Rockinger, M., Tawn, J., 2004. Extreme-value dependence in financial markets: diagnostics, models and financial implications. Review of Financial Studies 17 (2), 581–610.

Ramchand, L., Susmel, R., 1998. Volatility and cross correlation across major stock markets. Journal of Empirical Finance 5 (4), 397–416.

Sahu, S.K., Dey, D.K., Branco, M.D., 2001. A new class of multivariate skew distributions with applications to Bayesian regression models. Working Paper, University of Southampton, University of Connecticut, and University of São Paulo.

Schaller, H., van Norden, S., 1997. Regime switching in stock market return. Applied Financial Economics 7, 177–191.

Shih, J.H., Louis, T.A., 1995. Inferences on the association parameter in copula models for bivariate survival data. Biometrics 51, 1384–1399.

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.

Susmel, R., Engle, R.F., 1994. Hourly volatility spillovers between international equity markets. Journal of International Money and Finance 13 (1), 3–25.

Theodossiou, P., 1998. Financial data and the skewed generalized *T* distribution. Management Science 44 (12-1), 1650–1661.

Tse, Y.K., 2000. A test for constant correlations in a multivariate GARCH model. Journal of Econometrics 98 (1), 107–127.

Tse, Y.K., Tsui, A.K.C., 2002. A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. Journal of Business and Economics Statistics 20 (3), 351–362.

Vlaar, P.J.G., Palm, F.C., 1993. The message in weekly exchange rates in the European monetary system: mean reversion, conditional heteroscedasticity, and jumps. Journal of Business and Economic Statistics 11 (3), 351–360.