Unconditional Copula-Based Simulation of Tail Dependence for Co-movement of International Equity Markets

Wei Sun
Institute for Statistics and Mathematical Economics
University of Karlsruhe, Germany

Svetlozar Rachev*
Institute for Statistics and Mathematical Economics
University of Karlsruhe, Germany
and
Department of Statistics and Applied Probability
University of California, Santa Barbara, USA

Frank J. Fabozzi
Yale School of Management, USA

Petko S. Kalev
Department of Accounting and Finance
Monash University, Australia

*Corresponding author: Prof. Dr. S. Rachev, Institute for Statistics and Mathematical Economics, Uni-Karlsruhe, Postfach 6980, 76128 Karlsruhe, Germany. Email: rachev@statistik.uni-karlsruhe.de.

S. Rachev's research was supported by grants from Division of Mathematical, Life and Physical Science, College of Letters and Science, University of California, Santa Barbara, and the Deutschen Forschungsgemeinschaft. W. Sun's research was supported by grants from the Deutschen Forschungsgemeinschaft. P.S. Kalev's research was supported with a NCG grant from the Faculty of Business and Economics, Monash University. Data are supplied by Securities Industry Research Center of Asia-Pacific (SIRCA) on behalf of Reuters.

Unconditional Copula-Based Simulation of Tail Dependence

for Co-movement of International Equity Markets

Abstract

Analyzing equity market co-movements is important for risk diversification of an international portfolio. Copulas have several advantages compared to the linear correlation measure in modeling co-movement. This paper introduces a copula ARMA-GARCH model for analyzing the co-movement of international equity markets. The model is implemented with an ARMA-GARCH model for the marginal distributions and a copula for the joint distribution. After goodness of fit testing, we find that the Student's t copula ARMA(1,1)-GARCH(1,1) model with fractional Gaussian noise is superior to alternative models investigated in our study where we model the simultaneous co-movement of nine international equity market indexes. This model is also suitable for capturing the long-range dependence and tail dependence observed in international equity markets.

Key Words: copula, fractional Gaussian noise, high-frequency data, self-similarity, tail dependence

JEL Classification: C15, C46, C52, G15

1. Introduction

The co-movement of world equity markets is often used as a barometer of economic globalization and financial integration. Analyzing such co-movement is important for risk diversification of an international portfolio. The source of co-movement of international equity markets is the volatility-in-correlation effect found by Andersen *et al.* (2001) in individual stock returns and by Solnik *et al.* (1996) in international equity index returns. In fact, volatility-in-correlation effect could be explained by the tail dependence of underlying assets, which exhibits extreme events happening simultaneously. The co-movement, volatility-in-correlation, and tail dependence in a sense are interrelated when analyzing the dependence structure of international equity markets. It is also found that the correlations between consecutive returns decay slowly, that is, long-range dependence in returns is exhibited. Co-movement reflects intercorrelation between underlying asset returns (or returns in different markets) and long-range dependence exhibits autocorrelation within a single asset return (or return of a single market). Therefore, when analyzing international equity markets, we face two dependence structures: the correlation within a single market and the correlation between several markets.

When dealing with the dependence (i.e., long-range dependence) of a single market, we should take other stylized factors into account such as volatility clustering and distributional heavy-tails. It is necessary to treat long-range dependence, volatility clustering, and heavy-tailedness simultaneously in order to obtain more accurate predictions of market volatility. Rachev and Mittnik (2000) note that for modeling financial data, not only does model structure play an important role, but distributional assumptions influence modeling accuracy. The stable Paretian distribution can be used to capture characteristics of financial data since it is rich enough to encompass those stylized facts. Other researchers have shown the advantages of stable distributions in financial modeling (see, Fama (1963), Mittnik and Rachev (1993), Rachev (2003), and Rachev *et al.* (2005)). Several studies have reported that long-range dependence, self-similar processes, and stable distribution are very closely related (see, Taqqu and Samorodnitsky (1994), Rachev and Mittnik (2000), Rachev and Samorodnitsky (2001), Doukhan *et al.* (2003), and Racheva and Samorodnitsky (2003)). Long-range dependence processes are asymptotically second-order self-similar (see Willinger *et al.* (1998)). Second-order self-similarity describes the property that the correlation structure of a process is preserved irrespective of time scaling. Although self-similarity and long-range dependence are different concepts, in the case of second-order self-similarity, long-range dependence implies self-similarity and vice versa. As to this point, it is natural to employ specified self-similar processes in the study of the within-market dependence together with capturing volatility clustering and heavy-tailedness.

When dealing with the dependent structure between several markets, the usual linear correlation is often applied. But the usual linear correlation is not a satisfactory measure of the dependence among global equity markets for several reasons. First, when the variance of returns in those markets turns out to be infinite, that is, extreme events are frequently observed, the linear correlation between these markets is undefined. Second, linear correlation assumes that both marginal and joint distributions of returns in these markets are elliptical. In real-world markets, this assumption is unwarranted. Third, the linear correlation is not invariant under nonlinear strictly increasing transformations, implying that the return might be uncorrelated whereas the prices are correlated or vice versa. Fourth, linear correlation only measures the degree of dependence but does not clearly discover the structure of dependence. It has been widely observed that market crashes or financial crises often occur in different countries at about the same time period even when the correlation among those markets is fairly low. The structure of dependence also influences the diversification benefit gained based on a linear correlation measure. Embrechts *et al* (2003) and Rachev *et al* (2005) illustrate the drawbacks of using linear correlation to analyze dependency. A more prevalent approach which overcomes the disadvantages of linear correlation is to model dependency by using copulas. With the copula method, the nature of dependence that can be modeled is more general and the dependence of extreme events can be considered.

Based on a copula-ARMA-GARCH modeling structure for stock market indexes from nine different countries, in this paper we compare several candidate specifications using simulation methods. In our modeling structure, the marginal distribution captures the long-range dependence, heavy tails, and volatility clustering simultaneously in order to obtain more accurate predictions, and these marginal distributions are connected by a specified copula. The empirical results indicate that the Student's t copula and ARMA-GARCH model with fractional Gaussian noise dominate the alternative models tested in this study.

We organized the paper as follows: A brief introduction of copula considering tail dependence is provided in Section 2. The data and empirical methodology we used in our study are described in Section 3. In Section 4, we specify two self-similar processes: fractional Gaussian noise and fractional stable noise. Methods for estimating the parameters of underlying self-similar processes are introduced. In Section 5, the simulation methods applied in our empirical study are introduced. The empirical results based on high-frequency data at 1-minute level for nine international stock market indexes are reported in Section 6. In that section, we compare the goodness of fit for both the marginal distribution and joint distribution. We summarize our conclusions in Section 7.

2. Unconditional copulas and tail dependence

Copulas enable the dependence structure to be extracted from both the joint distribution function and the marginal distribution functions. From a mathematical viewpoint, a copula function C is a probability distribution function on the n-dimensional hypercube. Sklar (1959) has shown that any multivariate probability distribution function F_Y of some random vector $Y = (Y_1, ..., Y_n)$ can be represented with the help of a copula function C of the following form:

$$\begin{aligned} F_Y(y_1, ..., y_n) &= P(Y_1 \leq y_1, \ldots, Y_n \leq y_n) \\ &= C(P(Y_1 \leq y_1), \ldots, P(Y_n \leq y_n)) \\ &= C(F_{y_1}(y_1), \ldots, F_{y_n}(y_n)) \end{aligned} \tag{1}$$

where F_{y_i}, $i = 1, \ldots, n$ denote the marginal distribution functions of the random variables, Y_i, $i = 1, \ldots, n$.

When the variables are continuous, the density c associated with the copula is given by:

$$c(F_{y_1}(y_1), \ldots, F_{y_n}(y_n)) = \frac{\partial^n C(F_{y_1}(y_1), \ldots, F_{y_n}(y_n))}{\partial F_{y_1}(y_1), \ldots, \partial F_{y_n}(y_n)}. \tag{2}$$

The density function f_Y corresponding to the n-variate distribution function F_Y is

$$f_Y(y_1, ..., y_n) = c(F_{y_1}(y_1), \ldots, F_{y_n}(y_n)) \prod_{i=n}^{n} f_{y_i}(y_i), \tag{3}$$

where f_{y_i}, $i = 1, \ldots, n$ is the density function of F_{y_i}, $i = 1, \ldots, n$ (see, Joe (1997), Cherubini *et al.* (2004), and Nelsen (2006)).

Two commonly used unconditional copulas are the unconditional Gaussian copula and unconditional Student's t copula. Specification of these two copulas are given in the Appendix A.

In financial data, we can observe that extreme events happen simultaneously for different assets. In a time interval, several assets might exhibit extreme values. *Tail dependence* reflects the dependence structure between extreme events. It turns out that tail dependence is a copula property. Letting $(Y_1, Y_2)^T$ be a vector of continuous random variables with marginal distribution functions F_1, F_2, then the coefficient of the upper tail dependence of $(Y_1, Y_2)^T$ is

$$\lambda_U = \lim_{u \to 1} P\Big(Y_2 > F_2^{-1}(u) | Y_1 > F_1^{-1}(u)\Big), \tag{4}$$

and the coefficient of the lower tail dependence of $(Y_1, Y_2)^T$ is

$$\lambda_L = \lim_{u \to 0} P\Big(Y_2 < F_2^{-1}(u) | Y_1 < F_1^{-1}(u)\Big). \tag{5}$$

If $\lambda_U > 0$, there exists upper tail dependence and the positive extreme values can be observed simultaneously. If $\lambda_L > 0$, there exists lower tail dependence and the negative extreme values can be observed simultaneously. Embrechts *et al.* (2003) introduce some coefficients of tail dependence of different copulas.

Empirical studies have failed to support the assumption that return data follow a Gaussian distribution. Instead, return data often exhibit excess kurtosis and heavy tails. Nor is the multivariate Gaussian distribution warranted either when studying multi-dimensional return data because of tail dependence among them. The Student's t copula and Clayton copulas can characterize tail dependence in multi-dimensional return data but the Gaussian copula can not. In Clayton copulas, only Joe-Clayton copula can capture both the lower and upper tail dependence and the Joe-Clayton copula allows for asymmetric tail dependence. Unfortunately, the Joe-Clayton copula cannot be implemented for multi-dimensional return data because of computational difficulties, while the Student's t copula can be implemented to capture tail dependence in multi-dimensional return data (see, Embrechts *et al.* (2003), Patton (2005), and Nelsen (2006)).

3. Data and empirical methodology

3.1 Data

In previous studies of the co-movement of international equity markets, low-frequency data are usually examined. Because stock indexes change their composition quite often over time, it is difficult to find the impact of these changes in composition when analyzing the return history of stock indexes using low-frequency data. Dacorogna *et al.* (2001) calls this phenomenon the "breakdown of the permanence hypothesis". In order to overcome this problem, we use high-frequency data in our study.

Employing high-frequency data has several advantages compared with low-frequency data. First, with a very large amount of observations, high-frequency data offers a higher level of statistical significance. Second, high-frequency data are gathered at a low level of aggregation, thereby capturing the heterogeneity of players in financial markets. These players should be properly modeled in order to make valid inferences about market movements. Low-frequency data, say daily or weekly data, aggregate the heterogeneity in a smoothing way. As a result, many of the movements in the same direction are strengthened and those in the opposite direction cancelled in the process of aggregation. The aggregated series generally show smoother style than their components. The relationships between the observations in these aggregated series often exhibit greater smoothness than their components. For example, a curve exhibiting a one-week market movement based on daily return data might be a line with a couple of nodes. The smooth line segment veils the intra-daily fluctuation of the market. But high-frequency data can reflect such intra-daily fluctuations and the intra-daily co-movement can be taken into account. Third, using high-frequency data in analyzing the co-movement of international equity markets can consider both microstructure effects and macroeconomic factors. This is because information contained in high-frequency data can be resolved into a higher frequency part (i.e., the intra-daily fluctuation) and a lower frequency part (i.e., low-frequency smoothness). The information provided by the higher frequency part mirrors the microstructure effect of the equity markets and the information in the lower frequency part shows the smoothed trend that is usually influenced by macroeconomic factors in these markets.

Standard econometric techniques are based on homogeneous time series analysis. If a researcher uses analytic methods of homogeneous time series for inhomogeneous time series, the reliability of the results will be doubtful. Aggregating inhomogeneous tick-by-tick data to the equally spaced (homogeneous) time series is needed. Engle and Russell (1998) argue that for aggregating tick-by-tick data to a fixed time interval, if a short time interval is chosen, there will be many intervals in which there is no new information, and if choosing a wide interval, micro-structure features might be missing. Aït-Sahalia (2005) suggests keeping the data at the ultimate frequency level. In our empirical study, intra-daily data, which we refer to as the high-frequency data in this paper, at 1-minute level were aggregated from tick-by-tick data to investigate the co-movement of international equity markets.

The high-frequency data of the nine international stock indexes listed in Table 1 from January 8, 2002 to December 31, 2003 were aggregated to the 1-minute frequency level. The aggregation algorithm is based on the linear interpolation introduced by Wasserfallen and Zimmermann (1995). That is, given an inhomogeneous series with times t_i and values $\varphi_i = \varphi(t_i)$, the index i identifies the irregularly spaced sequence. The target homogeneous time series is given at times $t_0 + j\Delta t$ with fixed time interval Δt starting at t_0. The index j identifies the regularly spaced sequence. The time $t_0 + j\Delta t$ is bounded by two times t_i of the irregularly spaced series, $I = \max(\, i \,|t_i \le t_0 + j\Delta t)$ and $t_I \le t_0 + j\Delta t > t_{I+1}$. Data are interpolated between t_I and t_{I+1}. The linear interpolation shows that

$$\varphi(t_0 + j\Delta t) \;=\; \varphi_I \;+\; \frac{t_0 + j\Delta t - t_I}{t_{I+1} - t_I}\,(\varphi_{I+1} - \varphi_I). \tag{6}$$

Dacorogna *et al.* (2001) pointed out that linear interpolation relies on the future of time and Müller *et al.* (1990) suggests that linear interpolation is an appropriate method for stochastic processes with independent and identically distributed (i.i.d.) increments.

Empirical evidence has shown the seasonality in high-frequency data. In order to remove such disturbance, several methods of data adjusting have been adopted in modeling. Engle and Russell (1998) and other researchers adopt several methods to adjust the seasonal effect in data. In our study, seasonality is treated as one type of self-similarity. Consequently, it is not necessary to adjust for the seasonal effect in the data.

3.2 Empirical methodology

To investigate the co-movement of international stock markets, we use the market index for each country as the proxy for the market movement and propose the copula ARMA-GARCH model. This model is implemented with an ARMA-GARCH model for the marginal distributions and one copula for the joint distribution. Six GARCH models with different kinds of residuals (i.e., residuals with forms of white noise, fractional Gaussian noise, fractional stable noise, stable distribution, generalized Pareto distribution, and generalized extreme value distribution) for the marginal distributions are simulated. After goodness of fit testing, we use the best goodness of fit model for the marginal distributions with Gaussian copula and Student's t copula for the joint distribution to simulate the returns on the equity indexes. Then, the models will be tested with several goodness of fit test methods for a large dataset.

We define the ARMA-GARCH model for the conditional mean equation as:

$$y_t = \alpha_0 + \sum_{i=1}^{r} \alpha_i y_{t-i} + \varepsilon_t + \sum_{j=1}^{m} \beta_j \varepsilon_{t-j}. \qquad (7)$$

Let $\varepsilon_t = \sigma_t u_t$, where the conditional variance of the innovations, σ_t^2, is by definition

$$Var_{t-1}(y_t) = E_{t-1}(\varepsilon_t^2) = \sigma_t^2. \qquad (8)$$

The general GARCH(p,q) processes for the conditional variance of the innovation is then

$$\sigma_t^2 = \kappa + \sum_{i=1}^{p} \gamma_i \sigma_{t-i}^2 + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j}^2. \qquad (9)$$

Since $\varepsilon_t = \sigma_t u_t$, u_t could be calculated from ε_t / σ_t. Defining

$$\tilde{u_t} = \frac{\varepsilon_t^s}{\hat{\sigma_t}}, \qquad (10)$$

where ε_t^s is estimated from the sample and $\hat{\sigma_t}$ is the estimation of σ_t. In our study, ARMA(1,1)-GARCH(1,1) are parameterized as marginal distributions with different kinds of u_t (i.e., normal distribution, fractional Gaussian noise, fractional stable noise, stable distribution, generalized Pareto distribution, and generalized extreme value distribution).

From the goodness of fit testing for marginal distributions, we find the best fit model. Then taking the best fit model as marginal distributions for each stock index return, we simulate a multivariate Gaussian copula and Student's t copula for the dependence structure of the nine stock index returns. The simulation method adopted is introduced in Section 4.

The Kolmogorov-Simirnov distance (KS) and Anderson-Darling distance (AD) proposed by Rachev and Mittnik (2000) and Cramer Von Mises distance (CVM)[1] are used as the criterion for the goodness of fit testing. The major disadvantage of KS statistics researchers have argued is that they tend to be more sensitive near the center of the distribution than at the tails. But AD statistics can overcome this. The reliability of testing the empirical distribution will be increased with the help of these two statistics, with the KS distance focusing on the deviations around the median of the distribution and the AD distance focusing on the discrepancies in the tails.

4. Analysis of the marginal distribution

As mentioned in the previous section, we apply ARMA(1,1)-GARCH(1,1) with alternative distributions for residuals u_t in our empirical study to model the marginal distribution of each equity market return. The key point of our marginal distributions is to empower the residuals u_t to capture the stylized factors, such as, long-range dependence and heavy-tailedness. One of the powerful forms of u_t is the self-similar process. Two specified self-similar processes applied in our empirical study are the fractional Gaussian noise and the fractional stable noise. The reason why such self-similar processes are powerful is that they impose an index on quantifying the degree of long-range dependence and measuring self-similarity. In this section, based on the residuals u_t in ARMA(1,1)-GARCH(1,1) model, we introduce how to estimate the parameters of the two specified self-similar processes for u_t.

4.1 The self-similarity parameter

Self-similarity is defined by Samorodnitsky and Taqqu (1994) as follows. Let T be either $R, R_+ = \{t : t \geq 0\}$ or $\{t : t > 0\}$. The real-valued process $\{X(t), t \in T\}$ is self-similar with Hurst index $H > 0$, if for any $a > 0$ and $d \geq 1$, $t_1, t_2, ..., t_d \in T$, satisfying:

$$\Big(X(at_1), X(at_2), ..., X(at_d)\Big) \stackrel{d}{=} \Big(a^H X(t_1), a^H X(t_2), ...a^H X(t_d)\Big). \tag{11}$$

The Hurst index H plays a key role in such processes to capture long-range dependence.

Let $\phi(k)$ denote the kth-order autocovariance function of $\{X(t), t \in T\}$, for $0 < H < 1$, and $H \neq 0.5$, $\phi(k) \sim H(2H-1)k^{2(H-1)}$ holds. $\{X(t), t \in T\}$ is called long-range dependence if $\sum_{k=1}^{\infty} \phi(k) = +\infty$. $\{X(t), t \in T\}$ is called short-range dependence if $\sum_{k=1}^{\infty} \phi(k) < +\infty$. If $0 < H < 0.5$, then $\sum_{k=1}^{\infty} \phi(k) \sim \sum_{k=1}^{\infty} H(2H-1)k^{2(H-1)}$. Note that in this case, $2(H-1) < -1$, $\sum_{k=1}^{\infty} k^{2(H-1)} < +\infty$. Thus $X(t)$

[1] Specifically, these criterion are defined as follows:

$$KS = \sup_{x \in \Re} \left| F_s(x) - \tilde{F}(x) \right|,$$

$$AD = \sup_{x \in \Re} \frac{\left| F_s(x) - \tilde{F}(x) \right|}{\sqrt{\tilde{F}(x)(1-\tilde{F}(x))}},$$

and

$$CVM = \int_{-\infty}^{\infty} \Big(F_s(x) - \tilde{F}(x)\Big)^2 d\tilde{F}(x),$$

where $F_s(x)$ denotes the empirical sample distribution and $\tilde{F}(x)$ is the estimated distribution function.

exhibits short-range dependence. If $0.5 < H < 1$, then $2(H-1) > -1$, thus $\sum_{k=1}^{\infty} k^{2(H-1)} = +\infty$, $X(t)$ shows long-range dependence. For all $k \geq 1$, if $H = 0.5$, the autocovariance is zero and $X(t)$ is a random walk; if $H = 1$, then $\phi(k) = 1$ and we have the degenerate situation with no meaning; and if $H > 1$, then $\phi(k) > 1$ and that is impossible.

Several methods for estimating the Hurst index have been proposed (see, Beran (1994)). Applying the method of calculating R/S statistic proposed by Hurst (1951), we estimated the Hurst index of the nine international equity index returns and show the results in Table 1.

As we have shown above, the Hurst index $H \in (0,1)$ usually serves as a measure of the self-similarity of stochastic processes. It can be somewhat explained by considering the covariance of two consecutive increments. When $H \in (0, 0.5)$, the increments of a process tend to have opposite signs and thus are more zigzagging due to their negative covariance; when $H \in (0.5, 1)$, the covariance between these two increments is positive and that process is less zigzagging; when $H = 0.5$, the covariance between these two increments is zero. It can be stated as following: If the Hurst index is less than 0.5, the process displays "anti-persistence" which means that positive excess return is more likely to be reversed and the performance in the next period is likely to be below the average, or in the contrary, negative excess return is more likely to be reversed and the performance in the next period is likely to be above the average. If the Hurst index is greater than 0.5, the process displays "persistence" which means that positive excess return or negative excess return is more likely to be continued and the performance in the next period is likely to be the same as that in the current period. If the Hurst index is equal to 0.5, the process displays no memory, meaning the performance in the next period has equal probability of being below and above the performance in the current period. From Table 1, we find that the Hurst index has no value of 0.5. Clearly, the memory effect occurs for the equity index returns in our study.

4.2 Specification of the self-similar processses

In this section, specification of two self-similar processes used in our empirical study, fractional Gaussian noise and fractional stable noise, are introduced. Mandelbrot and Wallis (1968) first introduced the fractional Brownian motion (FBM) and Samorodnitsky and Taqqu (1994) clarified the definition of FBM as a Gaussian process having self-similarity index H and stationary increments. Mandelbrot and van Ness (1968) defined the stochastic representation

$$B_H(t) := \frac{1}{\Gamma(H+\frac{1}{2})}\left(\int_{-\infty}^{0}[(t-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}}]dB(s) + \int_{0}^{t}(t-s)^{H-\frac{1}{2}}dB(s)\right), \tag{12}$$

where $\Gamma(\cdot)$ represents the Gamma function:

$$\Gamma(a) := \int_0^\infty x^{a-1}e^{-x}dx,$$

and $0 < H < 1$ is the Hurst parameter. The integrator B is ordinary Brownian motion. The main difference between fractional Brownian motion and ordinary Brownian motion is that the increments in Brownian motion are independent while in fractional Brownian motion they are dependent. As for the fractional Brownian motion, Samorodnitsky and Taqqu (1994) define its increments $\{Y_j, j \in Z\}$ as fractional Gaussian noise (FGN), which is for $j = 0, \pm 1, \pm 2, ...$, $Y_j = B_H(j-1) - B_H(j)$.

Fractional Brownian motion can capture the effect of long-range dependence, but with less power to capture the heavy tailedness. The existence of abrupt discontinuities in return data, combined with the empirical observation of sample excess kurtosis and unstable variance, suggests that return series can best be described by a stable Paretian distribution (see, Mandelbrot (1963, 1983)). It is natural to introduce the stable Paretian distribution in self-similar processes in order to capture both long-range dependence and heavy tailedness. Samorodinitsky and Taqqu (1994) introduce the α-stable *H-sssi* processes $\{X(t), t \in R\}$ with $0 < \alpha < 2$. If $0 < \alpha < 1$, the values of the Hurst index are $H \in (0, 1/\alpha]$ and if $1 < \alpha < 2$, the values of the Hurst index are $H \in (0, 1]$. There are several extensions of fractional Brownian motion to the stable distribution. The most commonly used is the linear fractional stable motion (also called linear fractional Lévy motion), which is defined by Samorodinitsky and Taqqu (1994).[2] In this paper, if there is no special indication, fractional stable noise (fsn) is generated from linear fractional stable motion.

4.3 Estimation of the self-similarity parameter

Beran (1994) discusses the approximate maximum likelihood estimator (MLE)[3] of the self-similarity parameter. For fractional Gaussian noise, Y_t, let $f(\lambda; H)$ denote the power spectrum of Y after being normalized to have variance 1 and let $I(\lambda)$ denote the periodogram of Y_t; that is,

$$I(\lambda) = \frac{1}{2\pi N} \left| \sum_{t=1}^{N} Y_t \, e^{i\,t\,\lambda} \right|^2 . \tag{13}$$

The MLE of H is to find $\hat{H}$ that minimizes

$$g(\hat{H}) = \int_{-\pi}^{\pi} \frac{I(\lambda)}{f(\lambda; \hat{H})} d\lambda. \tag{14}$$

Stoev *et al.* (2002) proposed the least-squares (LS) estimator for the Hurst index based on the finite impulse response transformation (FIRT) and wavelet transform coefficients of the fractional stable motion. FIRT is a filter $v = (v_0, v1, ..., v_p)$ of real numbers $v_t \in \Re, t = 1, ..., p$, and length $p + 1$. It is defined for X_t by

$$T_{n,t} = \sum_{i=0}^{p} v_i \, X_{n(i+t)}, \tag{15}$$

where $n \geq 1$ and $t \in N$. The $T_{n,t}$ are the FIRT coefficients of X_t (i.e., the FIRT coefficients of the fractional stable motion). The indices n and t can be explained as "scale" and "location". If $\sum_{i=0}^{p} i^r v_i = 0$, for $r = 0, ..., q-1$, but $\sum_{i=0}^{p} i^q v_i \neq 0$, the filter v_i can be said to have $q \geq 1$ zero moments. If $\{T_{n,t}, n \geq 1, t \in N\}$ is the FIRT coefficients of fractional stable motion with the filter v_i that has at least one zero moment, Stoev *et al.* (2002) prove the following properties of $T_{n,t}$: (1) $T_{n,t+h} \stackrel{d}{=} T_{n,t}$, and (2) $T_{n,t} \stackrel{d}{=} n^H T_{1,t}$, where $h, t \in N$, $n \geq 1$. We assume that $T_{n,t}$ are available for the fixed scales n_j $j = 1, ..., m$ and locations $t = 0, ..., M_j - 1$ at the scale n_j, since only a finite number, say M_j, of the FIRT coefficients are available at the scale n_j. By using these two properties, we have

$$E \log |T_{n_j,0}| = H \log n_j + E \log |T_{1,0}|. \tag{16}$$

[2]Some properties of these processes have been discussed in Mandelbrot and Van Ness (1968), Maejima and Rachev (1987), Manfields *et al.* (2001), Rachev and Mittnik (2000), Rachev and Samorodnitsky (2001), Samorodnitsky (1994), and Samorodinitsky and Taqqu (1994).

[3]It is also called the Whittle estimator.

The left-hand side of this equation can be approximated by

$$Y_{\log}(M_j) = \frac{1}{M_j} \sum_{t=0}^{M_j-1} \log |T_{n_j,t}|. \tag{17}$$

Then we obtain

$$\begin{pmatrix} Y_{\log}(M_1) \\ \vdots \\ Y_{\log}(M_m) \end{pmatrix} = \begin{pmatrix} \log n_1 & 1 \\ \vdots & \vdots \\ \log n_m & 1 \end{pmatrix} \begin{pmatrix} H \\ E \log |T_{1,0}| \end{pmatrix} + \begin{pmatrix} \sqrt{M}\Big(Y_{\log}(M_1) - E \log |T_{n_1,0}|\Big) \\ \vdots \\ \sqrt{M}\Big(Y_{\log}(M_m) - E \log |T_{n_m,0}|\Big) \end{pmatrix}. \tag{18}$$

We can express equation (18) as follows

$$Y = X\theta + \frac{1}{\sqrt{M}}\varepsilon, \tag{19}$$

where ε is the vector showing the difference between $\sqrt{M}Y_{\log}(M_m)$ and $\sqrt{M}E(\log |T_{n_m,0}|)$. Equation (19) shows that the self-similarity parameter H can be estimated by a standard linear regression of the vector Y against the matrix X. Stoev *et al.* (2002) provide the details for implementing such a procedure.

4.4 The parameters of a stable Non-Gaussian distribution

A stable distribution requires four parameters for complete description: an index of stability $\alpha \in (0, 2]$ also called the tail index, a skewness parameter $\beta \in [-1, 1]$, a scale parameter $\gamma > 0$, and a location parameter $\zeta \in \Re$. There is unfortunately no closed-form expression for the density function and distribution function of a stable distribution. Rachev and Mittnik (2000) give the definition for the stable distribution: A random variable X is said to have a stable distribution if there are parameters $0 < \alpha \leq 2$, $-1 \leq \alpha \leq 1$, $\gamma \geq 0$ and ζ real such that its characteristic function has the following form:

$$E \exp(i\theta X) = \begin{cases} \exp\{-\gamma^\alpha |\theta|^\alpha (1 - i\beta(\sin\theta) \tan\frac{\pi\alpha}{2}) + i\zeta\theta\}, & if \quad \alpha \neq 1 \\ \exp\{-\gamma|\theta|(1 + i\beta\frac{2}{\pi}(\sin\theta) \ln|\theta|) + i\zeta\theta\}, & if \quad \alpha = 1 \end{cases} \tag{20}$$

and,

$$\sin\theta = \begin{cases} 1, & if \quad \theta > 0 \\ 0, & if \quad \theta = 0 \\ -1, & if \quad \theta < 0 \end{cases} \tag{21}$$

For $0 < \alpha < 1$ and $\beta = 1$ or $\beta = -1$, the stable density is only for a half line.

In order to estimate the parameters of a stable distribution, we use the ML method given in Rachev and Mittnik (2000). Given N observations, $X = (X_1, X_2, \cdots, X_N)'$ for the positive half line, the log-likelihood function is of the form

$$\ln(\alpha, \lambda; X) = N \ln \lambda + N \ln \alpha + (\alpha - 1) \sum_{i=1}^{N} \ln X_i - \lambda \sum_{i=1}^{N} X_i^\alpha, \tag{22}$$

which can be maximized using, for example, a Newton-Raphson algorithm. It follows from the first-order condition,

$$\lambda = N \left(\sum_{i=1}^{N} X_i^\alpha \right)^{-1} \tag{23}$$

that the optimization problem can be reduced to finding the value for α which maximizes the concentrated likelihood

$$\ln^*(\alpha; X) = \ln \alpha + \alpha\nu - \ln\left(\sum_{i=1}^{N} X_i^{\alpha}\right), \tag{24}$$

where $\nu = N^{-1}\Sigma_{i=1}^{N} \ln X_i$.

The information matrix evaluated at the maximum likelihood estimates, denoted by $I(\hat{\alpha}, \hat{\lambda})$, is given by

$$I(\hat{\alpha}, \hat{\lambda}) = \begin{pmatrix} N\hat{\alpha}^{-2} & \sum_{i=1}^{N} X_i^{\hat{\alpha}} \ln X_i \\ \sum_{i=1}^{N} X_i^{\hat{\alpha}} \ln X_i & N\hat{\lambda}^{-2} \end{pmatrix}.$$

It can be shown that under fairly mild conditions, the maximum likelihood estimates $\hat{\alpha}$ and $\hat{\lambda}$ are consistent and have asymptotically a multivariate normal distribution with mean $(\alpha, \lambda)'$(see Rachev and Mittnik (2000)).[4]

5 Simulating the co-movement of international equity markets

5.1 Simulation of the marginal distribution

Paxson (1997) gives a method to generate the fractional Gaussian noise by using the Discrete Fourier Transform of the spectral density. Bardet *et al.* (2003) give a concrete simulation procedure based on this method with respect to alleviating some of the problems faced in practice. The procedure is:

1. Choose an even integer M. Define the vector of the Fourier frequencies $\Omega = (\theta_1, ..., \theta_{M/2})$, where $\theta_t = 2\pi t/M$ and compute the vector $F = f_H(\theta_1), ..., f_H(\theta_{M/2})$, where

   $$f_H(\theta) = \frac{1}{\pi}\sin(\pi H)\Gamma(2H+1)(1-\cos\theta)\sum_{t\in\aleph}|2\pi t + \theta|^{-2H-1}$$

   $f_H(\theta)$ is the spectral density of FGN.

2. Generate $M/2$ i.i.d exponential $Exp(1)$ random variables $E_1, ..., E_{M/2}$ and $M/2$ i.i.d uniform $U[0,1]$ random variables $U_1, ..., U_{M/2}$.

3. Compute $Z_t = \exp(2i\pi U_t)\sqrt{F_t E_t}$, for $t = 1, ..., M/2$.

4. Form the M-vector: $\tilde{Z} = (0, Z_1, ...Z_{(M/2)-1}, Z_{M/2}, \overline{Z}_{(M/2)-1}, ..., \overline{Z}_1)$.

5. Compute the inverse fast Fourier transform of the complex Z to obtain the simulated sample path.

Stoev and Taqqu (2004) generate the approximation of fractional stable noise. They introduce parameters $n, N \in \aleph$, and let the fractional stable noise $Y(t)$ be expressed as

$$Y_{n,N}(t) := \sum_{j=1}^{nN}\Big((\frac{j}{n})_+^{H-1/\alpha} - (\frac{j}{n} - 1)_+^{H-1/\alpha}\Big)L_{\alpha,n}(nt - j), \tag{25}$$

[4] Other methods for estimating the parameters of a stable distribution (i.e., the method of moments based on the characteristic function, the regression-type method, and the fast Fourier transform method) are discussed in Stoyanov and Racheva-Iotova (2004a, 2004b, 2004c).

where $L_{\alpha,n}(t) := M_\alpha((j+1)/n) - M_\alpha(j/n)$, $j \in \Re$. The parameter n is mesh size and the parameter M is the cut-off of the kernel function. Stoev and Taqqu (2004) describe an efficient approximation involving the fast Fourier transform algorithm for $Y_{n,N}(t)$. Consider the moving average process $Z(m)$, $m \in \aleph$,

$$Z(m) := \sum_{j=1}^{nM} g_{H,n}(j) L_\alpha(m-j), \tag{26}$$

where

$$g_{H,n}(j) := \Big((\frac{j}{n})^{H-1/\alpha} - (\frac{j}{n} - 1)_+^{H-1/\alpha} \Big) n^{-1/\alpha}, \tag{27}$$

and where $L_\alpha(j)$ is the series of i.i.d standard stable Paretian random variables. Since $L_{\alpha,n}(j) \stackrel{d}{=} n^{-1/\alpha} L_\alpha(j), j \in \Re$, equations (25) and (26) imply $Y_{n,N}(t) \stackrel{d}{=} Z(nt)$, for $t = 1, ..., T$. Then, the computing of $Y_{n,N}(t)$ is transferred to focus on the moving average series $Z(m)$, $m = 1, ..., nT$. Let $\tilde{L}_\alpha(j)$ be the $n(N+T)$-periodic with $\tilde{L}_\alpha(j) := L_\alpha(j)$, for $j = 1, ..., n(N+T)$ and let $\tilde{g}_{H,n}(j) := g_{H,n}(j)$, for $j = 1, ..., nN$; $\tilde{g}_{H,n}(j) := 0$, for $j = nN+1, ..., n(N+T)$. Then

$$\{Z(m)\}_{m=1}^{nT} \stackrel{d}{=} \Big\{ \sum_{j=1}^{n(N+T)} \tilde{g}_{H,n}(j) \tilde{L}_\alpha(n-j) \Big\}_{m=1}^{nT}, \tag{28}$$

because for all $m = 1, ..., nT$, the summation in equation (26) involves only $L_\alpha(j)$ with indices j in the range $-nN \le j \le nT - 1$. Using a circular convolution of the two $n(N+T)$-periodic series $\tilde{g}_{H,n}$ and $\tilde{L}_\alpha$ computed by using the Stoev-Taqqu discrete Fourier transform, the variables $Z(n)$, $m = 1, ..., nT$ (i.e., the fractional stable noise), can be generated.

5.2 Simulation of the multi-dimensional copulas

Embrechts *et al.* (2003) suggest a simulation method for the n-dimension Gaussian copula and Student's t copula. For the Gaussian copula, the algorithm is:

1. Find the Cholesky decomposition A of correlation matrix ρ.
2. Simulate n independent random variates $y_1, \ldots, y_n$ from $\mathcal{N}(0,1)$.
3. Set $\mathbf{z} = A\mathbf{y}$.
4. Set $u_i = \Phi(z_i)$ for $i = 1, \ldots, n$.
5. $(u_1, \ldots, u_n)^T \sim C_\rho^N$.

For the Student's t copula, the algorithm is:

1. Find the Cholesky decomposition A of correlation matrix ρ.
2. Simulate n independent random variates $y_1, \ldots, y_n$ from $\mathcal{N}(0,1)$.
3. Simulate a random variate α from χ_ν^2 independent of $y_1, \ldots, y_n$.
4. Set $\mathbf{z} = A\mathbf{y}$.

5. Set $\mathbf{x} = \frac{\sqrt{\nu}}{\sqrt{\alpha}}\mathbf{y}$.

6. Set $u_i = t_\nu(x_i)$ for $i = 1, \ldots, n$.

7. $(u_1, \ldots, u_n)^T \sim C^t_{\nu,\rho}$.

These algorithms have been adopted in Section 6 for the empirical research in this paper.

6. Empirical results

Table 1 shows the descriptive statistics for the nine international stock indexes in our study. All returns for the indexes used in this study are calculated as

$$y_{i,t} = 100 \times \log\Big(\frac{P_{i,t}}{P_{i,t-1}}\Big).$$

From the statistics reported in this table, it can be seen that excess kurtosis exists. Figure 1 shows the movement of the nine stock indexes. From this figure, the co-movement can be observed.

For the return of each stock index in our study, we denote N as the sample length, sub-sample series that have been randomly selected by a moving window with length T ($1 \le T \le N$). Replacement is allowed in the sampling. In the empirical analysis, sub-sample length (i.e., the window length) of $T = 1$ month was chosen. A total of 1,800 (200 sub-samples for each stock index) sub-samples were randomly created.

Engle (1982) proposes a Lagrange-multiplier test for ARCH phenomenon. A test statistic for ARCH of lag order q is given by

$$X_q \equiv nR_q^2,$$

where R_q^2 is the non-centered goodness-of-fit coefficient of a qth-order autoregression of the squared residuals taken from the original regression

$$\hat{u}_t^2 = \omega_0 + \omega_1\hat{u}_{t-1}^2 + \omega_2\hat{u}_{t-2}^2 + \cdots + \omega_q\hat{u}_{t-q}^2 + e_t, \tag{29}$$

The $\hat{u}$ in the above equation is the residual in the original regression equation. Under the null hypothesis of the residuals of the original model being normally i.i.d., the ARCH statistic of lag order q follows a χ^2 distribution with q degree of freedom:

$$\lim_{n\to\infty} X_q \sim \chi_q^2.$$

Table 2 shows the test statistics and the critical values to reject the null hypothesis that there is no ARCH effect at different lag levels. It is clear from the results reported in the table that an ARCH effect is exhibited in these return series.

We use the Ljung-Box-Pierce Q-statistic based on the autocorrelation function to test for serial correlation (i.e., the memory effect). The Q-statistic is given as follows:

$$Q :\sim \chi_m^2 = N(N+2)\sum_{k=1}^{m}\frac{\rho_k^2}{N-k}, \tag{30}$$

where N denotes the sample size, m the number of autocorrelation lags included in the statistic, and ρ_k the sample autocorrelation at lag order k which is

$$\rho_k = \frac{\sum_{t=1}^{N-k} y_t y_{t+k}}{\sum_{t=1}^{N} y_t^2}. \tag{31}$$

Ljung and Box (1978) show that the Q-statistic follows an asymptotic chi-square distribution.

Table 3 shows that the null hypothesis that there is no serial correlation can be rejected at different lags. The table shows that the memory effect occurs for each index return series. In order to see when the memory effect vanishes, we compare the Q-statistic with its corresponding critical value. When the quotient of the Q-statistic and the corresponding critical value are less than 1, we cannot reject the null hypothesis that there is no serial correlation. From Table 3, we find that the quotient of the Q-statistic to its corresponding critical value exceeds unity. We can therefore reject the null hypothesis that there is no serial correlation and can say that long-range dependence is exhibited by our dataset.

Table 4 reports the parameters estimated from the ARMA(1,1)-GARCH(1,1) assuming that residuals are identically and independently normally distributed with zero mean and unit variance. Based on equation (9), we generate the empirical residuals. The descriptive statistic of the empirical residuals $\tilde{u}_t$ is shown in Table 5. The results reported in the table make it clear that excess kurtosis still exists and the residuals do not follow i.i.d. $N(0,1)$ distribution.

Table 6 shows the parameters estimated for empirical residuals $\tilde{u}_t$ based on the methods introduced in Section 4. As we mentioned in that section, the Hurst index for non-Gaussian stable processes has different bounds for "persistence" and "anti-persistence". For tail index $\alpha \in (0, 2)$, when $H \in (0, 1/\alpha)$, the processes exhibit "anti-persistence", and when $H \in (1/\alpha, 1)$, the processes exhibit "persistence". There is no long-range dependence when $\alpha \in (0, 1]$ because the Hurst index is bounded in the interval $(0, 1)$. When $H = 1/\alpha$, depending on the value of α, the processes exhibit either no memory or long-range dependence. From Tables 1 and 6, we find that the Hurst index has no value that is equal to $1/\alpha$. Therefore, we find that long-range dependence occurs in our dataset.

The AD and KS statistics were calculated for the six candidate distributional assumptions. Table 7 reports the descriptive statistics of the computed AD, KS, and CVM statistics. As can be seen in the table, ARMA-GARCH with a fractional Gaussian noise model exhibits a smaller mean value for the AD, KS, and CVM statistics than the other five models. Figure 2 shows the boxplot of AD statistics for the six alternative ARMA-GARCH models investigated. Figure 3 shows the boxplot of KS statistics and Figure 4 shows the boxplot of CVM statistics. ARMA-GARCH with fractional Gaussian noise model exhibits smaller mean and less outliers, demonstrating the advantage of this model.

As can be seen from Table 1, the index returns clearly do not follow the Gaussian distribution. Stable parameters in this table also exhibit the non-Gaussian characteristic since the stable parameter is equal to 2 for the Gaussian case. The heavy tailedness can be easily observed in the data. Accordingly it seems that the application of heavy-tailed distributions should perform better than the Gaussian distribution or fractional Gaussian noise. Our empirical result found by simulating the marginal distribution indicates that the fractional Gaussian noise subordinated in the ARMA-GARCH model fits better than other alternatives. We believe that there are two reasons for this. First, the heavy tailendness of index returns stems from the heavy tailedness of each index component stock. The equity market indexes aggregate those stocks that exhibit heavy tailedness. The aggregation of heavy-tailed distributions is asymptotically self-similar, and the fractional Gaussian noise is a typical stochastic process with self-similarity. The heavy-tailedness effect is considered in the self-similarity. Second, after the aggregation, although equity market indexes exhibit heavy tailedness, the influence of such effect (non-Gaussian and heavy tailedness) in the movement of the market index is weak compared to long-range dependence and volatility clustering.

From the goodness of fit testing for marginal distributions, we find the best fit model is ARMA-GARCH with a fractional Gaussian noise model. Then taking ARMA-GARCH with a fractional Gaussian noise model as marginal distributions for each stock index return, we simulate multivariate Gaussian copula (in our empirical study, 9 dimensions for our data) and Student's t copula for the dependence structure of these nine index returns. Table 8 shows the descriptive statistics of the computed AD, KS, and CVM statistics for 200 sub-sample matrices with nine marginal distributions as the column vectors. In this table, the Student's t copula exhibits smaller mean values for the computed AD, KS, and CVM statistics than the Gaussian copula, indicating the better performance.

7. Conclusions

There is considerable interest in the co-movement of international equity markets. The linear correlation measure is not satisfactory to discover the dependence structure between equity markets. With several advantages, copulas are regarded as the ideal measure to model both the degree and structure of dependence. Some works are based on the bivariate co-movement. In this paper, we use the copula ARMA-GARCH model to capture the multivariate co-movement among the international equity markets in our study.

In our empirical analysis, we investigate a ARMA-GARCH model with six forms for the residuals (fractional stable noise, fractional Gaussian noise, stable distribution, white noise, generalized Pareto distribution, and generalized extreme value distribution) for modeling the marginal distribution for the nine international equity market indexes. By using parameters estimated from the empirical series, we simulate a series for each index returns with these six different modeling structures. Then we compare the goodness of fit for these generated series to the empirical series by adopting three criteria for the goodness of fit test: the Kolmogorov-Simirnov distance, Anderson-Darling distance, and Cramer von Mises distance. Based on a comparison of these criteria, the empirical evidence shows that the ARMA-GARCH model with fractional Gaussian noise demonstrates better performance in modeling marginal distributions.

Using an ARMA-GARCH model with fractional Gaussian noise, we simultaneously simulate the nine index returns with both the Gaussian copula and Student's t copula. By using the same criteria of goodness of fit test in comparing marginal distributions, we find that the Student's t copula is better than the Gaussian copula when modeling the multivariate co-movement of these nine equity markets. The reason is that Student's t copula can capture the tail dependence among these index returns for both positive and negative extreme values, while the Gaussian copula cannot.

The findings reported in this paper should be taken into account in modeling the co-movement of global equity markets for several reasons. First, using multi-dimension copulas rather than bivariate copulas can reveal the simultaneous co-movement of several markets. Second, when modeling the marginal distribution of each market returns, our model can capture long-range dependence, heavy tails, and volatility clustering simultaneously. Third, using high-frequency data, the impact of both macroeconomic factors and microstructure effects on each market can be considered. Our model reveals that similar macroeconomic factors impact the co-movement of international markets and that investors' behaviors in each market are similar, especially their reactions in each market to world news are similar. With our model, more accurate prediction is possible for the simultaneous co-movement of several equity markets.

Appendix A

The unconditional Gaussian copula and unconditional Student's t copula are specified in this section. The multivariate version of these two copulas are given as follows. Let ρ be the correlation matrix which is a symmetric, positive definite matrix with unit diagonal, and Φ_ρ the standardized multivariate normal distribution with correlation matrix ρ. The unconditional multivariate Gaussian copula is then

$$C(u_1, \ldots, u_n; \rho) = \Phi_\rho\Big(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_n)\Big),$$

and the corresponding density is

$$c(u_1, \ldots, u_n; \rho) = \frac{1}{|\rho|^{1/2}} \exp\Big(-\frac{1}{2}\lambda^T(\rho^{-1} - I)\lambda\Big),$$

where $\lambda = (\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_n))^T$ and u_n is the margins.

The unconditional (standardized) multivariate Student's copula $T_{\rho,\nu}$ can be expressed as

$$T_{\rho,\nu}(u_1, \ldots, u_n; \rho) = t_{\rho,\nu}\Big(t_\nu^{-1}(u_1), \ldots, t_\nu^{-1}(u_n)\Big),$$

where $t_{\rho,\nu}$ is the standardized multivariate Student's t distribution with correlation matrix ρ and ν degrees of freedom and t_ν^{-1} is the inverse of the univariate cumulative density function (c.d.f) of the Student's t with ν degrees of freedom. The density of the unconditional multivariate Student's t copula is

$$c_{\rho,\nu}(u_1, \ldots, u_n; \rho) = \frac{\Gamma(\frac{\nu+n}{2})}{\Gamma(\frac{\nu}{2})|\rho|^{1/2}} \Big(\frac{\Gamma(\frac{\nu}{2})}{\Gamma(\frac{\nu+1}{2})}\Big)^n \left(\frac{\Big(1 + \frac{1}{\nu}\lambda^T\rho^{-1}\lambda\Big)^{-\frac{\nu+n}{2}}}{\prod_{j=1}^n \Big(1 + \frac{\lambda_j^2}{\nu}\Big)^{-\frac{\nu+1}{2}}}\right),$$

where $\lambda_j = t_\nu^{-1}(u_j)$ and and u_n is the margins.

References

[1] Aït-Sahalia, Y., 2004. Estimating diffusions with discretely and possibly randomly spaced data: A general theory. *Annals of Statistics* **32**: 2186-2222.

[2] Andersen T.G, Bollerslev T, Diebold F.X, Ebens H. 2001. The distribution of realized stock return volatility. *Journal of Financial Economics* **61**: 43-76.

[3] Baillie, R.T. 1996. Long memory processes and fractional integration in econometrics. *Journal of Econometrics* **73**: 5-59.

[4] Bardet J, Lang G, Oppenheim G, Philippe A, Taqqu M. 2003. Generators of long-range dependent processes: a survey. In *Theory and Applications of Long-range Dependence*, Doulkhan P, Oppenheim G, Taqqu M (eds). Birkhäuser: Boston.

[5] Beran, J. 1994. *Statistics for Long-Memory Processes*. Chapman & Hall: New York.

[6] Bollerslev, T., and J. H. Wright. 2000. Semiparametric estimation of long-memory volatility dependencies: The role of high-frequency data. *Journal of Econometrics* **98**: 81-106.

[7] Cherubini U, Luciano E, Vecchiato W. 2004. *Copula Methods in Finance*. John Wiley & Sons: West Sussex.

[8] Dacorogna M, Gençay R, Müller U.A, Olsen R, Pictet O. 2001. *An Introduction of High-Frequency Finance*. Academic Press: San Diego.

[9] Deheuvels P. 1979. La fonction de dependance empirique et ses proprietes, Un test non paramtrique d'independance. *Acad. Roy. Belg. Bull. C1. Sci* **65**: 274-292.

[10] Embrechts P, Lindskog F, McNeil A. 2003. Modelling dependence with copulas and applications to risk management. In *Handbook of Heavy Tailed Distributions in Finance*, Rachev S (ed). Elsevier: Amsterdam.

[11] Engle R, 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica* **50**: 987-1008.

[12] ——, Russell J. 1998. Autoregressive conditional duration: a new model for irregularly spaced transaction data. *Econometrica* **66**: 1127-1162.

[13] Fama E. 1963. Mandelbrot and the stable Paretian hypothesis. *Journal of Business* **36**: 420-429.

[14] ——. 1965. The behavior of stock market prices. *Journal of Business* **38**: 34-105.

[15] ——, Roll R. 1971. Parameter estimates for symmetric stable distributions. *Journal of the American Statistical Association* **66**: 331-338.

[16] Genest C, Rivest L. 1993. Statistical inference procedures for bivariate Archimedean copulas. *Journal of American Statistical Association* **88**: 1034-1043.

[17] Hols M, de Vries C. 1991. The limiting distribution of extremal exchange rate returns. *Journal of Applied Econometrics* **6**: 287-302.

[18] Hurst H. 1951. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers* **116**: 770-808.

[19] ——. 1955. Methods of using long-term storage in reservoirs. *Proceedings of the Institution of Civil Engineers* **1**: 519-577.

[20] Joe H. 1997. *Multivariate Models and Dependence Concepts*. Chapman& Hall: London.

[21] Mandelbrot B. 1963. New methods in statistical economics. *Journal of Political Economy* **71**: 421-440.

[22] ——. 1983. *The Fractal Geometry of Nature*. Freeman: San Francisco.

[23] ——, Van Ness J. 1968. Fractional Brownian motions, fractional noises and applications. *SIAM Review* **10**: 422-437.

[24] Mandelbrot B, Wallis J. 1968. Noah, Joseph and operational hydrology. *Water Resources Research* **4**: 909-918.

[25] Maejima M. 1983. On a class of self-similar processes. *Zeitschrift für Vahrscheinlichkeitstheorie und verwandte Gebiete* **62**: 235-245.

[26] ——, Rachev S. 1987. An ideal metric and the rate of convergence to a self-similar process. *Annals of Probability* **15**: 702-727.

[27] Mansfield P, Rachev S, Samorodnitsky G. 2001. Long strange segments of a stochastic process and long-range dependence. *The Annals of Applied Probability* **11**: 878-921.

[28] Mittnik S, Rachev S. 1993. Modeling asset returns with alternative stable models. *Econometric Reviews* **12**: 261-330.

[29] Müller U, Dacorogna M, Olsen R, Pictet O, Schwarz M, Morgenegg C. 1990. Statistical study of foreign exchange rates, empirical evidence of a price change scaling law, and intraday analysis. *Journal Banking and Finance* **14**: 1189-1208.

[30] Nelsen R. 2006. *An Introduction to Copulas*(Second Edition). Springer: New York.

[31] Patton A. 2005. Modelling asymmetric exchange rate dependence. *International Economic Review*, *forthcoming*.

[32] Paxson V. 1997. Fast, approximate synthesis of fractional Gaussian noise for generating self-similar network traffic. *Computer Communication Review* **27**: 5-18.

[33] Rachev S. 2003. *Handbook of Heavy Tailed Distributions in Finance*(ed). Elsevier: Amsterdam.

[34] ——, Mittnik S. 2000. *Stable Paretian Models in Finance*. Wiley: New York.

[35] Rachev S, Samorodnitsky G. 2001. Long strange segments in a long range dependent moving average. *Stochastic Processes and their Applications* **93**: 119-148.

[36] Rachev S, Menn C, Fabozzi F. 2005. *Fat-Tailed and Skewed Asset Return Distributions*. Wiley: New Jersey.

[37] Samorodnitsky G. 1994. Possible sample paths of self-similar alpha-stable processes. *Statistics and Probability Letters* **19**: 233-237.

[38] ——, Taqqu M. 1994. *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. Chapman & Hall/CRC: Boca Raton.

[39] Sklar A. 1959. Fonctions de Réparitition à n dimensions et Leurs Marges, *Publications de l'Institut de Statistique de l'Université de Paris* **8**: 229-231.

[40] Solnik B, Boucrelle C, Le Y.F. 1996. International market correlation and volatility. *Financial Analysts Journal* **52**: 17-34.

[41] Stoev S, Pipiras V, Taqqu M. 2002. Estimation of the self-similarity parameter in linear fractional stable motion. *Signal Processing* **82**: 1873-1901.

[42] Stoev S, Taqqu M. 2004. Simulation methods for linear fractional stable motion and FARIMA using the Fast Fourier Transform. *Fractals* **12**: 95-121.

[43] Stoyanov S, Racheva-Iotova B. 2004a. Univariate stable laws in the field of finance—approximation of density and distribution functions. *Journal of Concrete and Applicable Mathematics* **2**: 38-57.

[44] ——. 2004b. Univariate stable laws in the field of finance—parameter estimation. *Journal of Concrete and Applicable Mathematics* **2**: 24-49.

[45] ——. 2004c. Numerical methods for stable modeling in financial risk management. In *Handbook of Computational and Numerical Methods*, S. Rachev (ed). Birkhäuser: Boston.

[46] Wasserfallen W, Zimmermann H. 1985. The behavior of intradaily exchange rates. *Journal of Banking and Finance* **9**: 55-72.

[47] Willinger W, Paxson V, Taqqu M. 1998. Self-similarity and heavy tails: structural modeling of network traffic. In *A Practical Guide to Heavy Tails*, Adler R, Feldman R, Taqqu M (eds). Birkhäuser: Boston.

Table 1: Summary of the statistical characteristics of nine index returns.

	location	mean	std	kurtosis	skewness	mininum	maximum	Hurst	α
AORD	Austrilia	-5.38E-06	0.0213	1903.7000	-0.8007	-2.3168	2.3168	0.5182	1.3864
DAX	Germany	-2.31E-05	0.0371	448.7800	-0.4372	-4.4347	2.9733	0.4978	1.2222
FCHI	France	-5.08E-05	0.0506	732.5400	0.7020	-3.9898	3.9050	0.5065	1.3253
FTSE	UK	-1.53E-05	0.0222	846.8900	0.3852	-2.2313	2.7460	0.5110	1.3270
HSI	China	2.19E-05	0.0486	235.6000	1.3178	-2.3099	2.6161	0.5410	1.4970
KS200	South Korea	2.25E-05	0.0439	1404.1000	4.9465	-3.7141	4.3014	0.5106	1.2961
N225	Japan	3.31E-06	0.0759	121.5700	1.7980	-2.0686	3.4282	0.4821	1.2411
SPX	US	-3.76E-06	0.0169	4478.7000	-14.7220	-3.6071	2.3110	0.5313	1.1819
STOXX	Switzerland	-2.45E-05	0.0247	619.6600	-2.3390	-3.4999	2.7500	0.5010	1.1720

Table 2: Result of the ARCH-test for different lags at $\alpha = 0.05$.

	lag1	*lag2*	*lag5*	*lag10*	*lag15*	*lag20*	*lag25*	*lag30*
AORD	37152	37669	39206	39257	39265	39346	39356	39355
DAX	3910	4103	4565	6678	6745	6898	6914	6932
FCHI	10	12	19	23	29	60	69	75
FTSE	12854	12913	14395	14465	14512	15008	15472	15483
HSI	21	28	37	41	44	47	48	49
KS200	8	21	38	39	39	39	39	39
N225	66	73	633	642	644	646	647	647
SPX	7	11	16	17	17	17	17	18
STOXX	2342	3763	4831	4961	5005	5035	5060	5087
Critical Value	3.8415	5.9915	11.0700	18.3070	24.9960	31.4100	37.6520	43.7730

Table 3: The Ljung-Box-Pierce Q-test statistic for different lags at $\alpha = 0.05$.

	$10-min$	$30-min$	$1hour$	$2hours$	$4hours$	$1day$	$1week$	$1month$
AORD	3202	3276	3375	3516	4245	6493	16943	51385
DAX	43937	44127	44375	44698	45360	47259	60129	105330
FCHI	8220	8461	8737	90051	97211	12711	26178	70851
FTSE	34008	35637	35944	36465	37205	38854	53398	100850
HSI	10490	10851	11162	11619	12324	13623	22593	55679
KS200	1547	1687	1807	2019	2785	3890	15770	52254
N225	2748	2836	2931	3166	3769	4945	13352	43790
SPX	25670	257030	258100	258550	259790	262420	283190	349850
STOXX	167900	168360	168770	169160	1701405	172290	189380	246550
Critical Value	55.7585	146.5673	277.1376	532.0754	1033.1928	2023.0522	9829.0489	38856.9694

Table 4: Estimated parameters of the AMAR(1,1)-GARCH(1,1) model with residuals following normal distribution with zero mean and unit variance. Numbers in parentheses are the standard errors. These parameters are used in the empirical simulation.

	α_0	α_1	β_1	κ	γ_1	θ_1
AORD	3.9724E-07	-0.1952	0.1136	4.6260E-009	0.6486	0.3438
	(9.5130E-08)	*(1.1634E-11)*	*(1.1618E-11)*	*(3.3215E-12)*	*(1.2023E-12)*	*(1.6527E-11)*
DAX	-1.9424E-07	0.5559	-0.3766	1.3826E-008	0.6558	0.3442
	(3.5419E-08)	*(1.1621E-12)*	*(1.3573E-12)*	*(1.7137E-12)*	*(3.4729E-12)*	*(2.2729E-12)*
FCHI	-9.5146E-08	0.5869	-0.4720	2.7586E-08	0.8135	0.1227
	(1.4999E-07)	*(1.0717E-06)*	*(1.2618E-6)*	*(8.4707E-12)*	*(2.7037E-05)*	*(2.1210E-05)*
FTSE	-1.2392E-07	0.8232	-0.7568	6.5616E-09	0.5987	0.2812
	(3.3959E-08)	*(3.0773E-13)*	*(2.3373E-13)*	*(3.7753E-12)*	*(1.2205E-12)*	*(1.2504E-13)*
HSI	-9.8445E-10	0.5154	-0.6893	2.8409E-08	0.6931	0.2610
	(5.5655E-08)	*(2.5658E-04)*	*(3.3054E-04)*	*(4.7913E-11)*	*(3.3037E-04)*	*(3.2846E-04)*
KS200	4.3005E-06	0.0075	-0.2692	1.9238E-08	0.6582	0.3418
	(2.4262E-08)	*(8.6562E-06)*	*(9.6637E-06)*	*(3.0034E-12)*	*(2.5391E-05)*	*(2.1966E-05)*
N225	-4.8124E-06	0.4782	-0.2905	6.6660E-08	0.6170	0.3766
	(4.5034E-07)	*(8.2554E-04)*	*(7.5777E-04)*	*(2.5256E-10)*	*(8.2544E-04)*	*(8.0406E-04)*
SPX	2.5689E-07	0.5548	-0.0262	2.4089E-09	0.8386	0.0619
	(3.1323E-08)	*(3.1166E-12)*	*(2.4869E-15)*	*(4.4352E-14)*	*(2.7531E-11)*	*(5.2184E-11)*
STOXX	1.3438E-07	0.6101	-0.3995	5.8155E-09	0.6677	0.2631
	(7.3838E-08)	*(2.9345E-12)*	*(1.3513E-12)*	*(1.3865E-13)*	*(2.2687E-11)*	*(2.4868E-11)*

Table 5: Summary of the empirical $\tilde{u}_t$.

	mean	*variance*	*kurtosis*	*skewness*
AORD	-0.0027	0.9085	326.1807	2.5378
DAX	0.0017	0.9384	1034.6394	-2.2134
FCHI	-3.6688E-05	0.9196	1032.1787	2.0182
FTSE	0.0038	0.9337	84.3404	-0.5075
HSI	0.0023	0.9522	147.8326	0.7201
KS200	-0.0170	1.0982	1379.5626	6.9291
N225	0.0103	1.0463	74.6847	0.6791
SPX	-0.0026	1.0008	14390.3118	-40.7157
STOXX	-0.0011	0.9487	6051.9163	-6.9826

Table 6: Parameters estimated from the empirical $\tilde{u}_t$.

	$Hurst_{FGN}$	$Hurst_{fsn}$	α	β	γ	ζ
AORD	0.5366	0.5791	1.8777	-0.2610	0.4765	0.0063
DAX	0.5619	0.5492	1.3976	-0.0103	0.3765	-0.0026
FCHI	0.5476	0.5107	1.4688	-0.0153	0.3508	-0.0039
FTSE	0.5544	0.5112	1.3538	-0.0126	0.4032	0.0014
HSI	0.4588	0.6376	1.0914	-0.0016	0.2550	-9.8442E-04
KS200	0.5168	0.5594	1.4387	0.0147	0.3869	-0.0195
N225	0.5554	0.5260	1.3326	-0.0234	0.3629	0.0021
SPX	0.5278	0.5149	1.3533	-0.0113	0.3535	-0.0073
STOXX	0.5787	0.5604	1.3847	-5.3461E-04	0.3852	-0.0019

Table 7: Summary of the AD, KS and CVM statistics for alternative models for marginal distribution. Mean, median, standard deviation ("std"), maximum value ("max"), minimum value ("min") and range of the AD, KS and CVM statistics are presented in this table. "FGN" stands for fractional Gaussian noise, "fsn" for fractional stable noise, "normal" for white noise, "stable" for stable distribution, "gev" for generalized extreme value distribution, and "gpd" for generalized Pareto distribution.

	AD_{mean}	AD_{median}	AD_{std}	AD_{max}	AD_{min}	AD_{range}
$ARMA-GARCH_{FGN}$	54.8381	54.8220	0.3222	56.2561	53.5083	2.7482
$ARMA-GARCH_{fsn}$	54.8459	54.7861	0.7213	63.4861	51.7917	11.6952
$ARMA-GARCH_{normal}$	55.0686	54.9050	0.8139	66.1060	53.9640	12.1425
$ARMA-GARCH_{stable}$	55.7993	55.1920	2.2074	77.7650	52.8670	24.9051
$ARMA-GARCH_{gev}$	55.4836	55.2240	1.2978	73.6110	53.4447	20.1677
$ARMA-GARCH_{gpd}$	76.7643	70.2010	18.1530	109.5412	45.5832	63.9523
	KS_{mean}	KS_{median}	KS_{std}	KS_{max}	KS_{min}	KS_{range}
$ARMA-GARCH_{FGN}$	0.5017	0.5012	0.0025	0.5136	0.4966	0.0170
$ARMA-GARCH_{fsn}$	0.5032	0.5016	0.0058	0.5855	0.4945	0.0910
$ARMA-GARCH_{normal}$	0.5039	0.5020	0.0075	0.6035	0.4965	0.1070
$ARMA-GARCH_{stable}$	0.5116	0.5053	0.0202	0.7103	0.4948	0.2155
$ARMA-GARCH_{gev}$	0.5079	0.5052	0.0121	0.6721	0.4967	0.1754
$ARMA-GARCH_{gpd}$	0.7059	0.6476	0.1646	1.0000	0.4309	0.5691
	CVM_{mean}	CVM_{median}	CVM_{std}	CVM_{max}	CVM_{min}	CVM_{range}
$ARMA-GARCH_{FGN}$	502.4011	500.1612	5.4108	535.6220	498.5100	37.1120
$ARMA-GARCH_{fsn}$	502.4176	500.2272	5.6007	545.4330	497.8300	47.6120
$ARMA-GARCH_{normal}$	502.9730	500.3900	6.3452	568.2810	498.6210	69.6620
$ARMA-GARCH_{stable}$	505.6970	500.9900	16.8300	773.8670	497.6620	276.1900
$ARMA-GARCH_{gev}$	503.8650	500.8452	10.4230	684.7800	498.2801	186.4900
$ARMA-GARCH_{gpd}$	910.7413	627.1045	551.4111	1999.7110	388.6400	1611.1001

Table 8: Summary of the AD, KS and CVM statistics for alternative models for joint distribution. Mean, median, standard deviation ("std"), maximum value ("max"), minimum value ("min") and range of the AD, KS and CVM statistics are presented in this table.

	AD_{mean}	AD_{median}	AD_{std}	AD_{max}	AD_{min}	AD_{range}
Gaussian copula	0.9241	0.9374	0.0338	0.9718	0.8370	0.1348
Student's t copula	0.9237	0.9362	0.0340	0.9716	0.8382	0.1334
	KS_{mean}	KS_{median}	KS_{std}	KS_{max}	KS_{min}	KS_{range}
Gaussian copula	48.4519	55.5841	16.3230	67.9456	9.6306	58.3150
Student's t copula	48.4470	55.5060	16.3190	67.9740	9.9158	58.0580
	CVM_{mean}	CVM_{median}	CVM_{std}	CVM_{max}	CVM_{min}	CVM_{range}
Gaussian copula	785.6190	798.7101	24.5134	817.5083	729.5811	87.9272
Student's t copula	785.2964	798.1155	24.6323	817.9235	728.6673	89.2562

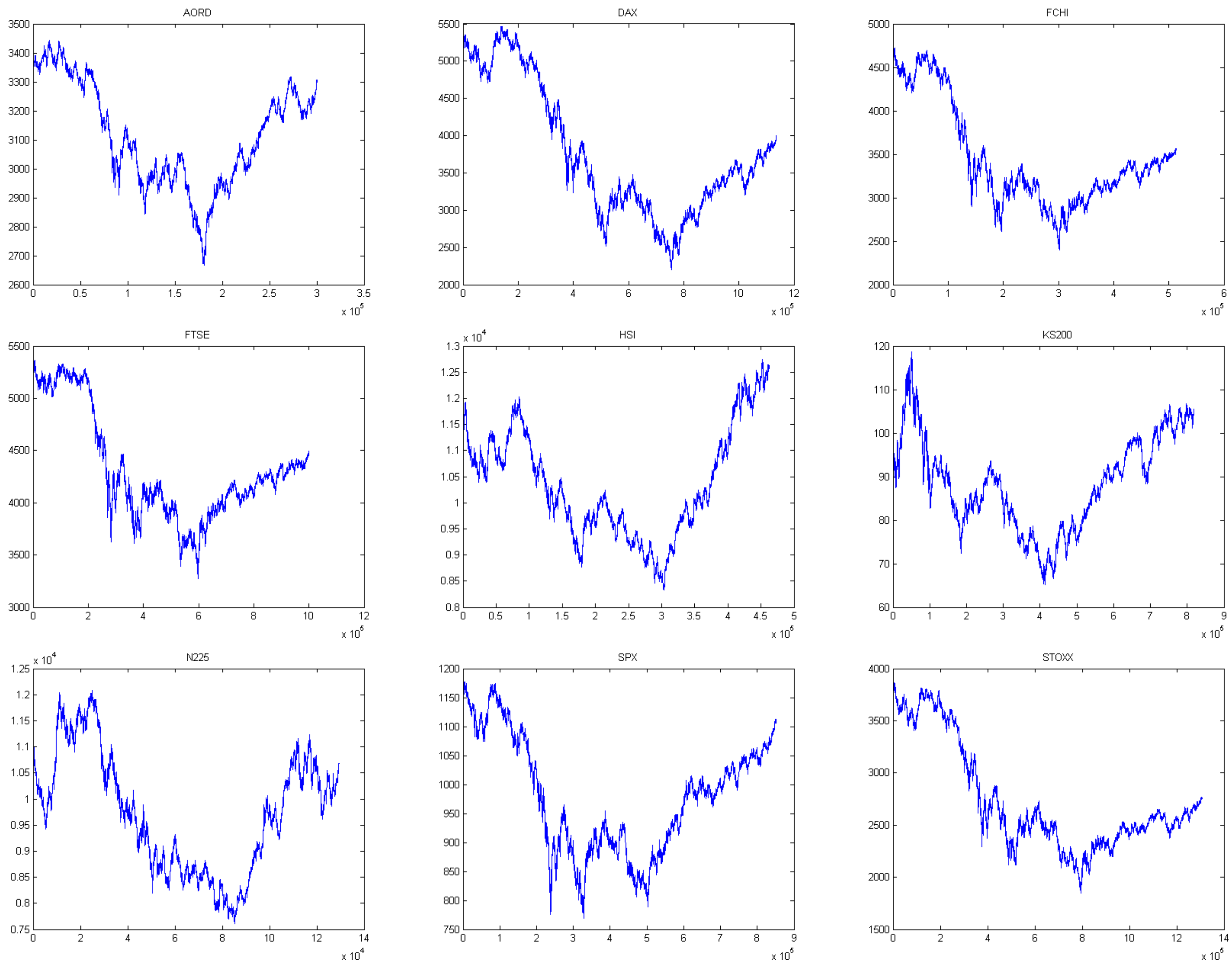

Figure 1: Plot of Index Movements

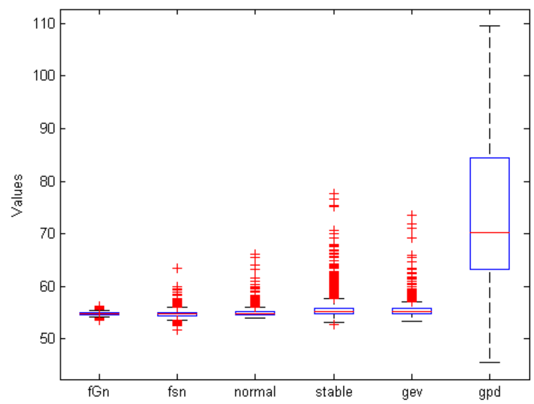

Figure 2: Boxplot of AD statistics of modeling marginal distribution with alternative residual distributions.

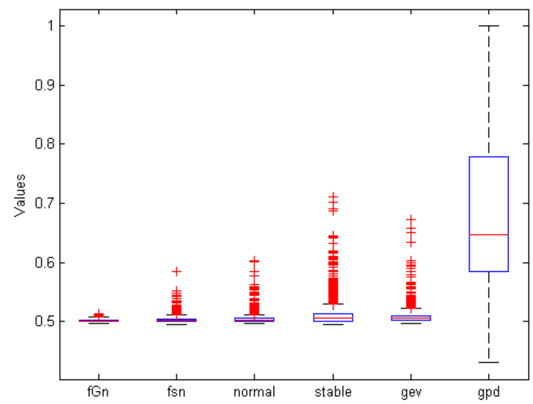

Figure 3: Boxplot of KS statistics of modeling marginal distribution with alternative residual distributions.

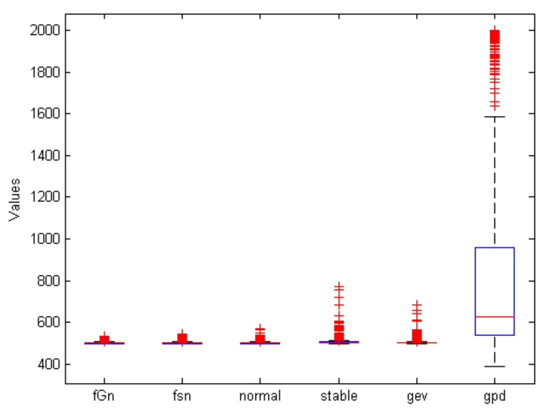

Figure 4: Boxplot of CVM statistics of modeling marginal distribution with alternative residual distributions.