

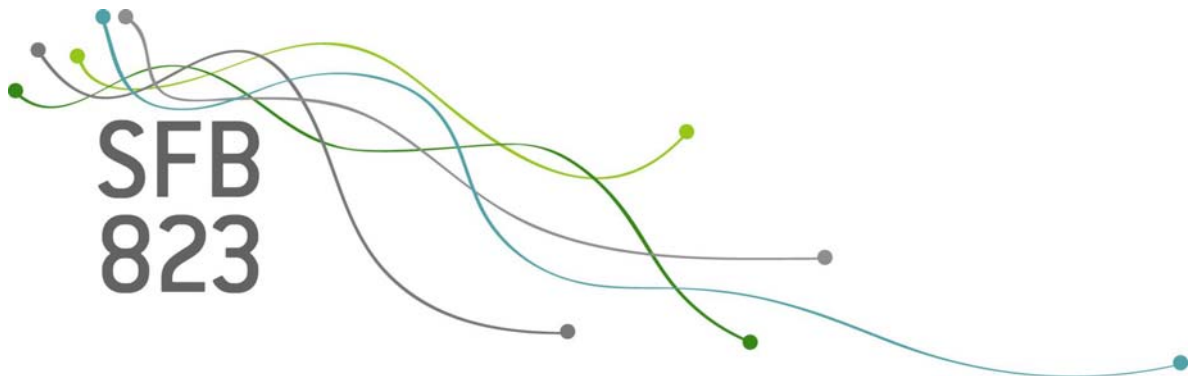
SFB  
823

# On the robustness of goodness-of-fit tests for copulas

Gregor N. F. Weiß

Nr. 40/2011

Discussion Paper





# On the robustness of goodness-of-fit tests for copulas

Gregor N.F. Weiß\*

Juniorprofessur für Finance, Technische Universität Dortmund

September 27, 2011

**Abstract** This paper is the first to analyze the robustness of goodness-of-fit for bivariate elliptical and archimedean copulas. To assess the tests' robustness, we consider perturbations and outliers both in the dependence structure and the observations from the joint distribution. The Monte Carlo simulations show that independent of the underlying true copula, the GoF-test or chosen test statistic, even minor contaminations of the data can lead to a significant decrease in the GoF-tests' power. In order to robustify the GoF-tests, several methods for the detection of multivariate outliers are applied to the contaminated data. The results show that the exclusion of outliers can have a beneficial effect on the power of the GoF-tests. Moreover, this robustification strategy improves the power of GoF-testing when used to identify the main component of a mixture copula. In the empirical risk management application, the practical usefulness of this strategy is exemplified for a set of bivariate portfolios.

**Keywords:** Dependence structures; Copulas; Goodness-of-fit-testing; robustness.

**JEL Classification Numbers:** G11,C13,C15.

**I:ME Insurance Branch Category:** IM12,IE43.

---

\*Address: Otto-Hahn-Str. 6a, D-44227 Dortmund, Germany, telephone: +49 231 755 4608, e-mail: [gregor.weiss@tu-dortmund.de](mailto:gregor.weiss@tu-dortmund.de). I thank participants of the 3rd International Conference on Computational and Financial Econometrics 2009 at the University of Cyprus and seminar participants at the University of Münster and the University of Ulm for their valuable comments on earlier versions of this paper as well as Janet Gabrysch, Sandra Gabrysch and Janina Mühlnickel for their outstanding research assistance. Support by the Collaborative Research Center "Statistical Modeling of Nonlinear Dynamic Processes" (SFB 823) of the German Research Foundation (DFG) is gratefully acknowledged.

# 1 Introduction

Parametric copulas have become the model of choice in statistics for analyzing the dependence structure inherent in multivariate random vectors. Starting with the seminal paper by Embrechts et al. (2002), the advantages and caveats of using copulas instead of linear correlation have now been addressed in several different areas of applied statistics and financial econometrics. During the last years, copulas have been used extensively in a risk management setting where they are being used for the analysis of risky portfolios of stocks (see e.g. Malevergne and Sornette, 2003; Junker and May, 2005; Fischer et al., 2009; Nikoloulopoulos et al., in press) or the modeling of credit default (see Li, 2000).

Recently, research has concentrated on deriving goodness-of-fit (GoF) test statistics in a copula setting (see e.g. Breyermann et al., 2003; Fermanian, 2005; Dobrić and Schmid, 2007; Savu and Trede, 2008; Kole et al., 2007; Genest et al. 2006, 2009, 2011; Genest and Rémillard, 2008; Kojadinovic and Yan, 2011) and comparing these GoF-statistics with each other (see Berg, 2009). While early approaches for copula model selection were based on AIC or variants of it (see e.g. Breyermann et al., 2003; Chen and Fan, 2005; and Huard et al., 2006) recent GoF-tests are based on some comparison of the hypothesized parametric copula and Deheuvels' empirical copula which converges uniformly to the true underlying copula (see Deheuvels, 1978, 1981). Surprisingly, the finite sample properties of these GoF-tests have only recently been analysed in two power studies by Genest et al. (2009) and Berg (2009). Their main result is that most tests perform rather well even for small sample sizes of  $n = 50$ . However, both simulation studies only consider uncontaminated data that is simulated from a prespecified copula model. As real data will usually be contaminated by outliers and measuring errors, the question which (if any) GoF-test for copulas is robust to these contaminations is of high practical importance. Moreover, recent results by Weiß (2011) show that GoF-tests for copulas are more or less unsuited for selecting the optimal parametric copula for the estimation of portfolio Value-at-Risk of Expected Shortfall giving rise to the idea that this lack of power could indeed be caused by a lack of the GoF-tests' robustness to data contaminations.

The literature on robust statistical inference for copulas is still in its infancy. To the best knowledge of the author, the only study that is concerned with the robustness of copula models is Mendes et al. (2007) (recent papers by Kim et al. (2007) and Fantazzini (2009a) focus on the

robustness of copula models to misspecified marginals; however, they do not consider contaminated data). In their work, Mendes et al. (2007) derive robust estimators for the parameters of copulas and analyse their finite sample properties by means of a simulation study. The two strategies they employ to robustify the estimation of copula parameters include (1) the inclusion of weight functions in Minimum-Distance estimators and (2) the identification and exclusion of outliers. The robustness of GoF-tests for copulas, however, has been addressed neither theoretically nor by means of a simulation study.

This paper focuses on filling this gap in the literature by performing a comprehensive simulation study on the robustness of copula GoF-tests as well as the performance of several strategies for robustifying these tests. Thus, this paper contributes significantly in several ways to the current state of the art: First, by analysing the finite sample properties of recently proposed copula GoF-tests when applied to contaminated data, this paper is the first one to answer the question which (if any) copula GoF-test is robust to outliers in the data. Second, this paper is the first to apply outlier detection methods from robust statistics before testing a copula model's goodness-of-fit. Third, five different methods are used for identifying outliers in the contaminated data before re-assessing the copula GoF-tests' power. While Mendes et al. (2007) only use the Donoho-Stahel projection based estimator of multivariate location and scatter in their simulation study, we compare the Minimum Covariance Determinant estimator of Rousseeuw (1985), the Donoho-Stahel estimator, Rocke's constrained M-estimator, the S-estimator based on Tukey's biweight function and the orthogonalized Gnanadesikan-Kettenring estimator with each other regarding their ability to robustify the results of the copula GoF-tests. Fourth, this paper is the first to analyse the effects of mistakenly assuming a single parametric copula when the data actually stems from a mixture copula.

The results presented in this paper show that independent of the underlying true copula, the GoF-test or chosen test statistic, even minor contaminations of the data can lead to a significant decrease in the GoF-tests' power. Conversely, the exclusion of outliers can have a beneficial effect on the power of the GoF-tests. Moreover, this robustification strategy improves the power of GoF-testing when used to identify the main component of a mixture copula. In the empirical risk management application, we show that excluding outliers can correct both false rejections and non-rejections of the copula GoF-tests for the majority of portfolios we consider.

The remainder of this article is structured as follows. Section 2 introduces copula-specific goodness-of-fit-tests. Section 3 discusses the different methods for identifying outliers used in this study. In Section 4 the results of the simulation study are presented and discussed. In Section 5, we conduct an empirical analysis for a set of bivariate financial portfolios. Concluding remarks are given in Section 6.

## 2 Testing the Goodness-of-fit of copula models

The purpose of this section is to shortly restate the outline of several goodness-of-fit tests that have been proposed recently for copula models. For ease of exposition, the formulae are restricted to the bivariate case. Extensions to the multivariate case are straightforward.

Consider a continuous bivariate random vector  $\mathbf{X} \equiv (X_1, X_2)$  with a joint cumulative distribution function (cdf)  $G$  and marginal cdfs  $F_1$  and  $F_2$ . Let  $C(u_1, u_2)$  be the unique copula such that  $G(x_1, x_2) = C(F_1(x_1), F_2(x_2))$  holds. In applications, the unique copula  $C$  is regularly assumed to come from a parametric family

$$\mathcal{C} := \{C_\theta | \theta \in \Theta\}$$

with  $\Theta \subset \mathbb{R}$ .

In goodness-of-fit testing for copula models, we are interested in testing the hypothesis  $H_0 : C \in \mathcal{C}$ , i.e. that the copula  $C$  belongs to a prespecified parametric family  $\mathcal{C}$ . In the following, three different approaches to copula GoF testing are described.

### 2.1 Goodness-of-fit tests based on the empirical copula process

The most intuitive approach to testing the fit of a parametric copula model is based on a comparison between Deheuvels' empirical and the hypothesized parametric copula. Given an i.i.d. sample of size  $n$ , Deheuvels' empirical copula is defined as

$$C_n(\mathbf{v}) \equiv \frac{1}{n} \sum_{i=1}^n \mathbf{1}(u_{i1} \leq v_1, u_{i2} \leq v_2), \quad \mathbf{v} \equiv (v_1, v_2)^T \in [0; 1]^2 \quad (1)$$

with  $\mathbf{1}(\cdot)$  being a logical indicator function (see Genest et al., 2009).

The empirical copula converges uniformly to the true underlying copula (see Deheuvels, 1978 and

1981) and thus constitutes a discontinuous approximation to the true copula. It is nonparametric in nature and goodness-of-fit tests based on computing a distance between the empirical and the hypothesized copula have been shown empirically to perform well in power studies (see Berg, 2009; Berg and Quessy, 2009; Genest et al., 2009).

Given a random sample  $\mathbf{X} = (x_{ij})$  ( $i = 1, \dots, n; j = 1, 2$ ), consider the transformed (pseudo-) sample  $(u_{ij})$  obtained from the original sample via

$$u_{ij} = \frac{n}{n+1} \hat{F}_j(x_{ij}) \quad (2)$$

where  $\hat{F}_j$  is the empirical cdf of the  $j$ th margin (see McNeil et al., 2005). The empirical copula process is then given by

$$\mathbb{C}_n \equiv \sqrt{n}(C_n - C_{\hat{\theta}}) \quad (3)$$

where  $C_n$  is Deheuvels' empirical copula and  $C_{\hat{\theta}}$  is the hypothesized copula from a parametric family parameterized by the parameter estimate  $\hat{\theta}$  obtained from the pseudo-sample  $(u_{ij})$ .

The convergence of  $\mathbb{C}_n$  under appropriate regularity conditions is established in Genest and Rémillard (2008). A simple Cramér-von-Mises test statistic based on  $\mathbb{C}_n$  is then given by

$$\rho_{emp}^{CvM} \equiv \int_{[0;1]^2} \mathbb{C}_n(\mathbf{v})^2 dC_n(\mathbf{v}) \quad (4)$$

with the empirical version of  $\rho_{emp}^{CvM}$  being (see Genest et al., 2009)

$$\hat{\rho}_{emp}^{CvM}(u_{ij}; \theta) \equiv \sum_{i=1}^n \{C_n((u_{i1}, u_{i2})) - C_{\hat{\theta}}((u_{i1}, u_{i2}))\}^2. \quad (5)$$

The asymptotic distribution of the test statistic is unknown and thus approximate p-values need to be obtained via parametric bootstrap (the same applies to the GoF-tests described below). Descriptions for the bootstrap procedures used in this study can be found in Genest et al. (2009). In the simulation study, we employ both Cramér-von-Mises as well as Kolmogorov-Smirnov test statistics.

## 2.2 Goodness-of-fit tests based on Kendall's dependence function

In the second approach for testing the fit of a given parametric copula model, Kendall's probability integral transform is used (see Savu and Tiede, 2008; and Genest et al., 2006). More specifically, Kendall's transform for an arbitrary random vector  $\mathbf{X}$  with joint cdf  $G$  and margins  $F_i$  ( $i \in \mathbb{N}_d$ ) is given by (see Genest et al., 2009)

$$\mathbf{X} \mapsto V = G(\mathbf{X}) = C(\mathcal{U}_1, \dots, \mathcal{U}_d), \quad (6)$$

where the joint cdf of  $\mathcal{U} = (\mathcal{U}_1, \dots, \mathcal{U}_d)$  is  $C$  and  $\mathcal{U}_i = F_i(X_i)$ . Let  $K$  be the cdf of the probability integral transform  $V$ . Then a nonparametric estimation of  $K$  based on the transformed sample  $V_i \equiv C_n((u_{i1}, u_{i2}))$  of size  $n$  is given by (see Genest and Rivest, 1993)

$$K_n(\omega) \equiv \frac{1}{n} \sum_{i=1}^n \mathbf{1}(V_i \leq \omega), \omega \in [0; 1]. \quad (7)$$

If  $\mathcal{U}$  is distributed as  $C_{\hat{\theta}}$ , a parametric estimation of  $K$  is given by the distribution  $K_{\hat{\theta}}$  of the Kendall transform  $C_{\hat{\theta}}(\mathcal{U})$ . Goodness-of-fit tests can then be based on the empirical process

$$\mathbb{K}_n \equiv \sqrt{n}(K_n - K_{\hat{\theta}}). \quad (8)$$

The convergence of the empirical process  $\mathbb{K}_n$  underlying these estimators is established in Genest et al. (2006) under appropriate regularity conditions. Here, we employ the Cramér-von-Mises test statistic

$$\rho_K^{CvM} \equiv \int_0^1 \mathbb{K}_E(\omega)^2 dK_{\hat{\theta}}(\omega) \quad (9)$$

and a respective Kolmogorov-Smirnov test.

## 2.3 Goodness-of-fit tests based on Rosenblatt's transform

Finally, we also employ tests based on Rosenblatt's probability integral transform (see Rosenblatt, 1952). The idea behind this approach for GoF-testing is the result that given the multivariate distribution of a sample, a set of dependent variables can be transformed into a set of indepen-



dent  $U([0; 1])$  variables. Specifically, for a given random vector  $\mathbf{X} \equiv (X_1, X_2)$  with marginal cdfs  $F_j(x_j)$  ( $j = 1, 2$ ) and conditional cdf  $F_{2|1}$ , Rosenblatt's transform of  $\mathbf{X}$  is given by (see Berg, 2009)  $\mathcal{R}(\mathbf{X}) \equiv (\mathcal{R}_1(X_1), \mathcal{R}_2(X_2))$  where

$$\begin{aligned}\mathcal{R}_1(X_1) &\equiv F_1(x_1), \\ \mathcal{R}_2(X_2) &\equiv F_{2|1}(x_2|x_1).\end{aligned}\tag{10}$$

Under a specified parametric null hypothesis copula, a given data sample can be transformed via Rosenblatt's transform. The closer the null copula is to the true underlying copula, the closer the data sample will be to being i.i.d.  $U([0; 1])^2$ . Consequently, Genest et al. (2009) propose to measure the distance between the empirical copula and the independence copula at each element of the transformed data matrix  $\mathbf{V} = (v_{ij}) \equiv (\mathcal{R}_1(u_{i1}); \mathcal{R}_2(u_{i2}))$  which is dependent on the null hypothesis copula  $C_{\hat{\theta}}$ . An empirical version for the Cramér-von-Mises test statistic is then given by

$$\hat{\rho}_{Ros}^{CvM}(\mathbf{V}) = \sum_{i=1}^n \{C_n((v_{i1}, v_{i2})) - C_{\perp}((v_{i1}, v_{i2}))\}^2.\tag{11}$$

where  $(v_{i1}, v_{i2})$  is obtained by transforming the pseudo-sample  $(u_{i1}, u_{i2})$  according to (10) (see Ghoudi and Rémillard, 2004; and Genest et al., 2009; for proofs concerning the asymptotic null behaviour of the underlying empirical process and the convergence of the test statistic).

### 3 Methods for detecting outliers

#### 3.1 Robust estimates of multivariate location and scale

A common approach to outlier identification is to use a metric measuring the distance between possible outliers and the center of the data points. Most commonly, a robust version of the Mahalanobis distance for a data point  $\mathbf{x}_i \in \mathbb{R}^d$  given by

$$MD_i = \sqrt{(\mathbf{x}_i - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu})}\tag{12}$$

where  $\boldsymbol{\mu}$  and  $\boldsymbol{\Sigma}$  are robust estimates of the location vector and covariance matrix respectively of the data set  $\mathbf{X} = (\mathbf{x}_i)_{i=1, \dots, n}$  is used in order to classify outliers.

Various methods have been reported in the literature for finding robust estimates of multivariate location and shape (see e.g. Campbell, 1980; Devlin et al., 1981; Donoho, 1982; Lopuhaä, 1989; Rousseeuw, 1985; Woodruff and Rocke, 1994; Becker and Gather, 1999, 2001). In this study, the Minimum Covariance Determinant (MCD) estimator (Rousseeuw, 1985; Lopuhaä and Rousseeuw, 1991), the constrained M-estimator (Kent and Tyler, 1996), the Stahel-Donoho estimator (Stahel, 1981; Donoho, 1982; Tyler, 1994), the orthogonalized Gnanadesikan-Kettenring (OGK) estimator (Gnanadesikan and Kettenring, 1972; Devlin et al., 1981; Maronna and Zamar, 2002) as well as the S-estimator based on Tukey's biweight function (see Lopuhaä, 1989; Rocke, 1996) are employed. In a copula setting, the only study concerned with the robust estimation of copula parameters is due to Mendes et al. (2007). In their work, only the Donoho-Stahel estimator is used for robustly estimating the location and scale of the multivariate data.

After the robust estimation of the multivariate location and scale, the detection (and subsequent exclusion) of outliers requires some knowledge of the distribution of the estimated Mahalanobis distances in (12). For data samples coming from a multivariate normally distributed random vector, the classic (non-robust) squared Mahalanobis distances based on the sample mean and covariance matrix follows a  $\chi_d^2$ -distribution. If robust estimators of multivariate location and scale are used instead, this result no longer holds thus necessitating the alternative use of a scaled  $F$ -distribution (see Hardin and Rocke, 2005). Both methods, however, seem to be impractical for the purposes of this paper due to two reasons. First, Becker and Gather (2001) show that the  $\chi_d^2$ -distribution is a rather bad approximation to the true distribution of the squared Mahalanobis distances when considering small sample sizes. Second, both approximations by the  $\chi_d^2$ - and the scaled  $F$ -distribution assume the uncontaminated data to be multivariate normal. For most applications in finance and risk management, this assumption regularly does not hold. Also, the assumption of multivariate normally distributed data would contradict the use of a copula model.

As a remedy, we follow Maronna and Zamar (2002) and Filzmoser et al. (2008) by weighting each data point  $i$  in the sample according to  $I(MD_i < MD_0)$  with  $MD_i$  being the Mahalanobis distance corresponding to the  $i$ th data point,  $I(\cdot)$  being an indicator function and  $MD_0$  being

$$MD_0 \equiv \frac{\chi_d^2(\beta)\text{med}(MD_1, \dots, MD_n)}{\chi_d^2(0.5)} \quad (13)$$

where  $\text{med}(\cdot)$  is the median function and  $\chi_d^2(\beta)$  is the  $\beta$ -quantile of the  $\chi_d^2$ -distribution. After excluding data points with  $MD_i \geq MD_0$ , the location and scale of the remaining data points are reestimated and the resulting Mahalanobis distances should then approximately follow a  $\chi_d^2$ -distribution so that the quantile  $\chi_d^2(\beta)$  can be used for detecting outliers in the original non-normal data.

### 3.2 Outlier detection and copula goodness-of-fit testing

When using copulas to model multivariate data, the choice of the parametric copula that fits the data best is of extremely high importance. Though the modeling of the univariate marginals can also have a significant impact on the overall fit of the multivariate model (see Fantazzini, 2009a), the task of finding a suitable copula is usually more demanding than fitting the univariate marginals. Recent simulation studies by Berg (2009) and Genest et al. (2009) suggest that copula goodness-of-fit tests are well suited for the task of identifying the best fitting parametric copula. Extending these results from simulated data to the case of real financial market data, a recent study by Weiß (2011) shows that the examined copula GoF-tests can no longer identify the best fitting parametric copula. One possible reason for this lack of power could be the fact that these GoF-tests are not robust to contaminations in the data.

At the same time, however, the robustification of copula goodness-of-fit tests by means of outlier detection could have severe drawbacks. By detecting and excluding possible outliers from the data sample, some data points in the tails of the multivariate distribution could be falsely excluded from the analysis. As copula modeling aims in particular at modeling the extreme tail behaviour of multivariate distributions, outlier detection could thus prove to be counterproductive. Consequently, there seems to exist a tradeoff between a robust test of the copula model's goodness-of-fit and an adequate modeling of the multivariate tail behaviour of the data. It is this tradeoff which will be of particular interest in both our simulation studies and real data example.

## 4 Simulation study

A large-scale simulation study was conducted on the robustness of goodness-of-fit tests based on the empirical copula process, Rosenblatt's and Kendall's transform for bivariate elliptical and

archimedean copulas. The aim of this simulation study is to compute and compare the nominal level and power of the different approaches for GoF-testing when used on contaminated and uncontaminated data. Especially, we are interested in comparing the losses in power of the three different GoF-tests when the data is contaminated. Conversely, we are interested in answering the question whether any detected loss in the GoF-tests' power can be prevented by identifying and excluding outliers prior to the testing of the copula models' fit.

In this study, two types of data contaminations will be considered. The first type of data contamination will consist of *perturbations* of the data points which are intended to mimic measurement errors. The second type tries to capture the model risk of falsely assuming a single parametric copula when the true dependence structure is given by a mixture copula (i.e. a convex combination of several copulas). This type of data contamination is somewhat closely related to the usual starting point in robust statistics where the majority of the data points is assumed to follow the true distribution with the remaining data points stemming from a contaminating distribution.

## 4.1 Robustness to perturbations of the data

In the first part of the simulation study, we examine the effects of perturbations and outliers in the data on the GoF-tests' power.

### 4.1.1 Design of the simulation study

In the following, the design of the first part of the simulation study is outlined in detail. After the description of the simulation steps, the different choices of parameters are given.

We examine bivariate distributions with one (fixed) marginal being distributed according to a univariate Normal distribution with mean  $\mu = 2$  and standard deviation  $\sigma = 0.5$  and the other (fixed) marginal being t-distributed with 4 degrees of freedom. As candidate parametric copulas we consider the Gaussian, Student's t, Clayton, Frank and Gumbel copula.

For each bivariate parametric copula family parameterised by the (true) parameter(s)  $\theta_k$  repeat the following steps  $K$  times where  $K$  is some large integer:

- (1) Simulate a sample  $(x_{ij})_{i=1,\dots,n;j=1,2}$  of size  $n$  from the joint distribution  $G(x_1, x_2) = C(F_1(x_1), F_2(x_2)|\theta_k)$  with  $\theta_k$  being the copula parameter(s) ( $k = 1, 2, \dots$ ) by first simulat-

ing a sample  $(u_{ij})_{i=1,\dots,n;j=1,2}$  of size  $n$  from the respective copula and then using the quantile functions of the known margins.

(2) Generate a contaminated data sample from the original sample  $(x_{ij})_{i=1,\dots,n;j=1,2}$  in the following way:

- (a) Set  $\epsilon = 0.1$ .
- (b) Generate random numbers  $(\eta_{ij})_{i=1,\dots,n;j=1,2}$  from the interval  $[1 - \epsilon; 1 + \epsilon]$  and compute  $(\tilde{u}_{ij}) := (u_{ij}\eta_{ij})_{i=1,\dots,n;j=1,2}$ . If  $\tilde{u}_{ij} \geq 1$ , set  $\tilde{u}_{ij} = 0.99$ .
- (c) Replace the components of  $n \cdot \epsilon$  pairs of copula observations  $(\tilde{u}_{ij})_{i=1,\dots,n;j=1,2}$  with random numbers from the unit interval yielding the contaminated copula sample  $(\hat{u}_{ij})_{i=1,\dots,n;j=1,2}$ .
- (d) Transform the sample  $(\hat{u}_{ij})_{i=1,\dots,n;j=1,2}$  via the quantile functions of the margins into a sample  $(\check{x}_{ij})_{i=1,\dots,n;j=1,2}$  of the joint distribution.
- (e) Contaminate the sample  $(\check{x}_{ij})_{i=1,\dots,n;j=1,2}$  in the same way as the original copula observations (i.e. scale all observations with a random number  $\tilde{\eta}_{ij} \in [1 - \epsilon; 1 + \epsilon]$  and replace the two components of  $n \cdot \epsilon$  observations with outliers. The two components of the outliers are chosen to be included in the interval  $[\min \{\check{x}_{ij}\} \cdot (1 - \epsilon); \max \{\check{x}_{ij}\} \cdot (1 + \epsilon)]$ ). The result is the final sample of contaminated data  $(\hat{x}_{ij})_{i=1,\dots,n;j=1,2}$ .

(3) Identify and exclude possible outliers in the contaminated data sample by the use of the five methods described above. As a result we get five (possibly different) data samples of different sizes  $\tilde{n}_l$  (with  $l = 1, \dots, 5$  being the index of the respective outlier detection method):

- (a)  $(\bar{x}_{ij}^{(MCD)})_{i=1,\dots,\tilde{n}_1;j=1,2}$  (Minimum Covariance Determinant estimator of Rousseeuw),
- (b)  $(\bar{x}_{ij}^{(DS)})_{i=1,\dots,\tilde{n}_2;j=1,2}$  (Donoho-Stahel estimator),
- (c)  $(\bar{x}_{ij}^{(M)})_{i=1,\dots,\tilde{n}_3;j=1,2}$  (Rocke's constrained M-estimator),
- (d)  $(\bar{x}_{ij}^{(S)})_{i=1,\dots,\tilde{n}_4;j=1,2}$  (S-estimator based on Tukeys biweight function) and
- (e)  $(\bar{x}_{ij}^{(OGK)})_{i=1,\dots,\tilde{n}_5;j=1,2}$  (orthogonalized Gnanadesikan-Kettenring estimator).

(4) Estimate the parameters of the five candidate copulas via canonical maximum-likelihood (see Kim et al., 2007) for the original sample  $(x_{ij})_{i=1,\dots,n;j=1,2}$ , the contaminated sample  $(\hat{x}_{ij})_{i=1,\dots,n;j=1,2}$  as well as the five samples from which outliers have been excluded.

- (5) Compute the Cramér-von-Mises and Kolmogorov-Smirnov test statistics of all three different approaches for goodness-of-fit testing (empirical copula process, Kendall's transform, Rosenblatt's transform) for the five estimated parametric candidate copulas at a significance level of  $\alpha$  percent. Estimate approximate p-values for the tests via parametric bootstrapping.
- (6) Compare the nominal levels and the rejection percentages of the different tests' decisions.

In this first simulation study, the number of simulated samples  $K$  was chosen to be 500. For the GoF-tests, the significance level  $\alpha$  was set to 5%. The procedure outlined above was then repeated

- for different sample sizes  $n$  with  $n \in \{50; 150\}$  to assess the improvement in the GoF-tests' power with increasing sample size and
- two different choices of copula parameters corresponding to Kendall's tau being either  $\tau = 0.25$  or  $\tau = 0.5$  to analyse the influence of the strength of the dependence on the performance of the tests.

To limit the computational complexity, only bivariate copulas were considered as the true copula. The choice of  $n$ ,  $K$  and  $\tau$  are comparable to similar studies like Berg (2009) and Genest et al. (2009).

All computations were performed in R version 2.11.1 on the HPC Compute Cluster of the *RWTH Aachen University* using the procedure *optimise*. For all estimates, the found optima were polished by the additional use of the function *optim*.

Obviously different procedures for generating the contaminations in the data could have also been used. The described perturbations and outliers in the data, however, can be seen as a very general scenario where both measurement errors (the perturbations of the data via the random scaling factors  $(\eta_{ij})_{i=1,\dots,n;j=1,2}$  and  $(\tilde{\eta}_{ij})_{i=1,\dots,n;j=1,2}$  respectively) as well as completely random outliers can distort the original data sample. Moreover, the contaminations are not limited to the sample of the joint distribution but also effect the copula and thus the dependence structure in an additional step of the data contamination.

Examples of the data contamination procedure are shown for a Clayton and a Student's t copula (sample size  $n = 500$ , Kendall's  $\tau = 0.75$ ) in Figure 1.

- insert Figure 1 about here -

From Figure 1 one can see that the contaminated data samples (red triangles) retain the general structure of the dependence structure in the original samples (black circles) while at the same time containing considerably more outliers in the extreme tails.

The joint effect of the contamination of both the simulated copula observations in the first step as well as the observations from the joint bivariate distribution in the second step are illustrated for a Gaussian and a Clayton copula (sample size  $n = 500$ , Kendall's  $\tau = 0.25$ ) in Figure 2.

- insert Figure 2 about here -

Again, one can see from Figure 2 that the contaminated data samples retain the general structure of the original samples with several outliers in the extreme tails of the distribution.

#### 4.1.2 Results

The results of the first simulation study on the level and power of the different GoF-tests and the different data samples are reported in Tables 1-3. The tables report the results for a sample size of  $n = 150$  and  $\tau = 0.50$  for the Cramér-von-Mises statistic for each of the three GoF-testing approaches (empirical copula process, Kendall's and Rosenblatt's transform). Furthermore, the results for the sample size of  $n = 50$  as well as the results for the Kolmogorov-Smirnov test statistics are omitted from the presentation though we comment on them in our discussion of the results. The complete set of results and Tables is available from the author upon request.

- insert Tables 1-3 about here -

In the following, we discuss several aspects of the results in more detail.

Concerning the *influence of the data perturbations*, we can see from Tables 1-3 (columns labeled "C") that the power of all GoF-tests decreases considerably when applied to the contaminated data sets. While all tests are able to hold their nominal level for the uncontaminated samples, the true copula is rejected in 10.8% to 34.4% of the contaminated samples. This finding is true for all five parametric copulas considered with the perturbations having the strongest effect on the samples simulated from the Gumbel copula. At the same time, the rejection rates for the remaining four (false) parametric copulas seem to increase for the majority of true copula/ $H_0$ -copula-pairs. It thus seems that the presence of outliers in both the copula as well as the joint distribution significantly

decreases the power of the GoF-tests. This result is even more striking when considering the fact that the contaminating parameter  $\epsilon = 0.1$  was relatively low.

In contrast to this result, *outlier detection and exclusion* seems to have a beneficial effect on the GoF-tests' power. For all copulas and GoF-tests we can see a considerable decrease in the rejection rates of the true copula. In several cases, the previous rejection rate from the uncontaminated data sample could be matched with the GoF-tests again holding their nominal level. Concerning the rejection rates for the false  $H_0$ -copulas, we can see from Tables 1-3 that the rejection rates for all parametric copula decrease considerably when excluding outliers. Even worse, the rejection rates for the false  $H_0$ -copulas decrease below the previous levels for the uncontaminated data. We can thus see that outlier exclusion can help decrease the type I error of the GoF-testing while at the same time increasing the chance of not rejecting a false  $H_0$ -copula. However, the results also show that the rejection rates for the false  $H_0$ -copulas remain well above 20% for most parametric copulas and GoF-tests.

Further results concerning the *choice of the respective outlier detection method* show that with the exception of the S-estimator, all four remaining methods yield comparable results. For the GoF-test based on the empirical copula process, we can see that the Minimum Covariance Determinant estimator yields rejection rates for the data samples from the  $H_0$ -copula which are slightly lower than the rates for the other outlier detection methods. Besides these minor differences and the poor performance of the S-estimator, however, the choice of the outlier detection method does not seem to have a significant impact on the GoF-tests' power.

Next, we address the question whether we can find different results for the *five types of parametric copulas*. For the Gaussian, Student's t, Clayton and Frank copulas, the results on both the contaminated samples as well as the samples from which outliers have been excluded are almost the same. The only striking fact is that perturbations of the data sample from the Gumbel copula seem to have an extremely negative effect on all three types of GoF-tests. In these cases, outlier detection can reverse some of the loss in the GoF-tests' power. However, the tests cannot hold their nominal level even after outliers have been excluded from the data samples (rejection rates for the true Gumbel copula remain well above 10 – 15% for all three tests and all five outlier detection methods).



From a risk management perspective, it is particularly interesting to analyse the changes in the GoF-tests' power for the Clayton copula (as it is far better suited to model e.g. financial portfolios than the upper tail dependent Gumbel copula). Though one could expect that outlier detection would increase the rejection rate for the Clayton copula as observations in the tails are more likely to be falsely excluded from the data sample, the opposite is true. The results show that the GoF-tests are less likely to falsely reject a true Clayton copula after excluding outliers. While rejecting the previous conjecture that outlier exclusion could delete too much valuable information on the bivariate distribution's tails, the exclusion of outliers comes at the price of an increased probability of misinterpreting the Clayton copula for an elliptical copula.

Finally, unreported results for the different choices for the *sample size and Kendall's tau* yielded results which were in line with our expectation. Both smaller sample sizes as well as a stronger dependence between the random variables as expressed by Kendall's tau led to a stronger influence of both the data contamination as well as the effects of outlier detection on the GoF-tests' performance while the overall power of the tests increased with the sample size.

## 4.2 Robustness to model risk

In the second part of the simulation study, we analyse the contaminating effect caused by mixing observations from a second parametric copula to the original dependence structure. To assess the loss in power of the copula GoF-tests, we consider four different mixture copulas consisting of two parametric copulas with mixture weights 0.9 and 0.1. The first parametric copula (with a respective weight of 0.9) will be considered the *true* copula while the contaminating effect is assumed to be caused by the second parametric copula. The second simulation study thus mimics typical applications in practice where one tries to model a more complex dependence structure with a single parametric copula.

### 4.2.1 Design of the simulation study

As before, the design of the second simulation study is outlined in detail below highlighting the steps of the simulation procedure as well as the parameter choices. We choose the same marginal models as in the first simulation study in order to keep the results from both studies comparable. As mixture copulas, we consider the following models with a mixing parameter  $\theta^{mix} = 0.9$ :

- $C_1^{mix}(u_1, u_2) = \theta^{mix} C^{Clayton} + (1 - \theta^{mix}) C^{Frank}$  (Clayton copula disturbed by a Frank copula)
- $C_2^{mix}(u_1, u_2) = \theta^{mix} C^t + (1 - \theta^{mix}) C^{Clayton}$  (t copula disturbed by a Clayton copula)
- $C_3^{mix}(u_1, u_2) = \theta^{mix} C^t + (1 - \theta^{mix}) C^{Frank}$  (t copula disturbed by a Frank copula)
- $C_4^{mix}(u_1, u_2) = \theta^{mix} C^t + (1 - \theta^{mix}) C^t$  (mixture of two t copulas with different parameters)

Our choice of particular mixture copulas is motivated by the wish to mix different types of tail dependence in order to cover a broad range of possible dependence structures. The simulation study then proceeds as follows:

For each mixture copula parameterized by the parameter(s)  $\theta_k$  and the mixing parameter  $\theta^{mix}$  repeat the following steps  $K$  times where  $K$  is some large integer:

- (1) Simulate a sample  $(x_{ij})_{i=1, \dots, n; j=1, 2}$  of size  $n$  from the joint distribution  $G(x_1, x_2) = C(F_1(x_1), F_2(x_2) | \theta_k, \theta^{mix})$  with  $\theta_k$  being the copula parameter(s) ( $k = 1, 2, 3, 4 \dots$ ) and  $\theta^{mix}$  being the mixing parameter by first simulating a sample  $(u_{ij})_{i=1, \dots, n; j=1, 2}$  of size  $n$  from the respective mixture copula and then using the quantile functions of the known margins.
- (2) Perform step (3) of the first simulation study in order to identify and exclude outliers from the original data sample  $(x_{ij})_{i=1, \dots, n; j=1, 2}$ .
- (3) Estimate the parameters of the five candidate copulas via canonical maximum-likelihood (see Kim et al., 2007) for the original sample  $(x_{ij})_{i=1, \dots, n; j=1, 2}$  as well as the five samples from which outliers have been excluded.
- (4) Perform steps (5) and (6) of the first simulation study for goodness-of-fit testing and a comparison of the nominal levels and the tests' power.

As in the first simulation study, the number of simulated samples  $K$  was chosen to be 500. The procedure outlined above was performed for sample sizes of  $n = 150$  and  $n = 500$ . The true copula in the mixture was parameterized with Kendall's tau being  $\tau = 0.25$  while the parameter(s) of the disturbing copula were chosen according to  $\tau = 0.5$ .

To illustrate the effect of the mixing of two parametric copulas, simulated samples from four mixture copulas are shown in Figure 3.

- insert Figure 3 about here -

As can be seen from the four scatterplots in Figure 3, distinctive features of the disturbing parametric copula in the mixture copula (e.g. tail dependence, high Kendall's tau) can be recognised from the data plots even though the mixing weight  $\theta^{mix}$  is relatively low (note e.g. the concentration of observations in the lower tail of the Student's t-Clayton mixture due to the disturbing copula's lower tail dependence or the concentration of observations around the middle of the distribution in the Student's t-Students' t-mixture due to the disturbing copula's high parameter value).

In contrast to the situation in the first simulation study on data perturbations, neglecting a disturbing parametric copula in a mixture copula does not necessarily have to result in outliers in the sample from the joint distribution. We illustrate this conjecture by simulating two data sets from joint distributions stemming from two mixture copulas. The scatterplots of the data are shown in Figure 4. While the left plot shows a bivariate sample simulated from a Student's t-Clayton mixture copula, the second plot on the right shows a sample distribution with a Student's t-Frank mixture copula (sample size  $n = 500$ , mixing parameter  $\theta = 0.9$  for the Student's t copula in both cases, Kendall's  $\tau = 0.25$  for the Student's t and  $\tau = 0.50$  for both Archimedean copulas).

- insert Figure 4 about here -

The plots given in Figure 4 show that while some outliers are visible in both plots, the general structure of the data is not as severely disturbed as in the first simulation study. It thus seems that the disturbing influence in form of a different parametric copula mixed into the true parametric copula does not change the characteristics of the joint distribution too much. Consequently, the identification and exclusion of outliers should not be able to greatly increase the power of the GoF-tests. On the contrary, outlier exclusion might prove counterproductive as distinctive features in the tails of the data are excluded from the analysis. We investigate this question in more detail in the following.

#### 4.2.2 Results

The results of the second simulation study on the GoF-tests' performance when used on mixture copula data are reported in Tables 4-6. The sample size and Kendall' tau were chosen as to equal

the choices from our first simulation study ( $n = 150$  and  $\tau = 0.50$ ). Results for the three GoF-testing approaches are given in separate tables. As before, we omit the results for a sample size of  $n = 50$  with the complete set of results being available from the author upon request.

- insert Tables 4-6 about here -

In the following, we discuss several aspects of the results in more detail.

Concerning the *influence of the data perturbations*, we can see from Tables 4-6 (columns labeled "C") that the power of all GoF-tests is relatively low for all types of mixture copulas we consider with the exception of the Student's t/Student's t-mixture. In all cases we consider, the tests are not able to hold their nominal level with the true copula being rejected in up to 16.8% of all simulations. Again, results do not differ across the parametric copulas we consider. The strongest negative effect on the GoF-tests' power can be observed for the mixtures of two copulas with different tail dependence behaviour (Clayton/Frank and t/Clayton).

Again, we can see that the *detection and exclusion of outliers* seems to increase the power of the GoF-tests: For all copulas and test approaches, we find significant decreases in the rejection rates of the true mixture copulas. Concerning the rejection rates for the false  $H_0$ -copulas, we can again observe decreasing rejection rates for the false  $H_0$ -copulas after excluding outliers. In some cases (e.g. the Clayton/Frank mixture and the Student's t-copula as the false  $H_0$ -copula), the exclusion of outliers even leads to a drop of the rejection rates down to 6%. Though the false parametric copulas become more likely to not be rejected by the GoF-test after the exclusion of outliers, rejection rates are still the lowest for the true copula.

The *choice of the respective outlier detection method* does not seem to have any significant influence on the outcome of excluding outliers. As all five methods yield comparable results, together with our previous results, the Minimum Covariance Determinant estimator appears to be the method of choice for identifying potential outliers in the data.

The mentioned results also hold for all *four different mixture copulas* we consider. Though small differences can be seen in the results presented in Tables 4-6, our overall assessment of the results given above remains valid for all four mixture copulas.

From a practical point of view, we can see that mixing a second parametric copula to another copula has the undesired effect of a decrease in the power of any of the three GoF-test approaches we

consider. In an empirical setting where dependence structures of financial time series are often too complex to be represented by a simple parametric copula model, mixture copulas offer a promising tool for capturing different facets of the underlying dependence structure. In this situation, however, GoF-tests do not seem to offer much guidance for choosing the major components of the mixture copula. The exclusion of outliers on the other hand comes with a tradeoff between a lower rejection rate of the true copula and (slightly) lower rejection rates for the remaining false copulas.

We will address the trade-off between these two results in our real-case study on financial market data.

## 5 Empirical application

### 5.1 Model

In this section, the practical usefulness of including outlier detection and exclusion methods in the modeling of financial returns is exemplified by forecasting and backtesting the value-at-risk and expected shortfall for a variety of bivariate portfolio. Though more commonly found in engineering application, data contaminations are also conceivable in a risk management setting. For example, Frésard et al. (2011) find that a large fraction of US and international banks uses data that include fees and commissions as well as intraday trading revenues in order to validate their market risk models. We do not analyze this question in detail here, but conjecture that contaminations in our data samples could be present. Moreover, the GoF-tests' power could possibly break down due to few extreme observations in our samples. The benefits and drawbacks of testing the goodness-of-fit of copula models in a robust manner are thus of high importance for risk managers.

We therefore forecast the VaR and ES of the portfolios by the use of copula-GARCH models and repeat the estimations for two data samples: one unaltered data sample including possible outliers and one with outliers excluded. As the results from the previous section did not show any significant qualitative difference between the methods for outlier detection, we employ the MCD-estimator for detecting possible outliers.

A stylized fact about financial data is the presence of conditional heteroscedasticity in financial returns. As the presence of conditional heteroscedasticity could bias the results and as the results for copula models described above only hold for i.i.d. data, we fit GARCH(1,1)-models to each

of the univariate marginals to account for time-varying volatility. Obviously, other more complex specifications for the marginals are possible irrespective of the modelling of the dependence structure. The GARCH(1,1) specification of the marginals is, however, in line with the vast majority of studies on copula models for VaR-estimation (Jondeau and Rockinger 2006; Fantazzini 2009b; Liu and Luger 2009; Aas and Berg 2009; Ausin and Lopes 2010; Hafner and Reznikova 2010). Moreover, results by Hansen and Lunde (2005) show that in many cases a GARCH (1,1) specification cannot be outperformed by more complex models thus underlining the suitability of this study's marginal models (for an analysis of a misspecification of the marginal models on VaR estimates see Fantazzini 2009a).

Then, let  $P_t$  ( $t = 0, 1, \dots, T$ ) be a time series of the prices of a financial asset. We use the log return  $R_t$  defined as  $R_t := \log(P_t/P_{t-1})$  for ( $t \geq 1$ ). We are then interested in modeling the joint distribution of the  $d$  assets with returns  $R_{t1}, \dots, R_{td}$ . As described above, we use GARCH(1,1)-models with Student's t distributed innovations for the marginals thus yielding the model:

$$R_{tj} = \mu_j + \sigma_{tj} Z_{tj}, \quad (14)$$

$$\sigma_{tj}^2 = \alpha_{0j} + \alpha_{1j} R_{t-1,j}^2 + \beta_j \sigma_{t-1,j}^2, \quad j = 1, \dots, d; \quad t = 1, \dots, T, \quad (15)$$

with independent and identically t-distributed innovations  $Z_{tj}$ . Next, we assume the vectors  $\mathbf{Z}_t = (Z_{t1}, \dots, Z_{td})$  ( $t = 1, \dots, T$ ) to be distributed according to

$$F_{\mathbf{Z}}(\mathbf{z}; \nu_1, \dots, \nu_d, \omega) = C [F_1(z_1; \nu_1), \dots, F_d(z_d; \nu_d); \omega] \quad (16)$$

with parameters  $\nu_1, \dots, \nu_d$  for the innovations' distribution and a copula with parameter(s)  $\omega$ .

While the parameters of the univariate GARCH-models are estimated in a standard fashion via Maximum-Likelihood, several different methods are available for the estimation of the copula parameters (see e.g. McNeil et al. 2005 for an overview of possible estimators). Here, the parameters are fitted by canonical maximum-likelihood (ML) using rank-transformed pseudo-observations. Recent simulation studies show that this estimator possesses some empirical optimality over the traditional full-maximum-likelihood or Inference for Margins procedures especially when the marginals are possibly misspecified (Kim et al. 2007).

In the empirical study we will consider equally-weighted portfolios of different assets with portfolio returns  $R_{p,t} = 2^{-1} \sum_{j=1}^2 R_{tj}$ . While the parameters of both the GARCH and the copula models are estimated using the  $T$  observations from the in-sample, the Value-at-Risk and Expected Shortfall of the portfolio are forecasted for day  $T+1$  via Monte Carlo simulation using the following algorithm (for an in-sample version of this algorithm, see Nikoloulopoulos et al., in press):

- For  $K = 10,000$ , simulate  $K$  observations  $u_{T+1,1}^{(k)}, u_{T+1,2}^{(k)}$  ( $k = 1, \dots, K$ ) from the fitted conditional copula.
- Convert  $u_{T+1,j}^{(k)}$  to  $z_{T+1,j}^{(k)}$  ( $j = 1, 2$ ) using the quantile function of the Student-t distribution.
- Transform  $z_{T+1,j}^{(k)}$  into the simulated return  $R_{T+1,j}^{(k)} = \hat{\mu}_j + \hat{\sigma}_{T+1,j} z_{T+1,j}^{(k)}$  where  $\hat{\sigma}_{T+1,j}$  and  $\hat{\mu}_j$  ( $j = 1, 2$ ) are the forecasted conditional volatility and mean values from the previously fitted marginal models.
- Compute the simulated portfolio return as  $R_{p,T+1}^{(k)} = 2^{-1} \sum_{j=1}^2 R_{j,T+1}^{(k)}$ .
- The simulated log returns are then smoothed by kernel density estimation using the Epanechnikov kernel (Pritsker 2006; Alexander and Sheedy 2008) and the VaR at the  $100(1 - \alpha)\%$  confidence level for day  $T+1$  is then computed as the  $\alpha$ -quantile of the kernel density estimate.
- Update the information set with the actual portfolio return  $R_{p,T+1}$ , reestimate all models and forecast the portfolio return for day  $T+2$  and so forth.

In order to formally test the performance of the different models, we use the test of conditional coverage proposed by Christoffersen (1998) and Christoffersen and Pelletier (2004). Approximate p-values for the test of conditional coverage are obtained via Monte-Carlo-simulation.

Moreover, we compute the estimate for the portfolios' sample Expected Shortfall on day  $t$  as the mean of the simulated returns beyond the estimated VaRs in the out-of-sample. The models' Expected Shortfall estimates are then backtested using the sample asymptotics derived by Wong (2008) under a standard normal null hypothesis.

## 5.2 Data

In the empirical study, we use daily returns on five different stock indices in order to achieve relatively heterogenous portfolios with ample opportunities for diversification. The data includes

the logarithmic returns on the DAX 30, S&P 500 Composite, the FTSE 100, the TOPIX and the NIKKEI 225 indices. The data we use is collected from *Thomson Reuters Financial Datastream* over the period September 14, 2006 to April 20, 2011. As portfolios, we consider all bivariate combinations of the five indices with both respective return series entering the portfolio with equal weights.

Excluding non-trading days, the sample consists of  $n = 1,200$  observations of the five indices. The full sample of log returns is then split into half (period 1 and period 2) in order to measure the results' robustness to different phases of market volatility. For each period, daily forecasting of VaR and ES is performed for the latter 100 days in that respective period with the copula-GARCH models being estimated from rolling windows of 500 observations preceding the day for which the portfolio returns are forecasted.

Table 7 presents summary statistics on the returns of the indices in the sample.

- insert Table 7 about here -

Over the sample period, all asset classes exhibit the classical stylised facts on financial market data. Whereas all indices yielded negligible mean log returns on average, the hypothesis of normally distributed log returns can be rejected for almost all return series as indicated by the skewness and fat-tails of the return series.

### 5.3 Results

Results for both the copula GoF-tests as well as the backtests are presented in Table 8.

- insert Table 8 about here -

For simplicity, only the different tests' decisions rather than the corresponding p-values and test statistics are reported in Table 8. The significance level of the VaR and ES estimates was 1%, the backtests as well as the GoF-test were performed at a significance level of 5%.

As a first result, we can see from Table 8 that the quality of the VaR and ES estimates is barely influenced by the exclusion of outliers for almost all portfolios in our sample. In contrast to this result, GoF-testing seems to be considerably affected by outlier exclusion. We find several cases where the use of the MCD-estimator for identifying and excluding outliers changes the copula



GoF-test’s decision (highlighted in bold type in Table 8). In the vast majority of these cases, the GoF-test’s decision is corrected by the exclusion of outliers. In these cases, both (false) rejections and non-rejections of GoF-test performed on the original sample were corrected. Only in one case (DAX/TOPIX, Student’s t copula-model) does the exclusion of outliers lead to a false rejection of a copula model that yielded acceptable VaR and ES results and which was not rejected previously by the GoF-test performed on the original sample. This result shows that the exclusion of outliers prior to the test of a copula model’s goodness-of-fit can aid in the selection of a VaR- or ES-optimal parametric copula family. At the same time, the (possibly incorrect) exclusion of extreme values in the tails of the portfolio returns does not seem to decrease the overall accuracy of the VaR forecasting.

It is interesting to note that the type of parametric copula family does not seem to have an influence on this result. Also, the main result that outlier exclusion can improve the accuracy of copula GoF-testing for VaR and ES forecasting seems to be robust to a change in time periods and thus volatility regimes. Concerning the differences in the results across the ten portfolios we consider, we can see that the results are similar for all portfolios with the exception of the combination of the TOPIX and NIKKEI indices for which no copula model produced acceptable risk forecasts (mainly due to the high linear correlation between the two assets).

In summary, we can see that the exclusion of outliers can considerably improve the identification of a parametric copula family which is optimal for forecasting the VaR and ES of a bivariate portfolio.

## 6 Conclusion

The aim of this study was to analyze the robustness of copula GoF-tests as well as the performance of several strategies for robustifying these tests. By performing two large simulation studies on the power of copula GoF-tests when performed on samples contaminated by either outliers or a contaminating second parametric copula, this study is the first to analyze in detail the robustness of copula GoF-tests to perturbations in the underlying data sample. The practical relevance for risk management purposes was exemplified in the empirical analysis of ten bivariate portfolios.

The results show that independent of the underlying true copula, the GoF-test approach or

chosen test statistic, even minor contaminations of the data can lead to a significant decrease in the GoF-test's power. Conversely, the exclusion of outliers can have a beneficial effect on the power of the GoF-tests. Moreover, the addition of a second parametric copula with even a small weight in the resulting mixture copula can significantly decrease a GoF-test's power. Again, the robustification strategy improves the power of GoF-testing when used to identify the main component of a mixture copula.

These results, however, have to be taken with care. While outlier exclusion decreases the danger of rejecting the true parametric copula, it also lowers rejection rates for the false candidate copulas. The gravity of this trade-off was examined closer in our real-case study on financial market data. In the empirical risk management application, we show that excluding outliers can correct false rejections and non-rejections of the copula GoF-tests for the majority of portfolios we consider. At the same time, the previously found decrease in the rejection rates of the false copulas does not seem to impair the VaR and ES forecasting accuracy.

Further research should concentrate on the question how robust versions of the copula GoF-tests used in this study could be derived. As this study concentrated on the bivariate case, it would also be interesting to answer the question whether the results found in this study also hold in the multivariate case for large portfolios.

## References

- Aas, K., Berg D., 2009. Models for construction of multivariate dependence a comparison study. *European Journal of Finance* 15, 639-659.
- Alexander, C., Sheedy E., 2008. Developing a stress testing framework based on market risk models. *Journal of Banking & Finance* 32, 2220-2236.
- Ausin, M., Lopes, H., 2010. Time-varying joint distributions through copulas. *Computational Statistics and Data Analysis* 54, 2383-2399.
- Becker, C., Gather, U., 2001. The largest nonidentifiable outlier: a comparison of multivariate simultaneous outlier identification rules. *Computational Statistics and Data Analysis* 36, 119-127.
- Becker, C., Gather, U., 1999. The masking breakdown point of multivariate outlier identification rules. *Journal of the American Statistical Association* 94, 947-955.
- Berg, D., 2009. Copula Goodness-of-fit Testing: An overview and power comparison. *European Journal of Finance* 15, 675-701.
- Berg, D., Quessy, J.-F., 2009. Local power analysis of goodness-of-fit tests for copulas. *Scandinavian Journal of Statistics* 36, 289-412.
- Breymann, W., Dias, A., Embrechts, P., 2003. Dependence structures for multivariate high-frequency data in finance. *Quantitative Finance* 3(1), 1-14.
- Campbell, N.A., 1980. Robust Procedures in Multivariate Analysis I: Robust Covariance Estimation. *Applied Statistics* 29, 231-237.
- Chen, X., Fan, Y., 2006. Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. *Journal of Econometrics* 135, 125-154.
- Christoffersen, P., 1998. Evaluating Interval Forecasts. *International Economic Review* 39, 841-862.
- Christoffersen, P., Pelletier, D., 2004. Backtesting Value-at-Risk: A Duration-Based Approach. *Journal of Financial Econometrics* 2, 84-108.
- Deheuvels, P., 1978. Caractérisation complète des Lois Extrêmes Multivariées et de la Convergence des Types Extrêmes. *Pub. l'Institut de Statist. l'Université de Paris* 23, 1-36.
- Deheuvels, P., 1981. A Nonparametric Test for Independence. *Université de Paris, Institut de Statistique*.
- Devlin, S.J., Gnanadesikan, R., Kettenring, J.R., 1981. Robust Estimation of Dispersion Matrices and Principal Components. *Journal of the American Statistical Association* 76, 354-362.
- Dobrić, J., Schmid, F., 2007. A goodness of fit test for copulas based on Rosenblatt's transformation. *Computational Statistics and Data Analysis* 51(9) 4633-4642.

- Donoho, D., 1982. Breakdown Properties of Multivariate Location Estimators. Ph.D. qualifying paper, Harvard University, Dept. of Statistics.
- Embrechts, P, McNeil, A., Straumann, D., 2002. Correlation and dependence in risk management: properties and pitfalls. <http://www.math.ethz.ch/~strauman/preprints/pitfalls.pdf>.
- Engle, R.F., 2002. Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. *Journal of Business and Economic Statistics* 20, 339-350.
- Fantazzini, D., 2009a. The Effects of Misspecified Marginals and Copulas on Computing the Value at Risk: A Monte Carlo Study. *Computational Statistics and Data Analysis* 53, 2168-2188.
- Fantazzini, D., 2009b. Market Risk Management for Emerging Markets: Evidence from the Russian Stock Market, in G. Gregoriou (ed.), *Emerging Markets: Performance, Analysis And Innovation*, 533-554. London: Chapman and Hall / CRC Finance.
- Fermanian, J.-D., 2005. Goodness-of-fit tests for copulas. *Journal of Multivariate Analysis* 95(1), 119-152.
- Filzmoser, P., Maronna, R., Werner, M., 2008. Outlier identification in high dimensions. *Computational Statistics and Data Analysis* 52(3), 1694-1711.
- Fischer, M., Köck, C., Schlter, S., Weigert, F., 2009. An empirical analysis of multivariate copula model. *Quantitative Finance* 7, 839-854.
- Frésard, L., Pérignon, C., Wilhelmsson, A., 2011. The pernicious effects of contaminated data in risk management. *Journal of Banking & Finance* 35, 2569-2583.
- C. Genest, I. Kojadinovic, Nešlehová, J., Yan, J., 2011. A goodness-of-fit test for bivariate extreme-value copulas. *Bernoulli* 17(1), 253-275.
- Genest, C., Rémillard, B., Beaudoin, D., 2009. Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics* 44, 199-213.
- Genest, C., Rémillard, B., 2008. Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Annales de l'Institut Henri-Poincaré* 44, 1096-1127.
- Genest, C., Quessy, J.-F., Rémillard, B., 2006. Goodness-of-fit procedures for copula models based on the probability integral transform. *Scandinavian Journal of Statistics* 33, 337-366.
- Genest, C., Rivest, L.-P., 1993. Statistical inference procedures for bivariate archimedean copulas. *Journal of the American Statistical Association* 88, 1034-1043.
- Ghoudi, K., Rémillard, B., 2004. Empirical processes based on pseudo-observations II: The multivariate case, in: Horváth, L., Szyszkowicz, B. (Eds.), *Asymptotic Methods in Stochastics: Festschrift for Miklós Csörgö*, Fields Institute Communications Series, Vol. 44, American Mathematical Society, pp. 381-406.

- Gnanadesikan, R., Kettenring, J.R., 1972. Robust estimates, residuals, and outlier detection with multiresponse data. *Biometrics* 28, 81-124.
- Hafner, C., Reznikova, O., 2010. Efficient estimation of a semiparametric dynamic copula model. *Computational Statistics and Data Analysis* 54(11), 2609-2627.
- Hansen, P.R., Lunde, A., 2005. A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH(1,1)? *Journal of Applied Econometrics* 20, 873-889.
- Hardin, J., Rocke, D., 2005. The Distribution of Robust Distances. *Journal of Computational and Graphical Statistics* 14(4), 928-946.
- Huard, D., Évin, G., Favre, A.-C., 2006. Bayesian copula selection. *Computational Statistics and Data Analysis* 51(2), 809-822.
- Jondeau, E., Rockinger, M., 2006. The Copula-GARCH model of conditional dependencies: An international stock market application. *Journal of International Money and Finance* 25(5), 827-853.
- Junker, M., May, A., 2005. Measurement of Aggregate Risk with Copulas. *Econometrics Journal* 8, 428-454.
- Kent, J.T., Tyler, D.E., 1996. Constrained M-estimation for multivariate location and scatter. *Annals of Statistics* 24, 1346-1370.
- Kim, G., Silvapulle, M., Silvapulle, P., 2007. Comparison of semiparametric and parametric methods for estimating copulas. *Computational Statistics and Data Analysis* 51, 2836-2850.
- Kojadinovic, I., Yan, J., 2011. A goodness-of-fit test for multivariate multiparameter copulas based on multiplier central limit theorems, *Statistics and Computing* 21(1), 17-30.
- Kole, E., Koedijk, K.C.G., Verbeek, M., 2007. Selecting Copulas for Risk Management. *Journal of Banking & Finance* 31, 2405-2423.
- Li, D., 2000. On Default Correlation: A Copula Function Approach. *Journal of Fixed Income* 9(4), 43-54.
- Liu, Y., Luger, R., 2009. Efficient estimation of copula-GARCH models. *Computational Statistics and Data Analysis* 53(6), 2284-2297.
- Lopuhaä H.P., 1989. On the relation between S-estimators and M-estimators of multivariate location and covariance. *Annals of Statistics* 17, 1662-1683.
- Lopuhaä, H.P., Rousseeuw, P.J., 1991. Breakdown points of affine equivariant estimators of multivariate location and covariance matrices. *Annals of Statistics* 19, 229-248.
- Malevergne, Y., Sornette, D., 2003. Testing the Gaussian Copula Hypothesis for Financial Assets Dependencies. *Quantitative Finance* 3, 231-250.
- Maronna, R., Zamar, R., 2002. Robust Estimation of Location and Dispersion for High-Dimensional Datasets. *Technometrics* 44(4), 307-317.

- McNeil, A., Frey, R., Embrechts, P., 2005. Quantitative Risk Management. Princeton University Press, Princeton.
- Mendes, B., de Melo, E., Nelsen, R., 2007. Robust Fits for Copula Models. *Communications in Statistics - Simulation and Computation* 36(5), 997-1017.
- Nikoloulopoulos, A., Joe, H., Li, H., 2011. Vine copulas with asymmetric tail dependence and applications to financial return data. *Computational Statistics and Data Analysis*, in Press.
- Pritsker, M., 2006. The hidden dangers of historical simulation. *Journal of Banking & Finance* 30, 561-582.
- Rocke, D.M., 1996. Robustness properties of S-estimators of multivariate location and shape in high dimension. *Annals of Statistics* 24, 1327-1345.
- Rosenblatt, M., 1952. Remarks on a multivariate transformation. *The Annals of Mathematical Statistics* 23, 470-472.
- Rousseeuw, P.J., 1985. Multivariate estimation with high breakdown point. In: Grossmann, W., PLug, G., Vincze, I., Wertz, W. (Eds.), *Mathematical Statistics and Applications*. Reidel, Dordrecht, 283-297.
- Savu, C., Trede, M., 2008. Goodness-of-fit tests for parametric families of Archimedean copulas. *Quantitative Finance* 8, 109-116.
- Stahel, W.A., 1981. Breakdown of covariance estimators. Research Report 31, Fachgruppe für Statistik, ETH Zürich.
- Tyler, D.E., 1994. Finite sample breakdown points of projection based multivariate location and scatter statistics. *Annals of Statistics* 22, 1024-1044.
- Weiß, G., 2011. Are Copula-GoF-tests of any practical use? Empirical evidence for stocks, commodities and FX futures. *The Quarterly Review of Economics and Finance* 51(2) 173-188.
- Wong, W.K., 2008. Backtesting trading risk of commercial banks using expected shortfall. *Journal of Banking & Finance* 32, 1404-1415.
- Woodruff, D.L., Rocke, D., 1994. Computable Robust Estimation of Multivariate Location and Shape in High Dimension Using Compound Estimators. *Journal of the American Statistical Association* 89, 888-896.

## Figures

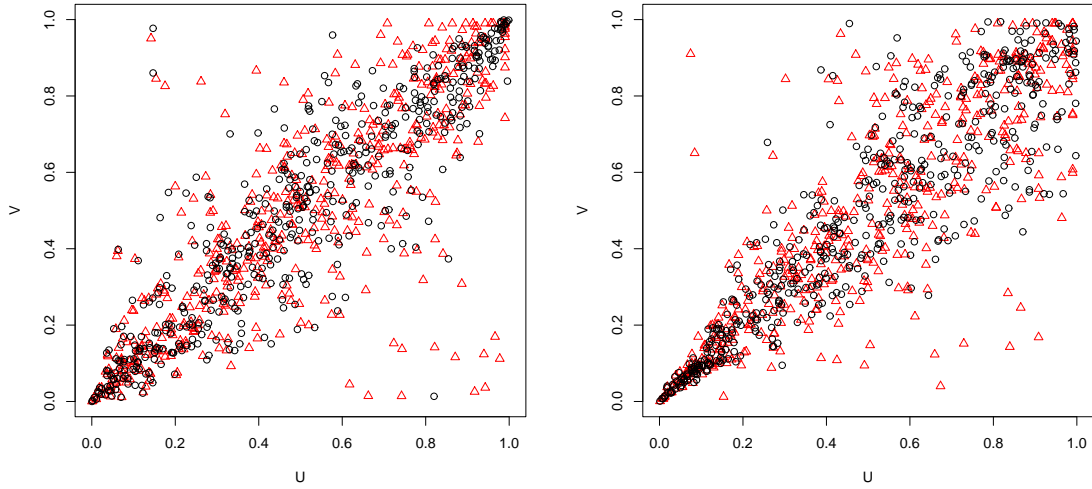


Figure 1: The figure shows two scatterplots for a bivariate Student's t copula (left half,  $\tau = 0.75$ ) and a Clayton copula (right half,  $\tau = 0.75$ ). The data of the original sample are shown in black (circles) while the contaminated data samples are shown in red color (triangles). The sample size is  $n = 500$  for both copulas.

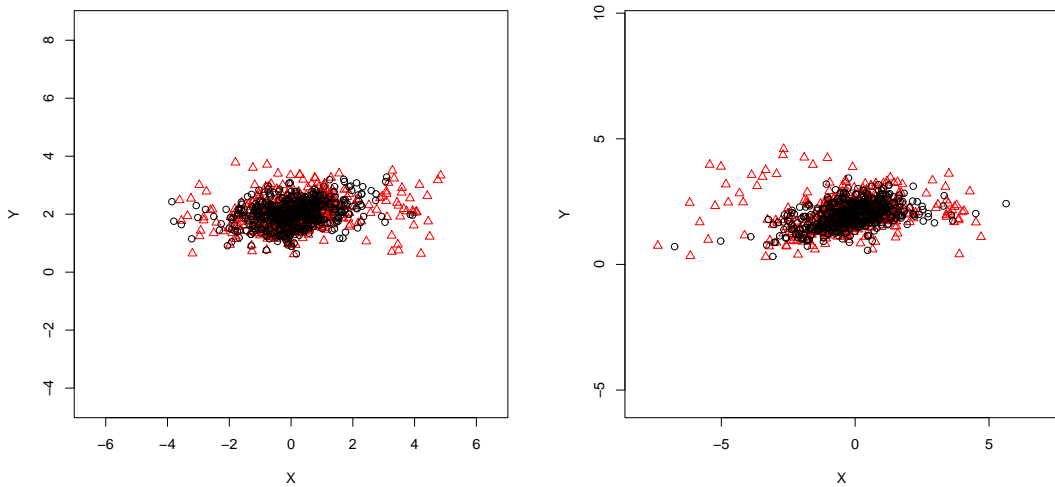


Figure 2: The figure shows two scatterplots for a bivariate distribution with a Gaussian copula (left half,  $\tau = 0.25$ ) and a Clayton copula (right half,  $\tau = 0.25$ ). The data of the original sample are shown in black (circles) while the contaminated data samples are shown in red color (triangles). The marginals are distributed according to  $X \sim N(\mu = 0; \sigma = 2)$  and  $Y \sim t(\nu = 4)$ . The sample size is  $n = 500$  for both distributions.

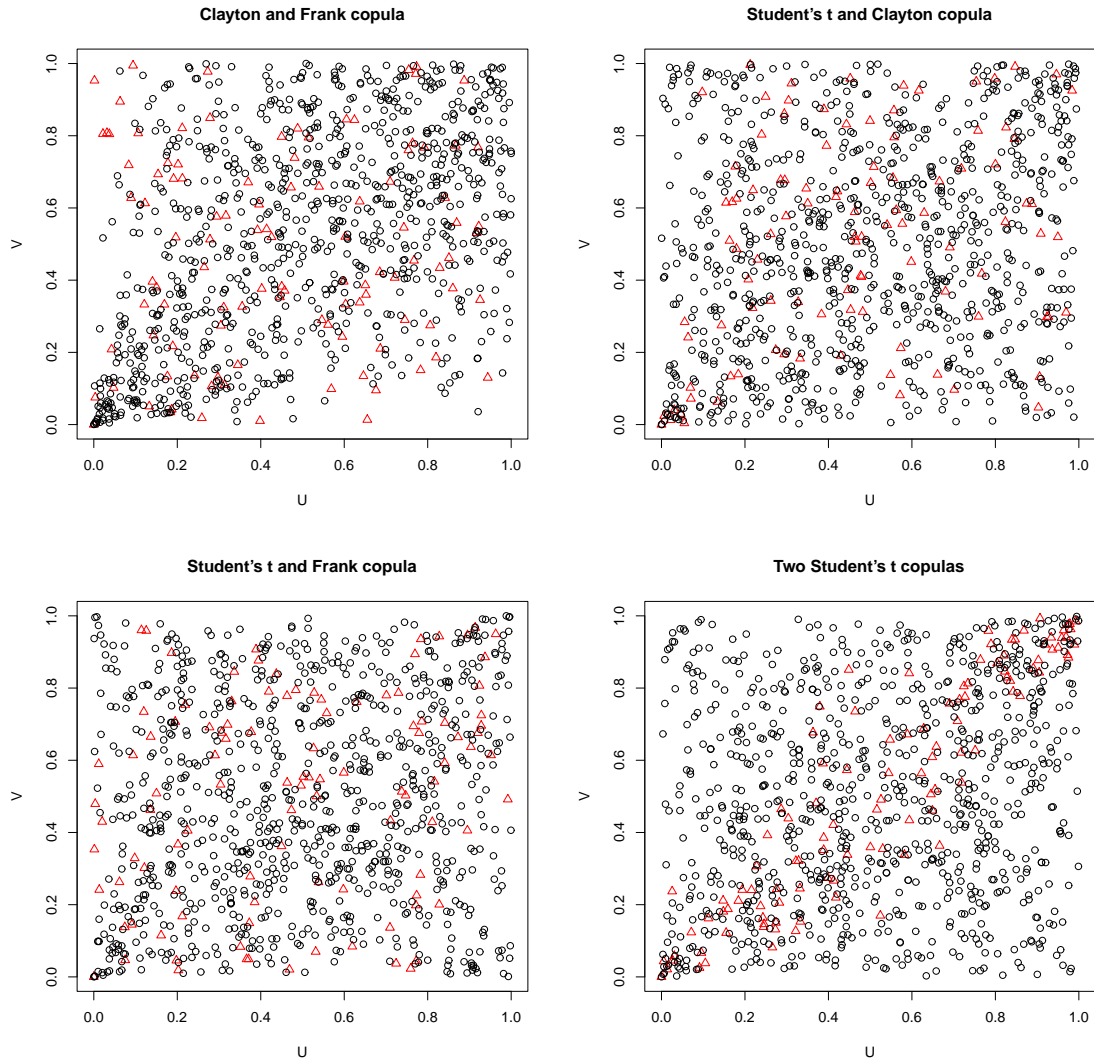


Figure 3: The figure shows four scatterplots for a bivariate mixture copula consisting of a Clayton copula disturbed by a Frank copula ( $\tau = 0.50$  for both copulas; upper left plot), a Student's t copula ( $\rho = 0.2$ ,  $df=3$ ) disturbed by a Frank copula ( $\tau = 0.25$ ; upper right plot), a Student's t copula ( $\rho = 0.2$ ,  $df=3$ ) disturbed by a Clayton copula ( $\tau = 0.25$ ; lower left plot) and a Student's t copula ( $\rho = 0.2$ ,  $df=3$ ) disturbed by a differently parameterized t copula ( $\rho = 0.9$ ,  $df=11$ ; lower right plot). The mixing parameter  $\theta^{mix}$  was chosen to be 0.9. The data of the sample simulated from the "true" copula (i.e. the copula entering the mixture with weight 0.9) are shown in black (circles) while the data samples stemming from the disturbing parametric copula in the mixture are shown in red color (triangles). The sample size is  $n = 500$  for both distributions.



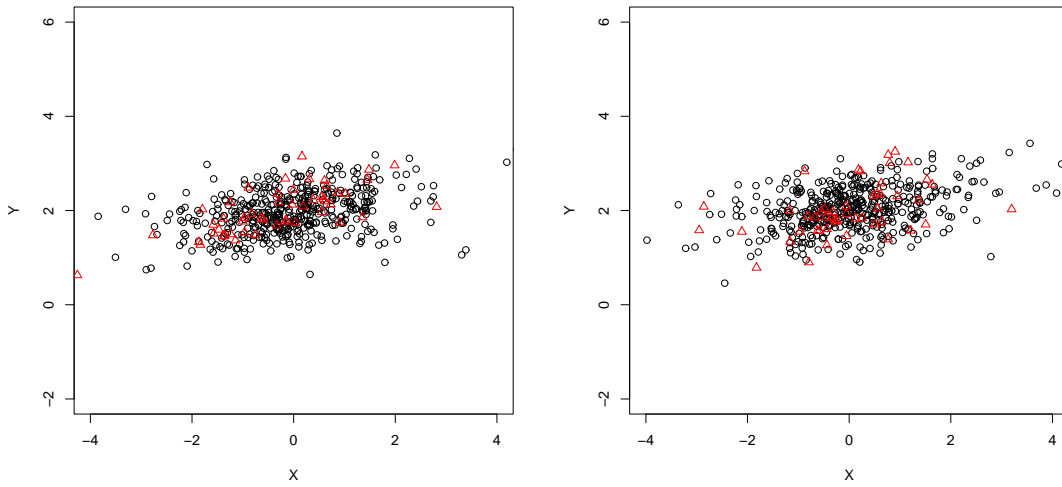


Figure 4: The figure shows two scatterplots for a bivariate distribution with a mixture copula consisting of a Student's t copula ( $\tau = 0.25$ ) and a Clayton copula (left half,  $\tau = 0.50$ ) or a Frank copula (right half,  $\tau = 0.50$ ). The data of the sample simulated according to the Student's t copula are shown in black (circles) while the data samples stemming from the Clayton or Frank copula in the mixture are shown in red color (triangles). The marginals are distributed according to  $X \sim N(\mu = 0; \sigma = 2)$  and  $Y \sim t(\nu = 4)$ . The sample size is  $n = 500$  for all four copulas.

**Table 1** Percentage of rejection of the  $H_0$ -copula by the GoF-test based on the empirical copula process (Cramér-von-Mises test statistic) for the uncontaminated and contaminated data samples. The samples of size  $n = 150$  are simulated from distributions with different parametric copulas with  $\tau = 0.25$ . The data samples are UC=uncontaminated, C=contaminated and the five samples with outliers excluded.

True copula	Copula under $H_0$	Data sample						
		UC	C	MCD	DS	M	S	OGK
Gaussian	Gaussian	<b>3.2</b>	<b>10.8</b>	<b>3.4</b>	<b>4.4</b>	<b>2.2</b>	<b>9.8</b>	<b>6.2</b>
	Student's t	2.4	9.0	3.2	4.6	3.8	4.4	3.4
	Clayton	48.8	42.4	35.8	37.4	36.0	41.8	38.4
	Frank	7.0	11.2	4.6	4.8	3.4	9.2	4.4
	Gumbel	23.2	24.4	18.8	198	20.0	24.6	21.2
Student's t	Gaussian	9.2	20.0	6.2	5.8	8.4	17.6	9.2
	Student's t	<b>3.8</b>	<b>17.8</b>	<b>3.8</b>	<b>4.2</b>	<b>5.4</b>	<b>8.6</b>	<b>8.6</b>
	Clayton	47.2	45.6	37.6	39.0	36.6	43.8	37.4
	Frank	22.8	21.2	7.6	9.4	9.2	18.8	9.8
	Gumbel	20.4	29.6	15.8	20.8	19.6	28.0	20.4
Clayton	Gaussian	41.4	28.2	8.4	12.8	12.4	27.2	15.8
	Student's t	37.8	17.0	9.2	12.0	12.2	14.6	11.4
	Clayton	<b>6.8</b>	<b>16.2</b>	<b>6.8</b>	<b>7.4</b>	<b>6.8</b>	<b>6.0</b>	<b>6.8</b>
	Frank	47.0	33.0	13.4	17.0	16.2	31.4	16.8
	Gumbel	70.4	62.2	47.2	53.0	52.0	61.8	52.2
Frank	Gaussian	7.0	10.8	3.4	4.6	3.0	12.8	5.4
	Student's t	2.4	15.4	3.8	5.6	6.2	8.4	5.8
	Clayton	61.6	59.8	48.4	53.2	52.2	58.2	49.6
	Frank	<b>3.8</b>	<b>17.2</b>	<b>4.2</b>	<b>5.0</b>	<b>4.4</b>	<b>6.2</b>	<b>4.4</b>
	Gumbel	37.0	36.4	25.0	24.4	24.4	36.2	27.2
Gumbel	Gaussian	14.2	28.2	8.6	7.6	9.2	20.6	8.4
	Student's t	16.2	22.0	7.4	9.6	8.8	17.0	8.8
	Clayton	99.8	99.2	97.8	98.6	98.4	99.2	98.0
	Frank	56.2	20.4	15.2	16.2	15.8	18.2	16.6
	Gumbel	<b>5.8</b>	<b>34.4</b>	<b>16.8</b>	<b>17.4</b>	<b>17.6</b>	<b>25.8</b>	<b>18.0</b>

**Table 2** Percentage of rejection of the  $H_0$ -copula by the GoF-test based on Kendall's transform (Cramér-von-Mises test statistic) for the uncontaminated and contaminated data samples. The samples of size  $n = 150$  are simulated from distributions with different parametric copulas with  $\tau = 0.25$ . The data samples are UC=uncontaminated, C=contaminated and the five samples with outliers excluded.

True copula	Copula under $H_0$	Data sample						
		UC	C	MCD	DS	M	S	OGK
Gaussian	Gaussian	<b>1.8</b>	<b>13.6</b>	<b>2.0</b>	<b>4.4</b>	<b>5.8</b>	<b>4.2</b>	<b>5.4</b>
	Student's t	2.0	8.8	2.6	3.2	4.2	5.6	4.0
	Clayton	33.4	23.8	25.8	25.8	26.6	23.8	25.2
	Frank	4.0	5.0	5.4	6.0	4.4	4.4	5.6
	Gumbel	40.0	28.0	33.4	35.8	36.4	27.0	35.4
Student's t	Gaussian	4.6	8.0	6.0	7.4	6.2	7.4	7.4
	Student's t	<b>6.4</b>	<b>17.4</b>	<b>6.4</b>	<b>5.4</b>	<b>5.8</b>	<b>8.2</b>	<b>6.8</b>
	Clayton	20.4	17.4	23.8	21.8	20.2	17.4	21.4
	Frank	6.2	6.0	3.6	4.8	4.4	6.6	3.2
	Gumbel	31.0	28.4	32.0	37.4	35.8	27.0	36.0
Clayton	Gaussian	36.6	22.0	14.4	15.4	12.6	17.8	15.6
	Student's t	41.2	23.0	16.0	13.2	19.4	18.8	15.6
	Clayton	<b>6.0</b>	<b>13.8</b>	<b>11.0</b>	<b>9.2</b>	<b>10.6</b>	<b>4.4</b>	<b>11.0</b>
	Frank	6.0	5.0	3.2	4.4	4.2	4.0	4.0
	Gumbel	93.6	74.8	68.8	69.8	74.2	74.0	70.4
Frank	Gaussian	3.8	4.6	6.4	6.4	6.8	5.8	6.6
	Student's t	7.2	10.8	5.0	7.0	4.6	11.8	9.4
	Clayton	53.4	39.2	40.4	41.4	40.8	38.2	41.0
	Frank	<b>3.4</b>	<b>15.2</b>	<b>6.8</b>	<b>4.8</b>	<b>7.2</b>	<b>5.6</b>	<b>3.0</b>
	Gumbel	41.6	25.8	30.4	34.0	33.4	27.0	35.0
Gumbel	Gaussian	12.0	11.4	7.4	7.4	8.4	10.0	7.4
	Student's t	16.4	16.0	6.8	7.2	9.2	11.8	7.4
	Clayton	98.8	93.4	91.6	90.6	91.0	93.2	91.2
	Frank	5.0	5.4	4.4	4.2	5.2	5.0	4.8
	Gumbel	<b>4.8</b>	<b>18.6</b>	<b>24.8</b>	<b>26.0</b>	<b>27.6</b>	<b>19.0</b>	<b>26.4</b>

**Table 3** Percentage of rejection of the  $H_0$ -copula by the GoF-test based on Rosenblatt's transform (Cramér-von-Mises test statistic) for the uncontaminated and contaminated data samples. The samples of size  $n = 150$  are simulated from distributions with different parametric copulas with  $\tau = 0.25$ . The data samples are UC=uncontaminated, C=contaminated and the five samples with outliers excluded.

True copula	Copula under $H_0$	Data sample						
		UC	C	MCD	DS	M	S	OGK
Gaussian	Gaussian	<b>3.8</b>	<b>23.4</b>	<b>2.4</b>	<b>2.2</b>	<b>2.4</b>	<b>20.4</b>	<b>2.2</b>
	Student's t	4.2	6.6	1.8	2.4	2.2	4.2	2.0
	Clayton	26.4	39.6	15.4	19.4	18.2	38.6	18.4
	Frank	5.2	10.0	4.0	3.6	3.4	7.8	3.4
	Gumbel	19.4	27.4	13.2	16.2	15.0	25.8	16.0
Student's t	Gaussian	15.0	49.4	5.0	6.6	5.0	44.4	6.6
	Student's t	<b>4.4</b>	<b>11.6</b>	<b>3.4</b>	<b>6.2</b>	<b>4.6</b>	<b>10.2</b>	<b>6.0</b>
	Clayton	40.6	58.6	20.0	21.0	19.0	54.4	21.8
	Frank	9.0	24.4	6.4	7.0	7.0	19.2	7.0
	Gumbel	21.2	37.4	13.8	17.6	17.4	34.2	18.6
Clayton	Gaussian	30.8	49.4	10.2	12.2	12.8	43.8	12.4
	Student's t	23.8	15.0	5.8	11.0	12.2	11.8	11.4
	Clayton	<b>4.4</b>	<b>16.4</b>	<b>3.4</b>	<b>5.2</b>	<b>4.6</b>	<b>12.2</b>	<b>5.0</b>
	Frank	28.4	25.6	10.0	12.8	14.0	24.4	13.6
	Gumbel	74.2	61.6	40.6	46.2	46.6	61.0	47.2
Frank	Gaussian	7.8	37.2	2.4	3.6	2.6	31.6	4.4
	Student's t	4.2	9.4	2.8	4.6	3.8	9.0	4.8
	Clayton	40.8	58.0	18.0	20.8	19.0	55.8	19.8
	Frank	<b>4.0</b>	<b>10.8</b>	<b>3.8</b>	<b>2.8</b>	<b>3.0</b>	<b>9.2</b>	<b>3.4</b>
	Gumbel	30.0	35.2	17.4	20.8	19.2	34.2	22.6
Gumbel	Gaussian	7.2	74.2	2.2	3.2	3.0	53.2	3.8
	Student's t	9.8	14.4	3.0	4.8	4.8	9.4	5.6
	Clayton	95.4	97.2	61.2	67.8	66.4	93.2	68.0
	Frank	12.2	25.4	6.4	6.6	7.4	15.6	7.4
	Gumbel	<b>5.8</b>	<b>52.0</b>	<b>10.2</b>	<b>13.4</b>	<b>12.2</b>	<b>38.2</b>	<b>12.6</b>

**Table 4** Percentage of rejection of the  $H_0$ -copula by the GoF-test based on the empirical copula process (Cramér-von-Mises test statistic) for the uncontaminated and contaminated data samples. The mixture weight  $\theta^{mix}$  was chosen to be 0.9. The true copula in the mixture was parameterized with Kendalls tau being  $\tau = 0.25$  while the parameter(s) of the disturbing copula were chosen according to  $\tau = 0.5$ . The samples of size  $n = 150$  are simulated from distributions with different parametric mixture copulas. The data samples are C=contaminated and the five samples with outliers excluded.

True copula	Copula under $H_0$	Data sample					
		C	MCD	DS	M	S	OGK
Clayton/Frank	Gaussian	39.8	11.6	13.4	13.8	27.2	14.2
	Student's t	33.6	12.4	14.8	14.4	29.0	14.2
	Clayton	<b>16.2</b>	<b>9.2</b>	<b>7.8</b>	<b>7.4</b>	<b>6.6</b>	<b>7.2</b>
	Frank	43.4	17.2	20.2	19.4	41.6	20.4
	Gumbel	81.8	54.0	62.0	60.8	79.6	62.4
t/Clayton	Gaussian	11.2	4.0	4.2	4.6	11.0	5.4
	Student's t	<b>15.6</b>	<b>0.6</b>	<b>3.4</b>	<b>4.0</b>	<b>5.4</b>	<b>4.2</b>
	Clayton	40.2	43.2	47.4	45.8	42.6	45.8
	Frank	20.4	8.4	11.6	11.0	18.4	11.6
	Gumbel	29.6	20.0	22.4	22.2	29.0	23.8
t/Frank	Gaussian	16.6	6.4	5.0	5.2	5.2	2.4
	Student's t	<b>11.8</b>	<b>2.0</b>	<b>2.4</b>	<b>5.0</b>	<b>2.8</b>	<b>7.8</b>
	Clayton	57.0	52.2	57.2	54.4	56.4	55.0
	Frank	16.0	6.2	9.0	9.6	15.8	11.2
	Gumbel	21.2	20.8	21.6	23.2	20.2	24.8
t/t	Gaussian	10.6	3.8	4.8	6.4	9.2	5.0
	Student's t	<b>9.6</b>	<b>3.2</b>	<b>5.2</b>	<b>4.2</b>	<b>5.6</b>	<b>4.4</b>
	Clayton	54.4	50.6	54.2	52.8	53.8	53.6
	Frank	21.6	8.0	10.8	11.0	19.2	10.6
	Gumbel	19.6	16.6	18.8	18.2	20.0	19.2

**Table 5** Percentage of rejection of the  $H_0$ -copula by the GoF-test based on Kendall's transform (Cramér-von-Mises test statistic) for the uncontaminated and contaminated data samples. The mixture weight  $\theta^{mix}$  was chosen to be 0.9. The true copula in the mixture was parameterized with Kendall's tau being  $\tau = 0.25$  while the parameter(s) of the disturbing copula were chosen according to  $\tau = 0.5$ . The samples of size  $n = 150$  are simulated from distributions with different parametric mixture copulas. The data samples are C=contaminated and the five samples with outliers excluded.

True copula	Copula under $H_0$	Data sample					
		C	MCD	DS	M	S	OGK
Clayton/Frank	Gaussian	29.8	13.2	16.6	14.4	32.4	15.4
	Student's t	34.6	14.6	14.8	15.8	34.8	17.6
	Clayton	<b>15.0</b>	<b>17.8</b>	<b>15.6</b>	<b>16.2</b>	<b>6.2</b>	<b>16.2</b>
	Frank	16.8	4.2	3.4	6.2	5.0	4.4
	Gumbel	94.2	78.6	81.8	81.2	93.8	83.6
t/Clayton	Gaussian	15.2	2.6	5.2	4.8	8.6	4.4
	Student's t	<b>16.8</b>	<b>4.0</b>	<b>5.6</b>	<b>4.0</b>	<b>7.4</b>	<b>4.4</b>
	Clayton	17.2	26.6	24.4	25.2	18.0	25.8
	Frank	4.6	3.8	5.2	4.0	6.2	4.4
	Gumbel	44.4	42.8	45.0	43.8	45.2	43.6
t/Frank	Gaussian	11.2	1.6	3.8	3.8	6.6	3.6
	Student's t	<b>13.4</b>	<b>4.2</b>	<b>2.2</b>	<b>3.0</b>	<b>6.8</b>	<b>4.6</b>
	Clayton	30.8	35.8	37.2	37.0	27.8	38.6
	Frank	3.4	5.0	6.6	9.6	3.2	3.0
	Gumbel	36.0	37.0	40.2	39.2	32.8	38.8
t/t	Gaussian	13.8	3.2	5.6	5.4	7.0	6.8
	Student's t	<b>8.8</b>	<b>4.6</b>	<b>4.2</b>	<b>4.6</b>	<b>4.6</b>	<b>6.8</b>
	Clayton	23.4	30.8	30.2	30.8	24.6	31.6
	Frank	6.2	6.0	5.2	5.2	5.6	6.4
	Gumbel	31.6	34.6	36.8	37.2	31.8	22.6

**Table 6** Percentage of rejection of the  $H_0$ -copula by the GoF-test based on Rosenblatt's transform (Cramér-von-Mises test statistic) for the uncontaminated and contaminated data samples. The mixture weight  $\theta^{mix}$  was chosen to be 0.9. The true copula in the mixture was parameterized with Kendalls tau being  $\tau = 0.25$  while the parameter(s) of the disturbing copula were chosen according to  $\tau = 0.5$ . The samples of size  $n = 150$  are simulated from distributions with different parametric mixture copulas. The data samples are C=contaminated and the five samples with outliers excluded.

True copula	Copula under $H_0$	Data sample					
		C	MCD	DS	M	S	OGK
Clayton/Frank	Gaussian	33.0	9.8	12.0	10.8	34.0	12.2
	Student's t	19.8	6.0	9.0	7.4	19.8	10.0
	Clayton	<b>15.8</b>	<b>6.8</b>	<b>8.2</b>	<b>8.6</b>	<b>5.6</b>	<b>9.6</b>
	Frank	25.8	9.4	12.8	11.6	24.8	14.0
	Gumbel	75.4	45.2	53.2	52.4	73.6	54.8
t/Clayton	Gaussian	21.6	2.6	3.4	3.0	19.8	4.2
	Student's t	<b>15.0</b>	<b>2.0</b>	<b>2.2</b>	<b>2.4</b>	<b>4.2</b>	<b>2.6</b>
	Clayton	32.4	14.0	18.0	18.2	31.4	20.2
	Frank	11.6	3.4	5.0	4.6	10.4	4.8
	Gumbel	29.8	16.0	17.2	16.4	27.8	18.2
t/Frank	Gaussian	17.8	2.0	4.2	3.8	16.8	3.6
	Student's t	<b>14.6</b>	<b>2.0</b>	<b>3.2</b>	<b>5.4</b>	<b>4.0</b>	<b>3.2</b>
	Clayton	40.2	18.2	25.2	24.8	36.0	23.2
	Frank	12.0	3.8	6.8	5.4	11.4	6.2
	Gumbel	26.4	18.0	21.0	19.6	25.0	21.2
t/t	Gaussian	17.4	2.0	4.4	3.4	16.6	4.4
	Student's t	<b>10.0</b>	<b>2.6</b>	<b>4.0</b>	<b>4.0</b>	<b>6.0</b>	<b>4.0</b>
	Clayton	44.2	22.4	25.0	24.2	42.0	27.0
	Frank	10.6	4.8	6.8	6.4	11.2	7.4
	Gumbel	23.2	12.8	15.2	15.8	22.6	16.0

**Table 7** Summary statistics for the log return series of the five stock indices.

	DAX 30	S&P 500	FTSE 100	TOPIX	NIKKEI 225
<i>Summary statistics</i>					
Mean	0.000171	0.000010	0.000019	-0.000531	-0.000411
Std. dev.	0.015680	0.016022	0.014894	0.016710	0.018162
Minimum	-0.074334	-0.094695	-0.092655	-0.100070	-0.121110
Maximum	0.107974	0.109572	0.093843	0.128646	0.132345
Kurtosis	7.6740	8.2569	6.9138	7.7527	8.7044
Skewness	0.2096	-0.2370	-0.0789	-0.3682	-0.5583
<i>Bravais-Pearson correlations</i>					
DAX 30	1	0.6337	0.8785	0.3247	0.3456
S&P 500		1	0.5731	0.1153	0.1215
FTSE 100			1	0.3533	0.3681
TOPIX				1	0.9744
NIKKEI 225					1

**Table 8** The table presents the results for the CvM-GoF test based on the empirical copula process as well as the backtesting decisions from the test of conditional coverage (Value-at-Risk) and the test by Wong (2008; Expected shortfall). Results are given separately for the two time windows and the ten bivariate portfolios considered. The confidence level for both the VaR and ES was 1%. For all three types of tests, the test decisions are given based on a significance level of 5%. Rejection (non rejection) of a test is symbolized by a  $\times$  ( $\checkmark$ ). All tests were performed on both the original sample (orig.) as well as the sample with outliers excluded (MCD) by the use of the MCD-estimator. Whenever the GoF-test produced conflicting decisions for the two samples, the corresponding line in the table was highlighted in bold type.

**Time window one: 14/09/2006-31/12/2008**

Copulas	DAX/S&P						DAX/FTSE						DAX/TOPIX					
	GoF		VaR		ES		GoF		VaR		ES		GoF		VaR		ES	
	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD
Gauss	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
Student's t	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\times$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
Clayton	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\times$	$\times$	$\times$	$\times$	$\checkmark$
Frank	$\times$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\times$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
Gumbel	$\times$	$\times$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\times$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\times$	$\times$	$\times$	$\checkmark$	$\times$

Copulas	DAX/NIKKEI						S&P/FTSE						S&P/TOPIX					
	GoF		VaR		ES		GoF		VaR		ES		GoF		VaR		ES	
	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD
Gauss	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
Student's t	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
Clayton	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\checkmark$	$\times$	$\times$	$\times$	$\times$	$\times$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
Frank	$\times$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\times$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
Gumbel	$\times$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\times$	$\times$	$\times$	$\times$	$\times$	$\checkmark$	$\times$	$\times$	$\times$	$\times$	$\times$

Copulas	S&P/NIKKEI						FTSE/TOPIX						FTSE/NIKKEI					
	GoF		VaR		ES		GoF		VaR		ES		GoF		VaR		ES	
	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD
Gauss	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
Student's t	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\times$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
Clayton	$\times$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\times$	$\times$	$\times$	$\times$	$\times$	$\checkmark$	$\times$	$\times$	$\times$	$\times$	$\times$
Frank	$\times$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\times$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
Gumbel	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$

Copulas	TOPIX/NIKKEI					
	GoF		VaR		ES	
	orig.	MCD	orig.	MCD	orig.	MCD
Gauss	$\checkmark$	$\times$	$\times$	$\times$	$\times$	$\times$
Clayton	$\checkmark$	$\checkmark$	$\times$	$\times$	$\times$	$\times$
Student's t	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$
Frank	$\checkmark$	$\times$	$\times$	$\times$	$\times$	$\times$
Gumbel	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$



Table 8 (continued).

Time window two: 1/1/2009-20/04/2011

Copulas	DAX/S&P						DAX/FTSE						DAX/TOPIX					
	GoF		VaR		ES		GoF		VaR		ES		GoF		VaR		ES	
	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD
Gauss	×	×	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Student's t	✓	×	×	×	×	×	✓	✓	✓	✓	✓	✓	✓	✓	×	✓	✓	✓
Clayton	✓	×	×	×	×	×	✓	×	✓	✓	✓	✓	✓	✓	×	✓	✓	✓
Frank	×	×	✓	✓	✓	✓	✓	×	✓	✓	✓	✓	×	✓	✓	✓	✓	✓
Gumbel	×	✓	✓	✓	✓	✓	×	×	✓	✓	✓	✓	×	✓	✓	✓	✓	✓

Copulas	DAX/NIKKEI						S&P/FTSE						S&P/TOPIX					
	GoF		VaR		ES		GoF		VaR		ES		GoF		VaR		ES	
	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD
Gauss	✓	✓	✓	✓	✓	✓	×	×	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Student's t	✓	✓	✓	×	✓	✓	✓	×	×	×	×	×	✓	✓	✓	✓	✓	✓
Clayton	×	✓	✓	✓	✓	✓	×	×	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Frank	×	✓	×	✓	✓	✓	×	×	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Gumbel	✓	×	×	×	×	×	×	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Copulas	S&P/NIKKEI						FTSE/TOPIX						FTSE/NIKKEI					
	GoF		VaR		ES		GoF		VaR		ES		GoF		VaR		ES	
	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD	orig.	MCD
Gauss	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	×	✓	✓	✓	✓	✓
Student's t	✓	×	×	×	×	×	✓	✓	✓	✓	✓	✓	×	✓	✓	✓	✓	✓
Clayton	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	×	×	×	×	×
Frank	✓	×	×	×	×	×	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Gumbel	✓	✓	✓	✓	✓	✓	✓	×	×	×	×	×	✓	×	✓	✓	✓	✓

Copulas	TOPIX/NIKKEI					
	GoF		VaR		ES	
	orig.	MCD	orig.	MCD	orig.	MCD
Gauss	✓	×	×	×	×	×
Clayton	✓	×	×	×	×	×
Student's t	×	×	×	×	×	×
Frank	✓	×	×	×	×	×
Gumbel	×	×	×	×	×	×





