SFB 823

Discussion Paper

A nonparametric constancy test for copulas under mixing conditions

Maarten van Kampen, Dominik Wied

Nr. 36/2010

A nonparametric constancy test for copulas under mixing conditions[☆]

Maarten van Kampen[a,b], Dominik Wied[b]

[a] *Ruhr Graduate School, D-45128 Essen, Germany*
[b] *Fakultät Statistik, Universität Dortmund, D-44221 Dortmund, Germany*

Abstract

This paper generalizes some recently proposed tests which examine if a copula is constant over time. The i.i.d. assumption underlying these tests is relaxed by imposing only strong mixing and it is shown that the resulting tests are consistent against the alternative of a structural break.

Keywords: copulas, mixing
JEL: C12, C14

1. Introduction

In econometric applications dependence measures such as linear correlations often change over time. A fortiori, the same applies to copulas. Patton (2006) and Jondeau and Rockinger (2006) examine if a time-varying copula model represents the dependence structure of the data better than a time-invariant copula. A serious drawback of their approach is that the results might depend on the choice of the functional form of the copula and the way the copula is allowed to change over time.

Recently, Busetti and Harvey (2008) and Krämer and Van Kampen (2009) proposed a nonparametric test to examine whether a copula is constant over time. The nonparametric test avoids the specification of a specific functional form as well as the specification of a transition mechanism. The test is based

[☆]Financial support from the Ruhr Graduate School and the Deutsche Forschungsgemeinschaft (SFB 823) is gratefully acknowledged.

Email addresses: `maarten.vankampen@tu-dortmund.de` (Maarten van Kampen), `wied@statistik.tu-dortmund.de` (Dominik Wied)

on the stationarity test of De Jong et al. (2007), who modified the original KPSS test (see Kwiatkowski et al. (1992)) by using indicators for whether the data is below or above the median instead of using deviations from the mean. Busettti and Harvey (2007) constructed a quantile constancy test which generalizes the previous idea for arbitrary quantiles. The underlying idea of the copula constancy test is to use the fact that a (bivariate) copula $C(\tau_1, \tau_2)$ gives the probability that the each of random variables takes values below their τ_i-quantile, $i = 1, 2$, and to construct suitable indicators for this event. This idea can easily be extended to more than two dimensions.

The copula constancy test has been developed under the assumption that the observations are independent and identically distributed. This assumption is often violated in empirical applications. Kwiatkowski et al. (1992) and De Jong et al. (2007) constructed their tests under the assumption that the observations are strong mixing, thereby allowing for weak dependence.

In this paper we likewise relax the i.i.d. assumption underlying the copula constancy test by imposing strong mixing conditions. We also show that the test has the same asymptotic null distribution for filtered observations. This result is useful if the marginal distributions are changing over time. Finally, we show that the resulting test is consistent against the alternative of a single structural break.

2. Testing for constancy under i.i.d. assumption

Consider the bivariate i.i.d. series $\{y_t\}_{t=1}^T$ with $y_t = (y_{1t}, y_{2t})$. Let $\xi_i(\tau_i)$ be the τ_i-quantile of y_{it} where $\tau_i \in (0, 1)$, $i = 1, 2$. The copula $C^{(t)}(\tau_1, \tau_2)$ gives the probability that each variable takes values below or equal to its τ_i-quantile

$$C^{(t)}(\tau_1, \tau_2) = P(y_{1t} \leq \xi_1(\tau_1), y_{2t} \leq \xi_2(\tau_2))$$

We examine if this probability changes over time. The hypothesis pair is

$$\begin{aligned} &H_0 : C^{(t)}(\tau_1, \tau_2) = C(\tau_1, \tau_2) && \text{for all } t = 1, \dots, T \\ &H_1 : C^{(t)}(\tau_1, \tau_2) \neq C^{(t+1)}(\tau_1, \tau_2) && \text{for some } t \in \{1, \dots, T-1\} \end{aligned}$$

where $C(\tau_1, \tau_2)$ is a time-invariant copula.

The test is based on indicators of the event $\{y_{1t} \leq \xi_1(\tau_1), y_{2t} \leq \xi_2(\tau_2)\}$. Let $\mathbb{I}(\cdot)$ be the indicator function taking the value 1 if the event between brackets is true and zero otherwise. Define

$$I(y_t, \xi(\tau)) := \mathbb{I}(y_{1t} \leq \xi_1(\tau_1), y_{2t} \leq \xi_2(\tau_2))$$

and let $C_T(\tau_1, \tau_2) := T^{-1} \sum_{t=1}^{T} I(y_t, \xi(\tau))$ be the empirical copula. Note that under the null hypothesis $I(y_t, \xi(\tau))$ is a Bernoulli variable with probability $C(\tau_1, \tau_2)$ and thus $Q_t := Q(y_t, \xi(\tau)) := C(\tau_1, \tau_2) - I(y_t, \xi(\tau))$ has expectation zero and variance $C(\tau_1, \tau_2)(1 - C(\tau_1, \tau_2))$.

Define $S_T(r) := 1/(\sigma\sqrt{T}) \sum_{t=1}^{[rT]} Q_t$ where $\sigma^2 = C(\tau_1, \tau_2)(1 - C(\tau_1, \tau_2))$, $r \in [0, 1]$ and $[rT]$ denotes the integer part of rT. Then, using a functional central limit theorem (FCLT), we have

$$S_T(\cdot) \xrightarrow{d} B(\cdot) \tag{1}$$

where B denotes a Brownian motion.

Replacing $C(\tau_1, \tau_2)$ by its empirical estimate $C_T(\tau_1, \tau_2)$ gives, using the terminology of Busetti and Harvey (2008), the bivariate $\tau-$quantics

$$BIQ(y_t, \xi(\tau)) := C_T(\tau_1, \tau_2) - I(y_t, \xi(\tau))$$

Note that these are the mean deviations of Q_t, i.e. $BIQ(y_t, \xi(\tau)) = Q_t - T^{-1} \sum_{t=1}^{T} Q_t$. Therefore, for $\tilde{S}_T(r) := 1/(\sigma\sqrt{T}) \sum_{t=1}^{[rT]} BIQ(y_t, \xi(\tau))$ we have

$$\tilde{S}_T(\cdot) \xrightarrow{d} V(\cdot)$$

where $V(r) := B(r) - rB(1)$ denotes a Brownian bridge.

The $BIQ(y_t, \xi(\tau))$ are unobserved since they depend on the population quantile $\xi(\tau)$. Let $\hat{\xi}(\tau)$ denote the sample quantile, let $\hat{C}_T(\tau_1, \tau_2) := T^{-1} \sum_{t=1}^{T} I(y_t, \hat{\xi}(\tau))$ be the empirical copula based on the sample quantiles and let $BIQ(y_t, \hat{\xi}(\tau)) = \hat{C}_T(\tau_1, \tau_2) - I(y_t, \hat{\xi}(\tau))$ be the corresponding bivariate τ-quantics. Define

$$\hat{S}_T(r) := 1/(\sigma\sqrt{T}) \sum_{t=1}^{[rT]} BIQ(y_t, \hat{\xi}(\tau)) \tag{2}$$

Then Busetti and Harvey (2008) show that

$$\sup_{r \in [0,1]} \left| \hat{S}_T(r) - \tilde{S}_T(r) \right| \xrightarrow{p} 0$$

The copula constancy tests are different functionals of $\hat{S}_T(\cdot)$. Using the continuous mapping theorem we obtain the asymptotic distribution under the null hypothesis. The test based on the *squares* is given by

$$\frac{1}{T^2 \hat{\sigma}_{iid}^2} \sum_{t=1}^{T} \left(\sum_{j=1}^{t} BIQ(y_j, \hat{\xi}(\tau)) \right)^2$$

where $\hat{\sigma}^2_{iid} := \hat{C}_T(\tau_1, \tau_2)(1 - \hat{C}_T(\tau_1, \tau_2))$ is the estimate of σ^2. The test is distributed as Cramér-von Mises and some useful critical values are 0.743 (1%), 0.461 (5%) and 0.347 (10%).

Krämer and Van Kampen (2009) propose complementary tests based on the *maximum* and the *range* of $\hat{S}_T(\cdot)$

$$\frac{1}{\sqrt{T}\hat{\sigma}_{iid}} \max_{t=1,\ldots,T} \left| \sum_{j=1}^{t} BIQ(y_j, \hat{\xi}(\tau)) \right|$$

$$\frac{1}{\sqrt{T}\hat{\sigma}_{iid}} \left[\max_{t=1,\ldots,T} \sum_{j=1}^{t} BIQ(y_j, \hat{\xi}(\tau)) - \min_{t=1,\ldots,T} \sum_{j=1}^{t} BIQ(y_j, \hat{\xi}(\tau)) \right]$$

Some useful critical values for the maximum test are 1.63 (1%), 1.36 (5%), 1.22 (10%) and for the range test are 2.001 (1%), 1.747 (5%) and 1.620 (10%).

3. Testing for constancy under a mixing assumption

In this section we relax the i.i.d. assumption by imposing strong mixing conditions. For $i = 1, 2$, the sequence $\{y_{it}\}_{t=-\infty}^{\infty}$ is said to be strong-mixing if $\lim_{m\to\infty} \alpha(m) = 0$, where

$$\alpha(m) := \sup_t \sup_{A \in \mathcal{F}_{-\infty}^t, B \in \mathcal{F}_{t+m}^{\infty}} |P(A \cap B) - P(A)P(B)|$$

and, $\mathcal{F}_{-\infty}^t$ and $\mathcal{F}_{t+m}^{\infty}$ are sigma-fields based on respectively $(\ldots, y_{i,t-1}, y_{it})$ and $(y_{i,t+m}, y_{i,t+m+1}, \ldots)$, see e.g. Davidson (1994, p.209). So a strong mixing sequence satisfies asymptotic independence.

To construct a copula constancy test, we adopt similar assumptions as in De Jong et al. (2007).

Assumption 1.

1. The observations y_{it} are strictly stationary and ξ_i is the unique population quantile of y_{it}.
2. y_{it} is strong mixing with mixing coefficient $\alpha(m) = O(m^{-p/(p-2)})$ for some finite $p > 2$ (see remark (i)).
3. $y_t - \xi$ has a continuous joint density $f_{12}(u_1, u_2)$ in a neighborhood $[-\eta, \eta]^2$ of 0 for some $\eta > 0$, and $\inf_{(u_1,u_2)\in[-\eta,\eta]^2} f_{12}(u_1, u_2) > 0$
4. Long run variance $\sigma^2 \in (0, \infty)$

Remark:

(i) Application of a FCLT for mixing variables requires that y_{it} is L_p- boundend, $E|y_{it}|^p < \infty$, for some finite $p > 2$ (see Davidson 1994, p.482).

(ii) The bound on the mixing coefficients is required to establish Lemma 1 in De Jong et al. (2007). This restriction allows, for example, for ARMA processes with Gaussian innovations, see Withers (1981). Lindner (2009, Theorem 8) gives conditions such that GARCH processes are strong mixing. However, in this case the copula constancy tests are subject to size distortions (see discussion below).

(iii) The joint density $f_{12}(u_1, u_2)$ can written as

$$f(u_1, u_2) = c(F_1(u_1), F_2(u_2))f_1(u_1)f_2(u_2)$$

where $f_i(\cdot)$ and $F_i(\cdot)$ are respectively the marginal density and distribution of $y_{it} - \xi_i$, and $c(\cdot, \cdot)$ is the copula density. Assumption 1.3 is satisfied if $c(F_1(u_1), F_2(u_2))$ and $f_i(u_i)$, $i = 1, 2$, are nonzero and continuous for $(u_1, u_2) \in [-\eta, \eta]^2$. Note that we do not require that the copula density is continuous on its complete domain $[0, 1]^2$.

Under Assumption 1, $S_T(\cdot)$ satisfies a functional central limit theorem. Provided $T^{-1}E(\sum_{t=1}^{T} Q_t)^2 \to \sigma^2$ with $0 < \sigma^2 < \infty$, we have $S_T(\cdot) \xrightarrow{d} B(\cdot)$, see e.g. Corollary 29.7 of Davidson (1994). In addition

$$\tilde{S}_T(\cdot) \xrightarrow{d} V(\cdot). \tag{3}$$

The HAC estimator, $\bar{\sigma}^2$, for σ^2 is given by

$$\bar{\sigma}^2 = T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} k((t-s)/\gamma_T) \cdot Q_t \cdot Q_s \tag{4}$$

where the bandwidth, γ_T, and the kernel, $k(\cdot)$, satisfy the following conditions:

Assumption 2.

1. $k(\cdot)$ satisfies $\int_{-\infty}^{\infty} |\psi(w)|dw < \infty$, where
$$\psi(w) = (2\pi)^{-1} \int_{-\infty}^{\infty} k(x) \exp(-iwx)dw.$$
2. $k(\cdot)$ is continuous at all but a finite number of points, $k(x) = k(-x)$, $|k(x)| \leq l(x)$ where $l(x)$ is non-increasing and $\int_0^{\infty} |l(x)|dx < \infty$, and $k(0) = 1$.
3. $\gamma_T/\sqrt{T} \to 0$, and $\gamma_T \to \infty$ as $T \to \infty$.

Remark:

(i) Assumption 2 ensures that the variance estimate remains nonnegative. The Bartlett, Parzen, Tukey-Hanning and Quadratic Spectral kernel satify this assumption. The truncated kernel does not satify this assumption, see De Jong and Davidson (2000).

(ii) Assumption 2.3 strengthens the rate of the bandwidth parameter compared to De Jong et al. (2007). Andrews (1991) points out that optimal growth rates of γ_T (in terms of a MSE criterion) are typically less than $o(T^{1/2})$. Imposing $o(T^{1/2})$ can therefore be regarded as a mild requirement.

The HAC estimate (4) is not feasible since Q_t depends on the true unobserved copula $C(\tau_1, \tau_2)$ and on the population quantile $\xi(\tau)$. Replacing $C(\tau_1, \tau_2)$ by the empirical copula $C_T(\tau_1, \tau_2)$ and $\xi(\tau)$ by the sample quantile $\hat{\xi}(\tau)$ gives the feasible HAC estimator
$$\hat{\sigma}^2 = T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} k((t-s)/\gamma_T) \cdot BIQ(y_t, \hat{\xi}(\tau)) \cdot BIQ(y_s, \hat{\xi}(\tau)). \qquad (5)$$

We make the following assumption on the empirical quantile process.

Assumption 3. $\sqrt{T}(\hat{\xi}_i(\tau) - \xi_i(\tau)) = O_p(1)$, for $i = 1, 2$.

Remark: Assumption 3 follows from asymptotic normality of $\sqrt{T}(\hat{\xi}_i(\tau) - \xi_i(\tau))$. Sufficient conditions for asymptotic normality are given by Koenker (2005, p.71-72) for the i.i.d. case, De Jong et al. (2007) for the strong mixing case but only for $\tau = 0.5$ and Sun and Lahiri (2006) for the general strong mixing case.

Theorem 1 establishes the result of the previous section for the case that the observations are strong mixing. The proof is given in Appendix A.

Theorem 1. *Under Assumptions 1, 2 and 3*

$$\hat{S}_T(\cdot) \xrightarrow{d} V(\cdot)$$

and $\hat{\sigma}^2 \xrightarrow{p} \sigma^2$.

Busetti and Harvey (2008) show that the test is subject to size distortions if the marginal distributions are changing over time. This is, for instance, the case if the series exhibit stochastic volatility. A solution is to use the standardized observations $y_{it}(\hat{\theta}_T) := x_{it}/h_{it}(\hat{\theta}_T)$ where x_{it} are the observations, $h^2_{it}(\hat{\theta}_T)$ is an estimate of the volatility and $\hat{\theta}_T$ is an $q \times 1$ vector of parameter estimates of the true population parameter, θ_0, at sample size T. This approach is only legitimate if we can substitute $y_t(\theta_0)$ by $y_t(\hat{\theta}_T)$ and $\sigma^2_t(\theta_0)$ by $\sigma^2_t(\hat{\theta}_T)$ in Theorem 1.

Since marginal distributions can also change for other reasons than stochastic volatility, we extend Theorem 1 to the general case where $y_t(\theta)$ depends on a parameter vector $\theta \in \Theta$, and where Θ denotes the compact parameter space. Let $\xi(\theta, \tau)$ denote the quantile function of $y_t(\theta)$ and let $\hat{S}_T(\theta, \cdot)$ be as $\hat{S}_T(\cdot)$ but with y_t replaced by $y_t(\theta)$.

We impose the following additional assumption:

Assumption 4.

1. $\sqrt{T}(\hat{\theta}_T - \theta_0) = O_p(1)$.
2. For $\varepsilon > 0$ and finite constants $c_{y,\varepsilon}$, $c_\xi > 0$, $\sup_{\theta\in\Theta} |\partial y(\theta)/\partial\theta| < c_{y,\varepsilon}$ with probability $1 - \varepsilon$ and $\sup_{\theta\in\Theta} |\partial\xi(\theta, \tau)/\partial\theta| < c_\xi$.
3. $y_t(\theta) - \xi(\theta, \tau)$ has a continuous differentiable joint density $f_{12}(u_1, u_2)$ in a neighborhood $[-\eta, \eta]^2$ of 0 for some $\eta > 0$, and $\inf_{(u_1,u_2)\in[-\eta,\eta]^2} f_{12}(u_1, u_2) > 0$.

Remark:

(i) Assumption 4.1 follows from asymptotic normality of $\sqrt{T}(\hat{\theta}_T - \theta_0)$ which is satisfied for GARCH models estimated by maximum likelihood (see e.g. Gouriéroux (1997, p.44)).

(ii) Suppose the volatility $h_{it}(\theta)$ (in the example of the main text) is estimated using a GARCH model. Then Assumption 4.2 is satisfied if $\partial h_{it}(\theta)/\partial\theta$ exist and the volatility is unequal to zero. Existence of $\partial h_{it}(\theta)/\partial\theta$ is also imposed to obtain the asymptotic variance-covariance matrix.

Theorem 2. *Under Assumptions 1, 2, 3 and 4 we have*

$$\hat{S}_T(\hat{\theta}_T, \cdot) \xrightarrow{d} V(\cdot)$$

and $\hat{\sigma}^2(\hat{\theta}_T) \xrightarrow{p} \sigma^2$.

An application of the continuous mapping theorem after Theorem 1 or 2 gives the same tests as established in the i.i.d. case.

Finally, we can evaluate the test at several τ values or construct an overall constancy test by combining BIQ values defined on different quadrants. Although we focus in the remainder of this paper on the asymptotic power and direct application of the copula constancy test described above, an interesting complementary approach is to make use of the relationship between dependence measures such as Spearman's ρ and the copula (see Nelson (2006, chapter 5)). Fluctuation tests for Spearman's ρ, such as the one proposed in Dehling et al. (2010), can be written as a properly scaled integral (with respect to τ) of the partial sums of the BIQ values as well and can so be seen as an alternative way to analyze if the copula is constant.

4. The asymptotic power of the test

4.1. Consistency

We consider first a *fixed* alternative of a single break in the copula at some fraction $z^* \in (0, 1)$ of the sample. Let $C(\tau_1, \tau_2)$ and $C^*(\tau_1, \tau_2)$ be two different bivariate copulas. The copulas $C(\tau_1, \tau_2)$ and $C^*(\tau_1, \tau_2)$ may come from the same family but should then have different parameter values.

The hypothesis pair is

$$\begin{aligned} H_0: &\quad C^{(t)}(\tau_1, \tau_2) = C(\tau_1, \tau_2) \\ H_1: &\quad C^{(t)}(\tau_1, \tau_2) = (1 - g(t, T))C(\tau_1, \tau_2) + g(t, T)C^*(\tau_1, \tau_2) \end{aligned} \tag{6}$$

where $g(t, T) = 0$ for $t/T \leq z^*$ and $g(t, T) = \omega$ for $t/T > z^*$, $\omega \in (0, 1]$.

Define

$$Q^1(y_t, \xi(\tau)) \;:=\; (1 - g(t, T))C(\tau_1, \tau_2) + g(t, T)C^*(\tau_1, \tau_2) - I(y_t, \xi(\tau))$$

Theorem 3. *Provided that* $T^{-1}E\left(\sum_{t=1}^{T} Q^1(y_t, \xi(\tau))\right)^2 \to \sigma_1^2$ *for* $\sigma_1^2 \in (0, \infty)$, *the copula constancy tests are consistent against the alternative* (6).

Remark:

(i) Note that under the fixed alternative (6), the variance of the terms in the partial sum process change over time.

(ii) In the special case when observations are i.i.d. the condition stated in the theorem is clearly satisfied since

$$T^{-1}E\left(\sum_{j=1}^{T} Q^1(y_j,\xi(\tau))\right)^2$$
$$= z^*C(\tau_1,\tau_2)[1-C(\tau_1,\tau_2)] + (1-z^*)C^1(\tau_1,\tau_2)[1-C^1(\tau_1,\tau_2)],$$

where $C^1(\tau_1,\tau_2) := (1-\omega)C(\tau_1,\tau_2) + \omega C^*(\tau_1,\tau_2)$.

4.2. Local Alternatives

Next consider a sequence of local alternatives

$$(1-g(t,T))C(\tau_1,\tau_2) + g(t,T)C^*(\tau_1,\tau_2), \tag{7}$$

where $g(t,T) : [0,T]\times\mathbb{R}_+ \to (0,1)$ is defined as $g(t,T) = T^{-1/2}h(t/T)$ for some function $h(t/T)$ satisfying $\sup_x h(x) < \infty$. Berg and Quessy (2009) use a similar setup to analyze the asymptotic behavior of goodness of fit tests for copulas.

Theorem 4. *Under local alternatives* (7)

$$\hat{S}_T(\cdot) \xrightarrow{d} \frac{\sigma_1}{\sigma}V(\cdot) + \frac{1}{\sigma}[C(\tau_1,\tau_2) - C^*(\tau_1,\tau_2)]\left(\int_0^{(\cdot)} h(s)ds - (\cdot)\int_0^1 h(s)ds\right)$$

and $\hat{\sigma}^2 \xrightarrow{p} \sigma_1^2$.

This shows that the copula constancy test is inconsistent against local alternatives (7) but does converge to a fixed limit. In section 5.2 we show by simulation that if the magnitude of ω and the difference between the copulas $C(\tau_1,\tau_2)$ and $C^*(\tau_1,\tau_2)$ is sufficiently large, then the power might still tend to 1.

5. Finite sample

In this section we examine the finite sample properties of the test. The results are generated using Ox (see Doornik (2005)) and the G@RCH package of Laurent and Peters (2006).

5.1. Size of test

To examine the size of the test, we simulate 50000 replications of 500 observations from the following Copula-ARMA-GARCH model

$$\begin{aligned} x_{it} &= \theta_1 x_{i,t-1} + \varepsilon_{i,t} + \theta_2 \varepsilon_{i,t-1} \\ \varepsilon_{it} &= h_{it} \varepsilon_{it}^{\dagger} \\ h_{it}^2 &= \theta_3 + \theta_4 \varepsilon_{it}^2 + \theta_5 h_{i,t-1}^2 \end{aligned} \tag{8}$$

where $\varepsilon_{it}^{\dagger} = \Phi(u_{it})$, $\Phi(\cdot)$ denotes the univariate normal CDF and $u_t = (u_{1t}, u_{2t})$ is simulated from a copula C with parameter such that Kendall's tau equals 0.25.

For the ARMA (and GARCH) recursion, we simulate 1000 additional observations and discard these afterwards. We examine the properties of the test using a Clayton, Gaussian and Student copula where we assume that the latter has 4 degrees of freedom. Following Kwiatkowski et al. (1992), we use the Bartlett window with respectively bandwidth rule $\gamma_{1T} = [4(T/100)^{1/4}]$ and $\gamma_{2T} = [12(T/100)^{1/4}]$ to calculate the HAC estimator of the variance.

First, we consider the size of the test if the DGP is not subject to stochastic volatility (i.e. $\theta_3 = 1$, $\theta_4 = 0$ and $\theta_5 = 0$). Table A.1 shows that the size of the test is close to its nominal value for the i.i.d. case (i.e. $\theta_1 = \theta_2 = 0$) but it exceeds the nominal value if there exists serial correlation in the data and we do not use a HAC estimator. If serial correlation is high then, even if we use a HAC estimator, there are still size distortions. As long as the data is not independently distributed, the size based on γ_{2T} is closer to its nominal value. These results are robust among the different copulas.

[Table A.1 about here]

Second, to illustrate the effect of stochastic volatility we set $\theta_1 = \theta_2 = 0$ and let θ_4 and θ_5 take positive values. Table A.2 shows the size of the test as applied to the original data and without HAS estimator. We also give the results for filtered data $y_{it}(\hat{\theta}) = x_{it}/h_{it}(\hat{\theta})$, where $\hat{\theta} = (\hat{\theta}_3, \hat{\theta}_4, \hat{\theta}_5)'$ are the ML estimates. In summary, we have that the test is subject to size distortions if the DGP contains stochastic volatility. Filtering as well as the use of a long-run variance estimator reduces the size distortions. The results based on filtered data are clearly better but we should take into account that in practice the GARCH model might be misspecified.

[Table A.2 about here]

5.2. The power of the test

We first we consider the power of the test against the fixed alternative (6). We assume that C and C^* are from the same copula family with copula parameter corresponding to kendall's tau = 0.25 and to kendall's tau = 0.1, 0.5 and 0.75, respectively. The break point fraction z^* takes the values 0.3, 0.5 and 0.7 and the break magnitude ω takes the values 0, 0.5 and 1. Note that $\omega = 0$ implies that the copula is time-invariant and $\omega = 1$ corresponds to a standard structural break in the copula parameter.

Table A.3 shows that the power is highest if the break occurs around half of the sample ($z^* = 0.5$). The power increases in w and in kendall's tau value of C^*. This is also expected since in both cases the deviation between the copula under the null and alternative hypothesis increases.

[Table A.3 about here]

Second, we simulate the local asymptotic power curves using 50000 replications of 2500 observations for the step pattern. The simulation setup is as above. Figure A.1 shows that, although the test is inconsistent against the local alternative (7), the power goes to 1 if the magnitude of the change in the copula parameter and the weight ω are sufficiently high.

[Figure A.1 about here]

6. Empirical application

Next, we consider stock returns from the US, UK, France, Germany and Japan. The dataset is provided by MSCI and consists of monthly returns from January, 1970 through November, 2009. Longin and Solnik (2001) consider a similar dataset but observed at a different period (January, 1959 through December, 1996).

We model the marginal distributions using a GARCH (1,1) model with Gaussian, Student and Skewed-Student distributed innovations. The model is like (8) but the mean equation only contains a constant term and no lag values. In addition, the innovations $\varepsilon_t^\dagger$ are modelled for each series using a Gaussian, Student or Skewed-Student distribution. All parameters are estimated using maximum likelihood. Using the AIC information criterium, we selected the GARCH model with skewed-student distributed innovations for all countries (except Japan; see below). The disturbances are from a symmetric student distribution if the logarithm of the asymmetry parameter (as reported in Table A.4) equals 0, see Laurent and Peters (2006).

Table A.4 contains the parameter estimates. For Japan we report the model with standard student distributed innovations, since the asymmetry parameter is insignificant. In summary, all reported coefficients are significant at the 5% level except the constant for Germany and θ_3 for the UK and Japan. The θ_4 parameter for the UK is only significant at the 10% level. The result for the UK might be affected by the severe spike at January 1975. Including a dummy variable in the mean equation improved the model. The results for the copula constancy test (not reported here) are almost the same as the ones below. Explicit results are available upon request.

We apply the Ljung-Box test to the standardized residuals as well as the squared standardized residuals. For all countries we do not reject the null of no serial correlation for the squared standardized residuals. For France, Germany and Japan the standardized residuals are serial correlated. As long as the dependence structure satisfies the mixing assumption made in section 3, Theorem 2 allows us to apply the copula constancy test to the standardized innovations.

[Table A.4 about here]

Since quantiles can also change for reasons different from stochastic volatility, we perform the quantile constancy test proposed by Busettti and Harvey (2007). Table A.5 shows that the GARCH (1,1) model performs reasonably well for all countries except Japan. For Japan we detect some time-varying behavior at the lower quantiles (at the 5% and 1% level). It is reasonable that results of the copula constancy tests for Japan are affected by this. Since the purpose of this section is solely to illustrate the effect of stochastic volatility we will not analyze more advanced models for Japan.

[Table A.5 about here]

We apply the copula constancy test to the original return series as well as to the standardized residuals of the GARCH(1,1) models. Table A.6 shows that we clearly reject the null hypothesis for some country pairs at the 5% significance level if we apply the test to the return series and we do not use a HAC estimate. In particular, the range test provides strong evidence against the null hypothesis. However, if we make use of a HAC estimator then we are hardly able to reject the null hypothesis at the 5% level. Applying the test to filtered observations gives a similar result. This example, therefore, clearly illustrates the importance of controlling for changes in the marginal distributions.

[Table A.6 about here]

Finally, we would like to emphasize that we should not conclude that this implies that for some country pairs the copula is time-invariant. Besides the fact that failing to reject the null hypothesis does not imply that the null hypothesis is true, we can indeed reject the null hypothesis if we consider other events than $\{y_{1t} \leq \xi_1(\tau_1), y_{2t} \leq \xi_2(\tau_2)\}$. In particular, using the Quadrant Association Test of Busetti and Harvey (2008) (which is based on the same idea but uses the events $\{y_{1t} \leq \xi_1(\tau_1), y_{2t} \leq \xi_2(\tau_2)\}$ as well as $\{y_{1t} > \xi_1(\tau_1), y_{2t} > \xi_2(\tau_2)\}$) we obtain, even if we control for stochastic volatility, strong evidence against the null hypthesis.

Appendix A. Proofs

The proof of Theorem 1 and 2 follows the one of De Jong et al. (2007). We extend their proof in two ways. First, in our case the indicator series depends on a vector series instead of a scalar series. Second, the indicator series depends on a parameter vector θ which needs to be estimated.

The structure of the proof is the following: Lemma 1 shows uniform convergence for some specific terms that occur in the proof of Theorem 1. To proof Lemma 1, we show pointwise convergence and stochastic equicontinuity in Lemma 2 and 3, respectively.

Lemma 1. *Write $y_t = y_t(\theta_0)$ and $\xi(\tau) = \xi(\theta_0, \tau)$. For $M > 0$, we have under Assumption 1*

$$\sup_{\phi\in[-M,M]^2} \sup_{r\in[0,1]} T^{-1/2} \sum_{t=1}^{[rT]} \left|d_t(\phi) - E[d_t(\phi)]\right| \xrightarrow{p} 0,$$

where

$$d_t(\phi) = I(y_t, \xi(\tau) + \phi T^{-1/2}) - I(y_t, \xi(\tau)) \tag{A.1}$$

Proof. The parameter space of ϕ is compact since it is closed and bounded. Compactness implies that it is also totally bounded (see e.g. Davidson (1994, Theorem 5.5)). Therefore, using Davidson (1994, Theorem 21.9) and noting that $[-M, M]^2$ is dense in the parameter space itself, it is sufficient to show that $\sup_{r\in[0,1]} T^{-1/2} \sum_{t=1}^{[rT]} |d_t(\phi) - E[d_t(\phi)]| \xrightarrow{p} 0$ for each $\phi \in [-M, M]^2$ and that the sequence $\{\sup_{r\in[0,1]} T^{-1/2} \sum_{t=1}^{[rT]} |d_t(\phi) - E[d_t(\phi)]|,\ T = 1, 2, \ldots\}$ is stochastically equicontinuous. Lemma 2 proves pointwise convergence and Lemma 3 proves stochastic equicontinuity. ∎

Lemma 2. *Let $M > 0$. Then, under Assumption 1, for each $\phi \in [-M, M]^2$*

$$\sup_{r\in[0,1]} T^{-1/2} \sum_{t=1}^{[rT]} \left|d_t(\phi) - E[d_t(\phi)]\right| \xrightarrow{p} 0$$

Proof. First, it is sufficient to show that $E \sup_{r\in[0,1]}(T^{-1/2} \sum_{t=1}^{[rT]} |d_t(\phi) - E[d_t(\phi)]|)^2 \to 0$ for $T \to \infty$.

Second, for $p > 2$ (see remark below Assumption 1) and $i = 1, 2$ we have that $\mathbb{I}(y_{it} \leq \xi_i(\tau_i))$ is strong mixing of size $-p/(p - 2)$, because the indicator function $\mathbb{I}(\cdot)$ is a measurable function (Theorem 3.27, Davidson (1994, p.53)) and every measurable transformation of y_{it} is also strong mixing with the same size as y_{it} (Theorem 14.1, Davidson (1994, p.210)). Using the same arguments, $I(y_t, \xi(\tau)) = \mathbb{I}(y_{1t} \leq \xi_1(\tau_1))\mathbb{I}(y_{2t} \leq \xi_2(\tau_2))$ is measurable (Theorem 3.33, Davidson (1994, p.56)) and thus also strong mixing with size $-p/(p-2)$. This implies that we can make use of Lemma 1 in De Jong et al. (2007).

Let $F(\cdot, \cdot)$ denote the joint distribution of $y_{1t} - \xi_1(\tau_1)$ and $y_{2t} - \xi_2(\tau_2)$ and let $F_i'(\cdot, \cdot)$ denote the derivative with respect to argument $i = 1, 2$.

Take $\phi \in [-M, M]^2$ arbitrary. For all $\eta > 0$ (as in Assumption 1.3) there exists a T_0 such that $MT^{-1/2} \leq \eta$ for all $T \geq T_0$. For $T \geq T_0$, we obtain using Lemma 1 of De Jong et al. (2007) and for some constants $c_1 > 0$, $c_2 > 0$ and $c_3 > 0$,

$$
\begin{aligned}
& E \sup_{r\in[0,1]} \left(T^{-1/2} \sum_{t=1}^{[rT]} |d_t(\phi) - E[d_t(\phi)]| \right)^2 \\
\leq \quad & c_1 T^{-1} \sum_{t=1}^{T} \| I(y_t, \xi(\tau) + \phi T^{-1/2}) - I(y_t, \xi(\tau)) \|_p^2 \\
\leq \quad & c_2 T^{-1} \sum_{t=1}^{T} \left(F(MT^{-1/2}, MT^{-1/2}) - F(-MT^{-1/2}, -MT^{-1/2}) \right)^{2/p} \\
\leq \quad & c_3 \left(\sup_{(a_1,a_2)\in[-\eta,\eta]^2} F_1'(a_1, a_2)(2MT^{-1/2}) + \sup_{(a_3,a_4)\in[-\eta,\eta]^2} F_2'(a_3, a_4)(2MT^{-1/2}) \right)^{2/p}
\end{aligned}
$$

where the last inequality follows using the mean value theorem. Since $F_i'(\cdot, \cdot)$, $i = 1, 2$, is finite under assumption 1, letting $T \to \infty$ gives the required result. ■

Lemma 3. *The sequence* $\{\sup_{r\in[0,1]} T^{-1/2} \sum_{t=1}^{[rT]} |d_t(\phi) - E[d_t(\phi)]|,\ T = 1, 2, \ldots\}$ *on the metric space* $([-M, M]^2, \rho)$ *with* $\rho(\phi, \ddot{\phi}) = |\phi_1 - \ddot{\phi}_1| + |\phi_2 - \ddot{\phi}_2|$ *is stochastically equicontinuous.*

Proof. Define

$$
v_T(\phi) := \sup_{r\in[0,1]} T^{-1/2} \sum_{t=1}^{[rT]} |d_t(\phi) - E[d_t(\phi)]|
$$

We have to show (see Davidson (1994, p336)) that for all $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$\limsup_{T\to\infty} P(\sup_{\phi\in[-M,M]^2} \sup_{\ddot{\phi}\in B_\rho(\phi,\delta)} \left|v_T(\phi) - v_T(\ddot{\phi})\right| \geq \varepsilon) < \varepsilon,$$

where $B_\rho(\phi,\delta) = \{\ddot{\phi} : \ddot{\phi} \in [-M,M]^2, \rho(\phi,\ddot{\phi}) < \delta\}$.
Write $\{\phi,\ddot{\phi} : |\phi_i - \ddot{\phi}_i| < \delta\} := \{\phi,\ddot{\phi} \in [-M,M]^2 : |\phi_1 - \ddot{\phi}_1| < \delta,\ |\phi_2 - \ddot{\phi}_2| < \delta\}$.
Then

$$\sup_{\phi\in[-M,M]^2} \sup_{\ddot{\phi}\in B_\rho(\phi,\delta)} \left|v_T(\phi) - v_T(\ddot{\phi})\right| \leq \sup_{\phi,\ddot{\phi}:|\phi_i-\ddot{\phi}_i|<\delta} \left|v_T(\phi) - v_T(\ddot{\phi})\right|$$

Subsequently, we have using the same arguments as in Lemma 2 of De Jong et al. (2007) that

$$\begin{aligned} &P(\sup_{\phi\in[-M,M]^2} \sup_{\ddot{\phi}\in B_\rho(\phi,\delta)} \left|v_T(\phi) - v_T(\ddot{\phi})\right| \geq \varepsilon) \\ &\leq o(1) + 2\mathbb{I}\left(\sup_{\phi,\ddot{\phi}:|\phi_i-\ddot{\phi}_i|<\delta} T^{-1/2}\sum_{j=1}^{T}\left|Ed_j(\phi) - Ed_j(\ddot{\phi})\right| > \varepsilon/4\right) \end{aligned}$$

Therefore, it is sufficient to show equicontinuity of $T^{-1/2}\sum_{j=1}^{T}|Ed_j(\phi) - Ed_j(\phi')|$.

For all $M > 0$ and for all $\eta > 0$ (as in Assumption 1.3) we can find an index in the sequence, T, such that $MT^{-1/2} \leq \eta$. Therefore,

$$\begin{aligned} &\sup_{\phi,\ddot{\phi}:|\phi_i-\ddot{\phi}_i|<\delta} T^{-1/2}\sum_{j=1}^{T}\left|Ed_j(\phi) - Ed_j(\ddot{\phi})\right| \\ &= \sup_{\phi,\ddot{\phi}:|\phi_i-\ddot{\phi}_i|<\delta} T^{-1/2}\sum_{j=1}^{T}\left|F(\phi_1 T^{-1/2},\phi_2 T^{-1/2}) - F(\ddot{\phi}_1 T^{-1/2},\ddot{\phi}_2 T^{-1/2})\right| \\ &\leq \sup_{\phi,\ddot{\phi}:|\phi_i-\ddot{\phi}_i|<\delta} T^{-1/2}\sum_{j=1}^{T}\Bigg(\sup_{(a_1,a_2)\in[-\eta,\eta]^2} F_1'(a_1,a_2)|\phi_1 T^{-1/2} - \ddot{\phi}_1 T^{-1/2}| \\ &\qquad + \sup_{(a_3,a_4)\in[-\eta,\eta]^2} F_2'(a_3,a_4)|\phi_2 T^{-1/2} - \ddot{\phi}_2 T^{-1/2}|\Bigg) \\ &\leq \delta\left(\sup_{(a_1,a_2)\in[-\eta,\eta]^2} F_1'(a_1,a_2) + \sup_{(a_3,a_4)\in[-\eta,\eta]^2} F_2'(a_3,a_4)\right) \end{aligned}$$

Since $F_i'(\cdot,\cdot)$ is finite under assumption 1, selecting δ sufficiently small gives the required result. ■

Proof of Theorem 1:
Define $\phi^* := T^{1/2}(\hat{\xi}(\tau) - \xi(\tau))$ and $d_t(\phi)$ as in (A.1). Then

$$
\begin{aligned}
&\frac{1}{\sigma T^{1/2}} \sum_{t=1}^{[rT]} BIQ(y_t, \hat{\xi}(\tau)) \qquad\qquad (A.2)\\
&= \frac{1}{\sigma T^{1/2}} \sum_{t=1}^{[rT]} BIQ(y_t, \xi(\tau)) - \frac{1}{\sigma T^{1/2}} \sum_{t=1}^{[rT]} d_t(\phi^*) + \frac{[rT]}{T} \frac{1}{\sigma T^{1/2}} \sum_{t=1}^{T} d_t(\phi^*)\\
&= \frac{1}{\sigma T^{1/2}} \sum_{t=1}^{[rT]} BIQ(y_t, \xi(\tau)) - \frac{1}{\sigma T^{1/2}} \sum_{t=1}^{[rT]} (d_t(\phi^*) - E[d_t(\phi^*)])\\
&\quad + \frac{[rT]}{T} \frac{1}{\sigma T^{1/2}} \sum_{t=1}^{T} (d_t(\phi^*) - E[d_t(\phi^*])
\end{aligned}
$$

Under Assumption 3 we have that for all $\epsilon > 0$ there exits a $M > 0$ such that $P(|\phi^*| \geq M) \leq \epsilon$. Therefore, using Lemma 1 and the triangle inequality the second and third term converge uniformly in probability to zero.

It remains to show that $\hat{\sigma}^2 \xrightarrow{p} \sigma^2$. Define the HAC estimate based on the empirical copula and the *population* quantiles as

$$
\tilde{\sigma}^2 := T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} k((t-s)/\gamma_T) \cdot BIQ(y_t, \xi(\tau)) \cdot BIQ(y_s, \xi(\tau)) \qquad (A.3)
$$

The proof consists of two steps. First we show that $\hat{\sigma}^2 \xrightarrow{p} \tilde{\sigma}^2$. Subsequently, we show that $\tilde{\sigma}^2$ is asymptotically equivalent to $\bar{\sigma}^2$ defined in (4). Finally, using Theorem 2.1 in De Jong and Davidson (2000) we have then that $\bar{\sigma}^2 \xrightarrow{p} \sigma^2$.
Step 1: Write

$$
BIQ(y_t, \hat{\xi}(\tau)) = BIQ(y_t, \xi(\tau)) - a_t + b_T \qquad (A.4)
$$

where $a_t := [d_t(\phi^*) - E(d_t(\phi^*))]$ and $b_T := \frac{1}{T}\sum_{k=1}^{T}[d_k(\phi^*) - Ed_k(\phi^*)]$. Then

$$
\hat{\sigma}^2 \;=\; T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} k((t-s)/\gamma_T)(BIQ(y_t, \xi(\tau)) - a_t + b_T)(BIQ(y_s, \xi(\tau)) - a_s + b_T)
$$

The cross products, except the ones consisting of $BIQ(y_t, \hat{\xi}(\tau)) \cdot BIQ(y_s, \hat{\xi}(\tau))$, converge to zero using arguments as in De Jong et al. (2007).

Step 2: It is sufficient that $|\tilde{\sigma}^2 - \bar{\sigma}^2| \xrightarrow{p} 0$. Write $\hat{I}_t := I(y_t, \hat{\xi}(\tau))$. Then

$$\begin{aligned}
|\tilde{\sigma}^2 - \bar{\sigma}^2| &= \Bigg| T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} k\left((t-s)/\gamma_T\right) \\
&\quad \times \Bigg\{ [C_T(\tau_1,\tau_2) - \hat{I}_t][C_T(\tau_1,\tau_2) - \hat{I}_s] - [C(\tau_1,\tau_2) - \hat{I}_t][C(\tau_1,\tau_2) - \hat{I}_s] \Bigg\} \Bigg| \\
&= \Bigg| T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} k\left((t-s)/\gamma_T\right) \times \Bigg\{ C_T^2(\tau_1,\tau_2) - C^2(\tau_1,\tau_2) \\
&\quad + \hat{I}_t(C(\tau_1,\tau_2) - C_T(\tau_1,\tau_2)) + \hat{I}_s(C(\tau_1,\tau_2) - C_T(\tau_1,\tau_2)) \Bigg\} \Bigg|
\end{aligned}$$

For some constant $c_1 > 0$

$$\begin{aligned}
&|T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} k\left((t-s)/\gamma_T\right) [(C(\tau_1,\tau_2) - C_T(\tau_1,\tau_2))\hat{I}_t]| \\
&= |\sqrt{T}(C(\tau_1,\tau_2) - C_T(\tau_1,\tau_2)) \cdot T^{-3/2} \sum_{t=1}^{T} \sum_{s=1}^{T} k\left((t-s)/\gamma_T\right) \hat{I}_t| \\
&\leq |T^{-1/2} \sum_{m=1}^{T} (C(\tau_1,\tau_2) - \hat{I}_m) \cdot c_1 T^{-3/2} \sum_{t=1}^{T} \sum_{j=-T}^{T} k(j/\gamma_T)| \\
&= |c_1 \cdot O_p(1) \cdot \frac{\gamma_T}{\sqrt{T}} \cdot \frac{1}{\gamma_T} \sum_{j=-T}^{T} k(j/\gamma_T)| \\
&= O_p(\gamma_T/\sqrt{T})
\end{aligned}$$

and likewise for some constant $c_2 > 0$

$$\begin{aligned}
&|T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} k\left((t-s)/\gamma_T\right) [C_T^2(\tau_1,\tau_2) - C^2(\tau_1,\tau_2)] \\
&= |[C_T(\tau_1,\tau_2) + C(\tau_1,\tau_2)][C_T(\tau_1,\tau_2) - C(\tau_1,\tau_2)]\sqrt{T} \cdot T^{-3/2} \sum_{t=1}^{T} \sum_{s=1}^{T} k\left((t-s)/\gamma_T\right)| \\
&\leq c_2 \cdot O_p(1) \cdot O_p(\gamma_T/\sqrt{T})
\end{aligned}$$

Noting that, under assumption 2, $\gamma_T/\sqrt{T} \to 0$ as $T \to \infty$ completes the proof. ∎

The following lemmas are used to prove Theorem 2. The structure of the proof is similar as for Theorem 1. Lemma 4 shows uniform convergence of some terms that occur in the proof of Theorem 2 below. To proof Lemma 4, we show pointwise convergence in Lemma 5 and stochastic equicontinuity in Lemma 6.

Lemma 4. *For $M > 0$ and $N > 0$ we have under Assumption 1 and 4*

$$\sup_{r\in[0,1]} \sup_{\phi\in[-M,M]^2} \sup_{\zeta\in\Upsilon} T^{-1/2} \sum_{t=1}^{[rT]} \left| d_t^*(\phi,\zeta) - E[d_t^*(\phi,\zeta)] \right| \xrightarrow{p} 0$$

where $\Upsilon := \{\zeta \in [-N,N]^q : \theta_0 + \zeta T^{-1/2} \in \Theta\}$ *and*

$$\begin{aligned} d_t^*(\phi,\zeta) &= I[y_t(\theta_0+\zeta T^{-1/2}), \xi(\theta_0+\zeta T^{-1/2},\tau) + \phi T^{-1/2}] \\ &\quad - I[y_t(\theta_0), \xi(\theta_0,\tau)] \end{aligned}$$

Proof. Write $d_t^*(\phi,\zeta) = d_t^\dagger(\phi,\zeta) + d_t(\phi)$, where

$$\begin{aligned} d_t^\dagger(\phi,\zeta) &= I[y_t(\theta_0+\zeta T^{-1/2}), \xi(\theta_0+\zeta T^{-1/2},\tau) + \phi T^{-1/2}] \\ &\quad - I[y_t(\theta_0), \xi(\theta_0,\tau) + \phi T^{-1/2}] \end{aligned}$$

and $d_t(\phi)$ as defined in (A.1).

Using the triangle inequality

$$\begin{aligned} & T^{-1/2} \sum_{t=1}^{[rT]} \left| d_t^*(\phi,\zeta) - E[d_t^*(\phi,\zeta)] \right| \\ & \quad \le T^{-1/2} \sum_{t=1}^{[rT]} \left| d_t^\dagger(\phi,\zeta) - Ed_t^\dagger(\phi,\zeta) \right| + T^{-1/2} \sum_{t=1}^{[rT]} \left| d_t(\phi) - Ed_t(\phi) \right| \end{aligned}$$

Since the second part converge uniformly to zero by Lemma 1, it is sufficient to prove that the first part converge uniformly to zero as well. Using the same arguments as in Lemma 1, it is sufficient to show that for each $(\phi,\zeta) \in [-M,M]^2 \times \Upsilon$, $\sup_{r\in[0,1]} \; T^{-1/2} \sum_{t=1}^{[rT]} \left| d_t^\dagger(\phi,\zeta) - Ed_t^\dagger(\phi,\zeta \right| \xrightarrow{p} 0$ and that the sequence $\{\sup_{r\in[0,1]} \; T^{-1/2} \sum_{t=1}^{[rT]} \left| d_t^\dagger(\phi,\zeta) - Ed_t^\dagger(\phi,\zeta) \right|,\; T = 1,2,\ldots\}$ is stochastically equicontinuous. Lemma 5 proves pointwise convergence and Lemma 6 proves stochastic equicontinuity. ∎

Lemma 5. *Let $M > 0$ and $N > 0$. Then, under Assumption 1 and 4, for each $(\phi, \zeta) \in [-M, M]^2 \times [-N, N]^q$*

$$\sup_{r\in[0,1]} T^{-1/2} \sum_{t=1}^{[rT]} \left| d_t^\dagger(\phi,\zeta) - Ed_t^\dagger(\phi,\zeta) \right| \xrightarrow{p} 0$$

Proof. By Lemma 1 in De Jong et al. (2007) and the mean value theorem we have for some points θ_1^* and θ_2^*, constants $c_1 > 0$, $c_2 > 0$ and $c_3 > 0$, and p as defined in Assumption 1

$$\begin{aligned}
&E \sup_{r\in[0,1]} \left| T^{-1/2} \sum_{t=1}^{[rT]} d_t^\dagger(\phi,\zeta) - Ed_t^\dagger(\phi,\zeta) \right| \\
&\le c_1 T^{-1} \sum_{t=1}^{T} \Bigg(E \Big| I[y_t(\theta_0 + \zeta T^{-1/2}), \xi(\theta_0 + \zeta T^{-1/2}, \tau) + \phi T^{-1/2}] \\
&\qquad - I[y_t(\theta_0), \xi(\theta_0, \tau) + \phi T^{-1/2}] \Big|^p \Bigg)^{2/p} \\
&= c_1 T^{-1} \sum_{t=1}^{T} \Bigg(E \Big| I[y_t(\theta_0) + \left.\frac{\partial y_t}{\partial \theta}\right|_{\theta_1^*} \cdot \zeta T^{-1/2}, \\
&\qquad \xi(\theta_0, \tau) + \left.\frac{\partial \xi}{\partial \theta}\right|_{\theta_2^*} \cdot \zeta T^{-1/2} + \phi T^{-1/2}] \\
&\qquad - I[y_t(\theta_0), \xi(\theta_0, \tau) + \phi T^{-1/2}] \Big|^p \Bigg)^{2/p} \\
&\le c_2 T^{-1} \sum_{t=1}^{T} \Bigg(F([M + c_3 N]T^{-1/2}, [M + c_3 N]T^{-1/2}) \\
&\qquad - F(-[M + c_3 N]T^{-1/2}, -[M + c_3 N]T^{-1/2}) \Bigg)^{2/p}
\end{aligned}$$

The last expression converges to zero as $T \to \infty$ using the same arguments as in Lemma 2. ∎

Lemma 6. *The sequence $\{\sup_{r\in[0,1]} T^{-1/2} \sum_{t=1}^{[rT]} \left| d_t^\dagger(\phi,\zeta) - Ed_t^\dagger(\phi,\zeta) \right|,\ T = 1, 2, \ldots\}$ on the metric space $([-M, M]^2 \times [-N, N]^q, \rho)$ with $\rho((\phi,\zeta), (\ddot{\phi}, \ddot{\zeta})) = |\phi_1 - \ddot{\phi}_1| + |\phi_2 - \ddot{\phi}_2| + \sum_{j=1}^{q} |\zeta_j - \ddot{\zeta}_j|$ is stochastically equicontinuous.*

Proof. Using the same arguments as in Lemma 3, it is sufficient to establish stochastic equicontinuity of $T^{-1/2}\sum_{j=1}^{T}\left|Ed_j^{\dagger}(\phi,\zeta)-Ed_j^{\dagger}(\ddot{\phi},\ddot{\zeta})\right|$, where $\ddot{\phi}\in B_\rho(\phi,\delta_\phi)$ and $\ddot{\zeta}\in B_\rho(\zeta,\delta_\zeta)$ and with scalars $\delta_\phi>0$, $\delta_\zeta>0$.

Define the $q\times 1$ vectors

$$c_j^* := -\left.\frac{\partial y_j(\theta)}{\partial\theta}\right|_{\theta_0}+\left.\frac{\partial\xi_j(\theta,\tau_j)}{\partial\theta}\right|_{\theta_0} \qquad j=1,\dots,T$$

and let F_i' and F_{ik}'' denote, respectively, the first and second derivative of F with respect to argument i and k where $i,k\in\{1,2\}$. Using the mean value theorem and a first order Taylor expansion

$$\begin{aligned}
&T^{-1/2}\sum_{j=1}^{T}\left|Ed_j^{\dagger}(\phi,\zeta)-Ed_j^{\dagger}(\ddot{\phi},\ddot{\zeta})\right|\\
&=T^{-1/2}\sum_{j=1}^{T}\Bigg|F\big(T^{-1/2}\zeta'c_j^*+\phi_1T^{-1/2},T^{-1/2}\zeta'c_j^*+\phi_2T^{-1/2}\big)-\ F(\phi_1T^{-1/2},\phi_2T^{-1/2})\\
&\qquad-\left\{F\big(T^{-1/2}\ddot{\zeta}'c_j^*+\ddot{\phi}_1T^{-1/2},T^{-1/2}\ddot{\zeta}'c_j^*+\ddot{\phi}_2T^{-1/2}\big)-\ F(\ddot{\phi}_1T^{-1/2},\ddot{\phi}_2T^{-1/2})\right\}\Bigg|\\
&=T^{-1/2}\sum_{j=1}^{T}\Bigg|T^{-1/2}\zeta'c_j^*[F_1'(\phi_1T^{-1/2},\phi_2T^{-1/2})+F_2'(\phi_1T^{-1/2},\phi_2T^{-1/2})]+O(T^{-1})\\
&\qquad-\left\{T^{-1/2}\ddot{\zeta}'c_j^*[F_1'(\ddot{\phi}_1T^{-1/2},\ddot{\phi}_2T^{-1/2})+F_2'(\ddot{\phi}_1T^{-1/2},\ddot{\phi}_2T^{-1/2})]+O(T^{-1})\right\}\Bigg|\\
&\le\sup_{j=1,\dots,T}|\zeta'c_j^*|\cdot\big|F_1'(\phi_1T^{-1/2},\phi_2T^{-1/2})+F_2'(\phi_1T^{-1/2},\phi_2T^{-1/2})\\
&\qquad\qquad-F_1'(\ddot{\phi}_1T^{-1/2},\ddot{\phi}_2T^{-1/2})-F_2'(\ddot{\phi}_1T^{-1/2},\ddot{\phi}_2T^{-1/2})\big|\\
&\quad+\sup_{j=1,\dots,T}|(\zeta'-\ddot{\zeta}')\cdot c_j^*|\cdot|F_1'(\ddot{\phi}_1T^{-1/2},\ddot{\phi}_2T^{-1/2})+F_2'(\ddot{\phi}_1T^{-1/2},\ddot{\phi}_2T^{-1/2})|+O(T^{-1/2})\\
&\le\sup_{j=1,\dots,T}|\zeta'c_j^*|\cdot\big\{|F_{11}''(b_1,b_2)+F_{21}''(b_3,b_4)|\cdot|\ddot{\phi}_1T^{-1/2}-\phi_1T^{-1/2}|\\
&\qquad\qquad+|F_{12}''(b_1,b_2)+F_{22}''(b_3,b_4)|\cdot|\ddot{\phi}_2T^{-1/2}-\phi_2T^{-1/2}|\big\}\\
&\quad+\sup_{j=1,\dots,T}|(\zeta'-\ddot{\zeta}')\cdot c_j^*|\cdot|F_1'(\ddot{\phi}_1T^{-1/2},\ddot{\phi}_2T^{-1/2})+F_2'(\ddot{\phi}_1T^{-1/2},\ddot{\phi}_2T^{-1/2})|+O(T^{-1/2})
\end{aligned}$$

where (b_1,b_2) and (b_3,b_4) are points between $(\phi_1T^{-1/2},\phi_2T^{-1/2})$ and $(\ddot{\phi}_1T^{-1/2},\ddot{\phi}_2T^{-1/2})$. Under Assumption 4, F_i', F_{ik}'' and c_j^* are bounded, so that for some constants

c_1 and c_2

$$T^{-1/2}\sum_{j=1}^{T}\left|Ed_j^{\dagger}(\phi,\zeta)-Ed_j^{\dagger}(\ddot{\phi},\ddot{\zeta})\right| \quad \leq \quad c_1T^{-1/2}\delta_{\phi}+c_2\delta_{\zeta}+O(T^{-1/2})$$

Selecting δ_{ϕ} and δ_{ζ} sufficiently small completes the proof. ∎

Proof of Theorem 2
Using $\phi^* = T^{1/2}(\hat{\xi}(\tau)-\xi(\tau))$ and $\zeta^* = T^{1/2}(\hat{\theta}_T-\theta_0)$, write

$$\begin{aligned}
&\frac{1}{\sigma\sqrt{T}}\sum_{t=1}^{[rT]} BIQ(y_t(\hat{\theta}_T),\hat{\xi}(\hat{\theta},\tau)) \\
&\quad = \frac{1}{\sigma T^{1/2}}\sum_{t=1}^{[rT]} BIQ(y_t(\theta_0),\xi(\theta_0,\tau)) - \frac{1}{\sigma T^{1/2}}\sum_{t=1}^{[rT]} d_t^*(\phi^*,\zeta^*)-E[d_t^*(\phi^*,\zeta^*)] \\
&\qquad + \frac{[rT]}{T}\frac{1}{\sigma T^{1/2}}\sum_{t=1}^{T} d_t^*(\phi^*,\zeta^*)-E[d_t^*(\phi^*,\zeta^*)]
\end{aligned}$$

Following Davidson (1994, Theorem 21.6), the second and third term converge to zero if (a) $\hat{\theta}_T \xrightarrow{p} \theta_0$ and (b) it converge uniformly. We have consistency by assumption and by selecting M and N sufficiently large we have, using the same arguments as Theorem 1, uniform convergence by Lemma 4.

Remains to show that $\hat{\sigma}^2(\hat{\theta})$ is asymptotically equivalent to σ^2. Write

$$\hat{\sigma}^2(\hat{\theta}_T)-\sigma^2 \quad = \quad \left(\hat{\sigma}^2(\hat{\theta}_T)-\hat{\sigma}^2\right)+\left(\hat{\sigma}^2-\sigma^2\right)$$

The last part converge to zero by Theorem 1. It is sufficient to show that

$$|\hat{\sigma}^2(\theta)-\hat{\sigma}^2| \xrightarrow{p} 0$$

Write

$$\begin{aligned}
BIQ(y_t(\hat{\theta}_T),\hat{\xi}(\hat{\theta}_T,\tau)) \quad &= \quad BIQ(y_t(\theta_0),\xi(\theta_0,\tau)) - [d_t^*(\phi^*,\zeta^*)-E(d_t^*(\phi^*,\zeta^*))] \\
&\qquad + \frac{1}{T}\sum_{k=1}^{T}[d_k^*(\phi^*,\zeta^*)-E(d_k^*(\phi^*,\zeta^*))] \\
&=: \quad BIQ(y_t(\theta_0),\xi(\theta_0,\tau)) - a_t(\hat{\theta}_T)+b_T(\hat{\theta}_T)
\end{aligned}$$

So that

$$|\hat{\sigma}^2(\hat{\theta}_T) - \hat{\sigma}^2| = \Big|T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T} k((t-s)/\gamma_T)$$
$$\times \{(BIQ(y_t(\theta_0), \xi(\theta_0,\tau)) - a_t(\hat{\theta}_T) + b_T(\hat{\theta}_T))(BIQ(y_s(\theta_0), \xi(\theta_0,\tau)) - a_s(\hat{\theta}_T) + b_T(\hat{\theta}_T))$$
$$- (BIQ(y_t(\theta_0), \xi(\theta_0,\tau)) - a_t + b_T)(BIQ(y_s(\theta_0), \xi(\theta_0,\tau)) - a_s + b_T)\}\Big| \tag{A.5}$$

where a_{Tt} and b_T are defined in (A.4). We show that the difference of the cross-products converge to zero. Write

$$|T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T} k((t-s)/\gamma_T)\{a_t(\hat{\theta}_T)a_s(\hat{\theta}_T) - a_t a_s\}|$$
$$= \Big|T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T} k((t-s)/\gamma_T) \tag{A.6}$$
$$\times \{\frac{1}{2}(a_t(\hat{\theta}_T) + a_t)(a_s(\hat{\theta}_T) - a_s) + \frac{1}{2}(a_t(\hat{\theta}_T) - a_t)(a_s(\hat{\theta}_T) + a_s)\}\Big|$$

Since for constants $c_1 > 0$ and $c_2 > 0$

$$|T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T} k((t-s)/\gamma_T)\frac{1}{2}(a_t(\hat{\theta}_T) - a_t)(a_s(\hat{\theta}_T) + a_s)|$$
$$\leq c_1|T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T} k((t-s)/\gamma_T(a_t(\hat{\theta}_T) - a_t)|$$
$$\leq c_2|T^{-1/2}\sum_{t=1}^{T}(a_t(\hat{\theta}_T) - a_t) \times T^{-3/2}\sum_{j=1}^{T} k(j/\gamma_T)|$$
$$= O_p(\gamma/\sqrt{T})$$

because the first term is $o_p(1)$ by Lemma 4. Using the same idea for the first term in (A.6) gives the required result.

Furthermore,

$$
\begin{aligned}
&|T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T}k((t-s)/\gamma_T)\{b_T^2(\hat{\theta}_T)-b_T^2\}| \\
&\quad = \left|T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T}k((t-s)/\gamma_T)\{(b_T(\hat{\theta}_T)+b_T)(b_T(\hat{\theta}_T)-b_T)\}\right| \\
&\qquad \le c_1|b_T(\hat{\theta}_T)-b_T|\cdot T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T}k((t-s)/\gamma_T) \\
&\qquad \le c_2T^{-1/2}|b_T(\hat{\theta}_T)-b_T|\cdot T^{-1/2}\sum_{j=-T}^{T}k(j/\gamma_T)
\end{aligned}
$$

where $T^{-1/2}|b_T(\hat{\theta}_T)-b_T| = o_p(1)$ by Lemma 4.

The other cross products in (A.5) follow a similar argument.

■

Proof of Theorem 3

Define the bivariate τ-quantics corresponding to the alternative hypothesis (6) as

$$
BIQ^1(y_t,\xi(\tau)) \quad := \quad Q^1(y_t,\xi(\tau)) - \frac{1}{T}\sum_{j=1}^{T}Q^1(y_j,\xi(\tau))
$$

Then

$$
BIQ(y_t,\xi(\tau)) = BIQ^1(y_t,\xi(\tau)) + \Delta_C\cdot[g(t,T) - \frac{1}{T}\sum_{j=1}^{T}g(j,T)] \qquad \text{(A.7)}
$$

where $\Delta_C := C(\tau_1,\tau_2) - C^*(\tau_1,\tau_2)$. Furthermore,

$$
\begin{aligned}
&1/(\sqrt{T})\sum_{t=1}^{[rT]}BIQ(y_t,\xi(\tau)) \\
&\quad = 1/(\sqrt{T})\sum_{t=1}^{[rT]}BIQ^1(y_t,\xi(\tau)) + \Delta_C\cdot\omega\frac{[z^*T]}{\sqrt{T}}\left(\frac{[rT]}{T}-1\right) \qquad \text{(A.8)}
\end{aligned}
$$

Provided that $T^{-1}E\left(\sum_{t=1}^{T}Q^1(y_t,\xi(\tau))\right)^2 \to \sigma_1^2$ for $\sigma_1^2 \in (0,\infty)$, we have under the alternative hypothesis that the first term on the right hand side is $O_p(1)$ and the last term $O_p(\sqrt{T})$ for $r \neq 1$.

Note that we used the population quantiles $\xi(\tau)$ instead of the sample quantiles $\hat{\xi}(\tau)$. This does not affect the result in view of the fact that we can rewrite $1/(\sqrt{T})\sum_{t=1}^{[rT]} BIQ(y_t, \hat{\xi}(\tau))$ like (A.2) in Appendix A and show using similar arguments as in the proof of Theorem 1 (but with $F(\cdot,\cdot)$ this time depending on t) that this is asymptotically equivalent to the $1/(\sqrt{T})\sum_{t=1}^{[rT]} BIQ(y_t, \hat{\xi}(\tau))$.

We show $\hat{\sigma}^2$ is $O_p(\sqrt{T})$ under the alternative. From (5) and (A.7)

$$\hat{\sigma}^2 = T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T} k((t-s)/\gamma_T)\left\{BIQ^1(y_t,\hat{\xi}(\tau)) + \tilde{g}_t\right\}\left\{BIQ^1(y_s,\hat{\xi}(\tau)) + \tilde{g}_s\right\}$$

where $\tilde{g}_t := \Delta_C \cdot [g(t,T) - \frac{1}{T}\sum_{j=1}^{T} g(j,T)]$. Then

$$\begin{aligned}\left|\hat{\sigma}^2 - \sigma_1^2\right| = \Bigg| T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T} k((t-s)/\gamma_T)\Bigg\{ & BIQ^1(y_t,\hat{\xi}(\tau))BIQ^1(y_s,\hat{\xi}(\tau)) \qquad (A.9)\\ - BIQ^1(y_t,\hat{\xi}(\tau))\tilde{g}_t - BIQ^1(y_s,\hat{\xi}(\tau))\tilde{g}_s + \tilde{g}_t\tilde{g}_s & - Q^1(y_t,\xi(\tau))Q^1(y_s,\xi(\tau))\Bigg|\Bigg\}\end{aligned}$$

Note

$$\begin{aligned}\left|T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T} k((t-s)/\gamma_T)\tilde{g}_t\tilde{g}_s\right| &\leq c_1|T^{-1}\sum_{t=1}^{T}\sum_{j=-T}^{T} k(j/\gamma_T)| \\ &= c_1|\sqrt{T}\frac{\gamma_T}{\sqrt{T}}\sum_{j=-T}^{T} k(j/\gamma_T)| \\ &= \sqrt{T}O_p(\gamma_T/\sqrt{T})\end{aligned}$$

is $O_p(\sqrt{T})$ since $\gamma_T/\sqrt{T} \to 0$ as $T \to \infty$ under assumption 2. The other cross products require similar arguments and thus $\hat{\sigma}^2$ is $O_p(\sqrt{T})$.

Combining the previous results we have that the tests are $O_p(\sqrt{T})$. In other words, the square, maximum and range statistics defined in section 2 become infinity large as $T \to \infty$, and thus the probability that the test statistic exceeds the critical value goes to 1 as $T \to \infty$.

■

Proof of Theorem 4

The asymptotic distribution follows from (A.7) and the FCLT. Setting $g(t/T) = T^{-1/2}h(t/T)$ in (A.9) gives $|\hat{\sigma}^2 - \sigma_1^2| = o_p(1)$ using arguments as in the proof of Theorem 3.

■

Table A.1: Size of *squares* test using no HAC estimator and a HAC estimator based on a Bartlett kernel with bandwidth γ_{1T} and γ_{2T}, respectively.

			no HAC				γ_{1T}				γ_{2T}			
	θ_1	θ_2	τ 0.1	0.25	0.5	0.75	0.1	0.25	0.5	0.75	0.1	0.25	0.5	0.75
	-0.5	0	0.053	0.044	0.021	0.017	0.050	0.052	0.046	0.044	0.046	0.047	0.045	0.045
	0.0	0	0.049	0.050	0.052	0.050	0.048	0.048	0.051	0.048	0.045	0.046	0.048	0.046
	0.3	0	0.078	0.096	0.122	0.122	0.053	0.055	0.061	0.059	0.046	0.047	0.050	0.049
	0.5	0	0.129	0.180	0.230	0.233	0.060	0.071	0.078	0.078	0.048	0.052	0.054	0.055
Cl	0.7	0	0.259	0.379	0.463	0.473	0.083	0.108	0.128	0.130	0.052	0.059	0.067	0.067
	0.9	0	0.631	0.833	0.913	0.918	0.193	0.309	0.375	0.379	0.072	0.117	0.143	0.145
	0.0	0.3	0.068	0.084	0.096	0.099	0.050	0.053	0.054	0.055	0.045	0.048	0.048	0.048
	0.0	0.5	0.085	0.104	0.119	0.121	0.052	0.054	0.056	0.057	0.045	0.047	0.047	0.049
	-0.5	0	0.054	0.043	0.018	0.016	0.051	0.052	0.044	0.043	0.046	0.047	0.043	0.043
	0.0	0	0.049	0.050	0.050	0.050	0.048	0.049	0.048	0.049	0.045	0.045	0.045	0.044
	0.3	0	0.071	0.100	0.122	0.125	0.051	0.057	0.060	0.061	0.046	0.049	0.048	0.049
	0.5	0	0.114	0.181	0.236	0.239	0.056	0.070	0.080	0.077	0.045	0.051	0.054	0.052
Ga	0.7	0	0.236	0.377	0.473	0.478	0.076	0.110	0.129	0.132	0.047	0.060	0.068	0.066
	0.9	0	0.597	0.827	0.911	0.917	0.181	0.309	0.383	0.381	0.066	0.117	0.147	0.147
	0.0	0.3	0.062	0.080	0.095	0.094	0.048	0.053	0.054	0.053	0.044	0.047	0.046	0.046
	0.0	0.5	0.077	0.102	0.121	0.120	0.049	0.055	0.057	0.055	0.044	0.049	0.049	0.047
	-0.5	0	0.052	0.043	0.019	0.019	0.051	0.050	0.046	0.045	0.045	0.045	0.046	0.044
	0.0	0	0.050	0.049	0.050	0.049	0.049	0.048	0.049	0.048	0.045	0.046	0.046	0.044
	0.3	0	0.073	0.096	0.122	0.123	0.052	0.054	0.059	0.058	0.045	0.046	0.049	0.048
	0.5	0	0.118	0.183	0.235	0.240	0.055	0.072	0.078	0.079	0.045	0.052	0.054	0.054
St4	0.7	0	0.241	0.370	0.468	0.477	0.077	0.106	0.129	0.132	0.049	0.059	0.067	0.069
	0.9	0	0.601	0.824	0.912	0.919	0.182	0.304	0.377	0.381	0.065	0.116	0.143	0.147
	0.0	0.3	0.065	0.080	0.096	0.096	0.049	0.051	0.054	0.054	0.045	0.046	0.047	0.048
	0.0	0.5	0.079	0.099	0.119	0.122	0.049	0.051	0.056	0.055	0.043	0.045	0.048	0.047

The DGP is given by (8) with $\theta_4 = \theta_5 = 0$ and $\theta_3 = 1$. The innovations are simulated from a Clayton (Cl), Gaussian (Ga) and Student (St4) copula with 4 degrees of freedom. The copula parameter corresponds to kendall's tau = 0.25. #obs. = 500, #rep. = 50000 and nominal size = 5%.

Table A.2: Size of *squares* test using no HAC estimator and a HAC estimator based on a Bartlett kernel with bandwidth γ_{2T}

					no HAC				γ_{2T}				no HAC (filtered data)			
	θ_3	θ_4	θ_5	τ	0.1	0.25	0.5	0.75	0.1	0.25	0.5	0.75	0.1	0.25	0.5	0.75
	0.2	0.1	0.70		0.071	0.064	0.051	0.062	0.049	0.049	0.047	0.048	0.048	0.049	0.050	0.050
	0.2	0.1	0.80		0.097	0.080	0.051	0.076	0.060	0.057	0.046	0.056	0.047	0.050	0.051	0.051
	1.0	0.1	0.70		0.071	0.062	0.050	0.059	0.049	0.048	0.046	0.047	0.048	0.048	0.050	0.050
	1.0	0.1	0.80		0.097	0.079	0.049	0.077	0.061	0.057	0.046	0.056	0.050	0.049	0.049	0.049
Cl	1.0	0.1	0.85		0.151	0.112	0.048	0.114	0.090	0.079	0.044	0.077	0.048	0.049	0.048	0.051
	1.0	0.2	0.70		0.143	0.103	0.051	0.105	0.069	0.064	0.047	0.063	0.048	0.049	0.051	0.052
	1.0	0.2	0.75		0.224	0.157	0.050	0.174	0.107	0.088	0.046	0.098	0.048	0.050	0.049	0.051
	0.2	0.1	0.70		0.066	0.064	0.050	0.062	0.047	0.049	0.046	0.048	0.049	0.050	0.050	0.049
	0.2	0.1	0.80		0.084	0.077	0.051	0.078	0.056	0.056	0.046	0.057	0.048	0.049	0.050	0.050
	1.0	0.1	0.70		0.065	0.064	0.051	0.063	0.048	0.051	0.045	0.049	0.048	0.052	0.050	0.050
	1.0	0.1	0.80		0.085	0.076	0.051	0.078	0.056	0.056	0.046	0.055	0.047	0.050	0.051	0.050
Ga	1.0	0.1	0.85		0.131	0.110	0.052	0.120	0.082	0.076	0.048	0.081	0.047	0.048	0.051	0.051
	1.0	0.2	0.70		0.122	0.102	0.051	0.106	0.063	0.063	0.046	0.064	0.048	0.050	0.051	0.050
	1.0	0.2	0.75		0.195	0.155	0.052	0.182	0.095	0.089	0.046	0.100	0.047	0.049	0.050	0.052
	0.2	0.1	0.70		0.067	0.061	0.051	0.066	0.048	0.048	0.046	0.049	0.048	0.048	0.050	0.049
	0.2	0.1	0.80		0.090	0.078	0.051	0.086	0.059	0.055	0.047	0.059	0.050	0.050	0.051	0.050
	1.0	0.1	0.70		0.066	0.062	0.050	0.066	0.048	0.048	0.046	0.050	0.049	0.049	0.049	0.052
	1.0	0.1	0.80		0.089	0.078	0.049	0.085	0.057	0.055	0.045	0.058	0.049	0.050	0.049	0.049
St4	1.0	0.1	0.85		0.139	0.112	0.050	0.137	0.086	0.076	0.047	0.090	0.048	0.050	0.050	0.052
	1.0	0.2	0.70		0.134	0.103	0.050	0.121	0.067	0.062	0.045	0.067	0.050	0.051	0.050	0.051
	1.0	0.2	0.75		0.212	0.157	0.051	0.206	0.103	0.089	0.048	0.106	0.048	0.049	0.049	0.050

The DGP is given by (8) with $\theta_1 = \theta_2 = 0$. The innovations are simulated from a Clayton (Cl), Gaussian (Ga) and Student (St4) copula with 4 degrees of freedom. The copula parameter corresponds to kendall's tau = 0.25. #obs. = 500, #rep. = 50000 and nominal size = 5%.

Table A.3: Power of *squares* test against the fixed alternative (6).

				kendall(C^*) = 0.1				kendall(C^*) = 0.5				kendall(C^*) = 0.75		
	z^*	w	τ	0.1	0.25	0.5	0.75	0.1	0.25	0.5	0.75	0.1	0.25	0.5	0.75
	0.3	0		0.044	0.045	0.046	0.047	0.044	0.045	0.046	0.044	0.045	0.047	0.044	0.046
	0.3	0.5		0.069	0.073	0.056	0.045	0.067	0.089	0.076	0.055	0.105	0.175	0.187	0.106
	0.3	1		0.158	0.159	0.094	0.054	0.137	0.212	0.187	0.088	0.267	0.509	0.576	0.310
	0.5	0		0.045	0.047	0.045	0.044	0.045	0.045	0.046	0.046	0.045	0.045	0.045	0.045
Cl	0.5	0.5		0.075	0.078	0.063	0.048	0.091	0.117	0.097	0.060	0.155	0.260	0.261	0.137
	0.5	1		0.179	0.198	0.119	0.060	0.211	0.318	0.257	0.103	0.432	0.709	0.748	0.420
	0.7	0		0.045	0.046	0.045	0.047	0.046	0.048	0.046	0.047	0.047	0.046	0.046	0.045
	0.7	0.5		0.059	0.064	0.057	0.048	0.084	0.096	0.082	0.056	0.134	0.202	0.195	0.103
	0.7	1		0.112	0.135	0.090	0.053	0.180	0.248	0.191	0.084	0.355	0.571	0.596	0.294
	0.3	0		0.044	0.047	0.044	0.046	0.046	0.046	0.047	0.045	0.045	0.046	0.045	0.046
	0.3	0.5		0.057	0.062	0.057	0.050	0.063	0.078	0.075	0.060	0.127	0.175	0.166	0.118
	0.3	1		0.092	0.111	0.095	0.065	0.117	0.174	0.171	0.116	0.326	0.510	0.517	0.354
	0.5	0		0.045	0.046	0.045	0.045	0.044	0.045	0.045	0.046	0.045	0.044	0.047	0.047
Ga	0.5	0.5		0.057	0.065	0.064	0.053	0.081	0.097	0.093	0.067	0.206	0.268	0.235	0.150
	0.5	1		0.098	0.135	0.118	0.077	0.185	0.259	0.234	0.146	0.544	0.718	0.689	0.470
	0.7	0		0.045	0.046	0.044	0.045	0.044	0.045	0.044	0.046	0.044	0.045	0.046	0.046
	0.7	0.5		0.049	0.056	0.057	0.050	0.081	0.087	0.076	0.060	0.180	0.208	0.175	0.113
	0.7	1		0.067	0.095	0.092	0.067	0.161	0.203	0.176	0.110	0.469	0.589	0.535	0.331
	0.3	0		0.044	0.045	0.045	0.045	0.045	0.047	0.045	0.046	0.045	0.046	0.045	0.046
	0.3	0.5		0.054	0.060	0.058	0.050	0.059	0.073	0.075	0.061	0.107	0.165	0.167	0.111
	0.3	1		0.079	0.101	0.096	0.066	0.101	0.164	0.169	0.109	0.272	0.476	0.518	0.334
	0.5	0		0.044	0.045	0.047	0.047	0.046	0.043	0.044	0.047	0.046	0.045	0.046	0.045
St4	0.5	0.5		0.052	0.064	0.062	0.053	0.075	0.096	0.093	0.067	0.164	0.250	0.237	0.144
	0.5	1		0.081	0.121	0.118	0.073	0.149	0.240	0.235	0.140	0.444	0.681	0.690	0.452
	0.7	0		0.044	0.046	0.046	0.046	0.047	0.045	0.046	0.044	0.044	0.047	0.046	0.046
	0.7	0.5		0.048	0.054	0.057	0.052	0.071	0.083	0.078	0.060	0.141	0.191	0.176	0.109
	0.7	1		0.061	0.091	0.091	0.065	0.133	0.189	0.175	0.106	0.371	0.551	0.532	0.315

The DGP is given by (8) with $\theta_1 = \theta_2 = \theta_4 = \theta_5 = 0$ and $\theta_3 = 1$. The parameter of the copula C corresponds to kendall's tau = 0.25 and the parameter of the copula C^* corresponds to kendall's tau as given in the table. #obs. = 500, #rep. = 50000

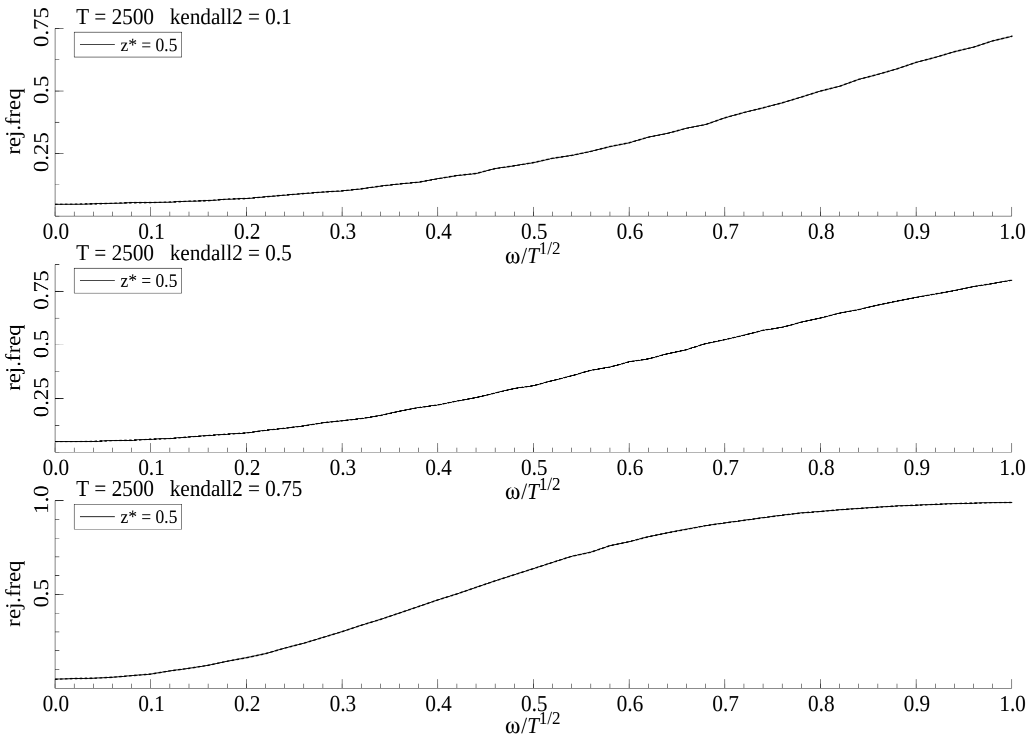

Figure A.1: Local asymptotic power curve. The DGP is given by (8) with $\theta_1 = \theta_2 = \theta_4 = \theta_5 = 0$ and $\theta_3 = 1$. The parameter of the copula C corresponds to kendall's tau = 0.25 and the parameter of the copula C^* corresponds to kendall's tau as given in the panel. #obs. = 500, #rep. = 50000.

Table A.4: Maximum Likelihood Estimates and Goodness-of-Fit statistics of a GARCH(1,1) model with skewed student-t innovations.

	US	UK	France	Germany	Japan
$const(mean)$	0.566^{***}	0.665^{***}	0.599^{**}	0.386^{*}	0.643^{***}
θ_3	1.027^{**}	1.502	3.756^{***}	2.333^{**}	0.517
θ_4	0.124^{***}	0.143^{*}	0.159^{***}	0.156^{***}	0.090^{**}
θ_5	0.831^{***}	0.815^{***}	0.745^{***}	0.783^{***}	0.901^{***}
$log(assym)$	-0.220^{***}	-0.232^{***}	-0.261^{***}	-0.199^{***}	-
$Tail$	7.512^{***}	5.474^{***}	10.101^{**}	6.918^{***}	5.778^{***}
AIC	5.731	6.033	6.319	6.181	6.117
Q(1)	0.735	0.238	7.955	4.245	8.572
	(0.391)	(0.626)	(0.005)	(0.039)	(0.003)
Q(2)	0.813	1.439	8.092	4.865	10.364
	(0.666)	(0.487)	(0.017)	(0.088)	(0.006)
Q(3)	1.321	1.448	9.381	5.452	12.491
	(0.724)	(0.694)	(0.025)	(0.142)	(0.006)
Q(6)	8.059	6.930	12.284	8.149	13.147
	(0.234)	(0.327)	(0.056)	(0.227)	(0.041)
Q(12)	10.357	9.754	19.428	14.140	18.726
	(0.585)	(0.637)	(0.079)	(0.292)	(0.095)
$Q^2(1)$	0.114	0.345	0.150	0.191	0.165
	(0.735)	(0.557)	(0.698)	(0.662)	(0.685)
$Q^2(2)$	0.561	0.405	0.775	0.586	0.497
	(0.755)	(0.817)	(0.679)	(0.746)	(0.780)

The table shows the parameter estimates of the GARCH(1,1) model with skewed student distributed innovations. For Japan we report the GARCH(1,1) with student distributed innovations. Significance levels denoted by: $1\%(^{***}), 5\%(^{**}), 10\%(^{*})$.
The statistics below the parameters are the Akaike Information Criteria (AIC) and the Ljung-Box statistics for serial correlation with p-values indicated in brackets (H_0 : no serial correlation). Q and Q^2 refer to the Ljung-Box statistic based on the standardized innovations and the squared standardized innovations, respectively.

Table A.5: quantile constancy test based on quantics. Significance levels denoted by: $1\%(^{***}), 5\%(^{**}), 10\%(^{*})$

quantile	US	UK	France	Germany	Japan
0.10	0.112	0.084	0.076	0.162	0.583**
0.25	0.128	0.150	0.081	0.115	0.831***
0.50	0.436*	0.135	0.287	0.190	0.355*
0.75	0.126	0.080	0.362*	0.340	0.394*
0.90	0.107	0.655**	0.172	0.134	0.131

Table A.6: Copula Constancy Test Statistics. Significance levels denoted by: $1\%(^{***}), 5\%(^{**}), 10\%(^{*})$.

			no HAC			γ_{2T}			γ_{2T} (filtered data)		
		τ	squares	max	range	squares	max	range	squares	max	range
		0.10	0.180	0.932	1.462	0.092	0.665	1.042	0.322	1.101	1.252
		0.25	0.086	0.695	1.256	0.057	0.562	1.017	0.126	0.952	1.455
US	UK	0.50	0.305	1.302*	1.923**	0.217	1.098	1.622*	0.167	1.001	1.518
		0.75	0.665**	1.715***	2.052***	0.378*	1.292*	1.546	0.089	0.873	1.227
		0.90	0.345	1.105	1.770**	0.226	0.894	1.432	0.286	1.083	1.428
		0.10	0.779***	1.630***	1.729*	0.347	1.088	1.154	0.382*	1.142	1.284
		0.25	0.156	0.842	1.119	0.105	0.692	0.920	0.204	0.844	1.010
US	Fr	0.50	0.220	1.203	1.681*	0.152	1.001	1.398	0.149	1.002	1.412
		0.75	0.661**	1.743***	2.076***	0.366*	1.297*	1.545	0.296	1.393**	1.567
		0.90	0.310	1.247*	2.159***	0.180	0.951	1.645*	0.209	0.954	1.654*
		0.10	1.394***	2.033***	2.180***	0.513**	1.233*	1.322	0.408*	1.260*	1.329
		0.25	0.477**	1.294*	1.694*	0.342	1.096	1.435	0.242	1.029	1.379
US	Ge	0.50	0.250	1.275*	1.823**	0.179	1.081	1.546	0.174	1.101	1.580
		0.75	0.179	1.129	1.734*	0.119	0.923	1.418	0.132	0.990	1.444
		0.90	0.302	1.316*	2.347***	0.147	0.918	1.637*	0.160	0.879	1.587
		0.10	0.403*	1.102	1.292	0.241	0.853	1.000	0.165	0.860	1.124
		0.25	0.227	0.805	1.545	0.172	0.700	1.343	0.252	1.012	1.257
US	Ja	0.50	0.203	1.103	1.493	0.129	0.878	1.188	0.128	0.879	1.189
		0.75	0.164	0.905	1.147	0.111	0.745	0.944	0.122	0.740	1.064
		0.90	0.071	0.693	1.122	0.066	0.667	1.080	0.090	0.703	0.992
		0.10	0.209	0.932	1.383	0.102	0.652	0.969	0.197	0.796	1.070
		0.25	0.088	0.824	1.158	0.062	0.693	0.974	0.070	0.726	1.052
UK	Fr	0.50	0.749***	1.561**	1.968**	0.501**	1.276*	1.610	0.426*	1.158	1.582
		0.75	1.944***	2.236***	2.426***	1.047***	1.641***	1.780**	0.521**	1.463**	1.650*

		0.90	0.915***	1.674***	2.646***	0.504**	1.243*	1.965**	0.384*	1.306*	1.847**
		0.10	0.649**	1.411**	1.504	0.329	1.005	1.071	0.391*	1.034	1.117
		0.25	0.461*	1.266*	1.570	0.303	1.026	1.273	0.199	1.022	1.312
UK	Ge	0.50	0.524**	1.357*	2.006***	0.347*	1.105	1.633*	0.305	1.039	1.572
		0.75	0.266	1.144	1.499	0.255	1.118	1.465	0.148	1.006	1.690*
		0.90	0.415*	1.261*	1.924**	0.294	1.060	1.618	0.382*	1.246*	1.640*
		0.10	0.328	0.985	1.176	0.218	0.804	0.959	0.256	0.921	1.074
		0.25	0.293	0.979	1.668*	0.164	0.732	1.247	0.325	1.119	1.393
UK	Ja	0.50	0.706**	1.673***	1.844**	0.489**	1.392**	1.535	0.407*	1.302*	1.444
		0.75	0.608**	1.439**	1.597	0.470**	1.266*	1.405	0.214	1.138	1.396
		0.90	0.478**	1.439**	1.755**	0.396*	1.310*	1.597	0.269	1.131	1.396
		0.10	0.969***	1.613**	1.721*	0.507**	1.166	1.244	0.554**	1.226*	1.343
		0.25	0.445*	1.174	1.267	0.323	0.999	1.078	0.266	1.036	1.199
Fr	Ge	0.50	0.376*	1.206	2.120***	0.256	0.995	1.749**	0.241	0.959	1.667*
		0.75	0.459*	1.356*	1.827**	0.313	1.120	1.509	0.254	1.103	1.599
		0.90	0.297	1.186	1.990**	0.172	0.902	1.513	0.271	1.035	1.525
		0.10	0.595**	1.453**	1.798**	0.325	1.074	1.329	0.323	1.062	1.314
		0.25	0.283	1.129	1.334	0.225	1.006	1.189	0.308	1.220	1.377
Fr	Ja	0.50	0.232	1.007	1.311	0.163	0.844	1.099	0.163	0.844	1.099
		0.75	0.500**	1.376**	1.655*	0.338	1.131	1.360	0.478**	1.190	1.417
		0.90	0.164	1.002	1.319	0.162	0.996	1.311	0.087	0.811	1.197
		0.10	1.127***	1.903***	1.996**	0.587**	1.374**	1.441	0.510**	1.344*	1.451
		0.25	1.182***	1.924***	2.288***	0.700**	1.481**	1.761**	0.449*	1.335*	1.590
Ge	Ja	0.50	0.333	1.172	1.568	0.247	1.010	1.352	0.202	0.941	1.289
		0.75	0.339	1.578**	2.315***	0.162	1.092	1.602	0.139	0.773	1.471
		0.90	0.126	0.996	1.602	0.094	0.860	1.383	0.067	0.675	1.287

References

Andrews, D. W. K., 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica 59 (3), 817–58.

Berg, D., Quessy, J.-F., 2009. Local power analyses of goodness-of-fit tests for copulas. Scandinavian Journal of Statistics 36 (3), 389–412.

Busetti, F., Harvey, A., 2008. When is a copula constant? A test for changing relationships. Cambridge Working Papers in Economics 0841, Faculty of Economics, University of Cambridge.

Busettti, F., Harvey, A., 2007. Tests of time-invariance. Cambridge Working Papers in Economics 0701, Faculty of Economics, University of Cambridge.

Davidson, J., 1994. Stochastic Limit Theory. Oxford University Press.

De Jong, R. M., Amsler, C., Schmidt, P., 2007. A robust version of the KPSS test based on indicators. Journal of Econometrics 137 (2), 311–333.

De Jong, R. M., Davidson, J., 2000. Consistency of kernel estimators of heteroscedastic and autocorrelated covariance matrices. Econometrica 68 (2), 407–424.

Dehling, H., Fried, R., Van Kampen, M., Vogel, D., Wied, D., 2010. A fluctuation test for constant correlation with spearman's rho. mimeo.

Doornik, J. A., 2005. Object-Oriented Matrix Programming Using Ox, 4th Edition. London: Timberlake Consultants Press.

Gouriéroux, C., 1997. ARCH models and financial applications. Springer-Verlag New York.

Jondeau, E., Rockinger, M., 2006. The copula-garch model of conditional dependencies: An international stock market application. Journal of International Money and Finance 25 (5), 827–853.

Koenker, R., 2005. Quantile Regression. Cambridge University Press.

Krämer, W., Van Kampen, M., 2009. A simple nonparametric test for structural change in joint tail probabilities. Tech. Rep. 04/09, Fakultät Statistik, Universität Dortmund.

Kwiatkowski, D., Phillips, P., Schmidt, P., Shin, Y., 1992. Testing the null hypothesis of stationarity against the alternative of a unit root. Journal of econometrics 54, 159–178.

Laurent, S., Peters, J. P., 2006. G@RCH 4.2, Estimating and Forecasting ARCH Models. London: Timberlake Consultants Press.

Lindner, A. M., 2009. Stationarity, mixing, distributional properties and moments of GARCH(p,q) processes. In: Andersen, T.G., Davis, R.A., Kreiß, J.P. and Mikosch, T. (Eds), Handbook of Financial Time Series, pp.481-496. Springer, Berlin.

Longin, F., Solnik, B., April 2001. Extreme correlation of international equity markets. Journal of Finance 56 (2), 649–676.

Nelson, R. B., 2006. An Introduction to Copulas, 2nd Edition. Springer Series in Statistics. Springer.

Patton, A. J., 2006. Modelling asymmetric exchange rate dependence. International Economic Review 47 (2), 527–556.

Sun, S., Lahiri, S. N., 2006. Bootstrapping the sample quantile of a weakly dependent sequence. Indian Journal of Statistics 68 (1), 130–166.

Withers, C. S., 1981. Conditions for linear processes to be strong-mixing. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 57, 477–480.