

SFB 649 Discussion Paper 2006-086

# Overreaction and Multiple Tail Dependence at the High-frequency Level — The Copula Rose

Wing Lon Ng\*



\* Institute of Econometrics and Economic Statistics,  
University of Münster, Germany

This research was supported by the Deutsche  
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

<http://sfb649.wiwi.hu-berlin.de>  
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin  
Spandauer Straße 1, D-10178 Berlin



SFB 649 ECONOMIC RISK BERLIN

# Overreaction and Multiple Tail Dependence at the High-frequency Level – The Copula Rose

Wing Lon Ng \*

University of Muenster

Institute of Econometrics and Economic Statistics

Am Stadtgraben 9, 48143 Muenster, Germany

E-mail: 05wing@wiwi.uni – muenster.de

December 17, 2006

## Abstract

This paper applies a non- and a semiparametric copula-based approach to analyze the first-order autocorrelation of returns in high frequency financial time series. Using the EUREX *D3047* tick data from the German stock index, it can be shown that the temporal dependence structure of price movements is not always negatively correlated as assumed in the stylized facts in the finance literature. Depending on the sampling frequency, the estimated copulas exhibit some kind of overreaction phenomena and multiple tail dependence, revealing patterns similar to the compass rose.

**Key Words:** high frequency data, non- and semiparametric copulas, overreaction, tail dependence, compass rose.

**JEL Classifications:** *C14, C22, G14.*

---

\*The author is very grateful to Mark Trede for his valuable comments and to seminar participants at the University of Muenster for their helpful discussion. This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 “Economic Risk” at the Humbolt University Berlin.

# 1 Introduction

The literature on financial econometrics and quantitative finance has been used to focus on the stochastic process of daily prices or returns on assets and their volatility. With the increased availability of ultra-high-frequency orderbook data in the last few years, researchers have now become interested in the price process and its realized volatility at transaction level (see, for example, Andersen, Bollerslev, and Meddahi (2005) and Russell and Engle (2005)).

However, one stylized fact that needs more accurate re-investigation concerns the negative first-order autocorrelation of returns, often observed in studies on financial time series of assets (Dacorogna, Gençay, Müller, B., and V. (2001)). According to the transaction model of Roll (1984) for stock markets, this negative autocorrelation is caused by the so-called bid-ask-bounce. Indeed, Goodhart (1991) have found empirical evidence for the existence of negative first-order correlation of returns. Bollerslev and Domowitz (1993), for example, describe this phenomena as an outcome of market makers skewing the spread into particular direction when they have order imbalances.

In fact, most former studies in this field only applied the common autocorrelation coefficient that can only measure the “aggregated” linear dependence and, thus, simply neglects other (possibly important) side-effects. Other attempts to discover nonlinearities in financial data are the so-called phase portraits, often used in dynamical systems in physical sciences to detect chaotic phenomena (Szpiro (1998)). In its simplest version, it represents a scatterplot, in which a time series is plotted against its lagged values (for a overview, see Wöhrmann (2005)). In contrast, this paper applies a flexible copula-based approach, a more general modeling method, for describing the returns of high-frequency EUREX tick data. For a better understanding of the dynamic behavior of the stochastic process, the main objective is to study the temporal dependence structure of price movements with non- and semi-parametric copulas in order to account for nonlinear and partial relationships as well.

Since time series generally can be seen as a drawing from a multivariate distribution, one may split this distribution into two components: (a) the marginal distributions and (b) the dependence structure determined by the copula. In this paper, the methodology focuses on the univariate stationary return process, in which the copulas control the temporal dependence of the time series, whereby the unconditional distributions are left unspecified, allowing all kinds of possible margins. In the first step, nonparametric copulas are applied to detect the first order temporal dependence of the data exploratively. In contrast to the compass rose approach, semiparametric cop-

ulas are estimated in the second step, not to measure the “quality” of the dependence pattern (Wang and Wang (2002)), but to quantify the different relationships of the consecutive returns, which is a matter of particular interest to economists. With these features at hand, this approach is able to capture more general nonlinear (and also partial) dependence of the stochastic process.

Finally, it is to emphasize that the effect of sampling frequency of financial time series data was often not taken into account (see Cai, Hudson, and Keasey (2003)). Regarding the results of Aït-Sahlia and Mykland (2003), this paper does not only investigate the original data observed at the 1 second interval, but also considers the return process at different aggregation levels in order to reveal possible effects of market microstructure noise.

The outline of this paper is structured as follows: In Section 2, the copula approach will be introduced. Section 3 describes the model estimation. In Section 4, the data and results are presented. Section 5 concludes.

## 2 The Copula Approach

Most existing papers in the finance literature using copulas are often interested in modeling the contemporaneous dependence between two or more several random variables (Fermanian and Scaillet (2005)). In contrast, this paper will focus on modeling the temporal dependence structure of a time series  $\{X_i\}_{i=1}^N$  via copulas (see Chen and Fan (2006a), Chen and Fan (2006b) and Patton (2006)).

A copula is a multivariate distribution, whose marginal distributions are uniform on the interval  $(0, 1)$ . The momentousness of copulas in modeling multivariate distributions has been stated in the famous theorem by Sklar (1959). Since this study is interested in the first-order autocorrelation of price differences, we only pay attention to the bivariate case. This section first briefly reviews the general copula theory and then extends this concept to the time series context. (For a comprehensive survey of the theory of copulas, the reader is referred to the textbooks of Joe (1997), Mari and Kotz (2001) and Nelsen (1990)).

Consider two random variables  $X$  and  $Y$  with continuous univariate distribution functions  $F_X(x) = P(X \leq x)$  and  $F_Y(y) = P(Y \leq y)$  and their joint distribution function  $F_{X,Y}(x, y) = P(X \leq x, Y \leq y)$ . Sklar’s theorem states that there exists a function called copula  $C$  that connects the univariate distributions  $F_X$  and  $F_Y$  to a bivariate distribution function.

**Theorem** *Let  $X$  and  $Y$  be random variables with marginal distribution functions  $F_X$  and  $F_Y$ , and joint distribution function  $F_{X,Y}$ . Then there exists*

a bivariate copula  $C : [0, 1]^2 \rightarrow [0, 1]$  such that for all  $x, y$  in  $\mathbb{R} \cup \{-\infty, +\infty\}$

$$F_{X,Y}(x, y) = C(F_X(x), F_Y(y)) \quad . \quad (1)$$

If the margins  $F_X$  and  $F_Y$  are continuous, then  $C$  is unique. Conversely, if  $C$  is a copula and  $F_X$  and  $F_Y$  are distribution functions, then the function  $F_{X,Y}$  defined in (1) is a joint distribution function with margins  $F_X$  and  $F_Y$ .

In case of continuous bivariate distributions, Sklar's theorem shows that the univariate margins can be separated from the copula which completely defines the dependence structure between them. In other words, the random variables  $X$  and  $Y$  have a copula  $C$  given by (1).

**Corollary** Let  $F_{X,Y}$  be a bivariate distribution function with continuous margins  $F_X$  and  $F_Y$  and copula  $C$ . Then for any  $u, v$  in  $[0, 1]^2$

$$C(u, v) = F_{X,Y}(F_X^{-1}(u), F_Y^{-1}(v)) \quad , \quad (2)$$

where  $F_X^{-1}(u)$  is the quantile function given by  $F_X^{-1}(u) = \inf \{x : F_X(x) \geq u\}$ , respectively for  $F_Y^{-1}(v)$ .

This corollary represents a construction method for bivariate distributions via the copula approach. The copula  $C$  is the bivariate joint distribution function of the transformed random variables  $U = F_X(X)$  and  $V = F_Y(Y)$ , i.e.

$$C(u, v) = P(U \leq u, V \leq v) \quad . \quad (3)$$

Deriving the copula function, one can obtain the conditional copula and the density of the copula: taking the first derivative of the copula function yields the conditional copula of  $U$  given  $V = v$

$$C_{U|V=v}(u) = \frac{\partial}{\partial v} C(u, v) \quad , \quad (4)$$

and if the copula function is twice differentiable, then the copula density is

$$c(u, v) = \frac{\partial^2 C(u, v)}{\partial u \partial v} \quad .$$

Similar to a common cumulative distribution function (cdf), it is also possible to define the survival copula function  $\bar{C}$  as a link between the univariate survival functions given by  $S_X(x) = P(X > x) = 1 - F_X(x)$ , respectively for  $S_Y(y)$ , and the joint survival function  $S_{X,Y}(x, y) = P(X > x, Y > y)$  in the following way:

$$S_{X,Y}(x, y) = \bar{C}(S_X(x), S_Y(y)) \quad .$$

The copula  $C$  and the survival copula  $\bar{C}$  are related through

$$\bar{C}(u, v) = u + v - 1 + C(1 - u, 1 - v)$$

and similarly the densities of the copula ( $c$ ) and the survival copula ( $\bar{c}$ ) through

$$\bar{c}(u, v) = c(1 - u, 1 - v) \quad .$$

As this study is only interested in the temporal dependence structure between consecutive returns, which is entirely captured by the copula, the specification of any parametric form for the marginal distributions is no longer necessary, allowing very flexible non- and semiparametric approaches.

Usually, most time series models consider general structures like  $X_i = g(X_{i-1}, X_{i-2}, \dots)$ , where the current variable is explained as a function of the past observations. For example, one possible form for the function  $g(\cdot)$  is the  $AR(1)$ -specification, e.g.  $X_i = \alpha X_{i-1} + \varepsilon_i$ , with error term  $\varepsilon_i$ . Following another attempt, this study solves the problem by specifying the function  $g(\cdot)$  as a copula. Applying the copula concept has the advantage that the temporal dependence structure of the stochastic process can be modeled in a more flexible way without restrictive assumptions such as linearity (see also Savu and Ng (2005)).

Instead of using different random variables  $X$  and  $Y$ , let  $X_i$  denote the  $i^{th}$  observation of the time series at time  $t_i$ , and  $X_{i-1}$  its lagged value, both with continuous marginal distribution functions  $F_i(x_i) = P(X_i \leq x_i)$ , respectively  $F_{i-1}(x_{i-1}) = P(X_{i-1} \leq x_{i-1})$ , where the joint distribution  $F(x_i, x_{i-1})$  is expressed via a copula function  $C$  as

$$\begin{aligned} F(x_i, x_{i-1}) &= P(X_i \leq x_i, X_{i-1} \leq x_{i-1}) \\ &= C_\theta(F_i(x_i), F_{i-1}(x_{i-1})) \quad . \end{aligned}$$

The copula  $C$  can also be seen as the joint distribution function of the transformed random variables  $U = F_i(X_i)$  and  $V = F_{i-1}(X_{i-1})$  with realizations  $u = F_i(x_i)$  and  $v = F_{i-1}(x_{i-1})$ . The copula parameter  $\theta$  controls the direction and the degree of dependence between  $X_i$  and  $X_{i-1}$ .

In most studies, researchers are used to work with parametric families of copulas  $C_\theta(u, v)$ , i.e. copula functions depending on a possibly  $q$ -dimensional vector of parameters  $\theta \in \Theta \subset \mathbb{R}^q$  controlling the direction and the degree of dependence. A very often used example for a parametric copula family is the Gaussian Normal copula

$$C_\theta(u, v) = C_\rho(u, v) = \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v))$$

where  $\rho \in [-1, 1]$ ,  $\Phi_\rho(\cdot)$  denotes the cdf of a bivariate standard normal variate with correlation coefficient  $\rho$  as the parameter of the copula and  $\Phi^{-1}(\cdot)$  the inverse cdf of a standard normal variate.

Another class of copulas, called Archimedean, also finds a wide range of applications in practice. These copulas are very easy to construct, many parametric families belong to this class and all commonly encountered Archimedean copulas have simple closed form expressions. Archimedean copulas are constructed (a) by a generator function  $\varphi : [0, 1] \rightarrow [0, \infty]$ , which is a continuous, strictly decreasing and convex function, such that  $\varphi(1) = 0$  and  $\varphi(0) = \infty$ , and (b) its inverse function  $\varphi^{-1}$ . They can be expressed in the form

$$C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v)) \quad .$$

In order to reveal the temporal structure of the high-frequency returns, a very flexible generalized Farlie-Gumbel-Morgenstern (FGM) Copula allowing multiple tail dependence with parameter  $\boldsymbol{\theta} = (a_1, a_2, b_1, b_2)$

$$\begin{aligned} C_{\boldsymbol{\theta}}(u, v) &= uv\{1 + (1-u)(1-v) \cdot \\ &\quad [a_1v(1-u) + a_2(1-u)(1-v) + b_1uv + b_2u(1-v)]\} \\ &= uv + (uv - u^2v - uv^2 + u^2v^2) \\ &\quad [(1-u)\{a_1v + a_2(1-v)\} + u\{b_1v + b_2(1-v)\}] \end{aligned}$$

is applied (see Nelsen (1990)). Deriving the conditional copula

$$\begin{aligned} C_{U|V=v}(u, v) &= \frac{\partial C_{\boldsymbol{\theta}}(u, v)}{\partial v} \\ &= u + (u - u^2 - 2uv + 2u^2v) \\ &\quad \cdot [(1-u)\{a_1v + a_2(1-v)\} + u\{b_1v + b_2(1-v)\}] \\ &\quad + (uv - u^2v - uv^2 + u^2v^2) \\ &\quad \cdot [(1-u)\{a_1 - a_2\} + u\{b_1 - b_2\}] \quad , \end{aligned}$$

the copula density is obtained by

$$\begin{aligned} c_{\boldsymbol{\theta}}(u, v) &= \frac{\partial^2 C_{\boldsymbol{\theta}}(u, v)}{\partial u \partial v} \\ &= 1 + (1 - 2u)(1 - 2v) \\ &\quad \cdot [(1-u)\{a_1v + a_2(1-v)\} + u\{b_1v + b_2(1-v)\}] \\ &\quad + (u - 2uv)(1-u)[(1-v)\{b_2 + a_2\} + v\{b_1 - a_1\}] \\ &\quad + (v - 2uv)(1-v)[(1-u)\{a_1 - a_2\} + u\{b_1 - b_2\}] \\ &\quad + uv(1-u)(1-v)[b_1 - b_2 - a_1 + a_2] \quad . \end{aligned}$$

The flexibility of this copula to model various association patterns (with different parameter settings) is shown in Figure 1 to 4.

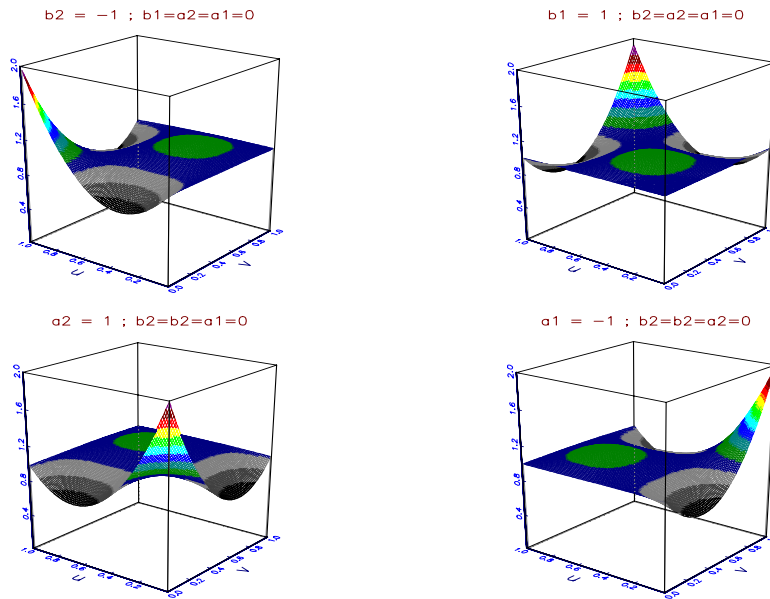


Figure 1: Examples of the FGM-Copula with different parameter settings

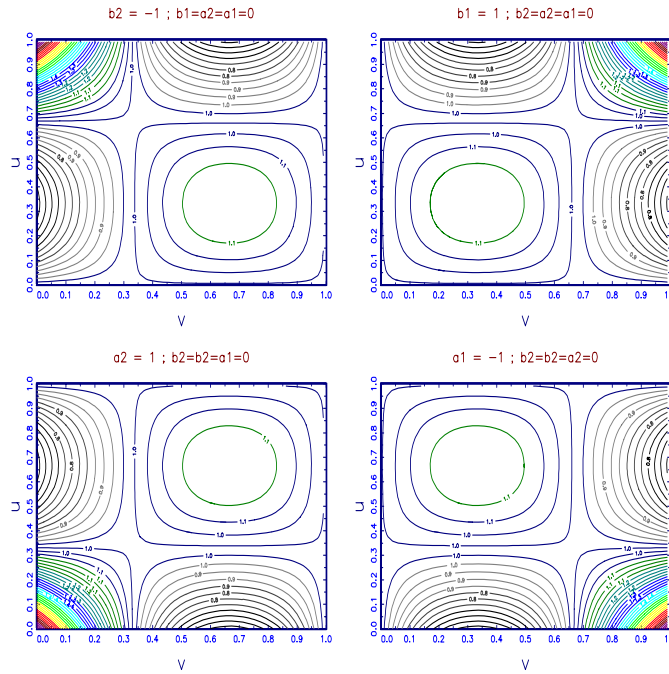


Figure 2: Contourplot of the FGM-Copula with different parameter settings



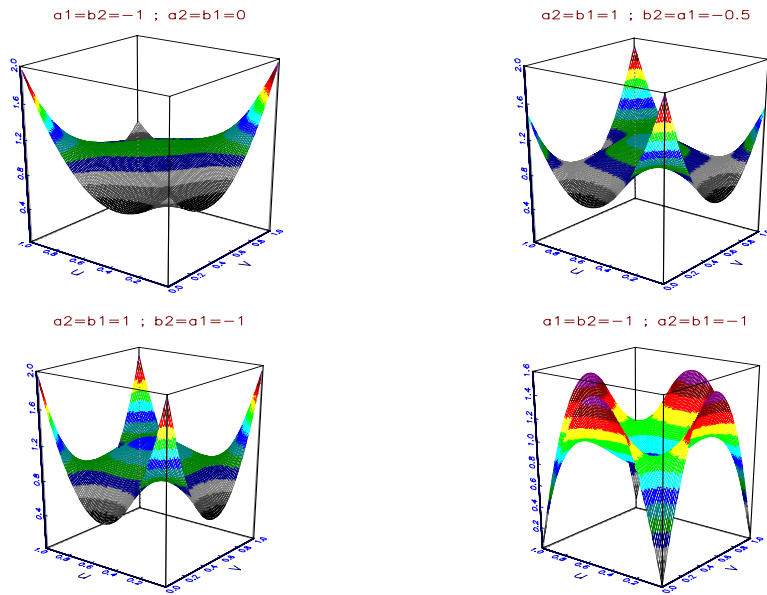


Figure 3: Examples of the FGM-Copula with different parameter settings

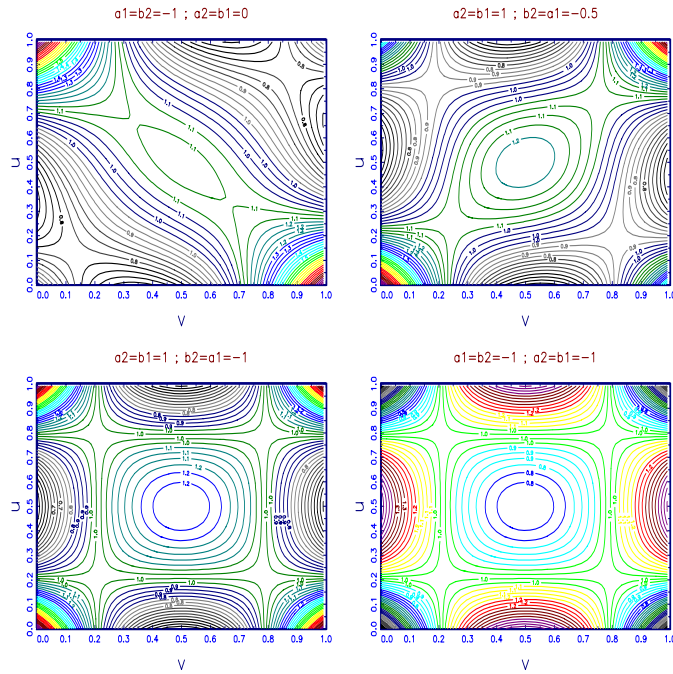


Figure 4: Contourplot of the FGM-Copula with different parameter settings

### 3 Estimation

As above mentioned, we are looking for copulas capable to model adequately the temporal dependence of high frequency time series data. The primary interest lies in the dependence function itself, no particular parametric form for the marginals is specified, avoiding misspecification and overfitting of the model. Since recent studies have shown that temporal aggregation and sampling frequency have an essential impact on the resulting stochastic process (see Lee, Gleason, and Mathur (1999), Cai, Hudson, and Keasey (2003) and Ait-Sahlia and Mykland (2003)), one must take these effects into account. Hence, the estimation is not only performed on the original data observed at the 1 second interval, but also on additional 120 different thinned return processes with increasing observation intervals 5 sec, 10 sec, 15 sec, ..., 600 sec .

Let  $P_i$  be the price of an asset a time  $t_i$ , observed at a certain sampling frequency, then  $\{R\}_{i=1}^N$  with  $R_i = \frac{R_i - R_{i-1}}{R_{i-1}}$  represents the return process. Drawing a random sample  $\{r_i, r_{i-1}\}_{i=2}^N$  of size  $N$  from the bivariate return vector  $(R_i, R_{i-1})$ , both non- and semiparametric copula estimations are performed in two stages. First of all, the marginal distributions are estimated nonparametrically using the empirical distribution

$$\hat{F}(r) = \frac{1}{N+1} \sum_{k=1}^N \mathbf{1}(R_k \leq r) \quad .$$

In the second step, the nonparametric copula for exploring the data can be estimated by means of any ordinary product kernel (Fermanian and Scaillet (2003)). However, since the copula density is bounded within the unit-square, one have to take the boundary bias into account that evolves when using common fixed symmetric kernel functions. To resolve this problem, one can use the mirroring technique supposed by Gijbels and Mielniczuk (1990), or, in order to save computation time, apply non-fixed Beta kernels for density functions as proposed by Chen (1999). Because the variance of the standard uniform distribution is  $\frac{1}{12}$ , the ‘‘asymptotic optimal’’ bandwidth of the kernel estimator  $K(\cdot)$  according to Scott’s rule is obtained by

$$h = \left( \sqrt[d+4]{12^3 N} \right)^{-1}$$

with  $d = 2$  (see Silverman (1986) and Scott (1992)). Let

$$K(p, q, u) = \frac{u^{p-1} (1-u)^{q-1}}{B(p, q)} = u^{p-1} (1-u)^{q-1} \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)}$$

be the density of the random variable  $U$  with  $Beta(p, q)$ -Distribution, the marginal density of the copula can be estimated via

$$\hat{f}(u) = \frac{1}{n} \sum_{i=1}^n K(u, h, U_i)$$

with

$$K(u, h, r) = \begin{cases} K\left(\varphi(u), \left(\frac{1-u}{h}\right), r\right) & \text{if } u \in [0, 2h) \\ K\left(\left(\frac{u}{h}\right), \left(\frac{1-u}{h}\right), r\right) & \text{if } u \in [2h, 1-2h] \\ K\left(\left(\frac{u}{h}\right), \varphi(1-u), r\right) & \text{if } u \in (1-2h, 1] \end{cases}$$

and

$$\varphi(u) = 2h^2 + 2.5 - \sqrt{4h^4 + 6h^2 + 2.25 - u^2} - \frac{u}{h}$$

(see Chen (1999)). Hence, the copula density can be obtained by

$$\hat{c}(u, v) = \frac{1}{n} \sum_{i=1}^n (K(u, h, U_i) \cdot K(v, h, V_i))$$

(see also Härdle, Müller, Sperlich, and Werwatz (2003) and Scott (1992)).

In order to estimate the semiparametric copula for quantifying the direction and the degree of the dependence, the canonical maximum likelihood method is adopted (Cherubini, Luciano, and Vecchiato (2004)). Since the copula can be written in the form

$$F(r_i, r_{i-1}; \boldsymbol{\theta}) = C\left(\hat{F}(r_i), \hat{F}(r_{i-1}); \boldsymbol{\theta}\right) \quad ,$$

the density of an observation  $(r_i, r_{i-1})$  is

$$f(r_i, r_{i-1}; \boldsymbol{\theta}) = c\left(\hat{F}(r_i), \hat{F}(r_{i-1}); \boldsymbol{\theta}\right) \quad .$$

Thus, the copula parameter vector  $\boldsymbol{\theta}$  can be estimated by maximizing the log-likelihood function

$$\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta}} \sum_{j=1}^N \ln c\left(\hat{F}(r_j), \hat{F}(r_{j-1}); \boldsymbol{\theta}\right) \quad ,$$

yielding the maximum likelihood estimator  $\hat{\boldsymbol{\theta}}$ , which is consistent and asymptotically normally distributed for time series data, as shown in Chen and Fan (2006b) (see also Chen and Fan (2006a) and Genest, Ghoudi, and Rivest (1995)).

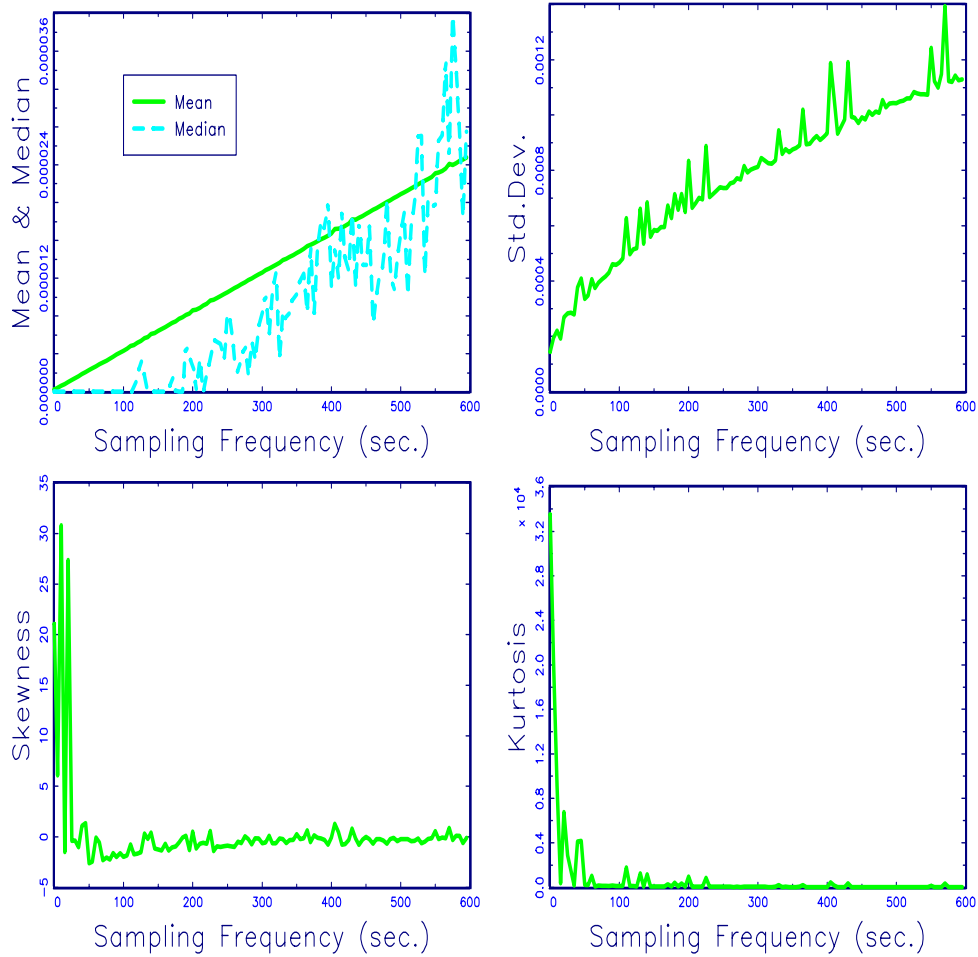


Figure 5: Description of the returns at different sampling frequencies

## 4 Empirical Results

The high frequency  $D3047$  data of the DAX performance index is extracted from the EUREX database. The sample includes 2612833 observations (at the 1 second sampling frequency) from  $2^{nd}$  January until  $28^{th}$  April 2006, observed for 90 trading days over 18 weeks. The daily trading phase starts at 9 a.m. and ends at 5.45 p.m. Descriptive statistics of the price returns at different sampling frequencies are given in Figure 5. The upper panels exhibit that the mean, the median and the standard deviation of the returns are slightly rising (albeit still close to zero) when decreasing the sampling frequency. In contrast, the lower panels show that skewness and the kurtosis

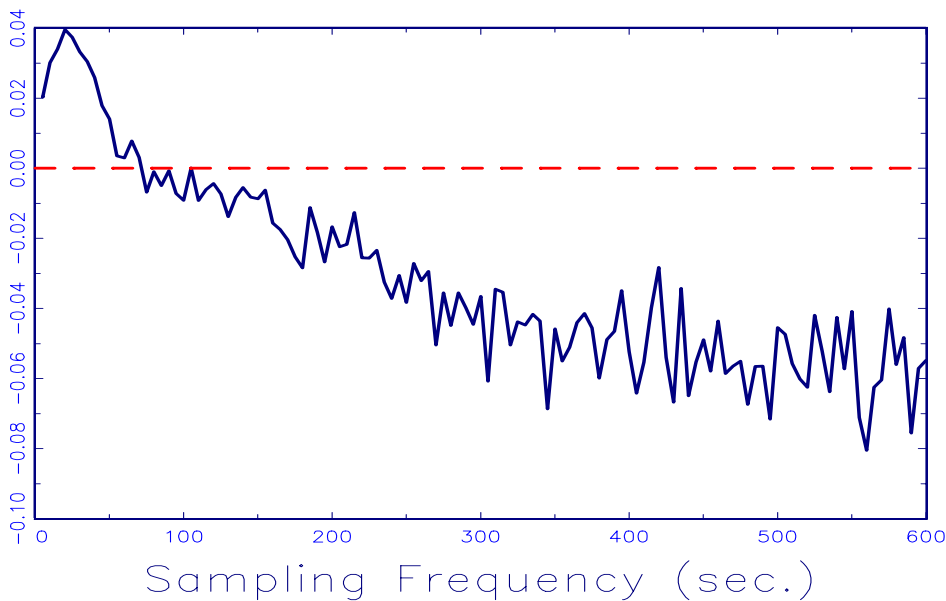


Figure 6: First order autocorrelation of returns at different sampling frequencies

are the smaller, the higher the aggregation level.

The first-order autocorrelation of the price process with respect to the sampling frequency is shown in Figure 6. Here, one can see that the correlation coefficient is positive in the high sampling frequencies and slowly decreases when increasing the observation interval. As expected, the entire dependence structure is not accurately captured by this measure when looking at the four nonparametrically estimated copula densities in Figures 7 to 14 (the CML-procedure of the Aptech software GAUSS 5.0 was used for the estimation of the copula models.).

Intriguingly, when analyzing the returns at the 1 second interval (see Figure 7 and 8), a nearly symmetric copula with one peak in the center and four orthogonal bumps is visible, similar to the well-known compass rose. According to the literature, this structure has several “rays” radiating from the origin with the thickest streams pointing towards the four major directions “north”, “east”, “south” and “west” of the compass. This pattern was first documented by Huang and Stoll (1994), later reinvestigated by Crack and Ledoit (1996). As shown by Krämer and Runde (1997) and Szpiro (1998), this phenomenon is mainly caused by the discreteness of price changes in financial markets, resulting in finite number of possible (often clustered) returns forming the rose. As long as the jumps take discrete ticks, this phenomenon

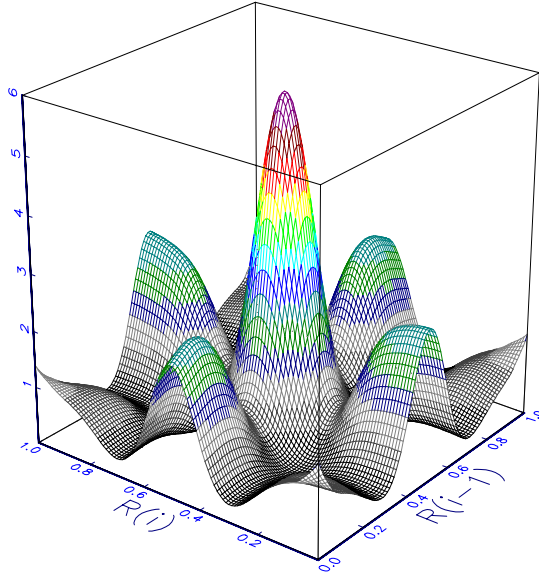


Figure 7: The Bivariate Copula Density of the Return Process with a sampling frequency of 1 sec .

also hold for portfolios and indices, due to their rounding “errors”.

Adopting the allegories of the literature, Figure 7 seem to resemble a “copula rose” with one “blossom” in the middle. The four major unfolding “petals” signalize that (a) the price changes are almost small during the short interval and that (b) those zero-returns (because the median is close to zero) have no predictive power concerning the non-zero-returns in the next period (Crack and Ledoit (1996)). Contrary, the minor “petals” in the four corners of the unit-square reveal that in a few cases extreme returns are somehow associated with *both* positive *and* negative extreme returns, which again implies unpredictability of stock prices.

When increasing the observation interval to 30 seconds (see Figure 9 and 10), one can see that the “blossom” and the four major “petals” immediately wither, whereas the density in the four corners rises, indicating multiple tail dependence: extreme large or extreme small price movements are now more likely than “moderate” ones. Moreover, the “petals” in the “south-west” and the “north-east” corner are the largest ones, which means that negative returns tend to be followed (but not definitely) by negative ones, and positive by positive ones. This pattern shows an overall positive dependence of consecutive returns (in each 30 seconds) and, thus, dissents the stylized fact

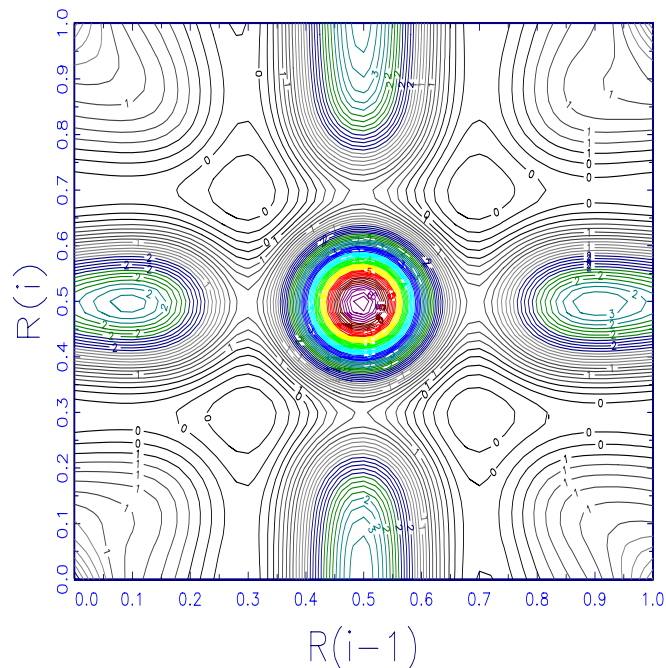


Figure 8: The Contourplot of the Copula Density for the Return Process with a sampling frequency of 1 sec .

in the literature that first-order autocorrelation are always negative. Taking a closer look at the middle of the copula, one can see that the dependence structure within the interquartile-square is negative. This result shows similarities to the so-called overreaction phenomena that has been widely studied in behavioral finance and financial psychology (see, for example, Bikhchandani, Hirshleifer, and Welch (1992) and Caginalp, Porter, and L. (2000)), but not at the high-frequency level.

Interestingly, all main relationships are switched, when increasing the observation interval once again (see Figure 12 to 14). The density in the “north-west” and the “south-east” corner are now higher, inducing a negative dependence. This, in fact, confirms the findings in the literature.

Comparing all nonparametric copulas, one can see that the “rose” is more visible the shorter the observation interval. This result is in line with the analysis of Wang, Hudson, and Keasey (2000), who found out “*that the compass rose becomes more apparent as the frequency of observations increases*”. But in contrast to phase portrait, where the pattern is sometimes not discernible due to the huge number of rays, the copula is always able to reveal the underlying dependence structure of the data. In general, the high den-

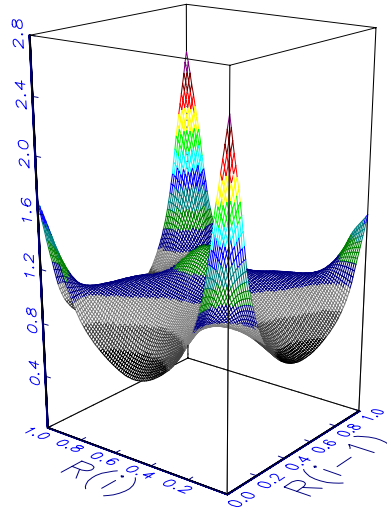


Figure 9: The Bivariate Copula Density of the Return Process with a sampling frequency of 30 sec .

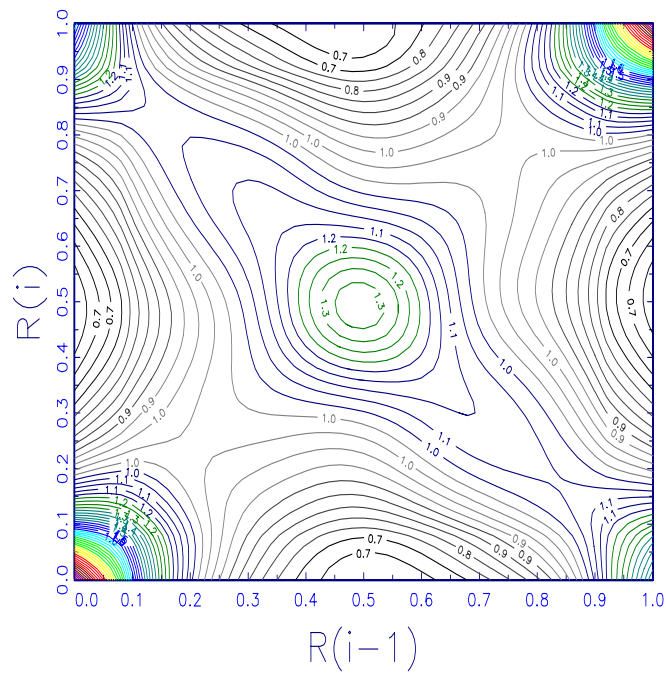


Figure 10: The Contourplot of the Copula Density for the Return Process with a sampling frequency of 30 sec .



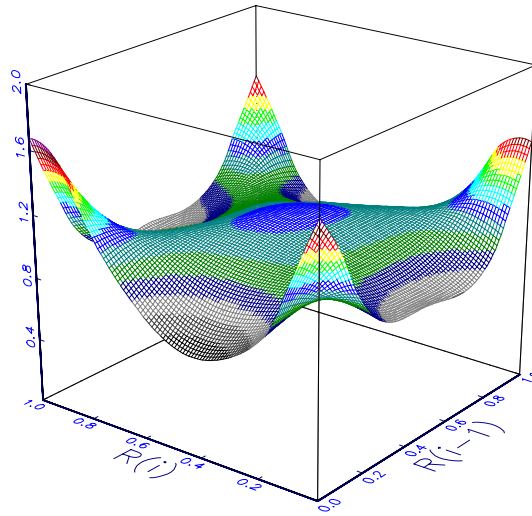


Figure 11: The Bivariate Copula Density of the Return Process with a sampling frequency of 300 sec .

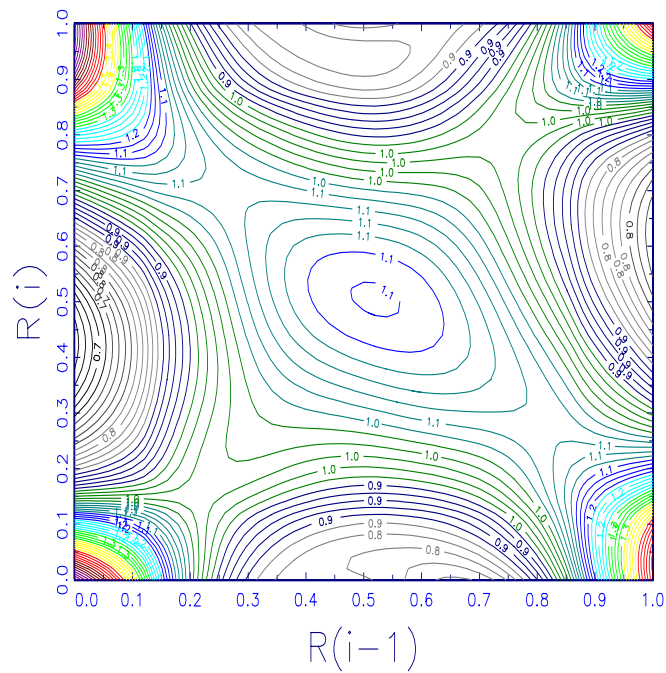


Figure 12: The Contourplot of the Copula Density for the Return Process with a sampling frequency of 300 sec .

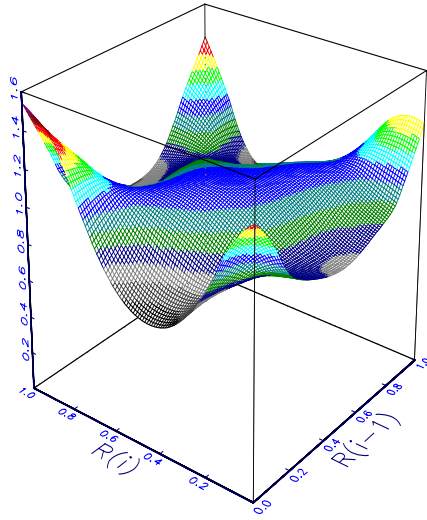


Figure 13: The Bivariate Copula Density of the Return Process with a sampling frequency of 600 sec .

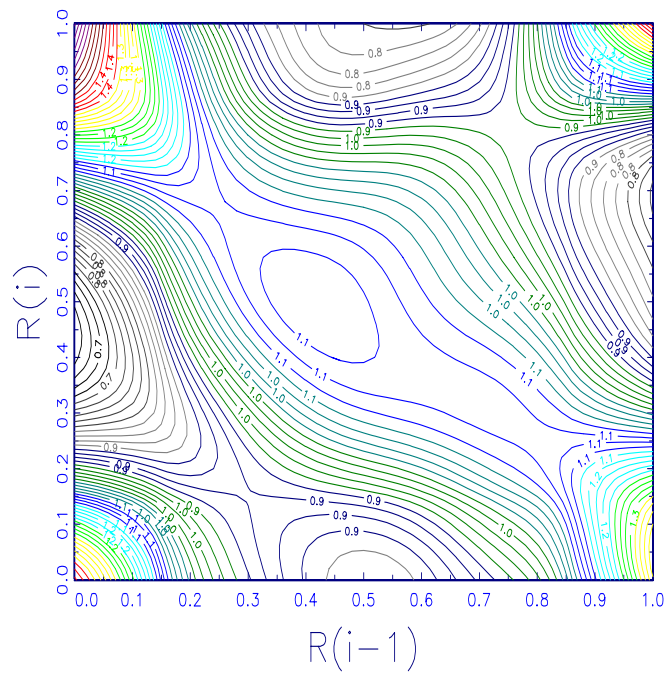


Figure 14: The Contourplot of the Copula Density for the Return Process with a sampling frequency of 600 sec .

sity in the center of the unit square has a neutral position and diminishes with decreasing sampling frequency, but never vanishes. Furthermore, when sampling the data at the higher frequencies, the positive relationship within the unit-square is stronger than the negative one (albeit slightly), whereas the entire dependence structure of the less frequently observed data is almost negative, although there are positive patterns. In other words, the overall “aggregated” dependence structure within the unit square is either negative or positive, but there are always “partial” dependence structures as well, signalling a opposite association. These two antagonistic effects cannot be discovered with common linear regression or correlation coefficients: extreme price differences are associated with outliers again, whereas price movements “with small jumps” seem to be uncorrelated. This relationship can be seen as an outcome of “informational overshooting” that causes booms and crashes (i.e. extreme price jumps see Zeira (1999)).

In contrast to phase portraits, these different “side-effects” can now be quantified with the generalized semiparametric FGM-Copula. In contrast to Wang and Wang (2002), the FGM copula measures the degree and the direction of the dependence, which is more interesting for economists than the quality of pattern. The estimation results are displayed in Figure 15. While the parameters  $\hat{a}_1$  and  $\hat{b}_2$  (responsible for a negative dependence) are often close to  $-1$ , the parameters  $\hat{a}_2$  and  $\hat{b}_1$  (responsible for a positive dependence) are  $+1$  at the beginning, but then decline and, thus, “allow” an overall negative relationship when moving to the higher aggregation levels. These results show that the first order dependence of high frequency returns is not necessarily negative as assumed in the literature, but strongly depends on the sampling frequency. Figure 16 reveals that all estimated parameters are significant at the 1%-level.

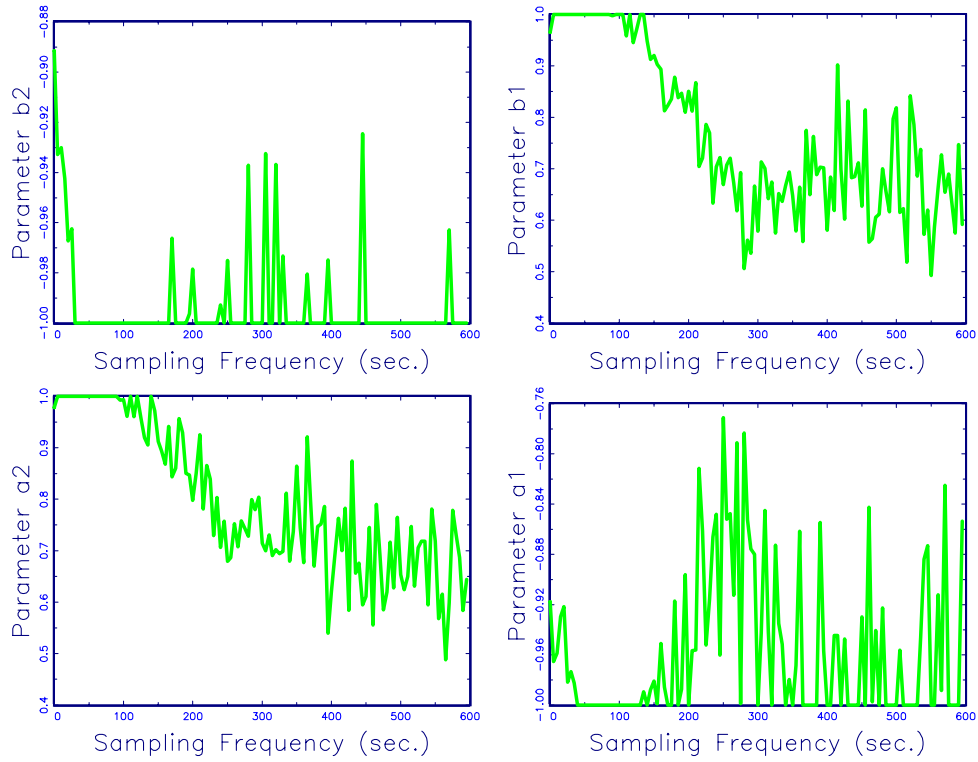


Figure 15: Estimated Parameters of the Generalized FGM Copula.

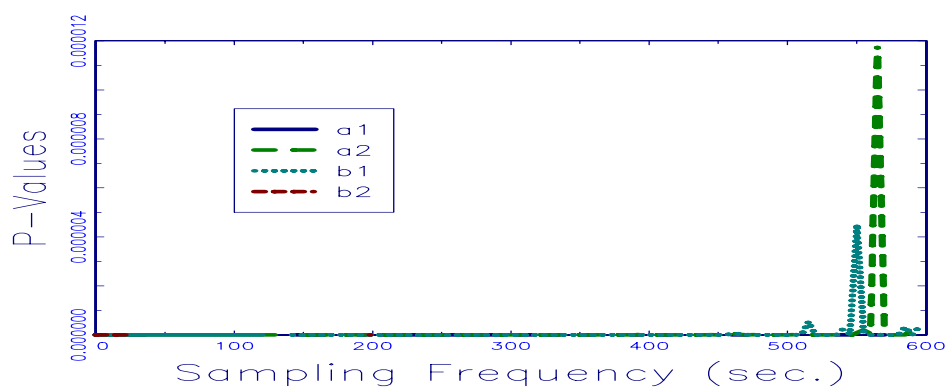


Figure 16: P-values of the estimated FGM-Parameters

## 5 Conclusion

This paper proposes a copula-based modeling framework for analyzing the return process of high-frequent EUREX tick data. The advantage of the copula model is the feature to separate the temporal dependence from the marginal distribution of the times series, enabling more flexibility in modeling. The idea is based on splitting the bivariate distribution of consecutive returns  $R_i$  and  $R_{i-1}$  into two components: (a) the marginal distribution of  $\{R_i\}_{i=1}^N$  without any parametric assumptions, and (b) the serial dependence of the return process captured in the copula.

Nonparametric copulas, which are used to explore the general dependence structure between consecutive returns, resemble a symmetric pattern with several partial sub-structures, forming a “rose”. In order to quantify the multiple tail dependence, a very flexible generalized Farlie-Gumbel-Morgenstern (*FGM*) copula with a 4-dimensional parameter vector was estimated. The results show that the first order dependence of high frequency returns is not always necessarily negative, but strongly depends on the sampling frequency that also influences the rose-structure of the copula. When sampling the data at a higher frequency, the positive relationship within the unit-square is stronger than the negative one, whereas the dependence of the less frequently observed data is almost negative, although there are positive patterns.

## References

- AÏT-SAHLIA, Y., AND P. MYKLAND (2003): “The Effects of Random and Discrete Sampling when Estimating Continuous-Time Diffusions,” *Econometrica*, 71(2), 483–549.
- ANDERSEN, T. G., T. BOLLERSLEV, AND N. MEDDAHI (2005): “Correcting the errors: Volatility Forecast Evaluation Using High-Frequency Data and Realized Volatilities,” *Econometrica*, 73(1), 279–296.
- BIKHCHANDANI, S., D. HIRSHLEIFER, AND I. WELCH (1992): “A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades,” *Journal of Political Economy*, 100(5), 992–1026.
- BOLLERSLEV, T., AND I. DOMOWITZ (1993): “Trading Patterns and Prices in the Interbank Foreign Exchange Market,” *Journal of Finance*, 48(4), 1421–1443.
- CAGINALP, G., D. PORTER, AND S. V. L. (2000): “Overreactions, Momentum, Liquidity, and Price Bubbles in Laboratory and Field Asset Markets,” *Journal of Psychology Financial Markets*, 1(1), 24–48.
- CAI, C., R. HUDSON, AND K. KEASEY (2003): “Trading frequency and the compass rose,” *Applied Economics Letters*, 10(8), 511–517.
- CHEN, S. X. (1999): “Beta Kernel Estimators for Density Functions,” *Computational Statistics & Data Analysis*, 31(2), 131–145.
- CHEN, X., AND Y. FAN (2006a): “Estimation and Model Selection of Semiparametric Copula-Based Multivariate Dynamic Models Under Copula Misspecification,” *Journal of Econometrics*, 135(1-2), 125–154.
- (2006b): “Estimation of Copula-Based Semiparametric Time Series Models,” *Journal of Econometrics*, *forthcoming*, 130(2), 307–335.
- CHERUBINI, U., E. LUCIANO, AND W. VECCHIATO (2004): *Copula Methods in Finance*. John Wiley & Sons, New York et al.
- CRACK, T., AND O. LEDOIT (1996): “Robust Structure without Predictability: The "Compass Rose" Pattern of the Stock Market,” *Journal of Finance*, 51, 751–762.
- DACOROGNA, M. M., R. GENÇAY, U. MÜLLER, O. R. B., AND P. O. V. (2001): *An Introduction to High-Frequency Finance*. Academic Press, New York.

- FERMANIAN, J.-D., AND O. SCAILLET (2003): “Nonparametric Estimation of Copulas for Time Series,” *Journal of Risk*, 5(4), 25–54.
- (2005): “Some Statistical Pitfalls in Copula Modeling for Financial Applications,” in *Capital Formation, Governance and Banking*, ed. by E. Klein, pp. 59–74. Nova Science Publ.
- GENEST, C., K. GHOUDI, AND L.-P. RIVEST (1995): “A Semiparametric Estimation Procedure of Dependence Parameters in Multivariate Families of Distributions,” *Biometrika*, 82(3), 543–552.
- GIJBELS, I., AND J. MIELNICZUK (1990): “Estimating the Density of a Copula Function,” *Communications in Statistics: Theory and Methods*, 19(2), 445–464.
- GOODHART, C.A.E. AND FIGLIUOLI, L. (1991): “Every minutes counts in financial markets,” *Journal of International Money and Finance*, 10(1), 23–52.
- HÄRDLE, W., M. MÜLLER, S. SPERLICH, AND A. WERWATZ (2003): *Nonparametric and Semiparametric Models*. Springer, Heidelberg.
- HUANG, R. D., AND H. R. STOLL (1994): “Market Microstructure and Stock Return Predictions,” *Review of Financial Studies*, 7, 179–213.
- JOE, H. (1997): *Multivariate Models and Dependence Concepts*, Monographs on Statistics and Applied Probability, 73. Chapman & Hall, London.
- KRÄMER, W., AND R. RUNDE (1997): “Chaos and the compass rose,” *Economics Letters*, 54(2), 113–118.
- LEE, C. I., K. C. GLEASON, AND I. MATHUR (1999): “A Comprehensive Examination of the Compass Rose Pattern in Future Markets,” *The Journal of Future Markets*, 19(5), 541–564.
- MARI, D. D., AND S. KOTZ (2001): *Correlation and Dependence*. Imperial College Press, London.
- NELSEN, R. B. (1990): *An Introduction to Copulas*, Lectures Notes in Statistics, 139. Springer, New York et al.
- PATTON, A. J. (2006): “Estimation of Multivariate Models for Time Series of Possibly Different Lengths,” *Journal of Applied Econometrics*, 21(2), 147–173.

- ROLL, R. (1984): “A Simple Explicit Measure of the Effective Bid-Ask Spread in an Efficient Market,” *Journal of Finance*, 39(4), 1127–1139.
- RUSSELL, J. R., AND R. F. ENGLE (2005): “A Discrete-State Continuous-Time Model of Financial Transactions Prices and Times: The Autoregressive Conditional Multinomial-Autoregressive Conditional Duration Model,” *Journal of Business & Economic Statistics*, 23(2), 166–180.
- SAVU, C., AND W. L. NG (2005): “The SCoD Model - Analyzing Durations with A Semiparametrics Copula Approach,” *International Review of Finance*, 5(1-2), 55–74.
- SCOTT, D. W. (1992): *Multivariate density estimation : Theory, Practice, and Visualization*, Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, New York et al.
- SILVERMAN, B. W. (1986): *Density Estimation for Statistics and Data Analysis*. Chapman & Hall, London, New York.
- SKLAR, A. (1959): “Fonctions de répartition à n dimensions et leurs marges,” *Publications de l’Institut de Statistique de Paris*, 8, 229–231.
- SZPIRO, G. G. (1998): “Tick size, the compass rose and market nanostructure,” *Journal of Banking and Finance*, 22(12), 1559–1569.
- WANG, E., R. HUDSON, AND K. KEASEY (2000): “Tick size and the compass rose: further insights,” *Economics Letters*, 68(2), 119–125.
- WANG, H., AND C. WANG (2002): “Visibility of the compass rose in financial asset returns: A quantitative study,” *Journal of Banking and Finance*, 26(6), 1099–1111.
- WÖHRMANN, P. (2005): “Predictability in Discrete Stock Prices,” Working Paper 245, Institute for Empirical Research in Economics, University of Zürich.
- ZEIRA, J. (1999): “Informational overshooting, booms, and crashes,” *Journal of Monetary Economics*, 43(1), 237–257.



## SFB 649 Discussion Paper Series 2006

For a complete list of Discussion Papers published by the SFB 649, please visit <http://sfb649.wiwi.hu-berlin.de>.

- 001 "Calibration Risk for Exotic Options" by Kai Detlefsen and Wolfgang K. Härdle, January 2006.
- 002 "Calibration Design of Implied Volatility Surfaces" by Kai Detlefsen and Wolfgang K. Härdle, January 2006.
- 003 "On the Appropriateness of Inappropriate VaR Models" by Wolfgang Härdle, Zdeněk Hlávka and Gerhard Stahl, January 2006.
- 004 "Regional Labor Markets, Network Externalities and Migration: The Case of German Reunification" by Harald Uhlig, January/February 2006.
- 005 "British Interest Rate Convergence between the US and Europe: A Recursive Cointegration Analysis" by Enzo Weber, January 2006.
- 006 "A Combined Approach for Segment-Specific Analysis of Market Basket Data" by Yasemin Boztuğ and Thomas Reutterer, January 2006.
- 007 "Robust utility maximization in a stochastic factor model" by Daniel Hernández-Hernández and Alexander Schied, January 2006.
- 008 "Economic Growth of Agglomerations and Geographic Concentration of Industries - Evidence for Germany" by Kurt Geppert, Martin Gornig and Axel Werwatz, January 2006.
- 009 "Institutions, Bargaining Power and Labor Shares" by Benjamin Bental and Dominique Demougine, January 2006.
- 010 "Common Functional Principal Components" by Michal Benko, Wolfgang Härdle and Alois Kneip, January 2006.
- 011 "VAR Modeling for Dynamic Semiparametric Factors of Volatility Strings" by Ralf Brüggemann, Wolfgang Härdle, Julius Mungo and Carsten Trenkler, February 2006.
- 012 "Bootstrapping Systems Cointegration Tests with a Prior Adjustment for Deterministic Terms" by Carsten Trenkler, February 2006.
- 013 "Penalties and Optimality in Financial Contracts: Taking Stock" by Michel A. Robe, Eva-Maria Steiger and Pierre-Armand Michel, February 2006.
- 014 "Core Labour Standards and FDI: Friends or Foes? The Case of Child Labour" by Sebastian Braun, February 2006.
- 015 "Graphical Data Representation in Bankruptcy Analysis" by Wolfgang Härdle, Rouslan Moro and Dorothea Schäfer, February 2006.
- 016 "Fiscal Policy Effects in the European Union" by Andreas Thams, February 2006.
- 017 "Estimation with the Nested Logit Model: Specifications and Software Particularities" by Nadja Silberhorn, Yasemin Boztuğ and Lutz Hildebrandt, March 2006.
- 018 "The Bologna Process: How student mobility affects multi-cultural skills and educational quality" by Lydia Mechtenberg and Roland Strausz, March 2006.
- 019 "Cheap Talk in the Classroom" by Lydia Mechtenberg, March 2006.
- 020 "Time Dependent Relative Risk Aversion" by Enzo Giacomini, Michael Handel and Wolfgang Härdle, March 2006.
- 021 "Finite Sample Properties of Impulse Response Intervals in SVECMs with Long-Run Identifying Restrictions" by Ralf Brüggemann, March 2006.
- 022 "Barrier Option Hedging under Constraints: A Viscosity Approach" by Imen Bentahar and Bruno Bouchard, March 2006.

**SFB 649, Spandauer Straße 1, D-10178 Berlin**  
**<http://sfb649.wiwi.hu-berlin.de>**

This research was supported by the Deutsche  
Forschungsgemeinschaft through the SFB 649 "Economic Risk".



- 023 "How Far Are We From The Slippery Slope? The Laffer Curve Revisited" by Mathias Trabandt and Harald Uhlig, April 2006.
- 024 "e-Learning Statistics – A Selective Review" by Wolfgang Härdle, Sigbert Klinke and Uwe Ziegenhagen, April 2006.
- 025 "Macroeconomic Regime Switches and Speculative Attacks" by Bartosz Maćkowiak, April 2006.
- 026 "External Shocks, U.S. Monetary Policy and Macroeconomic Fluctuations in Emerging Markets" by Bartosz Maćkowiak, April 2006.
- 027 "Institutional Competition, Political Process and Holdup" by Bruno Deffains and Dominique Demougin, April 2006.
- 028 "Technological Choice under Organizational Diseconomies of Scale" by Dominique Demougin and Anja Schöttner, April 2006.
- 029 "Tail Conditional Expectation for vector-valued Risks" by Imen Bentahar, April 2006.
- 030 "Approximate Solutions to Dynamic Models – Linear Methods" by Harald Uhlig, April 2006.
- 031 "Exploratory Graphics of a Financial Dataset" by Antony Unwin, Martin Theus and Wolfgang Härdle, April 2006.
- 032 "When did the 2001 recession *really* start?" by Jörg Polzehl, Vladimir Spokoiny and Cătălin Stărică, April 2006.
- 033 "Varying coefficient GARCH versus local constant volatility modeling. Comparison of the predictive power" by Jörg Polzehl and Vladimir Spokoiny, April 2006.
- 034 "Spectral calibration of exponential Lévy Models [1]" by Denis Belomestny and Markus Reiß, April 2006.
- 035 "Spectral calibration of exponential Lévy Models [2]" by Denis Belomestny and Markus Reiß, April 2006.
- 036 "Spatial aggregation of local likelihood estimates with applications to classification" by Denis Belomestny and Vladimir Spokoiny, April 2006.
- 037 "A jump-diffusion Libor model and its robust calibration" by Denis Belomestny and John Schoenmakers, April 2006.
- 038 "Adaptive Simulation Algorithms for Pricing American and Bermudan Options by Local Analysis of Financial Market" by Denis Belomestny and Grigori N. Milstein, April 2006.
- 039 "Macroeconomic Integration in Asia Pacific: Common Stochastic Trends and Business Cycle Coherence" by Enzo Weber, May 2006.
- 040 "In Search of Non-Gaussian Components of a High-Dimensional Distribution" by Gilles Blanchard, Motoaki Kawanabe, Masashi Sugiyama, Vladimir Spokoiny and Klaus-Robert Müller, May 2006.
- 041 "Forward and reverse representations for Markov chains" by Grigori N. Milstein, John G. M. Schoenmakers and Vladimir Spokoiny, May 2006.
- 042 "Discussion of 'The Source of Historical Economic Fluctuations: An Analysis using Long-Run Restrictions' by Neville Francis and Valerie A. Ramey" by Harald Uhlig, May 2006.
- 043 "An Iteration Procedure for Solving Integral Equations Related to Optimal Stopping Problems" by Denis Belomestny and Pavel V. Gapeev, May 2006.
- 044 "East Germany's Wage Gap: A non-parametric decomposition based on establishment characteristics" by Bernd Görzig, Martin Gornig and Axel Werwatz, May 2006.
- 045 "Firm Specific Wage Spread in Germany - Decomposition of regional differences in inter firm wage dispersion" by Bernd Görzig, Martin Gornig and Axel Werwatz, May 2006.

**SFB 649, Spandauer Straße 1, D-10178 Berlin**  
<http://sfb649.wiwi.hu-berlin.de>

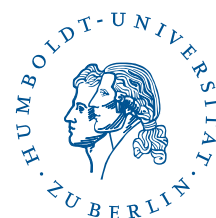
This research was supported by the Deutsche  
 Forschungsgemeinschaft through the SFB 649 "Economic Risk".



- 046 "Produktdiversifizierung: Haben die ostdeutschen Unternehmen den Anschluss an den Westen geschafft? – Eine vergleichende Analyse mit Mikrodaten der amtlichen Statistik" by Bernd Görzig, Martin Gornig and Axel Werwatz, May 2006.
- 047 "The Division of Ownership in New Ventures" by Dominique Demougin and Oliver Fabel, June 2006.
- 048 "The Anglo-German Industrial Productivity Paradox, 1895-1938: A Restatement and a Possible Resolution" by Albrecht Ritschl, May 2006.
- 049 "The Influence of Information Costs on the Integration of Financial Markets: Northern Europe, 1350-1560" by Oliver Volckart, May 2006.
- 050 "Robust Econometrics" by Pavel Čížek and Wolfgang Härdle, June 2006.
- 051 "Regression methods in pricing American and Bermudan options using consumption processes" by Denis Belomestny, Grigori N. Milstein and Vladimir Spokoiny, July 2006.
- 052 "Forecasting the Term Structure of Variance Swaps" by Kai Detlefsen and Wolfgang Härdle, July 2006.
- 053 "Governance: Who Controls Matters" by Bruno Deffains and Dominique Demougin, July 2006.
- 054 "On the Coexistence of Banks and Markets" by Hans Gersbach and Harald Uhlig, August 2006.
- 055 "Reassessing Intergenerational Mobility in Germany and the United States: The Impact of Differences in Lifecycle Earnings Patterns" by Thorsten Vogel, September 2006.
- 056 "The Euro and the Transatlantic Capital Market Leadership: A Recursive Cointegration Analysis" by Enzo Weber, September 2006.
- 057 "Discounted Optimal Stopping for Maxima in Diffusion Models with Finite Horizon" by Pavel V. Gapeev, September 2006.
- 058 "Perpetual Barrier Options in Jump-Diffusion Models" by Pavel V. Gapeev, September 2006.
- 059 "Discounted Optimal Stopping for Maxima of some Jump-Diffusion Processes" by Pavel V. Gapeev, September 2006.
- 060 "On Maximal Inequalities for some Jump Processes" by Pavel V. Gapeev, September 2006.
- 061 "A Control Approach to Robust Utility Maximization with Logarithmic Utility and Time-Consistent Penalties" by Daniel Hernández-Hernández and Alexander Schied, September 2006.
- 062 "On the Difficulty to Design Arabic E-learning System in Statistics" by Taleb Ahmad, Wolfgang Härdle and Julius Mungo, September 2006.
- 063 "Robust Optimization of Consumption with Random Endowment" by Wiebke Wittmüß, September 2006.
- 064 "Common and Uncommon Sources of Growth in Asia Pacific" by Enzo Weber, September 2006.
- 065 "Forecasting Euro-Area Variables with German Pre-EMU Data" by Ralf Brüggemann, Helmut Lütkepohl and Massimiliano Marcellino, September 2006.
- 066 "Pension Systems and the Allocation of Macroeconomic Risk" by Lans Bovenberg and Harald Uhlig, September 2006.
- 067 "Testing for the Cointegrating Rank of a VAR Process with Level Shift and Trend Break" by Carsten Trenkler, Pentti Saikkonen and Helmut Lütkepohl, September 2006.
- 068 "Integral Options in Models with Jumps" by Pavel V. Gapeev, September 2006.
- 069 "Constrained General Regression in Pseudo-Sobolev Spaces with Application to Option Pricing" by Zdeněk Hlávka and Michal Pešta, September 2006.

**SFB 649, Spandauer Straße 1, D-10178 Berlin**  
<http://sfb649.wiwi.hu-berlin.de>

This research was supported by the Deutsche  
 Forschungsgemeinschaft through the SFB 649 "Economic Risk".



- 070 "The Welfare Enhancing Effects of a Selfish Government in the Presence of Uninsurable, Idiosyncratic Risk" by R. Anton Braun and Harald Uhlig, September 2006.
- 071 "Color Harmonization in Car Manufacturing Process" by Anton Andriyashin, Michal Benko, Wolfgang Härdle, Roman Timofeev and Uwe Ziegenhagen, October 2006.
- 072 "Optimal Interest Rate Stabilization in a Basic Sticky-Price Model" by Matthias Paustian and Christian Stoltenberg, October 2006.
- 073 "Real Balance Effects, Timing and Equilibrium Determination" by Christian Stoltenberg, October 2006.
- 074 "Multiple Disorder Problems for Wiener and Compound Poisson Processes With Exponential Jumps" by Pavel V. Gapeev, October 2006.
- 075 "Inhomogeneous Dependency Modelling with Time Varying Copulae" by Enzo Giacomini, Wolfgang K. Härdle, Ekaterina Ignatieva and Vladimir Spokoiny, November 2006.
- 076 "Convenience Yields for CO<sub>2</sub> Emission Allowance Futures Contracts" by Szymon Borak, Wolfgang Härdle, Stefan Trück and Rafal Weron, November 2006.
- 077 "Estimation of Default Probabilities with Support Vector Machines" by Shiyi Chen, Wolfgang Härdle and Rouslan Moro, November 2006.
- 078 "GHICA - Risk Analysis with GH Distributions and Independent Components" by Ying Chen, Wolfgang Härdle and Vladimir Spokoiny, November 2006.
- 079 "Do Individuals Recognize Cascade Behavior of Others? - An Experimental Study -" by Tim Grebe, Julia Schmid and Andreas Stiehler, November 2006.
- 080 "The Uniqueness of Extremum Estimation" by Volker Krätschmer, December 2006.
- 081 "Compactness in Spaces of Inner Regular Measures and a General Portmanteau Lemma" by Volker Krätschmer, December 2006.
- 082 "Probleme der Validierung mit Strukturgleichungsmodellen" by Lutz Hildebrandt and Dirk Temme, December 2006.
- 083 "Formative Measurement Models in Covariance Structure Analysis: Specification and Identification" by Dirk Temme and Lutz Hildebrandt, December 2006.
- 084 "PLS Path Modeling – A Software Review" by Dirk Temme, Henning Kreis and Lutz Hildebrandt, December 2006.
- 085 "Relational Contracts and Inequity Aversion" by Jenny Kragl and Julia Schmid, December 2006.
- 086 "Overreaction and Multiple Tail Dependence at the High-frequency Level – The Copula Rose" by Wing Lon Ng, December 2006.

**SFB 649, Spandauer Straße 1, D-10178 Berlin**  
**<http://sfb649.wiwi.hu-berlin.de>**

This research was supported by the Deutsche  
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

