IWQW

Institut für Wirtschaftspolitik und Quantitative
Wirtschaftsforschung

Diskussionspapier
Discussion Papers

No. 09/2011

Dynamic copula-based Markov chains at work: Theory, testing and performance in modeling daily stock returns

Fabian Tinkl
University of Erlangen-Nuremberg

Katja Reichert
University of Erlangen-Nuremberg

ISSN 1867-6707

Friedrich-Alexander-Universität
IWQW
Institut für Wirtschaftspolitik und Quantitative Wirtschaftsforschung

Dynamic copula-based Markov chains at work: Theory, testing and performance in modeling daily stock returns

Working Paper, 21.7.2011

Fabian Tinkl[1]
University of Erlangen-Nuremberg
Fabian.Tinkl@wiso.uni-erlangen.de

Katja Reichert
KatjaReichert@t-online.de

Keywords and phrases: Dynamic copula models, Markov chains, score test, GARCH models

Abstract

We generalize the score test for time-varying copula parameters proposed by [Abegaz & Naik-Nimbalkar, 2008] to a setting where more than one-parametric copulas can be tested for time variation in at least one parameter. In a next step we model the daily log returns of the Commerzbank stock using copula-based Markov chain models. We found evidence that compared to usual GARCH models the copula-based Markov chain models perform worse when daily stock returns are estimated. Thus we do not see any advantage of this model type when daily returns from financial data are modeled.

[1]Correspondence author: Fabian Tinkl, Department of Statistics and Econometrics, University of Erlangen-Nuremberg, Lange Gasse 20, D-90403 Nuremberg, E-Mail: Fabian.Tinkl@wiso.uni-erlangen.de

1 Introduction

Since the works of [Embrechts et al., 2002], [McNeil et al., 2005], [Patton, 2002] and [Patton, 2006] among others, copulas are now common tools in investigating con-temporal dependency between assets like in portfolio or in quantitative risk management. [Joe, 1997] proposed to model the inter-temporal dependency of Markov time series models using copulas. As economic theory often does not tell us which kind of dependency to expect, the two questions naturally arising are: first, which copula to choose and second, how to model the parameter(s) of the copula. For instance, there is some empirical evidence that correlation between different assets varies over time, see [McNeil et al., 2005] p.123. Another stylized fact is that especially financial times series tend to so-called "volatility clustering" meaning that the conditional volatility, often measured by the conditional standard deviation, varies over time. Thus having answered question one in a way such we do not have evidence against the copula model chosen, the questions arises whether time variation of certain conditional moments of our observed time series lead to time varying parameters in our copula. [Abegaz & Naik-Nimbalkar, 2008] suggested a score test under the Null that there is no time variation in an one-parametric copula. They proved the standard χ^2 asymptotic under the Null and mixing conditions on the process. The test has reasonable power but has the shortfall of being applicable only to one parametric copula families. This leaves out several interesting interesting copula families like the student-t or Joe-Clayton copula. Thus, we focus on the question how to generalize the score test of [Abegaz & Naik-Nimbalkar, 2008] to copulas with more than just one parameter. We propose a transformation for the Joe-Clayton copula, as a bivariate extension of the univariate transformation for the Clayton copula used by [Abegaz & Naik-Nimbalkar, 2008]. The paper is organized as follows: section 2 reviews the idea of copula-based Markov chains, section 3 generalize the score test of [Abegaz & Naik-Nimbalkar, 2008], section 4 investigates the power of our test in finite samples by Monte-Carlo simulation. Section 5 shows the potentials of dynamic copula-based Markov chains for modeling log returns. We compare the models with a broadly used GARCH(1,1)-model. It will be seen that after some residual analysis, the standard GARCH(1,1)-model outperforms the dynamic copula-based Markov chain models and Section 6 concludes. The proofs can be found in the appendix.

2 Copula-based Markov chain models

Throughout the rest of the paper we write $\nabla_{\delta\delta} = \nabla'_{\delta}(\nabla_{\delta}), \nabla_{\theta\delta} = \nabla'_{\theta}(\nabla_{\delta}), \nabla_{\theta\theta} = \nabla'_{\theta}(\nabla_{\theta})$, where ∇ denotes the partial derivative of function with respect to the parameter(-vector) δ or θ. Moreover, vectors or matrices are shown in bold typeface.

This section will briefly review copula-based Markov models. For a deeper understanding of copula models we refer to the textbooks of [Nelsen, 2006], [McNeil et al., 2005] and [Joe, 1997]. A nice overview article of [Härdle & Okhrin, 2010] gives some possible applications of copula models for risk management. First, we give some definitions and restrict ourselves to the two dimensional case. The general case is straightforward.

Definition 1 *A copula is a function $C : [0,1] \times [0,1] \to [0,1]$, such that:*

1. *for every $u, v \in [0,1]$*
$$C(u,0) = 0 = C(0,v)$$
and
$$C(u,1) = u, \text{ and } C(1,v) = v.$$

2. *for every $u_1, u_2, v_1, v_2 \in [0,1]$ with $u_1 \le u_2$, $v_1 \le v_2$ there is*
$$C(u_2,v_2) - C(u_2,v_1) - C(u_1,v_2) + C(u_1,v_1) \ge 0.$$

The second property will be often called the two-increasing property. To sum it up, a copula is a distribution function with uniformly distributed margins. By Sklar's theorem we are able to separate the distribution function into its copula and the marginal distributions.

Theorem 1 *Let F_X and F_Y be the marginal distributions of some real valued, continuous random variables X and Y and G the joint distribution function of (X,Y). Then there exists a copula C such that, for all $(x,y) \in \mathbb{R}^2$:*

$$G(x,y) = C(F_X(x), F_Y(y)). \tag{1}$$

Moreover, if F_X and F_Y are continuous, then C is unique.
Conversely, if F_X and F_Y are the distributions of X and Y, respectively, the function G defined by (1) is a joint distribution function with marginal distributions F_X and F_Y.

Especially part two of the theorem is interesting for simulation or generating new distribution functions by simply combining some univariate distribution functions through copulas. To establish the main result of this section we need the concept of conditional copula functions.

Definition 2 *The conditional copula of V given $U = u$ is defined as:*

$$C_{2|1}(v|u) = P(V \le v | U = u) = \frac{\partial C(u,v)}{\partial u}. \tag{2}$$

A stationary first order Markov chain can be constructed as proposed by [Joe, 1997] p.245. We summarize his explanation in the following theorem.

Theorem 2 *Let $(X_t)_{t\in\mathbb{N}}$ be a stochastic process with absolutely continuous distribution function (cdf from now on) F, i.e. F has density function f. Then $F(x,y) = C(F(x), F(y))$. Let $C_{2|1}(v,u)$ the conditional Copula defined as in (2). Then the conditional cdf is given by:*

$$F(x_t|x_{t-1}) = C_{2|1}(F(x_t)|F(x_{t-1})). \tag{3}$$

We now give some examples for constructing first order Markov chains from a given copula and a given marginal distribution.

2.1 Examples of copula-based Markov models

2.1.1 Clayton copula

The cdf of the bivariate Clayton copula is given for $0 \le \delta < \infty$ by

$$C(u,v;\delta) = (u^{-\delta} + v^{-\delta} - 1)^{-1/\delta}$$

and the conditional copula by

$$C_{2|1}(u|v;\delta) = (1 + u^{\delta}(v^{-\delta} - 1))^{-1-1/\delta}.$$

Kendall's τ can be derived by $\tau = \frac{\delta}{\delta+2}$ and the lower tail dependence coefficient by $\lambda_L = 2^{-1\delta}$, which is increasing in δ. The upper tail coefficient is zero.

2.1.2 Gumbel copula

The bivariate Gumbel copula is defined by

$$C(u,v;\delta) = \exp(-[(-\ln u)^{-\delta} + (-\ln v)^{-\delta}]^{1/\delta}),$$

and the conditional copula by

$$C_{2|1}(u|v;\delta) = u^{-1}\exp\{[(-\ln u)^{\delta} + (-\ln v)^{\delta}]^{1/\delta}\}\left(1 + \left(\frac{\ln u}{\ln v}\right)^{\delta}\right)^{-1+1/\delta}.$$

Kendall's τ is given by $\tau = 1 - 1/\delta$ and contrary to the Clayton copula, the Gumbel copula has $\lambda_L = 0$ and the upper tail coefficient is $\lambda_U = 2 - 2^{1/\delta}$.

2.1.3 Joe-Clayton copula

For asymmetric modeling of both lower and upper tail dependence one may use the Joe-Clayton copula, which is defined by:

$$C(u,v;\delta_1,\delta_2) = 1 - (1 - ((1-\bar{u}^{\delta_1})^{-\delta_2} + (1-\bar{v}^{\delta_1})^{-\delta_2} - 1)^{-1/\delta_2})^{1/\delta_1},$$

for $\delta_1 \geq 0,\ \delta_2 > 0$ and $\bar{u} = 1-u,\ \bar{v} = 1-v$. The conditional copula is given by

$$C_{2|1}(v|u;\delta_1,\delta_2) = (1-w^{-1/\delta_2})^{1/\delta_1 - 1} w^{-1/\delta_2 - 1}(1-\bar{u}^{\delta_1})^{-\delta_2-1}\bar{u}^{\delta_1-1}$$

and $w = ((1-\bar{u}^{\delta_1})^{-\delta_2} + (1-\bar{v}^{\delta_1})^{-\delta_2} - 1)$. The tail dependence coefficients are given by $\lambda_U = 2 - 2^{1/\delta_1}$ and $\lambda_L = 2^{-1/\delta_1}$. Therefore, $\lambda_U \neq \lambda_L$ in general and in contrast to the above mentioned copulas, the Joe-Clayton copula is able to model the asymmetric tail behavior of financial data that is often observed.

2.2 Estimation of copula-based Markov models

We focuss on the IFM method proposed by [Joe, 1997] p.299 ff. for estimating the unknown parameters when we observe an *iid* sample. Given a parametric, copula-based model for the d-dimensional random variable X with absolutely continuous distribution function F, such that:

$$F(\boldsymbol{x};\boldsymbol{\theta},\boldsymbol{\delta}) = C(F_1(x_1,\theta_1),...,F_d(x_d,\theta_d);\boldsymbol{\delta}).$$

The parameters of interest are $\boldsymbol{\theta} = (\theta_1,...,\theta_d) \subset \Omega^{d1}$, a $d1$-dimensional parameter space. Note that $\theta_1,\theta_2,...$ need not to have the same dimension nor need $F_1,...F_d$ be distribution functions of the same type. Let c denote the pdf corresponding to C, then the density of X can be written as:

$$f(\boldsymbol{x};\boldsymbol{\theta},\boldsymbol{\delta}) = c(F_1(x_1,\theta_1),...,F_d(x_d,\theta_d);\boldsymbol{\delta})\prod_{j=1}^{d} f_j(x_j;\theta_j).$$

Denote

$$L_n = \prod_{t=1}^{n} f_t(\boldsymbol{x};\boldsymbol{\theta},\boldsymbol{\delta}) = \prod_{t=1}^{n} c_t(F_1(x_1,\theta_1),...,F_d(x_d,\theta_d);\boldsymbol{\delta})\prod_{j=1}^{d} f_j(x_j;\theta_j).$$

Taking logarithm on both sides, we have

$$LL_n = \underbrace{\sum_{t=1}^{n}\log c_t(F_1(x_1,\theta_1),...,F_d(x_d,\theta_d);\boldsymbol{\delta})}_{=:\,L_{2n}} + \underbrace{\sum_{t=1}^{n}\sum_{j=1}^{n}\log f_j(x_{t,j};\theta_j)}_{=:\,L_{1n}}$$

or in a short hand notation:

$$LL_n = L_{1n} + L_{2n}.$$

The two-stage maximum likelihood estimator can be found the following way: Let $\hat{\boldsymbol{\theta}}$ be solution of the maximization problem

$$\max_{\boldsymbol{\theta}\in\Omega^{d1}} L_{1n}(X_t, \boldsymbol{\theta}). \tag{4}$$

Then denote $\hat{\boldsymbol{\delta}}$ the solution of the second step

$$\max_{\boldsymbol{\delta}\in\Omega^{d2}} L_{2n}(X_t, \hat{\boldsymbol{\theta}}, \boldsymbol{\delta}). \tag{5}$$

For consistency see [White, 1994] theorem 3.10 and theorem 6.11 for the asymptotic distribution of this two-stage estimator. In the context of conditional copula models we refer to [Patton, 2002] p.77 ff.

3 The general score-test for time varying parameters

The score test for testing for time varying parameters in our copula model is based on the score test in [Rao, 1973] p.415 ff. which is actually a LM-Test. We restrict us for reasons of clarity to the case where the copula is specified by two-parameters. An extension to the n-variate parameter vector is straightforward.

Consider the following model for the copula parameter $\delta \in \mathbb{R}^2$:

$$\boldsymbol{\delta}_t = \boldsymbol{\delta} + \boldsymbol{\epsilon}_t = \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} + \begin{pmatrix} \epsilon_{t,1} \\ \epsilon_{t,2} \end{pmatrix}, \tag{6}$$

where $\boldsymbol{\epsilon}_t \overset{iid}{\sim} G(0,\Sigma)$, where G is some distribution function, with $E_G[\boldsymbol{\epsilon}_t] = 0$ and $\Sigma = \begin{pmatrix} \sigma_{\epsilon_1} & 0 \\ 0 & \sigma_{\epsilon_2} \end{pmatrix}$. We test under H_0 whether $\sigma_{\epsilon_{1,1}} = \sigma_{\epsilon_{2,1}} = 0$ against the alternative that at least one σ is greater than zero. The test statistic of the score-test is based on the score-function:

$$\boldsymbol{Z}_0 = \frac{1}{\sqrt{n}} \boldsymbol{S}(\boldsymbol{\theta}, \boldsymbol{\delta})\Big|_{H_0}, \tag{7}$$

where

$$\boldsymbol{S}(\boldsymbol{\theta}, \boldsymbol{\delta}) = \begin{pmatrix} \nabla_{\sigma^2_{\epsilon_1}} LL(\boldsymbol{\theta}, \boldsymbol{\delta}) \\ \nabla_{\sigma^2_{\epsilon_2}} LL(\boldsymbol{\theta}, \boldsymbol{\delta}) \end{pmatrix}.$$

The log-likelihood function $LL(\boldsymbol{\theta}, \boldsymbol{\delta})$ can be derived like in the previous section.

We may derive the test statistics using the standard LM-testing approach, like in [White, 2001] p.77 ff. Under validity of the null hypothesis, $\boldsymbol{Z}_0$ should be near zero. Using the arguments given in [White, 2001] and the assumptions listed in appendix A, we get the following result:

Theorem 3 *Under the null hypothesis and the assumptions A1-A6 listed in appendix A, the following result holds:*

$$n^{-1/2}\boldsymbol{S}(\hat{\theta},\hat{\delta}) \xrightarrow{d} N(0,\boldsymbol{\Sigma}), \tag{8}$$

and

$$n^{-1}\boldsymbol{S}(\hat{\theta},\hat{\delta})'\left(\boldsymbol{\Sigma}^{1/2}\boldsymbol{\Sigma}^{1/2}\right)^{-1}\boldsymbol{S}(\hat{\theta},\hat{\delta}) \xrightarrow{d} \chi^2(2), \tag{9}$$

where

$$\boldsymbol{S}(\hat{\theta},\hat{\delta}) = \nabla_{\boldsymbol{\sigma}^2_{\boldsymbol{\epsilon_1}}} LL(\boldsymbol{\theta},\boldsymbol{\delta})|_{\hat{\theta},\hat{\delta}}$$

and

$$\boldsymbol{\Sigma} = \frac{1}{4}\boldsymbol{\sigma}^2 - \boldsymbol{\sigma}_{\boldsymbol{\delta}}\boldsymbol{I}^{-1}_{\boldsymbol{\delta\delta}}\boldsymbol{\sigma}'_{\boldsymbol{\delta}} + (\boldsymbol{\sigma}_{\boldsymbol{\theta}} + \boldsymbol{\sigma}_{\boldsymbol{\delta}}\boldsymbol{I}_{\boldsymbol{\delta\theta}}\boldsymbol{I}^{-1}_{\boldsymbol{\delta\delta}})\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\theta}}(\boldsymbol{\sigma}_{\boldsymbol{\theta}} + \boldsymbol{\sigma}_{\boldsymbol{\delta}}\boldsymbol{I}_{\boldsymbol{\delta\theta}}\boldsymbol{I}^{-1}_{\boldsymbol{\delta\delta}})'.$$

In addition we set:

$$\begin{aligned}
\boldsymbol{\sigma}^2 &= E[W_t(\boldsymbol{\theta},\boldsymbol{\delta})W_t(\boldsymbol{\theta},\boldsymbol{\delta})'],\\
\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\theta}} &= (\boldsymbol{D}^{-1})\boldsymbol{V}(\boldsymbol{D}^{-1})',\ \textit{with}\\
\boldsymbol{D} &= E\left[\nabla_{\theta\theta}\log f(x_t;\theta)\right]\ \textit{and}\\
\boldsymbol{V} &= E\left[\nabla_\theta \log f(x_t;\boldsymbol{\theta})\nabla'_\theta \log f(x_t;\boldsymbol{\theta})\right] + 2\sum_{k=1}^{\infty} E\left[\nabla_\theta \log f(x_1;\boldsymbol{\theta})\nabla'_\theta \log f(x_{1+k};\boldsymbol{\theta})\right],\\
\boldsymbol{I}_{\boldsymbol{\delta\delta}} &= -E\left[\nabla_{\delta\delta}c(F(x_{t-1},\boldsymbol{\theta}),F(x_t;\boldsymbol{\theta});\boldsymbol{\delta})\right]\\
\boldsymbol{I}_{\boldsymbol{\delta\theta}} &= E\left[\nabla_{\delta\theta}\log c(F_{t-1}(x_{t-1};\boldsymbol{\theta}),F_t(x_t;\boldsymbol{\theta});\boldsymbol{\theta})\right]\\
\boldsymbol{\sigma}_{\boldsymbol{\theta}} &= E\left[\frac{1}{2}W_t(\boldsymbol{\theta},\boldsymbol{\delta})\nabla'_\theta \log c(F(x_{t-1};\boldsymbol{\theta}),F(x_t;\boldsymbol{\theta});\boldsymbol{\delta})\right]\\
\boldsymbol{\sigma}_{\boldsymbol{\delta}} &= E\left[\frac{1}{2}W_t(\boldsymbol{\theta},\boldsymbol{\delta})\nabla'_\delta \log c(F(x_{t-1};\boldsymbol{\theta}),F(x_t;\boldsymbol{\theta});\boldsymbol{\delta})\right].
\end{aligned}$$

Even though the formulas in theorem 3 are quite oblongly and may be confusing, one should have a detailed look at the different parts of $\boldsymbol{\Sigma}$. Like in the common LM-test setting, we actually test whether the constraints $\sigma_{\epsilon_1} = \sigma_{\epsilon_2} = 0$ are binding or not by testing if the Lagrange multiplier λ from the constrained estimation of the model is large enough to reject the null hypothesis of no time variation in the copula-parameter. Thus the first part is just the variance of $\hat{\lambda}$, the estimated Langrangian. Assuming that the parameter vector of interest $\Theta = (\boldsymbol{\theta},\boldsymbol{\delta})$ is separable, the information matrix has a block form say

$$I(\Theta) = \begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix},$$

Thus the variance of $\hat{\lambda}$ that corresponds to the restriction $\sigma_{\epsilon_1} = \sigma_{\epsilon_2} = 0$ is given by inverting I:

$$Var(\hat{\lambda}) = I_{22}(\boldsymbol{\delta}) - I_{21}(\boldsymbol{\delta}) I_{11}^{-1}(\boldsymbol{\delta}) I_{12}(\boldsymbol{\delta}).$$

Therefore, $\frac{1}{4}\boldsymbol{\sigma}^2 - \boldsymbol{\sigma}_{\boldsymbol{\delta}} \boldsymbol{I}_{\boldsymbol{\delta}\boldsymbol{\delta}}^{-1} \boldsymbol{\sigma}'_{\boldsymbol{\delta}}$ corresponds to the variance of $\hat{\lambda}$. The second part is due to the two-stage estimation procedure employed and can be derived using arguments of theorem 6.11 in [White, 1994]. Thus the test proposed by [Abegaz & Naik-Nimbalkar, 2008] and its test statistics can be treated as in the LM-test context, we "just" have to be more careful about the assumptions and restrictions we impose, as we are not dealing with the usual *iid* sample setting, but with Markov processes. Based on the assumptions proposed in appendix A the standard χ^2 asymptotic will hold.

To prove theorem 3, we will proceed in several different steps by making a first order taylor series expansion of $\boldsymbol{S}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\delta}})$ around the true vector $(\boldsymbol{\theta}_0, \boldsymbol{\delta}_0)$:

$$\begin{aligned} n^{-1/2}\boldsymbol{S}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\delta}}) &= n^{-1/2}\boldsymbol{S}(\boldsymbol{\theta}_0, \boldsymbol{\delta}_0) + n^{-1/2}\nabla'_\theta \boldsymbol{S}(\boldsymbol{\theta}_0, \boldsymbol{\delta}_0)(\hat{\boldsymbol{\theta}} - \theta_0) + n^{-1/2}\nabla'_\delta \boldsymbol{S}(\boldsymbol{\theta}_0, \boldsymbol{\delta}_0)(\hat{\boldsymbol{\delta}} - \delta_0) + o_P(1) \\ &\cong \underbrace{n^{-1/2}\boldsymbol{S}(\boldsymbol{\theta}, \boldsymbol{\delta})}_{\text{Lemma 1}} + \underbrace{n^{-1}\nabla'_\theta \boldsymbol{S}(\boldsymbol{\theta}_0, \boldsymbol{\delta}_0)}_{\text{Lemma 4}} \underbrace{n^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)}_{\text{Lemma 2}} + \underbrace{n^{-1}\nabla'_\delta \boldsymbol{S}(\boldsymbol{\theta}_0, \boldsymbol{\delta}_0)}_{\text{Lemma 4}} \underbrace{n^{1/2}(\hat{\boldsymbol{\delta}} - \boldsymbol{\delta}_0).}_{\text{Lemma 3}} \end{aligned}$$

$\cong$ means asymptotic equivalent, see for instance lemma 4.7. in [White, 2001]. The proof of theorem 3 will be split up into the following four lemmas following standard arguments: first we show that $n^{-1}\nabla'_\theta \boldsymbol{S}(\boldsymbol{\theta}_0, \boldsymbol{\delta}_0)$ converges in probability to its expectation (which is a constant) using some suitable law of large numbers, then we prove that the two-stage maximum likelihood estimator $n^{-1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$ will converge in distribution to a normal limit. Combining these results we make use of Slutzky's theorem to establish the asymptotic normality of the score test $\boldsymbol{S}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\delta}})$

Lemma 1 *Under the assumptions A1, A3 and A5 in appendix A we have:*

$$n^{-1/2}\boldsymbol{S}(\boldsymbol{\theta}, \boldsymbol{\delta}) \xrightarrow{d} N(0, 1/4\boldsymbol{\sigma}^2), \tag{10}$$

with $\boldsymbol{\sigma}$ *as in theorem 3.*

The proof can be found in appendix B. Note that lemma 1 provides the convergence law and the asymptotic variance of the actual score function. Whereas the convergence of the next two parts is due to the two-stage maximum likelihood estimation used to calibrate the model.

Lemma 2 *Under the assumptions A1-A4 in appendix A we have:*

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{d} N(0, \boldsymbol{\Sigma}_\theta^{-1}), \tag{11}$$

where $\boldsymbol{\Sigma}_\theta^{-1}$ *is just as in theorem 3.*

Next we establish the asymptotic distribution of the two-stage maximum likelihood estimator for δ:

Lemma 3 *Under the assumptions A1-A4 in appendix A we have:*

$$\sqrt{n}(\hat{\boldsymbol{\delta}} - \boldsymbol{\delta}_0) \xrightarrow{d} N(0, \boldsymbol{\Sigma}_\delta^{-1}), \tag{12}$$

where $\boldsymbol{\Sigma}_\delta^{-1} = \boldsymbol{I}_{\delta\delta}^{-1} + \boldsymbol{I}_{\delta\delta}^{-1}\boldsymbol{I}_{\delta\theta}\boldsymbol{\Sigma}_\theta^{-1'}\boldsymbol{I}_{\delta\theta}\boldsymbol{I}_{\delta\delta}^{-1'}$.

The last lemma provides that $n^{-1}\nabla_\theta \boldsymbol{S}(\boldsymbol{\theta}, \boldsymbol{\delta})$ and $n^{-1}\nabla_\delta \boldsymbol{S}(\boldsymbol{\theta}, \boldsymbol{\delta})$ will converge a.s. to the expectation of the hessian evaluated at the true parameter vector $(\boldsymbol{\theta}_0, \boldsymbol{\delta}_0)$:

Lemma 4 *Under the assumptions A1 and A3 in appendix A we have:*

1. $n^{-1}\nabla_\theta \boldsymbol{S}(\boldsymbol{\theta}, \boldsymbol{\delta}) \xrightarrow{a.s.} -\boldsymbol{\sigma}_\theta$
2. $n^{-1}\nabla_\delta \boldsymbol{S}(\boldsymbol{\theta}, \boldsymbol{\delta}) \xrightarrow{a.s.} -\boldsymbol{\sigma}_\delta$

Combining these results yields the proof of theorem 3.

4 Simulation study

To investigate the asymptotic power of the proposed test we carried out a simulation study with different copulas and parameter constellations. In this paper we focus on the two-parametric Joe-Clayton copula. For conclusions on the score-test with a one-parametric copula we refer to [Abegaz & Naik-Nimbalkar, 2008]. To generate observations $\{x_t : t = 1, \ldots, n\}$ following a first order Markov chain with a given copula $C(u_{t-1}, u_t; \delta)$ and margins $F(x_t; \theta)$ we used the algorithm in [Abegaz & Naik-Nimbalkar, 2008].
After estimating the parameters $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\delta}}$ from the obtained observations, the test statistic can be computed. We estimated each component of the variance-covariance matrix $\boldsymbol{\Sigma}$ in theorem 3 consistently. For the estimation of the component $\boldsymbol{V}$ we follow [Abegaz & Naik-Nimbalkar, 2008] by taking the following window estimator:

$$\boldsymbol{V} = \frac{1}{n}\left[\sum_{t=1}^{n} \nabla_\theta \log f(x_t; \hat{\boldsymbol{\theta}}) \nabla'_\theta \log f(x_t; \hat{\boldsymbol{\theta}}) + \sum_{t=k+1}^{b_n} d_n(k) \left(\sum_{t=k+1}^{n} \nabla_\theta \log f(x_t; \hat{\boldsymbol{\theta}}) \nabla'_\theta \log f(x_{t-k}; \hat{\boldsymbol{\theta}}) \right)\right].$$

d_n is a weight function with the Bartlett kernel $d_n(k) = 1 - (k/b_n + 1), k = 1, 2, \ldots, b_n$. Where b_n is a sequence of real, positive numbers, with $b_n \to \infty$ and $b_n/n^{1/4} \to 0$ if $n \to \infty$.

We generated 500, 1000 and 1500 observations from the Joe-Clayton copula with normal margins and did 1000 replications to investigate the finite sample properties of the test, especially the test power, i.e. 1-β, where β is frequency how often the null hypothesis of no dynamic is not rejected given the alternative is true. $\boldsymbol{\theta} = (\mu, \sigma) = (-3, 0.5)$ and $(1, 2)$ are chosen for the parameter of the normal distribution. The copula parameter $\boldsymbol{\delta}$ follows (6) and ϵ_t is bivariate log-normal distributed with mean 0 and variance-covariance matrix Σ. We chose $\boldsymbol{\delta} = (\delta_1, \delta_2) = (1, 0.5)$ and $(1.5, 1)$. To include the strength of variation in $\boldsymbol{\delta_t}$ we increased $\boldsymbol{\sigma}_\epsilon$ step by step, that is $\sigma_{\epsilon_1} \in (0, 2)$ and $\sigma_{\epsilon_2} \in (0, 1)$. Note that this type of variation is much smaller than in the article of [Abegaz & Naik-Nimbalkar, 2008] where $\sigma_\epsilon = (0, 25)$. A variation that is that big may be seen with pure looks, so we follow the question whether the test is also able to detect very small deviation from the constancy hypothesis. For $\sigma_{\epsilon_1} = 0$ and $\sigma_{\epsilon_2} = 0$ the α error (type I error) is obtained, because then $\boldsymbol{\delta_t}$ is constant. The results are presented in table 4 and for one parameter constellation in figure 1. As can be seen the asymptotic power is affected by the number of observations and the variation of $\boldsymbol{\delta_t}$. The α error lies between 5% and 7%. We see, that the significance level is achieved even in small samples (n=500). The test power increases with more observations and higher variation in the dynamic model. For small sample size (n=500) the selectivity of the test is only acceptable for $\sigma_{\epsilon_1} = 2$ and $\sigma_{\epsilon_2} = 1$ with a β error of 13% to 26%. For $n = 1000$ instead, the asymptotic distribution of the test seems to hold, even when the σs are smaller. When variations are small, e.g. $\sigma_{\epsilon_1} = 0.25$ and $\sigma_{\epsilon_2} = 0.125$ then the test doesn't detect this deviation from constancy of the copula parameter. We conclude, that the test holds the significance level and has reasonable power at least when sample size is large (n=1000) and the variation is not too small ($\sigma_{\epsilon_1} \approx 1.25$ and $\sigma_{\epsilon_2} \approx 0.5$).

5 Empirical analysis

In this section we investigate the potential of the dynamic copula-based Markov model for financial data compared to a usual GARCH(1,1) model.

5.1 Preliminary analysis

We chose daily log returns of Commerzbank from 18th April 2001 to 31th March 2010. To get an overview of the data we did some descriptive statistics and the KPSS test on stationarity and Jarque-Bera test for normality. As can be seen in table 1 the observations are skewed, leptokurtic, stationary and not normally distributed.

Table 1: Descriptive statistics of Commerzbank daily log returns

n	mean	std.dev.	skewness	kurtosis	p-value of KPSS	p-value of JB
1877	-10^{-5}	0.0320	-0.3939	13.1125	0.07688	10^{-16}

Since we are modeling dependence structures we computed some common dependence measures: the linear correlation ρ, Spearman's Rho ρ_S and Kendall's Tau τ. Following [Cont, 2001] we use $\rho_{[\alpha]} = \rho[|X_{t-1}|^\alpha, |X_t|^\alpha]$ as a measure of nonlinear dependence. For $\alpha = 2$ volatility clustering, already mentioned in the introduction, can be measured. The results are presented in table 2.

Table 2: Intertemporal dependence in Commerzbank daily log returns

ρ	ρ_S	τ	$\rho_{[1]}$	$\rho_{[2]}$
0.0791	0.0216	0.0134	0.3936	0.2475

The log returns show a slight slight positive correlation regarding *rho* ρ_S and τ. The nonlinear measures show a higher amount and especially for $\alpha = 2$ we can assume volatility clustering. This is not surprising because figure 2 pictures this phenomenon. Because GARCH models take the volatility clustering into account, our approach to compare the copula-based Markov model with a GARCH(1,1) model is supported.

5.2 Parameter estimation

We assumed the IFM method, thus the parameters are estimated in two steps.
First, we fit the marginal distribution and follow two approaches:

Parametric and empirical distribution
We fit a hyperbolic distribution

$$f(x) = \frac{\psi^2 - \eta^2}{2\psi\sigma K_1(\sigma\sqrt{\psi^2 - \eta^2})} \exp\left(-\psi\sqrt{\sigma^2 + (x-\mu)^2} + \eta(x-\mu)\right)$$

with $\psi > 0$, $0 \leq |\eta| < \psi$, $\mu \in \mathbb{R}$, $\sigma \geq 0$ and K_1 is the modified bessel function of 2nd order. As benchmark we use the gaussian distribution. The reason for choosing the hyperbolic distribution is due its ability in modeling skewness as well as heavy tails, which is a well-known so-called stylized fact of financial returns. The hyperbolic distribution has been investigated in financial market models as by [Jaschke, 2000],[Eberlein & Keller, 1995] or [Reimann, 2005] among others. Moreover, [Eberlein & Keller, 1995] found evidence for stock returns to follow a hyperbolic distribution. In addition to the full parametric approach, a semi-parametric is chosen, where the marginal distributions are estimated using the empirical distribution $F_n(x) = \frac{1}{n}\sum_{t=1}^{n} \mathbb{1}_{(-\infty,x]}(X_t)$ can be applied. This encompasses the approach of [Genest et al., 1995].

GARCH
A very well documented stylized fact for financial returns are volatility clusters. [Engle, 1982]

and [Bollerslev, 1986] suggested ARCH resp. GARCH models to capture this phenomena. Therefore, we adapt a standard GARCH(1,1) with gaussian innovations to the data and compare its performance with the copula-based Markov approach. Note that a GARCH(1,1) is clearly not Markovian, but a martingale difference. So we compare en passant two different types of stochastic processes, namely Markov chains and martingale differences.
The GARCH(1,1) process can be defined by $x_t|\mathcal{F}_{t-1} \sim \mathcal{N}(0, \sigma_t^2)$ with $\sigma_t^2 = \omega + \alpha x_{t-1}^2 + \beta\sigma_{t-1}^2$ and $\omega > 0, \alpha, \beta \geq 0$.
The residuals $u_t = \sigma_t^{-1} x_t$ with $u_t \sim \mathcal{N}(0, 1)$ are then GARCH-filtered log returns.

We proceed with the estimation of the copula parameters. For the hyperbolic margins and the margins estimated with the empirical distribution function the score test for dynamic copula parameters is performed. For the GARCH filtered innovations we also fit our copula-based Markov model and we test whether there is still some dynamics in the dependence structure. If this were the case, we would have evidence that there is time-variation not only in the conditional variances but also in other moments. If instead the null cannot be rejected, we can conclude that all sort of time variation is already captured by the GARCH(1,1)-model. The only way a copula may now help in modeling the filtered GARCH(1,1)-residuals is just due to the fact, that we have assumed gaussian innovations, which is usually not convenient. Instead other distributions like skewed t-distributions as proposed in [Hansen, 1994] or [Chen, 2007] would be preferable. But the focus lies not on modeling the conditional distribution as exact as possible, but to elaborate the benefit of dynamic copula-based Markov models over the standard GARCH(1,1) models.

If the null hypothesis of constant copula parameters is rejected, we model a dynamic copula parameter. For the score-test we assumed the variation in 6. The aim was to expose any kind of variation. In a next step we model $\boldsymbol{\delta}_t$ as a modified ARMA(1,k) process, as proposed in [Abegaz & Naik-Nimbalkar, 2008]:

$$\boldsymbol{\delta}_t = \exp\left(\boldsymbol{\omega} + \boldsymbol{\alpha}\log(\boldsymbol{\delta}_{t-1}) + \boldsymbol{\beta}\frac{1}{k}\sum_{i=1}^{k}|u_{t-1} - u_{t-i-1}|\right).$$

This approach includes an autoregressive term $\boldsymbol{\delta}_{t-1}$ and an error term for the mean absolute difference between u_t and u_{t-1}, which captures variation in the dependence structure. The u_ts are estimated by $\hat{u}_t = F(x_t; \hat{\boldsymbol{\theta}})$. $\boldsymbol{\delta}_1$ is assumed to be constant.

5.3 Results

The results for Commerzbank daily stock returns are summarized in table 5. Regarding to the Bayesian Information Criterion (BIC) for the marginal distributions it can be seen that the hyperbolic distribution performs better than the gaussian one. The GARCH(1,1)

model achieves the best fit due to BIC.
For the GARCH residuals we did some further investigation and tested them for autocorrelation with a Ljung-Box test and for normality with a Jarque-Bera test (see table 3). If we had significance against the Null of no serial correlation, we would still have some relevant information in our model, and thus we just re-identify our GARCH-model, by taking higher orders or an AR(p)-process for the conditional mean. But this would contradict economic theory, where the efficient market and the rational expectations hypothesis of the financial actors contains that the conditional return of an asset should be zero, i.e. $E[R_t|\mathfrak{F}_{t-1}] = 0$, because otherwise there is a systematic information about the behavior of the stock return and everyone will buy (if we had positive trend) or sell the asset. The null hypothesis of no correlation can be rejected for assuming different lags. As our times series includes 2363 observations we report the results of the Ljung-Box test of order 20 in table 3. Also the residuals are not normal distributed, the main focuss lies on the non-presence of autocorrelation in the residuals, the absolute values of the residuals and the squared residuals. Thus we can assume that the filtered residuals are not following a gaussian white noise process, but they are white noise and that the conditional distribution of the GARCH(1,1) model is misspecfied. The fit could have been improved by assuming student-t distributed residuals or a higher order GARCH process. To simplify matters we will not follow this.

Table 3: Tests of the GARCH residuals

p-value of Ljung-Box test			p-value of Jarque-Bera test
u_t	$\lvert u_t \rvert$	u_t^2	u_t
0.3353	0.392	0.7927	0.0001

Next, we estimated the parameters of the copulas, proposed in section 2.1. For the gaussian and the filtered GARCH innovations the clayton and for the hyperbolic and empirical distribution the Joe-Clayton copula performs best, regarding BIC. So our generalization of the score-test to two-parametric copulas seems to be helpful. The Gumbel copula performs worst and is inappropriate for our data set. We also computed Kendall's Tau from the estimated copula parameters $\boldsymbol{\delta}$. Next we applied the generalized score-test for the Joe-Clayton copula and the Clayton copula. The null hypothesis of constant copula parameters is rejected for all marginal distributions except the gaussian and the filtered GARCH innovations. Modeling dynamic copulas improves the fit in both other cases. To sum up we compare the dynamic Joe-Clayton copula with hyperbolic margins and 10 parameters, with an added up BIC of $-8,139.4$, to a simple GARCH(1,1) model with three parameters and a BIC of $-8,497.6$. It gets clear that there is no advantage of the (dynamic) copula-based Markov chain model for the log returns. Our approach to improve the usual GARCH(1,1) model by applying the copula-based Markov model on the residuals, brings few improvement of the fit. Note that the choice of the Clayton copula corresponds to the stylized fact, that (extreme) negative returns are more likely then positives. As the Clayton copula covers negative tail dependence, it can be concluded that the filtered GARCH residuals still exhibits significant lower tail dependence which encompasses the aforementioned stylized fact. Note that this phenomena could have been captured by taking a skewed conditional distribution in our GARCH model.

6 Conclusion

In our paper we gave a short review on dynamic copula-based Markov models and generalized the score-test proposed by [Abegaz & Naik-Nimbalkar, 2008]. The null hypothesis of no time variation in the parameters of our copula models is extended to variation in at least one of the possible multidimensional parameter sets. The significance level is maintained when a Joe-Clayton copula, which has two parameters, is investigated. The test power increases with more observations, but it is general lower than in one-dimensional parameter case. This result is not astonishing as we would reject the null, when there is time variation in at least one parameter. To show that the generalization is useful, we modeled the daily returns of the Commerzbank stock using different copulas. We found evidence that the very flexible Joe-Clayton copula outperforms the other one parametric copula models. After estimating the margins with hyperbolic distribution as well as the gaussian and the empirical distribution, the possible dynamics in the parameter is investigated. We see that the null is rejected for all copulas except for the gaussian and the. In a next step we set up as a benchmark model a standard GARCH(1,1) model with gaussian innovations. Analyzing the residuals, we found no evidence against white noise, but the distribution still exhibits skewness. This may be captured by adapting a more realistic distribution like a skewed-t or hyperbolic. The main advantage of the standard GARCH(1,1) lies in the fact that it is numerical preferable as it is fast and quite easy to estimate, has less parameters compared with dynamic copula-based Markov models. Moreover, it is able to capture volatility clusters and its easy applicability to VaR calculations, portfolio optimization and option-pricing. Therefore, the question may arise, whether Markov chains are suitable models for stock returns at all or that martingale difference equation like GARCH(1,1) are more able to capture stylized facts that stock returns exhibit.

Mathematical appendix

A: Assumptions

Following [Abegaz & Naik-Nimbalkar, 2008] and [White, 1994], resp. [Patton, 2002] p.112 ff. we post the following assumptions:

A1. The process $(X_t)_{t\in\mathbb{N}}$ is stationary and α-mixing, with mixing coefficient $\alpha(n)$, such that $\sum_{n=1}^{\infty}\alpha(n)^{\frac{\beta}{2(2+\beta)}} < \infty$, for $\beta > 0$.

A2. $(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\delta}})$ is the two-stage maximum likelihood estimator and thus solution of the maximization problem (4) resp. (5).

A3. (a) $f(x_t|x_{t-1};\Theta) > 0$ P-a.s. independent of Θ, and twice continuously differentiable on Ω.

(b) The copula densities are twice continuously differentiable on Ω^{d2}.

(c) There exists neighborhoods U_θ and U_δ such that we have for all $\boldsymbol{\theta} \in U_\theta \subset \Omega^{d1}$ and $\boldsymbol{\delta} \in U_\delta \Omega^{d2}$:

i. $E\left[\sup_{\boldsymbol{\theta}} |\nabla_\theta f(x_t;\boldsymbol{\theta})|\right] < \infty$, $E\left[\sup_{\boldsymbol{\delta}} |\nabla_\delta c(u_t, u_{t-1};\boldsymbol{\delta})|\right] < \infty$

ii. $E\left[\sup_{\boldsymbol{\theta}} |\nabla_{\theta\theta} f(x_t;\boldsymbol{\theta})|\right] < \infty$, $E\left[\sup_{\boldsymbol{\delta}} |\nabla_{\delta\delta} c(u_t, u_{t-1};\boldsymbol{\delta})|\right] < \infty$

iii. $E\left[\sup_{\boldsymbol{\theta}} |\nabla_{\theta\theta} \log f(x_t;\boldsymbol{\theta})|\right] < \infty$, $E\left[\sup_{\boldsymbol{\delta}} |\nabla_{\delta\delta} \log c(u_t, u_{t-1};\boldsymbol{\delta})|\right] < \infty$

A4. For all $\boldsymbol{\theta} \in U_\theta \subset \Omega^{d1}$ and $\boldsymbol{\delta} \in U_\delta \subset \Omega^{d2}$ we have, that $-E[\nabla_{\theta\theta}\log f(x_t;\boldsymbol{\theta})]$, $Var(\nabla_\theta \log f(x_t;\boldsymbol{\theta}))$, $-E[\nabla_{\delta\delta}\log c(u_t, u_{t-1};\boldsymbol{\delta})]$, $Var(\nabla_\delta \log c(u_t, u_{t-1};\boldsymbol{\delta}))$ are $O(1)$ and uniformly positive definite.

A5. $E[|W_{t,j}(\boldsymbol{\theta},\boldsymbol{\delta})|^{2+\beta}] < \infty$ for $\beta > 0$, $j = 1, 2$ and all $t \geq 1$ and

$$\lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} E[W_{t,j}^2(\boldsymbol{\theta},\boldsymbol{\delta})|\mathfrak{F}_{t-1}] = \sigma_j^2 > 0, \text{ a.s.}$$

Note the minor modification compared to [Abegaz & Naik-Nimbalkar, 2008], especially the assumptions in A3 that are needed to ensure consistency of the the two-stage maximum-likelihood estimator are less strong than in for instance [Joe, 1997] p.318. To obtain locally asymptotic normally distributed estimators the twice continuously differentiability assumption on the densities is sufficient, see [Ferguson, 1996] p.119 ff. The uniformly integrability for the score functions and the densities are needed, on the one hand side to make use of the weak law of large numbers for α- mixing processes, see for instance [White, 1984], and to interchange differentiation and integration, see [Ferguson, 1996] p.124.

B: Proofs

Proof of lemma 1 First note that $\boldsymbol{S}(\boldsymbol{\theta},\boldsymbol{\delta})$ is a martingale under A1 and A3 and that $\boldsymbol{S}(\boldsymbol{\theta},\boldsymbol{\delta}) = \frac{1}{2}\sum_{t=2}^{n} W_t(\boldsymbol{\theta},\boldsymbol{\delta})$. With A5 we employ the central limit theorem for stationary martingales in [Basawa & Rao, 1980] p.388 to conclude, that

$$n^{-1/2}\sum_{t=2}^{n} W_t(\boldsymbol{\theta},\boldsymbol{\delta}) \xrightarrow{d} N(0,\boldsymbol{\sigma}^2),$$

where $\boldsymbol{\sigma}^2 = E[W_t(\boldsymbol{\theta},\boldsymbol{\delta})W_t(\boldsymbol{\theta},\boldsymbol{\delta})']$. Finally we have:

$$\boldsymbol{S}(\boldsymbol{\theta},\boldsymbol{\delta}) = \frac{1}{2}\sum_{t=2}^{n} W_t(\boldsymbol{\theta},\boldsymbol{\delta}) \xrightarrow{d} N(0,1/4\boldsymbol{\sigma}^2).$$

□

Proof of lemma 2 Denote the maximum likelihood estimator by $\hat{\boldsymbol{\theta}}$ obtained in a first step by maximizing (4) as the solution of the corresponding score equation $\boldsymbol{S}_\theta(\theta) = 0$. Then we have under the conditions A3, that $\frac{1}{n}\boldsymbol{S}_\theta(\hat{\boldsymbol{\theta}}) \overset{a.s.}{\to} E[\boldsymbol{S}_\theta(\hat{\boldsymbol{\theta}})] = 0_{d1\times 1}$, where $0_{d1\times 1}$ is a vector of zeros with length $d1$. Making a first order Taylor series expansion of $\boldsymbol{S}(\hat{\boldsymbol{\theta}})$ around the true parameter vector $\boldsymbol{S}(\boldsymbol{\theta}_0)$ we get:

$$\begin{aligned}\boldsymbol{S}_\theta(\hat{\boldsymbol{\theta}}) &= \boldsymbol{S}_\theta(\boldsymbol{\theta}_0) + \nabla_\theta \boldsymbol{S}_\theta(\boldsymbol{\theta}_0)(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0) + o_P(1)\\ \frac{1}{n}\boldsymbol{S}_\theta(\hat{\boldsymbol{\theta}}) &\cong \frac{1}{n}\boldsymbol{S}_\theta(\boldsymbol{\theta}_0) + \frac{1}{n}\nabla_\theta \boldsymbol{S}_\theta(\boldsymbol{\theta}_0)(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0)\end{aligned}$$

Rearranging leads to:

$$n^{1/2}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0) = \left(-\frac{1}{n}\nabla_\theta \boldsymbol{S}_\theta(\boldsymbol{\theta}_0)\right)^{-1} n^{-1/2}\boldsymbol{S}_\theta(\boldsymbol{\theta}_0).$$

For the first equation we have

$$\left(-\frac{1}{n}\nabla_\theta \boldsymbol{S}_\theta(\boldsymbol{\theta}_0)\right)^{-1} \overset{a.s.}{\to} E[\nabla_\theta \boldsymbol{S}_\theta(\boldsymbol{\theta}_0)]^{-1} = \boldsymbol{D}^{-1},$$

which is just the inverse of hessian of the the log-likelihood problem (4). Due to assumption A4 the matrix $\boldsymbol{D}$ is invertible at least in U_θ. The second term $n^{-1/2}\boldsymbol{S}_\theta(\boldsymbol{\theta}_0)$ is just a continuous function of an α-mixing process and thus is itself α-mixing, see [Davidson, 1994] theorem 14.1. For completeness we present a result due to [Denker, 1986]:

Theorem 4 *Let $(X_n)_{n\in\mathbb{N}}$ be strictly stationary α-mixing sequence, i.e. $E[X_1] = 0$ and $E[|X_1|^{2+\delta}] < \infty$ with mixing coefficient $\alpha(n)$, s.t. $\sum_{n=1}^{\infty}\alpha(n)^{\frac{\delta}{2(2+\delta)}} < \infty$ and $\lim_{n\to\infty}\sigma_n^2 = \infty$. Set $S_n = \sum_{i=1}^{n} X_i$, then*

$$n^{1/2}S_n/\sigma_n \xrightarrow{d} N(0,\sigma^2),$$

where $\sigma^2 = E[X_1^2] + 2\sum_{n=1}^{\infty} E[X_1X_{1+n}]$, iff the sequence $\{S_n^2/\sigma_n^2, n \geq 1\}$ is uniformly integrable.

Under the assumptions imposed on the process $(X_t)_{t\in\mathbb{N}}$ we have that

$$\boldsymbol{S}_\theta(\boldsymbol{\theta}_0) = \sum_{t=1}^{n} \nabla_\theta \log f(x_t;\boldsymbol{\theta}_0)$$

fulfills the the assumptions of theorem 4, with $E[X_1^2] = E[\nabla_\theta \log f(x_1;\boldsymbol{\theta}_0)\nabla'_\theta \log f(x_1;\boldsymbol{\theta}_0)]$ and $\sum_{n=1}^{\infty} E[X_1 X_{1+n}] = \sum_{k=1}^{\infty} E[\nabla_\theta \log f(x_1;\boldsymbol{\theta}_0)\nabla'_\theta \log f(x_{1+k};\boldsymbol{\theta}_0)]$. Summarizing we have

$$n^{-1/2}\boldsymbol{S}_\theta(\boldsymbol{\theta}_0) \xrightarrow{d} N(0,V),$$

where

$$V = E[\nabla_\theta \log f(x_1;\boldsymbol{\theta}_0)\nabla'_\theta \log f(x_1;\boldsymbol{\theta}_0)] + 2\sum_{k=1}^{\infty} E[\nabla_\theta \log f(x_1;\boldsymbol{\theta}_0)\nabla'_\theta \log f(x_{1+k};\boldsymbol{\theta}_0)].$$

We finally get the result of lemma 2 by applying Slutzky's theorem.

□

Proof of lemma 3 Lemma 3 can be proven in exactly the same way as lemma 2 or seen as a direct consequence of theorem 6.11 in [White, 1994], therefore, it is omitted here. For a detailed proof see [Reichert, 2010] p.46 ff. If the model is correctly specified, the result for the asymptotic variance of the two-stage estimator simplifies to the form given in lemma 3. Also note that our parameter-vector $\Theta = (\boldsymbol{\theta}, \boldsymbol{\delta})$ is separable in the sense, that there is no dependency between $\boldsymbol{\theta}$ and $\boldsymbol{\delta}$, therefore, $E[S_\theta S_\delta] = 0$.

□

Proof of lemma 4 First we have for $i = 1, ..., d_1$ and $j = 1...d_2$:

$$\frac{1}{n}\frac{\partial S(\boldsymbol{\theta},\boldsymbol{\delta})_j}{\partial \theta_i} = \frac{1}{n}\left[\sum_{t=2}^{n} \frac{\partial^2}{\partial\theta_i \partial\epsilon_{1,j}^2} \underbrace{\log\left(c(u_{t-1},u_t;\boldsymbol{\delta}) + \frac{1}{2}\sum_{j=1}^{d_2}\frac{\partial^2 c(u_{t-1},u_t;\boldsymbol{\delta})}{\partial\delta_j^2}\sigma_{\epsilon_{1,j}}\right)}_{:= N_t(u_{t-1},u_t;\boldsymbol{\delta})}\right]_{\sigma_{\epsilon_1}^2=0}.$$

We have

$$E\left[-\frac{\partial^2 N_t(u_{t-1},u_t;\boldsymbol{\delta})}{\partial\delta_i\partial\sigma_{\epsilon_{1,j}}^2}\right]_{\sigma_{\epsilon_1}^2=0} = E\left[\frac{\partial N_t(u_{t-1},u_t;\boldsymbol{\delta})}{\partial\theta_i}\frac{\partial N_t(u_{t-1},u_t;\boldsymbol{\delta})}{\partial\sigma_{\epsilon_{1,j}}^2}\right]_{\sigma_{\epsilon_1}^2=0}.$$

Some calculation yields:

$$\begin{aligned}\frac{\partial N_t(u_{t-1},u_t;\boldsymbol{\delta})}{\partial\sigma_{\epsilon_{1,j}}^2} &= \frac{1}{2}W_{t,j}(\boldsymbol{\theta},\boldsymbol{\delta})\\ \frac{\partial N_t(u_{t-1},u_t;\boldsymbol{\delta})}{\partial\theta_i} &= \nabla_{\theta_i}\log c_t(\boldsymbol{\theta}).\end{aligned}$$

Using the ergodic theorem we can finally conclude that:

$$\begin{aligned}
\lim_{n\to\infty}\frac{1}{n}\frac{\partial S(\boldsymbol{\theta},\boldsymbol{\delta})_j}{\partial\theta_i} &= E\left[\frac{\partial N_t(u_{t-1},u_t;\boldsymbol{\delta})}{\partial\theta_i}\frac{\partial N_t(u_{t-1},u_t;\boldsymbol{\delta})}{\partial\sigma^2_{\epsilon_{1,j}}}\right]_{\sigma^2_{\epsilon_1}=0}\\
&= E\left[-\frac{1}{2}W_{t,j}(\boldsymbol{\theta},\boldsymbol{\delta})\nabla'_{\theta_i}\log c_t(\boldsymbol{\theta})\right]_{\sigma^2_{\epsilon_1}=0}\\
&:= \sigma_{\theta,ij}
\end{aligned}$$

And thus

$$\lim_{n\to\infty}\frac{1}{n}\frac{\partial S(\boldsymbol{\theta},\boldsymbol{\delta})}{\partial\boldsymbol{\theta}} = \boldsymbol{\sigma}_{\boldsymbol{\theta}},$$

where $\boldsymbol{\sigma}_{\boldsymbol{\theta}}$ is just as in theorem 3.

The second assertion of lemma 4 can be proven in exactly the same manner.

□

Proof of theorem 3 We have for $n\to\infty$:

$$n^{-1/2}\boldsymbol{S}(\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\delta}}) = n^{-1/2}\boldsymbol{S}(\boldsymbol{\theta}_0,\boldsymbol{\delta}_0) + \frac{1}{n}\nabla_\theta\boldsymbol{S}(\boldsymbol{\theta}_0,\boldsymbol{\delta}_0)n^{1/2}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0) + \frac{1}{n}\nabla_\delta\boldsymbol{S}(\boldsymbol{\theta}_0,\boldsymbol{\delta}_0)n^{1/2}(\hat{\boldsymbol{\delta}}-\boldsymbol{\delta}_0).$$

With the results from lemma 2-4 we have:

$$\begin{aligned}
&= n^{-1/2}\boldsymbol{S}(\boldsymbol{\theta}_0,\boldsymbol{\delta}_0) - \boldsymbol{\sigma}_\theta D^{-1}n^{-1/2}\boldsymbol{S}_\theta(\boldsymbol{\theta}_0) - \boldsymbol{\sigma}_\delta\boldsymbol{I}^{-1}_{\delta\delta}n^{-1/2}\boldsymbol{S}_\delta(\boldsymbol{\theta}_0,\boldsymbol{\delta}_0) - \boldsymbol{\sigma}_\delta\boldsymbol{I}^{-1}_{\delta\delta}\boldsymbol{I}_{\delta\theta}D^{-1}n^{-1/2}\boldsymbol{S}_\theta(\boldsymbol{\theta}_0)\\
&= n^{-1/2}\boldsymbol{S}(\boldsymbol{\theta}_0,\boldsymbol{\delta}_0) - \boldsymbol{\sigma}_\delta\boldsymbol{I}^{-1}_{\delta\delta}n^{-1/2}\boldsymbol{S}_\delta(\boldsymbol{\theta}_0,\boldsymbol{\delta}_0) - (\boldsymbol{\sigma}_\theta + \boldsymbol{\sigma}_\delta\boldsymbol{I}^{-1}_{\delta\delta}\boldsymbol{I}_{\delta\theta})\boldsymbol{D}^{-1}n^{-1/2}\boldsymbol{S}_\theta(\boldsymbol{\theta}_0).
\end{aligned}$$

Now we have to calculate the covariances between $\boldsymbol{S}(\boldsymbol{\theta}_0,\boldsymbol{\delta}_0), \boldsymbol{S}_\theta(\boldsymbol{\theta})$ and $\boldsymbol{S}_\delta(\boldsymbol{\theta}_0,\boldsymbol{\delta}_0)$. First, as there is no dependency between $\boldsymbol{\delta}$ and $\boldsymbol{\theta}$ we have $Cov(\boldsymbol{S}_\theta,\boldsymbol{S}_\delta)=0$. Moreover, we have:

$$\begin{aligned}
Cov[\boldsymbol{S}(\boldsymbol{\theta}_0,\boldsymbol{\delta}_0),\boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{\theta}_0)] &= E[\boldsymbol{S}(\boldsymbol{\theta}_0,\boldsymbol{\delta}_0)\boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{\theta}_0)'] - E[\boldsymbol{S}(\boldsymbol{\theta}_0,\boldsymbol{\delta}_0)]E[\boldsymbol{S}_{\boldsymbol{\theta}_0}(\boldsymbol{\theta}_0)']\\
&= \sum_{t=2}^{n}E\left[\frac{1}{2}\boldsymbol{W}_t(\boldsymbol{\theta}_0,\boldsymbol{\delta}_0)\nabla_\theta\log f(x_t;\boldsymbol{\theta}_0)'\right]\\
&- \sum_{t=2}^{n}E\left[\frac{1}{2}\boldsymbol{W}_t(\boldsymbol{\theta}_0,\boldsymbol{\delta}_0)\right]\sum_{t=2}^{n}E\left[\nabla_\theta\log f(x_t;\boldsymbol{\theta}_0)'\right]\\
&= \boldsymbol{0}.
\end{aligned}$$

Note that $E[\boldsymbol{W}_t(\boldsymbol{\theta}_0, \boldsymbol{\delta}_0)\nabla'_\theta \log f(x_t; \theta)] = 0$, because for every combination of $\{r, s, t \in 0, 1, ...\}$ of $E[\boldsymbol{W}_{r,s}(\boldsymbol{\theta}_0, \boldsymbol{\delta}_0)\nabla'_\theta \log f(x_t; \theta)] = 0$ we have:

$$
\begin{aligned}
&\int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}} \tfrac{1}{c(F(x_r;\boldsymbol{\theta}),F(x_s;\boldsymbol{\theta});\boldsymbol{\delta})} \tfrac{\partial^2 c(F(x_r;\boldsymbol{\theta}),F(x_s;\boldsymbol{\theta});\boldsymbol{\delta})}{\partial \boldsymbol{\delta}^2} \tfrac{\partial \log f(x_t,\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} f(x_r, x_s, x_t; \boldsymbol{\theta}, \boldsymbol{\delta}) dx_r dx_s dx_t \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}} \tfrac{1}{c(F(x_r;\boldsymbol{\theta}),F(x_s;\boldsymbol{\theta});\boldsymbol{\delta})} \tfrac{\partial^2 c(F(x_r;\boldsymbol{\theta}),F(x_s;\boldsymbol{\theta});\boldsymbol{\delta})}{\partial \boldsymbol{\delta}^2} \tfrac{1}{f(x_t;\boldsymbol{\theta})} \tfrac{\partial f(x_t;\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \\
&\times f(x_r;\boldsymbol{\theta}) f(x_s;\boldsymbol{\theta}) f(x_t;\boldsymbol{\theta}) c(F(x_r;\boldsymbol{\theta}), F(x_s;\boldsymbol{\theta});\boldsymbol{\delta}) c(F(x_s;\boldsymbol{\theta}), F(x_t;\boldsymbol{\theta});\boldsymbol{\delta}) dx_r dx_s dx_t \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}} \tfrac{\partial^2 c(F(x_r;\boldsymbol{\theta}),F(x_s;\boldsymbol{\theta});\boldsymbol{\delta})}{\partial \boldsymbol{\delta}^2} \tfrac{\partial f(x_t,\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} f(x_r;\boldsymbol{\theta}) f(x_s;\boldsymbol{\theta}) c(F(x_s;\boldsymbol{\theta}), F(x_t;\boldsymbol{\theta});\boldsymbol{\delta}) dx_r dx_s dx_t \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}} \tfrac{\partial f(x_t;\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} c(F(x_s;\boldsymbol{\theta}), F(x_t;\boldsymbol{\theta});\boldsymbol{\delta}) f(x_s;\boldsymbol{\theta}) \left\{ \tfrac{\partial^2 c(F(x_r;\boldsymbol{\theta}),F(x_s;\boldsymbol{\theta});\boldsymbol{\delta}) f(x_r;\boldsymbol{\theta})}{\partial \delta^2} \right\} dx_r dx_s dx_t \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} \tfrac{\partial f(x_t;\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} c(F(x_s;\boldsymbol{\theta}), F(x_t;\boldsymbol{\theta});\boldsymbol{\delta}) f(x_s;\boldsymbol{\theta}) \left\{ \int_{\mathbb{R}} \tfrac{\partial^2 c(F(x_r;\boldsymbol{\theta}),F(x_s;\boldsymbol{\theta});\boldsymbol{\delta}) f(x_r;\boldsymbol{\theta})}{\partial \boldsymbol{\delta}^2} dx_r \right\} dx_s dx_t \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} \tfrac{\partial f(x_t;\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} c(F(x_s;\boldsymbol{\theta}), F(x_t;\boldsymbol{\theta});\boldsymbol{\delta}) f(x_s;\boldsymbol{\theta}) \underbrace{\left\{ \tfrac{\partial^2}{\partial \boldsymbol{\delta}^2} \int_{\mathbb{R}} f(x_r|x_s;\boldsymbol{\theta},\boldsymbol{\delta}) dx_r \right\}}_{=\mathbf{0}} dx_s dx_t \\
&= \mathbf{0}.
\end{aligned}
$$

Finally we can conclude that

$$
\begin{aligned}
Cov(\boldsymbol{S}(\boldsymbol{\theta}, \boldsymbol{\delta}), \boldsymbol{S}_\delta(\boldsymbol{\theta}, \boldsymbol{\delta})) &= \sum_{t=2}^{n} E\left[\frac{1}{2} W_t(\boldsymbol{\theta}_0, \boldsymbol{\delta}_0)\nabla'_\delta \log c(u_t, u_{t-1}, \boldsymbol{\theta})\right] \\
&= (n-1)\boldsymbol{\sigma}_\delta.
\end{aligned}
$$

Putting the results together we get:

$$
\begin{pmatrix} n^{-1/2}\boldsymbol{S}(\boldsymbol{\theta}_0, \boldsymbol{\delta}_0) \\ n^{-1/2}\boldsymbol{S}_\delta(\boldsymbol{\theta}_0, \boldsymbol{\delta}_0) \\ n^{-1/2}\boldsymbol{S}_\theta(\boldsymbol{\theta}_0, \boldsymbol{\delta}_0) \end{pmatrix} \xrightarrow{d} N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \boldsymbol{\sigma}/4 & \boldsymbol{\sigma}_\delta/n & 0 \\ \boldsymbol{\sigma}_\delta/n & \boldsymbol{I}_{\delta\delta} & 0 \\ 0 & 0 & V \end{pmatrix} \right).
$$

Thus we finally get:

$$
n^{-1/2}\boldsymbol{S}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\delta}}) \xrightarrow{d} N(0, \boldsymbol{\Sigma}),
$$

with $\boldsymbol{\Sigma}$ just as in theorem 3.

□

Tables and Figures

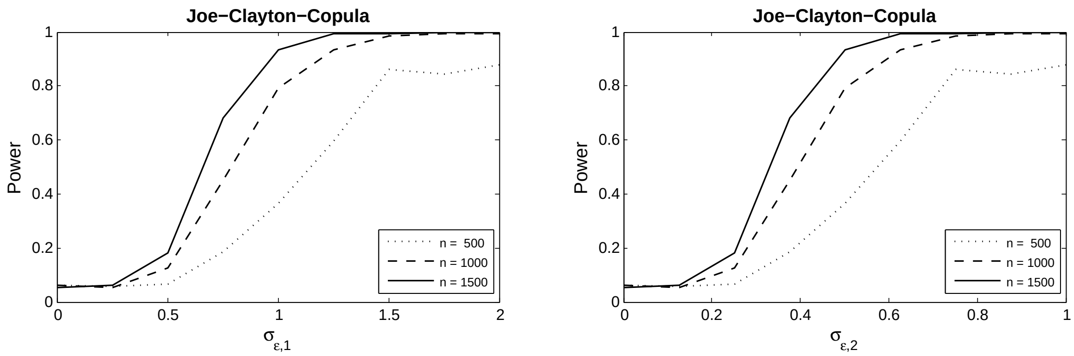

Figure 1: Power of the general score-test

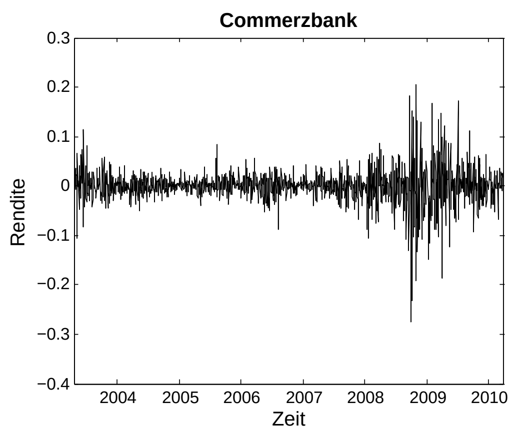

Figure 2: Daily log returns of Commerzbank

Table 4: Power of the score-test with Joe-Clayton copula and normal margins

						n=500		n=1000		n=1500	
μ	σ	δ_1	δ_2	σ_{ϵ_1}	σ_{ϵ_2}	p-value	power	p-value	power	p-value	power
-3	0.5	1.0	0.5	0.00	0.000	0.4803	0.0630*	0.4811	0.0620*	0.4738	0.0530*
				0.25	0.125	0.4739	0.0590	0.4758	0.0520	0.4700	0.0600
				0.50	0.250	0.4123	0.0650	0.3317	0.1250	0.2931	0.1800
				0.75	0.275	0.2731	0.1880	0.1313	0.4510	0.0673	0.6800
				1.00	0.500	0.1532	0.3650	0.0407	0.7930	0.0141	0.9320
				1.25	0.625	0.0823	0.5960	0.0130	0.9340	0.0023	0.9920
				1.50	0.750	0.0468	0.7600	0.0043	0.9840	0.0011	0.9950
				1.75	0.875	0.0317	0.8420	0.0020	0.9920	0.0007	0.9980
				2.00	1.000	0.0248	0.8780	0.0030	0.9920	0.0005	1.0000
		1.5	1.0	0.00	0.000	0.4937	0.0700*	0.4824	0.0690*	0.4728	0.0600*
				0.25	0.125	0.4934	0.0440	0.4927	0.0670	0.4726	0.0570
				0.50	0.250	0.4737	0.0460	0.4047	0.0690	0.3814	0.0980
				0.75	0.275	0.3770	0.0770	0.2507	0.2010	0.1662	0.3780
				1.00	0.500	0.2493	0.1790	0.1045	0.5050	0.0438	0.7610
				1.25	0.625	0.1528	0.3430	0.0397	0.8090	0.0129	0.9540
				1.50	0.750	0.0917	0.5210	0.0133	0.9430	0.0073	0.9790
				1.75	0.875	0.0622	0.6610	0.0082	0.9750	0.0049	0.9840
				2.00	1.000	0.0540	0.7310	0.0094	0.9700	0.0035	0.9890
1.0	2.0	1.0	0.5	0.00	0.000	0.4583	0.0710*	0.4730	0.0650*	0.4795	0.0640*
				0.25	0.125	0.4615	0.0460	0.4708	0.0610	0.4763	0.0470
				0.50	0.250	0.4299	0.0640	0.3491	0.1360	0.2698	0.2180
				0.75	0.275	0.2782	0.1670	0.1279	0.4470	0.0721	0.6720
				1.00	0.500	0.1514	0.3670	0.0339	0.8160	0.0096	0.9570
				1.25	0.625	0.0820	0.5710	0.0119	0.9440	0.0030	0.9870
				1.50	0.750	0.0456	0.7550	0.0051	0.9820	0.0020	0.9970
				1.75	0.875	0.0317	0.8360	0.0034	0.9870	0.0006	0.9980
				2.00	1.000	0.0278	0.8730	0.0036	0.9870	0.0004	0.9990
		1.5	1.0	0.00	0.000	0.4843	0.0680*	0.4751	0.0680*	0.4872	0.0610*
				0.25	0.125	0.4785	0.0540	0.4863	0.0420	0.4886	0.0430
				0.50	0.250	0.4384	0.0420	0.4180	0.0720	0.3922	0.0960
				0.75	0.275	0.3711	0.0960	0.2461	0.2140	0.1703	0.3640
				1.00	0.500	0.2490	0.1550	0.1000	0.5300	0.0474	0.7810
				1.25	0.625	0.1543	0.3380	0.0342	0.8000	0.0138	0.9410
				1.50	0.750	0.0951	0.5430	0.0167	0.9290	0.0079	0.9800
				1.75	0.875	0.0675	0.6460	0.0106	0.9640	0.0035	0.9890
				2.00	1.000	0.0485	0.7400	0.0061	0.9760	0.0051	0.9830

REMARK: The means of 1000 replications of the p-values of the genaral score-test and the power are displayed for different parameter constellations. The α errors are marked with * and describe how often the null hypothesis is rejected, although it is true.

Table 5: Estimates for the daily log returns of Commerzbank

		Marginal distribution						
		gaussian		hyperbolic		empiric		GARCH(1,1)
		$\hat{\mu}$	-0.0002	$\hat{\psi}$	50.7488*			
			(0.0007)		(1.1934)			
		$\hat{\sigma}$	0.0317*	$\hat{\eta}$	-1.4810		$\hat{\omega}$	0.0001*
			(0.0005)		(0.8446)			(0.0000)
				$\hat{\sigma}$	0.0002*		$\hat{\alpha}$	0.0656*
					(0.00001)			(0.0049)
				$\hat{\mu}$	0.0010		$\hat{\beta}$	0.9302*
					(0.00050)			(0.0045)
		BIC	-7312.5067	BIC	-8036.1388		BIC	-8497.5719
Constant copula								GARCH-filtered
Gauss	$\hat{\delta}$		0.0762		0.00479	0.0381		0.0429
			(0.0333)		(0.0226)	(0.0235)		(0.0239)
	$\hat{\tau}$		0.0504		0.0305	0.0243		0.0273
	BIC		-11.2744		-4.4591	-2.5736		-3.2651
Clayton	$\hat{\delta}$		0.0826*		0.1180	0.1241		0.0731*
			(0.0151)		(0.0200)	(0.0275)		(0.0249)
	$\hat{\tau}$		0.0397		0.0557	0.0584		0.0344
	BIC		-58.4801		-41.4359	-27.0222		-8.2544
Gumbel	$\hat{\delta}$		1.0554*		1.0535*	1.0001		1.0000
			(0.0158)		(0.0136)	(0.0171)		(0.0148)
	$\hat{\tau}$		0.0525		0.0507	0.0000		0.0000
	BIC		-34.9325		-31.4023	1430.2750		1491.3505
Joe-Clayton	$\hat{\delta}_1$		1.0312*		1.0442*	1.0541*		1.0000
			(0.0128)		(0.0133)	(0.0167)		(0.0208)
	$\hat{\delta}_2$		0.0964*		0.1066*	0.1043*		0.0708*
			(0.0182)		(0.0293)	(0.0273)		(0.0343)
	$\hat{\tau}$		0.0621		0.0730	0.0769		0.0342
	BIC		-13.473		-61.2031	-39.0727		-7.7324
Score-test for Joe-Clayton	p-value		0.0591†		0.0001	0.0001		0.3879†
Dynamic copula								
Joe-Clayton	$\hat{\omega}_1$		***		0.5860*	0.1802*		***
					(0.0007)	(0.0107)		
	$\hat{\omega}_2$				-1.4947*	-1.1684*		
					(0.0033)	(0.0019)		
	$\hat{\alpha}_1$				-0.7466*	-0.1379*		
					(0.0002)	(0.0330)		
	$\hat{\alpha}_2$				0.0983*	0.8718*		
					(0.0023)	(0.0035)		
	$\hat{\beta}_1$				-1.1579*	0.8432*		
					(0.0019)	(0.0160)		
	$\hat{\beta}_2$				-1.7954*	0.1368*		
					(0.0068)	(0.0049)		
	BIC				-103.2394	-90.0195		

REMARK: Standard errors are embraced. The parameter estimates marked with * are significant at 5%-level. The field marked with *** indicates that no dynamic copula parameter is estimated because the null hypothesis of the score-test cannot be rejected. † indicates that the test is performed for the Clayton copula.

References

[Abegaz & Naik-Nimbalkar, 2008] Abegaz, F. & Naik-Nimbalkar, U. (2008). Dynamic copula-based markow time series. *Communications in Statistics - Theory and Methods*, 37, 2447–2460.

[Basawa & Rao, 1980] Basawa, I. & Rao, B. (1980). *Statistical inference for stochastic processes*. London: Academic Press.

[Bollerslev, 1986] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31, 307–327.

[Chen, 2007] Chen, Y.-T. (2007). Moment-based copula tests for financial returns. *Journal of Business and Economic Statistics*, 24, 377–397.

[Cont, 2001] Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, 1, 223–236.

[Davidson, 1994] Davidson, J. (1994). *Stochastic limit theory: An introduction for econometricians*. Oxford: Oxford University Press.

[Denker, 1986] Denker, M. (1986). Uniform integrability and the central limit theorem for strongly mixing processes. In *Dependence in probability and statistics*, volume 11 (pp. 269–289). Boston: Birkhäuser Boston.

[Eberlein & Keller, 1995] Eberlein, E. & Keller, U. (1995). Hyperbolic distributions in finance. *Bernoulli*, 1, 281–299.

[Embrechts et al., 2002] Embrechts, P., McNeil, A., & Strautmann, D. (2002). *Correlation and dependence properties in risk management: Properties and pitfalls*. Risk Management: Value-at-risk and beyond. Cambridge: Cambridge University Press.

[Engle, 1982] Engle, R. (1982). Autoregressive conditional heteroscedasticity with estimates of variance of united kingdom inflation. *Econometrica*, 50, 987–1008.

[Ferguson, 1996] Ferguson, T. (1996). *A course in large sample theory*. London: Chapman & Hall.

[Genest et al., 1995] Genest, C., Ghoudi, K., & Rivest, L.-P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, 82, 543–552.

[Hansen, 1994] Hansen, B. (1994). Autoregressive conditional density estimation. *International Economic Review*, 35, 705–730.

[Härdle & Okhrin, 2010] Härdle, W. & Okhrin, O. (2010). De copulis non est disputandum. *Advances in Statistical Analysis*, 94, 1–32.

[Jaschke, 2000] Jaschke, S. (2000). A note on stochastic volatility, GARCH models, and hyperbolic distributions. *Working Paper*.

[Joe, 1997] Joe, H. (1997). *Multivariate models and dependence concepts*. New York: Chapman & Hall.

[McNeil et al., 2005] McNeil, A., Frey, R., & Embrechts, P. (2005). *Quantitative risk management*. Princton: Princton University Press.

[Nelsen, 2006] Nelsen, R. (2006). *An introduction to copula*. New York: Springer.

[Patton, 2002] Patton, A. (2002). *Applications of copula theory in financial econometrics*. PhD thesis, University of California, San Diego.

[Patton, 2006] Patton, A. (2006). Modelling asymmetric exchange rate dependence. *International Economic Review*, 47, 527–556.

[Rao, 1973] Rao, C. (1973). *Linear statistical inference and its application*. New York: John Wiley & Sons.

[Reichert, 2010] Reichert, K. (2010). Copulabasierte Modellierung intertemporärer Abhängigkeiten. Diploma thesis, Departement of Mathematics, Erlangen.

[Reimann, 2005] Reimann, S. (2005). Evidence for a hyperbolic-like distribution of asset returns drawn from a simple economical financial markets model. *FINRISK Working paper*, 212.

[White, 1984] White, H. (1984). Maximum likelihood estimation of misspecified dynamic models. In T. Dijkstra (Ed.), *Misspecification analysis*. New York: Springer.

[White, 1994] White, H. (1994). *Estimation, inference and specification analysis*. New York: Cambridge University Press.

[White, 2001] White, H. (2001). *Asymptotic theory for econometricians*. San Diego: Academic Press.

Diskussionspapiere 2011
Discussion Papers 2011

01/2011 **Klein, Ingo, Fischer, Matthias, Pleier, Thomas:** Weighted Power Mean Copulas: Theory and Application

02/2011 **Kiss, David:** The Impact of Peer Ability and Heterogeneity on Student Achievement: Evidence from a Natural Experiment

03/2011 **Zibrowius, Michael:** Convergence or divergence? Immigrant wage assimilation patterns in Germany

04/2011 **Klein, Ingo, Christa, Florian:** Families of Copulas closed under the Construction of Generalized Linear Means

05/2011 **Schnitzlein, Daniel**: How important is the family? Evidence from sibling correlations in permanent earnings in the US, Germany and Denmark

06/2011 **Schnitzlein, Daniel**: How important is cultural background for the level of intergenerational mobility?

07/2011 **Steffen Mueller**: Teacher Experience and the Class Size Effect - Experimental Evidence

08/2011 **Klein, Ingo:** Van Zwet Ordering for Fechner Asymmetry

Diskussionspapiere 2010
Discussion Papers 2010

01/2010 **Mosthaf, Alexander, Schnabel, Claus and Stephani, Jens:** Low-wage careers: Are there dead-end firms and dead-end jobs?

02/2010 **Schlüter, Stephan and Matt Davison:** Pricing an European Gas Storage Facility using a Continuous-Time Spot Price Model with GARCH Diffusion

03/2010 **Fischer, Matthias, Gao, Yang and Herrmann, Klaus:** Volatility Models with Innovations from New Maximum Entropy Densities at Work

04/2010 **Schlüter, Stephan, Deuschle, Carola:** Using Wavelets for Time Series Forecasting – Does it Pay Off?

05/2010 **Feicht, Robert, Stummer, Wolfgang:** Complete closed-form solution to a stochastic growth model and corresponding speed of economic recovery.

06/2010 **Hirsch, Boris, Schnabel, Claus:** Women Move Differently: Job Separations and Gender.

07/2010 **Gartner, Hermann, Schank, Thorsten, Schnabel, Claus:** Wage cyclicality under different regimes of industrial relations.

08/2010 **Tinkl, Fabian**: A note on Hadamard differentiability and differentiability in quadratic mean.

09/2010 **Tinkl, Fabian**: Asymptotic theory for M estimators for martingale differences with applications to GARCH models.

Diskussionspapiere 2009
Discussion Papers 2009

01/2009 **Addison, John T. and Claus Schnabel:** Worker Directors: A German Product that Didn't Export?

02/2009 **Uhde, André and Ulrich Heimeshoff:** Consolidation in banking and financial stability in Europe: Empirical evidence

03/2009 **Gu, Yiquan and Tobias Wenzel:** Product Variety, Price Elasticity of Demand and Fixed Cost in Spatial Models

04/2009 **Schlüter, Stephan:** A Two-Factor Model for Electricity Prices with Dynamic Volatility

05/2009 **Schlüter, Stephan and Fischer, Matthias**: A Tail Quantile Approximation Formula for the Student t and the Symmetric Generalized Hyperbolic Distribution

06/2009 **Ardelean, Vlad:** The impacts of outliers on different estimators for GARCH processes: an empirical study

07/2009 **Herrmann, Klaus:** Non-Extensitivity versus Informative Moments for Financial Models - A Unifying Framework and Empirical Results

08/2009 **Herr, Annika:** Product differentiation and welfare in a mixed duopoly with regulated prices: The case of a public and a private hospital

09/2009 **Dewenter**, Ralf, **Haucap**, Justus and **Wenzel**, Tobias: Indirect Network Effects with Two Salop Circles: The Example of the Music Industry

10/2009 **Stuehmeier**, Torben and **Wenzel**, Tobias: Getting Beer During Commercials: Adverse Effects of Ad-Avoidance

11/2009 **Klein**, Ingo, **Köck**, Christian and **Tinkl**, Fabian: Spatial-serial dependency in multivariate GARCH models and dynamic copulas: A simulation study

12/2009 **Schlüter**, Stephan: Constructing a Quasilinear Moving Average Using the Scaling Function

13/2009 **Blien,** Uwe, **Dauth,** Wolfgang**, Schank,** Thorsten and **Schnabel,** Claus**:** The institutional context of an "empirical law": The wage curve under different regimes of collective bargaining

14/2009 **Mosthaf**, Alexander, **Schank**, Thorsten and **Schnabel**, Claus: Low-wage employment versus unemployment: Which one provides better prospects for women?

Diskussionspapiere 2008
Discussion Papers 2008

01/2008 **Grimm, Veronika and Gregor Zoettl**: Strategic Capacity Choice under Uncertainty: The Impact of Market Structure on Investment and Welfare

02/2008 **Grimm, Veronika and Gregor Zoettl**: Production under Uncertainty: A Characterization of Welfare Enhancing and Optimal Price Caps

03/2008 **Engelmann, Dirk and Veronika Grimm**: Mechanisms for Efficient Voting with Private Information about Preferences

04/2008 **Schnabel, Claus and Joachim Wagner**: The Aging of the Unions in West Germany, 1980-2006

05/2008 **Wenzel, Tobias**: On the Incentives to Form Strategic Coalitions in ATM Markets

06/2008 **Herrmann, Klaus**: Models for Time-varying Moments Using Maximum Entropy Applied to a Generalized Measure of Volatility

07/2008 **Klein, Ingo and Michael Grottke:** On J.M. Keynes' "The Principal Averages and the Laws of Error which Lead to Them" - Refinement and Generalisation