

Hautsch, Nikolaus; Okhrin, Ostap; Ristig, Alexander

Working Paper

Modeling time-varying dependencies between positive-valued high-frequency time series

SFB 649 Discussion Paper, No. 2012-054

Provided in Cooperation with:

Collaborative Research Center 649: Economic Risk, Humboldt University Berlin

Suggested Citation: Hautsch, Nikolaus; Okhrin, Ostap; Ristig, Alexander (2012) : Modeling time-varying dependencies between positive-valued high-frequency time series, SFB 649 Discussion Paper, No. 2012-054, Humboldt University of Berlin, Collaborative Research Center 649 - Economic Risk, Berlin

This Version is available at:

<https://hdl.handle.net/10419/79591>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Modeling Time-Varying Dependencies between Positive-Valued High-Frequency Time Series

Nikolaus Hautsch *

Ostap Okhrin *

Alexander Ristig *



* Humboldt-Universität zu Berlin, Germany

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

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

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



Modeling Time-Varying Dependencies between Positive-Valued High-Frequency Time Series

Nikolaus Hautsch, Ostap Okhrin and Alexander Ristig

Abstract Multiplicative error models (MEM) became a standard tool for modeling conditional durations of intraday transactions, realized volatilities and trading volumes. The parametric estimation of the corresponding multivariate model, the so-called vector MEM (VMEM), requires a specification of the joint error term distribution, which is due to the lack of multivariate distribution functions on \mathbb{R}_+^d defined via a copula. Maximum likelihood estimation is based on the assumption of constant copula parameters and therefore, leads to invalid inference, if the dependence exhibits time variations or structural breaks. Hence, we suggest to test for time-varying dependence by calibrating a time-varying copula model and to re-estimate the VMEM based on identified intervals of homogenous dependence. This paper summarizes the important aspects of (V)MEM, its estimation and a sequential test for changes in the dependence structure. The techniques are applied in an empirical example.

Keywords: vector multiplicative error model, copula, time-varying copula, high-frequency data

JEL classification: C32, C51

Acknowledgements The financial support from the Deutsche Forschungsgemeinschaft via SFB 649 Ökonomisches Risiko, Humboldt-Universität zu Berlin is gratefully acknowledged.

Nikolaus Hautsch

Professor, Chair of Econometrics, Humboldt-Universität zu Berlin, School of Business and Economics, Germany, e-mail: nikolaus.hautsch@wiwi.hu-berlin.de

Ostap Okhrin

Assistant Professor, Ladislaus von Bortkiewicz Chair of Statistics, Humboldt-Universität zu Berlin, School of Business and Economics, Germany, e-mail: ostap.okhrin@hu-berlin.de

Alexander Ristig

Research Fellow, Ladislaus von Bortkiewicz Chair of Statistics, Humboldt-Universität zu Berlin, School of Business and Economics, Germany, e-mail: ristigal@hu-berlin.de

1 Multiplicative error models

MEMs are frequently applied to describe autocorrelated positive-valued processes. The multiplicative structure became popular in the context of (G)ARCH models, see [8, 2]. [9] adopted this multiplicative approach to analyze the conditional duration of irregularly spaced financial transaction data under the assumption that the error term follows an Exponential or Weibull distribution. This extends directly to MEMs, when other positive-valued random variables such as, e.g., trading volumes, are considered. As stressed by [9], a joint model including volumes, transaction prices and time variations in liquidity gives a better understanding of the fundamental mechanisms of stock markets than individual univariate analyses.

1.1 Univariate MEM

Let x_i be a non-negative univariate time series, with time index $i = 1, \dots, n$. The univariate MEM is defined as

$$\begin{aligned} x_i &= \mu_i \varepsilon_i \\ \mu_i &\stackrel{\text{def}}{=} \mathbb{E}(x_i | \mathcal{F}_{i-1}; \xi), \end{aligned} \quad (1)$$

where ξ denotes an m -dimensional vector of parameters and the scale factor μ_i is assumed to be measurable with respect to the information set \mathcal{F}_{i-1} . Furthermore, assume that ε_i follows an iid process with $\mathbb{E}(\varepsilon_i) = 1$ and density $f(\cdot)$. The conditional mean can be specified in several ways, e.g.,

$$\mu_i = \omega + \sum_{j=1}^P \alpha_j x_{i-j} + \sum_{j=1}^Q \beta_j \mu_{i-j}, \quad (2)$$

with $\omega \geq 0$, $\alpha_j \geq 0$ and $\beta_j \geq 0$, $\forall j$, $\xi = (\omega, \alpha_1, \dots, \alpha_P, \beta_1, \dots, \beta_Q)^\top$. Based on the filters $\phi(L) = \sum_{j=1}^R \phi_j L^j = \sum_{j=1}^R (\alpha_j + \beta_j) L^j$, $\beta(L) = \sum_{j=1}^Q \beta_j L^j$ and the martingale difference series $\eta_i = x_i - \mu_i$, (2) can be transformed to an ARMA(R, Q) model

$$x_i = \omega + \phi(L)x_i + \{1 - \beta(L)\} \eta_i, \quad (3)$$

where $R = \max(P, Q)$ and L denotes the lag operator with $L^j x_i = x_{i-j}$. According to standard time series arguments, (3) is guaranteed to be weakly stationary, if $\sum_{j=1}^P \alpha_j + \sum_{j=1}^Q \beta_j < 1$. Given the above set of assumptions, we implicitly assume an exponential decay of the autocorrelation function $\rho(\cdot)$, i.e., $\lim_{l \rightarrow \infty} \sum_{j=-l}^l |\rho(j)| < \infty$. However, in case of financial high-frequency data this assumption is often not fulfilled.

As such data typically reveal long memory, we provide a short review of the fractionally integrated MEM (FIMEM), which allows the autocorrelation function

of the underlying random variable to decay hyperbolically. Formally, x_i exhibits long memory if $\lim_{l \rightarrow \infty} \sum_{j=-l}^l |\rho(j)| = \infty$. Following [1], [14] specifies the FIMEM in the context of conditional durations by introducing the fractional difference operator $(1-L)^\delta$ to equation (3), i.e.,

$$\{1 - \phi(L)\} (1-L)^\delta x_i = \omega + \{1 - \beta(L)\} \eta_i, \quad (4)$$

with $\delta \in [0, 1]$ the fractional integration parameter. [13] defines the fractional difference operator by a binomial series:

$$(1-L)^\delta = \sum_{j=0}^{\infty} \binom{\delta}{j} (-1)^j L^j = \sum_{j=0}^{\infty} \pi_j L^j. \quad (5)$$

Substituting the martingale difference series defined above in (4) leads to

$$\begin{aligned} \{1 - \beta(L)\} \mu_i &= \omega + \left[1 - \beta(L) - \{1 - \phi(L)\} (1-L)^\delta \right] x_i \\ \mu_i &= \omega \{1 - \beta(1)\}^{-1} + \lambda(L) x_i, \end{aligned} \quad (6)$$

where $\lambda(L) = \sum_{j=1}^{\infty} \lambda_j L^j$. The linear filter $\lambda(L)$ implies an infinite number of parameter restrictions to guarantee the non-negativity of μ_i , i.e., $\lambda_j \geq 0, \forall j$. As a consequence, in practice the filter $\lambda(L)$ is truncated to a finite number of lags or one needs to apply Theorem 3 of [6] to verify that the combination of parameters of the FIMEM($P; \delta; Q$) are within the feasible parameter space. To emphasize this point, consider the following two extreme examples for which we assume that δ lies within the unit interval, such that $\pi_j < 0$, for $j > 0$. Then, (i) μ_i can become negative although all parameters are greater than zero and (ii) μ_i can be positive almost surely for all i , even though all parameters except δ are negative. Note, that these restrictions play a fundamental role for the validity of forecasts.

The first unconditional moment of x_i is not defined, since the fractional difference operator evaluated at $L = 1$ equals zero. As a result, the FIMEM is not covariance stationary. If the parameters are non-negative and $\sum_{j=1}^P \alpha_j + \sum_{j=1}^Q \beta_j < 1$, then the strict stationarity and ergodicity of the FIMEM can be deduced from the stationarity and ergodicity of the integrated MEM, since the infinite-order representation of (6) is dominated in an absolute value sense by the coefficients of the corresponding integrated MEM, cf. [3, 1]. Alternative covariance stationary long memory MEMs are discussed in [12].

In general, parametric ML estimation of univariate MEMs leads to asymptotically efficient and unbiased estimates, if the distribution of the innovations ε_i is specified correctly. Typical candidates to describe ε_i are the standard Exponential or Weibull distribution, but flexible distributions as the generalized Gamma or F distribution can also be considered. In a standard ML framework for time series models, where $\ell_i(\xi)$, $i = 1, \dots, n$, denotes the i -th contribution to the log likelihood $\ell(\xi) = \sum_{i=1}^n \ell_i(\xi)$, $H_n(\xi) = \sum_{i=1}^n \left\{ \frac{\partial^2}{\partial \xi \partial \xi^\top} \ell_i(\xi) \right\}$ denotes the Hessian matrix and $S_n(\xi) = \sum_{i=1}^n \left\{ \frac{\partial}{\partial \xi} \ell_i(\xi) \frac{\partial}{\partial \xi^\top} \ell_i(\xi) \right\}$ the outer product of scores, the limiting dis-

tribution of the estimator $\hat{\xi}$ is given by

$$\{H_n(\xi)^{-1}S_n(\xi)H_n(\xi)^{-1}\}^{-1/2}\sqrt{n}(\hat{\xi}-\xi)\xrightarrow{\mathcal{L}}N(0_m,I_m), \quad (7)$$

with identity matrix I_m . Statistical inference is based on the finite sample approximation of (7), i.e., the Hessian matrix and the outer score product are replaced by the consistent estimates $H_n(\hat{\xi})$ and $S_n(\hat{\xi})$.

Furthermore, [9] adopts the asymptotic theory of [16] and proposes a quasi-ML setup which leads to consistent estimates for the linear and integrated MEM even if the true error term distribution does not correspond to the assumed standard Exponential distribution. In this case, $\hat{\xi}$ converges under some regularity conditions to the asymptotic distribution of (7) as long as the conditional mean is correctly specified.

1.2 Vector MEM

[5] formalizes the VMEM as

$$x_i = \mu_i \odot \varepsilon_i, \quad (8)$$

where \odot denotes the Hadamard product and $x_i = (x_{i1}, \dots, x_{id})^\top$, $i = 1, \dots, n$, is the vector of positive valued processes. The multivariate scale factor $\mu_i \stackrel{\text{def}}{=} E(x_i | \mathcal{F}_{i-1}; \xi)$ and the vector of error terms ε_i are $(d \times 1)$ vectors. The natural multivariate extension of (6) is given by

$$[I_d - B(L)]\mu_i = \omega + [I_d - B(L) - \{I_d - \Phi(L)\}D]x_i, \quad (9)$$

with $\Phi(L) = A(L) + B(L)$ and A, B being $(d \times d)$ matrices. Short-run effects enter equation (9) through the linear filters $A(L)$ and $B(L)$ and ω denotes the vector of constants. The univariate fractional difference operator from (6) extends to the diagonal matrix $\text{diag}(D) = \{(1-L)^{\delta_1}, \dots, (1-L)^{\delta_d}\}$, which contains the individual fractional difference operators, with $\delta_j \in [0, 1]$, $j = 1, \dots, d$. By this restriction, we exclude deterministic low frequency patterns between the marginal time series. Note that the individual mean equations of (9) collapse to the univariate FIMEM (6), if A and B are diagonal and to the linearly parameterized MEM (2), if additionally $\delta_j = 0$, $j = 1, \dots, d$. Based on the diagonality assumption for A and B the model can be estimated equation by equation and is stationary.

For the full parametric specification of the VMEM we need to define an innovation process ε_i , $i = 1, \dots, n$, which must follow a distribution with only positive probabilities on $\mathbb{R}_+^d = [0, \infty)^d$ and $E(\varepsilon_{ij}) = 1$, $j = 1, \dots, d$. However, the distribution function of a univariate error term process does not have a natural multivariate equivalent. Therefore, the d marginal distributions are coupled together with a copula splitting a multivariate distribution function into its margins and a pure dependence component – the copula. Copulae are introduced in [23] stating that if F is an arbitrary d -dimensional continuous distribution function of the random variables

X_1, \dots, X_d , then the associated copula is unique and defined as a continuous function $C : [0, 1]^d \rightarrow [0, 1]$ which satisfies the equality

$$C(u_1, \dots, u_d) = F\{F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)\}, \quad u_1, \dots, u_d \in [0, 1], \quad (10)$$

where $F_1^{-1}(\cdot), \dots, F_d^{-1}(\cdot)$ are the quantile functions of the continuous marginal distribution functions $F_1(x_1), \dots, F_d(x_d)$. Based on the copula density $c(\cdot, \dots, \cdot; \theta)$, the log likelihood of the VMEM can be written as

$$\begin{aligned} \ell(\theta, \xi, \alpha | \mathcal{F}_{i-1}) &= \sum_{i=1}^n \sum_{j=1}^d [\log [\varepsilon_{ij}(\xi_j) f_j \{\varepsilon_{ij}(\xi_j); \alpha_j\}] - \log x_{ij}] \\ &\quad + \sum_{i=1}^n \log c[F_1\{\varepsilon_{i1}(\xi_1), \alpha_1\}, \dots, F_1\{\varepsilon_{id}(\xi_d), \alpha_d\}; \theta], \end{aligned} \quad (11)$$

with $x_i/\mu_i(\xi) | \mathcal{F}_{i-1} = \varepsilon_i(\xi) | \mathcal{F}_{i-1} \sim C[F_1\{\varepsilon_{i1}(\xi_1), \alpha_1\}, \dots, F_1\{\varepsilon_{id}(\xi_d), \alpha_d\}; \theta]$ having expectation one, where θ denotes the copula-, α the marginal- and ξ the mean-parameters, cf. [5]. Conversely to the Hadamard product, x_i/μ_i denotes element-wise division. The efficient approach to obtain parameter estimates is given by full ML estimation, as the multivariate density function is assumed to be known, i.e., the product of the marginal densities multiplied with the copula density. On the other hand, full ML estimation is difficult to implement even if the induced dependence is non-elliptical. E.g., if we assume a Vine- or hierarchical Archimedean copula (HAC), (see Section 2), the copula density varies with the structure of the underlying copula. Thus, the log likelihood must be optimized for each possible structure and the parameter vector generating the largest log likelihood value is selected as ML estimate. To avoid this computationally intensive method, a two-step procedure similar to [4] can be straightforwardly applied, since (11) can be decomposed into a marginal and a copula part as follows: First, the parameters of the mean equation are estimated to filter the residuals, for which only the information about the marginal distributions is used. Then, the copula is calibrated to the fitted values of the residuals' empirical distribution functions.

Similar to classical risk management applications, where several time-varying models for correlations and copulae are proposed, e.g., [7, 22], time-varying dependence cannot be excluded in our context and consequently, the copula estimated at the second step may contain time variations. Yet, the final target of VMEMs is not to predict, e.g., tail dependencies or risk measures, but to produce forecasts of μ_i , which crucially depend on precise parameter estimates and thus on the complete log likelihood and the most recent data for which the dependence between the variables is constant. Thus, we suggest to re-estimate the parameters of μ_i by maximizing (11) with fixed θ for time intervals at which the copula model calibrated at the second step supports constant dependence.

2 Hierarchical Archimedean Copulae

Among other important families, there exists the class of Archimedean copulae (AC), which (i) permits modeling non-elliptical dependencies, (ii) can describe different types of tail dependencies and (iii) has a closed form expression. Formally, AC are defined through the generator function $\phi_\theta \in \mathcal{L} = \{\phi_\theta : [0; \infty) \rightarrow [0, 1] \mid \phi_\theta(0) = 1, \phi_\theta(\infty) = 0; (-1)^i \phi_\theta^{(i)} \geq 0; i \in \mathbb{N}\}$ and $(-1)^i \phi_\theta^{(i)}(x)$ being non-decreasing and convex on $[0, \infty)$, for $x > 0$, which commonly depends on a single parameter θ , i.e.,

$$C(u_1, \dots, u_d; \theta) = \phi_\theta \{ \phi_\theta^{-1}(u_1) + \dots + \phi_\theta^{-1}(u_d) \}, \quad u_1, \dots, u_d \in [0, 1]. \quad (12)$$

Properties of Archimedean copulae are reviewed and investigated in [18, 15]. [19] discusses generator families depending on two parameters. The restricted dependence structure induced by Archimedean generators is the major disadvantage of d -dimensional ACs, since this assumption is mostly violated in practice.

To permit more flexibility, arguments of an AC can be replaced by further ACs leading to the concept of HAC, which can adopt arbitrary complicated structures denoted by s in the following. The generators of a single HAC, ϕ_j , can come from different generator families. However, if the ϕ_{θ_j} 's come from the same family, the required complete monotonicity of $\phi_{\theta_{j+1}}^{-1} \circ \phi_{\theta_j}$ imposes constraints on the parameters $\theta_1, \dots, \theta_{d-1}$. The flexibility induced by the structure is accompanied by larger amounts of parameters, as each generator composition corresponds to one additional parameter. Sufficient conditions on the generator functions guaranteeing that C is a copula are stated in [17]. It holds that if $\phi_{\theta_j} \in \mathcal{L}$, for $j = 1, \dots, d-1$, and $\phi_{\theta_{j+1}}^{-1} \circ \phi_{\theta_j}$ have completely monotone derivatives, then C is a copula for $d \geq 2$. The major advantage of HACs compared to ACs is the non-exchangeability of the arguments beyond a single node, which is imposed by the structure of a HAC. Similar to the dependence parameters, s is generally unknown and can be regarded as an additional parameter to estimate.

A sequential estimation procedure for HACs is discussed by [20] providing statistical inference for parametric and nonparametric estimated margins. The procedure uses Proposition 1 of [21] stating that HACs can be uniquely reconstructed from marginal distributions and bivariate copula functions. The estimation procedure can be summarized in the following way: at the first step, estimate all binary copula parameters of a specified Archimedean family under the assumption of known marginal distribution functions. Select the largest parameter and fix the binary copula as pseudo-variable. At next steps, assume the estimated margins and sub-copulae from lower levels are known and estimate all binary copula parameters by considering pairs of margins, pairs of pseudo variables and pairs of margins and pseudo variables. Then, choose the largest parameter and fix the corresponding copula as a pseudo variable. This procedure leads a binary approximation of an arbitrary HAC. Let $\varepsilon_i = \{\varepsilon_{i1}, \dots, \varepsilon_{id}\}^\top$ be the sample, $i = 1, \dots, n$, and $\theta = (\theta_1, \dots, \theta_{d-1})^\top$ be the copula parameters ordered from the lowest to the highest hierarchical level.

The multi-stage ML-estimator, $\hat{\theta}$, provides a solution for the following system of equations

$$\begin{aligned} & \left(\frac{\partial \ell_1}{\partial \theta_1}, \dots, \frac{\partial \ell_{d-1}}{\partial \theta_{d-1}} \right)^\top = \mathbf{0}, \quad (13) \\ \text{where } \ell_j &= \sum_{i=1}^n l_j(\epsilon_i), \text{ for } j = 1, \dots, d-1, \\ l_j(\epsilon_i) &= \log \left\{ c \left[\{\hat{F}_m(\epsilon_{im})\}_{m \in s_j}; s_j, \theta_j \right] \prod_{m \in s_j} \hat{f}_m(\epsilon_{im}) \right\} \\ & \text{for } i = 1, \dots, n, \end{aligned}$$

where ℓ denotes the copula part of (11) and s_j contains the indices, which are structured according to the fixed subcopulae (and margins) at lower hierarchical levels.

3 Change point detection

The time intervals for which the parameters of μ_i should be re-estimated are identified by calibrating a time-varying copula. In this context, [11] proposes a framework, which incorporates time-varying HAC parameters θ_i and s_i , and is closely related to the local change point (LCP) procedure applied in [24]. As a detailed discussion of this sophisticated method is beyond the scope of this paper, this section describes only the main ideas of the data driven LCP.

Let θ_i, s_i be the unknown time-varying parameters and structure of the HAC C . Let $I = [i_0 - m, i_0]$ denote an interval with reference point i_0 , $m > 0$ and let $\Delta_I(\theta, s) = \sum_{i \in I} \mathcal{K} \{c(\cdot, \theta_i, s_i), c(\cdot; \theta, s)\}$ be a random quantity, where $\mathcal{K}(\cdot, \cdot)$ denotes the Kullback-Leibler divergence. Furthermore, let $\Delta_I(\theta, s) \leq \Delta$ be the small modeling bias (SMB) condition with $\Delta \geq 0$ and constant parameters θ, s . As $\mathcal{K}(\cdot, \cdot)$ measures the discrepancy between two densities, the data generating process can be well approximated by the local constant copula $C(\cdot; \theta, s)$ on I in the sense of the SMB condition. Based on this condition [11] proposes testing whether a HAC with time-varying parameters and structure can be locally approximated by a HAC with constant parameters and structure.

Under the null hypothesis assume that the SMB condition holds for interval I and parameters $\{\theta, s\}$ and define the set of possible change points \mathcal{T}_I for interval I , which is tested for a single but unknown change point $\tau \in \mathcal{T}_I$. The test hypotheses are formalized as

$$\begin{aligned} H_0 : & \quad \forall \tau \in \mathcal{T}_I, \theta_i = \theta, s_i = s, \forall i \in I = J \cup J^C = [\tau, i_0] \cup [i_0 - m, \tau) \quad (14) \\ H_1 : & \quad \exists \tau \in \mathcal{T}_I, \theta_i = \theta_1, s_i = s_1, \forall i \in J = [\tau, i_0], \\ & \quad \text{and } \theta_i = \theta_2 \neq \theta_1 \text{ or } s_i = s_2 \neq s_1, \forall i \in J^C = [i_0 - m, \tau). \end{aligned}$$

The null hypothesis is rejected, if the likelihood ratio (LR) test statistic

$$T_I = \max_{\tau \in \mathcal{T}_I} \left[\max_{\theta_1, s_1} \{\ell_J(\theta_1, s_1)\} + \max_{\theta_2, s_2} \{\ell_{JC}(\theta_2, s_2)\} - \max_{\theta, s} \{\ell_I(\theta, s)\} \right], \quad (15)$$

exceeds the critical value z_I . In practice, the length of the homogenous interval and the parameters of interest $\{\theta, s\}$ are estimated simultaneously due to their relation through the test statistic. For a well performing choice of the critical value, which is found via a Monte-Carlo simulation from the local parametric model and implicitly defines the significance level of the test statistic T_I , we refer to [24].

4 Empirical analysis

The considered time span of NASDAQ trade data for Apple (AAPL) starts at the January 2nd and ends at December 31th, 2009. Similar to the cleaning of TAQ data sets as, e.g., applied in [12], all non-executed trades, trades with a price smaller or equal to zero and outliers are removed from our tick-by-tick high-frequency data set. To overcome the phenomenon of simultaneous observations, trades with the same time stamp are merged and the corresponding values are aggregated by their median. A cleaned tick-by-tick data set provides information about (i) the price series p_j , (ii) the amount of traded shares s_j and (iii) the time stamp of the trades t_j , $j = 1, \dots, n^*$, where n^* is the number of daily observations. To investigate the relationships between these series, we construct the series of high-low ranges (HL), average volumes (Vol) and the number of trades (NT) on a sampling frequency of 10 min, i.e.,

$$\begin{aligned} \text{HL}_i &= \max \{p_j | t_j \in (t_{i-1}, t_i]\} - \min \{p_j | t_j \in (t_{i-1}, t_i]\}, \\ \text{NT}_i &= \# \{t_j | t_j \in (t_{i-1}, t_i]\}, \\ \text{Vol}_i &= \text{NT}_i^{-1} \sum_{t_j \in (t_{i-1}, t_i]} s_j, \end{aligned} \quad (16)$$

for $i = 1, \dots, n$, where $\#$ counts the elements of the set $\{\cdot\}$. Note, that other proxies for price variations as, e.g., the 10 min realized volatility or the squared returns, can replace the high-low range.

To remove the U-shaped daily seasonal pattern provided by the variables defined above, the individual seasonal components are approximated by fitting cubic splines and each series is divided by the respective estimated seasonal factor. Then, model (8) with mean (9) is calibrated to the process, where $A(L)$ and $B(L)$ are restricted to be diagonal and to the first lag. The infinite sums of the mean equations of the FIMEMs are truncated to 400 lagged coefficients, i.e., $\sum_{l_j=0}^{400} \pi_{l_j} L^{l_j}$, since the parameters ξ_j are almost unaffected by including additional π_{l_j} 's, $j = 1, \dots, d$. Despite these restrictions, the estimated models produce uncorrelated residuals. Figure 1 presents scatterplots of the filtered residuals. The lower diagonal elements of Figure

1 do not reveal elliptical dependencies, thus the Gaussian copula is inappropriate in this case. In the following, we prefer an approximation of the dependence structure by the hierarchical or simple Archimedean Gumbel copula, since the bivariate contour plots indicate almost the same dependencies as the underlying scatterplots.

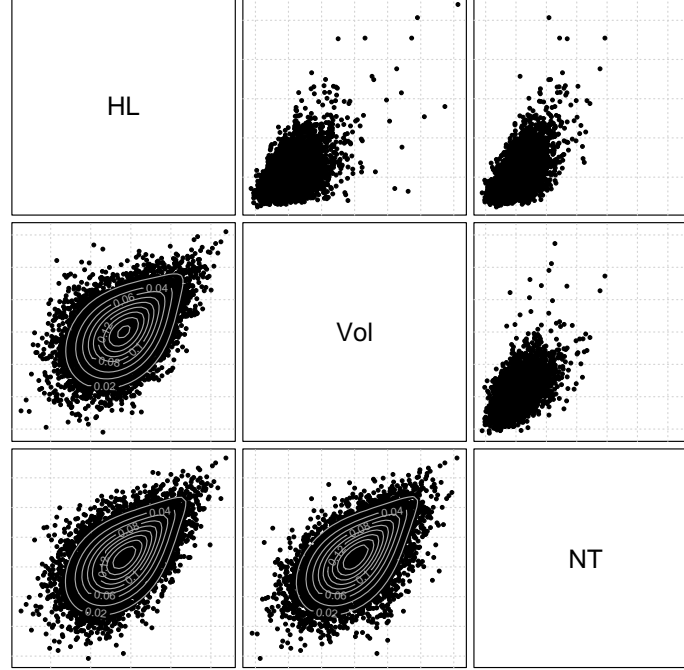


Fig. 1 The upper diagonal elements show the pairwise dependence between the filtered residuals. The lower diagonal elements present the values of the standard normal quantile applied to the values of the empirical distribution functions. Scales of the axes are not presented as they differ slightly. The origins of the coordinate planes of the upper diagonal elements correspond to zero.

The approach proposed in Section 3 considers only one single interval I , whose subintervals, defined through the set of possible change points \mathcal{T}_I , are tested for homogeneity. This method turns out to be time-varying, when it is applied as a sequential testing procedure. For this purpose, define the set \mathcal{I} , which contains the geometrically growing sequence of nested interval-candidates $I_0 \subset I_1 \subset \dots \subset I_k \subset \dots \subset I_K$, with $I_k = [i_0 - m_k, i_0]$, reference point i_0 , geometric grid $m_k = \lceil 1.25^k m_0 \rceil$, and the sets of possible change points $\mathcal{T}_{I_k} = [t_0 - m_{k-1}, t_0 - m_{k-2}]$ for all $I_k \in \mathcal{I}$. $[x]$ means the integer part of x and $m_0 = 40$. If the null hypothesis of constant dependence is not rejected for interval I_k , the interval length is extended and interval I_{k+1} is tested for homogeneity. This procedure is continued until a change point is identified or the largest interval I_K is accepted as interval of homogeneity. If a change point is detected at $k + 1$, the local adaptive estimates are given by $\hat{\theta} = \hat{\theta}_k$, $\hat{s} = \hat{s}_k$, where $\hat{\theta}_k, \hat{s}_k$ denote the ML-estimates from Section 2. While other time-varying methods

permit only the parameter(s) to vary over time, the structure of this time-varying HAC may change as well.

Based on the Gumbel family, we apply the LCP procedure to the filtered residuals, because an application of the LCP procedure to the full VMEM is cumbersome due to the large number of parameters. The first panel of Figure 2 shows the changing HAC-structure estimated for an accepted interval of homogeneity, whose length is shown in the fourth panel. The two thick solid lines (grey and black) in the second panel present the time-varying parameters in terms of Kendall's $\hat{\tau}$. For the relationship between bivariate Archimedean generators and Kendall's τ see [10]. Based on these results, we propose to re-estimate the parameters of the VMEM's scale function μ_i for at least three intervals separated by the dashed vertical lines, using full ML with fixed copula parameters. The first interval ending in the middle of March can be clearly identified, as the structure is constant and the estimates of Kendall's τ exhibit a certain distance. The HAC for this interval is given by $s_{HAC}^1 = ((\text{NT Vol})_{1.69} \text{HL})_{1.55}$ and the simple AC by $s_{AC}^1 = (\text{NT Vol HL})_{1.58}$, where the subscript is related to $\hat{\theta}$. The second interval is characterized by an alternating structure, while the values of Kendall's $\hat{\tau}$ can almost be distinguished. This makes it, however, difficult, to decide, whether a HAC or a simple AC should be used for re-estimating the VMEM. In general, the corresponding HAC, $s_{HAC}^2 = ((\text{NT Vol})_{1.63} \text{HL})_{1.40}$, and AC, $s_{AC}^2 = (\text{NT Vol HL})_{1.45}$, indicate a weaker dependence than the fitted copulas of the first interval. In the third interval beginning in June, the underlying copula corresponds with high probability to a simple AC, since the structure changes frequently and both parameters are very close to each other. The HAC of this interval, $s_{HAC}^3 = ((\text{NT HL})_{1.52} \text{Vol})_{1.40}$, shows a different structure and the AC, $s_{AC}^3 = (\text{NT Vol HL})_{1.42}$, a weaker dependence than the calibrated copulas of the first and second interval. We admit, at this point, that shorter interval specifications are possible, as the method provides a sensitive picture of the time-varying dependence. Note, that shorter time intervals are accompanied with less data and therefore, imply a loss in efficiency. The estimated HAC based on the entire sample is given by $s_{HAC} = ((\text{Vol NT})_{1.56} \text{HL})_{1.41}$ and the respective simple AC by $s_{AC} = (\text{NT HL Vol})_{1.45}$. We investigated the time-varying dependence for a few of other stocks and found similar results.

The third and fourth panels illustrate the performance of the LCP procedure. As proposed in Section 3, the LR test statistic measures the stability of the fitted model. Therefore, the length of the accepted intervals increase continuously in periods of a stable fit, whereas the interval length is typically short if the ML process is volatile. The dynamic of the ML process is presented in the third picture and allows to reproduce this relationship. The ML process exhibits a higher volatility in the last two months of the observed sample. This implies shorter intervals, for which the hypotheses of homogeneity are accepted, since the LR test statistics are smaller. [11] illustrates in a simulation study, that the procedure detects dependence changes with a short delay and [24] investigates the quality of the local adaptive estimators. A simple alternative approach is the rolling window method, which also allows for time-varying parameters but detects changes in the dependence with a larger delay.

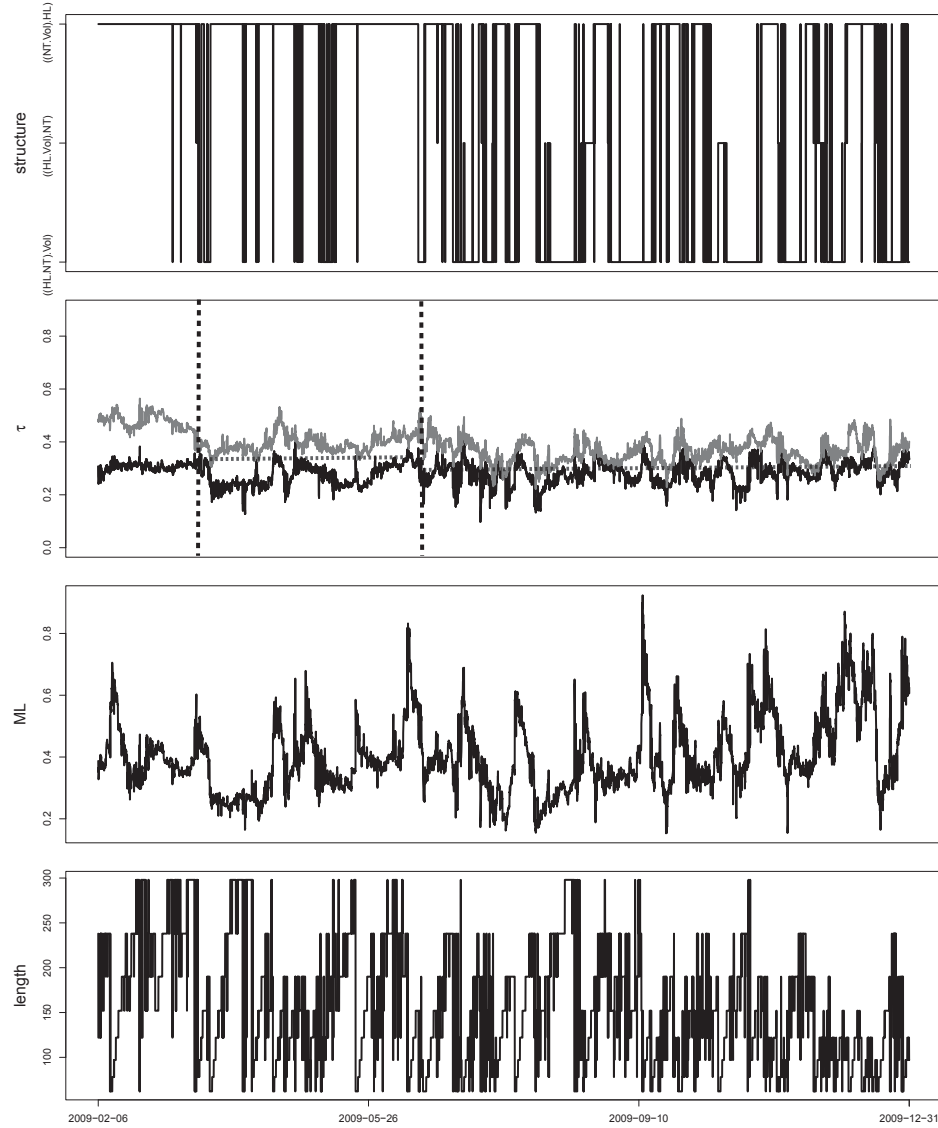


Fig. 2 Results of the LCP-procedure of AAPL. The first panel shows changes in the structure, the second the estimates of Kendall's τ and the third variations of the ML process for the intervals of homogeneity, whose varying length is presented in the lower panel.

References

1. Baillie, R.T., Bollerslev, T., Mikkelsen, H.O.: Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **74**, 3–30 (1996)
2. Bollerslev, T.: Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31**(3), 307–327 (1986)
3. Bougerol, P., Picard, N.: Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics* **52**(1-2), 115–127 (1992)
4. Chen, X., Fan, Y.: Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. *Journal of Econometrics* **135**(1-2), 125–154 (2006)
5. Cipollini, F., Gallo, G.M.: Automated variable selection in vector multiplicative error models. *Computational Statistics & Data Analysis* **54**(11), 2470–2486 (2010)
6. Conrad, C., Haag, B.R.: Inequality constraints in the fractionally integrated GARCH model. *Journal of Financial Econometrics* **4**(3), 413–449 (2006)
7. Engle, R.: Dynamic conditional correlation - a simple class of multivariate GARCH models. *Journal of Business and Economic Statistics* **20**(3), 339–350 (2002)
8. Engle, R.F.: Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica* **50**(4), 987–1007 (1982)
9. Engle, R.F., Russell, J.R.: Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica* **66**(5), 1127–1162 (1998)
10. Genest, C., Rivest, L.P.: Statistical inference procedures for bivariate Archimedean copulas. *Journal of the American Statistical Association* **88**, 1034–1043 (1993)
11. Härdle, W.K., Okhrin, O., Okhrin, Y.: Time varying hierarchical Archimedean copulae. SFB 649 discussion paper 2010, 018, SFB 649, Economic Risk, Berlin (2010)
12. Hautsch, N.: *Econometrics of Financial High-Frequency Data*. Springer, Berlin (2012)
13. Hosking, J.R.M.: Fractional differencing. *Biometrika* **68**(1), 165–176 (1981)
14. Jasiak, J.: Persistence in intertrade durations. *Finance* **19**, 166–195 (1998)
15. Joe, H.: Families of m -variate distributions with given margins and $m(m-1)/2$ bivariate dependence parameters. In: L. Rüschendorf, B. Schweizer, M. Taylor (eds.) *Distribution with Fixed Marginals and Related Topics*, IMS Lecture Notes – Monograph Series. Institute of Mathematical Statistics (1996)
16. Lee, S., Hansen, B.: Asymptotic theory for the GARCH(1, 1) quasi-maximum likelihood estimator. *Econometric Theory* **10**, 29–52 (1994)
17. McNeil, A.J.: Sampling nested archimedean copulas. *Journal of Statistical Computation and Simulation* **78**, 567–581 (2008)
18. McNeil, A.J., Nešlehová, J.: Multivariate Archimedean copulas, d -monotone functions and l_1 norm symmetric distributions. *The Annals of Statistics* **37**(5b), 3059–3097 (2009)
19. Nelsen, R.B.: *An Introduction to Copulas*, 2 edn. Springer Verlag, New York (2006)
20. Okhrin, O., Okhrin, Y., Schmid, W.: On the structure and estimation of hierarchical Archimedean copulas. Under Second Revision in *Journal of Econometrics* (2012)
21. Okhrin, O., Okhrin, Y., Schmid, W.: Properties of hierarchical Archimedean copulae. to appear in *Statistics and Risk Modeling* (Former *Statistics and Decisions*) (2012). DOI 10.1524/strm.2012.1107
22. Patton, A.J.: Modelling asymmetric exchange rate dependence. *International Economic Review* **47**, 527–556 (2006)
23. Sklar, A.: Fonctions de répartition à n dimension et leurs marges. *Publications de l’Institut de Statistique de l’Université de Paris* **8**, 299–231 (1959)
24. Spokoiny, V.: Multiscale local change point detection with applications to value-at-risk. *The Annals of Statistics* **37**(3), 1405–1436 (2009)

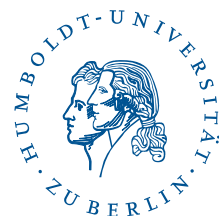
SFB 649 Discussion Paper Series 2012

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

- 001 "HMM in dynamic HAC models" by Wolfgang Karl Härdle, Ostap Okhrin and Weining Wang, January 2012.
- 002 "Dynamic Activity Analysis Model Based Win-Win Development Forecasting Under the Environmental Regulation in China" by Shiyi Chen and Wolfgang Karl Härdle, January 2012.
- 003 "A Donsker Theorem for Lévy Measures" by Richard Nickl and Markus Reiß, January 2012.
- 004 "Computational Statistics (Journal)" by Wolfgang Karl Härdle, Yuichi Mori and Jürgen Symanzik, January 2012.
- 005 "Implementing quotas in university admissions: An experimental analysis" by Sebastian Braun, Nadja Dwenger, Dorothea Kübler and Alexander Westkamp, January 2012.
- 006 "Quantile Regression in Risk Calibration" by Shih-Kang Chao, Wolfgang Karl Härdle and Weining Wang, January 2012.
- 007 "Total Work and Gender: Facts and Possible Explanations" by Michael Burda, Daniel S. Hamermesh and Philippe Weil, February 2012.
- 008 "Does Basel II Pillar 3 Risk Exposure Data help to Identify Risky Banks?" by Ralf Sabiwalsky, February 2012.
- 009 "Comparability Effects of Mandatory IFRS Adoption" by Stefano Cascino and Joachim Gassen, February 2012.
- 010 "Fair Value Reclassifications of Financial Assets during the Financial Crisis" by Jannis Bischof, Ulf Brüggemann and Holger Daske, February 2012.
- 011 "Intended and unintended consequences of mandatory IFRS adoption: A review of extant evidence and suggestions for future research" by Ulf Brüggemann, Jörg-Markus Hitz and Thorsten Sellhorn, February 2012.
- 012 "Confidence sets in nonparametric calibration of exponential Lévy models" by Jakob Söhl, February 2012.
- 013 "The Polarization of Employment in German Local Labor Markets" by Charlotte Senftleben and Hanna Wielandt, February 2012.
- 014 "On the Dark Side of the Market: Identifying and Analyzing Hidden Order Placements" by Nikolaus Hautsch and Ruihong Huang, February 2012.
- 015 "Existence and Uniqueness of Perturbation Solutions to DSGE Models" by Hong Lan and Alexander Meyer-Gohde, February 2012.
- 016 "Nonparametric adaptive estimation of linear functionals for low frequency observed Lévy processes" by Johanna Kappus, February 2012.
- 017 "Option calibration of exponential Lévy models: Implementation and empirical results" by Jakob Söhl und Mathias Trabs, February 2012.
- 018 "Managerial Overconfidence and Corporate Risk Management" by Tim R. Adam, Chitru S. Fernando and Evgenia Golubeva, February 2012.
- 019 "Why Do Firms Engage in Selective Hedging?" by Tim R. Adam, Chitru S. Fernando and Jesus M. Salas, February 2012.
- 020 "A Slab in the Face: Building Quality and Neighborhood Effects" by Rainer Schulz and Martin Wersing, February 2012.
- 021 "A Strategy Perspective on the Performance Relevance of the CFO" by Andreas Venus and Andreas Engelen, February 2012.
- 022 "Assessing the Anchoring of Inflation Expectations" by Till Strohsal and Lars Winkelmann, February 2012.

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

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



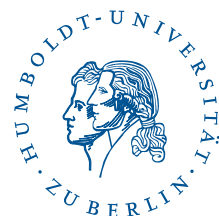
SFB 649 Discussion Paper Series 2012

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

- 023 "Hidden Liquidity: Determinants and Impact" by Gökhan Cebiroglu and Ulrich Horst, March 2012.
- 024 "Bye Bye, G.I. - The Impact of the U.S. Military Drawdown on Local German Labor Markets" by Jan Peter aus dem Moore and Alexandra Spitz-Oener, March 2012.
- 025 "Is socially responsible investing just screening? Evidence from mutual funds" by Markus Hirschberger, Ralph E. Steuer, Sebastian Utz and Maximilian Wimmer, March 2012.
- 026 "Explaining regional unemployment differences in Germany: a spatial panel data analysis" by Franziska Lottmann, March 2012.
- 027 "Forecast based Pricing of Weather Derivatives" by Wolfgang Karl Härdle, Brenda López-Cabrera and Matthias Ritter, March 2012.
- 028 "Does umbrella branding really work? Investigating cross-category brand loyalty" by Nadja Silberhorn and Lutz Hildebrandt, April 2012.
- 029 "Statistical Modelling of Temperature Risk" by Zografia Anastasiadou, and Brenda López-Cabrera, April 2012.
- 030 "Support Vector Machines with Evolutionary Feature Selection for Default Prediction" by Wolfgang Karl Härdle, Dedy Dwi Prastyo and Christian Hafner, April 2012.
- 031 "Local Adaptive Multiplicative Error Models for High-Frequency Forecasts" by Wolfgang Karl Härdle, Nikolaus Hautsch and Andrija Mihoci, April 2012.
- 032 "Copula Dynamics in CDOs." by Barbara Choroś-Tomczyk, Wolfgang Karl Härdle and Ludger Overbeck, May 2012.
- 033 "Simultaneous Statistical Inference in Dynamic Factor Models" by Thorsten Dickhaus, May 2012.
- 034 "Realized Copula" by Matthias R. Fengler and Ostap Okhrin, Mai 2012.
- 035 "Correlated Trades and Herd Behavior in the Stock Market" by Simon Jurkatis, Stephanie Kremer and Dieter Nautz, May 2012
- 036 "Hierarchical Archimedean Copulae: The HAC Package" by Ostap Okhrin and Alexander Ristig, May 2012.
- 037 "Do Japanese Stock Prices Reflect Macro Fundamentals?" by Wenjuan Chen and Anton Velinov, May 2012.
- 038 "The Aging Investor: Insights from Neuroeconomics" by Peter N. C. Mohr and Hauke R. Heekeren, May 2012.
- 039 "Volatility of price indices for heterogeneous goods" by Fabian Y.R.P. Bocart and Christian M. Hafner, May 2012.
- 040 "Location, location, location: Extracting location value from house prices" by Jens Kolbe, Rainer Schulz, Martin Wersing and Axel Werwatz, May 2012.
- 041 "Multiple point hypothesis test problems and effective numbers of tests" by Thorsten Dickhaus and Jens Stange, June 2012
- 042 "Generated Covariates in Nonparametric Estimation: A Short Review." by Enno Mammen, Christoph Rothe, and Melanie Schienle, June 2012.
- 043 "The Signal of Volatility" by Till Strohsal and Enzo Weber, June 2012.
- 044 "Copula-Based Dynamic Conditional Correlation Multiplicative Error Processes" by Taras Bodnar and Nikolaus Hautsch, July 2012

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

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



SFB 649 Discussion Paper Series 2012

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

- 045 "Additive Models: Extensions and Related Models." by Enno Mammen, Byeong U. Park and Melanie Schienle, July 2012.
- 046 "A uniform central limit theorem and efficiency for deconvolution estimators" by Jakob Söhl and Mathias Trabs, July 2012
- 047 "Nonparametric Kernel Density Estimation Near the Boundary" by Peter Malec and Melanie Schienle, August 2012
- 048 "Yield Curve Modeling and Forecasting using Semiparametric Factor Dynamics" by Wolfgang Karl Härdle and Piotr Majer, August 2012
- 049 "Simultaneous test procedures in terms of p-value copulae" by Thorsten Dickhaus and Jakob Gierl, August 2012
- 050 "Do Natural Resource Sectors Rely Less on External Finance than Manufacturing Sectors? " by Christian Hattendorff, August 2012
- 051 "Using transfer entropy to measure information flows between financial markets" by Thomas Dimpfl and Franziska J. Peter, August 2012
- 052 "Rethinking stock market integration: Globalization, valuation and convergence" by Pui Sun Tam and Pui I Tam, August 2012
- 053 "Financial Network Systemic Risk Contributions" by Nikolaus Hautsch, Julia Schaumburg and Melanie Schienle, August 2012
- 054 "Modeling Time-Varying Dependencies between Positive-Valued High-Frequency Time Series" by Nikolaus Hautsch, Ostap Okhrin and Alexander Ristig, September 2012

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

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

