Time varying Hierarchical Archimedean Copulae

Wolfgang Karl Härdle*
Ostap Okhrin*
Yarema Okhrin**

* Humboldt-Universität zu Berlin, Germany
** Universität Augsburg, Germany

BERLIN

ECONOMIC RISK

SFB 649

Time varying Hierarchical Archimedean Copulae[*]

Wolfgang Karl Härdle[†], Ostap Okhrin[‡], Yarema Okhrin[§]

17th February 2010

Abstract:

There is increasing demand for models of time-varying and non-Gaussian dependencies for multivariate time-series. Available models suffer from the curse of dimensionality or restrictive assumptions on the parameters and the distribution. A promising class of models are the hierarchical Archimedean copulae (HAC) that allow for non-exchangeable and non-Gaussian dependency structures with a small number of parameters. In this paper we develop a novel adaptive estimation technique of the parameters and of the structure of HAC for time-series. The approach relies on a local change point detection procedure and a locally constant HAC approximation. Typical applications are in the financial area but also recently in the spatial analysis of weather parameters. We analyse the time varying dependency structure of stock indices and exchange rates. We find that for stock indices the copula parameter changes dynamically but the hierarchical structure is constant over time. Interestingly in our exchange rate example both structure and parameters vary dynamically.

Keywords: copula; multivariate distribution; Archimedean copula; adaptive estimation.

JEL Classification: C13, C14, C50.

[†]C.A.S.E. - Center for Applied Statistics and Economics, Ladislaus von Bortkiewicz Chair of Statistics of Humboldt-Universität zu Berlin, Spandauer Straße 1, D-10178 Berlin, Germany and National Central University, Department of Finance, No. 300, Jhongda Rd., Jhongli City, Taoyuan County 32001, Taiwan (R.O.C.). Email: haerdle@wiwi.hu-berlin.de.

[‡]C.A.S.E. - Center for Applied Statistics and Economics, Ladislaus von Bortkiewicz Chair of Statistics of Humboldt-Universität zu Berlin, Spandauer Straße 1, D-10178 Berlin, Germany. Email: ostap.okhrin@wiwi.hu-berlin.de

[§]Department of Statistics, University of Augsburg, D-86135 Augsburg, Germany. Email: yarema.okhrin@wiwi.uni-augsburg.de

1 Introduction

The key difference between univariate and multivariate time series analysis is the fact that the future dynamics is affected not only by the univariate past but also by cross-sectional dependencies. These dependencies are not constant and vary in time. Their dynamics, form and strength are important in many applications. Risk diversification, asset allocation, financial spillovers illustrate this importance. The most straightforward and therefore well established approach to modeling dependencies is via the correlation (or covariance) matrix. The correlation matrix uniquely characterises the dependency, if the data is driven, for example, by a multivariate normal distribution. Similar arguments also hold for arbitrary elliptical distributions (as the multivariate t). Due to its simplicity the covariance matrix became the standard parametrisation of dependency. In many applications the dependency structure varies over time. Time varying conditional volatilities are modeled using e.g. GARCH-type processes. For a recent review of multivariate GARCH processes, including DCC, CCC, BEKK, among others, we refer to Silvennoinen and Teräsvirta (2009). These models though still assume parameters of the process that are constant over an entire estimation period. Such approach has been challenged even in the univariate case as the growing literature demonstrates, see Lamoureux and Lastrapes (1990). In practice, however, it is likely that the parameters characterising the dependency change with time in possibly a nonstationary manner.

Another disadvantage of covariance-based dependency modeling is the fact that it fails to capture important types of data features. First, covariances are measures of linear dependence and therefore fail to represent nonlinear relationships. As an alternative approach one may consider other measures such as Kendall's τ or Spearman's ρ, see Joe (1997). However, the extensions of these measures to higher dimensions is problematic, see e.g. Schmid and Schmidt (2006). Second, elliptical distributions postulate symmetric dependency, i.e. the strength of the relationship is the same for high and low values. This is, however, in some applications a too restrictive assumption. Third, the covariance matrix – used as a parameter of the multivariate normal distribution – fails to fit heavy tails typical for asset returns. An approach which, partially solves these problems, is based on copulae, proposed by Sklar and reviewed in Joe (1997) and Nelsen (2006). Copulae allow us to model dependency separately from marginal distributions and provide better fit for heavy tails, asymmetries, etc.

Time-varying copulae were considered recently by Patton (2004), Rodriguez (2007) and Giacomini, Härdle and Spokoiny (2009). Patton (2004) considers an asset-allocation prob-

lem with a time-varying parameter of bivariate copulae. Rodriguez (2007) studies financial contagion using switching-parameter bivariate copulae. In contrary to these papers Giacomini et al. (2009) use a novel method based on local adaptive estimation discussed in Spokoiny (2010). The idea of that approach is to determine a period of homogeneity where the parameter of a low-dimensional Archimedean copula can be approximated by a constant.

The online instantaneous selection of high dimensional dependency structures via multivariate copulae is still an open problem. Here we tackle this problem via multivariate hierarchical Archimedean copulae. Detailed analysis of this copula class is given in Okhrin, Okhrin and Schmid (2009). In contrary to simple Archimedean copulae, the HAC is characterised not only by its parameters, but also by the structure. The time-varying dependency therefore affects structure and parameters simultaneously. The variability of the parameters implies that the dependency becomes stronger or weaker; variability of the structure implies that there is a change not only in the strength of the dependency, but also in its form. The proposed technique allows us to determine the periods with local constant structure and parameters. It is based on the selection of an appropriate interval out of a set of candidate intervals. The procedure requires the calculation of a sequence of critical values (by simulations) that are used in testing local homogeneity. Local homogeneity is checked via a test against a change point alternative.

To assess the performance of the developed methodology we perform extensive simulations and empirical studies. Within the simulation study we show that this novel technique quickly reacts to shifts in the structure and in the parameters. The varying estimation window allows to increase the precision of the estimators in stable periods, but simultaneously to react quickly to changes if they occur. The detection delay clearly demonstrates the effectiveness of the procedure compared to a rolling window estimation. In the empirical study we demonstrate one example with changes only in parameters and the second example with changes both in parameters and in structure.

The paper is structured as follows. In the next section we give a short theoretical background of HAC with estimation and grouping techniques. Section 3 extends the local adaptive estimation procedures to copulae. Sections 4 and 5 deal with applications to simulated and real data.

2   Hierarchical Archimedean Copulae

The advantage of the copula is that it allows to split the multivariate distribution into the margins and a pure dependency component: it captures the dependency between variables eliminating the impact of the marginal distributions. Formally copulae where introduced in Sklar (1959). The main result states that if F is an arbitrary d-dimensional continuous distribution function of the random variables $X_1, \ldots, X_d$, then the associated copula is unique and defined as a continuous function $C : [0,1]^d \to [0,1]$ which satisfies the equality

$$C(u_1, \ldots, u_d) = F\{F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)\}, \quad u_1, \ldots, u_d \in [0,1],$$

where $F_1^{-1}(\cdot), \ldots, F_d^{-1}(\cdot)$ are the quantile functions of the corresponding marginal distributions $F_1(x_1), \ldots, F_d(x_d)$. If F belongs to the class of elliptical distributions, then this results in a so called elliptical copula. Note, however, that this copula cannot be given explicitly, because the distribution function F and the inverse marginal distributions F_i usually have only integral representations. One of the classes that overcomes this drawback of elliptical copulae is the class of Archimedean copulae

$$C(u_1, \ldots, u_k) = \phi\{\phi^{-1}(u_1) + \cdots + \phi^{-1}(u_d)\}, \quad u_1, \ldots, u_d \in [0,1], \tag{1}$$

where $\phi \in \mathfrak{L} = \{\phi : [0;\infty) \to [0,1] \,|\, \phi(0) = 1,\ \phi(\infty) = 0;\ (-1)^j \phi^{(j)} \geq 0;\ j = 1, \ldots, \infty\}$. The function ϕ is called the generator of the copula and commonly depends on a single parameter θ. For example, the Gumbel generator is given by $\phi = \exp(-x^{1/\theta})$ for $0 \leq x < \infty$, $1 \leq \theta < \infty$. The generator ϕ is required to be d-monotone, i.e. differentiable up to the order $d - 2$, with $(-1)^j \phi^{(j)}(x) \geq 0$, $j = 0, \ldots, d-2$ for any $x \in [0, \infty)$ and with $(-1)^{d-2} \phi^{(d-2)}(x)$ being nondecreasing and convex on $[0, \infty)$ (e.g. McNeil and Nešlehová (2009)). For simplicity we make a stronger assumption that ϕ is a completely monotone function, i.e. $(-1)^j \phi^{(j)}(x) \geq 0$ for all $j \geq 0$. A detailed review of the properties of Archimedean copulae can be found in McNeil and Nešlehová (2009) and Joe (1996).

A disadvantage of Archimedean copulae is the fact that the rendered dependency is symmetric with respect to the permutation of variables, i.e. the distribution is exchangeable. Moreover, the multivariate dependency structure is somewhat stiff, since it depends on a single parameter of the generator function ϕ. The Hierarchical Archimedean Copulae (HAC) overcome this problem by considering the compositions of simple Archimedean
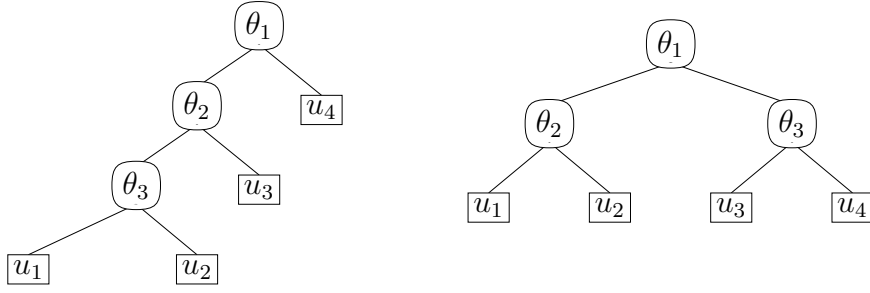
Figure 1: Fully and partially nested copulae of dimension $d = 4$ with structures $s = (((12)3)4)$ on the left and $s = ((12)(34))$ on the right

copulae. For example, the special case of HAC fully nested copula can be given by

$$
\begin{aligned}
C(u_1, \ldots, u_d) &= C_1\{C_2(u_1, \ldots, u_{d-1}), u_d\} = \phi_1\{\phi_1^{-1} \circ C_2(u_1, \ldots, u_{d-1}) + \phi_1^{-1}(u_d)\} \\
&= \phi_1\{\phi_1^{-1} \circ \phi_2(\phi_2^{-1}(C_3(u_1, \ldots, u_{d-2})) + \phi_2^{-1}(u_{d-1})) + \phi_1^{-1}(u_d)\}. \quad (2)
\end{aligned}
$$

The composition can be applied recursively using different segmentations of variables leading to more complex HACs. For notational convenience let the expression $s = \{(\ldots (i_1 \ldots i_{j_1}) \ldots (\ldots) \ldots)\}$ denote the structure of a HAC, where $i_\ell \in \{1, \ldots, d\}$ is a reordering of the indices of the variables. s_j denotes the structure of subcopulae with $s_d = s$. Further let the d-dimensional hierarchical Archimedean copula be denoted by $C(u_1, \ldots, u_d; s, \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ the set of copula parameters. For example the fully nested HAC (2) can be expressed as

$$
\begin{aligned}
C(u_1, \ldots, u_d; s = s_d, \boldsymbol{\theta}) &= C\{u_1, \ldots, u_d; ((s_{d-1})d), (\theta_1, \ldots, \theta_{d-1})^\top\} \\
&= \phi_{d-1,\theta_{d-1}}(\phi_{d-1,\theta_{d-1}}^{-1} \circ C\{u_1, \ldots, u_{d-1}; ((s_{d-2})(d-1)), (\theta_1, \ldots, \theta_{d-2})^\top\} \\
&+ \phi_{d-1,\theta_{d-1}}^{-1}(u_d)),
\end{aligned}
$$

where $s = \{(\ldots (12)3) \ldots)d)\}$. In Figure 1 we present the fully nested HAC with structure $s = (((12)3)4)$ and partially nested with $s = ((12)(34))$ in dimension $d = 4$.

HAC are thoroughly analysed in Joe (1997), Whelan (2004), Savu and Trede (2006), Embrechts, Lindskog and McNeil (2003).

Note that generators ϕ_i within a HAC can come either from a single generator family or from different generator families. If ϕ_i's belong to the same family, then the complete monotonicity of $\phi_i \circ \phi_{i+1}$ imposes some constraints on the parameters $\theta_1, \ldots, \theta_{d-1}$. Theorem 4.4 of McNeil (2008) provides sufficient conditions on the generator functions to

guarantee that C is a copula. It holds that if $\phi_i \in \mathfrak{L}$ for $i = 1, \ldots, d-1$ and $\phi_i \circ \phi_{i+1}$ has a completely monotone derivative for $i = 1, \ldots, d-2$ then C is a copula. For the majority of the generators a feasible HAC requires decreasing parameters from the lowest to the highest level. However, in the case of different families within a single HAC, the condition of complete monotonicity is not always fulfilled.

In general the structure s of the HAC can be arbitrary. On the one hand this makes it a very flexible and simultaneously parsimonious distribution model. If we use the same single-parameter generator function on each level, but with a different value of θ, we may specify the whole distribution with $d-1$ parameters. From this point of view, the HAC approach can be seen as a alternative to covariance driven models. On the other hand for each HAC not only the parameters are unknown, but also the structure has to be determined. One possible procedure is to enumerate and to estimate all possible HACs. Using a suitable goodness-of-fit test we then determine the optimal structure. This approach is however unrealistic in practice even in moderate dimensions. Okhrin, Okhrin and Schmid (2009) suggest a computationally efficient procedure, which allows to estimate the HAC recursively.

We constrain the discussion to binary copulae, i.e. at each level of the hierarchy only two variables are joined together. Joining more than two variables dramatically increases the number of formal candidate distributions and the needed computational power. At the lowest level we fit a bivariate copula to every couple of the variables. The estimation procedure is discussed below. We select the couple of variables with strongest fit and denote the respective estimator of the parameter at the first level by $\hat{\theta}_1$ and the set of indices of the variables by I_1. The selected couple is joined together to define the pseudo-variables $C\{(I_1); \hat{\theta}_1, \phi_1\}$. At the next level we proceed in the same way by considering the remaining variables and the new pseudo-variable. The considered procedure allows us to determine the estimated structure of the copula.

The multi-stage maximum-likelihood estimation is a convenient tool for recursive estimation. At the first stage we estimate the marginal distributions either parametrically or nonparametrically. At the next stage we estimate the parameter of the copula at the first level assuming that the marginal distributions are known. At further stages the next level copula parameter is estimated assuming that the margins as well as the copula parameters at lower levels are known.

Let $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)^\top$ be the parameters of copulae starting with the lowest up to the

highest level. The multistage ML estimator $\hat{\boldsymbol{\eta}}$ of $\boldsymbol{\eta} = (\boldsymbol{\alpha}^\top, \boldsymbol{\theta}^\top)^\top$ solves the system

$$\left(\frac{\partial \mathcal{L}_1}{\partial \boldsymbol{\alpha}_1^\top}, \ldots, \frac{\partial \mathcal{L}_d}{\partial \boldsymbol{\alpha}_d^\top}, \frac{\partial \mathcal{L}_{d+1}}{\partial \theta_1}, \ldots, \frac{\partial \mathcal{L}_{d+p}}{\partial \theta_p} \right)^\top = \mathbf{0}, \tag{3}$$

$$\text{where} \quad \mathcal{L}_j = \sum_{i=1}^{n} l_j(\mathbf{X}_i), \text{ for } j = 1, \ldots, d+p,$$

$$l_j(\mathbf{X}_i) = \log f_j(x_{ij}, \boldsymbol{\alpha}_j), \text{ for } j = 1, \ldots, d, i = 1, \ldots, n,$$

$$l_{j+d}(\mathbf{X}_i) = \log \left(c\big[\{F_m(x_{im}, \boldsymbol{\alpha}_m)\}_{m \in s_j}; s_j, \{\theta_\ell\}_{\ell=1,\ldots,p} \big] \prod_{m \in s_j} f_m(x_{im}, \boldsymbol{\alpha}_m) \right)$$

$$\text{for } j = 1, \ldots, p, \; i = 1, \ldots, n.$$

where $\hat{F}_i(\cdot)$ is an estimator of the marginal cdf F_i. If we estimate the margins parametrically then $\hat{F}_i(\cdot) = F_i(\cdot, \hat{\boldsymbol{\alpha}}_i)$. The marginal density $\hat{f}(\cdot)$ is estimated accordingly. Chen and Fan (2006) and Okhrin et al. (2009) provide asymptotic behavior of the estimates.

3 Inhomogeneous Dependence

Numerous models were proposed for time-varying correlation structure, with the multivariate GARCH model among the most popular. In these models the correlations are defined as functions of (lagged) explanatory variables which may influence the variation in the current dependency structure (measured via correlation). This implies that the conditional correlation changes at each moment of time, but the parameters of the conditioning functions are assumed to be constant. This approach suffers from two important drawbacks. First, the estimation of this type of the models is tedious because of the large number of parameters to be estimated. Second, there is evidence that the parameters do change with time, see e.g. Lamoureux and Lastrapes (1990). Neglecting this fact may lead to inconsistent estimators.

In order to cope with the time varying dependency structure we propose a parsimonious alternative that is based on a local constant HAC approximation. With a once and for all calculated set of critical values we determine periods of homogeneity instantaneously at each time point. Corresponding theory and applications may be found in Čížek, Härdle and Spokoiny (2009), Mercurio and Spokoiny (2004) and Chen, Härdle and Jeong (2008). The method can be virtually applied to any dependency model. However, applied to HACs, it allows us to control not only for the periods with constant parameters, but also for the periods with constant structure. Moreover, the method is applicable not

only to abrupt changes in the dependency, but also to smooth transitions in the model parameters.

Let s_t and $\boldsymbol{\theta}_t$ denote the time varying but unknown copula structure and parameters. The idea is to select for each time point t_0 an interval I at which $\boldsymbol{\theta}_t$ and s_t can be well approximated by constant $\boldsymbol{\theta}^*$ and s^*. The discrepancy between two copulae $C(\cdot; s, \boldsymbol{\theta})$ and $C(\cdot; s', \boldsymbol{\theta}')$ is measured by the Kullback-Leibler divergence $\mathcal{K}$:

$$\mathcal{K}\{C(\cdot; s', \boldsymbol{\theta}'), C(\cdot; s, \boldsymbol{\theta})\} = \mathsf{E}_{s', \boldsymbol{\theta}'} \log \frac{c(\cdot; s', \boldsymbol{\theta}')}{c(\cdot; s, \boldsymbol{\theta})}$$

where c is the copula density, see Nelsen (2006). The aim is to select I as close as possible to the so-called "oracle" choice interval. Define the "oracle" choice I_{k^*} of the interval as the largest interval $I = [t_0 - m_{k^*}; t_0]$, for which the small modelling bias condition (SMB) is fulfilled

$$\triangle_I(s, \boldsymbol{\theta}) = \sum_{t \in I} \mathcal{K}\{C(\cdot; s_t, \boldsymbol{\theta}_t), C(\cdot; s, \boldsymbol{\theta})\} \leq \triangle, \text{ for some } \triangle \geq 0, \ s, \boldsymbol{\theta}. \tag{4}$$

The unknown local constant copula parameter (at point t_0) can be best estimated on the largest interval $\operatorname{argmax}_I \triangle_I(s, \boldsymbol{\theta}) = [t_0 - m_{k^*}; t_0]$ fulfilling (4). This means, that the model s_t, $\boldsymbol{\theta}_t$ can be well approximated by locally constant $s, \boldsymbol{\theta}$ on interval I. Methods of estimation of the HAC discussed in Okhrin et al. (2009) maximise the ML with respect to the structure s and parameters $\boldsymbol{\theta}$, what leads to the best parametric fit to the underlying model on I, defined by $\tilde{s}_I$, $\tilde{\boldsymbol{\theta}}_I$. Recall that the Kullback-Leibler divergence plays a particular role in the analysis and estimation of misspecified models, see White (1982). In the case of minimising $\triangle_I(s, \boldsymbol{\theta})$ with respect to the length of the interval I, we minimise the loss caused by the ignorance of the time variation in the copula.

Note, that the true time-varying parameters $\boldsymbol{\theta}_t$ and s_t are unknown. Therefore also the "oracle" choice m_{k^*} is unknown. In a data-driven algorithm based on the Local Change Point (LCP) detection procedure, see Spokoiny (2010), we sequentially test whether $\boldsymbol{\theta}_t = \boldsymbol{\theta}^*$ and the structure of the HAC $s_t = s^*$ is constant within some interval I. Here the aim of the LCP technique is not to detect a change point, but rather to conveniently determine the period of constant dependency. Alternative techniques can be found in e.g. Čìžek et al. (2009).

The risk arising in the estimation of locally constant copulae under the SMB is bounded. Let $\mathcal{L}(s, \boldsymbol{\theta})$ denote the log-likelihood function based on the HAC with the parameters s and $\boldsymbol{\theta}$. Following Čìžek et al. (2009) let $\tilde{\boldsymbol{\theta}}_I$ and $\tilde{s}_I$ be any estimators on an interval I. If

the SMB holds for some I, s and $\boldsymbol{\theta}$ then

$$\mathsf{E}_{s_t,\boldsymbol{\theta}_t} \log \left\{ 1 + \frac{|\mathcal{L}(\tilde{s}_I, \tilde{\boldsymbol{\theta}}_I) - \mathcal{L}(s,\boldsymbol{\theta})|^r}{\mathfrak{R}_r(s,\boldsymbol{\theta})} \right\} \leq 1 + \triangle, \tag{5}$$

where $\mathfrak{R}_r(s,\boldsymbol{\theta})$ is an upper bound satisfying

$$\mathsf{E}_{s^*,\boldsymbol{\theta}^*} |\mathcal{L}(\tilde{s}_I, \tilde{\boldsymbol{\theta}}_I) - \mathcal{L}(s^*,\boldsymbol{\theta}^*)|^r \leq \mathfrak{R}_r(s^*,\boldsymbol{\theta}^*)$$

which is called a "propagation condition". We set $r = 0.5$ since this choice is also proposed in the literature. The bound given in (5) tells us that the risk in an estimated local constant model (under SMB) differs from the risk in the true constant model by an amount of $\triangle$.

The LCP is based on sequentially testing the hypotheses of homogeneity on intervals I_k. We select I_k with $k = -1, 0, 1, \ldots$ as the sequence of intervals $I_k \subset I_{k+1}$, starting with $k = 1$. If there are no change points in $\mathcal{T}_k \subset I_k \setminus I_{k-1}$ then we accept I_k as an interval with constant copula parameter and structure. At the next step we take $\mathcal{T}_{k+1}$ and test it for homogeneity. We repeat the steps until rejection or the largest possible interval I_K is accepted, leading to an interval $I_{\hat{k}}$.

Two sources of errors occur in practical applications. Let I_{k^*} denote the oracle choice. This implies that for I_k ($k < k^*$) the SMB holds. The first type of error ("false alarm") occurs if $\hat{k} < k^*$. In this case the estimation is based on a shorter data period and therefore implies higher variability. Let $\hat{s}_k$ and $\hat{\boldsymbol{\theta}}_k$ be the respective estimators and $\tilde{s}_k$ and $\tilde{\boldsymbol{\theta}}_k$ denote the corresponding fixed-sample estimators on I_k. Under the SMB condition on I_{k^*} and assuming that $\max_{k \leq k^*} \mathsf{E}_{s,\boldsymbol{\theta}} |\mathcal{L}(\tilde{s}_k, \tilde{\boldsymbol{\theta}}_k) - \mathcal{L}(s,\boldsymbol{\theta})|^r \leq \mathfrak{R}_r(s,\boldsymbol{\theta})$, we obtain by Theorem 4.2 of Čížek et al. (2009)

$$\mathsf{E}_{s_t,\boldsymbol{\theta}_t} \log \left\{ 1 + \frac{|\mathcal{L}(\tilde{s}_{\hat{k}}, \tilde{\boldsymbol{\theta}}_{\hat{k}}) - \mathcal{L}(s,\boldsymbol{\theta})|^r}{\mathfrak{R}_r(s,\boldsymbol{\theta})} \right\} \leq 1 + \triangle, \tag{6}$$

$$\mathsf{E}_{s_t,\boldsymbol{\theta}_t} \log \left\{ 1 + \frac{|\mathcal{L}(\tilde{s}_{\hat{k}}, \tilde{\boldsymbol{\theta}}_{\hat{k}}) - \mathcal{L}(\hat{s}_{\hat{k}}, \hat{\boldsymbol{\theta}}_{\hat{k}})|^r}{\mathfrak{R}_r(s,\boldsymbol{\theta})} \right\} \leq 1 + \triangle.$$

The inequalities (6) say, that if we observe a false alarm at the step $\hat{k} < k^*$, then the estimation risk measured by $|\mathcal{L}(\tilde{s}_{\hat{k}}, \tilde{\boldsymbol{\theta}}_{\hat{k}}) - \mathcal{L}(\hat{s}_{\hat{k}}, \hat{\boldsymbol{\theta}}_{\hat{k}})|^r$ is of the same order as the risk of a pure parametric estimation with fixed interval given by $I_{\hat{k}}$.

The second type of the error arises if $\hat{k} > k^*$. Outside the oracle interval we are exploiting data which does not support the SMB condition. This implies that the bounds in (6)

9

increase. Theorem 4.3 of Čížek et al. (2009) provides general bounds for the adaptive estimator, showing that

$$
\mathsf{E}_{s_t, \boldsymbol{\theta}_t} \log \left\{ 1 + \frac{|\mathcal{L}(\tilde{s}_{k^*}, \tilde{\boldsymbol{\theta}}_{k^*}) - \mathcal{L}(\hat{s}_{\hat{k}}, \hat{\boldsymbol{\theta}}_{\hat{k}})|^r}{\mathfrak{R}_r(s, \boldsymbol{\theta})} \right\} \leq 1 + \triangle + \log \left\{ 1 + \frac{\mathfrak{z}_{k^*}^r}{\mathfrak{R}_r(s, \boldsymbol{\theta})} \right\}, \tag{7}
$$

where $\mathfrak{z}_{\hat{k}}$ are the critical values of the test for homogeneity and are defined below. This statement implies that the copula based on $\hat{s}_{\hat{k}}$ and $\hat{\boldsymbol{\theta}}_{\hat{k}}$ belongs with high probability to the confidence interval of the oracle copula with $\tilde{s}_{k^*}$ and $\tilde{\boldsymbol{\theta}}_{k^*}$.

3.1 Local test of homogeneity

A local homogeneity test can now be performed. Let us fix some t_0 and let $I = [t_0 - m, t_0]$ be an interval candidate and $\mathcal{T}_I$ be a set of interval points within I. We estimate the copula parameter $\boldsymbol{\theta}$ and the structure s from observations in I, assuming the homogeneous model within I, i.e. using the notation from the previous section $\hat{\boldsymbol{\theta}}_{t_0} = \tilde{\boldsymbol{\theta}}_I$ and $\hat{s}_{t_0} = \tilde{s}_I$. The null hypothesis H_0 means that $\forall \tau \in \mathcal{T}_I : \boldsymbol{\theta}_\tau = \boldsymbol{\theta}, \ s_\tau = s$ i.e. the observations in I follow the model with the dependence parameter $\boldsymbol{\theta}$ and the structure s. The alternative (change point) hypothesis H_1 claims that $\exists \tau \in \mathcal{T}_I : \boldsymbol{\theta}_t = \boldsymbol{\theta}_1$ and $s_t = s_1$ for $t \in J = [\tau, t_0]$ and $\boldsymbol{\theta}_t = \boldsymbol{\theta}_2 \neq \boldsymbol{\theta}_1$ or $s_t = s_2 \neq s_1$ for $t \in J^c = [t_0 - m, \tau)$, i.e. either the parameter $\boldsymbol{\theta}$ or the whole structure s changed spontaneously at some intermediate point τ of the interval I. In other words

$$
\begin{aligned}
H_0 \ &: \ \forall \tau \in \mathcal{T}_I, \ \boldsymbol{\theta}_t = \boldsymbol{\theta}, \ s_t = s, \ \forall t \in I = J \cup J^c = [\tau, t_0] \cup [t_0 - m, \tau) \\
H_1 \ &: \ \exists \tau \in \mathcal{T}_I, \ \boldsymbol{\theta}_t = \boldsymbol{\theta}_1, \ s_t = s_1; \ \forall t \in J = [\tau, t_0], \\
&\qquad \text{and } \boldsymbol{\theta}_t = \boldsymbol{\theta}_2 \neq \boldsymbol{\theta}_1 \text{ or } \ s_t = s_2 \neq s_1, \forall t \in J^c = [t_0 - m, \tau).
\end{aligned}
$$

If $\mathcal{L}_I(s, \boldsymbol{\theta})$ and $\mathcal{L}_J(s_1, \boldsymbol{\theta}_1) + \mathcal{L}_{J^c}(s_2, \boldsymbol{\theta}_2)$ are the log-likelihood functions corresponding to H_0 and H_1 respectively, the likelihood ratio test for the single change point with known fixed location τ is given by

$$
\begin{aligned}
T_{I,\tau} \ &= \ \max_{s_1, \boldsymbol{\theta}_1, s_2, \boldsymbol{\theta}_2} \{\mathcal{L}_J(s_1, \boldsymbol{\theta}_1) + \mathcal{L}_{J^c}(s_2, \boldsymbol{\theta}_2)\} - \max_{s, \boldsymbol{\theta}} \mathcal{L}_I(s, \boldsymbol{\theta}) \\
&= \ ML_J + ML_{J^c} - ML_I.
\end{aligned}
$$

Since the point τ is unknown, one defines the test statistics:

$$
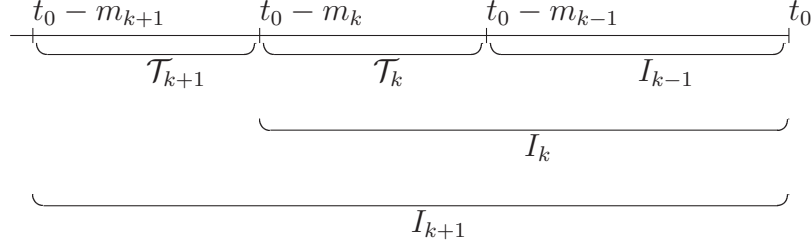T_I = \max_{\tau \in \mathcal{T}_I} T_{I,\tau}.
$$

Figure 2: Interval Selection

It tests the homogeneity hypothesis in I against a change point alternative with unknown location τ (in the set $\mathcal{T}_I$). The decision rule of the test requires to compare T_I with the critical value $\mathfrak{z}_I$. The critical value depends on the interval I, the dimension and the parameter of the copula. We reject the hypothesis of homogeneity if $T_I > \mathfrak{z}_I$. To run the test several parameters have to be specified. This includes the choice of the interval candidates (I_k) and internal points $\mathcal{T}_k = \mathcal{T}_{I_k}$ for each of this intervals and the choice of the critical values $\mathfrak{z}_k = \mathfrak{z}_{I_k}$. One possible example of the implementation is based on the choice of the interval candidates (I_k) in form of a geometric grid. If the length of the interval I_1 if fixed at m_1, then we define $m_0 = a_1 m_1$ and $m_{-1} = a_2 m_1$ for $a_1 > a_2 \in (0, 1)$ and $m_k = [m_1 c^{k-1}]$ for $k = 1, 2, \ldots, K$ and $c > 1$, where $[x]$ means the integer part of x. Further we set $I_k = [t_0 - m_k, t_0]$ and $\mathcal{T}_k = [t_0 - m_{k-1}, t_0 - m_{k-2}]$ for $k = 1, 2, \ldots, K$, see Figure 2.

For each particular copula model and for each sequence of intervals the critical values $\mathfrak{z}_k$ are determined from simulations. Under the null hypothesis of homogeneous dependence, the adaptive estimator should coincide with the largest allowed interval I_K. However, if the estimated interval is $I_{\hat{k}}$ with $\hat{k} < K$ then the test rejects a correct null hypothesis. The critical values are therefore determined not from the classical level condition, but relying on the precision of the parameter estimators. If $\hat{k}$ is small the volatility of the parameter estimator is high. This implies that false decisions with small $\hat{k}$ stronger deteriorate the impact of the test of homogeneity. To overcome this problem we select the critical values $\mathfrak{z}_k$ to ensure that

$$\mathsf{E}_{s^*, \theta^*} |\mathcal{L}(\tilde{s}_k, \tilde{\boldsymbol{\theta}}_k) - \mathcal{L}(\hat{s}_k, \hat{\boldsymbol{\theta}}_k)|^r \leq \rho_k \mathfrak{R}_r(s^*, \boldsymbol{\theta}^*),$$

where $\rho_k = \rho\, k/K \leq \rho$ and $\mathfrak{R}_r(s^*, \boldsymbol{\theta}^*) = \max_k |\mathcal{L}(\tilde{s}_k, \tilde{\boldsymbol{\theta}}_k) - \mathcal{L}(s^*, \boldsymbol{\theta}^*)|^r$. The parameter ρ plays the role of the level of significance and influences the sensitivity of the procedure to inhomogeneity. This means, that large values of ρ lead to larger critical values, and small ρ to smaller.

In this paper we used the sequential choice of critical values $\mathfrak{z}_k$ discussed in Spokoiny (2009). Assuming situation after k steps of the algorithm we distinguish between two cases. In the first, change point is detected at some step $\ell \leq k$ and in the other case no change point is detected. Following notation in Spokoiny (2009), let $\mathcal{B}_\ell$ be the event meaning the rejection of the null hypothesis at step ℓ

$$\mathcal{B}_\ell = \{T_1 \leq \mathfrak{z}_1, \ldots, T_{\ell-1} \leq \mathfrak{z}_{\ell-1}, T_\ell > \mathfrak{z}_\ell\},$$

and $(\hat{s}_k, \hat{\boldsymbol{\theta}}_k) = (\tilde{s}_{\ell-1}, \tilde{\boldsymbol{\theta}}_{\ell-1})$ on $\mathcal{B}_\ell$ for $\ell = 1, \ldots, k$. By Monte-Carlo simulations from some fixed parametric models discussed in Section 4 we found sequentially such a minimal value of $\mathfrak{z}_l$ that ensures the following inequality

$$\max_{k=l,\ldots,K} \mathsf{E}_{s^*,\theta^*} |\mathcal{L}(\tilde{s}_k, \tilde{\boldsymbol{\theta}}_k) - \mathcal{L}(\tilde{s}_{\ell-1}, \tilde{\boldsymbol{\theta}}_{\ell-1})|^r \mathbf{I}(\mathcal{B}_\ell) \leq \rho \mathfrak{R}_r(s^*, \boldsymbol{\theta}^*)k/K,$$

where $\mathbf{I}$ is the indicator function. For $\ell = 1$ this inequality depends only on $\mathfrak{z}_1$ in $\mathcal{B}_1 = \{T_1 > \mathfrak{z}_1\}$. For every $\ell \geq 2$ we take $\mathfrak{z}_1, \ldots, \mathfrak{z}_{\ell-1}$ being fixed from previous step, what means, that $\mathcal{B}_\ell$ is controlled only by $\mathfrak{z}_\ell$. r is fixed throughout the study and equals $r = 0.5$.

4 Simulation Study

How fast reacts the LCP to shifts in the parameters and/or in the structure? We consider a 3-dimensional HAC with Gumbel generators and uniform margins. To simulate from a HAC we used the algorithm of McNeil (2008). We consider samples of size 400, where a change in parameters and structure occurs at $t = 200$. The parameter changes are modeled as:

$$C_t(u_1, u_2, u_3; s, \boldsymbol{\theta}) = \begin{cases} C\{u_1, C(u_2, u_3; \theta_1 = 3.33); \theta_2 = 1.43\} & \text{for } 1 \leq t \leq 200, \\ C\{u_1, C(u_2, u_3; \theta_1 = 2.00); \theta_2 = 1.43\} & \text{for } 200 < t \leq 400; \end{cases}$$

(8)

$$C_t(u_1, u_2, u_3; s, \boldsymbol{\theta}) = \begin{cases} C\{u_1, C(u_2, u_3; \theta_1 = 3.33); \theta_2 = 1.43\} & \text{for } 1 \leq t \leq 200, \\ C\{u_1, C(u_2, u_3; \theta_1 = 3.33); \theta_2 = 2.00\} & \text{for } 200 < t \leq 400. \end{cases}$$

(9)

Via model (8) we investigate the sensitivity of a downward jump in θ_1, while (9) is designed for study of an upward jump in θ_2. The initial parameters $\theta_1 = 3.33$ and $\theta_2 = 1.43$ correspond to the Kendall τ's equal to 0.7 and 0.3 respectively. In (8) τ_1 decreases to 0.5
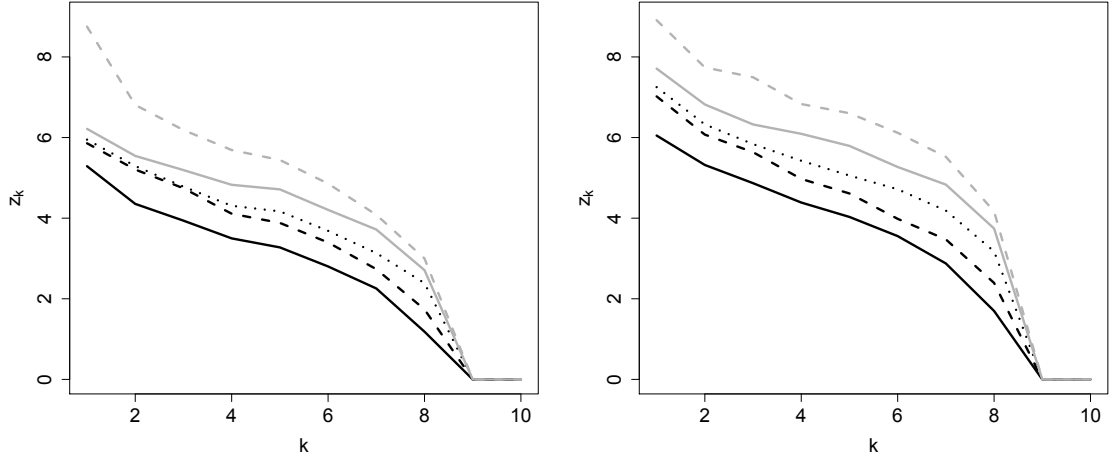
Figure 3: Simulated critical values of the 3-dimensional HAC as a function of k with the parameters of the geometric grid set to $m_0 = 20$ (left) and $m_0 = 40$ (right). ρ and τ_1 are fixed and equal 0.5 and 0.1 respectively, while τ_2 varies with $\tau_2 = 0.1$ (solid black), $\tau_2 = 0.3$ (dashed black), $\tau_2 = 0.5$ (dotted black), $\tau_2 = 0.7$ (solid grey), $\tau_2 = 0.9$ (dashed grey).

from 0.7, while in (9) τ_2 increases to 0.5 from 0.3. Note, that in both cases the difference between the parameters becomes smaller.

The change point in the structure is modeled in a similar way

$$C_t(u_1, u_2, u_3; s, \boldsymbol{\theta}) = \begin{cases} C\{u_1, C(u_2, u_3; \theta_1 = 3.33); \theta_2 = 1.43\} & \text{for } 1 \leq t \leq 200, \\ C\{C(u_1, u_2; \theta_1 = 3.33), u_3; \theta_2 = 1.43\} & \text{for } 200 < t \leq 400. \end{cases} \quad (10)$$

Our technique is implemented with a family of interval candidates of a geometric grid form defined by $m_0 = 20$, 40 and $c = 1.25$. The values of m_0 and c have turned out to provide stable results, which is confirmed in the literature cited earlier. Note that the simulated critical values are indifferent to the form of the initial structure $s_1 = ((12)3)$ or $s_2 = (1(23))$, but depend on the parameters. Using the fact that for the Gumbel copula the parameter $\theta \in [1; \infty)$ is unbounded from above, we define the grid based on the Kendall's τ by

$$\boldsymbol{\theta} = (\theta_1, \theta_2)^\top = \{\theta(\tau_1), \theta(\tau_2)\}^\top,$$

where

$$\{\tau_1, \tau_2\} \in \{0.1, 0.3, 0.5, 0.7, 0.9\}^2, \ \tau_1 \geq \tau_2.$$

This grid in correlation space corresponds to the grid in parameter space given by $\{\theta_1, \theta_2\} \in \{1.11, 1.43, 2, 3.33, 10\}^2$, $\theta_1 \geq \theta_2$. Thus, we simulate from copulae $C\{u_1, C(u_2, u_3; \theta_1 = $

3.33); $\theta_2 = 1.43$}, $C\{u_1, C(u_2, u_3; \theta_1 = 2.00); \theta_2 = 1.43\}$, etc. The case $\theta_1 = \theta_2$ corresponds to the simple 3-dimensional Archimedean copula $C(u_1, u_2, u_3, \theta_1)$. To estimate $\mathfrak{z}_k$ $k = 1, \ldots, K = 10$ we simulate $N = 10000$ samples of size $n = [m_0 c^K] + 1$ using the same geometric grid of the intervals. Figure 3 shows the behavior of the critical values as a function k for different values of τ_2.

In each change point model we simulate $n = [m_0 c^K] + 400$ observations, where the first $[m_0 c^K]$ values are used as a prerun for model estimation. For each $t = t_0$ starting from $[m_0 c^K] + 1$ we apply the LCP to the recent observations, i.e. we determine the interval with constant dependency and estimate the corresponding HAC. The results are shown in Figures 5, 6 and 7. m_0 is set to 20 in the left column and to 40 in the rights column, ρ is set to 0.5. The solid lines show the average values, the dashed line the median values and the grey areas show the interval containing 95 of 100 replications.

The shifts in the first parameter for (8) and in the second parameter for (9) are plotted in the first rows of Figures 5 and 6 respectively. The Figure 7 illustrates the application of LCP to the change-point model (10), where in the first row we show the changes in the structure and in the second row the changes in the parameters. For all three types of the shift, we observe that the average estimated parameter or structure smoothly moves from the value before the shift to the value after the shift. The delay reaction naturally depends on m_0. Smaller m_0 value let our procedure react more quickly to changes. On the other hand the precision of the estimation decreases with decreasing sample size.

The last two rows of all three figures show the dynamics of the average length of the estimated interval of homogeneity and the behavior of the maximum-likelihood. The estimation is initiated with the shortest available interval of homogeneity. Since the copula is stable and more observations become available, the length of the interval increases to the largest allowed value. After the shift the length of the interval decreases and increases only after the change-point leaves the smallest allowed interval.

The detection ability of the proposed procedure is conveniently characterized by the detection delay. Denote by γ_i the size of the jump at time $t = 200$, i.e. $\gamma_i = \theta_{it} - \theta_{i,t-1}$ with $i = 1$ for the model (8) and $i = 2$ for the model (9) The detection delay δ at rule $\mathfrak{r} \in [0, 1]$ is defined by

$$\delta(t, \gamma_i, \mathfrak{r}) = \min\{u \geq t : \hat{\theta}_{iu} \geq \theta_{i,t-1} + \mathfrak{r}\gamma_i\} - t$$

and shows the number of steps needed to detect the fraction $\mathfrak{r}$ of the jump in the true θ. For the model (10) we just look for the time point after $t = 200$ where the structure

14

model	$\mathfrak{r}$	$m_0 = 20$					$m_0 = 40$				
		Q1	Med.	Mean	Q3	SD	Q1	Med.	Mean	Q3	SD
	0.25	0.00	9.0	13.20	20.00	14.23	8.00	28.5	30.87	50.25	24.14
(8)	0.50	8.75	25.0	25.02	35.00	19.41	36.75	50.0	52.58	63.25	26.47
	0.75	19.00	31.0	35.19	47.00	25.77	50.00	63.5	72.71	87.00	35.18
	0.25	2.00	8.0	11.74	19.25	12.00	8.00	21.0	22.65	31.25	18.59
(9)	0.50	12.00	18.0	20.86	28.00	13.52	28.00	37.0	39.83	46.25	20.09
	0.75	16.75	27.0	30.75	39.00	22.51	37.00	52.5	59.53	73.75	31.83
(10)		15.00	18.0	17.78	21.00	5.23	28.00	32.0	32.45	37.00	5.90

Table 1: Detection delay statistics for LCP, $\rho = 0.5$.

model	$\mathfrak{r}$	rolling window				
		Q1	Med.	Mean	Q3	SD
	0.25	6.25	35.5	39.57	64.50	34.52
(8)	0.50	51.00	75.5	76.64	103.00	38.16
	0.75	85.25	113.0	109.70	128.20	35.90
	0.25	10.00	26.0	26.69	40.25	19.12
(9)	0.50	36.75	57.5	55.90	74.75	24.38
	0.75	77.00	95.5	95.53	112.50	29.82
(10)	0.25	68.00	75.5	75.53	84.25	11.40

Table 2: Detection delay statistics for rolling window.

$s_2 = (1(23))$ is obtained for the first time

$$\delta(t) = \min\{u \geq t : \hat{s}_u = s_{t-1}\} - t.$$

Spokoiny (2010) argued that the detection delays are proportional to the probability of the error of type II. Table 1 represents the descriptive statistics of the detection delay for different models (8), (9) and (10), $\mathfrak{r} \in \{0.25, 0.5, 0.75\}$ and $m_0 \in \{20, 40\}$. To detect half of the shift in the parameters, the procedure needs 20 to 25 observations for $m_0 = 20$ and 40 to 50 observations for $m_0 = 40$. The detection ability of the procedure for the changes in the structure is similar. Additionally we observe that the mean of detection delay is higher for upward jumps than for downward jumps. The mathematical reason for this is explained below. Table 2 contains the detection delays for the rolling window estimation. To make the comparison fair we set the length of the estimation window equal to the average length of the intervals of homogeneity in the LCP procedure. We observe that the flexible choice of the interval of homogeneity leads to substantially shorter detection delays, compared to the rolling estimation.

To get more insight into the detection delay we consider the difference

$$\mathcal{K}[C\{s_0; \theta(0.1, 0.2)^\top\}, C\{s_0; \theta(\tau_1, \tau_2)^\top\}] - \mathcal{K}[C\{s_0; \theta(\tau_1, \tau_2)^\top\}, C\{s_0; \theta(0.1, 0.2)^\top\}], \quad (11)$$

where $\theta(\tau_1, \tau_2)$ denotes the vector of parameters corresponding to the Kendall τ's given by τ_1 and τ_2. The first term in (11) denotes the KL divergence between the true copula with $\theta(\tau_1, \tau_2)$ and the misspecified copula with the same structure s_0 but with $\theta(0.1, 0.2)$. τ_1 and τ_2 take values between zero and one. Thus we observe in general an increase in the parameters from the true values $\theta(0.1; 0.2)$ to the misspecified values $\theta(\tau_1; \tau_2)$. In the second term in (11) the situation is opposite and we observe a decrease of the copula parameters from the true values $\theta(\tau_1, \tau_2)$ to the misspecified values $\theta(0.1, 0.2)$. The difference in (11) is plotted in Figure 4. The KL divergence is larger for increasing parameters and the difference becomes larger with increasing τ_1 and τ_2. This explains why the adaptive detection procedure based on the KL divergence reacts faster to an increase in parameters than to a decrease.
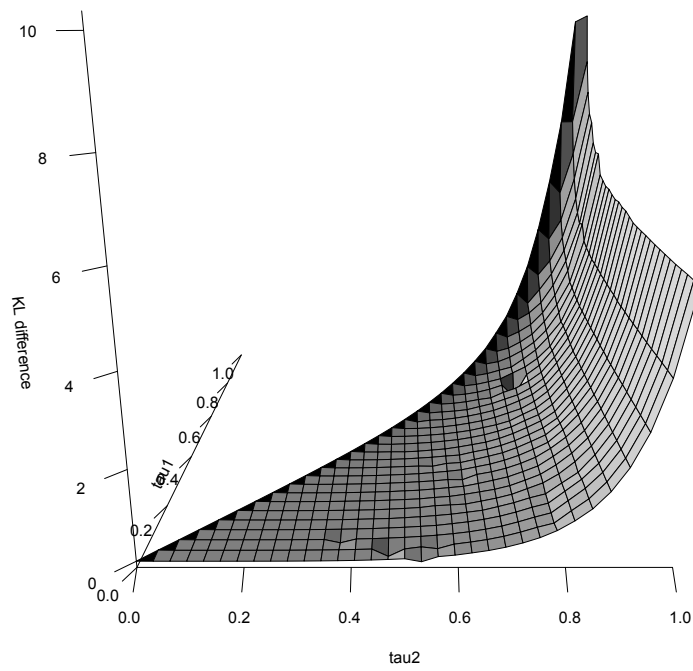
Figure 4: The difference between the KL divergences for misspecified models in (11).
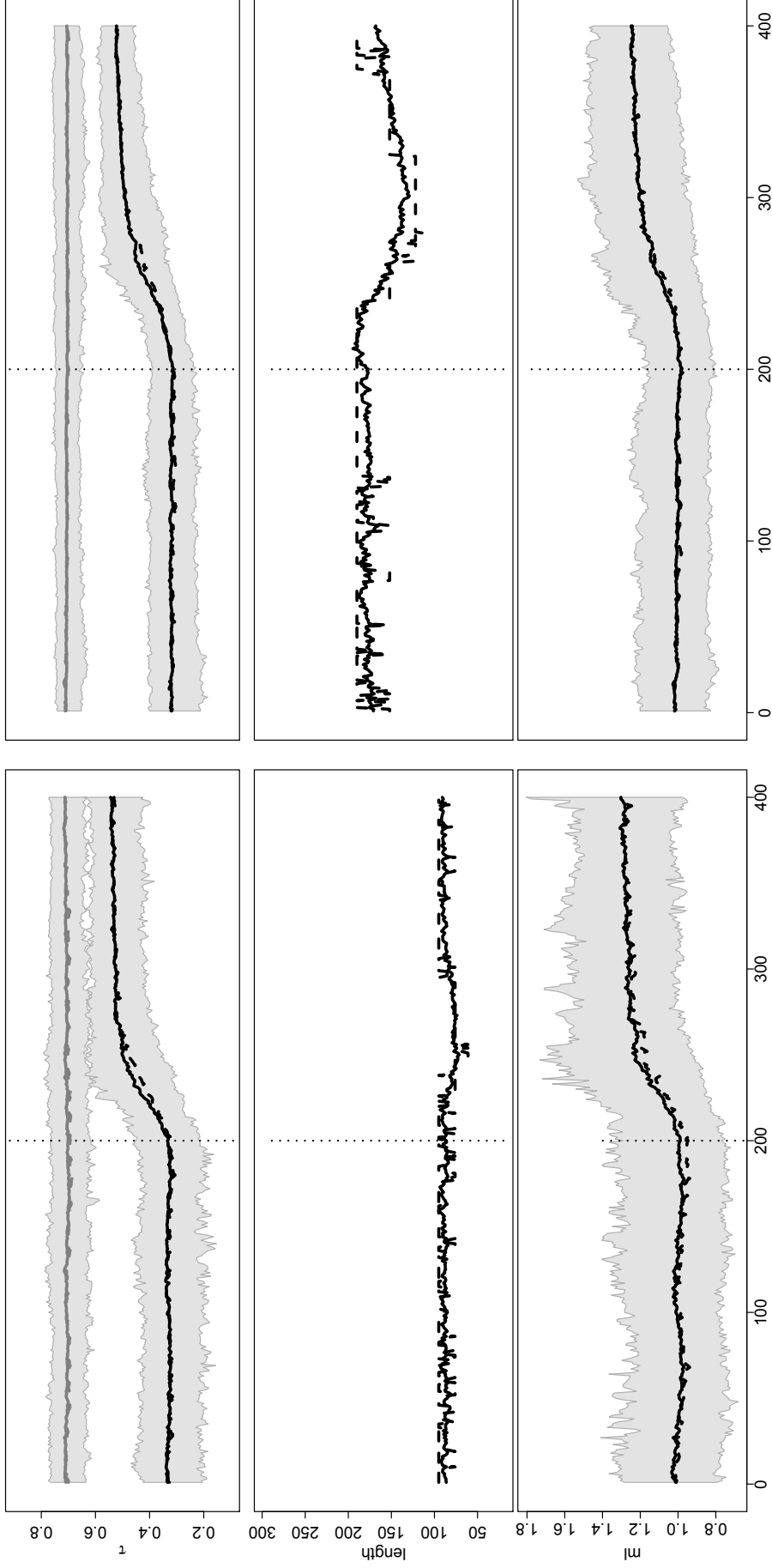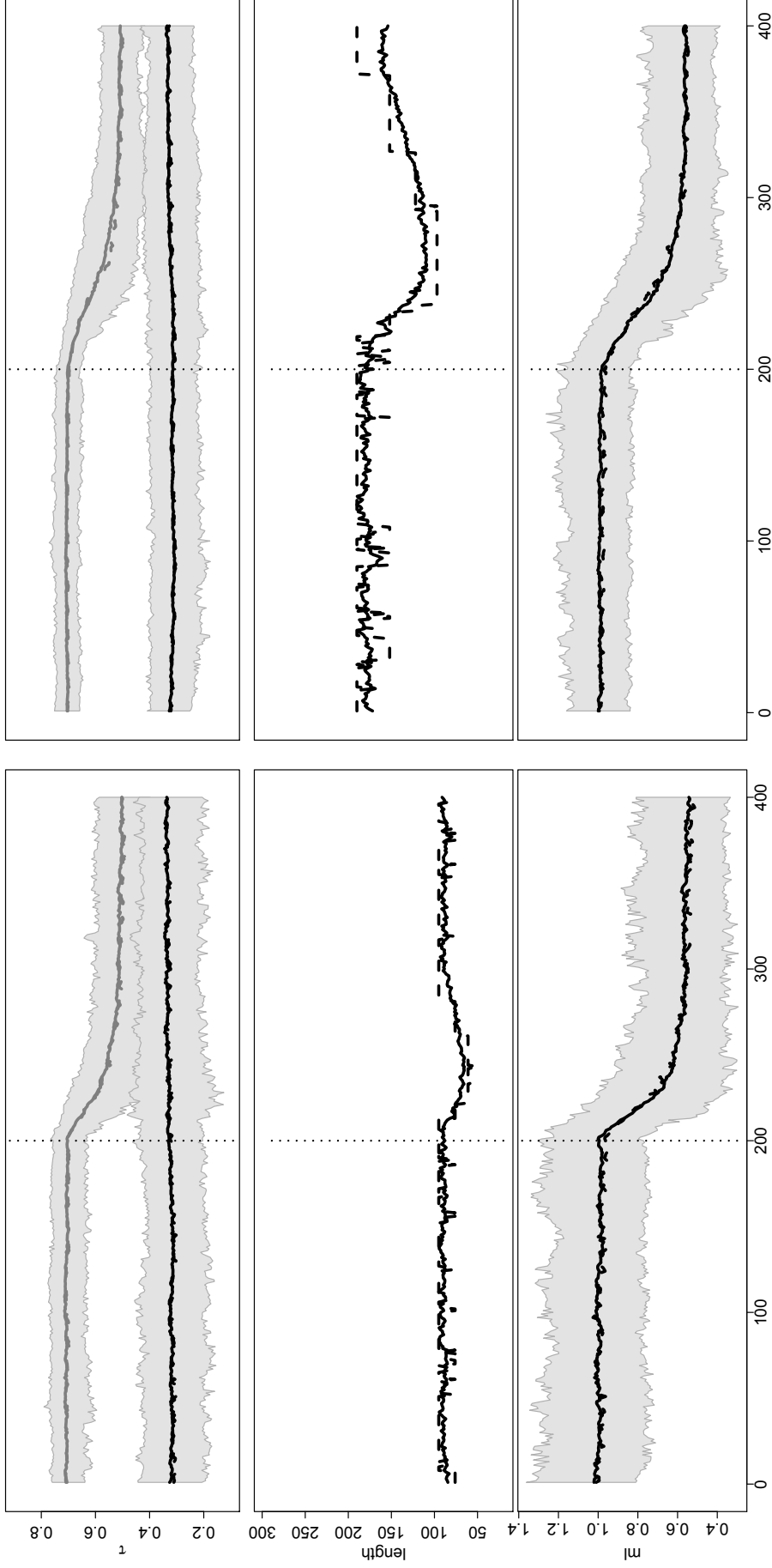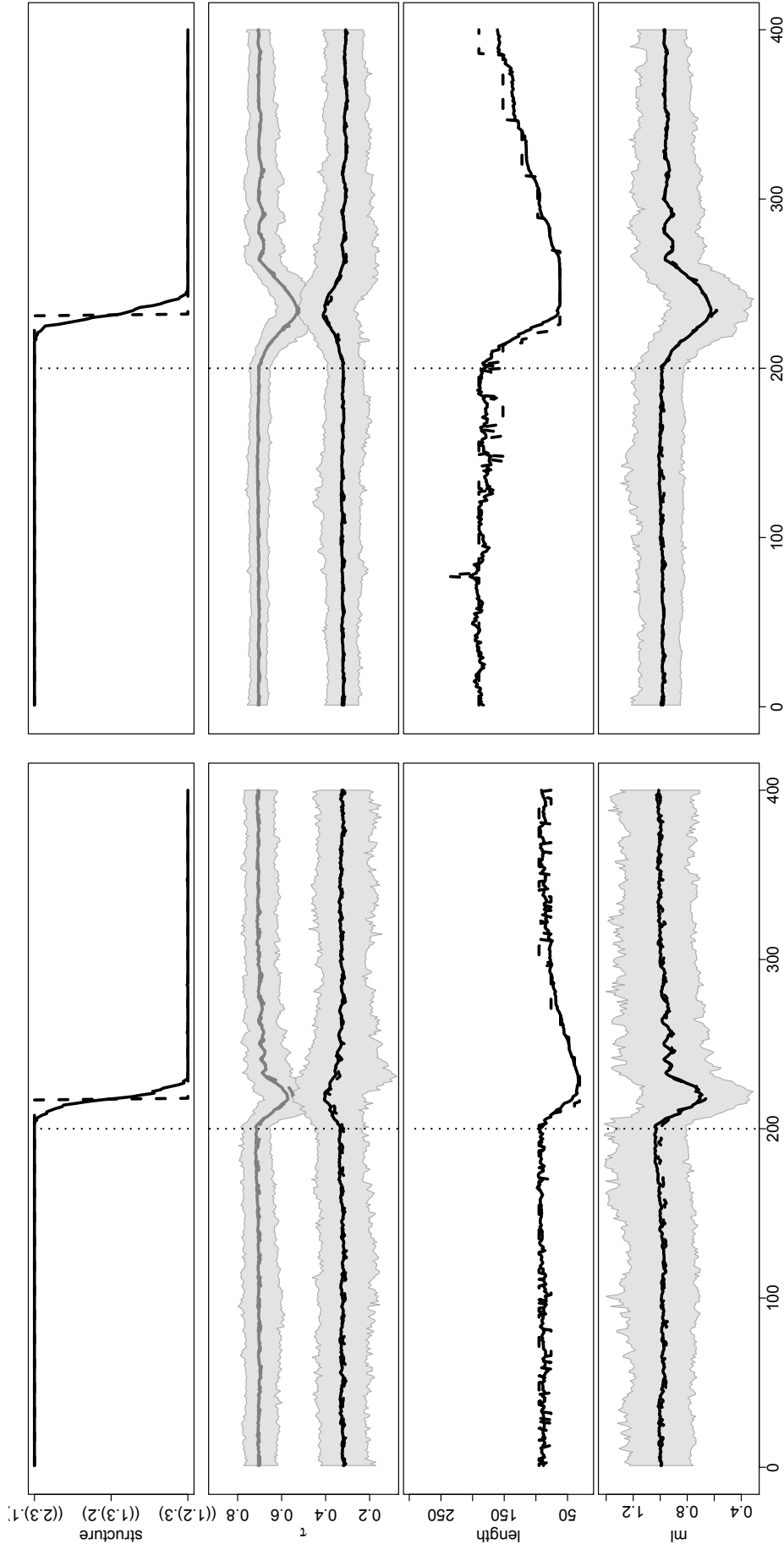
Figure 5: The average estimated parameters, the length of the interval of homogeneity and the ML criteria on the intervals of homogeneity for the model (8). The solid line depicts the average, the dashed line the median and grey shaded are the 90% confidence interval for the respective variables over 100 replications. The shift occurs at the observation 200. m_0 is set to 20 for the left panel and to 40 for the right panel, $\rho = 0.5$.

Figure 6: The average estimated parameters, the length of the interval of homogeneity and the ML criteria on the intervals of homogeneity for the model (9). The solid line depicts the average, the dashed line the median and grey shaded are the 90% confidence interval for the respective variables over 100 replications. The shift occurs at the observation 200. m_0 is set to 20 for the left panel and to 40 for the right panel, $\rho = 0.5$.

Figure 7: The average estimated structure, parameters, the length of the interval of homogeneity and the ML criteria on the intervals of homogeneity for the model (10). The solid line depicts the average, the dashed line the median and grey shaded are the 90% confidence interval for the respective variables over 100 replications. The shift occurs at the observation 200. m_0 is set to 20 for the left panel and to 40 for the right panel, $\rho = 0.5$.

5 Empirical Study

We now apply the developed local estimation procedure to multivariate data on stock indices and exchange rates. The data on indices contains the daily returns values for Dow Jones (DJ), DAX and NIKKEI, while the second data set consists of the daily values for the exchange rates JPN/EUR, GBP/EUR and USD/EUR. Both data sets are taken from `DataStream`. The indices are obtained for the period $[31.12.1986; 14.09.2009]$ resulting in 5903 observations, where exchange rates cover the period $[4.1.1999; 14.8.2009]$, resulting in 2771 observations.

To eliminate the intertemporal conditional heteroscedasticity we fit to each marginal time series of log-returns a univariate GARCH(1,1) process

$$X_{j,t} = \mu_{j,t} + \sigma_{j,t}\varepsilon_{j,t} \text{ with } \sigma_{j,t}^2 = \omega_j + \alpha_j \sigma_{j,t-1}^2 + \beta_j(X_{j,t-1} - \mu_{j,t-1})^2 \tag{12}$$

and $\omega > 0$, $\alpha_j \geq 0$, $\beta_j \geq 0$, $\alpha_j + \beta_j < 1$. The estimates of the parameters are given in Table 3. The parameters in (12) are significant. The table contains the p-values of the Box-Ljung test with 12 lags and Kolmogorov-Smirnov test applied to the residuals. The residuals exhibit the typical behavior: they are not normally distributed, which motivates nonparametric estimation of the margins. From the results of the Box-Ljung test we conclude that the autocorrelation of the residuals is strongly significant only for GBP/EUR rate and DJ. After this intertemporal correction we work only with the residuals.

5.1 Rolling Window Estimation

The dependency variation is measured by Kendall and Pearson's correlation coefficients: Figure 8 shows the behavior both coefficients calculated in a rolling window of width $r = 250$. Their dynamic behavior is similar, but not identical. This opens the door for a time varying copula based model. A key difference between the plots for indices and the exchange rates, it that the correlation paths do not intersect for the indices, but do intersect for the exchange rates. This implies that there is a change in the order of the bivariate dependencies at the intersection points of the lines for the exchange rates. Thus we may expect that the structure of HAC stays constant for the indices, but varies for the exchange rates.

To give a justification for the use of a copula-based distribution to model the residuals, we estimate additionally alternative parametric models using rolling window of the same

	$\hat{\mu}_j$	$\hat{\omega}_j$	$\hat{\alpha}_j$	$\hat{\beta}_j$	BL	KS
DAX	6.94e-04	4.17e-06	0.11	0.87	0.23	3.35e-05
	(1.39e-04)	(5.29e-07)	(8.90e-03)	(9.39e-03)		
DJ	5.96e-04	3.09e-06	0.11	0.87	0.02	1.58e-07
	(1.11e-04)	(3.38e-07)	(8.50e-03)	(9.40e-03)		
NIKKEI	5.62e-04	3.01e-06	0.11	0.88	0.78	2.45e-13
	(1.45e-04)	(5.18e-07)	(8.90e-03)	(8.71e-03)		
JPY	4.85e-05	2.99e-07	0.06	0.94	0.73	1.70e-05
	(1.15e-04)	(1.02e-07)	(7.49e-03)	(7.64e-03)		
GBP	6.34e-05	1.44e-07	0.06	0.93	0.01	2.10e-04
	(7.39e-05)	(5.11e-08)	(8.75e-03)	(9.12e-03)		
USD	1.76e-04	1.19e-07	0.03	0.97	0.87	1.65e-03
	(1.10e-04)	(5.92e-08)	(4.14e-03)	(4.28e-03)		

Table 3: Results of the fitting of univariate GARCH(1,1) to exchange rates and to indices. The standard deviation of the parameters are given in parentheses. The last two columns provide the p-values of the Box-Ljung test (BL) for autocorrelations with 12 lags and Kolmogorov-Smirnov test (KS) for normality applied to the residuals.

width. We consider the binary HAC with Gumbel generator; the 3-dimensional Gaussian and 3-dimensional simple Archimedean copula. The maximum-likelihood (ML) and the Bayes information criterion are used as goodness-of-fit measures.

The BIC criterion is computed by $BIC = -2ML + 2\log(m)$, where m is the number of the parameters to be estimated and ML is the corresponding maximum-likelihood criteria. Since the number of unknown parameters in the nonparametric case is unknown, it is incorrect to compare the models with nonparametrically and parametrically estimated margins using BIC. In such cases we consider only the parameters of the copula function. Figure 9 illustrates the dynamics of BIC for three multivariate models.

We verify if the variation in the dependency can be linked to some characteristics of the distribution. The dots in Figure 9 depict the time-points of changes in the bivariate HAC estimated using rolling window procedure. There is no visible relationship between the dynamics of the model fit measured by BIC and the changes in the structures. The thin grey line shows the dynamics of the $||\widehat{\Theta}_t - \widehat{\Theta}_{t-1}||_2$, where $\widehat{\Theta}_t$ denotes the matrix of copula parameters estimated at the time point t and $|| \cdot ||_2$ denotes the L_2 matrix norm. It is defined as $||\mathbf{A}||_2 = \sqrt{\lambda_{max}(\mathbf{A}^\top \mathbf{A})}$, where λ_{max} is the largest eigenvalue of the matrix
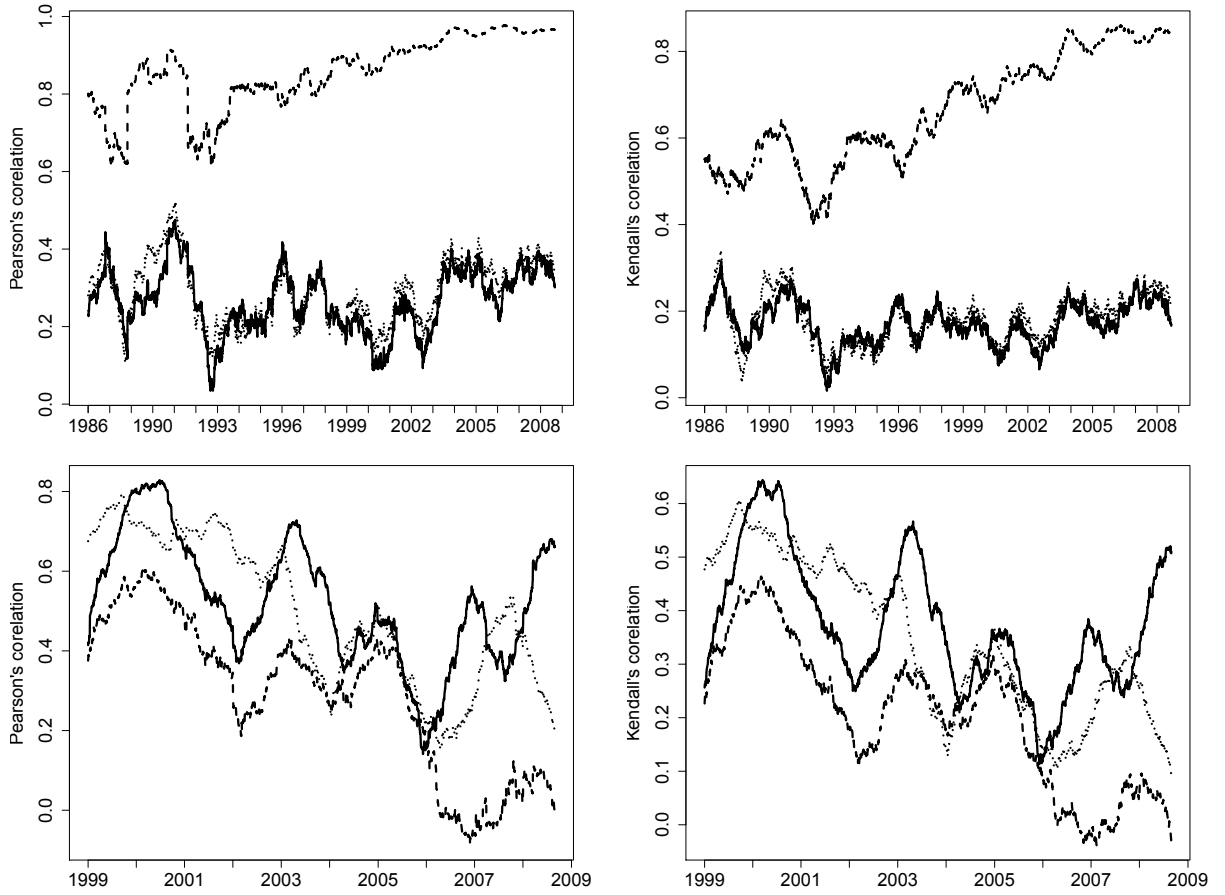
Figure 8: Rolling window estimators of Pearson's (left) and Kendall's (right) correlation coefficients between the GARCH(1,1) residuals of indices (top) and exchange rates (bottom). For the indices: DAX and NIKKEI (solid line), DAX and DJ (dashed line), DJ and NIKKEI (dotted line). For the exchange rates: JPY and USD (solid line), JPY and GBP (dashed line), GBP and USD (dotted line). The width of the rolling window is set to 250 observations.

$\mathbf{A}^{\top}\mathbf{A}$. Similarly as for BIC, there is no clear relationship between the changes in the structure and the variation in copula parameters. This implies that there is no obvious way how can we exploit the results from rolling window estimation to determine the intervals with homogeneous dependency.

5.2   Local window estimation

The previous section provided evidence on two important issues. First, the univariate marginal distributions are not normal and the joint distribution can be better modelled using a HAC based distribution. Second, the dependency is not constant and varies with time. Since we model the dependency by HAC, this implies that either the structure of
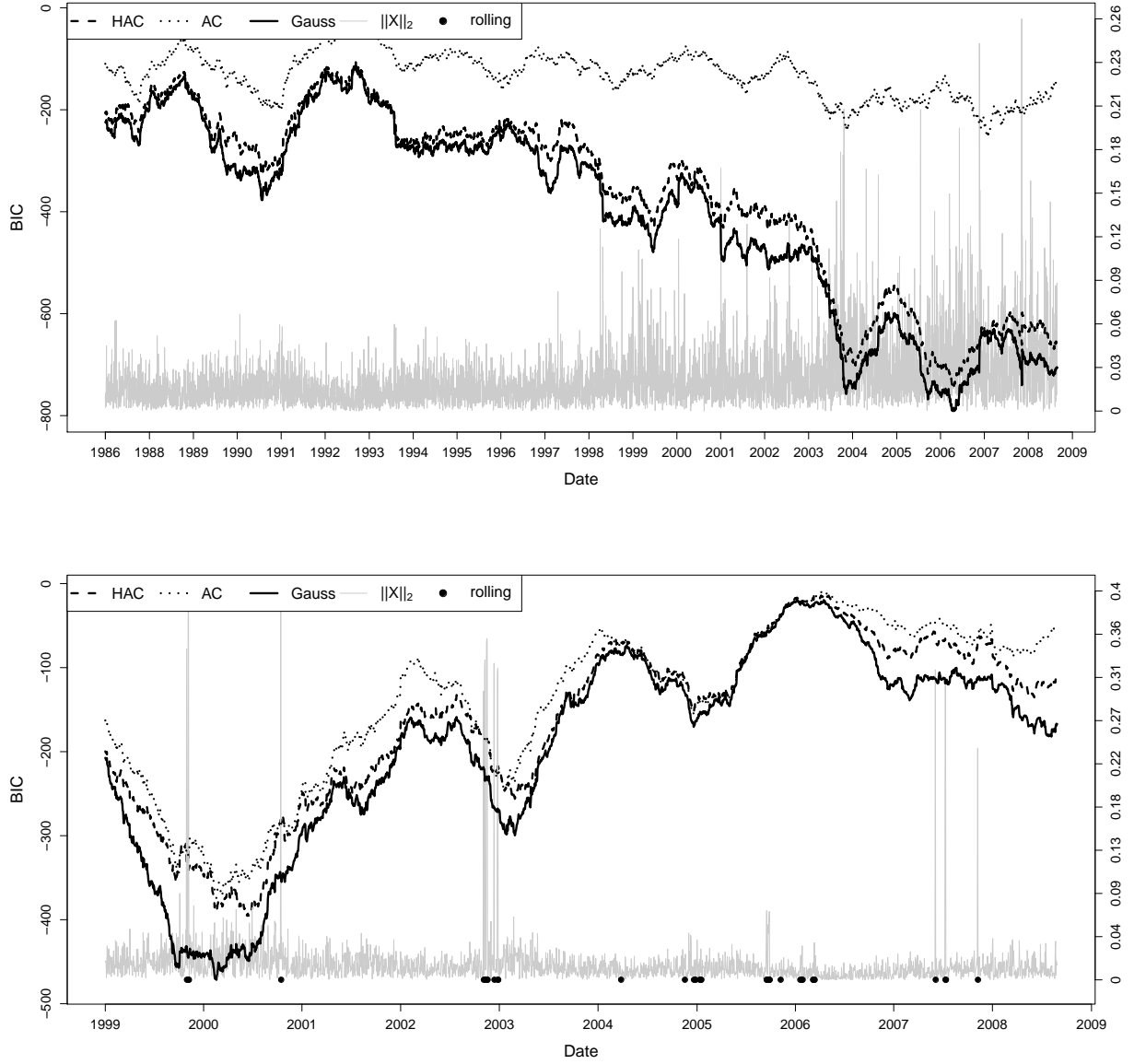
Figure 9: Rolling window estimator of BIC for HAC-based, based on simple Archimedean copula and multivariate normal distributions for indices (top) and exchange rates (bottom). The width of the rolling window is 250 observations. The grey line shows the variation of L_2 norm of the difference in the parameter matrices of the copulae. The dots mark the changes in the structure of binary copula using rolling window estimation.

HAC or the copula parameters are time-dependent. In this section we apply the local window procedure to compute a robust estimator of HAC.

The setup of our procedure is chosen as follows. The set of m_k's defining the length of I_k and $\mathcal{T}_k$ is determined by a geometric grid with $m_k = [m_0 c^k]$ for $k = 1, 2, \ldots, K$. The starting value are set to $m_0 = 40$, $\rho = 0.5$ and $c = 1.25$, where $[x]$ denotes the integer part of x. The critical values $\mathfrak{z}$ are taken from the simulation study. The structure estimated from the whole data sample is given for the indices by $s^* = (\text{DAX DJ})_{2.954} \text{ NIKKEI}_{1.222}$ and for the exchange rates $s^* = (\text{JPY USD})_{1.588} \text{ GBP}_{1.418}$.

Figure 10 and Figure 11 illustrate the results of the application. The upper plots show the changes in the structure. The structure ((1.2).3) is very robust for the indices, showing no shifts over the whole period. This fact is supported by the rolling window estimation of the correlation coefficients in the previous section. Moreover, this structure coincides with the structure estimated from the full sample. In the analysis of exchange rates we observe numerous shifts in the structure. There are intervals with stable structure and with frequently changing structure. The latter periods correspond to the intervals where the dependencies between different pairs of variables are similar. In this case the procedure can hardly distinguish between different, but similar structures. The second two pictures show the parameters estimators over the intervals of homogeneity. The grey line depicts the larger parameter, while the black line depicts the smaller parameter. The intersection point of two lines imply a change in the structure. We see that the algorithm captures even relatively small changes in parameters. The figure for the indices exhibits that the parameters become more distant with time. However, for the exchange rates the parameters show a stable comovement over the whole sample with minor drifts.

The third pictures indicate the dynamics of the ML criteria over the intervals of homogeneity. Recall that the local window procedure is based on the stability of the fit measured by maximum-likelihood. The overall fit of the HAC increases in time for the indices, but decreases for the exchange rates. Note, that neither the changes in the structure nor changes in the parameters can explain the variation in the ML. However, the overlapping of both shifts closely follows the drift in the ML criteria. The bottom figures present the length of the intervals of homogeneity. For the indices the intervals drops to $m_0 = 40$ after a larger drift in the parameters, but steadily increases to the maximum of $m_K = 300$ over stable periods. For the exchange rates the intervals of homogeneity are shorter. This is due to the fact that the decreases of the intervals of homogeneity are caused both by the changes in the structure and the shifts in the parameters. This renders shorter intervals,

25

but with similar steady increases over stable periods.

The analysis of the figures reveals two different behaviors of the distribution for indices and the exchange rates. For the indices the LCP procedure detects a clear transition in the parameters under virtually constant structure. However, for the exchange rates the local window estimation detects numerous changes in the structure keeping the variation in parameters low. Thus the suggested procedure is successful in both estimation of the structure and estimation of the parameters. Additionally this implies that the change in the distribution is substantial and cannot be neglected.

6   Conclusions

We propose a method of estimating the time-varying dependencies. The joint distribution of multivariate observations is modeled by a Hierarchical Archimedean copula. Using the Local Change Point detection procedure we determine the intervals with homogeneous dependency structure and homogeneous copula parameters. In contrary to non-copula-based distributions, where the functional form of the dependency is fixed, the adaptive estimation procedure allows for variations in the form of the copula. The procedure was evaluated in an extensive simulation study and compared to the classical rolling window estimation. Application to real data disclosed interesting features of the dynamics of dependencies.

References

Chen, X. and Fan, Y. (2006). Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspesification, *Journal of Econometrics* **135**: 125–154.

Chen, Y., Härdle, W. and Jeong, S.-O. (2008). Nonparametric risk management with generalized hyperbolic distributions, *Journal of the American Statistical Association* **103**(483): 910–923.

Embrechts, P., Lindskog, F. and McNeil, A. J. (2003). Modeling dependence with copulas and applications to risk management, *in* S. T. Rachev (ed.), *Handbook of Heavy Tailed Distributions in Finance*, Elsevier, North-Holland.
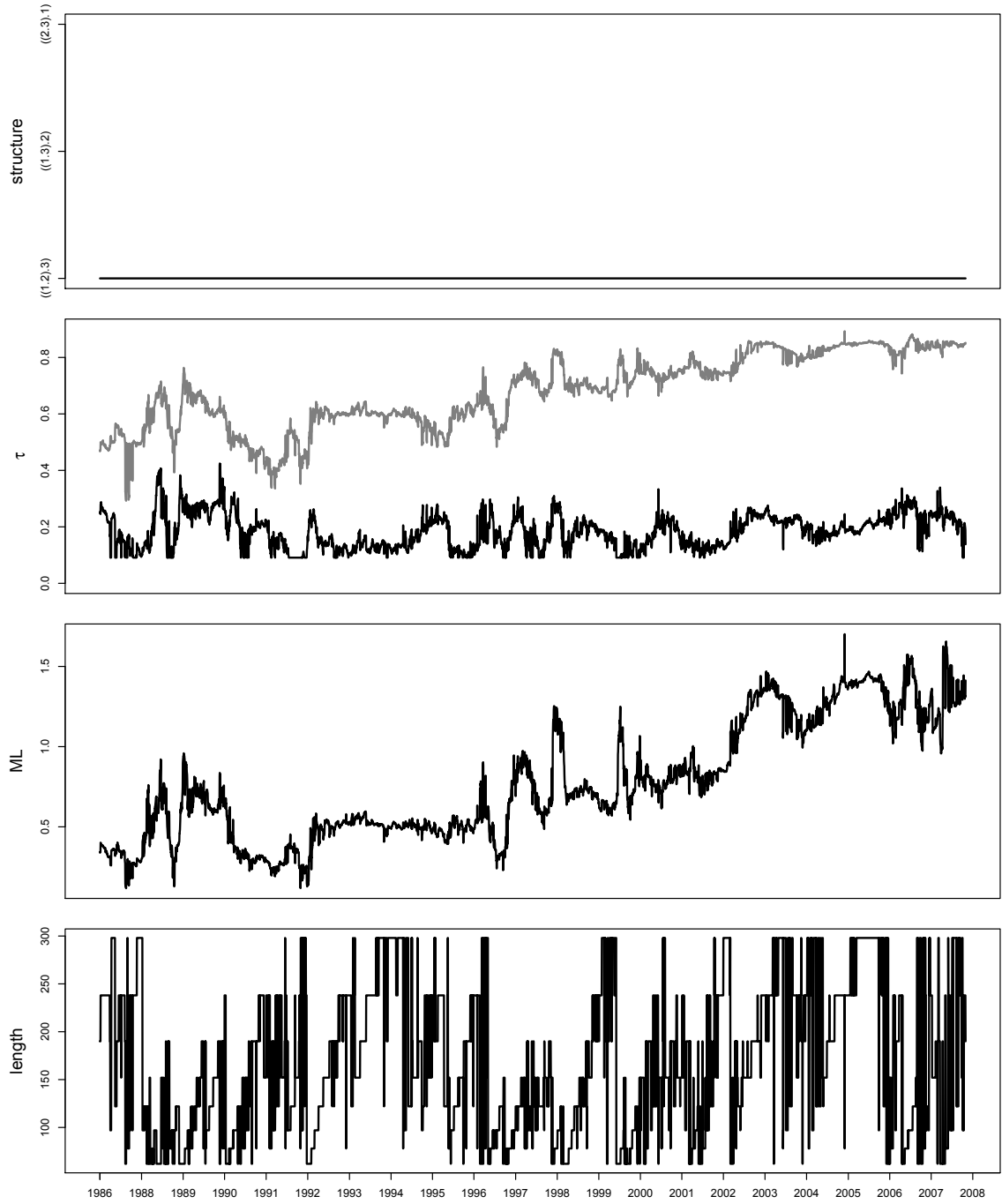
Figure 10: Changes in the structure, changes in the parameters and variation of the maximum-likelihood over the intervals of homogeneity for DJ, NIKKEI and DAX modeled with binary Gumbel HAC. m_0 is set to 40 and $\rho = 0.5$.
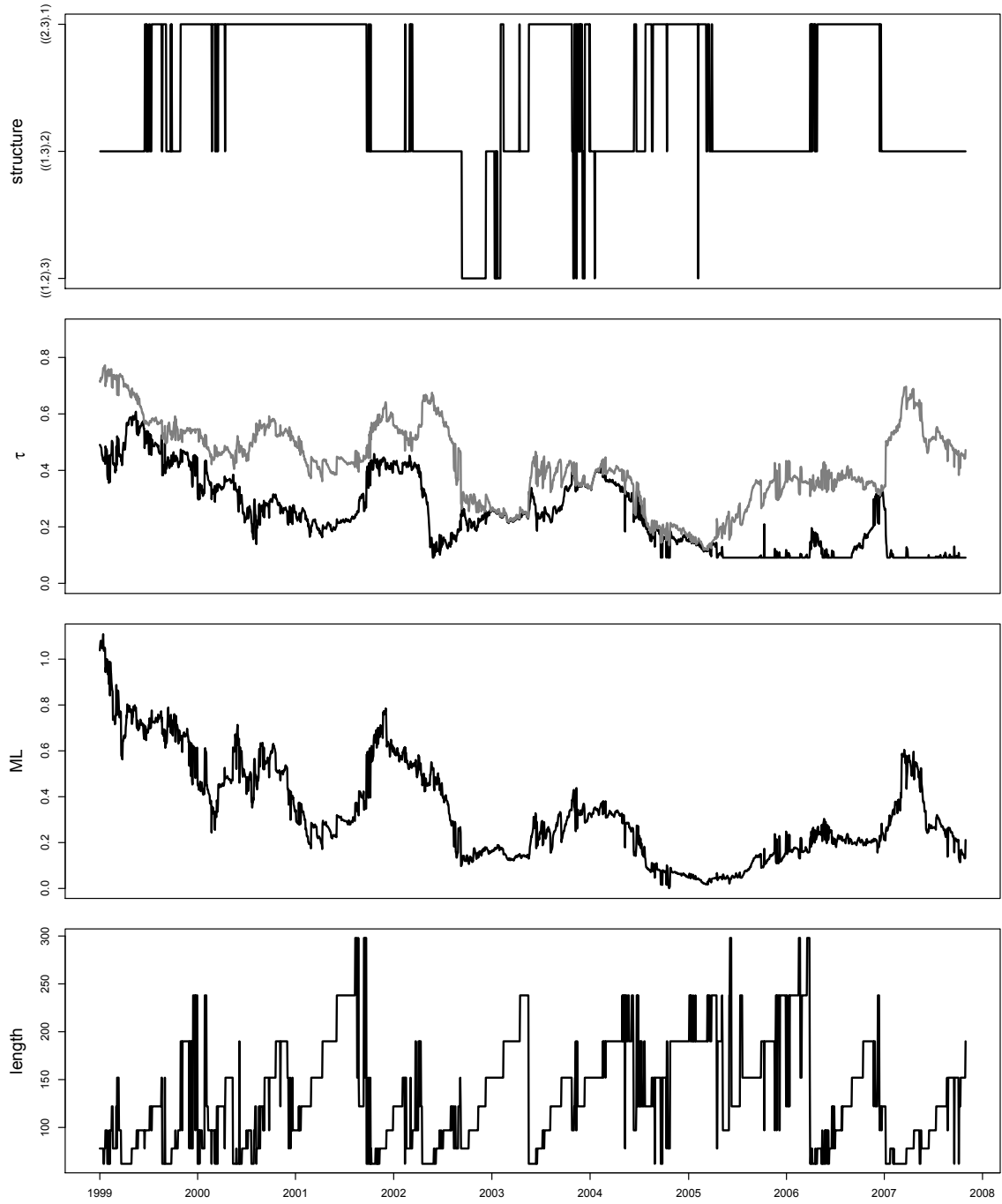
Figure 11: Changes in the structure, changes in the parameters and variation of the maximum-likelihood over the intervals of homogeneity for JPY, GBP, EUR modeled with binary Gumbel HAC. m_0 is set to 40 and $\rho = 0.5$.

Giacomini, E., Härdle, W. K. and Spokoiny, V. (2009). Inhomogeneous dependence modeling with time-varying copulae, *Journal of Business and Economic Statistics* **27**(2): 224–234.

Joe, H. (1996). Families of m-variate distributions with given margins and $m(m-1)/2$ bivariate dependence parameters, *in* L. Rüschendorf, B. Schweizer and M. Taylor (eds), *Distribution with fixed marginals and related topics*, IMS Lecture Notes – Monograph Series, Institute of Mathematical Statistics.

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

Lamoureux, C. G. and Lastrapes, W. D. (1990). Persistence-in-variance, structural change and the GARCH model, *Journal of Business and Economic Statistics* **8**: 225–234.

McNeil, A. J. (2008). Sampling nested Archimedean copulas, *Journal Statistical Computation and Simulation* **78**(6): 567–581.

McNeil, A. J. and Nešlehová, J. (2009). Multivariate Archimedean copulas, d-monotone functions and l_1 norm symmetric distributions, *Annals of Statistics* **37**(5b): 3059–3097.

Mercurio, D. and Spokoiny, V. (2004). Statistical inference for time-inhomogeneous volatility models, *Annals of Statistics* **32**(2): 577–602.

Nelsen, R. B. (2006). *An Introduction to Copulas*, Springer Verlag, New York.

Okhrin, O., Okhrin, Y. and Schmid, W. (2009). On the structure and estimation of hierarchical Archimedean copulas. under revision in Journal of Econometrics.

Patton, A. J. (2004). On the out-of-sample importance of skewness and asymmetric dependence for asset allocation, *Journal of Financial Econometrics* **2**: 130–168.

Rodriguez, J. C. (2007). Measuring financial contagion: a copula approach, *Journal of Empirical Finance* **14**: 401–423.

Savu, C. and Trede, M. (2006). Hierarchical Archimedean copulas, *Discussion paper*, University of Muenster.

Schmid, F. and Schmidt, R. (2006). Multivariate extensions of Spearman's rho and related statistics, *Statistics and Probability Letters* **77**(4): 407–416.

Silvennoinen, A. and Teräsvirta, T. (2009). Multivariate GARCH models, *in* T. G. Andersen, R. A. Davis, J.-P. Kreiß and T. Mikosch (eds), *Handbook of Financial Time Series*, Springer-Verlag, Berlin Heidelberg, pp. 201–233.

Sklar, A. (1959). Fonctions dé repartition á n dimension et leurs marges, *Publ. Inst. Stat. Univ. Paris* **8**: 299–231.

Spokoiny, V. (2009). Multiscale local change point detection with applications to value-at-risk, *The Annals of Statistics* **37**(3): 1405–1436.

Spokoiny, V. (2010). *Local Parametric Methods in Nonparametric Estimation*, Springer Verlag, Berlin, Heidelberg. forthcoming.

Čìžek, P., Härdle, W. and Spokoiny, V. (2009). Adaptive pointwise estimation in time-inhomogeneous conditional heteroscedasticity models, *Econometrics Journal* **12**(2): 248–271.

Whelan, N. (2004). Sampling from Archimedean copulas, *Quantitative Finance* **4**: 339–352.

White, H. (1982). Maximum likelihood estimation of misspecified models, *Econometrica* **50**: 1–25.

SFB 649 Discussion Paper Series 2010

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