Dependence Structures for Multivariate High–Frequency Data in Finance

Wolfgang Breymann*, Alexandra Dias†, Paul Embrechts

Department of Mathematics
ETHZ
CH – 8092 Zürich
Switzerland
www.math.ethz.ch/finance

Abstract

Stylised facts for univariate high–frequency data in finance are well–known. They include scaling behaviour, volatility clustering, heavy tails, and seasonalities. The multivariate problem, however, has scarcely been addressed up to now.

In this paper, bivariate series of high–frequency FX spot data for major FX markets are investigated. First, as an indispensable prerequisite for further analysis, the problem of simultaneous deseasonalisation of high–frequency data is addressed. In the bulk of the paper we analyse in detail the dependence structure as a function of the time scale. Particular emphasis is put on the tail behaviour, which is investigated by means of copulas and spectral measures.

1 Introduction

Numerous papers have studied statistical properties of one–dimensional return data in finance. Results like leptokurtosis, stochastic volatility effects, occurrence of extremes, seasonalities, and scaling behavior are now referenced to as stylised facts of empirical finance. The work by Olsen & Associates has extended these facts across sampling frequencies reaching from minutes to months; see for instance Dacorogna et al. (2001). Similar results for more–dimensional return data are however scarce. In Embrechts et al. (2002) some of the basic techniques for the analysis of dependence beyond linear correlation were introduced through the notion of copula. In this paper, the latter techniques will be applied to a two–dimensional high–frequence (hourly) data set of FX returns. As such, the change in dependence as a function of the sampling frequency will be established. Also, for each separate frequency, ellipticality will be tested. Finally, several statistical techniques for the study of extremal clustering in higher dimensions will be applied. As a necessary prerequisite for this analysis, a method for deseasonalising bivariate returns for time horizons up to one day will be presented.

The outline of the paper is as follows. Section 2 presents the transformation from the raw high–frequency (tick–by–tick) data to properly deseasonalised data. In Section 3, several families of copulas will be fitted to deseasonalised two–dimensional FX data, and this at several frequencies (from hourly to daily). Goodness–of–fit tests, including tests for ellipticality, are presented in Section 4. In Section 5,

*Research supported by Credit Suisse Group, Swiss Re, and UBS AG through RiskLab, Switzerland.

†Support from Fundação para a Ciência e a Tecnologia - FCT/POCTI and Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal, is gratefully acknowledged.

the problem of clustering of extremes in two dimensional data is discussed. Finally, Section 6 gives an outlook for further research.

2 The data

We investigate a high-frequency bivariate time series consisting of USD/DEM and USD/JPY spot rates. Before they are to be used for dependency analysis, the following preliminary steps have been performed:

- collection and filtering,
- regularisation and transformation to logarithmic middle prices, and
- deseasonalisation.

They are described in turn.

2.1 Collection and filtering

The data set consists of tick–by–tick data originating mainly from Reuters and collected and filtered by Olsen Data. It consists of a large part but not all of the quotes emitted in the market because the market coverage of the data providers it not complete and depends on the region of the world. The high-frequency series are irregularly spaced; they start February 1986 and end June 30, 2001. Since we are interested in USD/DEM, which ends December 31, 1998, we discard later data also for other currencies. A single quote at time t consists of a bid price, $p_{\alpha,t}^{Bid}$, and an ask price, $p_{\alpha,t}^{Ask}$, $\alpha \in \{\text{USD/DEM}, \text{USD/JPY}\}$. For both series, middle prices are displayed in Figure 1. In a first step the data are cleaned by means of a special filter, described in Dacorogna et al. (2001), that takes peculiarities of the financial market into account. Among others it corrects for decimal errors caused by the transmission line and removes automatically generated fake quotes during inactive periods used by market participants to test the transmission channel. Since only a small fraction of quotes is removed by the filter the filtered time series is still irregularly spaced, and the number of data points is very high (about 10 million for the USD/DEM series).

2.2 Regularisation and transformation to logarithmic middle prices

To reduce the data we regularise the time series to a regular series with step size $\delta = 5$ minutes by linear interpolation. Since we are not interested in effects related to the bid–ask spread, we will work with logarithmic middle prices $\xi_{\alpha,t}$ defined as

$$\xi_{\alpha,t} = \frac{\log\left(p_{\alpha,t}^{Bid} \cdot p_{\alpha,t}^{Ask}\right)}{2}. \tag{1}$$

Returns with respect to a time horizon ΔT are then defined as the difference of logarithmic middle prices:

$$r_{\alpha,t}[\Delta T] = \xi_{\alpha,t} - \xi_{\alpha,t-\Delta T}. \tag{2}$$

Hourly USD/DEM returns are displayed in Figure 2. The advantage of taking the logarithm is that returns of the inverted rate (e.g. DEM/USD) are just the negative of the corresponding returns of the original rate USD/DEM, as one intuitively expects.

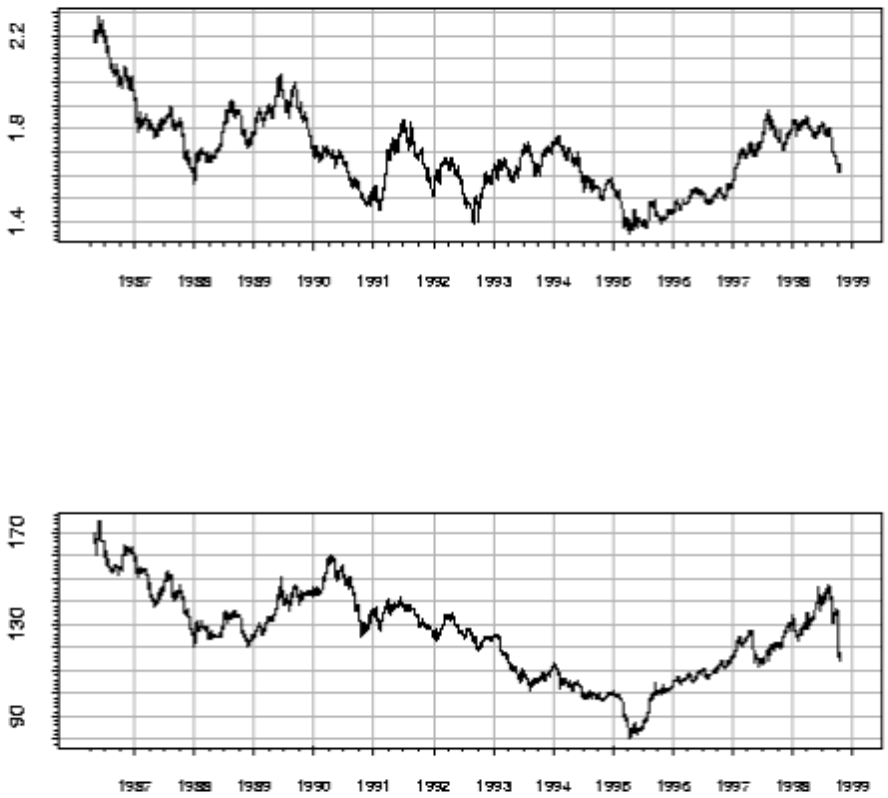

Figure 1: FX middle prices for USD/DEM (top) and USD/JPY (bottom).

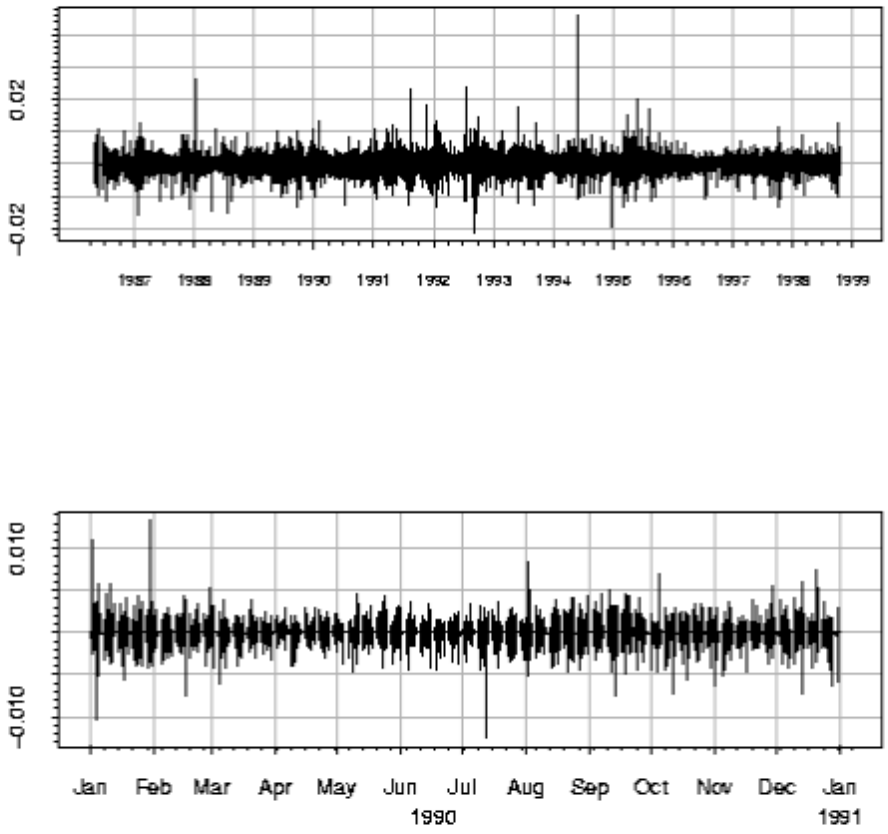

Figure 2: Hourly USD/DEM returns of the whole 11-year period (top) and for 1 year (bottom). Notice the weekly seasonalities.

2.3 Deseasonalisation of financial data

Practically all financial time series exhibit seasonalities. The most striking one is the absence of any activity during weekends, which causes a weekly seasonality in the autocorrelation function of lagged absolute returns. With high–frequency data the problem of seasonality becomes much more important and more difficult to handle because the entire form of the weekly activity pattern (mean computed conditional on the time of the week) has to be taken into account. In the autocorrelation function of hourly absolute returns, the weekly and daily periods can be distinguished, as shown in Figure 3. It also shows that the seasonal patterns of the two FX rates, though overall similar, exhibit differences in the details. Already for univariate time series deseasonalisation is not an easy task and there is not yet unanimity on how to solve it. In the multivariate case, deseasonalisation is still largely an open problem.

Before presenting our method of bivariate deseasonalisation we will summarise the two main approaches of deseasonalising univariate high–frequency time series. More detailed discussions can for instance be found in Dacorogna et al. (2001).

2.3.1 Approaches for deseasonalising univariate financial time series

There are two main approaches of deseasonalising univariate high-frequency time series: time transformation and volatility weighting by periodically varying weights. Eventually both approaches are based on a weekly conditional mean of volatility similar to the one shown in Figure 4. One of the earliest deseasonalisation methods has been developed by Olsen & Associates; see Dacorogna et al. (1993). It consists of a transformation from physical time to an activity–related time scale, the so–called ϑ-time scale, which is proportional to a measure of the market activity a_t:

$$\vartheta_{t_2} - \vartheta_{t_1} \propto \int_{t_1}^{t_2} a_t \, \mathrm{d}t. \tag{3}$$

The market activity is based on the volatility $\overline{v}_\tau$ estimated by a mean of absolute hourly returns, conditional on the time in the week:

$$\overline{v}_\tau[1\,\text{hour}] = E\left[|r_t[1\,\text{hour}]| \,|\, t \bmod (1\,\text{week}) = \tau\right]. \tag{4}$$

Volatility and market activity turn out to be empirically related by

$$a_t \sim (\overline{v}_t)^{1/h} \tag{5}$$

where h is the scaling exponent of absolute returns,

$$E[|r_t[\Delta T]|] \sim (\Delta T)^h. \tag{6}$$

While $h = 0.5$ for Brownian motion, values empirically observed are found to be around 0.58 for major currencies. In addition the activity a_t is decomposed into component activities $a_{i,t}$ of the mean regional markets (American, East-Asian, European), which is important for properly taking public holidays into account. For more details on (3)–(6), see Dacorogna et al. (2001).

Time transformation is appealing because it is intimately related to the concept of random time change in the theory of stochastic processes. In addition it conserves the aggregation property of returns,

$$r_t[\Delta T_1 + \Delta T_2] = r_{t-\Delta T_2}[\Delta T_1] + r_t[\Delta T_2], \tag{7}$$

which follows immediately from (2). However, since any coordinate of a multivariate series has its own activity–related time scale, it is currently unclear how to construct a time transformation for multivariate series without losing synchronicity of the different coordinates. In addition, there is not yet unanimity about the interpretation of (6); see for instance Barndorff-Nielsen and Prause (2001) for a critical

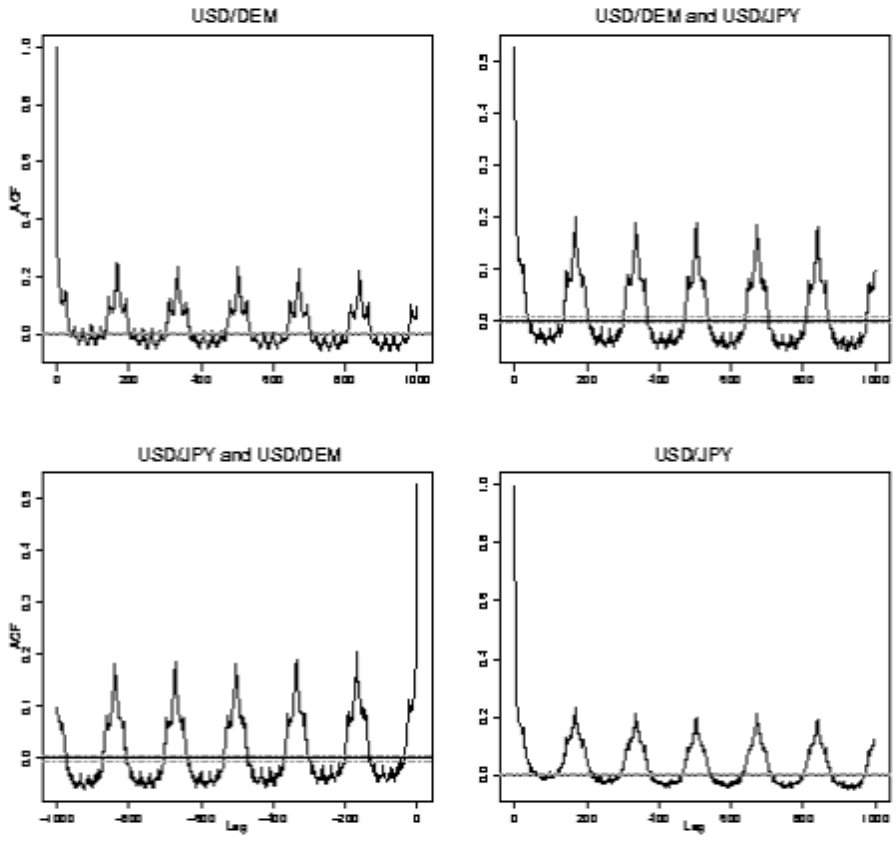

Figure 3: Autocorrelation function of bivariate absolute hourly returns (USD/DEM vs. USD/JPY). Notice the weekly and the daily seasonality. The latter is much more pronounced for USD/DEM rates than for USD/JPY rates.

discussion. Therefore (5) is at least questionable. For all these reasons we will not rely on (3)–(6) for deseasonalisation in this paper.

Alternative approaches based on volatility weighting have been proposed by Bollerslev and Ghysels (1996), Andersen and Bollerslev (1997), Taylor and Xu (1997), Andersen and Bollerslev (1998), Martens et al. (2002), Beltratti and Morana (1999). In this framework the return is written as $r_t[\delta] = \tilde{\sigma}_t[\delta]\, s_t[\delta]\, \epsilon_t$ for the generic intraday return at day t (Andersen and Bollerslev (1997), Andersen and Bollerslev (1998)). Here, $r_t[\delta]$ represents the intraday 5 minutes return, $\tilde{\sigma}_t[\delta]$ the deseasonalised volatility, ϵ_t denotes an i.i.d. (independent, identically distributed) mean zero, unit variance error term, and $s_t[\delta]$ essentially represents the seasonal pattern. Thus, the deseasonalised return reads:

$$x_t[\delta] = \tilde{\sigma}_t[\delta]\, \epsilon_t = \frac{r_t[\delta]}{s_t[\delta]}. \tag{8}$$

Different methods have been proposed to model the seasonal volatility $s_t[\delta]$ (see Andersen and Bollerslev (1997), Martens et al. (2002), Beltratti and Morana (1999), Andersen and Bollerslev (1998)).

2.3.2 Bivariate deseasonalisation

For the above-mentionned reasons no attempt towards a bivariate time transformation is made in this paper. Instead, we will use a weighting method similar to (8), which can be extended to the multivariate case in a straightforward way. Our definition of volatility is based on quadratic variations, which has the advantage that the theoretically expected scaling exponent 1 is also observed empirically, at least in the case of freely floating currencies of major markets. For any instrument α the deseasonalised return $x_{\alpha,t}[\Delta T]$, which will be analysed in detail in the following sections, is computed by

$$x_{\alpha,t}[\Delta T] = \begin{cases} \dfrac{r_{\alpha,t}[\Delta T]}{v_{\alpha,t}[\Delta T]} & \text{, if } v_{\alpha,t}[\Delta T] > 0, \\ \\ 0 & \text{, otherwise,} \end{cases} \tag{9}$$

where $\Delta T = n\delta$ is the time horizon of the return and $v_{\alpha,t}[\Delta T]$ is the expected volatility of instrument α. Note that $v_{\alpha,t}[\Delta T] = 0$ implies $r_{\alpha,t}[\Delta T] = 0$. In contrast to (8) the time horizon can be chosen freely, at least up to one day. This is important in the present context because we are interested in the variation of the dependence structure as a function of time.

In the remainder of this section we will explain how $v_{\alpha,t}[\Delta T]$ is computed. This is similar to the method described in Breymann (2000). To alleviate notation, the index α will be dropped whenever this is possible without ambiguity. In fact, quantities with α dropped can be interpreted as two element vectors, and the operations on them are understood to apply component-wise.

For reasons which will become clear below we prefer to use the integrated squared volatility

$$V_t^2 = \sum_{t' \le t} \left(v_{t'}[\delta]\right)^2 \tag{10}$$

where $\delta = 5$ minutes is the binning of the volatility histogram (see (16) below). The volatility $v_t[\Delta T]$ for an arbitrary horizon ΔT is recovered by

$$\left(v_t[\Delta T]\right)^2 = \left(\Delta V_t[\Delta T]\right)^2 = \sum_{i=0}^{n-1} \left(v_{t-i\delta}[\delta]\right)^2 \tag{11}$$

with $n = \Delta T/\delta$. Here the definition

$$\Delta V_t[\Delta T] := \sqrt{V_t^2 - V_{t-\Delta T}^2} \tag{12}$$

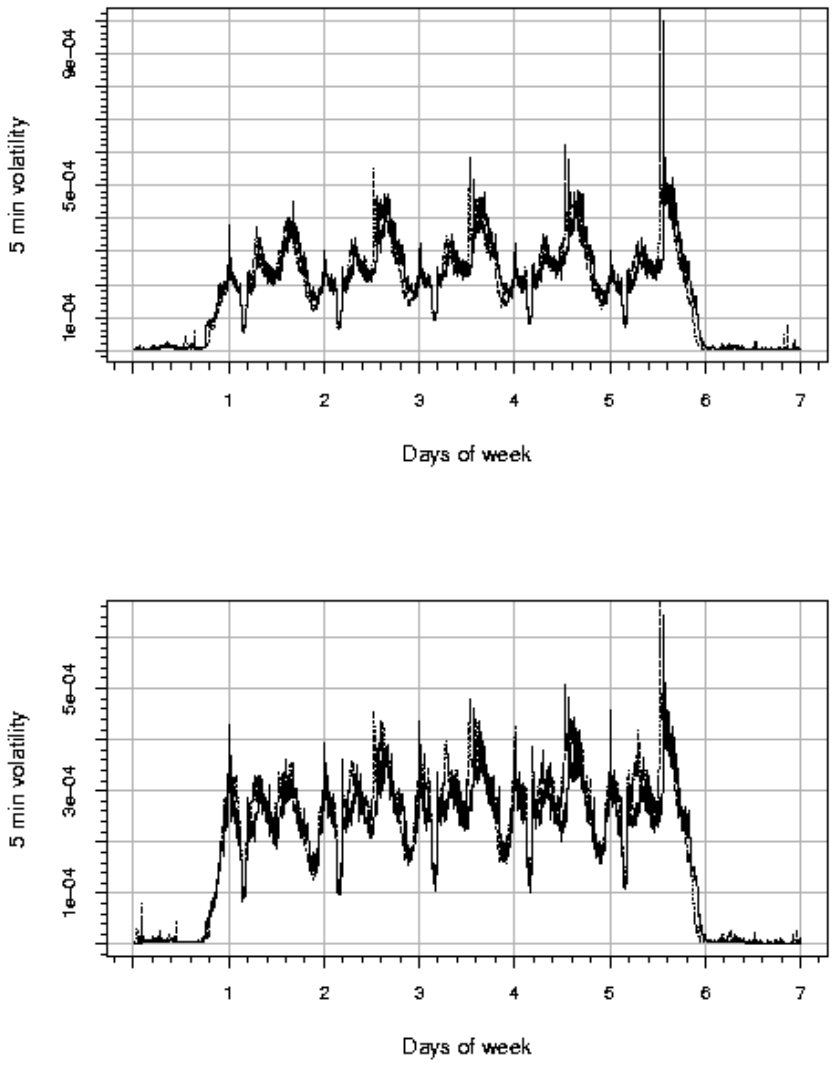

Figure 4: Mean weekly volatility pattern, averaged from 4/86 till 10/98. Top: USD/DEM, bottom: USD/JPY. Full lines: winter time, dotted lines: Summer time. Notice that during the periods where the European or the American Market are active, both patterns are slightly shifted (dotted line left of the full line) while no shift associated with Daylight Saving Time (DST) is present when only the Japanese market is active (most remarkably seen during the Japanese lunch break). The fact that at the beginning of the weekly activity period the dotted line is at the *right* of the full line stems from the Australian market, where a DST shift is present during the months October–April.

has been used. Thus, the deseasonalisation can be written as

$$X_t[\Delta T] = \frac{\xi_t - \xi_{t-\Delta T}}{\sqrt{V_t - V_{t-\Delta T}}}. \tag{13}$$

As already noted above, deseasonalised returns loose the aggregation property (7). Instead, from (7), (9), and (10) follows the relation

$$\begin{aligned} x_t[\Delta T] &= \frac{x_{t-\Delta T_2}[\Delta T_1]\, v_{t-\Delta T_2}[\Delta T_1] + x_t[\Delta T_2]\, v_t[\Delta T_2]}{v_t[\Delta T_1 + \Delta T_2]} \\ &= \frac{x_{t-\Delta T_2}[\Delta T_1]\ \Delta V_{t-\Delta T_2}[\Delta T_1] + x_t[\Delta T_2]\ \Delta V_t[\Delta T_2]}{\Delta V_t[\Delta T]} \end{aligned} \tag{14}$$

with $\Delta T = \Delta T_1 + \Delta T_2$. Thus, time aggregation of deseasonalised returns requires the knowledge of V_t^2.

2.3.3 The seasonal volatility pattern

Modelling requirements:

- The possibility to model arbitrary patterns with abrupt changes in market volatility as they for instance occur during the Japanese lunch break;
- The modelling of the slow dynamic behaviour of the activity pattern to take into account slow temporal changes in the habits of the market participants, institutional changes, etc.;
- Keeping track of Daylight Saving Time (DST) to take into account a partial 1 hour displacement of the average volatility patterns for DST and non-DST periods, one with respect to the other (Figure 4);
- The modelling of the geographical decomposition of market activity to take into account perturbations of the regular weekly activity pattern due to local holidays affecting only parts of the market.

The first requirement is met by using weekly volatility histograms with a 5 minutes time step, and the second and third requirements are met by treating the volatility pattern of different DST periods separately. To meet the last requirement we decompose the squared volatility $v_t^2[\delta]$ into the product

$$v_t^2[\delta] = a_t \left(v_\tau^{(d)}[\delta]\right)^2 \tag{15}$$

where a_t is a relative market activity factor and $v_\tau^{(d)}[\delta]$ is the volatility averaged over DST period d conditional to the time in the week, $\tau = t \bmod (1\text{ week})$:

$$\left(v_\tau^{(d)}[\delta]\right)^2 = \frac{1}{N_d} \sum_{i=1}^{N_d} (r_{t_i+\tau}[\delta])^2. \tag{16}$$

Here, $\tau \in \{0\,\mathrm{h}, \delta, 2\delta, \ldots, 168\,\mathrm{h} - \delta\}$ is the time in the week, t_i is the start of week i (always a Sunday, 00:00:00 UTC (GMT)), and N_d is the number of weeks in DST period d. In (15) the appropriate DST period d is selected by the condition $t \in d$. The market activity factor a_t is 1 for normal days, when the expected volatility is $v_\tau^{(d)}[\delta]$. If the expected volatility is lower, a_t assumes values strictly less than one. This happens for public holidays. Since public holidays are different in different regional markets, a_t is written as a sum of regional market activity factors, $a_t = \sum_i a_{i,t}$. We work with an American, an East-Asian, and a European market component. When a public holiday occurs in one of these components, the corresponding $a_{i,t}$ is set to zero.

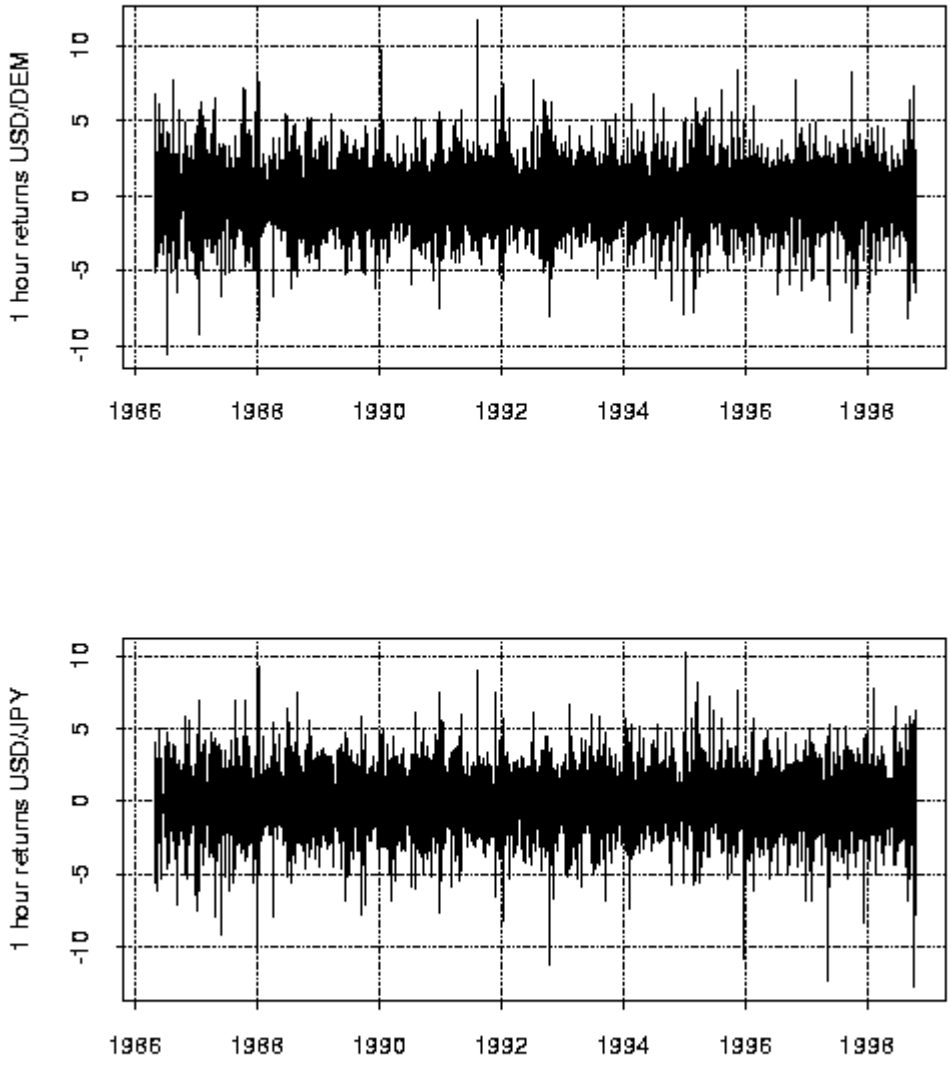

Figure 5: *The one hour deseasonalised returns.*

2.3.4 The Weekends

Weekends need special treatment because even though the mean activity is low it occasionally happens that a big price jump occurs between Friday evening closing and Monday morning opening. These price jumps cannot be properly accounted for by the mean weekend activity and, therefore, deteriorate the deseasonalisation results for short time horizons up to about one day. One way of dealing with this effect would be to drop the return between the last Friday evening price and the first Monday morning price. Beside the fact that this is difficult in an OTC market without well-defined opening and closing times it has the additional drawback that it is in conflict with the modified aggregation property (14). Therefore we decided to adopt the following procedure. An effective weekend volatility $v^{(w)}[\delta]$ is computed for every weekend w by

$$v^{(w)}[\delta] = \left| r_{t_w^{(end)}}[\Delta T_w] \right| \sqrt{\frac{\delta}{\Delta T_w}}, \tag{17}$$

where $\Delta T_w = t_w^{(end)} - t_w^{(start)}$ is the weekend length. Start and end of weekends are fixed to $t_i^{(start)} =$ Friday, 21:00:00 UTC and $t_i^{(end)} =$ Sunday, 21:00:00 UTC, respectively. The volatility $v_t[\delta]$ in (15) is set to this value if $t \in (t_w^{(start)}, t_w^{(end)}]$. In this way, weekend peaks are exactly compensated even if the weekend is dropped simultaneously from the series of logarithmic prices ξ_t and from the integrated volatility pattern V_t^2; notice, in particular, that (13) remains valid. If a jump occurs during the weekend, a corresponding jump will be present in V_t^2 such that deseasonalisation is assured.

Here, a note of caution is in order. First, our studies showed that the weekend weighting destroys the dependence structure of the returns for time horizons longer than about four days, while no significant effect could be seen in the margins. By construction of the weekend weighting the destruction of dependence only affects the returns reaching over a weekend. For horizons of the order of one hour this corresponds to less than 1% of the data. This ratio increases to 20% for daily data, and to 80% for a four day horizon. For time horizons from one week onwards all data are affected.

Second, deseasonalisation is no longer meaningful for time horizons being multiples of one week. This follows from the fact that by virtue of (10) and (16), $V_t^2 - V_{t-1\text{ week}}^2$ is constant within a given DST period. Even though this argument contains the assumption of no special weekend weighting, one can argue that for such time horizons, weekends are part of the regular dynamics and should not be treated separately. Thus, to treat time horizons between one day and one week adequately, a smooth transition from the deseasonalisation procedure for short horizons presented in this paper to no deseasonalisation for long horizons is needed. We emphasize that the need for such a transition only arises in the case of multivariate data.

2.4 Univariate properties of deseasonalised returns

The series of 1 hour deseasonalised returns is displayed in Figure 5. Even though volatility clustering is present no seasonalities can be distinguished. Figure 6 displays the autocorrelation function of absolute returns for a time horizon of 1 hour, which clearly shows the well-known long-range correlation of the volatility. This figure confirms also the efficient removal of seasonalities, as can be seen by comparison with Figure 3. To show the tail behavior, QQ–plots of the univariate series are displayed in figures 7 and 8 for time horizons ranging from 1 hour to 1 day. The well–known transition from pronounced heavy tails at short horizons to more thin–tailed distributions at long time horizons is clearly seen.

3 Dependence structure modelling

Given the bivariate deseasonalised returns from (9), we want to analyse the dependence structure at each of the different frequencies: one hour, two hours, four hours, eight hours, twelve hours and one day. In each case, we will fit parametric families of copulas using a two stage semi–parametric procedure (see

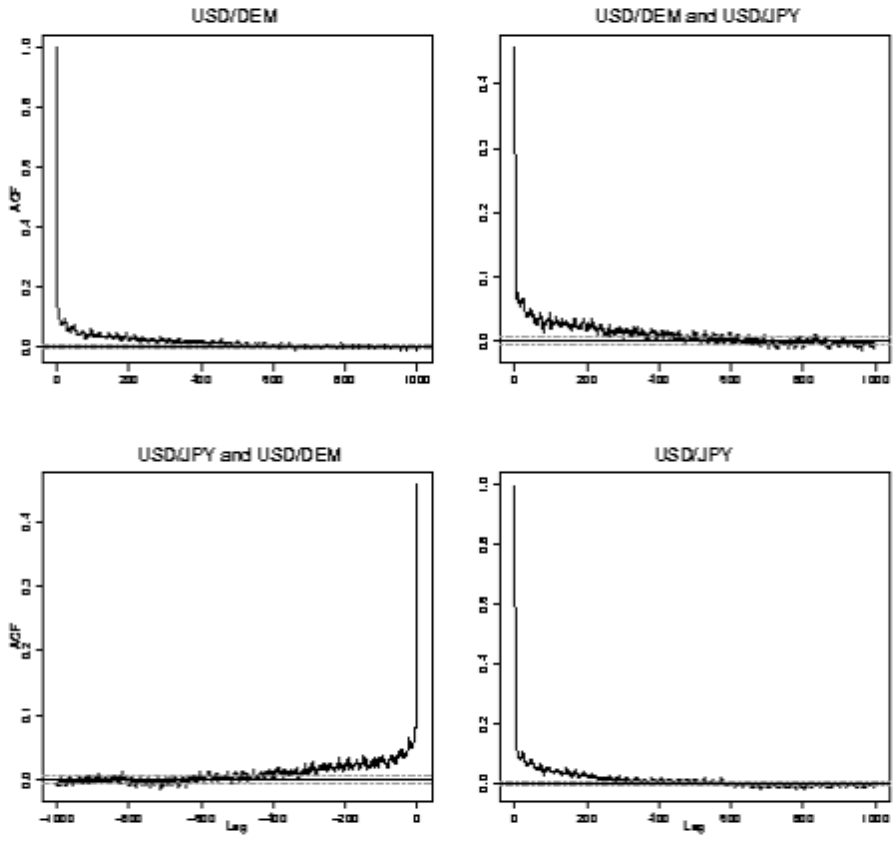

Figure 6: *Autocorrelation function of absolute deseasonalised 1 hour returns.*

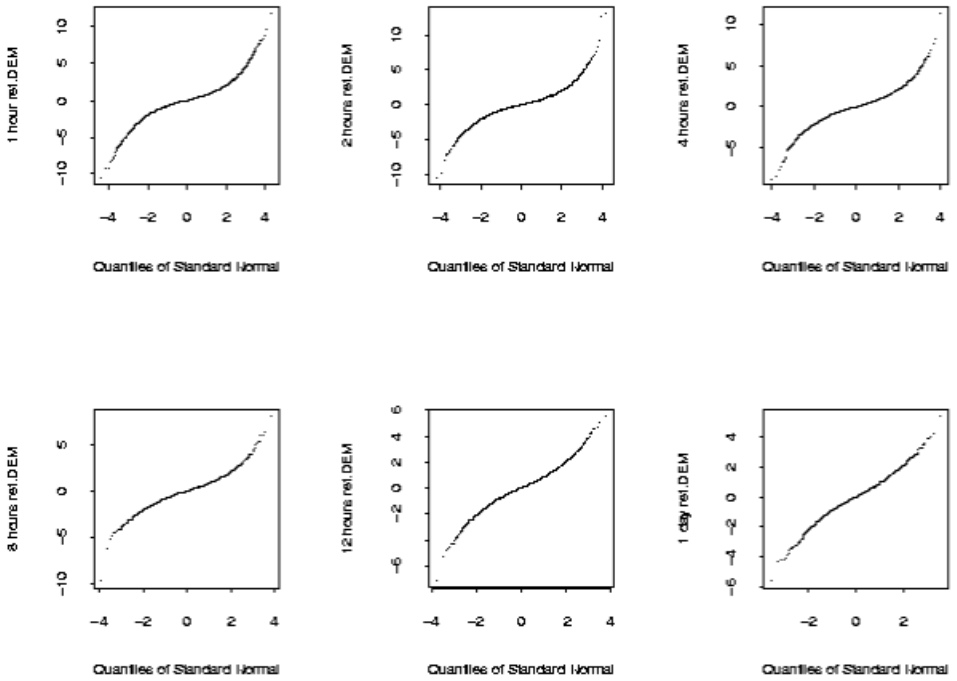

Figure 7: *The QQ–plots of the normal against the empirical quantiles for several frequencies of DEM deseasonalised returns.*

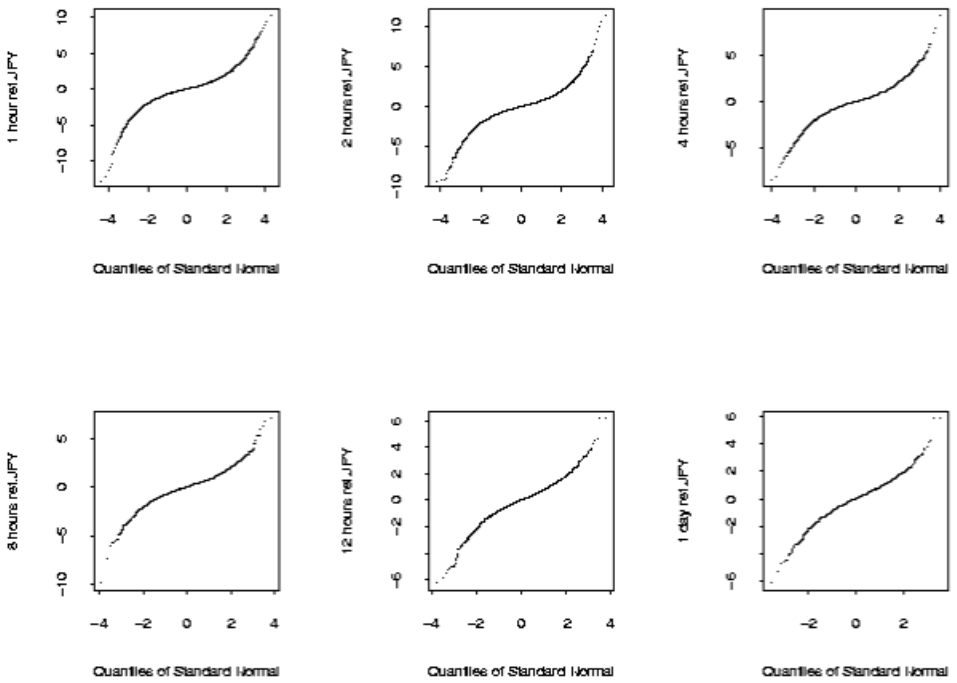

Figure 8: *The QQ–plots of the normal against the empirical quantiles for several frequencies of JPY deseasonalised returns.*

Genest et al. (1995)). This procedure consists of transforming the marginal observations into uniformly distributed vectors using the empirical distribution functions, in a first step. Then, the copula parameters are estimated by the maximisation of a pseudo log-likelihood function. Each one of these fitting phases is explained more in detail below.

For each considered frequency we have two vectors of observations, deseasonalised returns of FX rates quoted against the US Dollar, one for USD/DEM and another for USD/JPY. The scatter plots are shown in Figure 9. We denote the random variable for the deseasonalised USD/DEM returns by X_1 and for the deseasonalised USD/JPY returns by X_2. If $\{x_{i1}, x_{i2}, \ldots, x_{in}\}$ are the n observed univariate deseasonalised returns of the FX rate X_i $(i = 1, 2)$ for a given frequency, then

$$\hat{F}_{in}(x_{ij}) = \frac{1}{n+1} \sum_{k=1}^{n} \mathbf{1}_{\{x_{ik} \leq x_{ij}\}},$$

are pseudo–observations approximately uniformly distributed in $[0, 1]$. Figure 10 displays the scatter plots of the bivariate pseudo–observations

$$(\hat{F}_{1n}(\mathbf{x}_1), \hat{F}_{2n}(\mathbf{x}_2)),$$

and Figure 11 shows the same returns but plotted with standard normal margins. The number of points plotted in each panel varies a lot and that makes them harder to compare. But even so, we can still see in Figure 11 that there is an evolution from a diamond to an elliptic shape as the time frequency decreases. In Figure 10, for the one, two and four hour returns, we plotted sub–samples of the pseudo–observations otherwise the scatter plots of the full samples would be just three useless black squares.

Computing the bivariate pseudo–observations $(\hat{F}_{1n}(\mathbf{x}_1), \hat{F}_{2n}(\mathbf{x}_2))$ for each time frequency is the first step for the copula fitting. On these transformed data sets we can estimate the parameters for several copula families. Here we consider the Gaussian, the t, the Frank, the Gumbel and the Clayton copulas. The specifications of these distribution families can be found in Embrechts et al. (2002) and Nelsen (1999). See also Joe (1997).

In Table 1 the parameter estimates, the corresponding standard errors, and the Akaike information criterion values (AIC) are fitted for each of the models. For every time frequency the models are ordered by the AIC value, according to which lower values indicate a better fit. The first observation is that, for the five models considered, the t copula model has the best performance according to the AIC criterion. In Figure 12 we plotted, for each model and for each frequency, the AIC of the t copula minus the AIC of the model and divided this difference by the number of observations (in order to give the plots a comparable scale). The degrees of freedom estimated for the t copula are plotted in parentheses. We note that the degrees of freedom of the t copula increase from 4.3 for hourly returns to 5.7 for daily returns. This is similar to the behavior of the tail index estimates for univariate data as a function of the time horizon; see Müller et al. (1998). It raises the question of what happens to the tail dependence when the time frequency of the returns vary. If we assume the t copula as a reasonable model for the data, from the results in Table 1 we can then estimate the tail dependence coefficient at the different time frequencies. This coefficient is given by

$$\lim_{\alpha \to 1^-} P(X_2 > F_2^{-1}(\alpha) | X_1 > F_1^{-1}(\alpha)) = \lambda$$

as long as the limit $\lambda \in [0, 1]$ exists. F_1 and F_2 denote the distribution functions of the random variables X_1 and X_2 respectively. If $\lambda \in (0, 1]$, X_1 and X_2 are asymptotically dependent and if $\lambda = 0$ the two variables are said to be asymptotically independent. In the case of the t copula, λ takes the form

$$\lambda = 2\bar{t}_{\nu+1}\left(\sqrt{(\nu+1)(1-\rho)/(1+\rho)}\right),$$

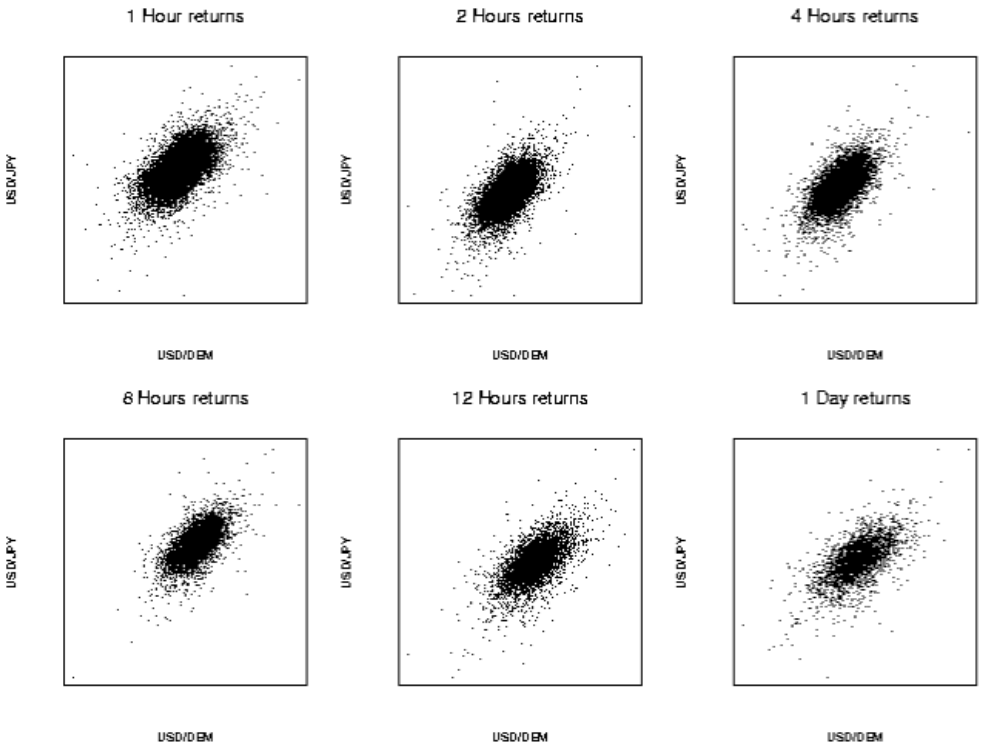

Figure 9: *Scatter plots of deseasonalised returns of DEM and JPY for different time frequencies.*

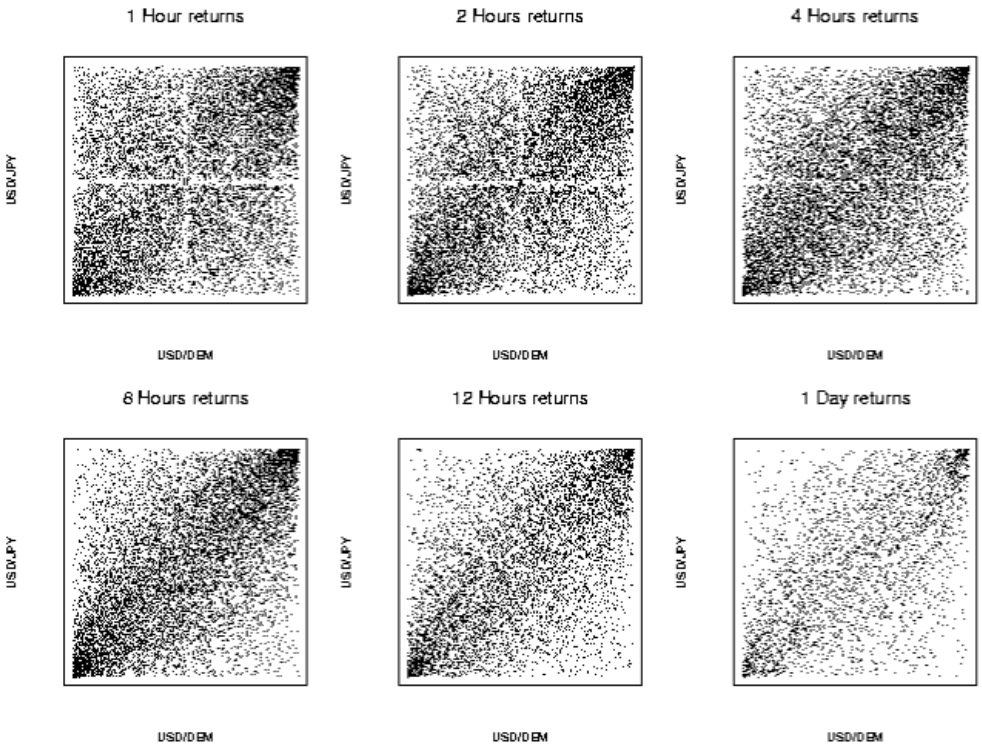

Figure 10: *Bivariate returns of DEM and JPY for different time frequencies mapped into the unit square by its marginal empirical distributions.*

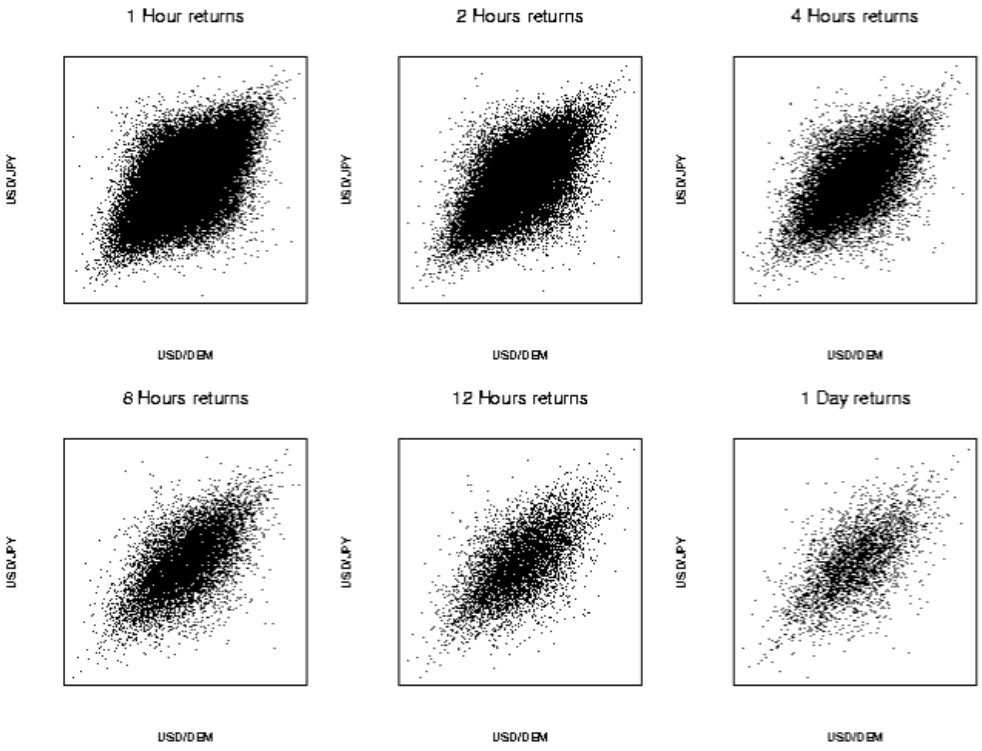

Figure 11: *Bivariate pseudo–observations for different time frequencies plotted with standard normal margins.*

Frequency	n. obs.	Model	Estimate	s.e.	AIC
1 hour	77 758	t	0.563; 4.339	0.003; 0.087	-32698
		Gumbel	1.577	0.005	-29052
		Gaussian	0.555	0.002	-28674
		Frank	4.030	0.025	-27275
		Clayton	0.880	0.007	-23997
2 hours	38 976	t	0.585; 4.269	0.004; 0.120	-17951
		Gumbel	1.622	0.007	-16066
		Gaussian	0.578	0.003	-15859
		Frank	4.252	0.036	-14983
		Clayton	0.944	0.010	-13267
4 hours	19 514	t	0.599; 4.282	0.005; 0.169	-9481
		Gumbel	1.652	0.009	-8530
		Gaussian	0.592	0.004	-8400
		Frank	4.402	0.051	-7957
		Clayton	0.978	0.014	-6936
8 hours	9 767	t	0.619; 4.833	0.007; 0.293	-5006
		Gumbel	1.688	0.014	-4561
		Gaussian	0.610	0.005	-4540
		Frank	4.633	0.073	-4347
		Clayton	1.020	0.020	-3697
12 hours	6 513	t	0.623; 5.438	0.008; 0.449	-3350
		Gaussian	0.617	0.007	-3111
		Gumbel	1.689	0.017	-3047
		Frank	4.680	0.089	-2953
		Clayton	1.037	0.025	-2518
1 day	3 259	t	0.624; 5.712	0.011; 0.714	-1675
		Gaussian	0.621	0.009	-1576
		Gumbel	1.689	0.024	-1525
		Frank	4.650	0.125	-1471
		Clayton	1.056	0.035	-1287

Table 1: Parameter estimates, standard errors and Akaike's information criterion values for the various copula models and time frequencies. For the t copula the first parameter estimate is the correlation and the second is the degrees of freedom and respectively for the s.e.'s.

where $\bar{t}_\nu$ denotes the tail of a standard univariate t distribution with ν degrees of freedom; see Embrechts et al. (2002). As we are dealing with the t copula, which has symmetric tails, we do not have to distinguish between upper and lower tail dependence. Table 2 shows the tail coefficient estimates for the time

Frequency	$\hat{\nu}$	$\hat{\rho}$	$\hat{\lambda}$
1 hour	4.339	0.563	0.273
2 hours	4.269	0.585	0.291
4 hours	4.282	0.599	0.299
8 hours	4.833	0.619	0.287
12 hours	5.438	0.623	0.264
1 day	5.712	0.624	0.254

Table 2: Tail coefficient estimates for the DEM and JPY bivariate returns for the different time frequencies considered.

frequencies considered for the bivariate DEM and JPY returns, assuming a t model for the dependence structure. The values obtained indicate that the bivariate returns DEM and JPY remain asymptotically dependent across the time frequencies considered. A confidence interval analysis can be worked out.

4 Goodness–of–fit tests

4.1 Test based on the probability integral transformation

A general goodness–of–fit test, valid for any copula family, can be performed using the following well known result, the probability integral transformation; see for instance Rosenblatt (1952).

Let $\mathbf{X} = (X_1, X_2, \ldots, X_d)$ be a random vector with absolutely continuous distribution function $F_{\mathbf{X}}(x_1, x_2, \ldots, x_d)$. Let $F_{X_i}(x_i) = P(X_i \leq x_i)$ be the distribution function of the univariate margins X_i, for $i = 1, \ldots, d$. Consider the d transformations

$$\begin{aligned} T(x_1) &= P(X_1 \leq x_1) = F_{X_1}(x_1), \\ T(x_2) &= P(X_2 \leq x_2 | X_1 = x_1) = F_{X_2|X_1}(x_2|x_1), \\ &\vdots \\ T(x_d) &= P(X_d \leq x_d | X_1 = x_1, \ldots, X_{d-1} = x_{d-1}) \\ &= F_{X_d|X_1,\ldots,X_{d-1}}(x_d|x_1, \ldots, x_{d-1}). \end{aligned}$$

Then the random variables $Z_i = T(X_i)$, for $i = 1, \ldots, d$ are uniformly and independently distributed on $[0, 1]^d$.

Suppose now that C is a copula such that

$$F_{\mathbf{X}}(x_1, x_2, \ldots, x_d) = C(F_{X_1}(x_1), \ldots, F_{X_d}(x_d)).$$

If $C_i(u_1, \ldots, u_i)$ denotes the joint i-marginal distribution

$$C_i(u_1, \ldots, u_i) = C(u_1, \ldots, u_i, 1, \ldots, 1), \qquad i = 2, \ldots, d-1,$$

of $(U_1, \ldots, U_i)$, with $C_1(u_1) = u_1$ and $C_d(u_1, \ldots, u_d) = C(u_1, \ldots, u_d)$, then the conditional distribution of U_i, given the values of $U_1, \ldots, U_{i-1}$, is

$$C_i(u_i|u_1, \ldots, u_{i-1}) = \frac{\partial^{i-1} C_i(u_1, \ldots, u_i)}{\partial u_1 \ldots \partial u_{i-1}} \Big/ \frac{\partial^{i-1} C_{i-1}(u_1, \ldots, u_{i-1})}{\partial u_1 \ldots \partial u_{i-1}},$$

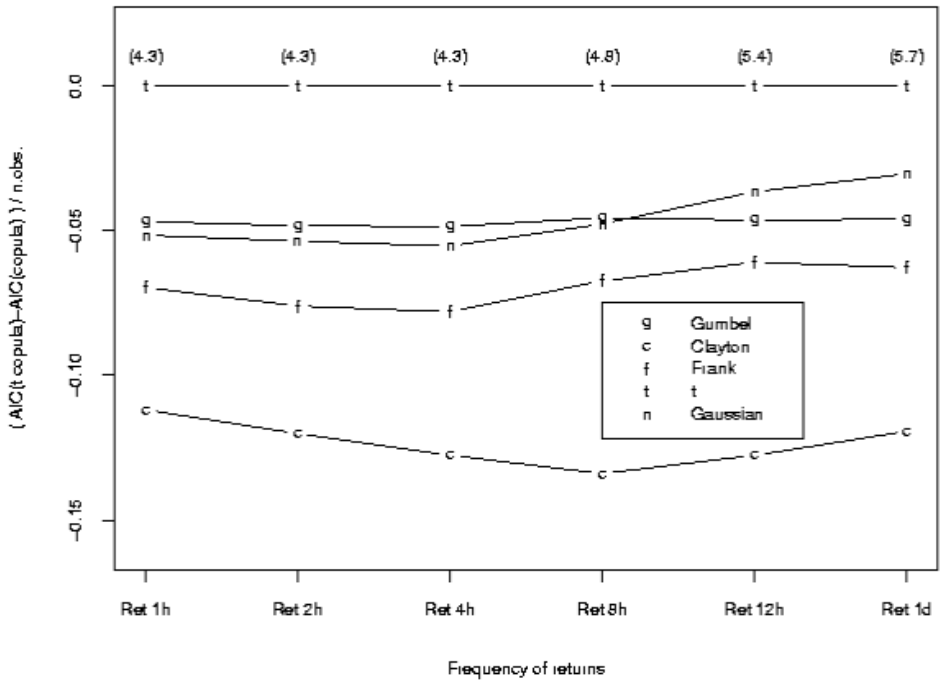

Figure 12: *Comparison of the AIC values for the different frequencies.*

for $i = 2, \ldots, d$. Hence we can write the variables Z_i, for $i = 2, \ldots, d$, using the conditional distributions C_i,

$$Z_i = C_i(F_{X_i}(X_i)|F_{X_1}(X_1), \ldots, F_{X_{i-1}}(X_{i-1})).$$

Hence if $(F_{X_1}(X_1), F_{X_2}(X_2), \ldots, F_{X_d}(X_d))$ has distribution function C, then $\Phi^{-1}(Z_i)$, $i = 1, \ldots, d$, are i.i.d., N$(0, 1)$ distributed. Consequently, $S = \sum_{i=1}^{d} \left(\Phi^{-1}(Z_i)\right)^2$ has a chi–square distribution with d degrees of freedom. Concretely, in the case $d = 2$,

$$S(X_1, X_2) = \left(\Phi^{-1}(F_{X_1}(X_1))\right)^2 + \left(\Phi^{-1}(C_2(F_{X_2}(X_2)|F_{X_1}(X_1)))\right)^2. \tag{18}$$

In performing the test, we do of course assume that the χ^2 distribution will not be significantly affected by the use of the empirical distribution functions used to transform the marginal data.

Other, related tests can for instance be found in Klugman and Parsa (1999) and Junker and May (2002).

4.2 Test of elliptical symmetry

The test of elliptical symmetry used here is due to Manzotti et al. (2002). Suppose that $\mathbf{X}$ is a d dimensional random vector with an elliptical distribution. $\mathbf{X}$ can be represented as

$$\mathbf{X} = \boldsymbol{\mu} + \mathbf{R}A\mathbf{U},$$

where $\boldsymbol{\mu} \in \mathbb{R}^d$, A is a non-singular $d \times d$ matrix, $\mathbf{R}$ is a real non negative random variable, $\mathbf{U}$ is uniformly distributed on the unit sphere $\mathbb{S}^{d-1} = \{\mathbf{x} \in \mathbb{R}^d : \| \mathbf{x} \| = 1\}$ and $\mathbf{R}$ and $\mathbf{U}$ are independent. Let $\Sigma = AA'$ be the shape matrix. The covariance matrix of $\mathbf{X}$, Σ_0 is proportional to Σ. Let $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ be an i.i.d. sample from a d dimensional distribution. The null hypothesis of the test is that the sample comes from an elliptically distributed population. Let $\bar{\mathbf{X}}$ and S denote the sample mean and covariance matrix, respectively. Consider the scaled residuals $\mathbf{Y}_k = S^{-1/2}(\mathbf{X}_k - \bar{\mathbf{X}})$ for $k = 1, \ldots, n$. Let $\mathbf{W}_k = \mathbf{Y}_k / \| \mathbf{Y}_k \|$ for $k = 1, \ldots, n$ be the projections of the scaled residuals on the unit sphere. If $\mathbf{X}$ is elliptically symmetric then $\mathbf{W}$ is approximately uniformly distributed on $\mathbb{S}^{d-1}$.

Consider $\varepsilon > 0$ fixed and let n_ε be the integer part of εn. Let q_n be the ε empirical quantile for the variables $\| \mathbf{Y}_1 \|, \| \mathbf{Y}_2 \|, \ldots, \| \mathbf{Y}_n \|$. With the average

$$Q_n(h) = \frac{1}{n} \sum_{k=1}^{n} h(\mathbf{W}_k) \mathbf{1}_{\{\|\mathbf{Y}_k\| > q_n\}},$$

where h is a function defined on $\mathbb{S}^{d-1}$, the test statistic Z_n^2 is given by

$$Z_n^2 = n \sum_{h \in \mathcal{J}_{jl}} Q_n^2(h)$$

for $j \geq 3$. The set of functions $\mathcal{J}_{jl}$ is the union $\mathcal{J}_{jl} = \bigcup_{j \leq i \leq l} \mathcal{H}_i$, where $\mathcal{H}_j$ denotes the set of spherical harmonics of degree j in the orthonormal basis as considered in Manzotti et al. (2002). This test statistic consists on averaging spherical harmonics over the projections of the $\mathbf{Y}_k$'s on the unit sphere. It will be useful to know that there are $N(d, j) = \binom{d+j-1}{j} - \binom{d+j-3}{j-2}$ linearly independent spherical harmonics of degree j in dimension d. Let N denote the number of functions in $\mathcal{J}_{jl}$. The main result of Manzotti et al.

(2002) states that independently of the unknown parameters of the distribution under the null hypothesis, the limiting distribution of Z_n^2 is that of $(1-\varepsilon)\chi^2$, where χ^2 is a variable with a chi-square distribution with N degrees of freedom.

With the $d = 2$ dimensional data we used $\varepsilon = 0.05$, $j = 3$ and $l = 6$. Then $N = 8$ and the orthonormal spherical harmonics used are

$$h_{1,j}(\mathbf{W}_k) = 2^{1/2}\cos(j\theta_k),\ h_{2,j}(\mathbf{W}_k) = 2^{1/2}\sin(j\theta_k),$$

for $3 \leq j \leq 6$ and $\mathbf{W}_k = (\cos(\theta_k), \sin(\theta_k))$.

4.2.1 The estimators

The estimation of the sample covariance matrix requires some care because in this setting, like elliptical distributions, we often have heavy tailed margins. In this case the standard estimators may have a poor performance. In order to test for ellipticality it is enough to estimate a matrix which is proportional to the covariance matrix. Indeed, if $\mathbf{X} = \boldsymbol{\mu} + \mathbf{R}A\mathbf{U}$ is elliptical then $\mathbf{Y} = \boldsymbol{\mu} + \mathbf{R}(cA)\mathbf{U}$, for $c > 0$, is also elliptical. Having this in mind, in the bivariate case, we can estimate the matrix

$$\sigma_1^{-2}\Sigma = \begin{pmatrix} 1 & \rho\frac{\sigma_2}{\sigma_1} \\ \rho\frac{\sigma_2}{\sigma_1} & \frac{\sigma_2^2}{\sigma_1^2} \end{pmatrix}$$

using more robust estimators. Here, σ_i^2 are the diagonal elements of Σ. For the linear correlation coefficient, under the elliptical assumption, the estimator based on Kendall's tau,

$$\hat{\rho}_\tau = \sin\left(\frac{\pi}{2}\hat{\tau}\right),$$

is more efficient and robust than Pearson's linear correlation estimator; see Lindskog et al. (2001). In order to estimate the ratio between the standard deviations we can use another dispersion estimator rather than the standard one, $S_n^2 = \sum_{i=1}^n (x_i - \overline{x})^2/n$, like the median absolute deviation, mad_n. The latter has some efficiency and bias problems but is very resistant to extreme observations coming from heavy tails; see for instance Pham-Gia and Hung (2001) or Hampel et al. (1986). The sample median absolute deviation is defined as follows:

$$mad_n(\mathbf{x}) = median(|\mathbf{x} - median(\mathbf{x})|),$$

where $median(\mathbf{x})$ denotes the sample median of the vector of observations $\mathbf{x}$. We then use

$$\widehat{\sigma_2/\sigma_1} = mad_n(\mathbf{x_2})/mad_n(\mathbf{x_1})$$

as an estimator more robust than the ratio of sample standard deviations.

4.3 Testing the results of the fittings

In Section 3, Table 1 we ranked the different models fitted according to their AIC values. Table 3 contains the p–values for the probability integral transformation goodness–of–fit test explained in Section 4.1. We have listed only the results for the best fitting models (minimal AIC) from Table 1. In the same table are also the p–values (fourth column) for the ellipticality test described in Section 4.2. To test whether

Time Frequency	Probability Integral test		P–values for the ellipticality test	
	Model	p-value	original margins	t margins
1 hour	t	0	0	0
2 hours	t	0	0	0
4 hours	t	0.01	0	0.092
8 hours	t	0.27	0	0.231
12 hours	t	0.19	0.034	0.369
1 day	t	0.74	0.821	0.675

Table 3: P–values of goodness–of–fit and ellipticality tests.

the values given by (18) come from a χ^2 distribution we use the Anderson–Darling goodness–of–fit test (see Anderson and Darling (1954)). Looking at columns three and four it is only at a frequency of one day that we cannot reject ellipticality and have strong support for a t copula. At higher frequencies, the situation is more subtle. The t copula fits well up to eight hours. For four hours and higher, a more careful analysis (and possibly more intricate copula) is needed, especially as at those frequencies very large sample sizes (e.g. $77'758$ bivariate hourly observations) are available. We will come back to this issue in forthcoming work. Note that the null hypothesis of ellipticality is rejected for frequencies higher than one day. In column five of Table 3, we perform the same ellipticality test, but now after transforming the marginals to a t distribution with the degrees of freedom coming from the columns three and four in Table 1. In this way we avoid that a rejection of the elliptical structure could come from non–elliptical margins. These new data are plotted in Figure 13. Ellipticality is now rejected only at the one and two hour frequency. Once more, a more detailed analysis, possibly also using non–parametric methodology, is called for. In the sequel of the paper, we concentrate on the important problem of clustering between extreme movements, in the literature also referred to as tail dependence.

5 Tail dependence

Several authors have looked at the issue of tail dependence in financial return data. See for instance Stărică (1999) for a more mathematical discussion. Other names encountered are contagion and spillover. Several references in Embrechts (2002) yield guidance towards the more economic oriented literature. By definition, the notion of tail dependence concerns bivariate rare events, hence limit theorems lie at the basis of any analysis. In order to investigate these problems on real data, a large number of observations is desirable. In Section 5.1 we start the study of the bivariate tail dependence with an estimation of the spectral measure for all the time horizons considered in the previous sections. We then concentrate in Section 5.2 on hourly data only. Bivariate extremes in these data will be analysed using the theory of multivariate regular variation, leading to a spectral analysis, and a statistical analysis of bivariate excesses over high thresholds.

5.1 Spectral measure estimation

The mathematics underlying this section is to be found in Resnick (2002) and Stărică (1999). Below we highlight the main definitions and notation. Let $\|\cdot\|$ denote the usual Euclidean L_2 norm on $\mathbb{R}^d$ and $\mathbb{S}^{d-1}$ be the unit sphere,

$$\mathbb{S}^{d-1} := \{\mathbf{x} \in \mathbb{R}^d : \| \mathbf{x} \| = 1\}.$$

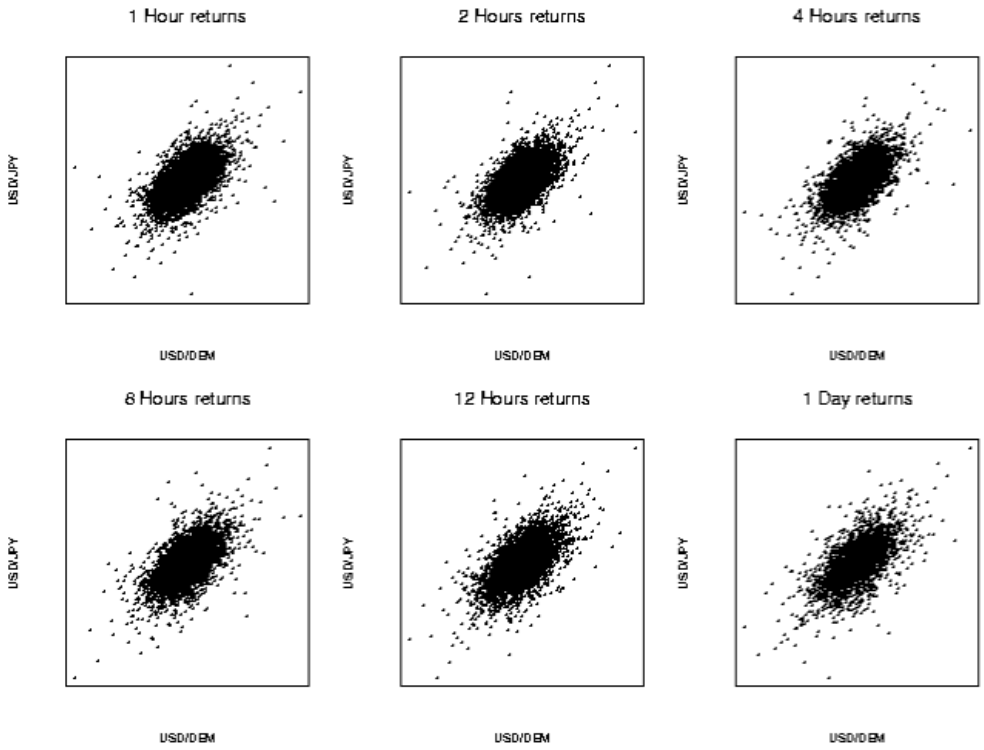

Figure 13: *Bivariate pseudo-observations for different time frequencies plotted with t margins.*

Suppose that the d-dimensional random vector $\mathbf{X}$ has a regularly varying tail distribution. This means that the tail behaviour of $\mathbf{X}$ is characterised by a tail index α and the limit

$$\frac{P(\| \mathbf{X} \| > tx, \mathbf{X}/\| \mathbf{X} \| \in \cdot)}{P(\| \mathbf{X} \| > t)} \quad \xrightarrow{v} \quad x^{-\alpha} P(\Theta \in \cdot), \tag{19}$$

where $x > 0$, $t \to \infty$ exists. The convergence is said to be vague and Θ is a random vector on the space $(\mathbb{S}^{d-1}, \mathcal{B}(\mathbb{S}^{d-1}))$. The distribution function of Θ is referred to as the spectral distribution of $\mathbf{X}$. Definition (19) is equivalent to the existence of a measure ν and a positive sequence (a_n), $a_n \to \infty$, such that for $n \to \infty$,

$$nP(a_n^{-1}\mathbf{X} \in \cdot) \quad \xrightarrow{v} \quad \nu(\cdot). \tag{20}$$

For a more precise and detailed treatment on this see for instance Resnick (1987). The measure ν has the following scaling property:

$$\nu(vS) = v^{-\alpha}\nu(S), \tag{21}$$

for any Borel set $S \subset [-\infty, \infty]^d \backslash \{0\}$. This property will be useful in order to find an estimator for the spectral distribution. Intuitively, α indicates the heaviness of the multivariate tails whereas Θ measures in which parts of the space extremes cluster.

Define for $\mathbf{x} \in \mathbb{R}^d$ and $B \in \mathcal{B}(\mathbb{R}^d)$

$$\epsilon_x(B) = \begin{cases} 1 & \text{if } x \in B, \\ 0 & \text{if } x \in B^c. \end{cases}$$

Then a consistent estimator of $c\nu$, for some $c > 0$, is given by

$$\nu_n := \frac{1}{k_n} \sum_{i=1}^{n} \epsilon_{\mathbf{X}_i / b(n/k_n)},$$

where $b(.)$ is the quantile function $b(t) := F^{\leftarrow}(1 - 1/t)$, for $t > 1$, of the random variable $\| \mathbf{X} \|$. As usual in extreme value theory, $k_n \to \infty$ and $k_n/n \to 0$ as $n \to \infty$; see Resnick (2002). If we estimate the quantile function with the corresponding empirical estimator

$$\hat{b}\left(\frac{n}{k_n}\right) = \| \mathbf{X} \|_{k_n, n},$$

where $\| \mathbf{X} \|_{k_n,n}$ is the k-th largest value of the one-dimensional $\{\| \mathbf{X}_i \| : 1 \le i \le n\}$ set, we obtain as estimator of the spectral measure

$$\hat{P}(\Theta \in S) = \frac{1}{k_n} \sum_{i=1}^{n} \epsilon_{\mathbf{X}_i / \| \mathbf{x} \|_{k_n, n}}(V(S)) \tag{22}$$

where $V(S) = \{\mathbf{x} \in \mathbb{S}_+^{d-1} : \mathbf{x}/\| \mathbf{x} \| \in S\}$ and $\mathbb{S}_+^{d-1} := \{\mathbf{x} : \| \mathbf{x} \| > 1\}$. The performance of this estimator very much depends on the choice of k_n. Here we use the scaling property (21) and choose k_n such that $\hat{\nu}_n(u\mathbb{S}_+^{d-1})/(u^{-\alpha}\hat{\nu}_n(\mathbb{S}_+^{d-1})) \approx 1$ for values of u in a neighbourhood of 1. We plot the set of values

$$\left\{ \left(u, \frac{\hat{\nu}_n(u\mathbb{S}_+^{d-1})}{u^{-\hat{\alpha}}\hat{\nu}_n(\mathbb{S}_+^{d-1})} \right) : 0 < u < 2 \right\}$$

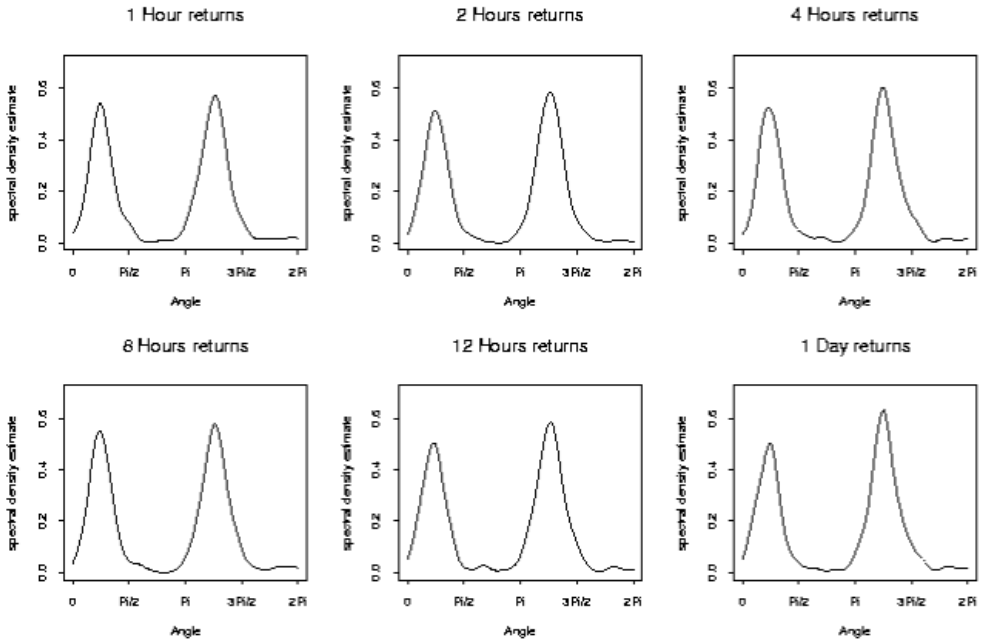

Figure 14: *The estimated spectral measures for the different time frequencies.*

for several values of k_n and choose the one corresponding to the plot for which these values are closer to 1 around $u = 1$. For more on this see Stărică (1999). We use the Hill estimator to get the tail index estimate $\hat{\alpha}$. The values obtained lie around 4.

We now estimate the spectral density of the bivariate returns $\mathbf{X}$ of the DEM and JPY data at a given time frequency using (22). First, choose k_n as described above and consider the points

$$\left\{\theta_i \in [0, 2\pi[: (\cos\theta_i, \sin\theta_i) = \frac{\mathbf{x}_i}{\| \mathbf{x}_i \|}, \ \| \mathbf{x}_i \| > \| \mathbf{x} \|_{k_n,n}, i = 1, \ldots, n\right\}.$$

We then plot a non–parametric density estimate for these angular observations using a smoothed kernel estimator with Gaussian weights and bandwidth 0.2π. In a more detailed analysis one could also work out confidence bands around the estimated functions. In Figure 14 we have plotted the estimated spectral densities for each time frequency. We would first like to point out that for a spherical distribution, a fairly constant spectral density would appear. Peakedness in one or other direction points at clustering of extremes in that direction. Also, the procedure discussed only uses data in the tails so that no information from the centre of the distribution enters. Figure 14 clearly shows clustering of extreme returns in the first and third quadrant, also referred to as positive dependence. This dependence persists at all frequencies and turns out to be fairly symmetric. Of course, a basic assumption concerns the property (19); as in the one–dimensional case, one can show, using extreme value theory, that the limit property (19) is very natural for multivariate financial return data. For an interesting paper leading to similar conclusions, see Hauksson et al. (2001).

5.2 Multivariate excesses

In the previous section, under the assumption (19), we modelled the ocurrence of joint extremes for bivariate returns by the density of the spectral random variable Θ. In this section we focus, for a given high threshold t, on the event $\{X_1 > t, X_2 > t\}$. Similarly we can define sets of returns simultaneously smaller than a given threshold. As the threshold t will be large (small) we will concentrate only on the one hour data which allows for sufficient data in these cases. In Figure 15 we plotted the bivariate excesses for different values of the threshold. For univariate random variables, the Balkema–de Haan–Pickands result (Theorem 3.4.13(b) in Embrechts et al. (1997)) yields the generalised Pareto distribution as a canonical model for the distribution function of conditional excesses. In Juri and Wüthrich (2002), a similar result is proved for bivariate excesses in the case of Archimedean copulas (note that in that case, one makes an assumption on the copula for the whole domain of the bivariate dependence structure model). We summarise below their main result which forms the basis for our statistical analysis. The copula $C(u, v)$, with $0 \leq u, v \leq 1$, of the random vector (U, V) is called Archimedean if there exists a continuous, strictly decreasing function, $\psi : [0, 1] \mapsto [0, \infty]$ with $\psi(1) = 0$, such that

$$C(u, v) = \psi^{[-1]}(\psi(u) + \psi(v)).$$

The function $\psi^{[-1]} : [0, \infty] \mapsto [0, 1]$ is defined by $\psi^{[-1]}(x) = \psi^{-1}(x)\mathbf{1}_{[0,\psi(0)]}(x)$ and is called the generator of the copula C. Denote by F_t the conditional distribution function

$$F_t(u) := P(U \leq u | U \leq t, V \leq t),\ 0 \leq u \leq 1.$$

The extreme tail dependence copula of the copula C relative to a threshold t is given by

$$C_t(u, v) = P(U \leq F_t^{-1}(u), V \leq F_t^{-1}(v) | U \leq t, V \leq t).$$

If C is an Archimedean copula having a regularly varying differentiable generator $\psi \in \mathcal{R}_{-\alpha}$ with $0 < \alpha < \infty$, then

$$\lim_{t \to 0^+} C_t(u, v) = C_\alpha^{Cl}(u, v), \tag{23}$$

for all $0 \leq u, v \leq 1$. The limit copula in this result, C_α^{Cl}, is the Clayton copula with parameter $\alpha > 0$ defined by

$$C_\alpha^{Cl}(u, v) = (u^{-\alpha} + v^{-\alpha} - 1)^{-1/\alpha},$$

for $0 < u\,, v \leq 1$. Juri and Wüthrich (2002) show that the condition $\psi \in \mathcal{R}_{-\alpha}$ is a very natural one which holds for several known examples.

For the one hour pseudo–returns of DEM and JPY, $(\hat{F}_{1n}(x_{1i}), \hat{F}_{2n}(x_{2i}))$ with $i = 1, \ldots, n$, we considered several thresholds t (both in the joint left as well as in the joint right tails) and fitted copula models C_t. The thresholds and the resulting data are to be found in Figure 16. Hence we want to model

$$C_{t^-}(u, v) = P(U \leq F_t^{-1}(u), V \leq F_t^{-1}(v) | U \leq t, V \leq t)$$

as well as

$$C_{t^+}(u, v) = P(U \leq F_t^{-1}(u), V \leq F_t^{-1}(v) | U \geq t, V \geq t).$$

In each case we fitted a list of copula models including the Gaussian, t, Gumbel, Frank, Clayton,

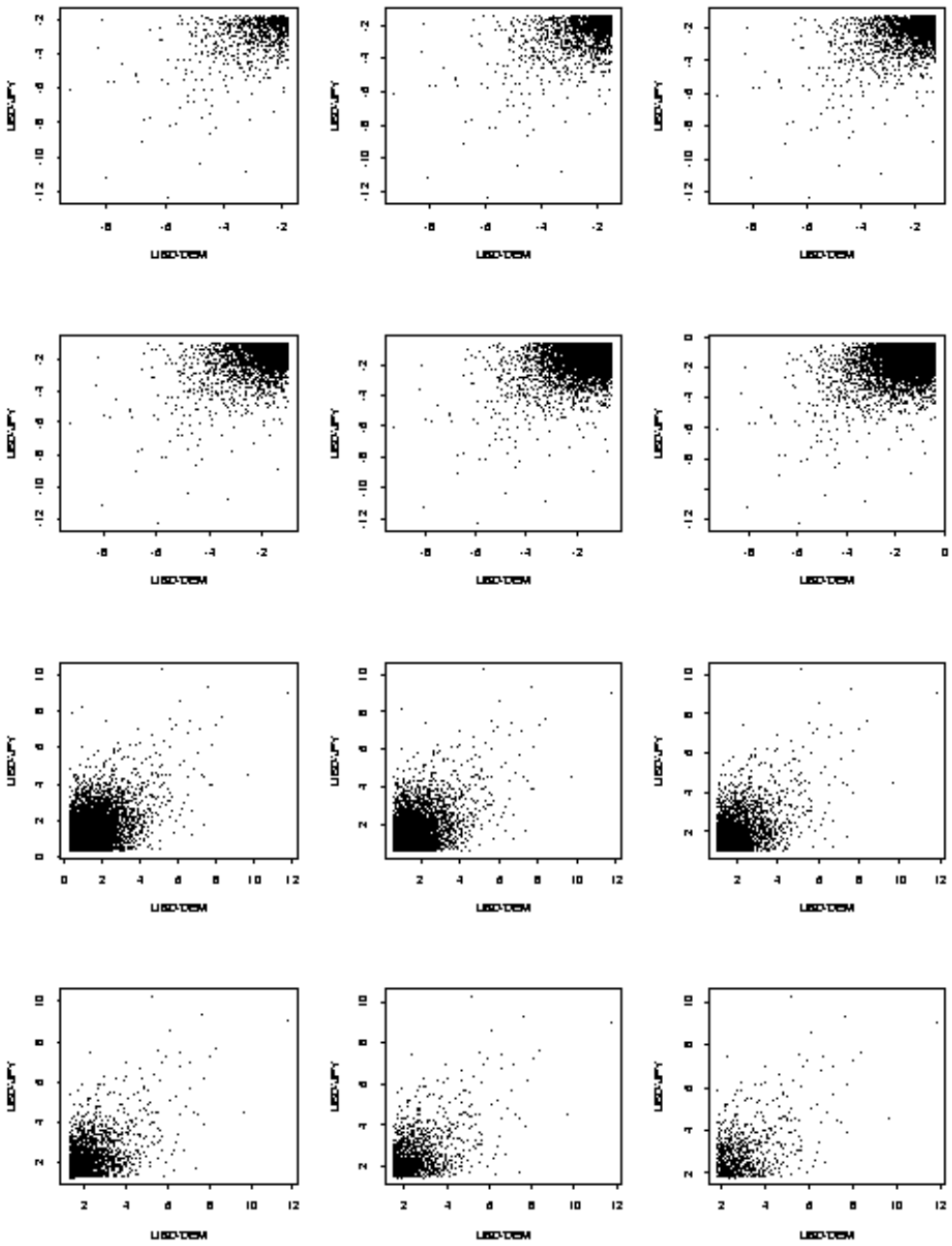

Figure 15: *The bivariate excesses of the one hour returns for different thresholds.*

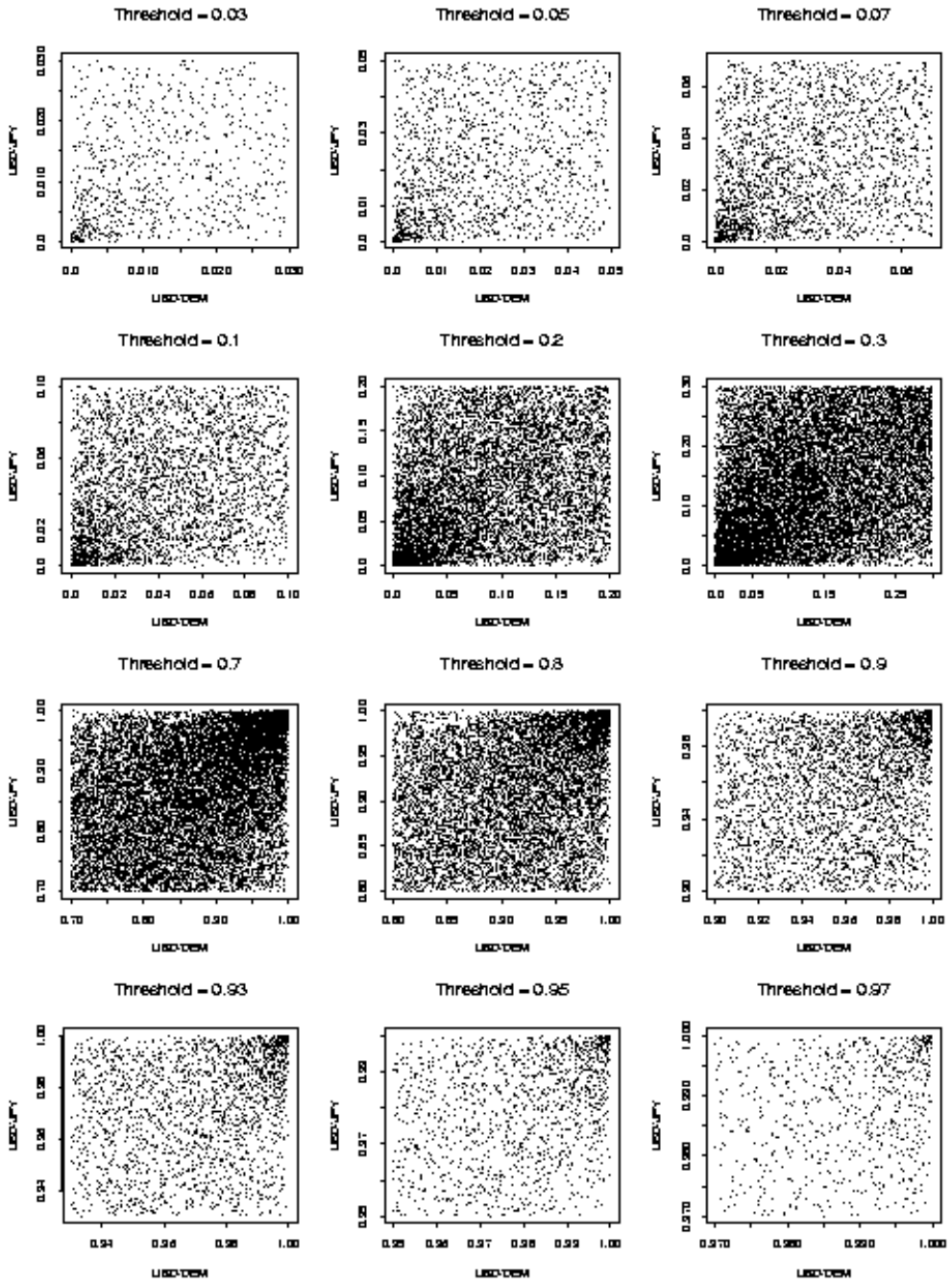

Figure 16: *The bivariate excesses of the one hour returns mapped into the unit square by the empirical distributions of the margins. Only the tail regions are shown.*

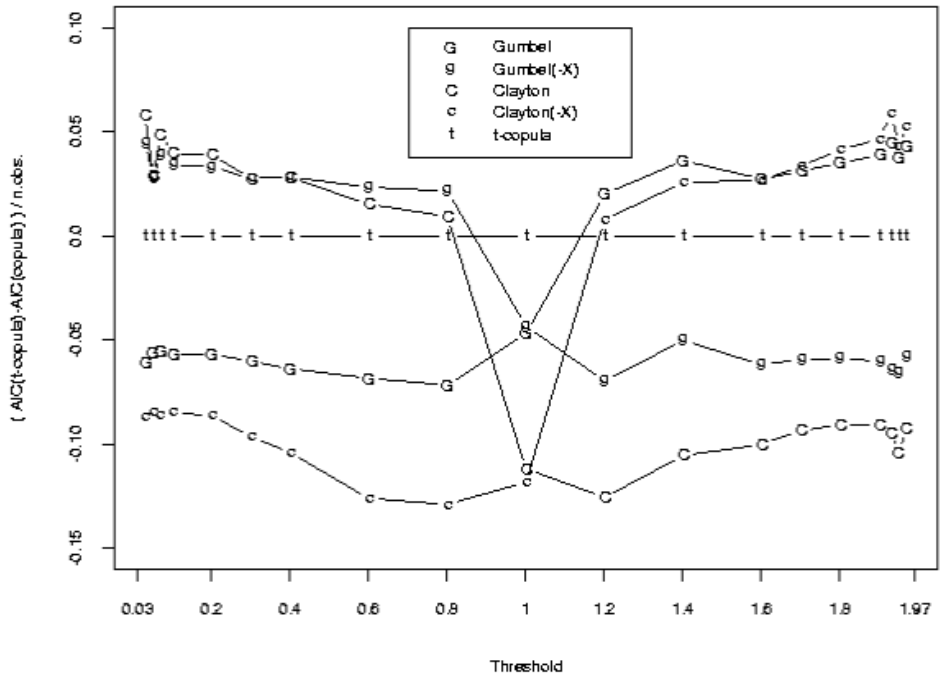

Figure 17: *Comparison of the AIC values for the different thresholds.*

survival Gumbel, survival Clayton and Farlie–Gumbel–Morgenstern copulas. For the definition of the copulas used, see Embrechts et al. (2002), Nelsen (1999) and Joe (1997). The "survival" stands for the copula applied to minus the random variables (hence left and right tails are exchanged). The results are reported in Tables 4 and 5 where the models fitted are ranked by Akaike's information criterion. In Table 4 are the results for C_{t^-}. The second column contains the number of observations below t in percentage of the total data. For the t copula, the first parameter estimate is the correlation and the second is the degrees of freedom. The Clayton copula is always the best one. Remember that the best fitting for the dependence structure of the full hourly data set was attained with the t copula, although without passing the goodness–of–fit test. Note that, for the considered thresholds, the Clayton parameter ranges from $\hat{\alpha} = 0.556$ to $\hat{\alpha} = 0.609$ which corresponds to a Kendall tau coefficient between $\hat{\tau} = 0.217$ and $\hat{\tau} = 0.233$ (for the Clayton copula with parameter α, $\tau = \alpha/(\alpha + 2)$). Table 5 contains the results for C_{t^+}. In this case, the survival Clayton copula yields the best fit. In this case the survival Clayton parameter varies between $\hat{\alpha} = 0.574$ and $\hat{\alpha} = 0.666$, corresponding to a Kendall tau from $\hat{\tau} = 0.223$ to $\hat{\tau} = 0.250$. Hence showing a slightly stronger dependence than for simultaneously small returns. Both tables are summarised graphically in Figure 17 (similar to Figure 12 for the full copula data).

Threshold	n. obs.	Model	Estimate	s.e.	AIC	p-value
		Clayton	0.583	0.059	-132.19	0.99
		Surv. Gumbel	1.295	0.034	-122.04	0.99
		t	0.335; 13.69	0.033; 8.219	-88.00	0.99
0.03^-	759	Gaussian	0.336	0.030	-86.79	0.92
	(0.98%)	Frank	2.031	0.227	-78.13	0.95
		F–G–M	0.873	0.086	-70.48	0.40
		Gumbel	1.195	0.033	-41.71	0.89
		Surv. Clayton	0.238	0.052	-21.81	0.45
		Clayton	0.561	0.043	-227.74	0.98
		Surv. Gumbel	1.298	0.025	-227.13	0.99
		t	0.342; 8.40	0.025; 2.31	-188.48	0.97
0.05^-	1 376	Gaussian	0.348	0.022	-173.46	0.68
	(1.77%)	Frank	2.070	0.169	-147.70	0.78
		F–G–M	0.863	0.062	-129.49	0.10
		Gumbel	1.223	0.024	-110.96	0.78
		Surv. Clayton	0.305	0.039	-72.38	0.23
		Clayton	0.556	0.035	-350.71	0.99
		Surv. Gumbel	1.283	0.020	-330.20	0.93
		t	0.330; 11.44	0.019; 3.35	-248.47	0.95
0.07^-	2 112	Gaussian	0.328	0.018	-236.91	0.79
	(2.72%)	Frank	1.999	0.135	-216.32	0.84
		F–G–M	0.903	0.054	-204.21	0.17
		Gumbel	1.200	0.019	-132.03	0.86
		Surv. Clayton	0.244	0.031	-66.80	0.27
		Clayton	0.558	0.028	-547.96	0.91
		Surv. Gumbel	1.289	0.016	-529.62	0.61
		t	0.340; 13.43	0.015; 3.62	-418.10	0.85
0.1^-	3 273	Gaussian	0.342	0.014	-403.71	0.81
	(4.21%)	Frank	2.046	0.108	-352.09	0.74
		F–G–M	0.897	0.040	-326.39	0.07
		Gumbel	1.209	0.015	-230.55	0.81
		Surv. Clayton	0.273	0.025	-140.27	0.15
		Clayton	0.556	0.018	-1302.15	0.96
		Surv. Gumbel	1.289	0.010	-1257.91	0.79
		t	0.340; 12.19	0.010; 1.90	-1000.15	0.82
0.2^-	7 807	Gaussian	0.339	0.009	-952.77	0.21
	(10.0%)	Frank	2.085	0.070	-870.88	0.47
		F–G–M	0.905	0.026	-803.05	0.001
		Gumbel	1.211	0.010	-555.75	0.50
		Surv. Clayton	0.271	0.016	-324.74	0.01
		Clayton	0.609	0.014	-2547.77	0.47
		Surv. Gumbel	1.330	0.008	-2546.12	0.68
		t	0.383; 12.59	0.007; 1.52	-2178.67	0.85
0.3^-	13 359	Gaussian	0.381	0.006	-2097.90	0.06
	(17.2%)	Frank	2.401	0.054	-1944.88	0.39
		F–G–M	0.982	0.015	-1750.99	0.00
		Gumbel	1.260	0.008	-1374.71	0.20
		Surv. Clayton	0.348	0.012	-889.42	0.00

Table 4: Fitting results for bivariate excesses on the third quadrant of one hour returns for different thresholds.

Threshold	n. obs.	Model	Estimate	s.e.	AIC	p-value
		Surv. Clayton	0.601	0.060	-137.23	0.94
		Gumbel	1.306	0.035	-130.25	0.97
		t	0.343; 8.42	0.034; 3.45	-98.15	0.98
0.97^+	745	Gaussian	0.350	0.030	-93.00	0.91
	(0.96%)	Frank	2.048	0.229	-77.25	0.96
		F–G–M	0.878	0.090	-68.27	0.33
		Surv. Gumbel	1.221	0.033	-54.98	0.93
		Clayton	0.278	0.054	-29.41	0.42
		Surv. Clayton	0.666	0.046	-288.12	0.92
		Gumbel	1.351	0.027	-280.52	0.57
		Gaussian	0.403	0.021	-230.68	0.94
0.95^+	1 331	t	0.400; 23.09	0.022; 16.97	-230.50	0.84
	(1.71%)	Frank	2.417	0.172	-196.73	0.71
		F–G–M	0.999	0.072	-189.17	0.25
		Surv. Gumbel	1.267	0.026	-142.60	0.87
		Clayton	0.359	0.041	-91.74	0.52
		Surv. Clayton	0.597	0.036	-378.41	0.99
		Gumbel	1.299	0.021	-350.57	0.92
		t	0.339; 11.25	0.020; 3.40	-260.50	0.95
0.93^+	2 014	Gaussian	0.344	0.018	-249.74	0.92
	2.59%)	Frank	2.032	0.139	-212.26	0.86
		F–G–M	0.896	0.054	-196.21	0.15
		Surv. Gumbel	1.203	0.020	-131.48	0.93
		Clayton	0.248	0.032	-69.92	0.30
		Surv. Clayton	0.583	0.028	-575.19	0.93
		Gumbel	1.299	0.017	-550.75	0.78
		t	0.345; 11.97	0.016; 3.13	-428.26	0.89
0.9^+	3 167	Gaussian	0.351	0.014	-413.18	0.82
	(4.07%)	Frank	2.050	0.110	-340.33	0.83
		F–G–M	0.896	0.043	-309.91	0.05
		Surv. Gumbel	1.215	0.016	-236.88	0.80
		Clayton	0.277	0.025	-141.00	0.10
		Surv. Clayton	0.574	0.018	-1376.35	0.99
		Gumbel	1.298	0.010	-1327.02	0.83
		t	0.345; 10.76	0.010; 1.56	-1055.79	0.87
0.8^+	7 765	Gaussian	0.348	0.009	-1001.24	0.27
	(9.99%)	Frank	2.091	0.071	-868.78	0.41
		F–G–M	0.910	0.026	-797.18	0.00
		Surv. Gumbel	1.218	0.010	-599.50	0.41
		Clayton	0.280	0.016	-352.28	0.01
		Surv. Clayton	0.594	0.014	-2459.31	0.84
		Gumbel	1.315	0.008	-2426.10	0.60
		t	0.366;11.63	0.007; 1.36	-2014.98	0.75
0.7^+	13 300	Gaussian	0.367	0.007	-1928.50	0.09
	(17.1%)	Frank	2.253	0.054	-1718.03	0.27
		F–G–M	0.954	0.018	-1561.01	0.00
		Surv. Gumbel	1.242	0.008	-1226.49	0.22
		Clayton	0.320	0.012	-770.77	0.00

Table 5: Fitting results for bivariate excesses on the first quadrant of one hour returns for different thresholds.

6 Conclusion and further work

In this paper we analysed the dependence structure within two–dimensional, high–density FX return data. The methods used are copula modelling together with statistical techniques for extremal clustering. An overall picture emerged that is as follows: At all time horizons the data can be fitted best with t-copulas with successively higher degrees of freedom as the time horizon increases. Note that the t-copula is rejected for the shortest horizons because of the large amount of data. This means that the t-copula has not enough structure to properly describe the details which can be discerned with such a large sample. The test for ellipticality is not rejected except for the 1 hour and 2 hours horizons if the margins are transformed to t-distribution with the number of degrees of freedoms adjusted to the result of the copula fit. With the empirical margins, ellipticality is rejected for horizons of 8 hours and shorter. The spectral measure, however, shows pronounced peaks in the diagonals for all time horizons. An analysis of the multivariate excesses of hourly returns shows that the lower left tails are best described with Clayton/survival Gumbel copulas while the upper right tails are best described with Gumbel/survival Clayton copula. These results are predicted by theory.

Our results extend the univariate stylized facts to the bivariate case and give valuable indications for time series models. However, further work is no doubt necessary at several levels. For instance:

- Though we introduced a method for multidimensional deseasonalisation, more work on this topic is needed. This not only for higher dimensional data, but also for different types of data. We thrust our results to be fairly insensitive with respect to changes in the deseasonalisation used for the time horizons investigated. For time horizons larger than one day deseasonalisation has to be done differently, as discussed in Section 2.3.4. Besides the method presented in Section 2.3, we analysed the data also using different time transformation tools.

- Throughout, we used a static stochastic model. As in the one–dimensional case, stationary models allowing for a richer volatility structure (as there are GARCH and stochastic volatility effects) are to be analysed. The methods introduced may then be used at the level of the residuals.

- There are several statistical issues which no doubt need a more detailed discussion. The large data size at the high–frequency level would allow for non–(or semi–) parametric modelling. Also an analysis taking a broader class of copulas into account could be useful; we analysed the data using several mixture classes of standard copulas (as for instance in Junker and May (2002)). The results obtained differed not significantly.

Acknowledgement

The authors take pleasure in thanking Guus Balkema, Filip Lindskog and Alexander McNeil for several helpful discussions. We thank Olsen Data for providing the FX data.

References

Andersen T. G. and Bollerslev T., 1997, *Intraday periodicity and volatility persistence in financial markets*, Journal of Empirical Finance, **4**(2-3), 115–158.

Andersen T. G. and Bollerslev T., 1998, *Deutsche Mark-Dollar volatility: intraday activity patterns, macroeconomic announcements, and longer run dependencies*, the Journal of Finance, **53**(1), 219–265.

Anderson T. W. and Darling D. A., 1954, *A test of goodness of fit*, Journal of the American Statistical Association, **49**, 765–769.

Barndorff-Nielsen O. E. and Prause K., 2001, *Apparent scaling*, Finance and Stochastics, **5**, 103–113.

Beltratti A. and Morana C., 1999, *Computing value at risk with high frequency data*, Journal of Empirical Finance, **6**(5), 431–455.

Bollerslev T. and Ghysels E., 1996, *Periodic autoregressive conditional heteroskedasticity*, Journal of Business & Economic Statistics, **14**(2), 139–150.

Breymann W., 2000, *Dynamic theta time: Algorithm, configuration, tests*, Internal document WAB.2000-07-31, Olsen & Associates, Seefeldstrasse 233, 8008 Zürich, Switzerland.

Dacorogna M. M., Gençay R., Müller U. A., Olsen R. B., and Pictet O. V., 2001, *An Introduction to High-Frequency Finance*, Academic Press, San Diego, CA.

Dacorogna M. M., Müller U. A., Nagler R. J., Olsen R. B., and Pictet O. V., 1993, *A geographical model for the daily and weekly seasonal volatility in the FX market*, Journal of International Money and Finance, **12**(4), 413–438.

Embrechts P., 2002, *Extremes in economics and the economics of extremes*, Available online from `http://www.math.ethz.ch/~embrechts/`.

Embrechts P., Klüppelberg C., and Mikosch T., 1997, *Modelling Extremal Events for Insurance and Finance*, Springer, Berlin.

Embrechts P., McNeil A. J., and Straumann D., 2002, *Correlation and dependence in risk management: Properties and pitfalls*, In: Risk Management: Value at Risk and Beyond (M. Dempster, Ed.) Cambridge University Press, Cambridge, 176–223.

Genest C., Ghoudi K., and Rivest L.-P., 1995, *A semiparametric estimation procedure of dependence parameters in multivariate families of distributions*, Biometrika, **82**(3), 543–52.

Hampel F. R., Ronchetti E. M., Rousseeuw P. J., and Stahel W. A., 1986, *Robust Statistics: The Approach Based on Influence Functions*, John Wiley and Sons, New York.

Hauksson H. A., Dacorogna M. M., Domenig T., Müller U. A., and Samorodnitsky G., 2001, *Multivariate extremes, aggregation and risk estimation*, Quantitative Finance, **1**, 79–95.

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

Junker M. and May A., 2002, *Measurement of aggregate risk with copulas*, Preprint of the Center of Advanced European Studies and Research, Bonn, Germany.

Juri A. and Wüthrich M. V., 2002, *Copula convergence theorems for tail events*, Insurance: Mathematics and Economics, **30**, 405–420.

Klugman S. A. and Parsa R., 1999, *Fitting bivariate loss distributions with copulas*, Insurance: Mathematics and Economics, **24**, 139–148.

Lindskog F., McNeil A. J., and Schmock U., 2001, *Kendall's Tau for Elliptical Distributions. Working paper, ETH Zürich*, Available online from `http://www.risklab.ch/Papers.html`.

Manzotti A., Perez F. J., and Quiroz A. J., 2002, *A statistic for testing the null hypothesis of elliptical symmetry*, Journal of Multivariate Analysis, **81**, 274–285.

Martens M., Chang Y.-C., and Taylor S. J., 2002, *A comparison of seasonal adjustment methods when forecasting intraday volatility*, The Journal of Financial Research, **15**(2), 283–299.

Müller U. A., Dacorogna M. M., and Pictet O. V., 1998, *Heavy tails in high-frequency financial data*, published in the book "A practical guide to heavy tails: Statistical Techniques for Analysing Heavy Tailed Distributions", edited by Robert J. Adler, Raisa E. Feldman and Murad S. Taqqu and published by Birkhäuser, Boston 1998, 55–77.

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

Pham-Gia T. and Hung T. L., 2001, *The mean and median absolute deviations*, Mathematical and Computer Modelling, **34**, 921–936.

Resnick S., 1987, *Extreme Values, Regular Variation, and Point Processes*, Springer, New York.

Resnick S., 2002, *On the foundations of multivariate heavy tail analysis*, Cornell Reports and Papers, Available online from `http://www.orie.cornell.edu/~sid`.

Rosenblatt M., 1952, *Remarks on a multivariate transformation*, The Annals of Mathematical Statistics, **23**, 470–472.

Stărică C., 1999, *Multivariate extremes for models with constant conditional correlations*, Journal of Empirical Finance, **6**, 515–553.

Taylor S. J. and Xu X., 1997, *The incremental volatility information in one million foreign exchange quotations*, the Journal of Empirical Finance, **4**(4), 317–340.