COPULA CONSTANCY TESTS

Maarten van Kampen

March 2012

A DISSERTATION
SUBMITTED TO THE FACULTY OF
BUSINESS, ECONOMICS AND SOCIAL SCIENCES
OF THE TECHNISCHE UNIVERSITÄT DORTMUND

(Prof. Dr. Walter Krämer)

(JProf. Dr. Gregor N. F. Weiß)

Contents

List of Tables

List of Figures

Acknowledgement

First, I would like to thank Walter Krämer for being my supervisor. His enthusiasm for econometrics is amazing and I am very thankful for his support. I also thank Gregor Weiß for being my supervisor and for the interesting discussions about copulas. The questions asked by Wolfgang Leininger greatly improved my dissertation.

In addition, I thank my colleagues at the TU Dortmund. In particular, I would like to thank Dominik Wied for being my co-author on several projects and my colleagues, Alesia Khudnitskaya, Editha Lockow, Matthias Arnold and Sebastian Voß, for the great time in *Raum 119*.

The financial support of the Ruhr Graduate School has been gratefully acknowledged. I thank my cohort members, Claudia Burgard, Frauke Dobnik, Regina Flake, Jonas Keil, Klemens Keldenich and Marcus Klemm for the collaboration during my first year in Essen, Vivien Procher for coordinating everything and Barbara Schilde for arranging all flights to the conferences that I was able to visit.

I would like to thank Siem Jan Koopman and Marius Ooms, from the Vrije Universiteit Amsterdam, for introducing me to the field of econometrics, and their support when I told them to continue my academic career.

Many thanks to my family. Pa, ma, Erik: *Dit was zeker niet gelukt zonder jullie steun. Bedankt voor alles.*

Thanks also to Chrisantha Blei and Christoph Kalka for the support in Essen and for improving my German. And last but definitely not least, I would like to thank Eddy Chan, Jie Chen, Herman Baars, Jacqueline Kerkvliet, Michiel van Dijk, Peter Boekel, René Schuit, Ilse Schuit-Hansen, Stefan de Boer and Marijn de Boer. I always enjoy the warm welcome when I come back to the Netherlands.

Chapter 1

Introduction

In financial econometrics we are often confronted with the measurement of the dependence between random variables and the construction of joint distribution functions. Consider for instance a risk manager who has to calculate the Value at Risk (VaR) of a portfolio consisting of several assets. Under the assumption that the asset returns are jointly normally distributed and given the linear correlation between the assets, the calculation of the portfolio VaR is straightforward (see e.g. Rosenberg and Schuermann (2004)). It is however well known that asset returns are not described very well by the normal distribution. In particular, they exhibit volatility clustering and the tails are fatter compared to the normal distribution (see e.g. Cont (2001)). Given some more elaborate specification for the returns, the question arises how to construct the joint distribution given the marginal distributions. Moreover, are marginal distributions and linear correlations sufficient to describe the joint distribution?

As another example, consider two countries and suppose that a market in one country is subject to a shock. To fix ideas, consider Thailand and Indonesia during the East Asian Crisis of 1997. On June 1 1997 the Thai market dropped and two months later the Thai and Indonesian market declined simultaneously (Forbes and Rigobon (2002, p.2242)). The question arises if there are so-called contagion effects. That is, does the inter-dependence between the two countries significantly increase after the shock? Moreover, if there does not exists a "correlation breakdown", does this imply that there are no contagion effects (see Forbes and Rigobon (2002), Rodriguez

(2007))?

Copulas provide a natural way to construct joint distribution functions and to measure dependence between random variable. Sklar's theorem states that the joint distribution, F_{12}, of two random variables X_1 and X_2 can be written in terms of the marginal distributions, $F_i(x) := P(X_i \le x)$, i= 1,2, and a copula function C such that

$$F_{12}(x_1, x_2) = C(F_1(x_1), F_2(x_2)). \tag{1.1}$$

If F_1 and F_2 are continuous, then C is unique (Nelson (2006)). The copula C basically *couples* the marginals together and, since F_i describes the marginal behaviour, we can interpret C as the dependence function.

As a corollary of the previous theorem, we have for the joint density f_{12} that

$$f_{12}(x_1, x_2) = c(F_1(x_1), F_2(x_2)) \cdot f_1(x_1) \cdot f_2(x_2),$$

where c is the density of the copula and f_i are the marginal densities, $i = 1, 2$. To illustrate this, suppose that the portfolio in the first example above consists of two standard normally distributed assets. Then, by the previous theorem, there exist a so-called Gaussian copula such that the portfolio is bivariate normally distributed. Figure 1.1 shows respectively the density of the Gaussian copula, the bivariate Gaussian density and the corresponding contour plot.

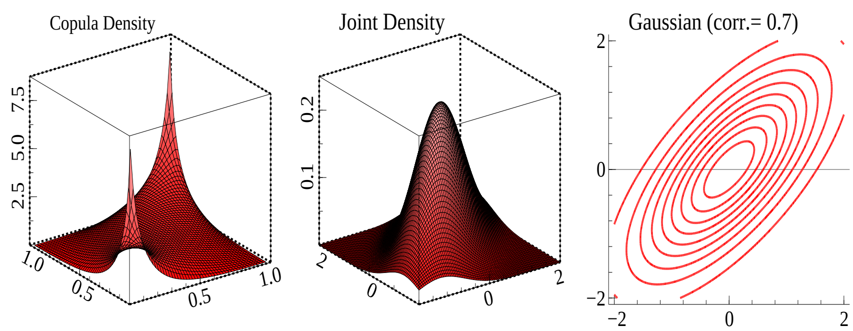

Figure 1.1: Construction of bivariate Gaussian density using copulas

Instead of using a Gaussian copula we can also pick another function C (which does have to satisfy certain mathematical properties, see e.g. Nelson (2006)). It is possible to choose this function such that the linear correlation remains the same, let say 0.7 (like Patton (2006)). Figure 1.2 shows the contour plots of the resulting joint density functions. The dissimilarities between the density functions clearly indicate that knowledge of the marginal distributions and the linear correlations between the assets is insufficient to uniquely describe the joint distribution function.

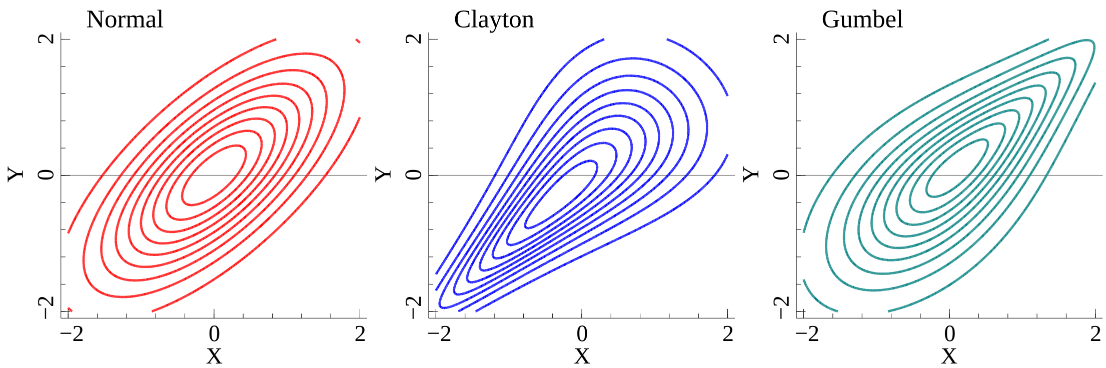

Figure 1.2: Joint density using Gaussian, Clayton and Gumbel copula. Marginals are standard normal and copula parameter is chosen such that the linear correlation equals 0.7.

Let $\xi_i(\tau_i)$ denote the τ_i quantile of X_i. Equation (1.1) can be rewritten as

$$C(\tau_1, \tau_2) = F_{12}(\xi_1(\tau_1), \xi_2(\tau_2)) = P(X_1 \leq \xi_1(\tau_1), X_2 \leq \xi_2(\tau_2)).$$

Hence, the copula C gives the probability that both random variables takes values below their marginal τ_i-quantiles (see Harvey (2010)).

As pointed out above, the copula is the dependence function between the random variables. It is in fact a joint distribution function with uniform marginals. This function is invariant under strictly increasing transformations. Scale-invariant measures of dependence, such as Kendall's tau and Spearman's rho can be expressed as function of the copula. Another dependence concept frequently used in financial econometrics is tail dependence. Contrary to the Gaussian copula, the Clayton and

Gumbel copulas can be tail dependent. In financial econometrics there exists considerable evidence that asset returns are lower tail dependent (see e.g. Longin and Solnik (2001)). Hence, the Clayton copula is often used in the literature.

In the second example above, we are interested in a change in the dependence structure. From the previous discussion it should be clear that solely examining linear correlations is insufficient; we would like to know if there is a change in the copula. The importance of changes in copulas is clear; changes in the dependence structure affect the VaR of a portfolio. Hence, investors would like to re-allocate their assets and risk managers have to adjust their capital buffer to cover unexpected losses. The extent to which a change is also economically significant has been analyzed by e.g. Patton (2004).

Changes in copulas have mainly been analyzed in a (semi-)parametric framework. Such an approach requires the functional form of the copula and often a specification for the transition of the copula over time. Jondeau and Rockinger (2006) consider changes in the dependence between four different stock indices. They propose a semi-parametric approach in which they partition the unit square in different quadrants and let the copula parameter depend on the location of the past realization in the unit square. The null hypothesis of a constant copula corresponds to the case that each part of the unit square has the same copula parameter. They compare their approach with two time-varying parameter specifications in which they explicitly describe the transition of the copula parameters over time. Alternative time-varying parameter specifications can be found in Patton (2006) and Creal et al (2008). Instead of modeling changes in the copula parameter, Rodriguez (2007) considers changes in the functional form of the copula. Here, the copula is a weighted average of three copulas where the weights depend on the state of the economy.

Dias and Embrechts (2004) propose a generalized likelihood ratio test which requires the estimation of the copula. The critical values of the test depend on the chosen copula but can be approximated using independent Brownian Bridges. In empirical applications the true copula is generally unknown and hence the application of the test requires some goodness-of-fit tests to validate the chosen copula.

Giacomini et al (2009) propose to model the copula parameter using a local change point procedure. Their method partitions the time interval such that the copula parameter is constant on each interval. To examine if a particular interval contains a change point they apply a likelihood ratio test. Hence, the procedure depends on the chosen copula function.

To examine possible changes in the copula, Harvey (2010) proposes a binary filtering approach. Unlike the majority of the papers cited above, the purpose is here is not to provide a model for the copula. The filter returns the predicted value of the copula given the past observations. Plotting the filtered series against time might indicate possible changes. This method does not depend on a chosen copula function but the drawback is of course that this is not a formal test.

In this thesis I introduce a nonparametric test that examines if a copula is constant over time. In the second chapter I show that under the assumption of independent and identically distributed (i.i.d.) variables the test outperforms a copula constancy test recently proposed in the literature provided that there are multiple breaks in the sample[1].

In time series analysis, the i.i.d. assumption is often violated and hence the question arises under what kind of dependence assumption we can derive the asymptotic distribution. In chapter 3[2], I characterize dependence using a strong mixing assumption. Such an assumption states, loosely speaking, that if the time separation between two events in the series increases (to infinity), than the events behave independent. In other words, the events are asymptotically independent. I show in the second chapter that under a suitable strong mixing assumption I can still derive the limiting distribution.

An important difference between the test described in chapter 2 and 3 concerns the estimation of the long run variance. If there exist serial correlation between the

[1]Chapter 2 is based on Krämer and Van Kampen (2011), A simple nonparametric test for structural change in joint tail probabilities, Economics Letters 110(3), pp.245-247. Krämer set up the manuscript. Van Kampen did most of the programming work and provided clear contributions to the theoretical part.

[2]Chapter 3 is based on Van Kampen and Wied (2010), A non-parametric constancy test for copulas under weak dependence. Tech. Rep. 36/10, Fakultät Statistik, Universität Dortmund. The idea has been developed together. Van Kampen set up the theoretical results and did most of the programming work. Wied provided numerous improvements and additional theoretical results. The paper has been submitted for publication.

observations we should replace the long run variance by a heteroskedasticity and autocorrelation consistent (HAC) estimate. The standard approach in the literature is to construct such an estimate using a kernel function that depends on a particular bandwidth parameter. To obtain a consistent estimate, the bandwidth should increase with the sample size at a suitable rate. In chapter 3, I make an assumption on the kernel and bandwidth such that this is indeed the case

In chapters 2 and 3 I propose to compare the test with the critical value obtained from the asymptotic distribution. I show that such a strategy results in size distortions in small samples. To improve the finite sample performance of the test, I replace the long run variance by an inconsistent estimate. Such an estimate introduces additional variability which might improve the finite sample properties. I show in chapter 4[3] that the resulting asymptotic distribution depends on the kernel and bandwidth. The question arises how to select the bandwidth parameter such that the power of the test is high and the size distortion of the test is low. In the spirit of Sun, Phillips and Jin (2008), I develop a bandwidth rule which minimizes a weighted average of the type I and II errors.

The previous test only examines the constancy of the copula in a particular point. The properties of the test (e.g. power) depend on the chosen point. Since it is unclear at which point we should apply the test, current practice is to apply the test to several points. In Chapter 5[4] I introduce a new test for examining the constancy of the complete copula function. This test does not require the *a priori* selection of a particular point. The asymptotic distribution of test depends on the copula. Hence, the critical values are simulated using a bootstrap algorithm.

[3]Chapter 4 is written without coauthors. I would like to thank, however, Walter Krämer and Dominik Wied for providing suggestions that clearly improved the paper.

[4]Chapter 5 has been added to Van Kampen and Wied (2010). The idea has been developed together. Wied set up the theoretical results. Van Kampen provided clear contributions to this and did most of the programming work. The paper has been submitted for publication.

Chapter 2

A simple nonparametric test for structural change in joint tail probabilities[1]

We propose a new test against a change in the probability of multivariate tail events. The test is based on partial sums of a suitably defined indicator function and detects multiple changes in joint tail probabilities better than a previously suggested competitor.

2.1 Introduction

In 2008, all major stock markets in the world fell by roughly 30% to 40%. The year before there were likewise some extreme events but there was no global downturn. The question arises whether such a downturn can be explained by chance or whether there was a structural change in joint tail probabilities sometime in between.

Campbell et al. (2002, 2008) and Forbes and Rigobon (2002) investigated possible changes in the dependence structure between stock returns. It is important to distinguish this from the concept of asymmetric dependence as examined by Ang and Chen (2002), Fortin and Kuzmicz (2002), De Melo Mendez (2005) and Sun et al. (2008).

[1]This chapter is based on Krämer and Van Kampen (2011), Economics Letters 110(3), pp.245-247.

They showed that joint stock returns exhibit larger dependence in the lower than in the upper tail. In this chapter we address the former issue.

In this chapter we propose a new test against a change in the probability of multivariate tail events. Following Busetti and Harvey (2011), we base our test on joint exceedances of certain quantiles of the marginal distributions. Instead of using sums of squares of a normalized indicator function, we propose two alternative test statistics. The first is based on the maximum of cumulative sums of the indicator variables, in the spirit of Ploberger and Krämer (1992). The second uses the range of the cumulative sums. We show via Monte Carlo simulation that no test is uniformly superior to the others. While the sums of squares version is more likely to detect gradual or continuous changes in probabilities, the max test and the range test are more successful with abrupt changes. None of the tests requires prior knowledge as to when a structural change occurs.

2.2 The test and its asymptotic null distribution

Following Busetti and Harvey (2011), we let $\xi(\tau)$ denote the τ-quantile of some univariate probability distribution. To avoid unnecessary notational complications, we consider continuous distributions only, so $\xi(\tau)$ is uniquely defined. For a bivariate series y_{1t} and y_{2t}, $t = 1, \dots, T$, let $\hat{\xi}_1(\tau_1)$ and $\hat{\xi}_2(\tau_2)$ denote the respective empirical quantiles, and let $C_T(\tau_1, \tau_2)$ be the proportion of observation where y_{1t} and y_{2t} are less than or equal to $\hat{\xi}_1(\tau_1)$ and $\hat{\xi}_2(\tau_2)$, respectively. $C_T(\tau_1, \tau_2)$ is an estimator of $P(y_{1t} \leq \hat{\xi}(\tau_1), y_{2t} \leq \hat{\xi}(\tau_2))$, which is assumed constant under our null hypothesis. Note that this probability is given by the true copula and, therefore, our test may be viewed as a procedure to check the constancy of a copula at a given point. For simplicity, we let $\tau_1 = \tau_2 = \tau$ from now on.

The basic input of our test is what Busetti and Harvey (2011) call the *bivariate τ-quantic*

$$BIQ(y_t, \hat{\xi}(\tau)) = C_T(\tau_1, \tau_2) - I(y_t, \hat{\xi}(\tau)), \qquad t = 1, \dots, T, \tag{2.1}$$

where $I(.)$ is the indicator function taking the value 1 if $y_{1t} \leq \hat{\xi}_1(\tau_1) \ \wedge \ y_{2t} \leq \hat{\xi}_2(\tau_2)$, and 0 otherwise. By definition, the $BIQ(y_t, \hat{\xi}(\tau))$ add to zero, and their partial sums should not deviate too much from zero if $P(y_{1t} \leq \hat{\xi}_1(\tau), y_{2t} \leq \hat{\xi}_2(\tau))$ remains constant across the sample. However, if the probability increases at $t = t_1$, then $BIQ(y_t, \hat{\xi}(\tau))$ will tend to be positive up to t_1 and negative from t_1 onwards. Therefore, the cumulated sum of the $BIQ(y_t, \hat{\xi}(\tau))$ will move away from zero farther than can be expected under the null hypothesis. This is illustrated in the second panel of figure 2.1. Alternatively, in case the probability decreases at $t = t_1$, the cumulative sum decreases up to t_1 and increases afterwards.

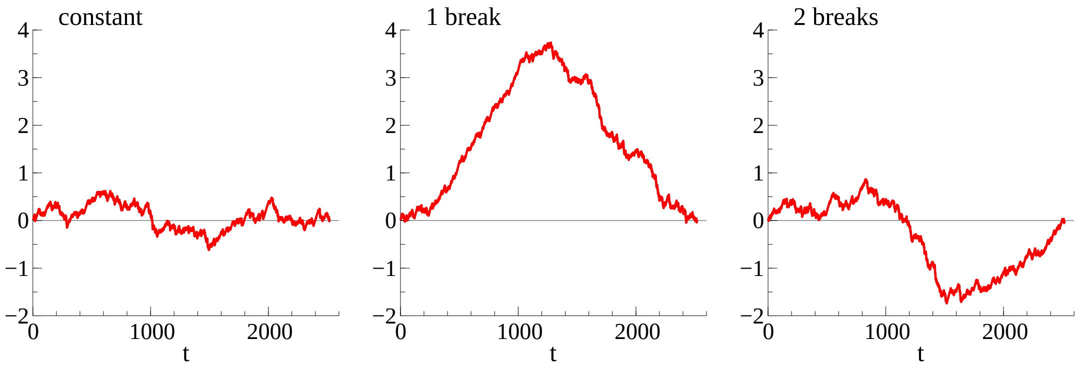

Figure 2.1: Partial sum process.

This motivates our choice of test statistic, which is a suitably normalized version of

$$\max_{t=1,\ldots,T} \left| \sum_{i=1}^{t} BIQ(y_i, \hat{\xi}(\tau)) \right|.$$

We show below that, under the null and whenever the events $(y_{1t} \leq \hat{\xi}(\tau), y_{2t} \leq \hat{\xi}(\tau))$ and $(y_{1s} \leq \hat{\xi}(\tau), y_{2s} \leq \hat{\xi}(\tau))$ are independent for all $t \neq s$, the stochastic process

$$\frac{1}{\sqrt{TC_T(\tau,\tau)(1 - C_T(\tau,\tau))}} \left[\sum_{i=1}^{[rT]} BIQ(y_i, \hat{\xi}(\tau)) \right] \qquad 0 \leq r \leq 1 \tag{2.2}$$

tends in distribution to a Brownian Bridge as $T \to \infty$, so the limiting null distribution of

$$\frac{1}{\sqrt{TC_T(\tau,\tau)(1 - C_T(\tau,\tau))}} \max_{t=1,\ldots,T} \left| \sum_{i=1}^{t} BIQ(y_i, \hat{\xi}(\tau)) \right|$$

is identical to that of the Kolmogorov-Smirnov test (see Ploberger and Krämer (1992)). Some useful critical values are 1.22 (α=10%), 1.36 (α=5%) and 1.63 (α=1%), where α denotes the significance level.

The right panel of figure 2.1 illustrates the case when the probability $P(y_{1t} \le \hat{\xi}(\tau), y_{2t} \le \hat{\xi}(\tau))$ increases at $t = t_1$ and subsequently decreases to its original level at $t = t_2$. This motivates an alternative test statistic based on the range of the cumulative sums, as in Krämer and Schotman (1992). In this case, the test statistic is

$$\frac{1}{\sqrt{TC_T(\tau,\tau)(1 - C_T(\tau,\tau))}} \left[\max_{t=1,\dots,T} \sum_{i=1}^{t} BIQ(y_i, \hat{\xi}(\tau)) - \min_{t=1,\dots,T} \sum_{i=1}^{t} BIQ(y_i, \hat{\xi}(\tau)) \right],$$

where the asymptotic null distribution is given by

$$P(X \le x) = 1 + 2 \sum_{k=1}^{\infty} \left(1 - 4k^2x^2\right) \exp(-2k^2x^2),$$

see e.g. Kennedy (1976). Some useful critical values are 1.620 ($\alpha = 10\%$), 1.747 ($\alpha = 5\%$) and 2.001 ($\alpha = 1\%$).

Of course, other functionals of the $BIQ(y_i, \hat{\xi}(\tau))$ such as the sum of absolute values might also be used as test statistics, but we focus here on the performance of the *maximum* and the *range* statistic (as compared to the sum of *squares* statistic proposed by Busetti and Harvey (2011)).

Convergence in Distribution

The convergence in distribution to a Brownian Bridge of (2.2) can be seen by first considering

$$Q(y_i, \xi(\tau)) = C(\tau,\tau) - I\left(y_{1t} \le \xi_1(\tau), y_{2t} \le \xi_2(\tau)\right). \tag{2.3}$$

This is an i.i.d. sequence with zero expectation, finite higher moments of all orders and variance $\sigma^2 = C(\tau,\tau)(1-C(\tau,\tau))$, so, by standard results from probability theory (see e.g. Billingsley (1986))

$$\frac{1}{\sqrt{T\sigma^2}} \sum_{i=1}^{[rT]} Q(y_i, \xi(\tau))$$

tends in distribution to a standard Wiener Process and

$$\frac{1}{\sqrt{T\sigma^2}} \sum_{i=1}^{[rT]} BIQ(y_i, \xi(\tau))$$

tends in distribution to a Brownian Bridge. The convergence to a Brownian Bridge of (2.2) then follows from the fact that $C_T(\tau,\tau)(1-C_T(\tau,\tau))$ is consistent for $\sigma^2 = p(1-p)$ and

$$\sup_{r\in[0,1]} \left| \frac{1}{\sqrt{T\sigma^2}} \sum_{i=1}^{[rT]} BIQ(y_i, \hat{\xi}(\tau)) - \frac{1}{\sqrt{T\sigma^2}} \sum_{i=1}^{[rT]} BIQ(y_i, \xi(\tau)) \right| \xrightarrow{p} 0. \tag{2.4}$$

A formal proof of the latter, under weaker conditions, can be found in section 3.

2.3 Finite sample properties

2.3.1 Base case scenario

Following Busetti and Harvey (2011), we examine the performance of the proposed tests using simulated values from the Clayton copula. We explicitly analyze the effect of multiple breaks in the copula parameter. The results in this section are generated using Ox (see Doornik (2005)).

Suppose that there are m breakpoints denoted by $t_1, \ldots, t_m$. Let θ_j denote copula parameter on segment $j = 1, \ldots, m+1$. The bivariate time series y_{1t} and y_{2t}, $t = t_{j-1}+1, \ldots, t_j$ are drawn from a Clayton copula $C(u, v; \theta_j)$ with parameter θ_j. In our base case scenario, we simulate 50000 replications of time series consisting of 2520 observations. Note that this is much higher than the series simulated by Busetti and Harvey (2011) (between 200 and 400 observations). In the simulation we restrict the number of copula parameters such that $\theta_1 \equiv \theta_{2k+1}$ and $\theta_2 \equiv \theta_{2k}$, $k = 0, 1, \ldots$. Intuitively, the series consist of periods of low dependence and periods of high dependence. Finally, we apply the test statistics to the 0.05, 0.1, 0.25 and 0.50 quantile.

Table 2.4 (in section 2.B) shows the rejection frequencies for 2 and 3 intervals. Note that the squares test outperforms the maximum and range tests if there is a single break in the copula parameter. However, the power of our test is higher in the case of two structural breaks, and the test based on the range outperforms the test based on the maximum.

2.3.2 Sensitivity analysis

The following paragraphs provide some robustness checks with respect to the obtained rejection frequencies. In particular, we investigate the sensitivity of the rejection frequencies with respect to the number of breaks, the number of observations and the copula type.

Number of breaks

To analyze the effect of the number of breaks on the rejection frequency (given a fixed sample size of 2520 observations) we perform simulations up to 9 breaks. To examine the effect of the magnitude of the break and the particular quantile, we give the results for a relatively high and low quantile and a relatively high and low break in the copula parameter. Figure 2.2 shows the results for an increase, at the odd-numbered break points, in the Clayton copula parameter.

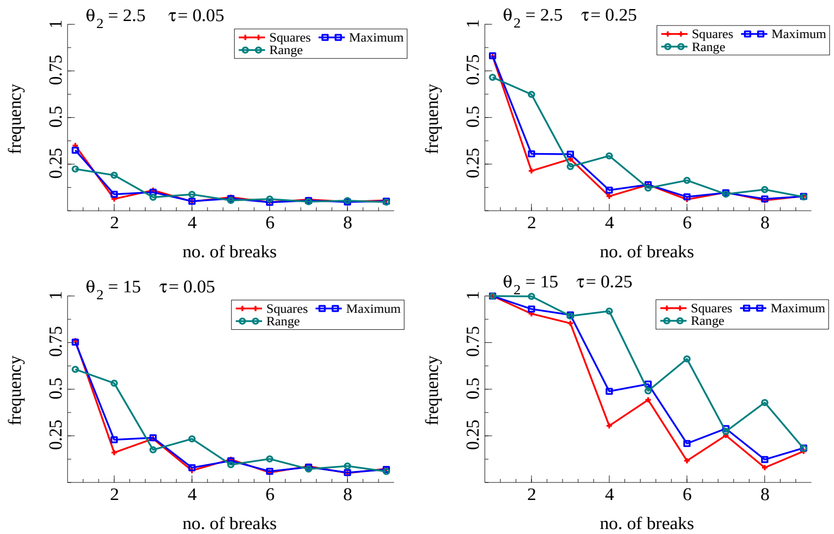

Figure 2.2: Rejection frequency versus number of breaks for a clayton copula with $\theta_1 = 1$. Results obtained using 50.000 replications of 2520 observations.

The following two results should be clear: First, the rejection frequency decreases as the number of breaks increases. Note that if the number of breaks increases the deviations of the cumulative sums from zero are more likely to be smaller (see Figure 2.1). Therefore, given a fixed sample size, it becomes more difficult for the test to reject the null hypothesis and the power of the test will be lower. Second, the range test performs better if there are an even number of breaks. This result is mainly due to the setup of the simulation. Given an odd number of breaks the cumulative sums remain almost everywhere positive (negative) given an increase (decrease) in the probability at odd numbered break points and a decrease (increase) at the even numbered break points. This is due to the fact that, in our simulation setup, all intervals have equal length and the magnitude in the increase of the copula parameter is equal to the magnitude of the decrease in the copula parameter.

Number of observations

The previous result depends on the assumed sample size of 2520 observations. Figure 2.3 shows the sensitivity of the rejection frequency if we adjust the sample size while holding the number of breaks fixed. That is, in case of 1 break in three samples consisting of 250, 500 and 750 observations, the breaks points are at $t = 125$, $t = 250$ and $t = 375$, respectively.

The upper panels show the results for the case of 1 break. We focus on the cases that the parameter and quantile are both relatively low and the case that they are both relatively high. Combining different quantiles and parameters (as in Figure 2.2) gives similar results. From the upper panels of figure 2.3 we conclude that the square test outperforms the other two tests. The lower panels show that the range tests outperforms the other two tests if there are two breaks. These results are robust with respect to the number of observations.

Copula type

The previous analysis took the Clayton copula as the benchmark copula. The question that arises is to what extent the results depend on the copula type. Therefore, we also examine the performance for the Gaussian bivariate distribution. The parameter values are taken from Busetti and Harvey (2011). Table 2.5 (in section 2.B) shows

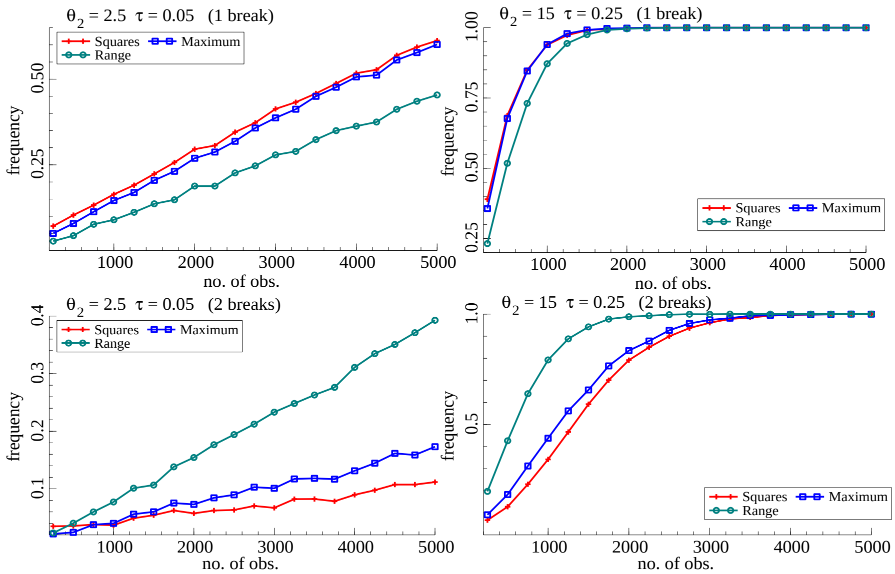

Figure 2.3: Rejection frequency versus number of observations for a Clayton copula with $\theta_1 = 1$. Results obtained using 5.000 replications.

that the range test again outperforms the square test if there are two breaks in the sample.

To improve the comparability between the results of the Clayton and Gaussian copula we might set the magnitude of break such that it is the same for both copulas. Table 2.1 shows the relationship between the probability of the 0.25-quantile and the parameters of the Gaussian and Clayton copula.

Table 2.1: Relationship between the joint probability and the copula parameters

C(0.25,0.25)	Gaussian	Clayton
0.10	0.34112	0.36484
0.12	0.49797	0.62063
0.14	0.63595	0.94584
0.16	0.75363	1.38176
0.18	0.84964	2.01544

If for a given quantile τ, the initial probability and the change in the probability of a particular quadrant are the same, the resulting rejection frequency is also the same. For example, the power of test is similar for a change in the Gaussian copula from 0.34112 to 0.75363 as a change in the Clayton copula parameter from 0.36484 to 1.38176, *if we evaluate the test at the 0.25-quantile.* If we evaluate the test at $\tau \neq 0.25$ this is not true. We abstain from a detailed analysis.

2.4 Empirical application

We illustrate our tests for the Kuala Lumpur stock exchange in Malaysia and the Hang-Seng index in Hong-Kong. The data have been obtained from EconStats and consist of daily observations from December, 8, 1993 through May, 19, 2009. The corresponding return series is calculated as $y_t = 100 \times log(x_t/x_{t-1})$, where x_t denotes the index at time $t = 1, \ldots, T$. In the analysis below we keep all dates at which both return series are observed. This reduces the sample of Malaysia and Hong-Kong from respectively, 3809 and 3831 observations, to 3679 observations.

For each series we estimate an AR(1)-GARCH(1,1) model with Student-t innovations:

$$\begin{aligned} y_t &= \mu + \phi y_{t-1} + \varepsilon_t \\ \varepsilon_t &= z_t \sigma_t \\ \sigma_t^2 &= \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2, \end{aligned}$$

where z_t is Student-t with υ degrees of freedom (to be estimated from the data). Table 2.3 shows the parameter estimates. Note that $\alpha + \beta$ is close to one for Hong-Kong and slightly exceeds one for Malaysia. To examine the sensitivity of our results we also estimated a GARCH model with Gaussian disturbances and an IGARCH model in which we explicitly restricted $\alpha + \beta = 1$. Based on the AIC criteria, we see that the model with Gaussian disturbances performs less good but the resulting IGARCH model behaves similar as the original model. Choudhry (1995) found similar parameter values for some European stock markets. As also pointed out by him, $\alpha + \beta = 1$ basically implies that shocks persists indefinitely on the conditional variance.

Table 2.2: Maximum likelihood estimates for Malaysia and Hong-Kong

Hong-Kong	GARCH(1,1)-T	IGARCH(1,1)-T	GARCH(1,1)-N	IGARCH(1,1)-N
μ	0.059***	0.059***	0.057***	0.057***
ϕ	0.035**	0.035**	0.051***	0.051***
ω	0.012***	0.010***	0.018***	0.014***
α	0.063***	0.065***	0.077***	0.080***
β	0.935***	$0.935^{(NA)}$	0.919***	$0.920^{(NA)}$
ν	7.445***	7.230***		
AIC	3.566	3.565	3.594	3.594
Malaysia	GARCH(1,1)-T	IGARCH(1,1)-T	GARCH(1,1)-N	IGARCH(1,1)-N
μ	0.021	0.021	0.037**	0.038**
ϕ	0.145***	0.145***	0.173***	0.173***
ω	0.017***	0.017***	0.015***	0.017***
α	0.143***	0.140***	0.144***	0.139***
β	0.860***	0.860***	0.861***	0.861***
ν	5.317***	5.417***		
AIC	2.904	2.904	2.977	2.976

Significance levels: 1% (***), 5% (**) and 10% (*); NA = not available.

Subsequently, the different tests are applied to the standardized empirical innovation series. Since the standardized residuals still contain serial correlation, we replaced the variance of the BIQ series by a long-run estimator with 9 lags. The number of lags is based on the bandwidth rule $b = 4(T/100)^{1/4}$. Table 2.3 shows that only the maximum and the range test are able to reject the null hypothesis at the 5% level for some quantiles.

2.A Copulas: functional forms and simulation methods

We follow the simulation methods proposed in Cherubini et al. (2004, p181). This appendix is solely included to facilitate the replication of the simulation study in section 2.3.

Table 2.3: Test statistics based on standardized innovations of an AR(1)-(I)GARCH(1,1)-T model and a long run variance estimate based on 9 lags. Significance is denoted by the superscripts 1% (^{a}), 5% (^{b}) and 10% (^{c}).

	resid. GARCH(1,1)-T			resid. IGARCH(1,1)-T		
τ	squares	cusum	range	squares	cusum	range
0.1	0.299	1.370**	2.027***	0.290	1.370**	2.027***
0.25	0.207	1.231*	1.638*	0.188	1.182	1.633*
0.5	0.416*	1.378**	2.069***	0.416*	1.378**	2.069***
0.75	0.072	0.627	1.168	0.063	0.598	1.117
0.9	0.135	0.863	1.190	0.154	0.904	1.230

Gaussian copula

The Gaussian copula is given by

$$C(u_1, u_2) = \Phi_{XY}(\Phi^{-1}(u_1), \Phi^{-1}(u_2); \rho),$$

where Φ_{XY} is the bivariate normal distribution with linear correlation parameter ρ and Φ is the standard (univariate) distribution function. We can simulate a pair (u_1, u_2) of observations as follow: First, construct a pair $(v_1, v_2) = Rz$ where R is the lower triangular matrix such that RR' equals the correlation matrix and $z = (z_1, z_2)'$ with z_i simulated from the standard normal distribution. Second, we have $(u_1, u_2) = (\Phi(v_1), \Phi(v_2))$.

Clayton copula

The Clayton copula is given by

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

A pair (u_1, u_2) of observations can be obtained using the conditional sampling method. The idea is to simulate a pair (u_1, v_2) from the uniform distribution on $[0, 1]$ and set $u_2 = C_{u_1}^{-1}(v_2)$ where $C_{u_1}(v_2) := P(V \leq v_2 | U = u_1)$ and $C_{u_1}^{-1}(\cdot)$ is the inverse function.

For the Clayton copula we have

$$u_2 = \left(u_1^{-\theta}(v_2^{-\theta/(\theta+1)+1})\right)^{-1/\theta}.$$

2.B Tables

Table 2.4: Clayton Copula with structural breaks in dependence: empirical rejection frequencies (T=2520, Rep = 50000, θ_1 = 1)

m	test	τ	θ_2	1	2.5	7.5	15
		0.05		0.05	0.35	0.68	0.76
1	Squares	0.10		0.05	0.57	0.93	0.96
		0.25		0.05	0.83	1.00	1.00
		0.50		0.05	0.81	1.00	1.00
		0.05		0.04	0.33	0.67	0.75
1	Maximum	0.10		0.05	0.56	0.93	0.97
		0.25		0.05	0.83	1.00	1.00
		0.50		0.05	0.81	1.00	1.00
		0.05		0.04	0.22	0.53	0.61
1	Range	0.10		0.04	0.42	0.85	0.92
		0.25		0.04	0.72	1.00	1.00
		0.50		0.04	0.69	1.00	1.00
		0.05		0.05	0.06	0.13	0.16
2	Squares	0.10		0.05	0.10	0.33	0.42
		0.25		0.05	0.21	0.79	0.91
		0.50		0.05	0.20	0.91	0.99
		0.05		0.05	0.09	0.19	0.23
2	Maximum	0.10		0.05	0.16	0.42	0.51
		0.25		0.05	0.30	0.84	0.93
		0.50		0.05	0.29	0.94	0.99
		0.05		0.04	0.19	0.46	0.53
2	Range	0.10		0.04	0.36	0.78	0.86
		0.25		0.04	0.62	0.99	1.00
		0.50		0.04	0.58	1.00	1.00
		0.05		0.05	0.11	0.20	0.24
3	Squares	0.10		0.05	0.17	0.37	0.45
		0.25		0.05	0.29	0.74	0.86
		0.50		0.05	0.27	0.88	0.97
		0.05		0.04	0.10	0.20	0.24
3	Maximum	0.10		0.04	0.17	0.41	0.49
		0.25		0.05	0.31	0.81	0.90
		0.50		0.05	0.29	0.92	0.98
		0.05		0.04	0.08	0.15	0.17
3	Range	0.10		0.04	0.12	0.33	0.41
		0.25		0.04	0.24	0.78	0.89
		0.50		0.05	0.23	0.92	0.99

Table 2.5: Gaussian Copula with structural breaks in correlation: empirical rejection frequencies (T=2520, Rep = 50000, θ_1 = 0.5)

m	test	τ	θ_2	0.10	0.25	0.50	0.75	0.90
		0.05		0.59	0.29	0.05	0.40	0.87
1	Squares	0.10		0.84	0.46	0.05	0.57	0.97
		0.25		0.96	0.63	0.05	0.72	1.00
		0.50		0.92	0.56	0.05	0.69	0.99
		0.05		0.56	0.26	0.04	0.37	0.87
1	Maximum	0.10		0.83	0.44	0.05	0.56	0.97
		0.25		0.96	0.62	0.05	0.72	1.00
		0.50		0.92	0.55	0.05	0.68	1.00
		0.05		0.38	0.16	0.03	0.25	0.76
1	Range	0.10		0.70	0.31	0.04	0.42	0.93
		0.25		0.91	0.48	0.04	0.59	0.99
		0.50		0.84	0.42	0.04	0.55	0.98
		0.05		0.13	0.09	0.05	0.06	0.24
2	Squares	0.10		0.20	0.11	0.05	0.10	0.46
		0.25		0.36	0.13	0.05	0.16	0.68
		0.50		0.30	0.11	0.05	0.15	0.62
		0.05		0.17	0.10	0.04	0.09	0.31
2	Maximum	0.10		0.29	0.14	0.04	0.15	0.54
		0.25		0.48	0.20	0.05	0.23	0.75
		0.50		0.41	0.17	0.05	0.22	0.71
		0.05		0.24	0.11	0.03	0.22	0.69
2	Range	0.10		0.54	0.22	0.04	0.36	0.88
		0.25		0.82	0.38	0.04	0.50	0.97
		0.50		0.75	0.33	0.04	0.45	0.96
		0.05		0.16	0.10	0.05	0.12	0.30
3	Squares	0.10		0.27	0.13	0.05	0.16	0.47
		0.25		0.44	0.19	0.05	0.23	0.65
		0.50		0.37	0.16	0.05	0.21	0.63
		0.05		0.13	0.08	0.04	0.11	0.31
3	Maximum	0.10		0.28	0.12	0.04	0.16	0.51
		0.25		0.49	0.19	0.04	0.24	0.72
		0.50		0.41	0.17	0.05	0.22	0.70
		0.05		0.09	0.05	0.03	0.08	0.23
3	Range	0.10		0.21	0.09	0.04	0.12	0.42
		0.25		0.40	0.14	0.04	0.18	0.67
		0.50		0.33	0.13	0.05	0.17	0.64

Chapter 3

A non-parametric constancy test for copulas under weak dependence[1]

This chapter extends 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 strong mixing conditions.

3.1 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 (2011) and Krämer and Van Kampen (2011) 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

[1]This chapter is based on Van Kampen and Wied (2010)

the specification of a transition mechanism. The test is based 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. Busetti 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. The resulting test is consistent against the alternative of a single structural break. 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.

3.2 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

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

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

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

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

$$S_T(\cdot) \xrightarrow{d} \sigma B(\cdot), \tag{3.1}$$

where $\sigma^2 = C(\tau_1,\tau_2)(1-C(\tau_1,\tau_2))$ and 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 (2011), 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 $\tilde{S}_T(r) := 1/\sqrt{T}\sum_{t=1}^{[rT]} BIQ(y_t,\xi(\tau))$ we have

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

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

$$\hat{S}_T(r) := 1/\sqrt{T}\sum_{t=1}^{[rT]} BIQ(y_t, \hat{\xi}(\tau)). \tag{3.2}$$

Then Busetti and Harvey (2011) 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 $\hat{\sigma}_{iid}^2 := \hat{C}_T(\tau_1,\tau_2)(1-\hat{C}_T(\tau_1,\tau_2))$ 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 (2011) propose complementary tests based on the *maximum* and the *range* of $\hat{S}_T(\cdot)$

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

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

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%).

3.3 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 $(\dots, y_{i,t-1}, y_{it})$ and $(y_{i,t+m}, y_{i,t+m+1}, \dots)$, 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 3.1.

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. The copula constancy tests is, however, subject to size distortions if the series are serially correlated or exhibit stochastic volatility patterns (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 3.1.3 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 3.1, $S_T(\cdot)$ 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 $S_T(\cdot) \xrightarrow{d} \sigma B(\cdot)$, see e.g. Corollary 29.7 of Davidson (1994). In addition

$$\tilde{S}_T(\cdot) \xrightarrow{d} \sigma V(\cdot). \tag{3.3}$$

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

$$\bar{\sigma}^2 = T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} k((t-s)/\gamma_T) \cdot Q_t \cdot Q_s, \tag{3.4}$$

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

Assumption 3.2.

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. $\gamma_T \to \infty$ and $\gamma_T/T \to 0$ as $T \to \infty$.

Remark:

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

(ii) The bandwith rate of Assumption 2.3 is similar to the one in De Jong et al. (2007) and Andrews (1991).

The HAC estimate (3.4) 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

$$\hat{\sigma}^2 = T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T} k((t-s)/\gamma_T)\cdot BIQ(y_t,\hat{\xi}(\tau))\cdot BIQ(y_s,\hat{\xi}(\tau)). \tag{3.5}$$

We make the following assumption on the empirical quantile process.

Assumption 3.3. $\sqrt{T}(\hat{\xi}_i(\tau)-\xi_i(\tau)) = O_p(1)$, for $i = 1,2$.

Remark: Assumption 3.3 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 3.1 establishes the result of the previous section for the case where the observations are strong mixing. The proof is given in 3.A.

Theorem 3.1. *Under Assumptions 3.1, 3.2 and 3.3*

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

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

Busetti and Harvey (2011) 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 conditional 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 3.1.

Since marginal distributions can also change for reasons other than stochastic volatility, we extend Theorem 3.1 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 3.4.

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 $\inf_{(u_1,u_2)\in[-\eta,\eta]^2} f_{12}(u_1, u_2) > 0$.

4. $\gamma_T \to \infty$ and $\gamma_T/\sqrt{T} \to 0$ as $T \to \infty$.

Remark:

(i) Assumption 3.4.1 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 3.4.2 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.

(iii) Assumption 3.4.4 strengthens the rate of the bandwidth parameter. 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.

Theorem 3.2. *Under Assumptions 3.1, 3.2, 3.3 and 3.4 we have*

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

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

An application of the continuous mapping theorem after Theorem 3.1 or 3.2 gives the same tests as established in the i.i.d. case.

3.4 The asymptotic power of the test

3.4.1 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

$$\begin{aligned} H_0: &\quad C^{(t)}(\tau_1, \tau_2) = C(\tau_1, \tau_2) \\ H_1: &\quad C^{(t)}(\tau_1, \tau_2) = (1 - g(t,T))C(\tau_1, \tau_2) + g(t,T)C^*(\tau_1, \tau_2), \end{aligned} \tag{3.6}$$

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 3.3. *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* (3.6)*.*

Remark:

(i) Note that under the fixed alternative (3.6), the variance of the terms in the partial sum process changes over time.

(ii) In the special case where 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 $C^1(\tau_1, \tau_2) := (1 - \omega)C(\tau_1, \tau_2) + \omega C^*(\tau_1, \tau_2)$.

3.4.2 Local Alternatives

Next consider a sequence of local alternatives

$$(1 - g(t, T))C(\tau_1, \tau_2) + g(t, T)C^*(\tau_1, \tau_2), \tag{3.7}$$

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.

Under H_1 the copula (weight) depends on the sample size T. Hence, to analyze the local power of the test we should formally work with triangular arrays $y_{tT} = \{y_{itT}\}_{i=1,2}$, $t = 1, 2, \ldots, T$, $T \in \mathbb{N}$. For notational simplicity, the partial sums $\hat{S}_T$ are like before but implicitly depending on y_{tT} instead of y_t.

By making use of a functional central limit theorem for triangular arrays we are able to show the limiting behavior of the test under the sequence of local alternatives (3.7).

Theorem 3.4. *Under local alternatives* (3.7)

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

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

This shows that the copula constancy test is inconsistent against local alternatives (3.7) but does converge to a fixed limit.

3.5 Finite sample properties

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).

3.5.1 Size of test

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

$$\begin{aligned} x_{it} &= \theta_1 x_{i,t-1} + \varepsilon_{i,t} + \theta_2 \varepsilon_{i,t-1} \\ \varepsilon_{it} &= h_{it} \varepsilon_{it}^{\dagger} \\ h_{it}^2 &= \theta_3 + \theta_4 \varepsilon_{it}^2 + \theta_5 h_{i,t-1}^2, \end{aligned} \tag{3.8}$$

where $\varepsilon_{it}^{\dagger} = \Phi^{-1}(u_{it})$, $\Phi(\cdot)$ 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 rules $\gamma_{1T} = [4(T/100)^{1/4}]$ 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 3.1 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.

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 3.2 shows the size of the test as applied to the original data and without HAC 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.

3.5.2 The power of the test

We consider the power of the test against the fixed alternative (3.6). 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 3.3 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 3.1: 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 (3.8) 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 3.2: 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 (3.8) 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 3.3: Power of *squares* test against the fixed alternative (3.6).

			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 (3.8) 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

3.6 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 (3.8) 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 3.4) equals 0, see Laurent and Peters (2006).

Table 3.4 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 in 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.

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 3.3, Theorem 3.2 allows us to apply the copula constancy test to the standardized innovations.

Since quantiles can also change for reasons different from stochastic volatility, we perform the quantile constancy test proposed by Busetti and Harvey (2007). Table 3.5

Table 3.4: 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.

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 3.5: 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

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 3.6 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.

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 $\{y_{1t} \le \xi_1(\tau_1), y_{2t} \le \xi_2(\tau_2)\}$. In particular, using the Quadrant Association Test of Busetti and Harvey (2011) (which is based on the same idea but uses the events $\{y_{1t} \le \xi_1(\tau_1), y_{2t} \le \xi_2(\tau_2)\}$ 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.

Table 3.6: 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

3.A Appendix: Proof Theorem 1 and 2

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 3.5. *Write $y_t = y_t(\theta_0)$ and $\xi(\tau) = \xi(\theta_0, \tau)$. For $M > 0$, we have under Assumption 3.1*

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

where

$$d_t(\phi) = I(y_t, \xi(\tau) + \phi T^{-1/2}) - I(y_t, \xi(\tau)). \tag{3.9}$$

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 3.6 proves pointwise convergence and Lemma 3.7 proves stochastic equicontinuity. ∎

Lemma 3.6. *Let $M > 0$. Then, under Assumption 3.1, 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 3.1) 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 $c_3 > 0$,

$$
\begin{aligned}
E \sup_{r\in[0,1]} & \left(T^{-1/2} \sum_{t=1}^{[rT]} |d_t(\phi) - E[d_t(\phi)]| \right)^2 \\
\leq \quad & c_1 T^{-1} \sum_{t=1}^{T} \| I(y_t, \xi(\tau) + \phi T^{-1/2}) - I(y_t, \xi(\tau)) \|_p^2 \\
\leq \quad & c_2 T^{-1} \sum_{t=1}^{T} \left(F(MT^{-1/2}, MT^{-1/2}) - F(-MT^{-1/2}, -MT^{-1/2}) \right)^{2/p} \\
\leq \quad & c_3 \Bigg(\sup_{(a_1,a_2)\in[-\eta,\eta]^2} F_1'(a_1,a_2)(2MT^{-1/2}) \\
& \qquad + \sup_{(a_3,a_4)\in[-\eta,\eta]^2} F_2'(a_3,a_4)(2MT^{-1/2}) \Bigg)^{2/p},
\end{aligned}
\tag{3.10}
$$

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 3.7. *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, p.336)) 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,

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

$$\leq \delta \left(\sup_{(a_1,a_2)\in[-\eta,\eta]^2} F_1'(a_1,a_2) + \sup_{(a_3,a_4)\in[-\eta,\eta]^2} F_2'(a_3,a_4) \right). \tag{3.11}$$

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

Proof of Theorem 3.1:
Define $\phi^* := T^{1/2}(\hat{\xi}(\tau) - \xi(\tau))$ and $d_t(\phi)$ as in (3.9). Then

$$\begin{aligned}
&\frac{1}{T^{1/2}} \sum_{t=1}^{[rT]} BIQ(y_t, \hat{\xi}(\tau)) \qquad (3.12)\\
&= \frac{1}{T^{1/2}} \sum_{t=1}^{[rT]} BIQ(y_t, \xi(\tau)) - \frac{1}{T^{1/2}} \sum_{t=1}^{[rT]} d_t(\phi^*) + \frac{[rT]}{T} \frac{1}{T^{1/2}} \sum_{t=1}^{T} d_t(\phi^*)\\
&= \frac{1}{T^{1/2}} \sum_{t=1}^{[rT]} BIQ(y_t, \xi(\tau)) - \frac{1}{T^{1/2}} \sum_{t=1}^{[rT]} (d_t(\phi^*) - E[d_t(\phi^*)])\\
&\quad + \frac{[rT]}{T} \frac{1}{T^{1/2}} \sum_{t=1}^{T} (d_t(\phi^*) - E[d_t(\phi^*)]) .
\end{aligned}$$

Under Assumption 3.3 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

$$\tilde{\sigma}^2 := T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} k((t-s)/\gamma_T) \cdot BIQ(y_t, \xi(\tau)) \cdot BIQ(y_s, \xi(\tau)). \tag{3.13}$$

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 $\bar{\sigma}^2$.
Step 1: Write

$$BIQ(y_t, \hat{\xi}(\tau)) = BIQ(y_t, \xi(\tau)) - a_t + b_T, \tag{3.14}$$

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: Note that $-BIQ(y_s,\xi(\tau))$ are the OLS residuals of regressing $I(y_t,\xi(\tau))$ on a constant. These residuals (as function of the parameter vector) satisfy Assumption B and C in Andrews (1991). Hence, his Theorem 1b gives the required result. ■

The following lemmas are used to prove Theorem 2. The structure of the proof is similar to Theorem 1. Lemma 3.8 shows uniform convergence of some terms that occur in the proof of Theorem 2 below. To prove Lemma 3.8, we show pointwise convergence in Lemma 3.9 and stochastic equicontinuity in Lemma 3.10.

Lemma 3.8. *For $M > 0$ and $N > 0$ we have under Assumption 3.1 and 3.4*

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

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 (3.9).

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 3.5, it is sufficient to prove that the first part converge uniformly to zero as well. Like Lemma 3.5, 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 3.9 proves pointwise convergence and Lemma 3.10 proves stochastic equicontinuity. ■

Lemma 3.9. *Let $M>0$ and $N>0$. Then, under Assumption 3.1 and 3.4, 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 3.1

$$
\begin{aligned}
&E\sup_{r\in[0,1]}\left|T^{-1/2}\sum_{t=1}^{[rT]}d_t^\dagger(\phi,\zeta)-Ed_t^\dagger(\phi,\zeta)\right| \\
&\quad\le c_1T^{-1}\sum_{t=1}^{T}\left(E\left|I[y_t(\theta_0+\zeta T^{-1/2}),\xi(\theta_0+\zeta T^{-1/2},\tau)+\phi T^{-1/2}]\right.\right. \\
&\qquad\qquad\left.\left.-\ I[y_t(\theta_0),\xi(\theta_0,\tau)+\phi T^{-1/2}]\right|^p\right)^{2/p}
\end{aligned}
$$

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

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

Lemma 3.10. *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 3.7, 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

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

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

$$
\begin{aligned}
& T^{-1/2}\sum_{j=1}^{T}\left|Ed_j^{\dagger}(\phi,\zeta)-Ed_j^{\dagger}(\ddot{\phi},\ddot{\zeta})\right| \\
&= T^{-1/2}\sum_{j=1}^{T}\left|F\Big(T^{-1/2}\zeta' c_j^* + \phi_1 T^{-1/2}, T^{-1/2}\zeta' c_j^* + \phi_2 T^{-1/2}\Big) - \ F(\phi_1 T^{-1/2},\phi_2 T^{-1/2})\right. \\
&\qquad \left. -\left\{F\Big(T^{-1/2}\ddot{\zeta}' c_j^* + \ddot{\phi}_1 T^{-1/2}, T^{-1/2}\ddot{\zeta}' c_j^* + \ddot{\phi}_2 T^{-1/2}\Big) - \ F(\ddot{\phi}_1 T^{-1/2},\ddot{\phi}_2 T^{-1/2})\right\}\right| \\
&= T^{-1/2}\sum_{j=1}^{T}\left|T^{-1/2}\zeta' c_j^*[F_1'(\phi_1 T^{-1/2},\phi_2 T^{-1/2}) + F_2'(\phi_1 T^{-1/2},\phi_2 T^{-1/2})] + O(T^{-1})\right. \\
&\qquad \left. -\left\{T^{-1/2}\ddot{\zeta}' c_j^*[F_1'(\ddot{\phi}_1 T^{-1/2},\ddot{\phi}_2 T^{-1/2}) + F_2'(\ddot{\phi}_1 T^{-1/2},\ddot{\phi}_2 T^{-1/2})] + O(T^{-1})\right\}\right| \\
&\le \sup_{j=1,\dots,T}|\zeta' c_j^*|\cdot\Big|F_1'(\phi_1 T^{-1/2},\phi_2 T^{-1/2}) + F_2'(\phi_1 T^{-1/2},\phi_2 T^{-1/2}) \\
&\qquad\qquad - F_1'(\ddot{\phi}_1 T^{-1/2},\ddot{\phi}_2 T^{-1/2}) - F_2'(\ddot{\phi}_1 T^{-1/2},\ddot{\phi}_2 T^{-1/2})\Big| \\
&\quad + \sup_{j=1,\dots,T}|(\zeta'-\ddot{\zeta}')\cdot c_j^*|\cdot|F_1'(\ddot{\phi}_1 T^{-1/2},\ddot{\phi}_2 T^{-1/2}) + F_2'(\ddot{\phi}_1 T^{-1/2},\ddot{\phi}_2 T^{-1/2})| + O(T^{-1/2}) \\
&\le \sup_{j=1,\dots,T}|\zeta' c_j^*|\cdot\big\{|F_{11}''(b_1,b_2) + F_{21}''(b_3,b_4)|\cdot|\ddot{\phi}_1 T^{-1/2} - \phi_1 T^{-1/2}| \\
&\qquad\qquad + |F_{12}''(b_1,b_2) + F_{22}''(b_3,b_4)|\cdot|\ddot{\phi}_2 T^{-1/2} - \phi_2 T^{-1/2}|\big\} \\
&\quad + \sup_{j=1,\dots,T}|(\zeta'-\ddot{\zeta}')\cdot c_j^*|\cdot|F_1'(\ddot{\phi}_1 T^{-1/2},\ddot{\phi}_2 T^{-1/2}) + F_2'(\ddot{\phi}_1 T^{-1/2},\ddot{\phi}_2 T^{-1/2})| + O(T^{-1/2}),
\end{aligned}
$$

where (b_1,b_2) and (b_3,b_4) are points between $(\phi_1 T^{-1/2},\phi_2 T^{-1/2})$ and $(\ddot{\phi}_1 T^{-1/2},\ddot{\phi}_2 T^{-1/2})$. Under Assumption 3.4, F_i', F_{ik}'' and c_j^* are bounded, so that for some constants c_1 and c_2

$$
T^{-1/2}\sum_{j=1}^{T}\left|Ed_j^{\dagger}(\phi,\zeta)-Ed_j^{\dagger}(\ddot{\phi},\ddot{\zeta})\right| \ \le \ c_1 T^{-1/2}\delta_\phi + c_2\delta_\zeta + O(T^{-1/2}).
$$

Selecting δ_ϕ and δ_ζ sufficiently small completes the proof. ∎

Proof of Theorem 3.2

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

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

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 3.1, uniform convergence by Lemma 3.8.

It 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 converges to zero by Theorem 3.1. It is sufficient to show that

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

Write

$$\begin{aligned}
BIQ(y_t(\hat{\theta}_T), \hat{\xi}(\hat{\theta}_T, \tau)) &= BIQ(y_t(\theta_0), \xi(\theta_0, \tau)) - [d_t^*(\phi^*, \zeta^*) - E(d_t^*(\phi^*, \zeta^*))] \\
&\qquad + \frac{1}{T} \sum_{k=1}^{T} [d_k^*(\phi^*, \zeta^*) - E(d_k^*(\phi^*, \zeta^*))] \\
&=: BIQ(y_t(\theta_0), \xi(\theta_0, \tau)) - a_t(\hat{\theta}_T) + b_T(\hat{\theta}_T). \qquad (3.15)
\end{aligned}$$

Using the definition of the HAC estimator (3.5) and equations (3.14) and (3.15)

$$
\begin{aligned}
&|\hat{\sigma}^2(\hat{\theta}_T) - \hat{\sigma}^2| \qquad (3.16)\\
&= \Big|T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T} k((t-s)/\gamma_T)\{(BIQ(y_t(\theta_0),\xi(\theta_0,\tau)) - a_t(\hat{\theta}_T) + b_T(\hat{\theta}_T))\\
&\qquad \times (BIQ(y_s(\theta_0),\xi(\theta_0,\tau)) - a_s(\hat{\theta}_T) + b_T(\hat{\theta}_T))\\
&- (BIQ(y_t(\theta_0),\xi(\theta_0,\tau)) - a_t + b_T)(BIQ(y_s(\theta_0),\xi(\theta_0,\tau)) - a_s + b_T)\}\Big|.
\end{aligned}
$$

We show that the difference of the cross-products converge to zero. Write

$$
\begin{aligned}
&|T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T} k((t-s)/\gamma_T)\{a_t(\hat{\theta}_T)a_s(\hat{\theta}_T) - a_t a_s\}|\\
&\quad = \big|T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T} k((t-s)/\gamma_T) \qquad (3.17)\\
&\qquad \times \{\frac{1}{2}(a_t(\hat{\theta}_T) + a_t)(a_s(\hat{\theta}_T) - a_s) + \frac{1}{2}(a_t(\hat{\theta}_T) - a_t)(a_s(\hat{\theta}_T) + a_s)\}\big|.
\end{aligned}
$$

For constants $c_1 > 0$ and $c_2 > 0$, we have

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

because the first term is $o_p(1)$ by Lemma 3.8. Using the same idea for the first term in (3.17) gives the required result.

In addition,

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

where $T^{-1/2}|b_T(\hat\theta_T)-b_T| = o_p(1)$ by Lemma 3.8.

The other cross products in (3.16) follow a similar argument.

■

Proof of Theorem 3.3

Define the bivariate τ-quantics corresponding to the alternative hypothesis (3.6) 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

$$
BIQ(y_t,\xi(\tau)) = BIQ^1(y_t,\xi(\tau)) + \Delta_C\cdot[g(t,T)-\frac{1}{T}\sum_{j=1}^{T}g(j,T)], \tag{3.18}
$$

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

$$
\begin{aligned}
&1/(\sqrt{T})\sum_{t=1}^{[rT]}BIQ(y_t,\xi(\tau)) \\
&\quad= 1/(\sqrt{T})\sum_{t=1}^{[rT]}BIQ^1(y_t,\xi(\tau)) + \Delta_C\cdot\omega\frac{[z^*T]}{\sqrt{T}}\left(\frac{[rT]}{T}-1\right).
\end{aligned} \tag{3.19}
$$

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 $\hat{\xi}(\tau)$. Therefore, it remains to show that under the alternative hypothesis

$$\sup_{r\in[0,1]}\left|\frac{1}{\sqrt{T\sigma^2}}\sum_{t=1}^{[rT]}BIQ(y_t,\hat{\xi}(\tau))-\frac{1}{\sqrt{T\sigma^2}}\sum_{t=1}^{[rT]}BIQ(y_t,\xi(\tau))\right|\xrightarrow{p}0. \tag{3.20}$$

Recall that we showed this, under the null hypothesis, in Theorem 3.1.

So again we rewrite $1/(\sqrt{T})\sum_{t=1}^{[rT]}BIQ(y_t,\hat{\xi}(\tau))$ like (3.12). Note, however, that under the alternative the joint distribution $F(\cdot,\cdot)$ depends on time t. The application of Lemma 3.5, 3.6 and 3.7 is not allowed since the proof of the latter two makes use of the fact that $F(\cdot,\cdot)$ is constant over time. We can, however, adjust the proof by replacing $F(\cdot,\cdot)$ by $F_t(\cdot,\cdot)$ and adding $\sup_t$ to equations (3.10) and (3.11).

We now show that $\hat{\sigma}^2$ is $O_p(\sqrt{T})$ under the alternative. From (3.5) and (3.18)

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

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

$$\begin{aligned}\left|\hat{\sigma}^2-\sigma_1^2\right| = \Bigg|T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T}k((t-s)/\gamma_T)\Bigg\{&BIQ^1(y_t,\hat{\xi}(\tau))BIQ^1(y_s,\hat{\xi}(\tau)) \\ -BIQ^1(y_t,\hat{\xi}(\tau))\tilde{g}_t-BIQ^1(y_s,\hat{\xi}(\tau))\tilde{g}_s+\tilde{g}_t\tilde{g}_s&-Q^1(y_t,\xi(\tau))Q^1(y_s,\xi(\tau))\Bigg|\Bigg\}.\end{aligned} \tag{3.21}$$

Note that

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

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 3.2 become infinity large as $T \to \infty$, and thus the probability that the test statistic exceeds the critical value goes to 1 as $T \to \infty$.

$\blacksquare$

Proof of Theorem 3.4

The asymptotic distribution follows from (3.18) and the FCLT. Setting $g(t/T) = T^{-1/2}h(t/T)$ in (3.21) gives $|\hat{\sigma}^2 - \sigma_1^2| = o_p(1)$ using arguments as in the proof of Theorem 3.

$\blacksquare$

Chapter 4

Optimal Bandwidth Selection in Robust Copula Constancy Tests

4.1 Introduction

In this chapter we consider a new copula constancy tests that performs better in finite samples compared to the recently proposed copula constancy test of Busetti and Harvey (2011). The difference between our test and the existing test concerns the estimation of the long-run variance. The conventional approach is to replace the long run variance by a heteroskedasticity and autocorrelation consistent (HAC) variance estimate and to construct such an estimate using a kernel-based approach (see Den Haan, and Levin (1997) for an overview). The kernel depends on a bandwidth parameter, γ_T, which should grow slower than the sample size, T, to obtain consistent estimates. Kiefer and Vogelsang (2002, 2005) construct inconsistent estimates by setting bandwidth equal or proportional to the sample size, i.e. $\gamma_T = bT$ with $b \in (0, 1]$. The idea is that the additional variability, due to the inconsistency of the estimate, might improve the size of the test (see also Jansson (2004)). Phillips, Sun and Jin (2006, 2007) introduce an alternative class of inconsistent estimates that are based on exponentiated kernel functions with bandwidth equal to sample size.

The copula constancy test follows from ideas established by De Jong et al. (2007)

for the indicator KPSS test. Setting the bandwidth equal to the sample size is somewhat problematic in the stationarity test proposed by Kwiatkowski et al. (1992), since their test becomes inconsistent for a fixed b (see Kiefer and Vogelsang (2002), Müller (2005)). Amsler et al. (2009) show, however, that the resulting asymptotic theory provides a better approximation to the finite sample distribution of the statistic.

The remainder of this chapter is structured as follows. First, we introduce the copula constancy test based on the contracted kernel and give the nonstandard asymptotic distribution. Subsequently, we derive the optimal bandwidth rule. Finally, we examine the performance of the test and bandwidth rule using simulations and we illustrate them using an empirical application of MSCI stock returns.

4.2 The copula constancy test

Let $y_t = \{y_{it}\}_{i=1}^2$ be a bivariate time series, $t = 1, \dots, T$ and let $\xi(\tau) = (\xi_1(\tau_1), \xi_2(\tau_2))$ denote the vector of marginal τ_i-quantiles, $\tau_i \in (0,1)$, $i = 1, 2$. The copula $C^{(t)}(\tau_1, \tau_2)$ gives the probability that y_{1t} takes values below its τ_1-quantile and y_{2t} takes values below its τ_2-quantile

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

We examine the constancy of the copula in the point (τ_1, τ_2) (see section 3). The hypothesis pair is given by

$$\begin{aligned} &H_0 : C^{(t)}(\tau_1, \tau_2) = C(\tau_1, \tau_2) \\ &H_1 : C^{(t)}(\tau_1, \tau_2) = [1 - h(t/T)T^{-1/2}]C(\tau_1, \tau_2) + h(t/T)T^{-1/2}C^*(\tau_1, \tau_2), \end{aligned} \tag{4.1}$$

where $C(\tau_1, \tau_2)$ and $C^*(\tau_1, \tau_2)$ are two different time-invariant copulas and

$$h(t/T) = \begin{cases} 0 & \text{if } t/T \le z^* \\ \delta & \text{if } t/T > z^*, \end{cases}$$

with jump magnitude $\delta \in (0, 1]$ and time fraction $z^* \in (0, 1)$.

Busetti and Harvey (2011) and Krämer and Van Kampen (2011) propose copula constancy tests based on the partial sums of an indicator series that takes the value one if the event $\{y_{1t} \leq \xi_1(\tau_1), y_{2t} \leq \xi_2(\tau_2)\}$ occurs and zero otherwise.

Let $I(\cdot)$ denote the indicator function and define

$$
\begin{aligned}
I_t &:= I(y_{1t} \leq \xi_1(\tau_1),\ y_{2t} \leq \xi_2(\tau_2)), \\
Q_t &:= Q(y_t, \xi(\tau)) := C(\tau_1, \tau_2) - I_t, \\
BIQ_t &:= BIQ(y_t, \xi(\tau)) := Q_t - T^{-1} \sum_{i=1}^{T} Q_t.
\end{aligned}
$$

Then $BIQ_t = C_T(\tau_1, \tau_2) - I_t$ where $C_T(\tau_1, \tau_2) = T^{-1} \sum_{t=1}^{T} I_t$ is the empirical copula. We define $\widehat{BIQ_t}$ as BIQ_t above but using the sample quantiles $\hat{\xi}(\tau)$ instead of the population quantiles $\xi(\tau)$. In addition, write

$$
S_T(r) := \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} Q_t, \qquad r \in [0,1].
$$

We make the following assumption:

Assumption 1.

$$
S_T(r) \xrightarrow{d} \sigma B(r) \qquad r \in [0,1],
$$

where $\sigma^2 = \lim_{T\to\infty} E\left(T^{-1/2} \sum_{t=1}^{T} Q_t\right)^2$ denotes the long run variance and B denotes Brownian Motion.

Under Assumption 1

$$
1/(\sqrt{T}) \sum_{t=1}^{[Tr]} BIQ_t \xrightarrow{d} \sigma V(r) \qquad r \in [0,1], \tag{4.2}
$$

where $V(r) := B(r) - rB(1)$ denotes a Brownian Bridge. Van Kampen and Wied (2010) provide lower level conditions under which assumption 1 is satisfied (see also section 3). Furthermore, they show that (4.2) also holds for the partial sums of $\widehat{BIQ_t}$.

The copula constancy tests are functionals of the partial sum process. Following

Busetti and Harvey (2011), we consider the test statistic

$$\eta_T(\sigma^2) := \frac{1}{T^2\hat{\sigma}^2}\sum_{t=1}^{T}\left(\sum_{j=1}^{t}\widehat{BIQ_{tj}}\right)^2, \tag{4.3}$$

where $\hat{\sigma}^2$ is a consistent estimate of the long run variance σ^2. Krämer and Van Kampen (2011) proposed complementary tests based on the maximum and the range of the partial sum process in (4.2).

Under H_0 we have

$$\eta_T(\hat{\sigma}^2) \xrightarrow{d} \int_0^1 V(r)^2 dr.$$

Under the local alternative (4.1) we have

$$\begin{aligned}\eta_T(\hat{\sigma}^2) \quad \xrightarrow{d} \quad & \int_0^1 \Bigg[V(r) + \frac{1}{\sigma_1}\{C(\tau_1,\tau_2) - C^*(\tau_1,\tau_2)\} \\ & \qquad\qquad \times \left(\int_0^r h(s)ds - r\int_0^1 h(s)ds\right)\Bigg]^2 dr,\end{aligned}$$

where $\sigma_1^2 = \lim_{T\to\infty} E\left(T^{-1/2}\sum_{t=1}^{T}[C^{(t)}(\tau_1,\tau_2) - I_t]\right)^2$.

The present chapter is concerned with the implications that different estimates of σ^2 have on the properties of the test.

4.3 Long run variance estimation

4.3.1 conventional HAC estimates

The HAC estimate of σ^2 is given by

$$\tilde{\sigma}^2 = T^{-1}\sum_{i=1}^{T}\sum_{j=1}^{T} k((i-j)/\gamma_T)\cdot Q_i \cdot Q_j, \tag{4.4}$$

where $k(\cdot)$ is a kernel function and γ_T is the bandwidth parameter. Some examples for $k(\cdot)$ are the Bartlett, Parzen and Quadratic Spectral (QS) kernels, which are given by

$$
\begin{aligned}
k_{BT}(x) &= \begin{cases} 1-|x| & \text{for } |x| \le 1 \\ 0 & \text{otherwise} \end{cases} \\
k_{PR}(x) &= \begin{cases} 1-6x^2+6|x|^3 & \text{for } 0 \le |x| \le 1/2 \\ 2(1-|x|)^3 & \text{for } 1/2 \le |x| \le 1 \\ 0 & \text{otherwise} \end{cases} \\
k_{QS}(x) &= \frac{25}{12\pi^2 x^2}\left(\frac{\sin(6\pi x/5)}{6\pi x/5} - \cos(6\pi x/5)\right),
\end{aligned}
$$

see e.g. Priestley (1981) for a complete discussion of these kernels.

Andrews (1991) and De Jong and Davidson (2000) provide conditions under which (4.4) is a consistent estimate for σ^2. In particular, they require that $\gamma_T/T \to 0$ as $\gamma_T \to \infty$ and $T \to \infty$. Andrews (1991) showed that the QS kernel is preferred to the other kernels mentioned above, in the sense that it minimizes the asymptotic mean square error (MSE) of the estimate of the long run variance. In the context of hypothesis testing, the use of the mean square error as optimality criterion is somewhat questionable. In particular, it might be possible that deviations in $\hat{\sigma}^2$ from σ^2 are partially offset by the deviations between the finite sample distribution and the limit distribution.

Note that the estimate (4.4) is not feasible since Q_t relies on the true unobserved copula $C(\tau_1, \tau_2)$ and on the population quantile $\xi(\tau)$. Therefore, we replace $C(\tau_1, \tau_2)$ by the empirical copula $C_T(\tau_1, \tau_2)$ and $\xi(\tau)$ by its sample estimate $\hat{\xi}(\tau)$. The feasible HAC estimator is given by

$$\hat{\sigma}^2 = T^{-1} \sum_{i=1}^{T} \sum_{j=1}^{T} k((i-j)/\gamma_T) \cdot \widehat{BIQ}_i \cdot \widehat{BIQ}_j. \tag{4.5}$$

Van Kampen and Wied (2010) show consistency of (4.5) for $\gamma_T = o(T)$.

In line with the KPSS stationarity test, the copula constancy test is subject to size distortions if y_{it} is highly autocorrelated. Kiefer and Vogelsang (2002, 2005) and Phillips, Sun and Jin (2006, 2007) argued that the use of inconsistent estimates results in tests with better size performance.

4.3.2 Fixed-b asymptotics

Kiefer and Vogelsang (2002, 2005) proposed a HAC estimator based on kernels that have bandwidth equal or proportional to sample size, i.e. $\gamma_T = bT$, $b \in (0,1]$. Equivalently, we could define the kernel as $k_b := k(x/b)$ with bandwidth equal to sample size. The kernel k_b is referred to as a contracted kernel (Phillips, Sun and Jin (2007)). Kiefer and Vogelsang (2005) show that HAC estimates based on k_b are inconsistent for fixed b and regression t-tests based on these estimates have a nonstandard limiting distribution.

Let $\hat{\sigma}^2_{k_b}$ be the estimate (4.5) based on the kernel k_b. The limit distribution under the null hypothesis follows immediately from arguments as in Theorem 1 of Amsler et al. (2009):

$$\eta_T(\hat{\sigma}^2_{k_b}) \xrightarrow{d} \int_0^1 V(r)^2 dr \Xi_b^{-1},$$

where Ξ_b is defined as follows:

(i) if $k(x)$ is twice continuously differentiable everywhere

$$\Xi_b := -\frac{1}{b^2}\int_0^1\int_0^1 k''\left(\frac{r-s}{b}\right)V(r)V(s)drds;$$

(ii) if $k(x) = 0$ for $|x| \geq 1$, $k(x)$ is twice continuously differentiable everywhere except possibly at $|x| = 1$ and $k'(1) = \lim_{h\to 0}\frac{k(1)-k(1-h)}{h}$

$$\Xi_b := -\frac{1}{b^2}\iint_{|r-s|\leq b} k''\left(\frac{r-s}{b}\right)V(r)V(s)drds + \frac{2}{b}k'(1)\int_0^{1-b} V(r)V(r+b)dr;$$

(iii) if $k(x)$ is equal to the Bartlett kernel

$$\Xi_b := \frac{2}{b}\left[\int_0^1 V(r)^2 dr - \int_0^{1-b} V(r)V(r+b)dr\right].$$

An example of a kernel that satisfies (i) is the QS kernel. The Parzen kernel satisfies (ii). Note that the distribution is nonstandard and depends on the choice of the kernel and the value of b. Amsler et al. (2009) provide critical values for the Bartlett kernel and the QS kernel.

From Van Kampen and Wied (2010) we have that the test is inconsistent against the local alternative (4.1). Following Amsler et al. (2009) we only use fixed-b critical values to improve the *finite sample* properties of the test and we select b using conventional bandwidth rules. Given a finite sample of size T, empirical researchers have to select a fixed bandwidth value, γ_T, which can be related to b by $b = \gamma_T/T$. Amsler et al. (2009) show for the KPSS stationarity test that using the fixed-b critical values corresponding to such a b reduces size distortions. Using their line of reasoning a similar result can the be obtained for the copula constancy test.

Given the bandwidth rules $\gamma_T(m) = int[m(T/100)^{1/4}]$, $m = 0, 4, 12, 25$ and 50, they choose m such that the size distortion of the test using conventional as well as fixed-b critical values falls below some specific threshold. It is, however, completely unclear if such a rule is optimal. Therefore, we derive an optimal bandwidth rule in section 4.4.

Finally, we would like to mention that an alternative copula constancy test can be obtained using the exponentiated kernels considered in Phillips, Sun and Jin (2006, 2007). These kernels are defined as

$$k_\rho(x) := k^\rho(x), \tag{4.6}$$

with bandwidth equal to sample size

Phillips, Sun and Jin (2006, 2007) show that taking ρ fixed results in an inconsistent HAC estimate while taking $\rho \to \infty$ as $T \to \infty$ results in a consistent HAC estimate. Let $\hat{\sigma}^2_\rho$ denote the HAC estimate of σ^2. Then for fixed ρ we have

$$\hat{\sigma}^2_\rho \xrightarrow{d} \sigma^2 \Big(\int_0^1 \int_0^1 k_\rho(r-s) dV(r) dV(s) \Big). \tag{4.7}$$

From (4.2) and (4.7) we obtain

$$\eta_T(\hat{\sigma}^2_{k_\rho}) \xrightarrow{d} \int_0^1 V(r)^2 dr \Big(\int_0^1 \int_0^1 k_\rho(u-s) dV(u) dV(s) \Big)^{-1}. \tag{4.8}$$

Here, it is not immediately obvious how to select ρ using the approach in Amsler, Schmidt and Vogelsang (2009). For comparison purposes it might be useful to relate the values of ρ and b to each other using a first order expansion of the kernel around the origin. We leave the application of exponentiated kernels as a topic for further research.

4.4 Optimal bandwidth selection

Sun, Phillips and Jin (2008) suggest to select b by minimizing a weighted average of the type I and type II error. An optimal rule for ρ using this approach has been derived in Sun, Phillips and Jin (2010). The idea is to make suitable expansions of the limit and finite sample distributions and express the size distortion and power of the test using these expansions. Subsequently, they minimize a weighted average of the type I and type II error. In this section we derive a similar rule for the copula constancy test.

We assume that the kernel satisfies Assumption 2 in Sun, Phillips and Jin (2008). This assumption ensures that the kernel is positive semidefinite and sufficiently smooth. The Bartlett, Parzen and QS kernels considered before satisfy this assumption.

4.4.1 Expansion of the limit distribution

Lemma 1 shows that the numerator of the limit distribution can be written as an infinite weighted sum of (noncentral) χ_1^2 distributed variables.

Lemma 1.

$$\eta_T(\hat{\sigma}_b^2) \xrightarrow{d} \left(\sum_{n=1}^{\infty} [q_{1,n}(\zeta_n + q_{4,n})^2] + q_\infty \right) \Xi_b^{-1}, \tag{4.9}$$

where ζ_n are i.i.d. N(0,1) distributed variables and

(i) under H_0 we have $q_{1,n} = (n\pi)^{-2}$, $q_{2,n} = 0$ and $q_\infty = 0$.

(ii) under H_1 we have $q_{1,n} = (\pi n)^{-2}$, $q_{4,n} = -2^{1/2} a \sin(n\pi z^*)/(n\pi z^*)$ and $q_\infty := \lim_{p\to\infty} q_p := \lim_{p\to\infty} \left[a^2(\frac{1}{3}z^{*2} - \frac{2}{3}z^* + \frac{1}{3}) - \sum_{n=1}^{p} q_{4,n}^2 \right]$, with constant $a = \sigma_1^{-1}\delta z^* \times [C(\tau_1, \tau_2) - C^*(\tau_1, \tau_2)]$.

We approximate the infinite sums in (4.9) by the sum of the first p terms. Durbin and Knott (1972) considered a similar approximation.

Let $F_\delta(z) := P\left(\left[\sum_{n=1}^p \{q_{1,n}(\zeta_n + q_{4,n})^2\} + q_p\right]\Xi_b^{-1} \leq z\right)$ be the limit distribution and let $G_\delta(\cdot)$ denote the cdf of $\sum_{n=1}^p [q_{1,n}(\zeta_n + q_{4,n})^2]$, which is a weighted sum of noncentral $\chi_1^2(q_{4,n}^2)$ distributed random variables with noncentrality parameter $q_{4,n}^2$, $n = 1, \ldots, p$. Theorem 2 provides an asymptotic expansion of $F_\delta(z)$ around $G_\delta(z)$.

Theorem 2.

$$
\begin{aligned}
F_\delta(z) \quad = \quad & G_\delta(z - q_p) + [c_2 G''_\delta(z - q_p)z^2 - c_1 G'_\delta(z - q_p)z]b \\
& - [G'_\delta(z - q_p)zc_3 - \frac{1}{2}G''_\delta(z - q_p)z^2(2c_4 - c_1^2) + G'''_\delta(z - q_p)z^3 c_1 c_2]b^2 \\
& + o(b^2), \qquad (4.10)
\end{aligned}
$$

where

$$
\begin{aligned}
c_1 &= \int_{-\infty}^{\infty} k(x)dx \qquad c_2 = \int_{-\infty}^{\infty} k^2(x)dx \\
c_3 &= \int_{-\infty}^{\infty} k(x)|x|dx \qquad c_4 = -\int_{-\infty}^{\infty} k^2(x)|x|dx.
\end{aligned}
$$

Note that the expansion only depends on the copula through the weights and noncentrality parameters of the noncentral χ_1^2 distributions. We conjecture that the expansions derived in this section are, more generally, applicable to tests (with consistent variance estimates) which are distributed as a weighted sum of noncentral χ_1^2 distributed variables. For example, the expansion of the limit distribution of the regression F-test in Sun, Phillips and Jin (2008) can be obtained as a special case.

The expansion (4.10) allows us to obtain an analytical expression for the critical values of the nonstandard limit distribution. Following Sun, Phillips and Jin (2008), we define the second-order corrected critical value as the critical value that is correct to $O(b)$ and the third-order corrected critical value as the critical value that is correct to $O(b^2)$. Let $D(\cdot) = G_0(\cdot)$. For given level α, define $z_\alpha \in \mathbb{R}^+$ such that $D(z_\alpha) = 1 - \alpha$ and define $z_{\alpha,b} \in \mathbb{R}^+$ such that $F_0(z_{\alpha,b}) = 1 - \alpha$.

Corollary 3. The second-order corrected critical values are given by

$$z_{\alpha,b} = z_\alpha + k_1 b + o(b)$$

and the third-order corrected critical values are given by

$$z_{\alpha,b} = z_\alpha + k_1 b + k_2 b^2 + o(b^2),$$

where

$$k_1 = -\frac{1}{D'(z_\alpha)}[D''(z_\alpha)z_\alpha^2 c_2 - D'(z_\alpha)z_\alpha c_1] \tag{4.11}$$

and

$$\begin{aligned} k_2 = & -\frac{1}{D'(z_\alpha)}\bigg[-D'(z_\alpha)z_\alpha c_3 + \frac{1}{2}D''(z_\alpha)z_\alpha^2(2c_4 - c_1^2) - D'''(z_\alpha)z_\alpha^3 c_1 c_2 \\ & +[c_2 D'''(z_\alpha)z_\alpha^2 + 2c_2 D''(z_\alpha)z_\alpha - c_1 D''(z_\alpha)z_\alpha - c_1 D'(z_\alpha)]k_1 \\ & +\frac{1}{2}D''(z_\alpha)k_1^2\bigg]. \end{aligned} \tag{4.12}$$

To obtain feasible estimates for the coefficients (4.11) and (4.12), we need closed-form expressions for the distribution $D(\cdot)$ and its derivatives. In general, these are not available. We propose two solutions. First, numerical values for the distribution can be obtained using the approach proposed by Imhof (1961). Subsequently, numerical differentiation gives the values for $D'(\cdot)$ and $D''(\cdot)$. Second, we can approximate the distribution using more conventional distributions. An overview of approximating the distribution can be found in Ullah (2004, chapter 3). Here, we consider the following approximation proposed by Zhang (2005):

$$G_\delta(z) \approx P(\chi^2_{d_\chi} \leq (z - b_\chi)/a_\chi) =: \tilde{G}_\delta(f(z)), \tag{4.13}$$

where $\chi^2_{d_\chi}$ is gamma distributed with parameters $\alpha_g := d_\chi/2$ and $1/2$, $f(z) = (z - a_\chi)/b_\chi$

and

$$
\begin{aligned}
a_\chi &= \frac{\sum_{n=1}^p q_{1,n}^3(1+3q_{4,n}^2)}{\sum_{n=1}^p q_{1,n}^2(1+2q_{4,n}^2)}, && (4.14)\\
b_\chi &= \sum_{n=1}^p q_{1,n}(1+q_{4,n}^2) - \frac{\{\sum_{n=1}^p q_{1,n}^2(1+2q_{4,n}^2)\}^2}{\sum_{n=1}^p q_{1,n}^3(1+3q_{4,n}^2)} && (4.15)
\end{aligned}
$$

and

$$d_\chi = \frac{\{\sum_{n=1}^p q_{1,n}^2(1+2q_{2,n}^2)\}^3}{\{\sum_{n=1}^p q_{1,n}^3(1+3q_{2,n}^2)\}^2}. \tag{4.16}$$

The first approach requires the numerical evaluation of some complex integral. We regard this as infeasible for standard empirical work. Therefore, we only use the numerical approach to validate the approximation based on the gamma distribution.

For corollary 4 below, we define

$$w_1 := \frac{1}{a_\chi}\left[\frac{\alpha_g - 1}{f(z_\alpha)} - \frac{1}{2}\right] \tag{4.17}$$

and

$$w_2 := \frac{1}{a_\chi^2}\left[\frac{(\alpha_g-1)(\alpha_g-2)}{f(z_\alpha)^2} - \frac{\alpha_g - 1}{f(z_\alpha)} + \frac{1}{4}\right]. \tag{4.18}$$

Corollary 4. The coefficients (4.11) and (4.12) based on the gamma approximation (4.13) are given by

$$k_1 = c_1 z_\alpha - w_1 c_2 z_\alpha^2$$

and

$$
\begin{aligned}
k_2 &= (c_3 + c_1^2) z_\alpha + (c_1^2 - c_4 - 3c_1c_2) w_1 z_\alpha^2 \\
&\quad + 2c_2^2 w_1^2 z_\alpha^3 + \left[w_1 w_2 - \frac{1}{2} w_1^3\right] c_2^2 z_\alpha^4.
\end{aligned}
$$

Table 4.1 gives the values of k_1 and k_2 for several kernels. The coefficients are relatively stable for different values of p.

Table 4.1: Coefficients of corrected critical values using the approximation of Zhang (2005)

		Bartlett		Parzen		QS	
α	p	k_1	k_2	k_1	k_2	k_1	k_2
	50	0.874	1.395	0.686	0.907	1.224	3.010
0.1	100	0.874	1.394	0.686	0.906	1.224	3.009
	1000	0.874	1.394	0.687	0.906	1.225	3.008
	50	1.339	2.080	1.056	1.356	1.894	4.507
0.05	100	1.340	2.078	1.057	1.355	1.894	4.505
	1000	1.340	2.077	1.057	1.355	1.895	4.502

We can obtain the corrected critical values derived in Sun, Phillips and Jin (2008)[1] from Corollary 4. In that case we have $p = 1$ and $c_{1,n} = 1$. Hence, the coefficients of the gamma approximation are $a_\chi = 1$, $b_\chi = 0$ and $d_\chi = 1$. In other words, the approximation is exact. Furthermore, from equations (4.17) and (4.18) we have that $w_1 = -1/2z_\alpha^{-1} - 1/2$ and $w_2 = 3/4z_\alpha^{-2} + 1/2z_\alpha^{-1} + 1/4$. Substituting in Corollary 4 gives

$$\begin{aligned} k_1 &= \left(c_1 + \frac{1}{2}c_2\right) z_\alpha + \frac{1}{2}c_2 z_\alpha^2, \\ k_2 &= \left(\frac{1}{2}c_1^2 + \frac{3}{2}c_1c_2 + \frac{3}{16}c_2^2 + c_3 + \frac{1}{2}c_4\right) \\ &\quad + \left(-\frac{1}{2}c_1^2 + \frac{3}{2}c_1c_2 + \frac{9}{16}c_2^2 + \frac{1}{2}c_4\right) z_\alpha^2 + \frac{5}{16}c_2^2 z_\alpha^3 - \frac{1}{16}c_2^2 z_\alpha^4. \end{aligned}$$

Finally, to facilitate the derivation of the optimal bandwidth below, we write the expansion of Theorem 2 also in terms of the gamma distribution.

$$F_\delta(z) = A_{1,\delta} + A_{2,\delta}b + A_{3,\delta}b^2 + o(b^2), \tag{4.19}$$

[1]Note that z_α^2 in Sun, Phillips and Jin (2008) corresponds to z_α in this chapter. Furthermore, we found a small error in their Theorem 1 and Corollary 2. To correct their results it is sufficient to replace c_4 by $2c_4$ and to replace c_1 by c_1^2 in the coefficient of z_α^4. This has been confirmed by one of the authors.

where

$$\begin{aligned} A_{1,\delta} &= \tilde{G}_\delta(f(z-q_p)), \\ A_{2,\delta} &= c_2\tilde{G}''_\delta(f(z-q_p))(f'(z-q_p))^2z^2 - c_1\tilde{G}'_\delta(f(z-q_p))f'(z-q_p)z \end{aligned}$$

and

$$\begin{aligned} A_{3,\delta} &= -\Bigg[\tilde{G}'_\delta(f(z-q_p))f'(z-q_p)zc_3 - \frac{1}{2}\tilde{G}''_\delta(f(z-q_p))(f'(z-q_p))^2z^2(2c_4-c_1^2) \\ &\quad + \tilde{G}'''_\delta(f(z-q_p))[f'(z-q_p)]^3z^3c_1c_2\Bigg]. \end{aligned}$$

Here we used that $f''(\cdot)$ and $f'''(\cdot)$ are zero.

4.4.2 Expansion of the finite sample distribution

In this section we develop an expansion of the finite sample distribution $F_{T,\delta}(z) := P\left(\eta_T(\hat{\sigma}_b^2) \le z\right)$. Define

$$e_t = C^{(t)}(\tau_1,\tau_2) - I_t \tag{4.20}$$

and

$$d_t = \sum_{j=1}^{t}[C_T(\tau_1,\tau_2) - C^{(j)}(\tau_1,\tau_2)]. \tag{4.21}$$

Then after some algebraic calculations which are given in the proof of Lemma 5 below:

$$\sum_{t=1}^{T}\left(\sum_{j=1}^{t}(C_T - I_j)\right)^2 = e'Ve + 2e'd^\dagger + c_d, \tag{4.22}$$

with $e = (e_1,\ldots,e_T)'$, $V = \{T+1-\max(s,t)\}_{s,t=1,\ldots,T}$, $d^\dagger = \left(\sum_{j=1}^{T} d_j, \sum_{j=2}^{T} d_j, \ldots, d_T\right)'$ and $c_d = \sum_{t=1}^{T} d_t^2$.

We make the following assumption (see Sun, Phillips and Jin (2008)):

Assumption 2. e_t is a mean zero covariance-stationary ARMA process with NID innovations η_t and $\sum_{h=-\infty}^{\infty} h^2|\Gamma_h| < \infty$, where $\Gamma_h = Ee_te_{t-h}$.

The ARMA assumption allows us to rewrite the long run variance of e_t in terms of the long run variance of η_t: $\sigma_e^2 = g(\theta)\sigma_\eta^2$ were $g(\theta)$ is a function of the ARMA parameter vector θ. For example, in case of an $AR(p)$ specification we have $g(\theta) := 1/(1-\sum_{i=1}^{p}\theta_i)^2$, see e.g. Sul, Phillips and Choi (2005).

Let $\sigma_{e,T}^2 = \mathrm{Var}(T^{-1/2}\sum_{t=1}^{T} e_t)$ and let $\sigma_{\eta,T}^2$ denotes variance of the error terms. Note that $\sigma_{e,T}^2$ and $g(\theta)\sigma_{\eta,T}^2$ are asymptotically the same but might be different in finite samples. For example, in the case of an AR(1) model we have for $n = 100$, and $\theta = 0.0, 0.3$ and 0.7, that the ratio $\sigma_{e,T}^2/(g(\theta)\sigma_{\eta,T}^2)$ is respectively 1.0, 0.99 and 0.97.

Define $\Gamma_0^{1/2} = QD^{1/2}Q'$ where D denotes the diagonal matrix of eigenvalues of Γ_0 and Q is the matrix of corresponding eigenvectors. Then

$$\begin{aligned} e'Ve &= e'\Gamma_0^{-1/2}\Gamma_0^{1/2}V\Gamma_0^{1/2}\Gamma_0^{-1/2}e \\ &= e'\Gamma_0^{-1/2}PP'\Gamma_0^{1/2}V\Gamma_0^{1/2}PP'\Gamma_0^{-1/2}e \\ &= \tilde{e}'\Lambda\tilde{e}, \end{aligned} \tag{4.23}$$

where $\tilde{e} = P'\Gamma_0^{-1/2}e$, P an orthogonal matrix of eigenvectors of $\Gamma_0^{1/2}V\Gamma_0^{1/2}$ such that $P'\Gamma_0^{1/2}V\Gamma_0^{1/2}P = \Lambda$ and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_T)$ with $\lambda_1, \ldots, \lambda_T$ the corresponding eigenvalues (see Ullah (2004, p.53)).

The following lemma follows from (4.22) and (4.23):

Lemma 5. The finite sample distribution can be written as

$$F_{T,\delta}(z) = P\left(\sum_{t=1}^{T}\tilde{\lambda}_t(\tilde{e}_t + \tilde{d}_t/\tilde{\lambda}_t)^2 \le \sum_{t=1}^{T}\tilde{d}_t^2/\tilde{\lambda}_t - c_d\right), \tag{4.24}$$

where $\tilde{\lambda}_t = \lambda_t - Tz\varsigma_{bT}\tilde{g}(\theta)$, with $\varsigma_{bT} = \hat{\sigma}_b^2/\sigma_{e,T}^2$, $\tilde{d} = (\tilde{d}_1, \ldots, \tilde{d}_T)' = P'\Gamma_0^{1/2}d^\dagger$ and $\tilde{g}(\theta) = g(\theta)\cdot\sigma_{e,T}^2/(g(\theta)\sigma_{\eta,T}^2) = \sigma_{e,T}^2/\sigma_{\eta,T}^2$.

As in the previous section, we will approximate the distribution of the weighted sum of χ_1^2 distributed variables using the approach of Zhang (2005). Let

$$\tilde{H}_{T,\delta}(z, \varsigma_{bT}) := \tilde{G}_{T,\delta}(f_T(z, \varsigma_{bT}); \alpha_{g,T}(\varsigma_{bT})) := P(\chi^2_{d_T(\varsigma_{bT})} \le f_T(\varsigma_{bT}))$$

be the gamma approximation of the true distribution with parameters $\alpha_{g,T}(\varsigma_{bT}) =$

$d_T(\varsigma_{bT})/2$ and $1/2$, and let

$$f_T(z, \varsigma_{bT}) = \left(\sum_{t=1}^{T} \tilde{d}_t^2 / \tilde{\lambda}_t(\varsigma_{bT}) - c_d - b_T(\varsigma_{bT}) \right) / a_T(\varsigma_{bT}).$$

Here, $a_T(\varsigma_{bT})$, $b_T(\varsigma_{bT})$ and $d_T(\varsigma_{bT})$ are defined as in (4.14), (4.15) and (4.16) but with weights and noncentrality parameter as in (4.24).

For Theorem 6 below, let q be the Parzen characteristic exponent defined by

$$q = \max \left\{ q_0 : \; q_0 \in \mathbb{Z}^+, \; g_{q_0} = \lim_{x \to 0} \frac{1 - k(x)}{|x|^{q_0}} < \infty \right\}$$

and

$$d_{qT} \;\; = \;\; \sigma_{e,T}^{-2} \sum_{h=-\infty}^{\infty} |h|^q E(e_t e_{t-h}).$$

For the Bartlett, Parzen and QS kernel we have $(q, g_q) = (1,1), (2,6)$ and $(2,1.421)$, respectively (see Andrews (1991)). The value of d_{qT} can be obtained using a plug-in estimate. For example, if e_t follows an AR(1) process $e_t = \phi e_{t-1} + \eta_t$ with $\eta_t \sim NID(0, \sigma_\eta^2)$, then (as $T \to \infty$) $d_{qT} = 2\phi/(1-\phi^2)$ if $q = 1$ and $d_{qT} = 2\phi/(1-\phi)^2$ if $q = 2$. If e_t follows an MA(1) process $e_t = \eta_t + \theta \eta_{t-1}$ then $d_{qT} = 2\theta/(1+\theta)^2$ for $q = 1, 2$. In the expressions for d_{qT} we substitute the estimates for ϕ and θ.

Let $\tilde{H}'_{T,\delta}$ and $\tilde{H}''_{T,\delta}$ be the first and second derivative of $\tilde{H}_{T,\delta}(z, \varsigma_{bT})$ with respect to ς_{bT}. Theorem 6 gives the expansion of $F_{T,\delta}(z)$ around $\tilde{H}_{T,\delta}(\cdot,\cdot)$.

Theorem 6.

$$F_{T,\delta}(z) = B_{1,\delta} + B_{2,\delta} b - g_q d_{qT} B_{3,\delta} (bT)^{-q} + o(b + (bT)^{-q}) + O(T^{-1}), \tag{4.25}$$

where

$$\begin{aligned} B_{1,\delta} &= \tilde{H}_{T,\delta}(z; 1), \\ B_{2,\delta} &= c_2 \tilde{H}''_{T,\delta}(z, \mu_b) - c_1 \tilde{H}'_{T,\delta}(z, 1), \\ B_{3,\delta} &= \tilde{H}'_{T,\delta}(z; 1). \end{aligned}$$

Given the expansion of the limit distribution (4.19) and the expansion of the finite sample distribution (4.25), we can derive the size distortion and the power of the test against the local alternative (4.1).

Corollary 7.

(i) The size distortion of the test based on second order critical values is given by

$$\begin{aligned} 1 - F_{T,0}(z_{\alpha,b}) - \alpha &= -(B_{1,0} - A_{1,0}) - (B_{2,0} - A_{2,0})b + g_q d_{qT} B_{3,0}(bT)^{-q} \\ &\quad + o(b + (bT)^{-q}) + O(T^{-1}). \end{aligned}$$

(ii) The power of the test based on second order critical values is given by

$$1 - F_{T,\delta}(z_{\alpha,b}) = 1 - B_{1,\delta} - B_{2,\delta}b + g_q d_{qT} B_{3,\delta}(bT)^{-q} + o(b + (bT)^{-q}) + O(T^{-1}).$$

4.4.3 Optimal bandwidth rule

We follow Sun, Phillips and Jin (2008, 2010) to find the optimal parameter b by balancing the type I and type II error probabilities. The analytical expressions for the type I and II errors follow from the size distortion and the power.

The type I error is given by

$$e_T^I = \alpha - (B_{1,0} - A_{1,0}) - (B_{2,0} - A_{2,0})b + g_q d_{qT} B_{3,0}(bT)^{-q} \tag{4.26}$$

and the type II error is given by

$$e_T^{II} = B_{1,\delta} + B_{2,\delta}b - g_q d_{qT} B_{3,\delta}(bT)^{-q}. \tag{4.27}$$

We define the loss function

$$L(b; \delta, T, z_\alpha) = \frac{w_T(\delta)}{1 + w_T(\delta)} e_T^I + \frac{1}{1 + w_T(\delta)} e_T^{II}, \tag{4.28}$$

where $w_T(\delta)$ is a function that determines the relative weight on the type I and type II error probabilities.

We consider a single alternative δ and write $w_T = w_T(\delta)$. Assume that $d_{qT} > 0$, which is true if the series exhibit positive serial correlation. Minimizing the resulting loss function with respect to b gives the optimal bandwidth rule

$$b = \left\{ \frac{q g_q d_{qT}(w_T B_{3,0} - B_{3,\delta})}{B_{2,\delta} - w_T(B_{2,0} - A_{2,0})} \right\}^{\frac{1}{q+1}} T^{-q/(q+1)}. \tag{4.29}$$

Note that the bandwidth rule goes to zero, as $T \to \infty$, at the same rate as the bandwidth rule developed in Sun, Phillips and Jin (2008). This is due to the fact that we used the same order for the expansion of the limit and finite sample distributions as in their paper. The major difference is that, in our case, the scaling factor depends explicitly on the difference between the coefficients A_i and B_i of the expansion of the limit and finite sample distribution.

Before we examine the finite sample properties of the resulting bandwidth rule, we have to make a few remarks. First, the optimal bandwidth rule (4.29) does not exist for each weight w_T. In particular, we found that the term in curly brackets might become negative. In the simulation study below we set the bandwidth equal to zero if this happens.

Second, the second-order approximations of the finite-sample and limit distribution do not fully reflect the true distributions (see the discussion below). This results in some counterintuitive behaviour for the type I and II error, and the weights. In particular, for weights $w_T \geq 10$ we observe that higher weights result in smaller values for b. Following, Sun, Phillips and Jin (2008) we will use weights $w_T \in [10, 40]$ in our empirical example. For these weights the decline in b turns out to be small.

Finally, in the next section we also show that the loss function (4.28) has the disadvantage that it might be optimal to choose the conventional bandwidth γ_T sufficiently large. In that case the power of the test is low. To correct for this behavior we might consider the sum of squares of the errors as an alternative loss function

$$L'(b; \delta, T, z_\alpha) = \frac{w_T(\delta)}{1 + w_T(\delta)}[e_T^I]^2 + \frac{1}{1 + w_T(\delta)}[e_T^{II}]^2. \tag{4.30}$$

We also tried to minimize this loss function with respect to b. The first order conditions are a third order polynomial for which we don't have an explicit solution. In the next section we will solely use this loss function to compare fixed-b asymptotic with the conventional approach.

4.5 Finite sample properties

In this section we examine the size and the power of the copula constancy tests for finite samples of size $T = 100$. We simulate observations from a Clayton copula and subsequently transform the values using the inverse normal distribution function. Let ε_{it} denote the resulting series. The observations are constructed as $y_{it} = \rho_i y_{it-1} + \varepsilon_{it}$. The HAC estimator is based on the Bartlett kernel. Following Amsler et al. (2009), we consider bandwidth rules $\gamma_T(m) = integer[m(T/100)^{1/4}]$, $m = 1, 2, \ldots$. The conventional bandwidth rules analysed by Amsler et al. (2009) are $m = 0, 4, 12, 25$ and 50.

Figure 4.1 and 4.2 show the size and power of the test. In line with the results for the KPSS test, the use of fixed-b critical values in combination with sufficiently high bandwidth values results in a clear reduction of the size distortion. We do not observe this behaviour for the standard critical values. However, selecting higher bandwidth values reduces the power of the test.

Amsler et al. (2009) select a conventional bandwith rule $m \in \{0, 4, 12, 25, 50\}$ such that the size distortion of the test, using standard as well as fixed-b critical value, falls below some threshold (say 0.1). Given the selected bandwidth rule, it is clear from the lower panels in figure 4.1 and 4.2 that the power of the test under fixed-b critical values is higher.

This approach of examining the test under fixed-b critical values versus conventional critical values is somewhat questionable if there is serial correlation in the data. From figure 4.1 it is clear that for fixed value m, the power as well as the size distortion of the test are higher if we use fixed-b critical values and select the bandwidth sufficiently small. Hence, to obtain the same level of size distortion we have to select higher bandwidth values for the test based on fixed-b critical values

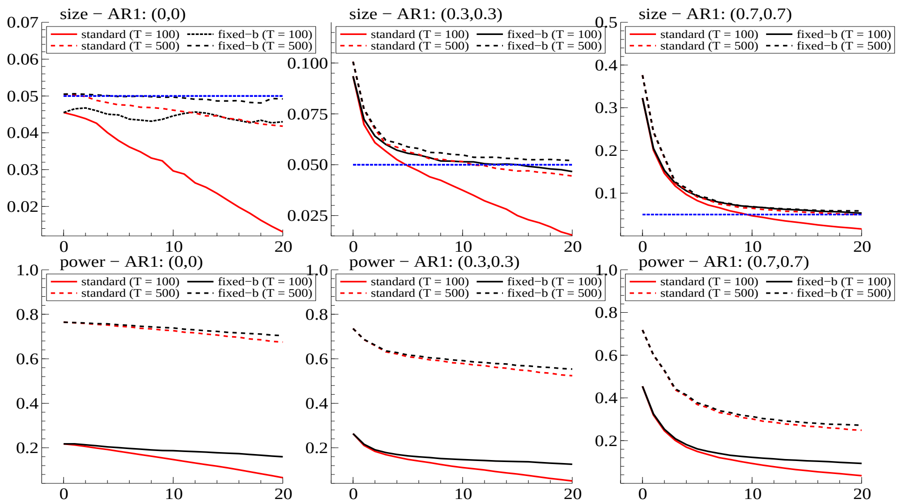

Figure 4.1: Size and power of the copula constancy test using standard and fixed-b critical values. The horizontal axis shows m for bandwith rule $\gamma_T(m) = integer[m(T/100)^{1/4}]$ and the vertical axis shows the rejection frequency. DGP: Clayton copula with Kendall's tau 0.25 under H_0 and 0.75 in second half of the sample under H_1. Bartlett kernel, #rep = 20000.

compared to the test based on standard critical values. Comparison based on single values $m \in \{0, 4, 12, 25, 50\}$ seems therefore unreasonable. Furthermore, it is not necessary to restrict m to be in the subset $\{0, 4, 12, 25, 50\}$ and, therefore, the question arises if there exists a value of m which is optimal in some sense. Moreover, we can ask if bandwidth rules of the form $\gamma_T(m) = integer[m(T/100)^{1/4}]$ are optimal at all.

The finite sample evidence presented above clearly emphasizes the need for the optimal bandwidth rules described in section 4.4. To examine the performance of the fixed-b critical values in combination with the optimal bandwidth rule we simulate the average loss (4.28) and compare it to the loss based on standard critical values and conventional bandwidth rules.

Table 4.2 shows the simulated loss. The loss using fixed-b critical values is generally higher than the loss using standard critical values. Furthermore, in case of fixed-b critical values we only select values $m > 0$ if the weight assigned to the type I error is sufficiently high. The new bandwidth rule performs reasonable compared

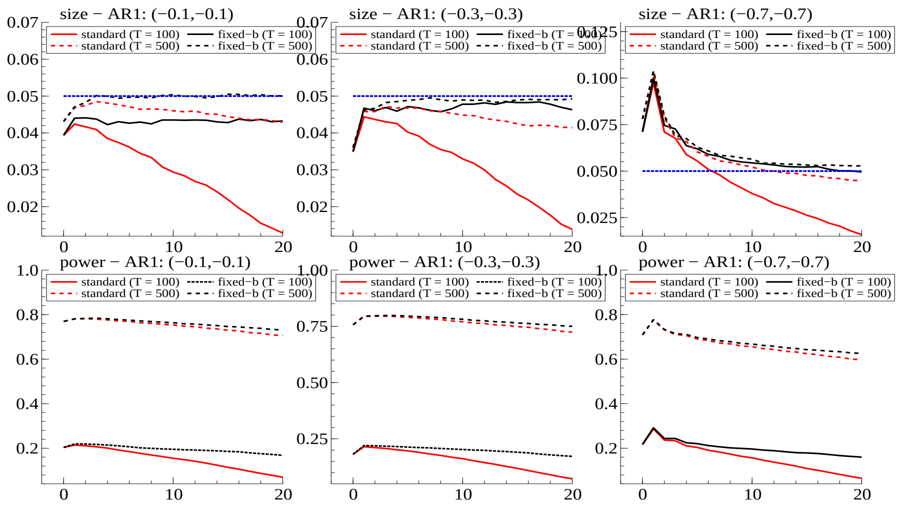

Figure 4.2: Size and power of the copula constancy test using standard and fixed-b critical values. The horizontal axis shows m for bandwith rule $\gamma_T(m) = integer[m(T/100)^{1/4}]$ and the vertical axis shows the rejection frequency. DGP: Clayton copula with Kendall's tau 0.25 under H_0 and 0.75 in second half of the sample under H_1. Bartlett kernel, #rep = 20000.

to the loss based on fixed-b critical values. It does, however, not always attain the minimum value and it does not always outperform the conventional bandwidth rules. We emphasize that in practice it might be unclear which value of m to choose and selecting m too small results in higher loss values.

Table 4.2: Loss (4.28) based on 5000 replications and T = 100. Clayton copula; Bartlett kernel; $\tau_1 = \tau_2 = 0.25$.

w	ρ_1,ρ_2	$\gamma_T(0)$	$\gamma_T(4)$	$\gamma_T(12)$	$\gamma_T(25)$	b(0)	b(4)	b(12)	b(25)	b^*
	0.0	0.114	0.109	0.105	0.093	0.114	0.112	0.116	0.120	0.109
10	0.3	0.148	0.120	0.109	0.093	0.148	0.123	0.122	0.124	0.122
	0.7	0.342	0.163	0.117	0.095	0.342	0.169	0.135	0.132	0.136
	0.0	0.066	0.058	0.049	0.029	0.066	0.062	0.066	0.067	0.057
40	0.3	0.105	0.067	0.052	0.028	0.105	0.071	0.069	0.068	0.069
	0.7	0.327	0.115	0.059	0.030	0.327	0.122	0.080	0.075	0.080

The fact that for conventional bandwidth rules it becomes optimal to select high bandwidth values is somewhat questionable. To address this issue we consider the weighted sum of squared errors as loss criterion. Table 4.3 shows the results. The loss obtained using fixed-b critical values is now somewhat lower compared to the loss base on conventional critical values.

Table 4.3: Loss (4.30) based on 5000 replications. Clayton copula; Bartlett kernel; $\tau_1 = \tau_2 = 0.25$.

w	ρ_1, ρ_2	$\gamma_T(0)$	$\gamma_T(4)$	$\gamma_T(12)$	$\gamma_T(25)$	b(0)	b(4)	b(12)	b(25)
	0.0	0.057	0.060	0.068	0.086	0.057	0.059	0.061	0.068
10	0.3	0.057	0.066	0.074	0.087	0.057	0.065	0.068	0.073
	0.7	0.121	0.070	0.078	0.089	0.121	0.070	0.074	0.079
	0.0	0.017	0.017	0.019	0.023	0.017	0.017	0.018	0.020
40	0.3	0.021	0.019	0.020	0.023	0.021	0.019	0.020	0.021
	0.7	0.108	0.026	0.022	0.024	0.108	0.027	0.023	0.023

Based on the simulation evidence provided in Table 4.2 and 4.3 we conclude that the use of fixed-b critical values is useful if we consider a loss criterion based on the sum of squared errors. In case of a simple sum, the optimal bandwidth rule performs reasonable compared to the standard fixed-b results. It does, however, not always outperform the conventional bandwidth rules.

To examine the sensitivity of our results, we also performed the previous simulation using a QS kernel instead of a Bartlett kernel. Table 4.4 provides the results. The results are in line with the ones obtained using a Bartlett kernel; the optimal bandwidth rule performs reasonable compared to the loss based on fixed-b critical values but does not always outperform the loss based on conventional bandwidth rules.

To examine the sensitivity of our results with respect to the copula, we replace the Clayton copula by a Gaussian copula. In line with previous simulations we consider jumps in Kendall's tau from 0.25 to 0.75. Table 4.5 provides the results. The results are similar to ones for the Clayton copula.

Table 4.4: Loss (4.28) based on 5000 replications. Clayton copula; QS kernel; $\tau_1 = \tau_2 = 0.25$.

w	ρ_1,ρ_2	$\gamma_T(0)$	$\gamma_T(4)$	$\gamma_T(12)$	$\gamma_T(25)$	b(0)	b(4)	b(12)	b(25)	b^*
	0.0	0.114	0.111	0.109	0.129	0.114	0.113	0.118	0.133	0.108
10	0.3	0.148	0.120	0.113	0.117	0.148	0.122	0.123	0.123	0.121
	0.7	0.342	0.155	0.112	0.106	0.342	0.159	0.120	0.111	0.128
	0.0	0.066	0.060	0.054	0.068	0.066	0.063	0.066	0.073	0.056
40	0.3	0.105	0.066	0.055	0.055	0.105	0.069	0.068	0.062	0.068
	0.7	0.327	0.105	0.052	0.042	0.327	0.109	0.062	0.047	0.070

Table 4.5: Loss (4.28) based on 5000 replications. Gaussian copula; Bartlett kernel; $\tau_1 = \tau_2 = 0.25$.

w	ρ_1,ρ_2	$\gamma_T(0)$	$\gamma_T(4)$	$\gamma_T(12)$	$\gamma_T(25)$	b(0)	b(4)	b(12)	b(25)	b^*
	0.0	0.115	0.112	0.104	0.093	0.115	0.113	0.115	0.120	0.109
10	0.3	0.148	0.122	0.111	0.095	0.148	0.126	0.121	0.119	0.122
	0.7	0.330	0.158	0.116	0.097	0.330	0.164	0.135	0.125	0.131
	0.0	0.066	0.061	0.048	0.029	0.066	0.063	0.064	0.066	0.058
40	0.3	0.106	0.070	0.054	0.031	0.106	0.075	0.068	0.063	0.070
	0.7	0.315	0.110	0.058	0.032	0.315	0.116	0.081	0.067	0.075

4.6 Empirical application

To illustrate the use of fixed-b critical values and the new bandwith rule, we consider MSCI stock returns from the US, UK, France, Germany and Japan. The dataset 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).

In chapter 3 (see also Van Kampen and Wied (2010)) we give a detailed analysis of the estimation of the marginal distributions and the application of the copula constancy test. In summary, we will use the standardized residuals from a GARCH (1,1) model with skewed-distributed innovations for the US, UK, France and Germany. For Japan we estimate a GARCH(1,1) with student-t distributed innovations.

We assume in the calculation of the optimal bandwidth rule that under the alternative hypothesis, Kendall's tau increases by 0.2. Subsequently, we transform Kendall's tau to the parameter of the Clayton copula. Under the null hypothesis, Kendall's tau corresponds to the empirical estimate of Kendall's tau. Note that the optimal bandwidth also depends on the confidence level α. In the analysis below we assume, as usual, that $\alpha = 0.05$. To examine the sensitivity of our results, we take weights $w = 10$ and $w = 40$.

Table 4.6 shows the test statistics where b is calculated using conventional bandwidth rules and using the optimal bandwidth rule. The values of b corresponding to $\gamma_T(4)$ and $\gamma_T(12)$ are 0.010 and 0.035, respectively. The optimal bandwidth values are zero or close to zero. This makes sense since the amount of serial correlation is quite low. Note also that the optimal bandwidth is rather insensitive to chosen values for w.

The results suggest that the bandwidth rule $\gamma_T(12)$ from the previous chapter, is relatively high. Although the final result are mainly in line with the results of chapter 3, we do however find some additional evidence that the copula is not constant over time; we also reject the null hypothesis for the country pair France-Germany.

To examine the sensitivity of our results, we also calculated the optimal bandwith rule for the alternative hypothesis that Kendall's tau increases by 0.4. This, however, did not change the results significantly.

4.7 Conclusion

This chapter introduces a new copula constancy test which is based on an inconsistent estimate of the long run variance. The resulting distribution is nonstandard and depends on the bandwidth parameter b. In the spirit of Sun, Phillips and Jin (2008), we derived an optimal bandwidth rule that minimizes a weighted average of the type I and type II error probabilities. Monte Carlo simulations suggest that the new copula constancy test improves the finite sample properties of the test if we consider a sum of squared errors as loss criterion. The new bandwidth rule performs reasonable but does not always attain the minimum loss value. We conjecture that this is due the

Table 4.6: Copula constancy test statistics using fixed-b critical values. $b_T(\cdot)$ refers to the test with b calculated using the convential bandwith rule $\gamma_T(\cdot)$ and $b^*_T(w = \cdot)$ refers to the test calculated via the optimal bandwidth rule using weight w. The value in brackets gives the optimal value of b and ** denotes that we are able to reject H_0 at the 5% level.

Country Pair	τ	$b_T(4)$	$b_T(12)$	$b^*_T(w = 10)$		$b^*_T(w = 40)$	
US - UK	0.10	0.380	0.322	0.339	[0.000]	0.342	[0.000]
	0.25	0.144	0.126	0.138		0.142	
	0.50	0.224	0.167	0.242		0.242	
US - France	0.10	0.490**	0.382	0.379	[0.000]	0.387	[0.000]
	0.25	0.211	0.204	0.223		0.223	
	0.50	0.188	0.149	0.150		0.156	
US - Germany	0.10	0.571**	0.408	0.418	[0.022]	0.426	[0.019]
	0.25	0.243	0.242	0.242		0.243	
	0.50	0.226	0.174	0.202		0.207	
US - Japan	0.10	0.176	0.165	0.165	[0.000]	0.165	[0.000]
	0.25	0.272	0.252	0.241		0.241	
	0.50	0.164	0.128	0.133		0.138	
UK - France	0.10	0.240	0.197	0.197	[0.023]	0.201	[0.020]
	0.25	0.077	0.070	0.072		0.074	
	0.50	0.568**	0.426	0.532**		0.543**	
UK - Germany	0.10	0.496**	0.391	0.406	[0.000]	0.420	[0.000]
	0.25	0.203	0.199	0.193		0.192	
	0.50	0.420	0.305	0.381		0.395	
UK - Japan	0.10	0.284	0.256	0.256	[0.028]	0.262	[0.025]
	0.25	0.404	0.325	0.342		0.360	
	0.50	0.523**	0.407	0.420		0.430	
France - Germany	0.10	0.687**	0.554	0.561**	[0.033]	0.569**	[0.029]
	0.25	0.231	0.266	0.258		0.253	
	0.50	0.302	0.241	0.248		0.268	
France - Japan	0.10	0.358	0.323	0.320	[0.021]	0.321	[0.018]
	0.25	0.316	0.308	0.314		0.311	
	0.50	0.183	0.163	0.162		0.161	
Germany - Japan	0.10	0.572**	0.510	0.507**	[0.017]	0.503**	[0.015]
	0.25	0.503**	0.449	0.515**		0.510**	
	0.50	0.221	0.202	0.199		0.201	

insufficient approximation of the finite sample and limit distributions (see also the remarks at the end of section 4.4.3).

A point for further research concerns the order of the expansions of the distribution functions. Comparing the approximations with the simulated values shows that the difference for the Bartlett kernel is large if $b \to 1$ (see Appendix 4.C). Although the limit theorems still hold for $b \to 0$, the second (and third) order critical values require improvement.

4.A Proofs

Proof of Lemma 1:
Under $H_1: C^{(t)}(\tau_1,\tau_2) = [1-h(t/T)T^{-1/2}]C(\tau_1,\tau_2) + h(t/T)T^{-1/2}C^*(\tau_1,\tau_2)$ we have

$$\eta_T(\hat{\sigma}_b^2) \xrightarrow{d} \int_0^1 [V(r) + f(r,\delta)]^2 dr \Xi_b^{-1}, \tag{4.31}$$

where

$$f(r,h) := \sigma_1^{-1}[C(\tau_1,\tau_2) - C^*(\tau_1,\tau_2)]\left(\int_0^r h(s)ds - r\int_0^1 h(s)ds\right). \tag{4.32}$$

When we have a single break at z^* of magnitude δ then

$$\int_0^r h(s)ds = \begin{cases} 0 & \text{if } r \leq z^* \\ \delta(r-z^*) & \text{it } r > z^* \end{cases} \qquad \text{and} \qquad r\int_0^1 h(s)ds = r(1-z^*)\delta.$$

Hence,

$$f(r,h) = f(r,\delta) = \begin{cases} \sigma_1^{-1}[C(\tau_1,\tau_2) - C^*(\tau_1,\tau_2)]\delta r(z^*-1) & \text{if } r \leq z^* \\ \sigma_1^{-1}[C(\tau_1,\tau_2) - C^*(\tau_1,\tau_2)]\delta z^*(r-1) & \text{if } r > z^*, \end{cases} \tag{4.33}$$

where we rewrite $f(r,h)$ as $f(r,\delta)$ to explicitly reflect the dependence on δ.

Substituting in (4.31) gives

$$\eta_T(\hat{\sigma}_b^2) \xrightarrow{d} \left[\int_0^{z^*} [V(r) + (1-1/z^*)ar]^2 dr + \int_{z^*}^1 [V(r) + ar - a]^2 dr\right]\Xi_b^{-1}, \tag{4.34}$$

where $a := \sigma_1^{-1}[C(\tau_1,\tau_2) - C^*(\tau_1,\tau_2)]\delta z^*$.

From Gikhman and Skorokhod (2004, p230) we have that

$$V(r) = \sqrt{2}\sum_{n=1}^{\infty} \zeta_n \frac{\sin n\pi r}{n\pi}, \tag{4.35}$$

where ζ_n is a sequence of independent standard normal distributed variables.

Using (4.35) it can be shown that

$$\begin{aligned}
\int_0^1 V(r)^2 dr &= \sum_{n=1}^{\infty} \zeta_n^2 (n\pi)^{-2} \qquad (4.36)\\
\int_x^y rV(r)dr &= \sqrt{2}\sum_{n=1}^{\infty} \zeta_n (n\pi)^{-2}\big\{-y\cos(n\pi y) + x\cos(n\pi x)\\
&\qquad + (n\pi)^{-1}\sin(n\pi y) - (n\pi)^{-1}\sin(n\pi x)\big\}\\
\int_x^y V(r)dr &= \sqrt{2}\sum_{n=1}^{\infty} \zeta_n (n\pi)^{-2}\big\{-\cos(n\pi y) + \cos(n\pi x)\big\},
\end{aligned}$$

where the formal proof of (4.36) can found in Kac and Siegert (1947).

Therefore,

$$\eta_T(\hat{\sigma}_b^2) \xrightarrow{d} \left(\sum_{n=1}^{\infty}[q_{1,n}\zeta_n^2 + q_{2,n}\zeta_n] + q_3\right)\Xi_b^{-1}, \qquad (4.37)$$

where $q_{1,n} = (\pi n)^{-2}$, $q_{2,n} := -2^{3/2}a(n\pi)^{-2}\sin(n\pi z^*)/(n\pi z^*)$ and $q_3 := a^2(\frac{1}{3}z^{*2} - \frac{2}{3}z^* + \frac{1}{3})$.

Under H_0 we have that $q_{1,n} = (\pi n)^{-2}$ and $q_{2,n} = q_3 = 0$.

Factor the polynomial as

$$q_{1,n}\zeta_n^2 + q_{2,n}\zeta_n = q_{1,n}(\zeta_n + q_{4,n})^2 - (q_{2,n}/(2q_{1,n}))^2, \qquad (4.38)$$

where $q_{4,n} = q_{2,n}/(2q_{1,n}) = -2^{1/2}a\sin(n\pi z^*)/(n\pi z^*)$. Then

$$\eta_T(\hat{\sigma}_b^2) \xrightarrow{d} \left(\sum_{n=1}^{\infty}[q_{1,n}(\zeta_n + q_{4,n})^2 + q_\infty]\right), \qquad (4.39)$$

where

$$q_\infty := \lim_{p\to\infty} q_p := \lim_{p\to\infty}\left\{q_3 - \sum_{n=1}^{p} q_{4,n}^2)\right\}. \qquad (4.40)$$

Proof of Theorem 2:

Below, we refer to Sun, Phillips and Jin (2008) as SPJ (2008).

Define $\mu_b := E(\Xi_b)$ and $\alpha_m := E(\Xi_b - \mu_b)^m$, $m = 1, 2, \ldots$. From SPJ (2008, eq. A.55, A.57, A.58) we have

$$\mu_b = 1 - \int_0^1\int_0^1 k_b(r-s)drds \qquad (4.41)$$

and

$$\begin{aligned}\alpha_2 &= 2\Big(\int_0^1 k_b(r-s)drds\Big)2 + 2\int_0^1\int_0^1 k_b^2(r-s)drds \\ &\quad - 4\int_0^1\int_0^1\int_0^1 k_b(r-p)k_b(r-q)drdpdq.\end{aligned}$$

Furthermore, SPJ (2008, eq. A.64, A.65 and A.67) show that

$$\begin{aligned}\int_0^1\int_0^1 k_b(r-s)drds &= bc_1 + b^2c_3 + o(b^2) \qquad (4.42)\\ \int_0^1\int_0^1 k_b^2(r-s)drds &= bc_2 + b^2c_4 + o(b^2)\\ \int_0^1\int_0^1\int_0^1 k_b(r-p)k_b(r-q)drdpdq &= c_1^2b^2.\end{aligned}$$

Hence,

$$\alpha_2 = 2bc_2 + 2b^2(c_4 - c_1^2) + o(b^2). \qquad (4.43)$$

Note that the result for α_2, stated in SPJ (2008,eq. A.69) is incorrect.

Also (see SPJ (2008, eq.A.55))

$$\begin{aligned}\alpha_3 &= o(b^2), \qquad (4.44)\\ \alpha_4 &= O(b^3). \qquad (4.45)\end{aligned}$$

Hence,

$$\begin{aligned}F_\delta(z) &:= P\left(\left(\sum_{n=1}^p[q_{1,n}(\zeta_n+q_{4,n})^2]+q_p\right)\Xi_b^{-1}<z\right)\\ &= E\{G_\delta(z\Xi_b-q_p)\}\\ &= E\Big\{G_\delta(z\mu_b-q_p)+G'_\delta(z\mu_b-q_p)z(\Xi_b-\mu_b)+\frac{1}{2}G''_\delta(z\mu_b-q_p)z^2(\Xi_b-\mu_b)^2\\ &\quad+\frac{1}{6}G'''_\delta(z\mu_b-q_p)z^6(\Xi_b-\mu_b)^3+\frac{1}{24}G_\delta^{(4)}(z\mu_b^*-q_p)z^8(\Xi_b-\mu_b)^4\Big\},\end{aligned}$$

with μ_b^* between Ξ_b and μ_b.

Using (4.44) and (4.45)

$$
\begin{aligned}
F_\delta(z) &= G_\delta(z\mu_b - q_p) + \frac{1}{2}G''_\delta(z\mu_b - q_p)z^2\alpha_2 + o(b^2) \\
&= G_\delta(z - q_p) + G'_\delta(z - q_p)z(\mu_b - 1) + \frac{1}{2}G''_\delta(z - q_p)z^2(\mu_b - 1)^2 \\
&\quad + \frac{1}{6}G'''_\delta(z\mu_b^{**} - q_p)z^3(\mu_b - 1)^3 + \frac{1}{2}\big\{G''_\delta(z - q_p)z^2\alpha_2 \\
&\quad + G'''_\delta(z - q_p)(\mu_b - 1)z^3\alpha_2 + \frac{1}{2}G^{(4)}_\delta(z\mu_b^{***} - q_p)(\mu_b - 1)^2 z^4\alpha_2\big\} + o(b^2),
\end{aligned}
$$

with μ_b^{**} and μ_b^{***} between μ_b and 1.

Finally, using (4.41) and (4.43)

$$
\begin{aligned}
F_\delta(z) &= G_\delta(z - q_p) + G'_\delta(z - q_p)z(-bc_1 - b^2c_3) + \frac{1}{2}G''_\delta(z - q_p)z^2(-bc_1)^2 \\
&\quad + \frac{1}{2}\big\{G''_\delta(z - q_p)z^2(2bc_3 + 2b^2(c_4 - c_1^2)) + G'''_\delta(z - q_p)z^3(-2b^2c_1c_2)\big\} + o(b^2) \\
&= G_\delta(z - q_p) + [c_2G''_\delta(z - q_p)z^2 - c_1G'_\delta(z - q_p)z]b \\
&\quad - [G'_\delta(z - q_p)zc_3 - \frac{1}{2}G''_\delta(z - q_p)z^2(2c_4 - c_1^2) + G'''_\delta(z - q_p)z^3c_1c_2]b^2 + o(b^2).
\end{aligned}
$$

Proof of Corollary 3:

This follows immediately from results already established in Sun, Phillips and Jin (2008) and the correction stated above; proof included to make the document self-contained.

Under H_0, we obtain

$$
\begin{aligned}
F_0(z_{\alpha,b}) &= D(z_{\alpha,b}) + [-D'(z_{\alpha,b})z_{\alpha,b}c_1 + D''(z_{\alpha,b})z_{\alpha,b}^2c_2]b \\
&\quad + [-D'(z_{\alpha,b})z_{\alpha,b}c_3 + \frac{1}{2}D''(z_{\alpha,b})z_{\alpha,b}^2(2c_4 - c_1^2) - D'''(z_{\alpha,b})z_{\alpha,b}^3c_1c_2]b^2 + o(b^2) \\
&= D(z_\alpha) + D'(z_\alpha)(z_{\alpha,b} - z_\alpha) + \frac{1}{2}D''(z_\alpha)(z_{\alpha,b} - z_\alpha)^2 \\
&\quad + [-D'(z_\alpha)z_\alpha c_1 + D''(z_\alpha)z_\alpha^2c_2]b \\
&\quad + [-D''(z_\alpha)z_\alpha c_1(z_{\alpha,b} - z_\alpha) - D'(z_\alpha)c_1(z_{\alpha,b} - z_\alpha) \\
&\qquad + D'''(z_\alpha)z_\alpha^2c_2(z_{\alpha,b} - z_\alpha) + 2D''(z_\alpha)z_\alpha c_2(z_{\alpha,b} - z_\alpha)]b \\
&\quad + [-D'(z_\alpha)z_\alpha c_3 + \frac{1}{2}D''(z_\alpha)z_\alpha^2(2c_4 - c_1^2) - D'''(z_\alpha)z_\alpha^3c_1c_2]b^2 + o(b^2), \quad (4.46)
\end{aligned}
$$

where we anticipated that $z_{\alpha,b} = z_\alpha + k_1 b + k_2 b^2 + o(b^2)$, i.e. higher order terms in (4.46) are $o(b^2)$.

Using $D(z_\alpha) = 1-\alpha$ and $F_0(z_{\alpha,b}) = 1-\alpha$

$$\begin{aligned}
0 &= D'(z_\alpha)(z_{\alpha,b}-z_\alpha) + \frac{1}{2}D''(z_\alpha)(z_{\alpha,b}-z_\alpha)^2 \\
&\quad + [-D'(z_\alpha)z_\alpha c_1 + D''(z_\alpha)z_\alpha^2 c_2]b + [-D''(z_\alpha)z_\alpha c_1 - D'(z_\alpha)c_1 \\
&\qquad + D'''(z_\alpha)z_\alpha^2 c_2 + 2D''(z_\alpha)z_\alpha c_2](z_{\alpha,b}-z_\alpha)b \\
&\quad + [-D'(z_\alpha)z_\alpha c_3 + \frac{1}{2}D''(z_\alpha)z_\alpha^2(2c_4 - c_1^2) - D'''(z_\alpha)z_\alpha^3 c_1 c_2]b^2 + o(b^2) \\
&= D'(z_\alpha)k_1 b + \frac{1}{2}D''(z_\alpha)k_1^2 b^2 + D'(z_\alpha)k_2 b^2 \\
&\quad + [-D'(z_\alpha)z_\alpha c_1 + D''(z_\alpha)z_\alpha^2 c_2]b \\
&\quad + [-D''(z_\alpha)z_\alpha c_1 - D'(z_\alpha)c_1 + D'''(z_\alpha)z_\alpha^2 c_2 + 2D''(z_\alpha)z_\alpha c_2]k_1 b^2 \\
&\quad + [-D'(z_\alpha)z_\alpha c_3 + \frac{1}{2}D''(z_\alpha)z_\alpha^2(2c_4 - c_1^2) - D'''(z_\alpha)z_\alpha^3 c_1 c_2]b^2 + o(b^2).
\end{aligned}$$

Solve for k_1 with $k_2 = 0$ (i.e. obtain second order critical values)

$$\begin{aligned}
& 0 = D'(z_\alpha)k_1 b + [-D'(z_\alpha)z_\alpha c_1 + D''(z_\alpha)z_\alpha^2 c_2]b \\
\Leftrightarrow \quad & k_1 = -\frac{1}{D'(z_\alpha)}[D''(z_\alpha)z_\alpha^2 c_2 - D'(z_\alpha)z_\alpha c_1]. \qquad (4.47)
\end{aligned}$$

Solve for k_2 with k_1 given in (4.47) already correcting all terms linear in b

$$\begin{aligned}
0 &= \frac{1}{2}D''(z_\alpha)k_1^2 b^2 + D'(z_\alpha)k_2 b^2 \\
&\quad + [-D''(z_\alpha)z_\alpha c_1 - D'(z_\alpha)c_1 + D'''(z_\alpha)z_\alpha^2 c_2 + 2D''(z_\alpha)z_\alpha c_2]k_1 b^2 \\
&\quad + [-D'(z_\alpha)z_\alpha c_3 + \frac{1}{2}D''(z_\alpha)z_\alpha^2(2c_4 - c_1^2) - D'''(z_\alpha)z_\alpha^3 c_1 c_2]b^2 \\
\Leftrightarrow \quad k_2 &= -\frac{1}{D'(z_\alpha)}\Bigg[- D'(z_\alpha)z_\alpha c_3 + \frac{1}{2}D''(z_\alpha)z_\alpha^2(2c_4 - c_1^2) - D'''(z_\alpha)z_\alpha^3 c_1 c_2 \\
&\quad + [c_2 D'''(z_\alpha)z_\alpha^2 + 2c_2 D''(z_\alpha)z_\alpha - c_1 D''(z_\alpha)z_\alpha \\
&\quad - c_1 D'(z_\alpha)]k_1 + \frac{1}{2}D''(z_\alpha)k_1^2\Bigg]. \qquad (4.48)
\end{aligned}$$

Proof of Corollary 4:
We have (note that $\tilde{D}'(x)$ is the density of the gamma distribution with parameters α_g and $1/2$)

$$\begin{aligned}
\tilde{D}'(x) &:= \frac{2^{-\alpha_g}}{\Gamma(\alpha_g)} x^{\alpha_g - 1} \exp(-\frac{1}{2}x), \\
\tilde{D}''(x) &= [(\alpha_g - 1)/x - 1/2]\, \tilde{D}'(x), \\
\tilde{D}'''(x) &= \left[(\alpha_g - 1)(\alpha_g - 2)/x^2 - (\alpha_g - 1)/x + 1/4\right] \tilde{D}'(x)
\end{aligned}$$

and

$$f(x) := (x - b_\chi)/a_\chi, \quad f'(x) = 1/a_\chi, \quad f''(x) = 0, \quad f'''(x) = 0.$$

Hence, the derivatives of $D(x) = \tilde{D}(f(x))$ are given by

$$\begin{aligned}
D'(x) &= \tilde{D}'(f(x))f'(x) = \tilde{D}'(f(x))/a_\chi \\
D''(x) &= \tilde{D}''(f(x))(f'(x))^2 + \tilde{D}'(f(x))f''(x) \\
&= [(\alpha_g - 1)/f(x) - 1/2]\, \tilde{D}'(f(x))/a_\chi^2 \\
D'''(x) &= \tilde{D}'''(f(x))(f'(x))^3 + 3\tilde{D}''(f(x))f'(x)f''(x) + \tilde{D}'(f(x))f'''(x) \\
&= \left[(\alpha_g - 1)(\alpha_g - 2)/f(x)^2 - (\alpha_g - 1)/f(x) + 1/4\right] \tilde{D}'(f(x))/a_\chi^3.
\end{aligned}$$

Furthermore,

$$\frac{D''(x)}{D'(x)} = [(\alpha_g - 1)/f(x) - 1/2]\, /a_\chi$$

and

$$\frac{D'''(x)}{D'(x)} = \left[(\alpha_g - 1)(\alpha_g - 2)/f(x)^2 - (\alpha_g - 1)/f(x) + 1/4\right] /a_\chi^2.$$

The coefficients k_1 and k_2 follow then from (4.47) and (4.48)

$$\begin{aligned} k_1 &= -\frac{1}{a_\chi}\left[(\alpha_g - 1)/f(x) - 1/2\right] z_\alpha^2 c_2 + z_\alpha c_1 \\ &= c_1 z_\alpha - w_1 c_2 z_\alpha^2 \end{aligned} \tag{4.49}$$

and

$$\begin{aligned} k_2 &= (c_3 + c_1^2) z_\alpha + [-\frac{1}{2}(2c_4 - c_1^2) - 3c_1c_2 + \frac{1}{2}c_1^2] w_1 z_\alpha^2 \\ &\quad + 2c_2^2 w_1^2 z_\alpha^3 + (w_1 w_2 - \frac{1}{2}w_1^3) c_2^2 z_\alpha^4. \end{aligned} \tag{4.50}$$

Proof of Lemma 5:
Using (4.20) and (4.21) we obtain

$$\begin{aligned} \sum_{t=1}^{T}\left(\sum_{j=1}^{t}(C_T - I_j)\right)^2 &= \sum_{t=1}^{T}\left(d_t + \sum_{j=1}^{t} e_j\right)^2 \\ &= \sum_{t=1}^{T}\left(\sum_{j=1}^{t} e_j\right)^2 + 2\sum_{t=1}^{T}\left(d_t \sum_{j=1}^{t} e_j\right) + \sum_{t=1}^{T} d_t^2 \\ &= \sum_{t=1}^{T}\left(\sum_{j=1}^{t} e_j\right)^2 + 2\sum_{t=1}^{T}\left(e_t \sum_{j=t}^{T} d_j\right) + \sum_{t=1}^{T} d_t^2 \\ &= \sum_{t=1}^{T}\sum_{s=1}^{T}[T + 1 - \max(s,t)] e_s e_t + 2\sum_{t=1}^{T}\left(e_t \sum_{j=t}^{T} d_j\right) + \sum_{t=1}^{T} d_t^2 \\ &=: e'Ve + 2e'd^\dagger + c_d. \end{aligned}$$

From (4.23) we have $e'Ve = \tilde{e}\Lambda'\tilde{e}$. Hence, the finite sample distribution is given by

$$\begin{aligned} F_{T,\delta}(z) &:= P\left(\frac{1}{\hat{\sigma}_b^2 T^2}\sum_{t=1}^{T}\left(\sum_{j=1}^{t}(C_T - I_j)\right)^2 \le z\right) \\ &= P\left(\frac{1}{\sigma_{e,T}^2 T^2}\left\{\sum_{t=1}^{T}\lambda_t \tilde{e}_t^2 + 2\sum_{t=1}^{T}\tilde{d}_t\tilde{e}_t + c_d\right\} \le z\hat{\sigma}_b^2/\sigma_{e,T}^2\right) \\ &= P\left(\sum_{t=1}^{T}[\lambda_t - Tz(\hat{\sigma}_b^2/\sigma_{e,T}^2)\tilde{g}(\theta)]\tilde{e}_t^2 + 2\sum_{t=1}^{T}\tilde{d}_t\tilde{e}_t + c_d \le 0\right) \end{aligned}$$

$$= \quad P\left(\sum_{t=1}^{T}\tilde{\lambda}_t(\tilde{e}_t+\tilde{d}_t/\tilde{\lambda}_t)^2 \le \tilde{d}_t^2/\tilde{\lambda}_t - c_d\right),$$

where $\tilde{\lambda}_t = \lambda_t - Tz(\hat{\sigma}_b^2/\sigma_{e,T}^2)\tilde{g}(\theta)$.

Proof of Theorem 6:

Let $\mu_{bT} := E(\varsigma_{bT})$ and $\alpha_{m,T} := E(\varsigma_{bT}-\mu_{bT})^m$, $m = 1,2,3,4$. Like Lemma 3 in SPJ (2008) (see also their eq. A.91) we have

$$\mu_{bT} = \mu_b - (bT)^{-q}q_q d_{qT}(1+o(1)) + O(T^{-1}) \tag{4.51}$$

and

$$\begin{aligned}
\alpha_{2,T} &= 2bc_2(1+o(1)) + O(T^{-1}), \qquad (4.52)\\
\alpha_{3,T} &= O(b^2) + O(T^{-1})\\
\alpha_{4,T} &= O(b^2) + O(T^{-1}).
\end{aligned}$$

Then

$$\begin{aligned}
F_{T,\delta}(z) &= E\left[\tilde{H}_{T,\delta}(z;\varsigma_{bT})\right]\\
&= E\Bigg[\tilde{H}_{T,\delta}(z;\mu_{bT}) + \tilde{H}'_{T,\delta}(z;\mu_{bT})(\varsigma_{bT}-\mu_{bT})\\
&\qquad + \frac{1}{2}\tilde{H}''_{T,\delta}(z;\mu_{bT})(\varsigma_{bT}-\mu_{bT})^2\Bigg] + o(b)\\
&\overset{(4.52)}{=} \tilde{H}_{T,\delta}(z;\mu_{bT}) + \frac{1}{2}\tilde{H}''_{T,\delta}(z;\mu_{bT})\cdot 2bc_2 + o(b) + O(T^{-1})\\
&= \tilde{H}_{T,\delta}(z;\mu_b) + \tilde{H}'_{T,\delta}(z;\mu_{bT})(\mu_{bT}-\mu_b) + \tilde{H}''_{T,\delta}(z;\mu_b)bc_2 + o(b) + O(T^{-1}).
\end{aligned}$$

We have

$$\begin{aligned}
\tilde{H}_{T,\delta}(z;\mu_b) &= \tilde{H}_{T,\delta}(z;1) + \tilde{H}'_{T,\delta}(z;1)(\mu_b-1) + o(b)\\
&\overset{(4.41),(4.42)}{=} \tilde{H}_{T,\delta}(z;1) - bc_1\tilde{H}'_{T,\delta}(z;1) + o(b)
\end{aligned}$$

and

$$\tilde{H}'_{T,\delta}(z;\mu_b)(\mu_{bT}-\mu_b) \overset{(4.51)}{=} -g_q d_{qT}\tilde{H}'_{T,\delta}(z;1)(bT)^{-q}(1+o(1))+o(b)+O(T^{-1}).$$

Hence,

$$\begin{aligned} F_{T,\delta}(z) &= \tilde{H}_{T,\delta}(z;1)-bc_1\tilde{H}'_{T,\delta}(z;1)-g_q d_{qT}\tilde{H}'_{T,\delta}(z;1)(bT)^{-q} \\ &\quad + \tilde{H}''_{T,\delta}(z;\mu_b)bc_2+o(b+(bT)^{-q})+O(T^{-1}) \\ &=: B_{1,\delta}+B_{2,\delta}b-g_q d_{qT}B_{3,\delta}(bT)^{-q}+o(b+(bT)^{-q})+O(T^{-1}). \quad (4.53) \end{aligned}$$

Proof of Corollary 7:
The size distortion is given by

$$\begin{aligned} &1-F_{T,0}(z_{\alpha,b})-\alpha \\ &\overset{(4.53)}{=} 1-B_{1,0}-B_{2,0}b+g_q d_{qT}B_{3,0}(bT)^{-q}-\alpha+o(b+(bT)^{-q})+O(T^{-1}) \\ &= F_0(z_{\alpha,b})-B_{1,0}-B_{2,0}b+g_q d_{qT}B_{3,0}(bT)^{-q}+o(b+(bT)^{-q})+O(T^{-1}) \\ &\overset{(4.10)}{=} A_{1,0}-B_{1,0}+(A_{2,0}-B_{2,0})b+g_q d_{qT}B_{3,0}(bT)^{-q}+o(b+(bT)^{-q})+O(T^{-1}). \end{aligned}$$

The power is given by

$$1-F_{T,\delta}(z_{\alpha,b}) \overset{(4.53)}{=} 1-B_{1,\delta}-B_{2,\delta}b+g_q d_{qT}B_{3,\delta}(bT)^{-q}+o(b+(bT)^{-q})+O(T^{-1}).$$

4.B Additional results

Proof (result SPJ (2008) from corollary 4):
If $p=1$ and $c_{1,n}=1$ for $n=1$ then it follows immediately that $a_\chi=1$, $b_\chi=0$ and $d_\chi=1$. Hence, $w_1=-\frac{1}{2}z_\alpha^{-1}-\frac{1}{2}$ and $w_2=\frac{3}{4}z_\alpha^{-2}+\frac{1}{2}z_\alpha^{-1}+\frac{1}{4}$. So that

$$\begin{aligned} k_1 &\overset{(4.49)}{=} c_1 z_\alpha-(-\frac{1}{2}z_\alpha^{-1}-\frac{1}{2})c_2 z_\alpha^2 \\ &= c_1 z_\alpha+\frac{1}{2}c_2 z_\alpha+\frac{1}{2}c_3 z_\alpha^2 \end{aligned}$$

and

$$
\begin{aligned}
k_2 \overset{(4.50)}{=} \quad & z_\alpha(c_3 + c_1^2) + z_\alpha^2[-\frac{1}{2}(c_4 - c_1^2) - 3c_1c_2 + \frac{1}{2}c_1^2] \cdot [-\frac{1}{2}z_\alpha^{-1} - \frac{1}{2}] \\
& + z_\alpha^3[2c_2^2(\frac{1}{4}z_\alpha^{-2} + \frac{1}{2}z_\alpha^{-1} + \frac{1}{4})] + z_\alpha^4\Big\{[-\frac{1}{2}z_\alpha^{-1} - \frac{1}{2}] \cdot [\frac{3}{4}z_\alpha^{-2} + \frac{1}{2}z_\alpha^{-1} + \frac{1}{4}] \\
& + -\frac{1}{2}[-\frac{1}{8}z_\alpha^3 - \frac{1}{4}z_\alpha^{-2} - \frac{1}{8}z_\alpha^{-1} - \frac{1}{8}z_\alpha^{-2} - \frac{1}{4}z_\alpha^{-1} - \frac{1}{8}]\Big\}c_2^2 \\
= \quad & z_\alpha[c_3 + c_1^2 + \frac{1}{4}(c_4 - c_1^2) + \frac{3}{2}c_1c_2 - \frac{1}{4}c_1^2 + \frac{1}{2}c_2^2] \\
& + z_\alpha^2[\frac{1}{4}(c_4 - c_1^2) + \frac{3}{2}c_1c_2 - \frac{1}{4}c_1^2 + c_2^2] \\
& + z_\alpha^3\frac{1}{2}c_2^2 + z_\alpha^4\{-\frac{3}{8}z_\alpha^{-3} - \frac{1}{4}z_\alpha^{-2} - \frac{1}{8}z_\alpha^{-1} - \frac{3}{8}z_\alpha^{-2} - \frac{1}{4}z_\alpha^{-1} - \frac{1}{8} \\
& \quad + \frac{1}{16}z_\alpha^{-3} + \frac{1}{8}z_\alpha^{-2} + \frac{1}{16}z_\alpha^{-1} + \frac{1}{16}z_\alpha^{-2} + \frac{1}{8}z_\alpha^{-1} + \frac{1}{16}\}c_2^2 \\
= \quad & z_\alpha[c_3 + \frac{1}{2}c_1^2 + \frac{1}{4}c_4 + \frac{3}{2}c_1c_2 + \frac{3}{16}c_2^2] + z_\alpha^2[\frac{1}{4}c_4 - \frac{1}{2}c_1^2 + \frac{3}{2}c_1c_2 + \frac{9}{16}c_2^2] \\
& + \frac{5}{16}z_\alpha^3 - \frac{1}{16}c_2^2z_\alpha^4.
\end{aligned}
$$

Proof (optimal b):

For the loss function L (sum of errors) we have

$$
\begin{aligned}
\frac{\partial L}{\partial b} = 0 \quad & \Leftrightarrow \quad B_{2,\delta} - w_T(B_{2,0} - A_{2,0}) - q\{w_T B_{3,0} - B_{3,\delta}\}q_q d_{qT} b^{-q-1}T^{-q} = 0 \\
& \Leftrightarrow \quad b^{-q-1} = \frac{B_{2,\delta} - w_T(B_{2,0} - A_{2,0})}{qq_q d_{qT}\{w_T B_{3,0} - B_{3,\delta}\}T^{-q}} \\
& \Leftrightarrow \quad b = \left\{\frac{qg_q d_{qT}\{w_T B_{3,0} - B_{3,\delta}\}}{B_{2,\delta} - w_T(B_{2,0} - A_{2,0})}\right\}^{\frac{1}{q+1}} T^{-\frac{q}{q+1}}.
\end{aligned}
$$

4.C Additional figures

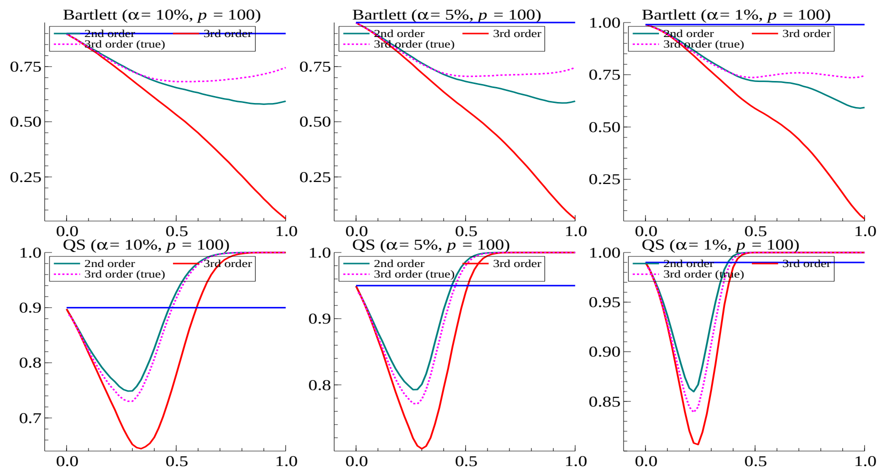

Figure 4.3: Expansion of limit distribution. Probability (y-axis) against bandwidth parameter b (x-axis).

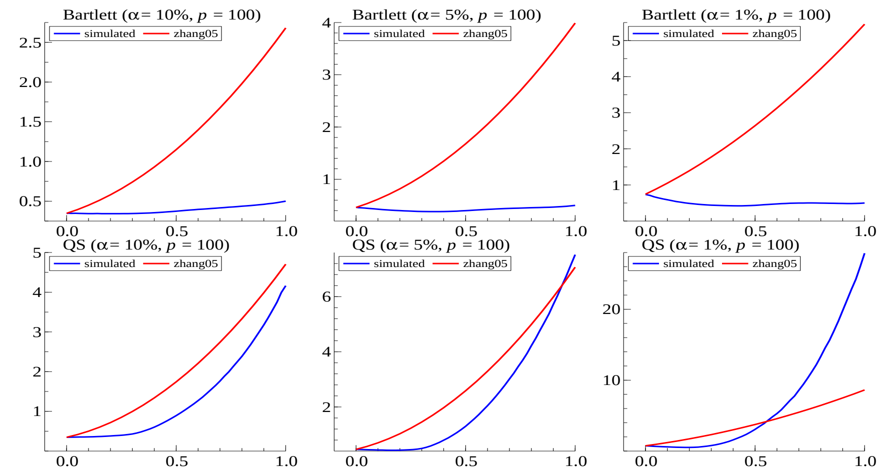

Figure 4.4: Simulated and 3rd order corrected critical values. Critical value (y-axis) against bandwidth parameter b (x-axis).

Chapter 5

A nonparametric overall copula constancy test*

In this chapter we introduce a new test that examines the constancy of the copula on the complete unit square. To test has a nonstandard limit distribution. We propose a bootstrap procedure to obtain the critical values. We show in a simulation study that the test performs good compared to some recent proposed tests.

5.1 Introduction

Recently, Busetti and Harvey (2011) and Krämer and Van Kampen (2011) propose several tests to examine the constancy of the copula in a particular point $(\tau_1, \tau_2) \in [0,1]^2$. In this chapter we construct an overall copula constancy test which does not require the *a priori* selection of a point $(\tau_1,\ \tau_2)$. Like the aforementioned authors we base our test on the partial sums of suitable indicator variables. The test has a nonstandard asymptotic distribution that differs from the point tests. To obtain critical values we use the bootstrap method proposed by Inoue (2001). A short simulation study shows that our test outperforms a recently proposed test that examines changes in Spearman's rankcorrelation but is not uniformly better as the

*This chapter is based on Van Kampen and Wied (2010).

point test. To illustrate our test, we apply it to the MSCI stock returns of the US, UK, France, Germany and Japan. In addition, we consider several US stock indices. For several pairs we are able to reject the hypothesis of a constant copula.

5.2 An overall copula constancy test

Let y_t be a bivariate series of observations and let $\xi(\tau) = (\xi_1(\tau_1), \xi_2(\tau_2))$ denote the marginal τ_i-quantiles, $i = 1, 2$. Like the point tests, we base our overall copula constancy test on the partial sums of the bivariate τ-quantics

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

where

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

and $C_T(\tau_1, \tau_2) = T^{-1} \sum_{t=1}^{T} I(y_t, \xi(\tau))$.

The idea of the overall copula constancy test is to summarize the information at the unit square by taking the maximum over all points (τ_1, τ_2). Note that we could also consider other functional forms (see discussion below) to construct an overall copula constancy test.

Define

$$S_T(r, \tau) := \frac{1}{\sqrt{T}} \sum_{t=1}^{[rT]} BIQ(y_t, \xi(\tau)) \qquad \tau \in [0,1]^2, \ r \in [0,1].$$

In the previous chapters we took $\tau = \tau^0$ fixed and considered the weak convergence of the random function $S_T(r) \equiv S_T(\tau^0, r)$ to a Brownian bridge. These random functions are in the function space, $D([0,1])$, of functions that are right continuous but may have discontinuities on the left. In this chapter, we do not fix τ and $S_T(r, \tau)$ is a random function in $D([0,1]^{d+1})$ with d the dimension of the vector y_t. For sake of simplicity we assume $d = 2$, but all arguments below also hold for higher dimensions.

We make the following assumption:

Assumption 5.1.

(i) $\{y_{it}\}$ is strong mixing with mixing coefficients that satisfy, for some $\gamma \in (0,2)$,

$$\sum_{j=1}^{\infty} j^2 \alpha(j)^{\gamma/(4+\gamma)}.$$

(ii) $\{y_{it}\}$ is a strictly stationary process.

Assumption 5.1(i) is similar to Inoue (2001). Recall that the point tests require that the mixing coefficients are $\alpha(m) = O(m^{-p/(p-2)})$ for $p > 2$. This implies that $\sum_{m=1}^{\infty} \alpha(m)^{\frac{1}{\phi_0}} < \infty$ for $0 < \phi_0 < p/(p-2)$. Hence, Assumption (i) for the overall copula constancy test is stronger.

Assumption (ii) is required for Theorem 2.1 in Inoue (2001), see the proof in the Appendix.

To obtain the asymptotic null distribution of our test we need an invariance principle for the multivariate rank process as defined in Appendix 5.A. Rüschendorf (1976) derives this under the assumption that the copula has continuous partial derivatives. More recently, Bücher and Volgushev (2011) provide an invariance principle without this condition. This is important since many copulas (such as the Clayton and Gumbel copula) do not have continuous derivatives in the corner points, see e.g. Segers (2010).

The following theorem gives the asymptotic null distribution of our test:

Theorem 5.1. *Under Assumption 5.1*

$$\sup_{0 \le r \le 1} \sup_{\tau \in [0,1]^2} |S_T(\cdot,\cdot)| \xrightarrow{d} \sup_{0 \le r \le 1} \sup_{\tau \in [0,1]^2} |A_0(\cdot,\cdot)| =: Z, \tag{5.1}$$

where

$$A_0(r,\tau) := V_0(r,\tau) - rV_0(1,\tau),$$

with $V_0(r,\tau)$ *an almost surely continuous, centered Gaussian process whose covariance* *structure is given by*

$$E[V_0(r,\tau)V_0(r',\tau')] = \\ \min(r,r') \sum_{j=\infty}^{\infty} E\big[I(y_t,\xi(\tau))I(y_{t+j},\xi(\tau')) - C(\tau_1,\tau_2)C(\tau_1',\tau_2')\big].$$

After some rewriting it can be shown that the test is basically the weighted Kolmogorov-Smirnov test presented in Inoue (2001). The difference between his paper and our paper is that the indicator functions in our case depend on an estimated quantity $\hat{\xi}(\tau)$ instead of a known quantity.

The process $V_0(r,\tau)$ is referred to as a Kiefer process. The challenge is the derivation of critical values of the random variable Z. For this, we use the bootstrap method of Inoue (2001) (which is based on Hansen (1996)). Analogously we search for a simulation version of the test statistic, i.e. a process that converges to $V_0(\cdot,\cdot)$ conditioned on the data when T and the so called block length l converge to ∞. This will be the process

$$V_T^*(r,\tau,\omega) = \frac{1}{\sqrt{T}} \sum_{t=1}^{[rT]-l+1} z_t \sum_{i=t}^{t+l-1} BIQ(y_t,\hat{\xi}(\tau)), \tag{5.2}$$

where z_t are $NID(0,1/l)$ random variables and ω denotes the particular sample, such that more formally $y_{it}(\omega)$ denotes the realization of a random variable. For notational convenience, we keep the simplified notation y_{it}.

With a modification of Theorem 2.3 in Inoue (2001) (see Appendix 5.A) we get

$$V_T^*(\cdot,\cdot,\omega) \xrightarrow{d} V_0(\cdot,\cdot) \qquad \omega - \text{almost surely.} \tag{5.3}$$

Hence, a simulated version of the test becomes

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[rT]-l+1} z_t \sum_{i=t}^{t+l-1} BIQ(y_t,\hat{\xi}(\tau)) - r\frac{1}{\sqrt{T}} \sum_{t=1}^{T-l+1} z_t \sum_{i=t}^{t+l-1} BIQ(y_t,\hat{\xi}(\tau)). \tag{5.4}$$

Let J denote the number of simulation replications. The following procedure can then be used to examine the overall constancy of the copula:

1. Calculate the test statistic (5.1).

2. For $j = 1, \dots, J$, draw $z_t^{(j)}$ from $N(0, 1/l)$ and calculate the test statistic (5.4) using $z_t^{(j)}$.

3 Calculate the $1-\alpha$ quantile of J simulated test statistics. If the resulting value is larger than the original test statistic, reject H_0.

Finally, an interesting complementary approach to the test proposed above 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 Wied 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 viewed as an alternative way to analyze if the copula is constant.

In the next section we examine the finite sample properties of our test using Monte Carlo simulations. We show that the overall copula constancy test of Theorem 5.1 outperforms the rankcorrelation test of Wied et al. (2010) if there are simple shifts in the copula parameter.

5.3 Finite sample properties

To examine the properties of the test in finite samples, we simulate observations from a Clayton copula with parameter $\psi = 1$. The marginals follow a AR(1) process with parameter ϕ_i, $i = 1, 2$, and standard normal distributed innovations. We consider samples of size $T = 500$ and 1000. The number of bootstrap replications is set at 199 and the number of Monte Carlo simulations is set at 1000. For computational reasons, we evaluate the test on a subset of points in the unit square. That is, we construct a grid $[1/T, 1]^2$ with step size 1/10, and take the maximum over these points. To examine the sensitivity of the block length parameter we choose $l = 10, 30, 50$. For comparison purposes we also report the results for the rankcorrelation test of Wied et al. (2010) based on the Bartlett kernel and conventional bandwidth rules $\gamma_T(m) = int[m(T/100)^{1/4}]$, $m = 0, 4, 12$. The rankcorrelation test and point test are based on 5000 replications.

Table 5.1 shows the size of the test. In case the series are independent, we have to choose relative low values of l to make the test of appropriate size. If the series are serial correlated we need higher values of l. For the Spearman test we need higher bandwidth values if the series are serial correlated. Since block lengths in the maximum test and bandwidth values in the Spearman test represent different quantities, comparing changes in block length values with changes in bandwidth values is somewhat difficult. It should however be clear that, using conventional block length values and bandwidth rules, the size of the Spearman test is much more sensitive with respect to the bandwidth rule than the size of the maximum test is with respect to the block length.

Table 5.1: Size of maximum and Spearman's rankcorrelation test. Nominal size is 0.05.

			Maximum			Spearman		
T	ϕ_1	ϕ_2	$l = 10$	$l = 30$	$l = 50$	$\gamma_T(0)$	$\gamma_T(4)$	$\gamma_T(12)$
500	0.0	0.0	0.045	0.029	0.023	0.038	0.049	0.097
	0.3	0.3	0.059	0.036	0.028	0.198	0.061	0.072
	0.5	0.5	0.080	0.041	0.037	0.444	0.100	0.065
	0.7	0.7	0.114	0.045	0.032	0.770	0.187	0.062
1000	0.0	0.0	0.048	0.040	0.029	0.044	0.047	0.083
	0.3	0.3	0.058	0.040	0.032	0.213	0.058	0.068
	0.5	0.5	0.079	0.043	0.034	0.450	0.084	0.060
	0.7	0.7	0.125	0.059	0.044	0.800	0.142	0.062

To examine the power of the test we include one or two breaks in the sample such that each interval is of equal length (last interval is slightly shorter in case of 500 observations and 2 breaks). We assume that the observations are serially uncorrelated ($\phi_1 = \phi_2 = 0$) and set $l = 10$ in the maximum test and $\gamma_T(0) = 0$ in the Spearman test.

Table 5.2 shows the results for a shift of $\psi = 1$ to respectively 2.5, 7.5 and 15 (values taken from Busetti and Harvey (2011)). The test has good power in case of 1 break. In line with the point tests, we see that the power of the test decreases with the number of breaks and increases with the sample size.

Our test outperforms the Spearman test in the simulation study but is not uniformly better as the point tests. In particular, our test outperforms the point test at

$\tau = 0.10$ but not at $\tau = 0.5$.

Table 5.2: Power of maximum test ($l = 10$), Spearman based test ($\gamma_T(0) = 0$) and the point test evaluated at $\tau = 0.1$ and $\tau = 0.5$, respectively. Copula = Clayton; Nominal size = 0.05.

#breaks	T	ψ_1	ψ_2	Maximum	Spearman	$\tau = 0.1$	$\tau = 0.5$
1	500	1	2.5	0.154	0.110	0.154	0.239
			7.5	0.588	0.237	0.317	0.721
			15	0.788	0.280	0.351	0.834
	1000	1	2.5	0.298	0.186	0.269	0.422
			7.5	0.920	0.459	0.570	0.947
			15	0.986	0.530	0.643	0.984
2	500	1	2.5	0.062	0.053	0.039	0.058
			7.5	0.114	0.074	0.049	0.149
			15	0.164	0.082	0.058	0.208
	1000	1	2.5	0.094	0.074	0.055	0.086
			7.5	0.271	0.134	0.096	0.348
			15	0.390	0.158	0.119	0.529

To examine the sensitivity of our results, we also performed the analysis using a Gaussian copula. Under the null hypothesis the Gaussian copula parameter equals 0.5 and under the alternative the copula parameter jumps to 0.1, 0.25, 0.75 and 0.9, respectively. These parameter values are in line with Busetti and Harvey (2011). Table 5.6 (in Appendix 5.B) presents the results. The results are mainly in line with the results of the Clayton Copula. However, we observe sometimes slightly higher rejection frequencies for the Spearman's test, compared to the maximum test, if there is a decrease in Gaussian copula parameter.

5.4 Empirical applications

In this section we consider two applications to illustrate our test. First, we apply the test to MSCI series analyzed in the previous chapter. Second, we consider return series of the Dow-Jones, NYSE and S&P 500. These US series are also analyzed in Inoue (2001).

5.4.1 MSCI stock index

We illustrate the test using the MSCI return series analyzed in the previous chapters. Again, we apply the test to the raw return series and the standardized innovations of the GARCH(1,1) model (see section 3.6 for details).

To determine l, note that Kendall's tau falls between 0.2 and 0.5 and the first autocorrelation coefficient between 0.0 and 0.2. Therefore, it is reasonable to select low values for the blocklength l. For computational reasons, we evaluate the test on a subset of points in the unit square. That is, we construct a grid $[1/T, 1]^2$ with step size $10/T$, and take the maximum over these points. Table 5.3 presents the resulting p-values. The results are mainly in line with results of the point test; for most of the series we are not able to reject the null hypothesis.

Table 5.3: P-values of the overall copula constancy test applied to the original return series and the standardized innovations of the GARCH model.

		original return series				standardized innovations			
countries	l =	10	20	30	40	10	20	30	40
US - UK		0.080	0.221	0.296	0.347	0.362	0.422	0.568	0.558
US - France		0.101	0.176	0.276	0.251	0.070	0.131	0.236	0.191
US - Germany		0.035	0.080	0.156	0.236	0.151	0.276	0.347	0.382
US - Japan		0.050	0.151	0.156	0.186	0.106	0.241	0.241	0.296
UK - France		0.010	0.030	0.075	0.116	0.020	0.055	0.095	0.146
UK - Germany		0.070	0.196	0.251	0.296	0.166	0.307	0.327	0.372
UK - Japan		0.035	0.055	0.090	0.090	0.055	0.146	0.126	0.121
France - Germany		0.090	0.146	0.226	0.271	0.116	0.171	0.261	0.347
France - Japan		0.050	0.080	0.075	0.131	0.090	0.186	0.276	0.266
Germany - Japan		0.020	0.050	0.075	0.111	0.156	0.266	0.307	0.342

5.4.2 US stock index

In our second application, we applied the test to the Dow-Jones, NYSE and S&P stock index, which are also analyzed in Inoue (2001). The time series consists of the Wednesday returns from 1973 through 1996 (1252 observations). We know that the test is subject to size distortions if the series exhibit stochastic volatility patterns. Therefore, we estimate a GARCH(1,1) model and apply the test to the standardized innovations. Table 5.4 shows the maximum likelihood estimates of a GARCH (1,1)

Table 5.4: Maximum likelihood estimates

	Dow Jones			NYSE			S&P500		
	normal	student	skew-st.	normal	student	skew-st.	normal	student	skew-st.
μ	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***
ω	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
α	0.084**	0.067***	0.069***	0.113***	0.086***	0.086***	0.098***	0.080***	0.080***
β	0.883***	0.910***	0.908***	0.860***	0.891***	0.893***	0.877***	0.897***	0.898***
asym			-0.163***			-0.191***			-0.170***
tail		9.637***	10.604***		10.746***	11.962***		11.252***	12.490***
AIC	-4.921	-4.950	-4.960	-5.038	-5.063	-5.078	-4.997	-5.020	-5.031
LB(1)	0.063	0.017	0.019	0.087	0.149	0.155	0.150	0.101	0.105
LB(2)	0.084	0.057	0.064	0.107	0.174	0.186	0.151	0.103	0.105
LB(4)	0.481	0.484	0.472	1.219	1.364	1.354	1.270	1.323	1.290
$LB^2(1)$	3.921	5.333	4.953	2.045	3.582	3.527	2.177	3.350	3.177
$LB^2(2)$	3.925	5.476	5.037	2.110	3.947	3.875	2.290	3.700	3.493

The statistics are the Akaike Information Criteria (AIC) and the Ljung-Box statistics for serial correlation (no. of lags between brackets). Significance levels: 1% (***), 5% (**) and 10% (*).

model with Gaussian, Student and skewed-student distributed innovations. Based on the AIC, we select the GARCH model with skewed-student distributed innovations.

The Ljung-Box statistics in table 5.4 shows that we cannot reject the null hypothese of no serial correlation. Hence, we select the block-length l in the overall test and bandwidth rule $\gamma_T(m)$ in the point test relatively low.

Table 5.5 gives the results of the copula constancy test. The results provide some evidence that the dependence function between the series is not constant over time. In line with simulation study, we see that the results of the point test heavily depend on the chosen quantile. The results of the Spearman tests are in line with our overall constancy test. An application of the quantile constancy test of Busetti and Harvey (2007) shows that the lower quantiles are not constant over time. The results presented above are likely affected by this. A point for further research might be to investigate the application of different models for the marginal distribution.

Table 5.5: Copula constancy test statistics for the standardized residuals of a GARCH(1,1) model with skewed-student distributed innovations.Significance levels: 1% (***), 5% (**) and 10% (*).

		DJ-NYSE	DJ-S&P500	NYSE-S&P500
overall	$l = 10$	0.005	0.000	0.030
(p-values)	$l = 20$	0.000	0.000	0.045
	$l = 30$	0.000	0.000	0.050
	$l = 40$	0.000	0.005	0.030
point	$\tau = 0.10$	0.820***	0.837***	0.646**
	$\tau = 0.25$	1.390***	1.572***	0.865***
	$\tau = 0.50$	0.672**	0.701**	0.425*
	$\tau = 0.75$	0.173	0.155	0.047
	$\tau = 0.90$	0.145	0.162	0.338
Spearman	$\gamma_T(0)$	1.678***	1.738***	1.548**
	$\gamma_T(4)$	1.707***	1.792***	1.577**

5.A Proofs

Proof of Theorem 5.1

Define the multivariate sequential empirical process

$$V_T(r,\tau) := \frac{1}{\sqrt{T}} \sum_{t=1}^{[rT]} (I(y_t, \xi(\tau)) - C(\tau_1, \tau_2))$$

and the multivariate rank order process

$$L_T(r,\tau) := \frac{1}{\sqrt{T}} \sum_{t=1}^{[rT]} (I(y_t, \hat{\xi}(\tau)) - C(\tau_1, \tau_2)).$$

Note that

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[rT]} BIQ(y_t, \hat{\xi}(\tau)) \quad = \quad \frac{[rT]}{T} L_T(1,\tau) - L_T(r,\tau) =: -A_T(r,\tau).$$

Since $V_T \xrightarrow{d} V_0$ (Theorem 2.1 in Inoue (2001)), condition 3.1 in Bücher and Volgushev (2011) is satisfied, and we have from their corollary 3.3a that

$$A_T \xrightarrow{d} A_0(\cdot,\cdot).$$

Finally, using the continuous mapping theorem, we have

$$\max_{0\le r\le 1} \max_{\tau\in[0,1]^2} \left| \frac{1}{\sqrt{T}} \sum_{t=1}^{[rT]} BIQ(y_t, \hat{\xi}(\tau)) \right| \xrightarrow{d} \max_{0\le r\le 1} \max_{\tau\in[0,1]^2} |A_0(r,\tau)|.$$

■

Proof equation (5.3)

Let $C_T(\tau_1, \tau_2; \omega)$ denote the empirical copula based on $y_t(\omega)$ and let $F_{i,T}$ denote the marginal distributions, $i = 1, 2$. Define

$$V_T^{**}(r,\tau,\omega) := \frac{1}{\sqrt{T}} \sum_{t=1}^{[rT]-l+1} z_t \sum_{i=t}^{t+l-1} (I(y_t(\omega), \xi(\tau)) - C_T(\tau_1, \tau_2; \omega))$$

and note that

$$V_T^*(r,\tau,\omega) = V_T^{**}(r,(F_{1,T}\circ\hat{\xi}(\tau_1),F_{2,T}\circ\hat{\xi}(\tau_2)),\omega).$$

Theorem 2.3 in Inoue (2001) gives

$$V_T^{**}(\cdot,\cdot,\omega) \overset{d}{\to} V_0(\cdot,\cdot) \qquad \omega - \text{almost surely.} \tag{5.5}$$

Hence,

$$\begin{aligned}
&\sup_{r,\tau}|V_T^*(r,\tau,\omega) - V_0(r,\tau,\omega)| \\
&\quad \le \sup_{r,\tau}\Big|V_T^{**}(r,(F_{1,T}\circ\hat{\xi}(\tau_1),F_{2,T}\circ\hat{\xi}(\tau_2)),\omega) \\
&\qquad\qquad - V_0(r,(F_{1,T}\circ\hat{\xi}(\tau_1),F_{2,T}\circ\hat{\xi}(\tau_2)),\omega)\Big| \\
&\qquad + \sup_{r,\tau}\Big|V_0(r,(F_{1,T}\circ\hat{\xi}(\tau_1),F_{2,T}\circ\hat{\xi}(\tau_2)),\omega) - V_0(r,\tau,\omega)\Big|
\end{aligned}$$

is $o_p(1)$ by (5.5) and the uniform convergence of $\hat{\xi}(\cdot)$ to $\xi(\cdot)$. ∎

5.B Additional results

Table 5.6: Power of maximum test, Spearman's rankcorrelation test and the point test evaluated at $\tau = 0.1$ and $\tau = 0.5$, respectively. Copula = Gausssian; Nominal size = 0.05.

#breaks	T	ψ_1	ψ_2	Maximum	Spearman	$\tau = 0.1$	$\tau = 0.5$
1	500	0.5	0.1	0.219	0.239	0.230	0.310
			0.25	0.109	0.114	0.131	0.144
			0.75	0.138	0.102	0.146	0.184
			0.9	0.389	0.219	0.394	0.496
	1000	0.5	0.1	0.420	0.435	0.443	0.562
			0.25	0.179	0.179	0.203	0.251
			0.75	0.288	0.185	0.273	0.336
			0.9	0.748	0.401	0.671	0.799
2	500	0.5	0.1	0.086	0.082	0.088	0.083
			0.25	0.048	0.056	0.068	0.062
			0.75	0.053	0.053	0.038	0.059
			0.9	0.091	0.073	0.046	0.096
	1000	0.5	0.1	0.105	0.142	0.104	0.118
			0.25	0.069	0.079	0.076	0.073
			0.75	0.073	0.068	0.052	0.072
			0.9	0.166	0.118	0.110	0.199

Chapter 6

Conclusion and Further Research

In the second chapter of this thesis we introduce a new copula constancy test. The test is based on a suitable indicator series. The test outperforms a recently proposed test of Busetti and Harvey (2011) if there are two breaks in the sample. The power of the test increases if the sample size increases but decreases if the number of breaks increases. To illustrate the test we apply it to a long time series of the stock indices of Hong Kong and Malaysia. The time series is characterized by several high volatility periods (East Asian crisis and the recent financial crisis). The proposed test indicates that the copula is not constant over time.

The test developed in chapter 2 assumes that the observations are independent and identically distributed. This assumption is often violated in the time series literature. In the third chapter we show that the asymptotic null distribution of the test remains the same under a suitable weak dependence (strong mixing) assumption.

The test is, however, subject to size distortions if the time series exhibits serial correlation and stochastic volatility patterns. Current practice suggests to apply the tests to the standardized residuals of some ARMA/GARCH model. We provide in chapter 3 sufficient conditions under which this is allowed.

To illustrate the importance of controlling for stochastic volatility patterns, we apply the tests to time series of the MSCI stock index of the US, UK, France, Germany and Japan. Our results show that if we do not control for stochastic volatility patterns, we often reject the null hypothesis of a constant copula. Application of the tests to the residual series shows that we are often not able to reject the null hypothesis. Although we can of course not conclude that the copula is constant, it illustrates the importance of controlling for changes in the marginal distributions.

The disadvantage of filtering, as suggested above, is that the resulting model might be misspecified. In particular, the residuals might be serially correlated. To control for this, the long run variance of the test can be replaced by some heteroscedasticity and autocorrelation consistent estimate. In chapter 4 we improve the finite sample performance of the test by replacing the long run variance by an inconsistent estimate. We construct this estimate using a kernel based approach. The resulting asymptotic distribution of the test depends on the kernel and a bandwidth parameter. A simulation study shows that this approach improves the finite sample performance of the test, for a loss criterion based on the sum of squares of the type I and II errors.

Since the asymptotic distribution depends on the bandwidth parameter, the question arises if we can select this parameter such that it minimizes the loss function. To this extend we approximate the finite sample and limit distribution to construct a new bandwidth rule. The resulting bandwidth rule performs better in some special cases but is not uniformly better. A reason for this is that the second order approximations do not fully reflect the true distributions. Especially, in case of the Bartlett kernel the difference is too large.

A drawback of the proposed tests is that they solely examine the constancy of the copula in a particular point. In chapter 5 we construct an overall copula constancy test. The test is based on the same partial sum process as the point test but examines the maximum deviation on the complete unit square. To obtain the critical values of the test we use a bootstrap algorithm. A simulation study shows that the test has nontrivial power. As before, the power decreases if the number of breaks increases. We also show that the test outperforms a recently proposed test that examines the constancy of Spearman's rankcorrelation coefficient.

The previous chapters provide a clear contribution to the current string of literature but still some important questions remain.

Functionals of partial sums

In the fifth chapter we show how to examine the overall constancy of a copula. The proposed test considers the maximum of the partial sum process at several points in the unit square. More generally, we could have used other functionals of the same process. Wied et al. (2010) consider e.g. the integral which can then be related to dependence measures such as Spearman's rho. We conjecture that a similar test can be derived for Kendall's tau.

Optimal copula constancy tests

The question remains to what extend the developed tests are optimal against a particular alternative. For example, against the alternative that the copula follows a mixture of copulas with a break in the weights. Optimality can then be defined in terms of some weighted average power criterion. That is, we assign weights (probabilities) to different alternatives and we maximize the weighted average of the power functions (see Andrews and Ploberger (1994), Andrews, Lee and Ploberger (1996)).

Finite sample performance

The bandwidth rule developed in chapter 4 only outperforms the conventional bandwidth rule in some special case but is not generally better. A reason for this is that the second order approximations are insufficient to fully describe the limit and finite sample distribution. An interesting topic for further research might therefore be to derive a bandwidth rule under higher order approximations.

The use of fixed-b critical values shows good performance for the loss function based on the sum of squared errors. An alternative method to improve the finite sample performance of the test might be to bootstrap the critical values. The question arises which method is the most favorable one.

Finally, the question remains which kernel function is optimal. Andrews (1991) showed that under an asymptotic mean square error criterion we should use the quadratic spectral kernel. It is unclear if this result still holds under the weighted sum of (squared) errors criterion of chapter 4.

Bibliography

Amsler, C., Schmidt, P. and Vogelsang, T. J., 2009. The KPSS Test Using Fixed-b Critical Values: Size and Power in Highly Autocorrelated Time Series. Journal of Time Series Econometrics 1 (1), 1–42.

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

Andrews, D. W. K. and Ploberger, W., 1994. Optimal tests when a nuisance parameter is present only under the alternative. Econometrica 62 (6), 1383-1414.

Andrews, D. W. K., Lee, I. and Ploberger, W., 1996. Optimal changepoint tests for normal linear regression. Journal of Econometrics 70, 9-38.

Ang, A., Chen, J., 2002. Asymmetric correlations of equity portfolios. Journal of Financial Economics 63 (3), 443-494.

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

Billingsley, P., 1986. Convergence of Probability Measures. New York, Wiley.

Bücher, A. and Volgushev, S., 2011. Empirical and sequential empirical copula processes under serial dependence. Working Paper.

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

Busetti, F. and Harvey, A., 2011. When is a copula constant? A test for changing relationships. Journal of Financial Econometrics 9 (1), 106–131.

Cambell, R., Koedijk, K. and Kofman, P., 2002. Increased correlation in bear markets. Financial Analysts Journal 58 (1), 87-94.

Cambell, R. A. J., Forbes, C. S., Koedijk, K. G. and Kofman, P., 2008. Increasing correlations or just fat tails? Journal of Empirical Finance 15, 287-309.

Choudhry, T., 1995. Integrated-GARCH an non-stationary variances: Evidence from European stock markets during the 1920s and 1930s. Economics Letters 48, 55–59.

Cherubini, U., Luciano, E. and Vecchiato, W., 2004. Copula methods in finance. Wiley Finance.

Cont, R., 2001. Empirical properties of asset returns: stylized facts and statistical issues. Quantitative Finance 1, 223–236.

Creal, D., Koopman, S.J. and Lucas, A., 2008. A general framework for observation driven time-varying parameter models, Tinbergen Institute Discussion Paper 108/4.

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

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

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

De Melo Mendez, B. V., 2005. Asymmetric extreme interdependence in emerging equity markets. Applied Stochastic Models in Business and Industry 21, 89–107.

Den Haan, W. J. and Levin, A., 1997. A practitioner's guide to robust variance estimation. in Handbook of Statistics. Vol 15. ed. by G.S. Maddala and C.R. Rao. New York: North Holland, 299–342.

Dias, A. and Embrechts, P., 2004. Change-point analysis for dependence structures in finance and insurance. In: Risk Measures for the 21st Century, ed. Giorgio Szegoe. Wiley Finance Series, 321–335.

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

Durbin, J. and Knott M., 1972. Components of Cramerér-von Mises Statistics. I. Journal of the Royal Statistical Society Ser.B 34, 290–307.

Forbes, K.J. and Rigobon, R., 2002. No contagion, only interdependence: measuring stock market comovements. Journal of Finance 57 (5), 2223–2261.

Fortin, I., Kuzmicz, C., 2002. Tail dependence in stock returns. Journal of Intelligent Systems in Accounting, Finance and Management 11, 89-107.

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

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

Hansen, B.E., 1996. Inference when a nuisance parameter is not identified under the null hypothesis. Econometrica 64 (2), 413–430.

Harvey, A., 2010. Tracking a changing copula. Journal of Empirical Finance 17, 485-500.

Imhof, J. P., 1961. Computing the Distribution of Quadratic Forms in Normal Variables. Biometrika 48(3/4), 419-426.

Inoue, A., 2001. Testing for distributional change in time series. Econometric Theory 17, 156-187.

Jansson, M., 2004. The Error on Rejection Probability of simple autocorrelation robust tests. Econometrica 72(3), 937–946.

Jondeau, E. and 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.

Kac, M. and Siegert, A. J. F., 1947. An Explicit Representation of a Stationary Gaussian Process. Annals of Mathematical Statistics 18(3), 438–442.

Kennedy, D., 1976. The distribution of the maximum Brownian excursion. Journal of Applied Probability 13, 371–376.

Kiefer, N. M. and Vogelsang, T. J., 2002. Heteroskedasticity-Autocorrelation Robust Testing using Bandwidth equal to Sample Size. Econometric Theory 18, 1350–1366.

Kiefer, N. M. and Vogelsang, T. J., 2005. A new asymptotic theory for heteroskedasticity-autocorrelation robust tests. Econometric Theory 21, 1130–1164.

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

Krämer, W. and Schotman, P., 1992. Range vs. maximum in the ols-based version of the cusum test. Economics Letters 40 (4), 379–381.

Krämer, W. and Van Kampen, M., 2011. A simple nonparametric test for structural change in joint tail probabilities. Economics Letters 110 (3), 245-247.

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

Laurent, S. and 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. and Solnik, B., 2001. Extreme correlation of international equity markets. Journal of Finance 56 (2), 649–676.

Müller, U. K., 2005. Size and Power of Tests of Stationarity in Highly Autocorrelated Time Series. Journal of Econometrics 128, 195–213.

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

Patton, A. J., 2004. On the Out-of-Sample Importance of Skewness and Asymmetric Dependence for Asset Allocation. Journal of Financial Econometrics, Vol. 2, No. 1, 130–168.

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

Phillips, P.C.B., Sun, Y. and Jin, S., 2007. Long run variance estimation and robust regression testing using sharp origin kernels with no truncation. Journal of Statistical Planning and Inference 137, 985–1023.

Phillips, P.C.B., Sun, Y. and Jin, S., 2006. Spectral density estimation and robust hypothesis testing using steep origin kernels without truncation. International Economic Review 47(3), 837–893.

Ploberger, W. and Krämer, W., 1992. The cusum test with ols residuals. Econometrica 60 (2), 271–85.

Priestley, M.B., 1981. Spectral Analysis and Time Series, Academic Press Inc.

Rodriguez, J. C., 2007. Measuring financial contagion: A copula approach. Journal of Empirical Finance 14 (3), 401–423.

Rosenberg, J.V. and Schuermann, T., 2004. A general approach to integrated risk management with skewed, fat-tailed risk. Federal Reserve Bank of New York, staff report 185.

Rüschendorf, L., 1976. Asymptotic Distributions of Multivariate Rank Order Statistics. Annals of Statistics 4 (5), 912–923.

Segers, J., 2010. Asymptotics of empirical copula processes under nonrestrictive smoothness conditions. Working paper.

Sul, D. Phillips, P.C.B. and Choi, C., 2005. Prewhitening Bias in HAC estimation. Oxford Bulletin of Economics and Statistics 67 (4), 517–546.

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

Sun, Y., Phillips, P.C.B. and Jin, S., 2008. Optimal bandwidth selection in heteroskedasticity-autocorrelation robust testing. Econometrica 76(1), 175–194.

Sun, Y., Phillips, P.C.B. and Jin, S., 2010. Power maximization and size control in heteroskedasticity and autocorrelation robust tests with exponentiated kernels. Cowles Foundation Discussion Paper No.1749, Yale University.

Sun, W., Rachev, S., Stoyanov, S., Fabozzi, F., 2008. Multivariate skewed student's t copula in the analysis of nonlinear and asymmetric dependence in the german equity market. Studies in Nonlinear Dynamics & Econometrics 12 (2), 1572–1572.

Ullah, A., 2004. Finite Sample Econometrics, Oxford University Press.

Van Kampen, M. and Wied, D., 2010. A non-parametric constancy test for copulas under weak dependence. Tech. Rep. 36/10, Fakultät Statistik, Universität Dortmund.

Wied, D., Dehling, H., Van Kampen, M. and Vogel, D., 2010. A fluctuation test for constant correlation with spearman's rho. Tech. Rep. 16/11, Fakultät Statistik, Universität Dortmund.

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

Zhang, J.T., 2005. Approximate and Asymptotic Distributions of Chi-Squared-Type Mixtures With Applications. Journal of the American Statistical Association 100, 273–285.