Time Varying Hierarchical Archimedean Copulae (*HALOC*)

Wolfgang Härdle
Ostap Okhrin
Yarema Okhrin

Ladislaus von Bortkiewicz
Chair of Statistics
C.A.S.E. – Center for Applied Statistics and Economics
Humboldt-Universität zu Berlin

Chair of Statistics
Universität Augsburg

Simple AC over time

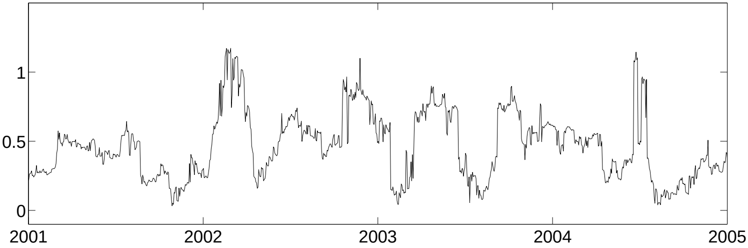


Figure 1: Estimated copula dependence parameter $\widehat{\theta}_t$ with the Local Change Point method for 6-dimensional data: DC, VW, Bayer, BASF, Allianz and Münchener Rückversicherung. Clayton Copula.
Giacomini et. al (2009)

Time Varying Structures

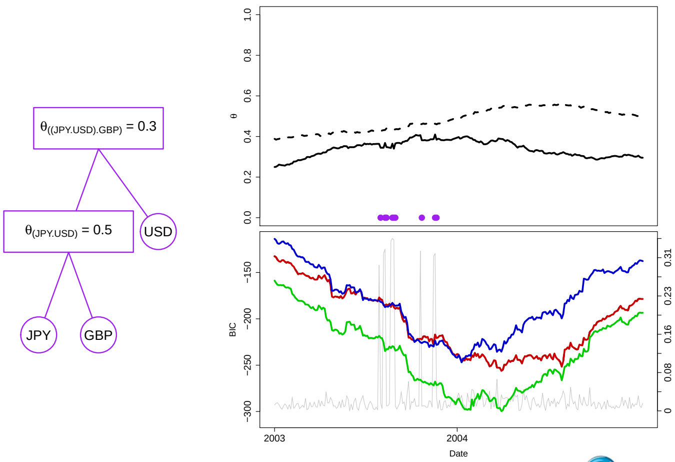


Figure 2: Film of changing structures over time.

Main Idea

- combine interpretability with flexibility of copulae
- determine the structure of HAC for a given time series
- identify time varying dependencies
- apply to risk pattern analysis
- reduce dimension of dependency

Outline

Copula

For a distribution function F with marginals $F_{X_1} \ldots, F_{X_d}$. There exists a copula $C : [0,1]^d \to [0,1]$, such that

$$F(x_1, \ldots, x_d) = C\{F_{X_1}(x_1), \ldots, F_{X_d}(x_d)\} \qquad (1)$$

for all $x_i \in \overline{\mathbb{R}}$, $i = 1, \ldots, d$. If $F_{X_1}, \ldots, F_{X_d}$ are cts, then C is unique. If C is a copula and $F_{X_1}, \ldots, F_{X_d}$ are cdfs, then the function F defined in (1) is a joint cdf with marginals $F_{X_1}, \ldots, F_{X_d}$.

Archimedean Copulae

Multivariate Archimedean copula $C : [0,1]^d \to [0,1]$ defined as

$$C(u_1, \dots, u_d) = \phi\{\phi^{-1}(u_1) + \dots + \phi^{-1}(u_d)\}, \qquad (2)$$

where $\phi : [0, \infty) \to [0,1]$ is continuous and strictly decreasing with $\phi(0) = 1$, $\phi(\infty) = 0$ and ϕ^{-1} its pseudo-inverse.

Example

$$\phi_{Gumbel}(u, \theta) = \exp\{-u^{1/\theta}\}, \text{ where } 1 \leq \theta < \infty$$

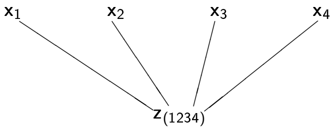
$$\phi_{Clayton}(u, \theta) = (\theta u + 1)^{-1/\theta}, \text{ where } \theta \in [-1, \infty) \backslash \{0\}$$

Disadvantages: too restrictive, single parameter, exchangeable

Hierarchical Archimedean Copulae

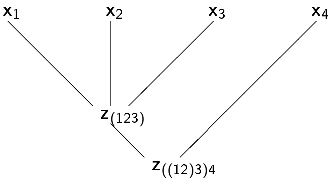
Simple AC with s=(1234)

$$C(u_1, u_2, u_3, u_4) = C_1(u_1, u_2, u_3, u_4)$$



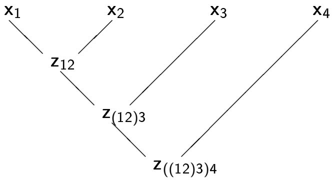
AC with s=((123)4)

$$C(u_1, u_2, u_3, u_4) = C_1\{C_2(u_1, u_2, u_3), u_4\}$$



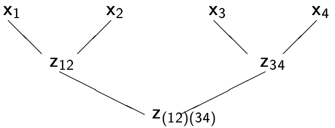
Fully nested AC with s=(((12)3)4)

$$C(u_1, u_2, u_3, u_4) = C_1[C_2\{C_3(u_1, u_2), u_3\}, u_4]$$



Partially Nested AC with s=((12)(34))

$$C(u_1, u_2, u_3, u_4) = C_1\{C_2(u_1, u_2), C_3(u_3, u_4)\}$$

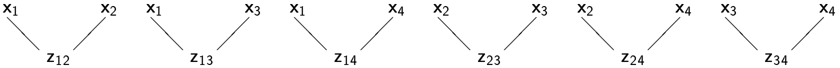


Hierarchical Archimedean Copulae

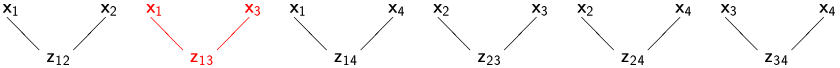
Advantages of HAC:

- flexibility and wide range of dependencies:
  for $d = 10$ more than $2.8 \cdot 10^8$ structures
- dimension reduction:
  $d - 1$ parameters to be estimated
- subcopulae are also HAC

Criteria for grouping based on θ's

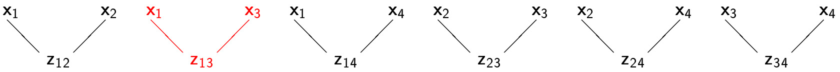


Criteria for grouping based on θ's

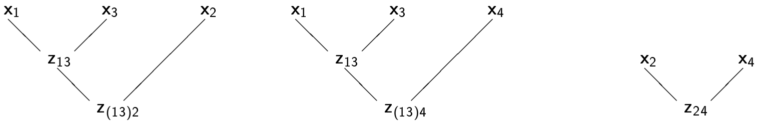


$$\max\{\widehat{\theta}_{12}, \widehat{\theta}_{13}, \widehat{\theta}_{14}, \widehat{\theta}_{23}, \widehat{\theta}_{24}, \widehat{\theta}_{34}\} = \widehat{\theta}_{13} \quad \Rightarrow$$

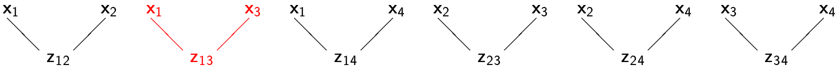
Criteria for grouping based on θ's



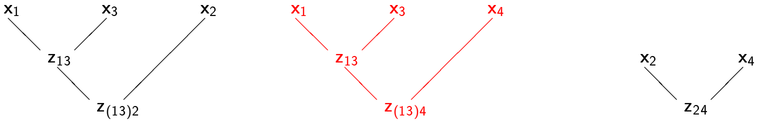
$$\max\{\widehat{\theta}_{12}, \widehat{\theta}_{13}, \widehat{\theta}_{14}, \widehat{\theta}_{23}, \widehat{\theta}_{24}, \widehat{\theta}_{34}\} = \widehat{\theta}_{13} \quad \Rightarrow$$



Criteria for grouping based on θ's

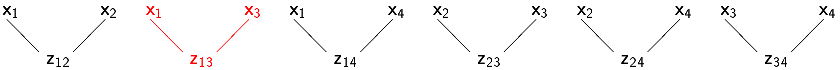


$$\max\{\widehat{\theta}_{12}, \widehat{\theta}_{13}, \widehat{\theta}_{14}, \widehat{\theta}_{23}, \widehat{\theta}_{24}, \widehat{\theta}_{34}\} = \widehat{\theta}_{13} \quad \Rightarrow$$

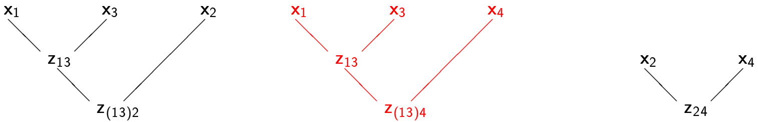


$$\max\{\widehat{\theta}_{(13)2}, \widehat{\theta}_{(13)4}, \widehat{\theta}_{24}\} = \widehat{\theta}_{(13)4} \quad \Rightarrow$$

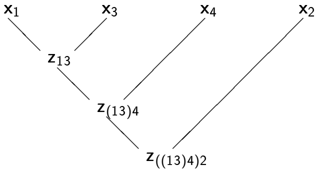
Criteria for grouping based on θ's



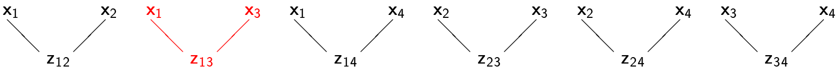
$$\max\{\widehat{\theta}_{12}, \widehat{\theta}_{13}, \widehat{\theta}_{14}, \widehat{\theta}_{23}, \widehat{\theta}_{24}, \widehat{\theta}_{34}\} = \widehat{\theta}_{13} \quad \Rightarrow$$



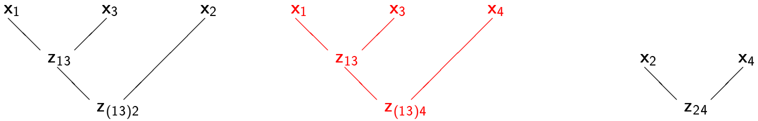
$$\max\{\widehat{\theta}_{(13)2}, \widehat{\theta}_{(13)4}, \widehat{\theta}_{24}\} = \widehat{\theta}_{(13)4} \quad \Rightarrow$$



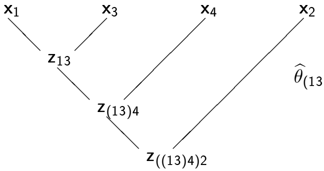
Criteria for grouping based on θ's



$$\max\{\widehat{\theta}_{12}, \widehat{\theta}_{13}, \widehat{\theta}_{14}, \widehat{\theta}_{23}, \widehat{\theta}_{24}, \widehat{\theta}_{34}\} = \widehat{\theta}_{13} \quad \Rightarrow$$

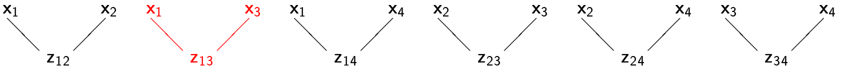


$$\max\{\widehat{\theta}_{(13)2}, \widehat{\theta}_{(13)4}, \widehat{\theta}_{24}\} = \widehat{\theta}_{(13)4} \quad \Rightarrow$$

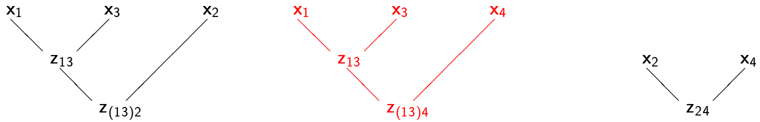


$$\widehat{\theta}_{(13)} \approx \widehat{\theta}_{((13)4)} \Rightarrow$$

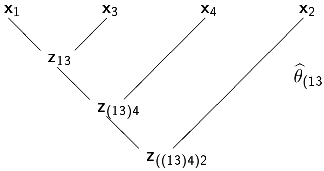
Criteria for grouping based on θ's



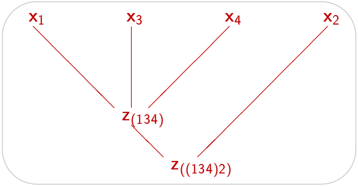
$$\max\{\widehat{\theta}_{12}, \widehat{\theta}_{13}, \widehat{\theta}_{14}, \widehat{\theta}_{23}, \widehat{\theta}_{24}, \widehat{\theta}_{34}\} = \widehat{\theta}_{13} \quad \Rightarrow$$



$$\max\{\widehat{\theta}_{(13)2}, \widehat{\theta}_{(13)4}, \widehat{\theta}_{24}\} = \widehat{\theta}_{(13)4} \quad \Rightarrow$$



$$\widehat{\theta}_{(13)} \approx \widehat{\theta}_{((13)4)} \Rightarrow$$



Local Change Point Detection

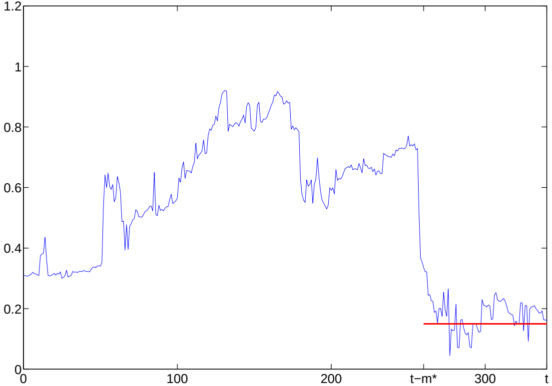


Figure 3: Dependence over time for DaimlerChrysler, Volkswagen, Bayer, BASF, Allianz and Münchener Rückversicherung, 20000101-20041231. Giacomini et. al (2009)

Adaptive Copula Estimation

- adaptively estimate largest interval where homogeneity hypothesis is accepted
- *Local Change Point* detection (LCP): sequentially test $\theta_t,\ s_t$ are constants (i.e. $\theta_t = \theta,\ s_t = s$) within some interval I (local parametric assumption).

- "*Oracle*" choice: largest interval $I = [t_0 - m_{k^*}, t_0]$ where small modelling bias condition (SMB)

$$\triangle_I(s, \boldsymbol{\theta}) = \sum_{t \in I} \mathcal{K}\{C(\cdot; s_t, \boldsymbol{\theta}_t), C(\cdot; s, \boldsymbol{\theta})\} \leq \triangle.$$

holds for some $\triangle \geq 0$. m_{k^*} is the ideal scale, $(s,\ \theta)^\top$ is ideally estimated from $I = [t_0 - m_{k^*}, t_0]$ and $\mathcal{K}(\cdot, \cdot)$ is the *Kullback-Leibler* divergence

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

Under the SMB condition on I_{k^*} and assuming that $\max_{k \le k^*} \boldsymbol{E}_{s,\boldsymbol{\theta}} |\mathcal{L}(\widetilde{s}_k, \widetilde{\boldsymbol{\theta}}_k) - \mathcal{L}(s, \boldsymbol{\theta})|^r \le \mathcal{R}_r(s, \boldsymbol{\theta})$, we obtain

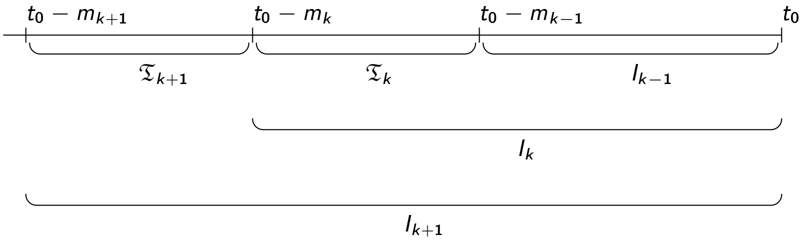
$$\boldsymbol{E}_{s_t, \boldsymbol{\theta}_t} \log \left\{ 1 + \frac{|\mathcal{L}(\widetilde{s}_{\widehat{k}}, \widetilde{\boldsymbol{\theta}}_{\widehat{k}}) - \mathcal{L}(s, \boldsymbol{\theta})|^r}{\mathcal{R}_r(s, \boldsymbol{\theta})} \right\} \le 1 + \triangle,$$

$$\boldsymbol{E}_{s_t, \boldsymbol{\theta}_t} \log \left\{ 1 + \frac{|\mathcal{L}(\widetilde{s}_{\widehat{k}}, \widetilde{\boldsymbol{\theta}}_{\widehat{k}}) - \mathcal{L}(\widehat{s}_{\widehat{k}}, \widehat{\boldsymbol{\theta}}_{\widehat{k}})|^r}{\mathcal{R}_r(s, \boldsymbol{\theta})} \right\} \le \rho + \triangle,$$

where $\widehat{a}_I$ is an adaptive estimator on I and $\widetilde{a}_I$ is any other parametric estimator on I.

Local Change Point Detection

1. define family of nested intervals
   $I_0 \subset I_1 \subset I_2 \subset \ldots \subset I_K = I_{K+1}$ with length m_k as
   $I_k = [t_0 - m_k, t_0]$
2. define $\mathfrak{T}_k = [t_0 - m_k, t_0 - m_{k-1}]$



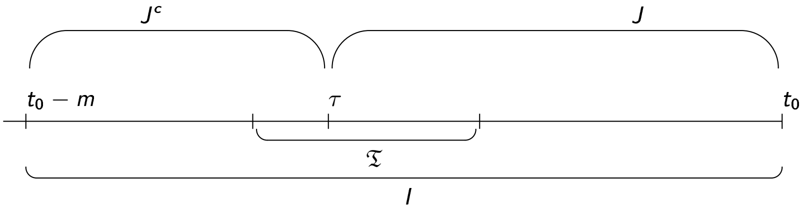
Local Change Point Detection

1. test homogeneity $H_{0,k}$ against the change point alternative in $\mathfrak{T}_k$ using I_{k+1}
2. if no change points in $\mathfrak{T}_k$, accept I_k. Take $\mathfrak{T}_{k+1}$ and repeat previous step until $H_{0,k}$ is rejected or largest possible interval I_K is accepted
3. if $H_{0,k}$ is rejected in $\mathfrak{T}_k$, homogeneity interval is the last accepted $\widehat{I} = I_{k-1}$
4. if largest possible interval I_K is accepted $\widehat{I} = I_K$

Test of homogeneity

Interval $I = [t_0 - m, t_0], \mathfrak{T} \subset I$

$$
\begin{aligned}
H_0 \quad &: \quad \forall \tau \in \mathfrak{T},\ \theta_t = \theta,\ s_t = s, \\
&\quad\ \ \forall t \in J = [\tau, t_0], \forall t \in J^c = I \setminus J \\
H_1 \quad &: \quad \exists \tau \in \mathfrak{T},\ \theta_t = \theta_1,\ s_t = s_1;\ \forall t \in J, \\
&\quad\ \ \theta_t = \theta_2 \neq \theta_1;\ s_t = s_2 \neq s_1, \forall t \in J^c
\end{aligned}
$$



Test of homogeneity

Likelihood ratio test statistic for fixed change point location:

$$\begin{aligned} T_{I,\tau} &= \max_{\theta_1,\theta_2}\{L_J(\theta_1) + L_{J^c}(\theta_2)\} - \max_{\theta} L_I(\theta) \\ &= ML_J + ML_{J^c} - ML_I \end{aligned}$$

Test statistic for unknown change point location:

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

Reject H_0 if for a critical value ζ_I

$$T_I > \zeta_I$$

Selection of I_k and ζ_k

- set of numbers m_k defining the length of I_k and $\mathfrak{T}_k$ are in the form of a geometric grid
- $m_k = [m_0 c^k]$ for $k = 1, 2, \ldots, K$, $m_0 \in \{20,\ 40\}$, $c = 1.25$ and $K = 10$, where $[x]$ means the integer part of x
- $I_k = [t_0 - m_k, t_0]$ and $\mathfrak{T}_k = [t_0 - m_k, t_0 - m_{k-1}]$ for $k = 1, 2, \ldots, K$

(Mystery Parameters)

Sequential choice of ζ_k

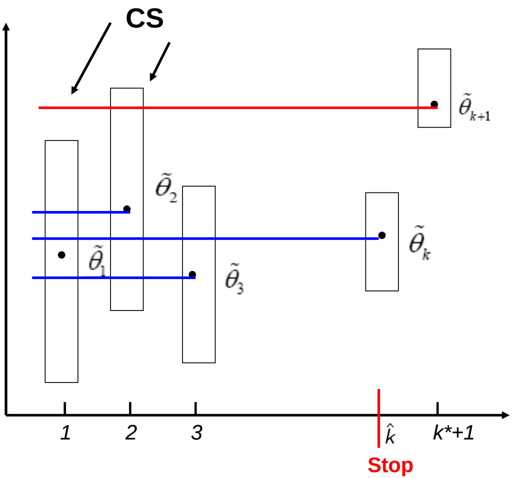
- after k steps are two cases: change point at some step $\ell \le k$ and no change points.
- let $\mathcal{B}_\ell$ be the event meaning the rejection at step ℓ

$$\mathcal{B}_\ell = \{T_1 \le \zeta_1, \ldots, T_{\ell-1} \le \zeta_{\ell-1}, T_\ell > \zeta_\ell\},$$

and $(\widehat{s}_k, \widehat{\boldsymbol{\theta}}_k) = (\widetilde{s}_{\ell-1}, \widetilde{\boldsymbol{\theta}}_{\ell-1})$ on $\mathcal{B}_\ell$ for $\ell = 1, \ldots, k$.
- we find sequentially such a minimal value of ζ_ℓ that ensures following inequality

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

Illustration



Sequential choice of ζ_k

1. pairs of Kendall's τ: $\forall\{\tau_1, \tau_2\} \in \{0.1, 0.3, 0.5, 0.7, 0.9\}^2, \; \tau_1 \geq \tau_2$
2. simul. from $C_{\theta_i,\theta_j}(u_1, u_2, u_3) = C\{C(u_1, u_2; \theta_1), u_3; \theta_2\}, \; \theta = \theta(\tau)$
3. run sequential algorithm for each sample

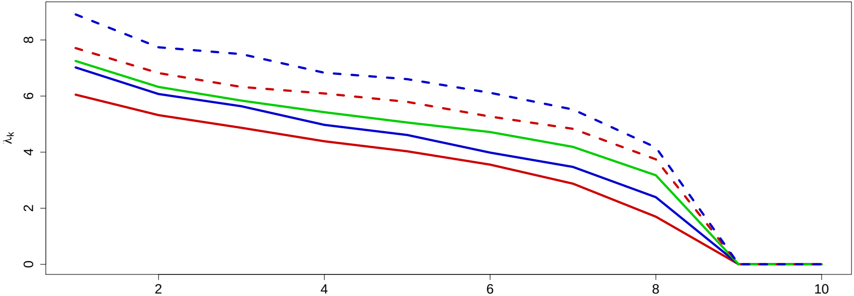
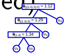


Figure 4: ζ_k of the 3-dimensional HAC as a function of k with the fixed $m_0 = 40$, $\rho = 0.5$, $r = 0.5$, $\tau_1 = 0.1$ and for different τ_2. $\tau_2 = 0.1$ (solid), $\tau_2 = 0.3$ (solid), $\tau_2 = 0.5$ (solid), $\tau_2 = 0.7$ (dashed), $\tau_2 = 0.9$ (dashed).



Simulation: Change in θ_1, I

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

1. $N = 400$ and 100 runs
2. LCP based on the same critical values

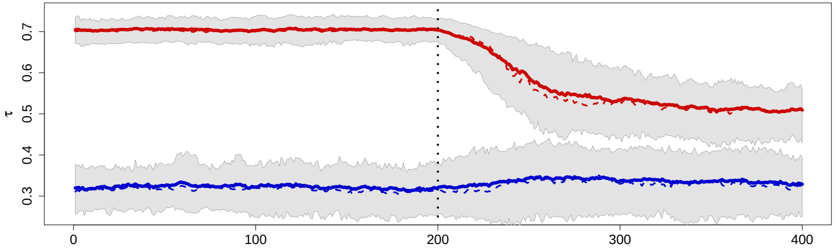


Figure 5: θ_1 and θ_2 on the intervals of homogeneity (median - dashed line, mean - solid line).

Simulation: Change in θ_1, II

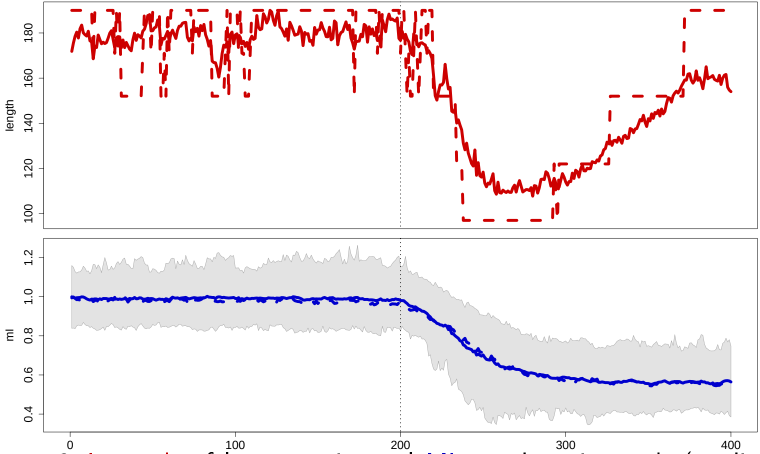


Figure 6: Intervals of homogeneity and ML on these intervals (median - dashed line, mean - solid line)

Simulation: Change in θ_2, I

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

1. $N = 400$ and 100 runs
2. LCP based on the same critical values

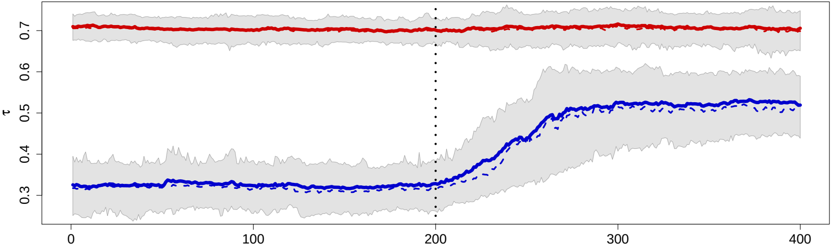


Figure 7: θ_1 and θ_2 on the intervals of homogeneity (median - dashed line, mean - solid line).

Simulation: Change in θ_2, II

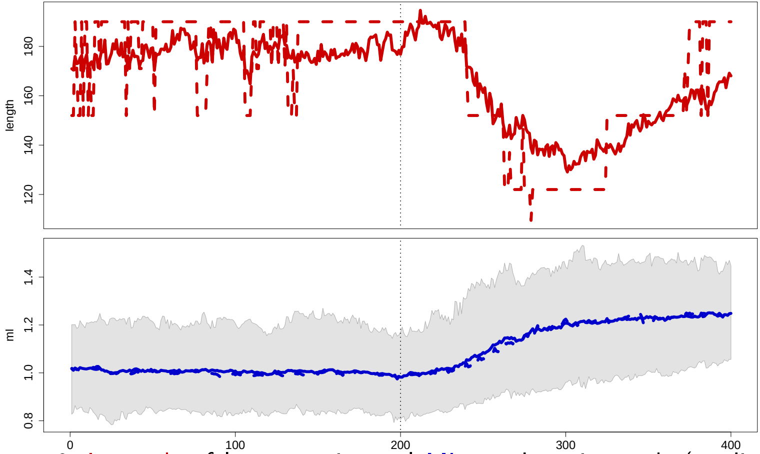


Figure 8: Intervals of homogeneity and ML on these intervals (median - dashed line, mean - solid line)

Simulation: Change in the Structure, I

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

1. $N = 400$ and 100 runs
2. LCP based on the same critical values

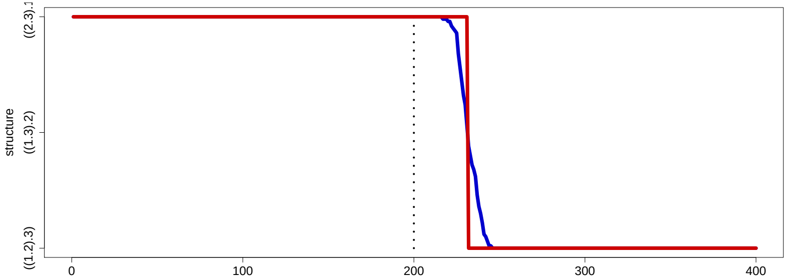


Figure 9: The structure of the est. HAC on the intervals of homogeneity (median - dashed line, mean - solid line)

Simulation: Change in the Structure, II

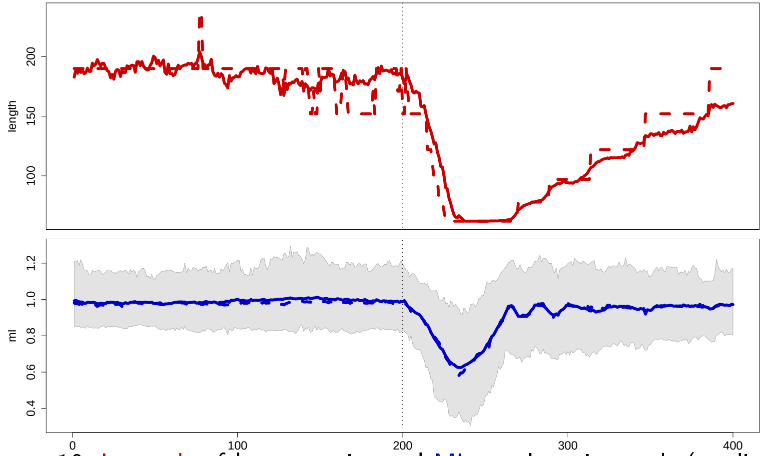


Figure 10: Intervals of homogeneity and ML on these intervals (median - dashed line, mean - solid line)

Data and Copula

daily values for the exchange rates
JPN/EUR, GBP/EUR and USD/EUR

timespan = [4.1.1999; 14.8.2009] ($n = 2771$)

$\{\phi = \exp(-u^{1/\theta})\}$ - Gumbel generator

Data and Copula

- a univariate GARCH(1,1) process on log-returns

$$X_{j,t} = \mu_{j,t} + \sigma_{j,t}\varepsilon_{j,t} \text{ with } \sigma^2_{j,t} = \omega_j + \alpha_j \sigma^2_{j,t-1} + \beta_j (X_{j,t-1} - \mu_{j,t-1})^2$$
$$\varepsilon_t \sim C\{F_1(x_1), \ldots, F_d(x_d); \theta_t\}$$

- estimated copula from the whole sample

$$s^* = (\text{JPY USD})_{1.588}\ \text{GBP}_{1.418}$$

	$\widehat{\mu}_j$	$\widehat{\omega}_j$	$\widehat{\alpha}_j$	$\widehat{\beta}_j$	BL	KS
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 1: Estimation results univariate time series modelling.

Rolling window

$$ML = \sum_{i=1}^{n} \log\{f(u_{i1}, \dots, u_{id}, \widehat{\theta})\},$$

where f denotes the joint multivariate density function.

$$AIC = -2ML + 2m, \qquad BIC = -2ML + 2\log(m),$$

where m is the number of parameters to be estimated.

$\boldsymbol{\Theta}_t(d \times d)$ - matrix of the pairwise θ based on the 250 days before t

$$||\widehat{\boldsymbol{\Theta}}_t - \widehat{\boldsymbol{\Theta}}_{t-1}||_2 = \sqrt{\lambda_{\max}\{(\widehat{\boldsymbol{\Theta}}_t - \widehat{\boldsymbol{\Theta}}_{t-1})(\widehat{\boldsymbol{\Theta}}_t - \widehat{\boldsymbol{\Theta}}_{t-1})^\top\}}$$

Copulae over time

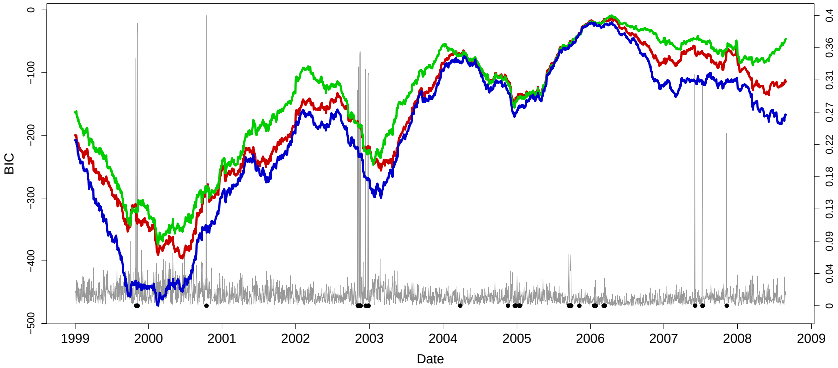


Figure 11: Time-varying HAC: BIC for the AC, Gaussian copula and HAC. Difference Matrix and points of the changes of the structure.

LCP for HAC to real Data

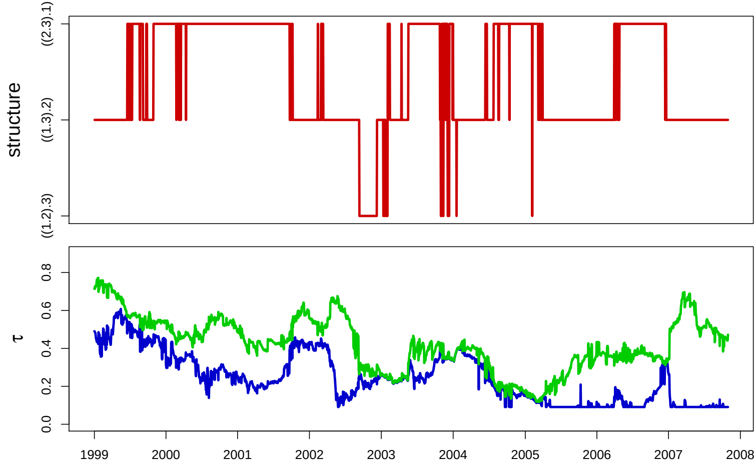
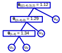


Figure 12: Structure, τ_1 and τ_2 of the HAC on the intervals of homogeneity



LCP for HAC to real Data

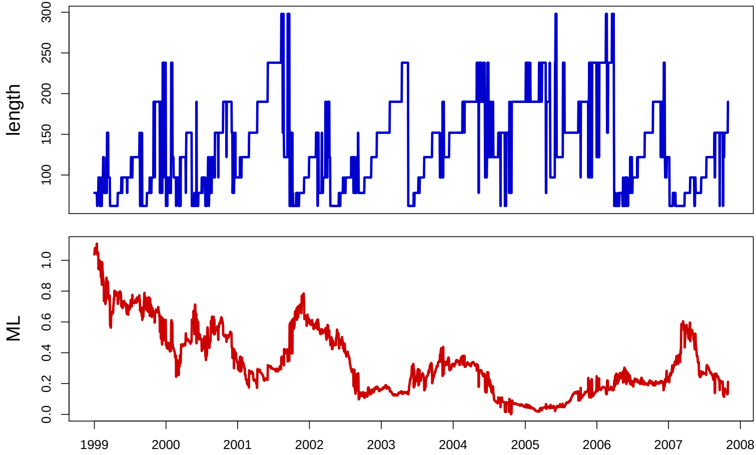


Figure 13: Intervals of homogeneity and ML on these intervals

Data and Copula

daily returns values for Dow Jones (DJ), DAX and NIKKEI
timespan = [4.1.1999; 14.8.2009] ($n = 2771$)
$\{\phi = \exp(-u^{1/\theta})\}$ - Gumbel generator
estimated copula from the whole sample

$$s^* = (\text{DAX DJ})_{2.954}\ \text{NIKKEI}_{1.222}$$

	$\widehat{\mu}_j$	$\widehat{\omega}_j$	$\widehat{\alpha}_j$	$\widehat{\beta}_j$	BL	KS
DAX	6.94e-04	4.17e-06	0.11	0.87	0.23	3.35e-05
	(1.39e-04)	(5.29e-07)	(0.01)	(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)	(0.01)	(9.40e-03)		
NIKKEI	5.62e-04	3.01e-06	0.12	0.88	0.78	2.45e-13
	(1.45e-04)	(5.18e-07)	(0.01)	(8.71e-03)		

Table 2: Estimation results univariate time series modelling.

Copulae over time

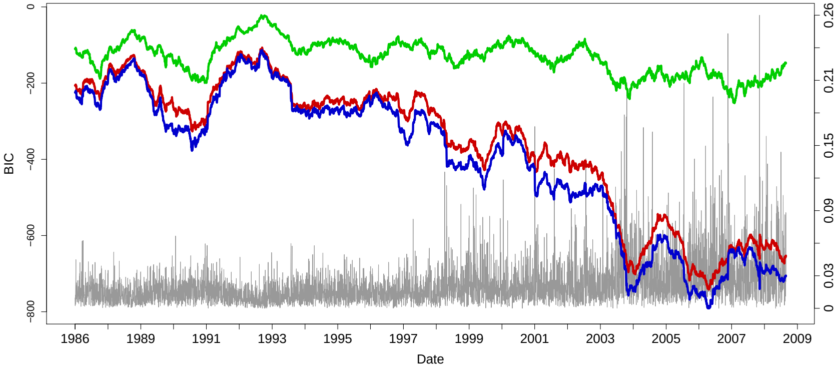


Figure 14: Time-varying HAC: BIC for the AC, Gaussian copula and HAC. Difference Matrix and points of the changes of the structure.

LCP for HAC to real Data

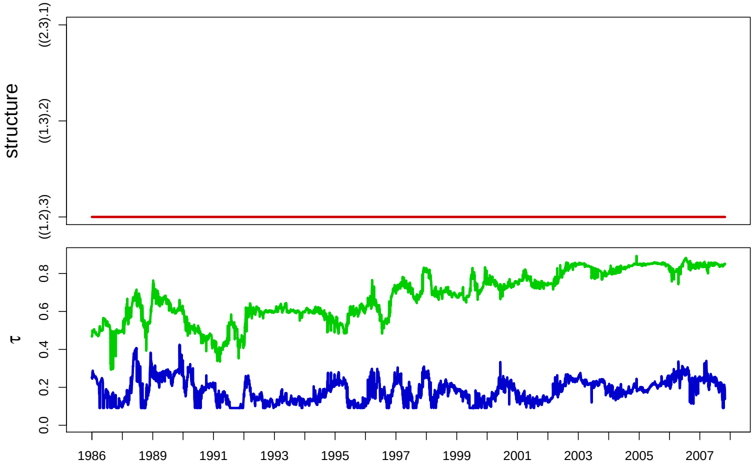


Figure 15: Structure, τ_1 and τ_2 of the HAC on the intervals of homogeneity

LCP for HAC to real Data

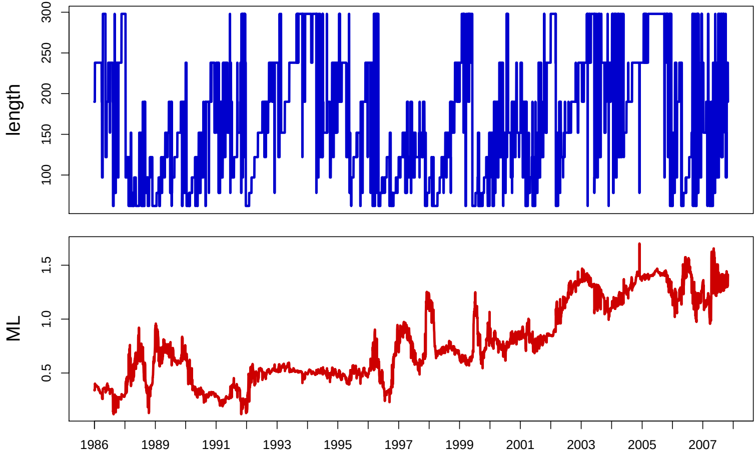


Figure 16: Intervals of homogeneity and ML on these intervals

Time Varying Hierarchical Archimedean Copulae (*HALOC*)

Wolfgang Härdle
Ostap Okhrin
Yarema Okhrin

Ladislaus von Bortkiewicz
Chair of Statistics
C.A.S.E. – Center for Applied Statistics and Economics
Humboldt-Universität zu Berlin

Chair of Statistics
Universität Augsburg

H. Joe
Multivariate Models and Concept Dependence
Chapman & Hall, 1997

V. Spokoiny
Local Parametric Methods in Nonparametric Estimation
Springer Verlag, 2009

E. Giacomini, W. Härdle and V. Spokoiny
Inhomogeneous Dependence Modeling with Time-Varying Copulae
Journal of Business and Economic Statistics, 27(2), 2009

O. Okhrin and Y. Okhrin and W. Schmid
On the Structure and Estimation of Hierarchical Archimedean Copulas
under revision in *Journal of Econometrics*, 2009