Erasmus University Rotterdam
Erasmus School of Economics

Standard Errors in Semiparametric Copula-based Univariate Time Series Models

***B**achelor Thesis - Quantitative Finance*

Tobias Hoogteijling
408976

Supervisor: J.W.N. Reuvers

Second assessor: R. Paap

July 7, 2019

Abstract

This paper investigates the estimation of standard errors in semipametric univariate copula-based time series models. In such models, the marginal distributions are estimated with the Empirical Distribution Function (EDF), which needs to be taken into account in the computation of (correct) MSML standard errors of copula parameter estimates. Patton (2012) compares MSML standard errors to 'naive' standard errors, which ignore this estimation of the marginal distribution, and attributes the difference between the naive and MSML standard errors to this naivety. We show that 'naive' standard errors are not only naive in this aspect, but also ignore autocorrelation in observations and possible misspecification. We define the alternative Truly Naive standard error, which only ignores the use of the EDF, and the Doubly Naive standard error, which only ignores the use of the EDF and the autocorrelation. For three data sets and a simulation study, we find that the Truly Naive standard errors are (almost) equal to the correct MSML standard errors. We conclude that any difference between the naive and MSML standard errors cannot be attributed to estimation of the marginal distribution.

Table of Contents

1 Introduction **1**

2 Literature review **3**

2.1 Copulas in broader economic context 4

3 Data **6**

4 Methodology **7**

4.1 General copula models 7

4.2 Univariate semiparametric copula-based time series models 7

4.3 Multi Stage Maximum Likelihood estimation 8

4.4 Naive copula parameter standard errors 11

4.5 Empirical verification 12

4.6 Rank autocorrelations 14

4.7 Estimation 15

4.8 Monte Carlo Simulation 16

5 Results **17**

5.1 Rank autocorrelations 17

5.2 Parameter estimation 18

5.3 Monte Carlo simulation 22

6 Conclusion **25**

A Appendix **31**

A.1 Plots of index values 31

A.2 Plots of scores 32

A.3 Matlab code 33

1 Introduction

Recently, more and more research has been devoted to the use of copula-based models for economic time series. A copula is a function that describes the dependence relation between random variables. It has been shown by Sklar (1959) that any joint n-dimensional distribution can be decomposed into an n-dimensional copula function and n univariate distributions, and also reversely, that we can easily create a valid joint distribution using a combination of any range of marginal distributions and any copula. An advantage of the use of copulas is that a very wide range of different joint distributions can be created. For the analysis of a created joint distribution, we can draw upon the preponderance of knowledge of the marginal distributions that have been used to create the joint distribution (Patton, 2012). Furthermore, decomposing a joint distribution into a copula and some marginal distributions allows for the use of different estimation methods in different parts of the model (Patton, 2012). A class of copula models that makes use of this feature is called semiparametric models. They use a nonparametric model based on the Empirical Distribution Function (EDF), for the estimation of the marginal distributions, but a parametric model for the copula. The estimation of the parameter(s) is then often done with maximum likelihood, and is usually referred to as Canonical Maximum Likelihood Estimation (Patton, 2012). This estimation is generally done in two stages (the estimation of the marginal distribution using the EDF precedes the copula estimation) and is therefore also referred to as Multi Stage Maximum Likelihood (MSML). The fact that the marginal distribution is estimated needs to be incorporated in the computation of the copula parameter standard errors.

A specific application for such semiparametric copula models is the stationary first-order Markov chain. Instead of modelling the dependence between two different random variables, in this case the copula models the dependence between two succeeding observations in a univariate time series. An empirical illustration is given by Patton (2012), who considers the S&P 100 index returns over the period 1995-2011. He considers four different copula models (Normal, Student's t, Half Rotated Clayton and Half Rotated Gumbel) and three different estimates of the standard errors of their copula parameters. One of those estimates is the MSML standard error, which takes the estimation error due to the use of the EDF into account, and is therefore asymptotically correct. The second, so-called 'naive' estimate, is simply based on the inverse of the Hessian of the likelihood function. Those estimates are naive, because they ignore the estimation error due to the use of the EDF to estimate the marginal distributions. Thirdly, the standard errors are estimated using 1000 simulations of the copula model, following the method proposed by Rémillard (2017), which also provides asymptotically correct standard errors.

This paper takes a critical stand on the results in Patton (2012). First, we cast doubt on the standard errors of the rank autocorrelations reported by Patton, by obtaining those in three different ways and arriving at three sets of standard errors that are very similar to each other, but different from the ones reported by Patton. Secondly, we consider the estimation of standard errors of copula parameters. We provide a derivation of the naive standard error and show that they are 'naive' in more than one way. Not only do they ignore the estimation error induced by the estimation of the marginal distributions with the EDF, they also ignore autocorrelation in the residuals and possible misspecification of the copula. To determine the impact of each of these issues, we define two alternative standard errors. The Truly Naive standard error only ignores the use of the EDF to estimate the marginal distribution, but not possible misspecification or autocorrelation in the observations. Any difference between the Truly Naive standard error and the MSML standard error is thus solely attributable to ignoring the use of the EDF. We also define the Doubly Naive standard error, which ignores both the use of the EDF and possible autocorrelation, but not possible misspecification. Differences between the Doubly Naive standard error and the Truly Naive standard error are thus attributable to autocorrelation in the returns, and differences between the Doubly Naive standard error and the naive standard error to misspecification.

Thirdly, we compute the naive standard errors, the Truly Naive standard errors, the Doubly Naive standard errors and the correct MSML standard errors for the S&P 100 index returns. We find that the Truly Naive standard errors are a lot closer (and often even equal) to the MSML standard errors than the naive standard errors, implying that the difference between the naive standard errors and the MSML standard errors cannot, contrarily to the interpretation by Patton, be attributed to the ignoring the use of the EDF. Furthermore, it appears that the two copulas that do not show signs of misspecification hardly show a difference between Doubly Naive and naive standard errors. This supports our claim that differences between the naive and the correct MSML standard errors have three causes. Fourthly, we extend our analysis to the Belgian Bel 20 and the Japanese Nikkei 225 indices and again find that the Truly Naive standard errors are much closer to the MSML standard errors than the Doubly Naive and naive standard errors. This indicates that estimating the marginal distribution has little impact on the size of the standard errors. Finally, we simulate data from our four copulas, for four different parameter values and for 6 different sample sizes. We again find that almost all Truly Naive standard errors are equal to the MSML standard errors.

We conclude that estimation error due to the use of the EDF has very little impact on the size of the standard errors, and that differences between the naive standard error and the

MSML standard error should be attributed to other factors. In this paper, the following research question will be investigated:

What is the effect of ignoring the estimation error resulting from the EDF on the size of copula parameter standard errors?

The research is done according to the following sub questions:

1. *To which factors can differences between the naive and the MSML standard errors be attributed?*
2. *Can we replicate Patton's empirical findings on naive and MSML standard errors?*
3. *How do the Truly Naive and Doubly Naive standard errors compare to naive standard errors and MSML standard errors?*
4. *Can similar patterns in standard errors be found in other times series?*
5. *Do the patterns in the standard errors also show in a Monte Carlo simulation study?*

The remainder of this paper is organised as follows. Section 2 provides a review of relevant literature and a discussion of the use of copulas in broader economic perspective. The data are described in Section 3 and the methodology in Section 4. The results are discussed in Section 5 and the conclusion is found in Section 6.

2 Literature review

Copulas are a major field of research in economics, statistics and econometrics. Joe (1997) and Nelsen (2003) provide extensive discussions and background information on the use of copulas. A general survey of the use of copulas in academic literature is given by Jaworski et al. (2010) and Choroś et al. (2010). Copulas are most commonly used in the estimation of multivariate, often bivariate, time series. It has been well established in the literature that the conditional volatility in economic time series is time varying (see e.g. Engle (1982) and Bollerslev (1986)). However the GARCH model that fits univariate time series with changing volatility well, is notoriously hard to generalize to multivariate time series. Copulas are a standard tool to model the dependence between variables with GARCH-like marginal distributions. Such an approach is called fully parametric, as both the copula and the marginal distributions are estimated parametrically. MSML estimation is the common approach for fully parametric copula models, and is discussed in e.g. Joe and Xu (1996) and Newey and McFadden (1994). Joe (2005) and

Patton (2006) perform simulation studies to investigate the efficiency of MSML estimation and find that its efficiency is often not far from one-stage efficient estimation.

Semiparametric estimation of copula-based time series models makes use of the fact that estimation of the marginal distributions can be done separately from the estimation of the copula. They use a nonparametric approach for the estimation of the marginal distributions, usually the Empirical Distribution Function, and a parametric specification of the copula. Although the marginal distributions are not specified nonparametrically, their conditional means and variances are often still modelled parametrically. The distribution of the copula parameter estimator is investigated by Genest and Favre (2007) and Chan et al. (2009). Chen and Fan (2006a) show that the asymptotic variance of the Maximum Likelihood estimator of the copula parameter does not depend on estimation error in the conditional means and variances of the marginal distributions, but only on estimation error resulting from the use of the EDF. Patton (2012) empirically shows that in the multivariate case, the naive standard errors that ignore this estimation error are indeed too small compared to the correct MSML or simulation standard errors. More empirical research is provided by Patton (2013). Nonparametric estimation of multivariate copula models is studied by for instance Genest and Rivest (1993) and Genest et al. (2011).

In this paper, we investigate the use of copulas to model a stationary first-order Markov chain. Instead of modelling the dependence between different variables, we model the dependence between consecutive observations. A fully parametric approach is discussed in Joe (1997). The semiparametric approach is discussed by Beare (2010), Beare (2012), Patton (2012), Patton (2013) and Bouyé and Salmon (2013). Chen and Fan (2006b) provide a method for the computation of the MSML standard errors and show that the naive standard errors of the copula parameter estimators are incorrect and relate those to the correct MSML standard errors. Our paper focuses on the research in Patton (2012). The author empirically studies the univariate copula-based time series model using daily data from the S&P 100 Index. The author finds empirical evidence that in the univariate case, in stark contrast to the multivariate case, the naive standard errors, based on the inverse Hessian of the likelihood, are not very different from the correct MSML or simulation errors.

2.1 Copulas in broader economic context

Copulas are widely used in financial modelling (McNeil et al., 2005), most often to model the dependence between two economic time series[1]. Copulas are very suitable for such purposes,

[1]Copulas of higher order are more complicated to model and bring about various kinds of issues. There has however been an increasing amount of research in this direction, most of which involving pairing bivariate copulas

because they allow for the separate estimation of the individual distributions of two random variables, and the copula that describes how those relate. For example, it has been well established in the literature that Generalized Autoregressive Conditional Heteroskedascity (GARCH) models are very suitable to model the movements of univariate time series, but that the generalization to a Multivariate GARCH model is hard (McNeil et al., 2005). Copulas can be used to model the relation between two different variables, whilst modelling their individual distributions as a GARCH model. Moreover, copulas can easily be used to create a very wide range of valid joint distributions by combining two known marginal distributions with a copula function. For the analysis of such multivariate distributions we can then make use of the known properties of used marginal distributions (Patton, 2012).

A specific example of the use of copulas is the modelling of tail dependence. It has been well established in the literature that the correlation between economic variables need not be constant (Hartmann et al., 2004). For instance, in times of economic downturn, all stock prices tend to move down together. The phenomenon that in extreme cases (the tails of the distributions), the dependence between variables becomes stronger is referred to as tail dependence (McNeil et al., 2005). Copulas can be used to model tail dependence, which can even be asymmetric (for example, stock prices mostly tend to show left tail dependence, but not so much right tail dependence).

In the early 00's, David Li introduced copulas to the modelling of Collateralized Debt Obligations (CDOs). A CDO is a financial derivative that can be created by pooling together various collateral-backed loans (usually mortgages), restructuring them and slicing them in smaller tranches. Each such a tranch is then called a CDO. Li used the Gaussian copula for the pricing of CDOs, a method that was quickly adapted by many financial institutions. Unfortunately, like was the case with many other financial instruments in the years preceding the financial crisis of 2008-2009, very few people that used copulas for pricing CDOs fully understood the model (Jones, 2009). Even with copula models, the pricing of CDOs was still based on historical data. When house prices stopped increasing in the summer of 2008, the copula models suddenly failed to describe the risks of securities well and, like all markets, the CDO market collapsed. The fact that the use of the Normal Copula for the pricing of CDOs had been so widespread is the reason why copulas are sometimes referred to as 'the formula that killed Wall Street' (MacKenzie and Spears, 2014). Since the crisis, copulas are applied to financial modelling with more care. To take into account the changing circumstances in financial markets, more and more research is devoted to time-varying copulas. According to

in smart ways. For examples, see Daul et al. (2003) or Smith et al. (2012).

Patton (2012), time-varying copulas are one of the two most promising frontiers of copula research, together with higher dimensional copulas.

3 Data

Following Patton (2012), we consider the returns of the S&P 100 index of largest US companies, and the S&P 600 Index of stocks with smaller capitalization, during the period of 17 August, 1995 till 30 May, 2011. Furthermore, we also consider the Japanese Nikkei 225 Index (NI) and the Belgian Bel 20 Index (Bel) in this same period. The reason why we use these specific indices, is that they show a first order dependence in the returns (which is shown in Section 5.1). The data are obtained from Yahoo Finance and descriptive statistics can be found in Table 1. Plots of the values of the four indices can be found in Appendix A.1.

Table 1: Descriptive statistics of daily log returns for the S&P 100, S&P 600 Nikkei 225 and Bel 20 indices for the period August 1995 till May 2011

	S&P 100	**S&P 600**	**Nikkei**	**Bel**
Mean ($\times 10^{-4}$)	2.02	3.38	-2.29	1.49
Min	-0.092	-0.117	-0.102	-0.083
Max	0.107	0.081	0.099	0.093
St. dev.	0.013	0.014	0.018	0.013
Skewness	-0.151	-0.303	-0.152	0.027
Kurtosis	10.03	7.97	6.98	9.32
Jarque-Bera*	8204	4148	1416	6674
Stocks	100	600	225	20
Observations	3975	4081	4042	4094

**All significant at 0.001%*

The S&P 100, S&P 600 and Bel 20 Indices show an average positive rate of return, whereas the returns are on average negative for the Nikkei 225, reflecting the stagnation of the Japanese economy during the 90s and 00s. Figure 2 shows that the index returns follow global economic trends, displaying large declines during the dot-com crash in the early 00's and the financial crisis of 2008-2009. The S&P and Nikkei indices show the small negative skewness that stock returns typically exhibit, but Bel 20's skewness is remarkably close to zero. This difference follows from the fact that, as can be observed in Figure 2, the Bel 20 Index does not have a lot of large negative returns, which lead to negative skewness. Nonetheless, all four index returns are clearly not normally distributed, with kurtoses much higher than 3 and Jarque-Bera statistics that reject normality at 0.001%. The volatilities of the four indices do not differ much, the standard deviation of the Nikkei 225 being the highest with 0.018 and that of the S&P 100 and Bel 20 being the lowest with 0.013. Note that we do not need to models these skewness and kurtosis explicitly, as they are characteristics of the marginal distribution and will be captured by the Empirical Distribution Function.

4 Methodology

4.1 General copula models

It was shown by Sklar (1959) that an n-dimensional joint distribution can always be decomposed into n marginal distributions and an n-dimensional copula. Let $\mathbf{X} = (X_1, \ldots, X_n)^T$ be a random vector with cumulative distribution function F and let F_i denote the marginal distribution of X_i, $\forall i \in \{1, \ldots, n\}$. The copula C is a mapping of the univariate marginal distributions $F_1, \ldots, F_n$ to the joint distribution F. That is, we have $C : [0,1]^n \to [0,1]$ such that $\forall\ \mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$:

$$F(\mathbf{x}) = C\ \{F_1(x_1), \ldots, F_n(x_n)\}.$$

Let the probability integral transforms of the original variables be defined as $U_i = F_i(X_i)$. Then $\mathbf{U} = (U_1, \ldots, U_n)^T \sim C$. If X_i is continuous, $U_i \sim \mathcal{U}(0,1)$ (Wackerly et al., 2014), so the copula function is a joint distribution with uniform univariate marginal distributions. The reverse of Sklar's theorem is also true, i.e. we can obtain valid joint distributions by combining univariate distributions with a copula. When decomposing a joint distribution in a copula and some marginal distributions, it is possible to use different estimation methods for the copula and the marginal distributions.

4.2 Univariate semiparametric copula-based time series models

Semiparametric copula models make use of the possibility to employ different estimation techniques for the copula and the marginal distributions. They estimate the marginal distribution nonparametrically using the Empirical Distribution Function (EDF), but estimate the copula parametrically. Such an estimation approach cannot only be used to model the dependence between different random variables, but also to model the dependence between consecutive observations in a univariate time series. We consider a stationary first order Markov Chain and denote it by $\{Y_t\}$. Instead of directly specifying $H(Y_{t-1}, Y_t)$, the joint distribution of Y_{t-1} and Y_t, we separately specify the marginal distribution of Y_t, $G^*(Y_t)$, and the copula d. To take possible misspecification into account, we estimate d with $C(\cdot, \cdot; \alpha^*)$, where α^* is the pseudo-true copula parameter. By 'pseudo-true' we mean that α^* gets $C(\cdot, \cdot; \alpha)$ 'as close' to d as possible, in terms of log likelihood (Geyer, 2003). If the model is correctly specified, α^* reduces to α. Note that we now have $U_t = G^*(Y_t) \sim \mathcal{U}(0,1)$. The advantage of this decomposition is that we can freely and separately specify the marginal properties of Y_t (such as skewness), and the temporal dependence between consecutive observations. The conditional distribution of Y_t on Y_{t-1} can be written as

$$h^*(y_t|y_{t-1}) = g^*(y_t)c(G^*(Y_{t-1}), G^*(Y_t); \alpha^*), \tag{1}$$

where $c(\cdot, \cdot; \alpha^*)$ is the copula distribution of $C(\cdot, \cdot; \alpha^*)$ and $g^*(\cdot)$ the probability density function corresponding to $G^*(\cdot)$ (Chen and Fan, 2006b).

4.3 Multi Stage Maximum Likelihood estimation

The estimation of the copula parameter is usually done using maximum likelihood, then called Canonical Maximum Likelihood or Multi Stage Maximum Likelihood (MSML) (Patton, 2012). We will now derive the MSML standard errors of the copula in the same way as Chen and Fan (2006b), p9-11. We also make use of the derivations provided in Geyer (2003). The log likelihood function of the conditional density function in Equation 1 is given by:

$$\begin{aligned} L(\alpha) &= \log H^*(Y_t|Y_{t-1}) \\ &= \frac{1}{n}\sum_{t=1}^{n} \log g^*(Y_t) + \frac{1}{n}\sum_{t=2}^{n} \log c(G^*(Y_{t-1}), G^*(Y_t); \alpha). \end{aligned}$$

In the semiparametric approach, we estimate the unknown $G^*(\cdot)$ by the empirical distribution function $G_n(\cdot)$:

$$G_n(y) = \frac{1}{n+1}\sum_{t=1}^{n} I\{Y_t < y\}, \tag{2}$$

where $I(\cdot)$ is the indicator function. We can ignore the first term in Equation 2, as it does not depend on α. When replacing $G^*(\cdot)$ with $G_n(\cdot)$, we obtain the estimator $\hat{\alpha}$ of α^*:

$$\hat{\alpha} = \text{argmax } L_n(\alpha),$$

with

$$L_n(\alpha) = \frac{1}{n}\sum_{t=2}^{n} \log c(G_n(Y_{t-1}), G_n(Y_t); \alpha).$$

Let $\nabla L_n(\alpha)$ be the gradient of $L_n(\alpha)$ and $\nabla^2 L_n(\alpha)$ its Hessian. Then the Taylor expansion states that

$$\nabla L_n(\hat{\alpha}) \approx \nabla L_n(\alpha^*) + (\hat{\alpha} - \alpha^*)^T \nabla^2 L_n(\alpha^*). \tag{3}$$

Let, following the notation of Chen and Fan (2006b):

$$
\begin{aligned}
l(v_1, v_2; \alpha) &\equiv \log\, c(v_1, v_2; \alpha), && (4)\\
l_\alpha(v_1, v_2; \alpha) &\equiv \frac{\partial l(v_1, v_2; \alpha)}{\partial \alpha}, && (5)\\
l_{\alpha,\alpha}(v_1, v_2; \alpha) &\equiv \frac{\partial^2 l(v_1, v_2; \alpha)}{\partial \alpha \partial \alpha'}, && (6)\\
l_{\alpha,j}(v_1, v_2; \alpha) &\equiv \frac{\partial^2 l(v_1, v_2; \alpha)}{\partial \alpha \partial v_j} \quad \forall j = 1, 2.. && (7)
\end{aligned}
$$

As $\hat{\alpha}$ is the Maximum Likelihood estimate, the First Order Condition states that the left side of Equation 3 equals 0. We can thus rearrange to arrive at:

$$
\begin{aligned}
(\hat{\alpha} - \alpha^*)^T \nabla^2 L_n(\alpha^*) &\approx -\nabla L_n(\alpha^*)\\
\Rightarrow (\hat{\alpha} - \alpha^*) &\approx -(\nabla^2 L_n(\alpha^*))^{-1} \nabla L_n(\alpha^*)\\
\Rightarrow \sqrt{n}(\hat{\alpha} - \alpha^*) &\approx -(\frac{1}{n}\nabla^2 L_n(\alpha^*))^{-1} \frac{1}{\sqrt{n}} \nabla L_n(\alpha^*)\\
&= -(\frac{1}{n}\nabla^2 L_n(\alpha^*))^{-1} \frac{1}{\sqrt{n}} \sum_{t=2}^{n} l_\alpha(G^*(Y_{t-1}), G^*(Y_t); \alpha^*).
\end{aligned} \tag{8}
$$

However, we have now not yet incorporated the fact that we estimate $G^*(\cdot)$ by $G_n(\cdot)$. Chen and Fan (2006b) show that this done by adding two extra terms to the gradients:

$$
\begin{aligned}
W_1(U_{t-1}) &\equiv \int_0^1 \int_0^1 [I\{U_{t-1} \leqslant v_1\} - v_1] l_{\alpha,1}(v_1, v_2; \alpha^*) c(v_1, v_2; \alpha^*)\, dv_1\, dv_2, && (9)\\
W_2(U_t) &\equiv \int_0^1 \int_0^1 [I\{U_t \leqslant v_2\} - v_2] l_{\alpha,2}(v_1, v_2; \alpha^*) c(v_1, v_2; \alpha^*)\, dv_1\, dv_2. && (10)
\end{aligned}
$$

Then (8) becomes:

$$
\begin{aligned}
\sqrt{n}(\hat{\alpha} - \alpha^*) &\approx -(\frac{1}{n}\nabla^2 L_n(\alpha^*))^{-1} \frac{1}{\sqrt{n}} \sum_{t=2}^{n} [l_\alpha(G_n(Y_{t-1}), G_n(Y_t); \alpha^*) + W_1(G_n(Y_{t-1})) + W_2(G_n(Y_t))]\\
&= -(\frac{1}{n}\nabla^2 L_n(\alpha^*))^{-1} \frac{1}{\sqrt{n}} \sum_{t=2}^{n} [l_\alpha(U_{t-1}, U_t; \alpha^*) + W_1(U_{t-1}) + W_2(U_t)].
\end{aligned} \tag{11}
$$

The terms $W_1(U_{t-1})$ and $W_2(U_t)$ that are introduced in Equation 9 and 10 correct for the fact that $G^*(\cdot)$ is unknown and is estimated with $G_n(\cdot)$. If distribution $G^*(\cdot)$ would be known, those terms would disappear from A_n^*. By the Law of Large Numbers,

$$
\begin{aligned}
-\frac{1}{n}\nabla^2 L_n(\alpha^*) &\xrightarrow{p} -\frac{1}{n} E(L_n(\alpha^*))\\
&= -E(l_{\alpha,\alpha}(U_{t-1}, U_t; \alpha^*)).
\end{aligned}
$$

We define B to equal this last term:

$$B = -E(l_{\alpha,\alpha}(U_{t-1}, U_t; \alpha^*)), \tag{12}$$

and we also define

$$A_n^* = \frac{1}{n-1} \sum_{t=2}^{n} [l_\alpha(U_{t-1}, U_t; \alpha^*) + W_1(U_{t-1}) + W_2(U_t)]. \tag{13}$$

If we now combine Equations 8, 12 and 13, we obtain:

$$\sqrt{n}(\hat{\alpha} - \alpha^*) \approx B^{-1} \sqrt{n} A_n^*. \tag{14}$$

Furthermore, let

$$\Sigma = \lim_{n \to \infty} Var(\sqrt{n} A_n^*). \tag{15}$$

Then by the Central Limit Theorem:

$$\sqrt{n}(\hat{\alpha} - \alpha^*) \xrightarrow{d} N(B^{-1} \Sigma B^{-1}) \tag{16}$$

However, B is unknown and has to be estimated. We use the sample estimate $\hat{B}$:

$$\hat{B} = -\frac{1}{n} \sum_{t=2}^{n} l_{\alpha,\alpha}(U_{t-1}, U_t; \hat{\alpha}).$$

To estimate a consistent estimator $\hat{\Sigma}$, we need to take into account that there is autocorrelation in the series U_t. We thus use a Heteroskedasticity & Autocorrelation Consistent (HAC) estimator of the variance as proposed by Chen and Fan (2006b), page 15. We use the Newey-West standard errors (Newey and West, 1986):

$$\hat{\Sigma} = \hat{Var}^{NW}\left(\frac{\sqrt{n}}{n-1} \sum_{t=2}^{n} [l_\alpha(U_{t-1}, U_t; \hat{\alpha}) + W_1(U_{t-1}) + W_2(U_t)]\right), \tag{17}$$

where $\hat{Var}^{NW}$ denotes the Newey-West Covariance Estimator. Consequently the MSML variance of $\hat{\alpha}$ is estimated as

$$\hat{Var}(\tilde{\alpha}) = \frac{1}{n} \hat{B}^{-1} \hat{\Sigma} \hat{B}^{-1}.$$

4.4 Naive copula parameter standard errors

There are three important things to note concerning the correct MSML standard errors. First of all, the terms terms $W_1(U_{t-1})$ and $W_2(U_t)$ in Equation 11 are introduced to correct for the fact that the marginal distributions are estimated with the EDF. Secondly, U_t and U_{t-1} are correlated, and then generally so are $\nabla l_n(U_{t-1}, U_t, \hat{\theta}_n)$ and $\nabla l_n(U_t, U_{t+1}, \hat{\theta}_n)$. This implies we need to use a Heteroskedasticity and Autocorrelation Consistent estimate of this variance. We, following (Patton, 2012), use Newey-West standard errors. Thirdly, the MSML standard errors are also correct under misspecification of the copula.

Patton (2012) defines the so-called naive variance as an alternative to the MSML standard error, which is simply the inverse of the Hessian of the log likelihood function[2]:

$$\begin{aligned} \hat{Var}^{Naive}(\tilde{\alpha}) &= \frac{1}{n}\hat{B}^{-1} \\ &= [\sum_{t=1}^{n} l_{\alpha,\alpha}(U_{t-1}, U_t; \hat{\alpha})]^{-1}. \end{aligned}$$

This naive variance is derived in Patton (2013), p 22-24. According to Patton, the 'naive' standard errors are naive, because they ignore the error resulting from estimating the marginal distributions with the EDF[3]. In other words, they ignore the terms $W_1(U_{t-1})$ and $W_2(U_t)$ in Equation 11. However, to derive the naive standard errors, we need not only ignore this issue, but all three issues mentioned above.

First, we ignore $W_1(U_{t-1})$ and $W_2(U_t)$ and obtain

$$A_n^{*Naive} = \frac{1}{n-1}\sum_{t=2}^{n} l_\alpha(U_{t-1}, U_t; \alpha^*). \tag{18}$$

and, using that the scores are zero on average[4]:

$$\begin{aligned} \Sigma^{Naive} =& \lim_{n\to\infty} \mathrm{Var}(\frac{\sqrt{n}}{n-1}\sum_{t=2}^{n} l_\alpha(U_{t-1}, U_t; \alpha^*)) \\ =& \lim_{n\to\infty} E[(\frac{\sqrt{n}}{n-1}\sum_{t=2}^{n} l_\alpha(U_{t-1}, U_t; \alpha^*)) - E(\frac{\sqrt{n}}{n-1}\sum_{t=2}^{n} l_\alpha(U_{t-1}, U_t; \alpha^*))] \\ & [(\frac{\sqrt{n}}{n-1}\sum_{t=2}^{n} l_\alpha(U_{t-1}, U_t; \alpha^*)) - E(\frac{\sqrt{n}}{n-1}\sum_{t=2}^{n} l_\alpha(U_{t-1}, U_t; \alpha^*))]^T \\ =& \lim_{n\to\infty} E[(\frac{\sqrt{n}}{n-1}\sum_{t=2}^{n} l_\alpha(U_{t-1}, U_t; \alpha^*))(\frac{\sqrt{n}}{n-1}\sum_{t=2}^{n} l_\alpha(U_{t-1}, U_t; \alpha^*))^T] \end{aligned} \tag{19}$$

[2]Patton (2012), p 11, last paragraph
[3]Patton (2012), p 11, last paragraph
[4]First Order Condition of maximum likelihood estimation

Then we ignore the correlation between U_t and U_{t-1} and instead of taking a HAC variance estimate, we can omit the multiplication cross terms of Equation 21:

$$\begin{aligned}\Sigma^{Naive} &= \lim_{n\to\infty}\frac{n}{(n-1)^2}\sum_{t=2}^{n} E[l_\alpha(U_{t-1},U_t;\alpha^*))(l_\alpha(U_{t-1},U_t;\alpha^*))^T] \\ &= E[l_\alpha(U_{t-1},U_t;\alpha^*))(l_\alpha(U_{t-1},U_t;\alpha^*))^T]\end{aligned} \tag{20}$$

The Information Matrix Equality then states that, if the model is correctly specified, the expected value of the outer product of the scores equals the negative of the expected value of the Hessian, see Papadopoulos (2014), Equation 2. This implies

$$\begin{aligned}\Sigma^{Naive} &= E[l_\alpha(U_{t-1},U_t;\alpha^*))(l_\alpha(U_{t-1},U_t;\alpha^*))^T] \\ &= -E[l_{\alpha,\alpha}(U_{t-1},U_t;\alpha^*)] \\ &= B.\end{aligned} \tag{21}$$

From this it follows that

$$\begin{aligned}Var^{Naive}(\tilde{\alpha}) &= \frac{1}{n}B^{-1}\Sigma^{Naive}B^{-1} \\ &= \frac{1}{n}B^{-1}BB^{-1} \\ &= \frac{1}{n}B^{-1}.\end{aligned}$$

Note that this simplification only holds if the model is correctly specified.

4.5 Empirical verification

We have shown that there are three issues with naive standard errors that make them different from the correct MSML standard errors. We can verify our claim by constructing two sandwich estimators that suffer from only one or two of those issues respectively, and comparing those to the MSML and naive standard errors.

First of all, we define new standard errors that are only 'naive' in the sense that they ignore estimation error due to the EDF, but not the dependence between consecutive returns or possible misspecification, and refer to those as the Truly Naive standard errors. The difference between the Truly Naive standard errors and the MSML standard errors can solely be attributed to the use of the EDF. The Truly Naive covariance matrix can obtained using the usual sandwich expression in Equation 4.3, after dropping $W_1(U_{t-1})$ and $W_2(U_t)$ from Equation 17. We define:

$$\begin{aligned}
Var^{TNaive}(\tilde{\alpha}) &= \frac{1}{n} B^{-1} \Sigma^{TNaive} B^{-1} \\
\Sigma^{TNaive} &= \lim_{n\to\infty} Var(\sqrt{n} A_n^{*TNaive}) \\
A_n^{*TNaive} &\equiv \frac{1}{n-1} \sum_{t=2}^{n} l_\alpha(U_{t-1}, U_t; \alpha^*).
\end{aligned}$$

We can estimate this consistently as

$$\begin{aligned}
\hat{Var}^{TNaive}(\tilde{\alpha}) &= \frac{1}{n} \hat{B}^{-1} \hat{\Sigma}^{TNaive} \hat{B}^{-1} \\
\hat{\Sigma}^{TNaive} &= \hat{Var}^{NW}(\frac{\sqrt{n}}{n-1} \sum_{t=2}^{n} l_\alpha(U_{t-1}, U_t; \hat{\alpha})).
\end{aligned}$$

Note that any difference between the Truly Naive standard errors and the MSML standard errors can always solely be attributed to the use of the EDF. If $l_\alpha(U_{t-1}, U_t; \alpha^*)$ and $l_\alpha(U_{s-1}, U_s; \alpha^*)$ are independent and the model is correctly specified, the Truly Naive standard error (asymptotically) reduces to Patton's naive standard error. If the Truly Naive standard errors are different from the naive standard errors, this is an indication that the naive standard errors are naive in more aspects than solely the use of the EDF.

Secondly, we define standard errors that are naive both in ignoring the use of the EDF and in ignoring autocorrelation in consecutive observations, and refer to those as Doubly Naive standard errors. The Doubly Naive covariance matrix can obtained using the usual sandwich expression in Equation 4.3, after dropping $W_1(U_{t-1})$ and $W_2(U_t)$ from Equation 17, and not using a HAC variance estimator of Σ. We define:

$$\begin{aligned}
Var^{DNaive}(\tilde{\alpha}) &= \frac{1}{n} B^{-1} \Sigma^{DNaive} B^{-1} \\
\Sigma^{DNaive} &= \lim_{n\to\infty} Var(\sqrt{n} A_n^{*DNaive}) \\
A_n^{*DNaive} &\equiv \frac{1}{n-1} \sum_{t=2}^{n} l_\alpha(U_{t-1}, U_t; \alpha^*).
\end{aligned}$$

We can estimate this consistently as

$$\begin{aligned}
\hat{Var}^{DNaive}(\tilde{\alpha}) &= \frac{1}{n} \hat{B}^{-1} \hat{\Sigma}^{DNaive} \hat{B}^{-1} \\
\hat{\Sigma}^{DNaive} &= \frac{1}{n} \sum_{t=2}^{n} (l_\alpha(U_{t-1}, U_t; \hat{\alpha}))(l_\alpha(U_{t-1}, U_t; \hat{\alpha}))^T.
\end{aligned}$$

The difference between Truly Naive and Doubly Naive standard errors is that the Truly Naive standard errors do not assume that observations are uncorrelated. This means that if there is no autocorrelation in the observations, the Truly Naive and Doubly Naive standard errors asymptotically have the same values. Furthermore, the difference between the Doubly Naive standard errors and the naive standard errors is that the Doubly Naive standard errors do not assume that the model is correctly specified. If we know that the model is not misspecified, we the Doubly Naive and naive standard errors asymptotically have the same values. Note that the original naive standard errors can now be seen as 'the most naive' or 'triply naive'.

4.6 Rank autocorrelations

For a time series to be suitable to be modeled as a first order Markov Chain, it should exhibit a first order dependence. To confirm that our data sets have this dependence, we compute the rank autocorrelations of the first 10 lags. Furthermore, we also need to know if the dependence relation to be modeled is positive or negative[5]. To obtain standard errors for the autocorrelations, we use the stationary block bootstrap proposed by Politis and Romano (1994). This stationary bootstrap maintains the temporal dependence properties in the data set, unlike the common *iid* bootstrap. The data set is split up in blocks, and a bootstrapped sample is created by randomly shuffling the blocks of data, without changing the order within the blocks. This corresponds to the following pseudo code. An ordinary block bootstrap has a fixed block length, but the Politis & Romano bootstrap has stochastic block sizes, following a Geometric distribution. Like Patton (2012), we use $B = 1000$ replications and an average block length of $w = 30$ observations. The following pseudo code describes how we obtain a $T \times B$ matrix I of shuffled time indices. Each bootstrap sample then is a sequence of the observations of the original data that corresponds to these time indices.

Algorithm 1 Stationary Block Bootstrap

```
procedure STATIONARY BLOCK BOOTSTRAP
    Set I(1,j) to uniformly random numbers between 1 and T for J = 1,...,B
    for j = 1,...,B do:
        for t = 2,...,T do:
            Set I(t,j) to I(t−1,j)+1 with probability 1/w and to a uniformly random number
             between 1 and T with probability 1 − 1/w
    Set I to I modulo T
```

To replicate the results as closely as possible, we have, to a certain extent, made use of the Matlab toolbox provided by the author[6]. For comparison, we have also obtained two other codes

[5]The Clayton and Gumbel copula only facilitate a positive relation. To accommodate a negative relation, they require a half rotation, which amounts to replacing U_t by $1 - U_t$ or U_{t-1} by $1 - U_{t-1}$ (Patton, 2012).

[6]To be found at `https://public.econ.duke.edu/~ap172/code.html`. This toolbox contains code used for a different paper (Patton, 2013)

of the stationary bootstrap, one provided in the Financial Econometrics toolbox of Sheppard[7], and the other provided by Mathworks[8]. These bootstraps have again been done with an average block length of 30 and 1000 replications.

4.7 Estimation

We consider the following 4 copulas:

$$C^{Normal}(u,v) = \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v)),$$

where Φ is the joint cumulative bivariate normal distribution function with correlation coefficient ρ and Φ^{-1} is the inverse cumulative normal distribution;

$$C^{Clayton}(u,v) = [\max(u^{-\kappa} + v^{-\kappa} - 1, 0)]^{-\frac{1}{\kappa}};$$

$$C^{Gumbel}(u,v) = \exp(-((-ln(u))^{\kappa} - ln(v))^{\kappa})^{\frac{1}{\kappa}});$$

$$C^{T}(u,v) = T_{\rho,\nu}(T_\nu^{-1}(u), T_\nu^{-1}(v)),$$

where T is the joint cumulative bivariate t distribution function with correlation coefficient ρ and degrees of freedom ν and T_ν^{-1} is the inverse cumulative t distribution with ν degrees of freedom. If our time series show evidence of a first order dependence relation, we can model them as a copula-based time series. Those indices with a negative rank autocorrelation should be modeled with a negative dependence relation between consecutive returns. The Normal and Student's t copula easily admit a negative relation, but the Clayton and Gumbel Copulas require a half rotation:

$$(1 - U_t, U_{t-1})^T \sim C(F, F)$$

or

$$(U_t, 1 - U_{t-1})^T \sim C(F, F).$$

To determine which of these two half rotations provides a better fit, we look at the log likelihood.

Another issue to deal with is that evaluating the double integrals $W_1(U_{t-1})$ and $W_2(U_t)$ at each in point in time is computationally very intensive. We therefore make use of the fact that $c(v_1, v_2; \alpha^*)$ is simply the likelihood function, and the fact that $\int\int q(v_1, v_2) c(v_1, v_2; \alpha^*)\, dv_1\, dv_2 =$

[7]`https://www.kevinsheppard.com/MFE_Toolbox`
[8]`https://nl.mathworks.com/matlabcentral/fileexchange/53701-bootstrapping-time-series`

$E[q(v_1, v_2; \alpha^*)]$ for any function $q(v_1, v_2)$:

$$W_1(U_{t-1}) = \int_0^1 \int_0^1 [I\{U_{t-1} \leqslant v_1\} - v_1] l_{\alpha,1}(v_1, v_2; \alpha^*) c(v_1, v_2; \alpha^*) \; dv_1 \; dv_2$$
$$= E[(I\{U_{t-1} \leqslant v_1\} - v_1) l_{\alpha,1}(v_1, v_2; \alpha^*)].$$

We then replace this with its sample estimate:

$$W_1(U_{t-1}) \approx \frac{1}{n} \sum_{s=2}^{n} [(I\{U_{t-1} \leqslant U_{s-1}\} - U_{s-1}) l_{\alpha,1}(U_{s-1}, U_s; \hat{\alpha})].$$

Similarly, we estimate:

$$W_2(U_t) \approx \frac{1}{n} \sum_{s=2}^{n} [(I\{U_t \leqslant U_s\} - U_s) l_{\alpha,1}(U_{s-1}, U_s; \hat{\alpha})].$$

4.8 Monte Carlo Simulation

To confirm that our results do not only hold for our data sets, we also perform a simulation study. We simulate data from our four copulas, and use a Gaussian marginal distribution. For the Gaussian copula, we consider the parameter values $\rho = -0.15, \; -0.05, \; 0.05, \; 0.15$. For the Clayton Copula and the Half Rotated Clayton Copula, we consider $\kappa = 0.15, \; 0.3$. For the Gumbel Copula and the Half Rotated Gumbel Copula, we have $\kappa = 1.1, 1.3$. Finally, for Student's t Copula, we consider $\rho = -0.05, \; 0.1, \; 0.05, \; 0.1$, all with $\nu^{-1} = 0.25$.

To generate our times series Y_t, we generate the time series of uniformly distributed variables U_t according to the copula likelihood and then use the Probability Integral Transform $Y_t = G^{*-1}(U_t)$. To simulate U_t, we use the rule of conditional probability:

$$F(U_t | U_{t-1}) = \frac{C(U_{t-1}, U_t; \alpha)}{F(U_{t-1})},$$

where $F(U_t)$ is the cumulative density function of U_t. Note that $U_t \sim \mathcal{U}(0, 1)$, so $F(U_{t-1}) = U_{t-1}$. However, there will hardly be any software that facilitates randomly sampling directly from $F(U_t | U_{t-1})$. Instead we will sample random uniformly distributed variables O_t (not to be confused with the uniformly distributed U_t) and again perform a Probability Integral Transform $F^{-1}_{U_t | U_{t-1}}(O_t) = U_t$. Finding an exact expression for this inverse will be very hard or even impossible, so we use a numerical approximation instead. At each point t, we set U_t at a very low value (such that $F(U_t | U_{t-1}) < O_t$) and then keep increasing it with small increments ϵ until $F(U_t | U_{t-1}) \geqslant O_t$. The value of U_t at that point is then the generated value U_t. This corresponds to the following pseudo code:

Algorithm 2 Copula Simulation

1: **procedure** SIMULATION
2: Generate O_t for $t = 1, \dots, n$
3: Generate Y_1 from $g^*(Y_t)$.
4: **for** $t = 2, \dots, n$ **do**:
5: $Y_t \leftarrow$ a low value
6: **while** $O_t \leqslant \frac{C(U_{t-1}, U_t; \alpha)}{F(U_{t-1})}$ **do**
7: $Y_t = Y_t + \epsilon$

5 Results

5.1 Rank autocorrelations

The rank autocorrelations and their standard errors can be found in Table 2.

Table 2: Rank autocorrelations in returns

Lags	**S&P 100**	**S&P 600**	**Nikkei**	**Bel**
1	$-5.06^{**/**}$ (1.54/1.96)	3.12 (2.02/2.53)	-3.46** (1.43)	6.77** (1.76)
2	- $3.45^{**/-}$ (1.65/2.85)	$-3.27^{*/-}$ (1.79/2.03)	-2.28 (1.41)	-0.42 (1.52)
3	-1.53 (1.92/3.17)	2.27 (1.61/1.92)	-0.39 (1.71)	-0.91 (1.65)
4	0.56 (1.72/2.58)	0.05 (1.80/2.85)	-1.34 (1.58)	-0.26 (1.79)
5	-2.24 (1.54/2.49)	$-2.56^{*/-}$ (1.53/2.49)	-0.92 (1.79)	-2.74* (1.57)
6	-0.55 (1.61/2.47)	$-3.45^{*/-}$ (1.78/2.77)	-0.51 (1.62)	-3.06* (1.79)
7	-2.30 (1.70/2.62)	-1.10 (1.65/2.24)	-1.04 (1.56)	-1.05 (1.63)
8	-1.25 (1.81/3.61)	-0.11 (1.72/2.69)	-1.70 (1.78)	2.94 (1.81)
9	0.37 (1.54/2.05)	-1.11 (1.49/1.84)	-1.93 (1.74)	1.92 (1.67)
10	$3.17^{*/-}$ (1.74/2.43)	-1.06 (1.54/2.26)	4.50** (1.56)	0.33 (174)

*This table presents sample rank autocorrelations ($\times 10^{-2}$) for the first 10 lags of the returns on the S&P 100, S&P 600, Nikkei 225 and Bel 20 Indices. Standard errors ($\times 10^{-2}$) are between brackets and have been obtained using the stationary bootstrap of Politis & Romano. For the S&P 100 and S&P 600, the second number between brackets is the standard error reported by Patton (2012). ** and * indicates significance at 5% and 10% level respectively.*

For the S&P 100 and S&P 600, the same values were found for the rank autocorrelations, but, surprisingly, the standard errors are different. As bootstrapping is a random process, some small variations are to be expected, but not of this magnitude. As the autocorrelations obtained by Patton are exactly the same as ours, the difference must stem from the stationary bootstrap. To determine which standard errors are correct, we compare the results to standard errors obtained using the bootstrap of Sheppard and of Mathworks. The results can be found in Table 3.

Table 3: Standard errors for rank autocorrelations in returns

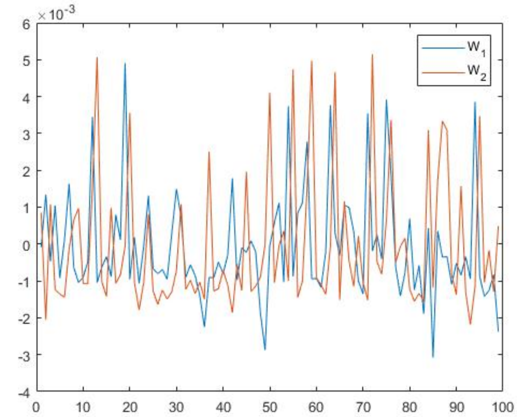
	S&P 100		**S&P 600**	
Lags	Sheppard	Mathworks	Sheppard	Mathworks
1	1.53	1.54	2.07	2.04
2	1.58	1.65	1.72	1.72
3	1.89	1.88	1.63	1.60
4	1.76	1.67	1.83	1.69
5	1.56	1.55	1.53	1.52
6	1.62	1.50	1.79	1.69
7	1.62	1.63	1.65	1.60
8	1.75	1.77	1.75	1.70
9	1.58	1.50	1.61	1.60
10	1.70	1.69	1.54	1.53

This table presents the standard errors on the sample rank autocorrelations ($\times 10^{-2}$) for the first 10 lags of the returns on the S&P 100 and S&P 600 Indices. They have been obtained using the stationary bootstrap of Politis & Romano, using code by Sheppard and by Mathworks.

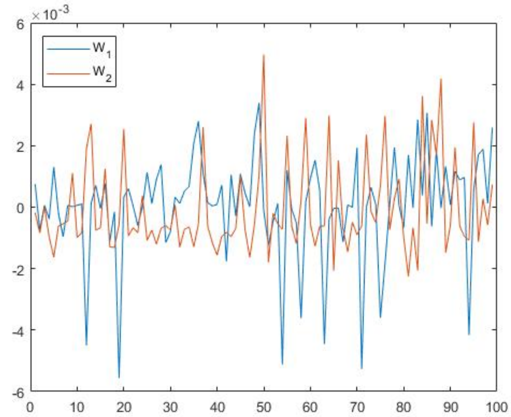
We observe that the standard errors obtained using the codes of Sheppard and Mathworks are very close to our own standard errors, the small variation being attributable to the randomness of the bootstrap. This casts some doubt on the bootstrap implementation by Patton. Nonetheless, most conclusions drawn by Patton (2012) based on these results remain correct for our findings. The first order rank autocorrelations are still significant for the S&P 100 and not for the S&P 600 index returns. This supports modelling the S&P 100 returns as a first order Markov Chain, but not the S&P 600. However, the second order rank autocorrelation is now also significant for the S&P 100. This might indicate a second order dependence relation, but we interpret this as the lagged effect of the strong first order dependence. We discard the S&P 600 for the remainder of this paper and continue our research with the S&P 100, the Nikkei 225 and the Bel 20.

5.2 Parameter estimation

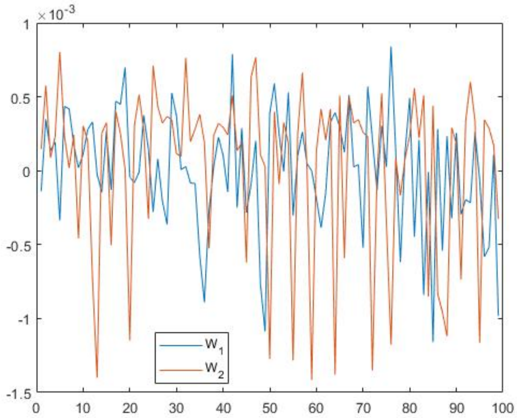
In Figure 1, $W_1(U_{t-1})$ and $W_2(U_t)$ are shown for the first 100 observations of the S&P 100 index for four copulas. We observe that the factors $W_1(U_{t-1})$ and $W_2(U_t)$ are all very close to zero. They appear to be rather heteroskedastic, but centered around 0. The scores are about 1000 times larger in absolute magnitude, plots of which are shown in Appendix A.2. This implies that the squares of $W_1(U_{t-1})$ and $W_2(U_t)$ and the squares of the scores, and thus their variances, have an even larger difference in magnitude, which suggests that $\hat{\Sigma}$ will be almost fully dominated by the variance in the scores. This provides a strong indication that the Truly Naive standard error will be very close to the correct MSML standard error.



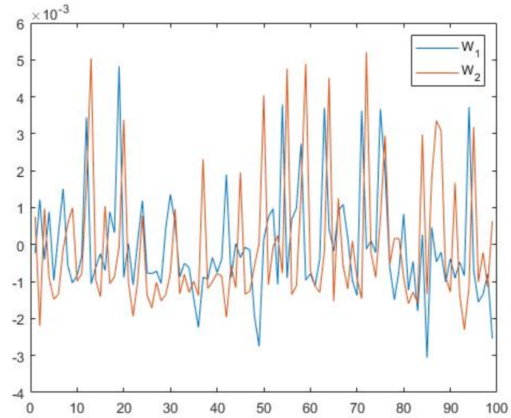
(a) Normal Copula



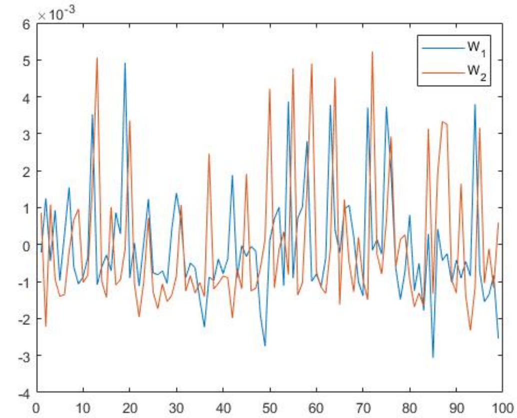
(b) Half Rotated Clayton Copula



(c) Half Rotated Gumbel Copula



(d) Student's t Copula (ρ)



(e) Student's t Copula (ν^{-1})

Figure 1: Plots of $W_1(U_{t-1})$, $W_2(U_t)$ for the first 100 observations of the S&P 100 index for four copulas.

We now proceed with the parameter estimation itself. The parameter estimates, with their corresponding Naive, Doubly Naive, Truly Naive and MSML standard errors and log likelihood for four copulas have been reported in Table 4. The results obtained by Patton (2012) are reported between parentheses. They are almost identical, and small differences can be attributed to the numerical optimization. We observe that the Truly Naive standard errors are equal to the correct MSML standard errors for all four copulas. As the only difference between those standard errors is that the Truly Naive standard errors ignore the estimation error resulting from the use of the EDF, the effect of this estimation error is negligible in this case. Furthermore, this implies that the differences between the naive standard errors and the MSML standard errors cannot be attributed to ignoring the use of the EDF. This is a remarkable result, which is contrary to the interpretation in Patton (2012), who attributes this difference to the use of the EDF.

When comparing the Doubly Naive standard errors with the Truly Naive standard errors, we observe that the differences are sizeable. This implies that the HAC estimate of the variance of the sum of the scores used in the computation of the Truly Naive standard error is quite different from the simple variance estimate used for the Doubly Naive standard errors. Correspondingly, the autocorrelation in the observations has a rather large influence on the differences between the standard errors.

The differences between the Doubly Naive and naive standard errors are also sizeable for the Normal copula and the Half Rotated Clayton copula. Interestingly, for the Half Rotated Gumbel copula the difference is quite small and for both parameters of Student's t copula, the difference is even smaller. As the difference between the Doubly Naive and naive standard errors is that the former takes possible misspecification into account, this suggest that the Half Rotated Gumbel and Student's t copulas are not misspecified. This corroborates with the fact that those two copulas provide the best fit in terms of log likelihood. Even more interestingly, both copulas were not rejected by the Kolgomorov-Smirnov and Cramér-von-Mises Goodness-of-Fit tests in the empirical analysis in Patton (2012), whereas the Normal and Half Rotated Clayton copulas were. This provides some confirmation that indeed the difference between the Doubly Naive and naive standard errors is that the latter don't take possible misspecification into account. By extension, this is also one of the differences between naive and MSML standard errors. For none of the four copulas, the difference between the naive and the MSML standard errors is very big, and for the Normal copula they are even almost equal.

Table 4: Copula parameter estimates for the S&P 100

	Normal	Half Rotated Clayton	Half Rotated Gumbel	Student's t	
	$\hat{\rho}$	$\hat{\kappa}$	$\hat{\kappa}$	$\hat{\rho}$	$\hat{\nu}^{-1}$
Estimate	-0.0565 (-0.0566)	0.1293 (0.1292)	1.0663 (1.0665)	-0.0533 (-0.0534)	0.2062 (0.2070)
Naive s.e.	0.0158 (0.0159)	0.0189 (0.0189)	0.0100 (0.0100)	0.0180 (0.0180)	0.0194 (0.0194)
Doubly Naive s.e.	0.0190	0.0213	0.0103	0.0179	0.0192
Truly Naive s.e.	0.0159	0.0173	0.0089	0.0160	0.0242
MSML s.e.	0.0159 (0.0159)	0.0173 (0.0174)	0.0089(0.0089)	0.0160 (0.0160)	0.0242 (0.0244)
ln $\mathcal{L}$	6.350 (6.376)	30.872 (30.617)	38.675 (38.815)	77.089 (76.883)	

This table presents the copula parameter estimates of four different univariate copula models on the S&P 100 returns, with the corresponding standard errors and log likelihood. The values between brackets are those reported by Patton (2012).

We repeat our estimation for the Nikkei 225 and Bel 20 Indices, the results of which can be found in Tables 5 and 6. As the Nikkei index also showed a negative first order dependence, the Clayton and Gumbel copulas again require a half rotation. Half Rotated Clayton I and Half Rotated Gumbel I are the copulas also used for the S&P 100, and have a rotation in U_t and U_{t-1} respectively. Considering their log likelihoods, we again observe that those rotations provide a much better fit than their mirrored counterparts. The best fit is provided by Student's t copula. The Truly Naive standard errors are identical to the MSML standard errors for five of the six copulas, and only differ a bit for Student's t copula. This again confirms that ignoring the use of the EDF is not very influential. Like for the S&P 100, the Truly Naive and Doubly Naive standard errors do differ, implying that the autocorrelation in observations again played a role. Finally, we observe that the Half Rotated Gumbel copula I and Student's t copula provide the best fit in terms of log likelihood. For those copulas, the difference between the Doubly Naive and naive standard errors is small, which we expect in case the copulas are correctly specified. For the copulas with a lower log likelihood, we also observe a larger difference.

Table 5: Copula parameter estimates for the Nikkei 225

	Normal	Half Rotated Clayton I	Half Rotated Clayton II	Half Rotated Gumbel I	Half Rotated Gumbel II	Student's t	
	$\hat{\rho}$	$\hat{\kappa}$	$\hat{\kappa}$	$\hat{\kappa}$	$\hat{\kappa}$	$\hat{\rho}$	$\hat{\nu}^{-1}$
Estimate	-0.0379	0.0928	0.0232	1.0468	1.0263	-0.0373	0.1476
Naive s.e.	0.0160	0.0182	0.0168	0.0093	0.0088	0.0178	0.0183
Doubly Naive s.e.	0.0187	0.0205	0.0190	0.0095	0.0092	0.0175	0.0182
Truly Naive s.e.	0.0160	0.0169	0.0182	0.0084	0.0087	0.0160	0.0276
MSML s.e.	0.0160	0.0169	0.0182	0.0084	0.0087	0.0161	0.0276
ln $\mathcal{L}$	2.7796	16.2658	1.0471	21.7026	6.5312	44.4920	

This table presents the copula parameter estimates of four different univariate copula models on the Nikkei 225 returns, with the corresponding standard errors and log likelihood.

The results are very similar for the Bel 20 Index. Student's t copula convincingly provides the best fit and the normal copula the worst. The Truly Naive standard errors are all equal to the MSML standard errors. Like for the S&P 100 and Nikkei 225 indices, the copula with the best fit, Student's t, has Doubly Naive and naive standard errors that hardly differ from each other. However, for the other three copulas, those errors differ more.

Table 6: Copula parameter estimates for the Bel 20

	Normal	Clayton	Gumbel	Student's t	
	$\hat{\rho}$	$\hat{\kappa}$	$\hat{\kappa}$	$\hat{\rho}$	$\hat{\nu}^{-1}$
Estimate	0.0803	0.1561	1.0717	0.0778	0.3428
Naive s.e.	0.0157	0.0196	0.0107	0.0187	0.0210
Doubly Naive s.e.	0.0202	0.0231	0.0121	0.0188	0.0209
Truly Naive s.e.	0.0180	0.0224	0.0116	0.0176	0.0241
MSML s.e.	0.0180	0.0224	0.0116	0.0176	0.0241
$\ln \mathcal{L}$	12.9620	42.9186	33.5734	179.4438	

This table presents the copula parameter estimates of four different univariate copula models on the Bel 20 returns, with the corresponding standard errors and log likelihood.

5.3 Monte Carlo simulation

Tables 7, 8, 9 and 10 present the estimation results from simulations from the Normal, Clayton, Gumbel and Student's t copula. As is to be expected, the standard errors decrease as the sample becomes larger. Generally, we observe that all standard errors seem to be rather close to each other for the larger sample sizes. This is an indication that none of the three issues with naive standard errors pose a very big problem in our case. The Truly Naive standard errors are again all either equal to, or almost equal to the MSML standard error. This is also what we have seen in our empirical study, and once more confirms that ignoring that the EDF has been used to estimate the marginal distribution has little influence. Secondly, the Doubly Naive standard errors are generally close to the naive standard errors, indicating that misspecification was not an issue. This too is conform expectations, as we use the known true copulas for our estimations.

A more striking result is that the Doubly Naive and Truly Naive estimates are now also rather close to each other, indicating that also the autocorrelation in the returns did not have a big impact on the standard error estimations. This is an interesting difference with the empirical analysis, where we have seen clear differences between the Truly Naive standard errors and the Doubly Naive standard errors. We can explain this when going back to the rank autocorrelations in Table 2. Apart from the first order autocorrelations, quite a few others are also significantly different from 0. On the other hand, our simulated data was generated to only have a first order dependence. The HAC variance estimator takes all autocorrelations into account, and not just the first order autocorrelation. This can explain why, for the simulated data, the Doubly Naive and Truly Naive standard errors are often close to each other, whereas for the empirical data, they are not. Furthermore, it could also be the case that there is heteroskedasticity in the scores, leading to a difference between the HAC estimator of the variance and the sum of outer products estimator of the variance. It has been well established that stock returns exhibit heteroskedasticity (Engle, 1982), although this need not necessarily translate to heteroskedasticity in the copula scores.

Although all standard errors are generally close to each other for the large sample, the MSML and Truly Naive standard errors are identical in almost all cases, even for the smallest sample. This is a further indication that any difference between the naive standard error and the MSML standard error, even if it is only small, should not be attributed to the estimation of the marginal distribution using the EDF.

Table 7: Copula parameter estimates for simulated data from a Normal Copula with a Gaussian marginal distribution

	n	100	500	1000	2000	4000	5000
	$\hat{\rho}$	-0.0976	-0.0493	-0.0311	-0.0420	-0.0288	-0.0401
	naive s.e.	0.0984	0.0450	0.0318	0.0224	0.0159	0.0141
$\rho = -0.05$	Doubly Naive s.e.	0.0811	0.0449	0.0323	0.0225	0.0161	0.0143
	Truly Naive s.e.	0.0859	0.0453	0.0312	0.0224	0.0171	0.0147
	MSML s.e.	0.0860	0.0453	0.0312	0.0224	0.0171	0.0147
	$\ln \mathcal{L}$	0.4340	0.5993	0.4789	1.7620	1.6581	4.0183
	$\hat{\rho}$	-0.0022	0.0176	0.0130	0.0229	0.0457	0.0293
	naive s.e.	0.2506	0.0454	0.0319	0.0225	0.0158	0.0142
$\rho = 0.05$	Doubly Naive s.e.	0.5851	0.0459	0.0327	0.0225	0.0157	0.0142
	Truly Naive s.e.	0.4834	0.0415	0.0304	0.0211	0.0147	0.0136
	MSML s.e.	0.4834	0.0415	0.0304	0.0211	0.0147	0.0136
	$\ln \mathcal{L}$	-0.0007	0.0808	0.0866	0.5272	4.1859	2.1514
	$\hat{\rho}$	-0.2330	-0.1679	-0.1514	-0.1563	-0.1086	-0.1083
	naive s.e.	0.0930	0.0434	0.0309	0.0217	0.0156	0.0139
$\rho = -0.15$	Doubly Naive s.e.	0.1090	0.0445	0.0317	0.0224	0.0159	0.0142
	Truly Naive s.e.	0.1326	0.0406	0.0276	0.0208	0.0156	0.0137
	MSML s.e.	0.1328	0.0406	0.0276	0.0208	0.0156	0.0137
	$\ln \mathcal{L}$	2.6722	7.1097	11.565	24.704	23.718	29.476
	$\hat{\rho}$	0.1681	0.0775	0.0665	0.0743	0.0908	0.0973
	naive s.e.	0.0943	0.0447	0.0316	0.0223	0.0157	0.0140
$\rho = 0.15$	Doubly Naive s.e.	0.0955	0.0447	0.0310	0.0222	0.0156	0.0139
	Truly Naive s.e.	0.0762	0.0446	0.0309	0.0222	0.0157	0.0141
	MSML s.e.	0.0762	0.0446	0.0309	0.0222	0.0157	0.0141
	$\ln \mathcal{L}$	1.4812	1.5140	2.2203	5.5387	16.574	23.776

This table presents the results of a random simulation of n observations from the Normal Copula with true copula parameter ρ, with a Gaussian marginal distribution. Corresponding standard errors and log likelihood are also reported.

Table 8: Copula parameter estimates for simulated data from a Clayton copula with a Gaussian marginal distributions

	n	100	500	1000	2000	4000	5000
	$\hat{\kappa}$	0.0661	0.1101	0.1156	0.1515	0.1722	0.1706
$\kappa = 0.15$	naive s.e.	0.1205	0.0534	0.0376	0.0280	0.0202	0.0180
(Half	Doubly Naive s.e.	0.1509	0.0490	0.0362	0.0271	0.0199	0.0177
Rotated)	Truly Naive s.e.	0.1520	0.0469	0.0354	0.0259	0.0181	0.0167
	MSML s.e.	0.1519	0.0469	0.0354	0.0259	0.0181	0.0167
	$\ln \mathcal{L}$	0.1690	2.553	5.704	17.984	45.645	56.513
	$\hat{\kappa}$	0.0342	0.1421	0.1552	0.1990	0.2157	0.2189
	naive s.e.	0.0779	0.0500	0.0384	0.0286	0.0207	0.0185
$\kappa = 0.15$	Doubly Naive s.e.	0.0485	0.0477	0.0371	0.0283	0.0205	0.0185
	Truly Naive s.e.	0.0503	0.0550	0.0386	0.0271	0.0206	0.0191
	MSML s.e.	0.0503	0.0550	0.0386	0.0271	0.0206	0.0191
	$\ln \mathcal{L}$	0.1082	5.5095	10.481	31.506	71.256	91.916
	$\hat{\kappa}$	0.1311	0.3477	0.3132	0.2914	0.2440	0.2451
$\kappa = 0.30$	naive s.e.	0.1223	0.0657	0.0449	0.0313	0.0213	0.0191
(Half	Doubly Naive s.e.	0.1273	0.0709	0.0451	0.0314	0.0211	0.0191
Rotated)	Truly Naive s.e.	0.1187	0.0620	0.0427	0.0290	0.0207	0.0184
	MSML s.e.	0.1186	0.0619	0.0426	0.0290	0.0207	0.0184
	$\ln \mathcal{L}$	0.7072	18.514	32.007	57.124	86.477	108.67
	$\hat{\kappa}$	0.2505	0.4975	0.4334	0.3975	0.4078	0.4094
	naive s.e.	0.1115	0.0690	0.0471	0.0329	0.0234	0.0210
$\kappa = 0.30$	Doubly Naive s.e.	0.0996	0.0756	0.0482	0.0334	0.0235	0.0211
	Truly Naive s.e.	0.0901	0.0721	0.0466	0.0308	0.0225	0.0202
	MSML s.e.	0.0893	0.0711	0.0463	0.0306	0.0225	0.0202
	$\ln \mathcal{L}$	4.1122	37.532	59.921	101.79	213.39	267.90

This table presents the results of a random simulation of n observations from the (Half Rotated) Clayton copula with true copula parameter κ, with a Gaussian marginal distribution. Corresponding standard errors and log likelihood are also reported.

Table 9: Copula parameter estimates for simulated data from a Gumbel copula with a Gaussian marginal distributions

	n	100	500	1000	2000	4000	5000
	$\hat{\kappa}$	1.0527	1.0416	1.0411	1.0437	1.0338	1.0343
$\kappa = 1.1$	naive s.e.	0.0711	0.0268	0.0196	0.0143	0.0101	0.0090
(Half	Doubly Naive s.e.	0.0793	0.0277	0.0203	0.0144	0.0103	0.0092
Rotated)	Truly Naive s.e.	0.0754	0.0257	0.0189	0.0145	0.0103	0.0091
	MSML s.e.	0.0754	0.0257	0.0189	0.0145	0.0103	0.0091
	$\ln \mathcal{L}$	0.3110	1.5516	2.7865	5.9041	6.6367	8.5397
	$\hat{\kappa}$	1.0682	1.0297	1.0231	1.0276	1.0245	1.0335
	naive s.e.	0.0736	0.0258	0.0177	0.0133	0.0093	0.0086
$\kappa = 1.1$	Doubly Naive s.e.	0.0643	0.0251	0.0175	0.0133	0.0093	0.0085
	Truly Naive s.e.	0.0676	0.0240	0.0173	0.0129	0.0088	0.0084
	MSML s.e.	0.0676	0.0240	0.0173	0.0129	0.0088	0.0084
	$\ln \mathcal{L}$	0.4870	0.9063	1.1135	2.8025	4.5867	10.411
	$\hat{\kappa}$	1.0519	1.0546	1.0762	1.0887	1.0686	1.0720
$\kappa = 1.3$	naive s.e.	0.0604	0.0280	0.0217	0.0159	0.0108	0.0098
(Half	Doubly naive s.e.	0.0595	0.0272	0.0210	0.0158	0.0106	0.0097
Rotated)	Truly Naive s.e.	0.0688	0.0300	0.0236	0.0170	0.0111	0.0099
	MSML s.e.	0.0688	0.0300	0.0236	0.0170	0.0111	0.0099
	$\ln \mathcal{L}$	0.463	2.468	9.055	22.413	26.911	35.873
	$\hat{\kappa}$	1.1299	1.0241	1.0860	1.0980	1.0792	1.0798
	naive s.e.	0.0933	0.0299	0.0232	0.0168	0.0114	0.0103
$\kappa = 1.3$	Doubly Naive s.e.	0.0993	0.0306	0.0239	0.0173	0.0119	0.0107
	Truly Naive s.e.	0.1106	0.0323	0.0252	0.0178	0.0119	0.0106
	MSML s.e.	0.1109	0.0323	0.0252	0.0178	0.0119	0.0106
	$\ln \mathcal{L}$	1.0846	0.3591	9.5308	23.910	32.027	39.704

This table presents the results of a random simulation of n observations from the (Half Rotated) Gumbel copula with true copula parameter κ, with a Gaussian marginal distribution. Corresponding standard errors and log likelihood are also reported.

Table 10: Copula parameter estimates for simulated data from a Student's t copula with a Gaussian marginal distributions

	n	100	500	1000	2000	4000	5000
	$\hat{\rho}$	0.0532	-0.0007	-0.0208	0.0157	-0.0092	-0.0056
	naive s.e.	0.1037	0.1027	0.0355	0.0257	0.0184	0.0154
$\rho = -0.05$	Doubly Naive s.e.	0.1059	0.2174	0.0362	0.0260	0.0189	0.0150
	Truly Naive s.e.	0.0962	0.2055	0.0340	0.0273	0.0198	0.0157
	MSML s.e.	0.0962	0.2055	0.0340	0.0273	0.0198	0.0158
	ν^{-1}	0.0001	0.1172	0.1420	0.1936	0.1774	0.1743
	naive s.e.	0.0000	0.0613	0.0397	0.0272	0.0194	0.0173
$\nu^{-1} = 0.25$	Doubly Naive s.e.	0.0000	0.0720	0.0400	0.0287	0.0209	0.0184
	Truly Naive s.e.	0.0000	0.0725	0.0404	0.0317	0.0232	0.0205
	MSML s.e.	0.0000	0.0725	0.0404	0.0318	0.0232	0.0205
	$\ln \mathcal{L}$	0.1444	2.0057	7.5585	36.8141	55.1442	66.1724
	$\hat{\rho}$	0.0250	0.0873	0.0870	0.0906	0.1035	0.0984
	naive s.e.	0.1124	0.0500	0.0351	0.0245	0.0176	0.0158
$\rho = 0.05$	Doubly Naive s.e.	0.1290	0.0506	0.0354	0.0246	0.0174	0.0156
	Truly Naive s.e.	0.1197	0.0516	0.0345	0.0240	0.0183	0.0162
	MSML s.e.	0.1198	0.0517	0.0345	0.0240	0.0183	0.0163
	ν^{-1}	0.0872	0.1514	0.1445	0.1281	0.1654	0.1672
	naive s.e.	0.1414	0.0581	0.0404	0.0279	0.0197	0.0175
$\nu^{-1} = 0.25$	Doubly Naive s.e.	0.2192	0.0632	0.0416	0.0281	0.0199	0.0174
	Truly Naive s.e.	0.2740	0.0606	0.0414	0.0276	0.0223	0.0192
	MSML s.e.	0.2739	0.0606	0.0414	0.0276	0.0223	0.0192
	$\ln \mathcal{L}$	0.2814	6.8393	12.9603	23.3723	74.5842	94.0517
	$\hat{\rho}$	0.0134	-0.0519	-0.0379	-0.0122	-0.0120	-0.0294
	naive s.e.	0.1081	0.0502	0.0356	0.0247	0.0179	0.0158
$\rho = -0.1$	Doubly Naive s.e.	0.1069	0.0502	0.0355	0.0245	0.0180	0.0158
	Truly Naive s.e.	0.1309	0.0529	0.0358	0.0233	0.0182	0.0160
	MSML s.e.	0.1310	0.0529	0.0358	0.0233	0.0182	0.0160
	ν^{-1}	0.0627	0.1656	0.1674	0.1326	0.1636	0.1673
	naive s.e.	0.1254	0.0528	0.0376	0.0252	0.0188	0.0168
$\nu^{-1} = 0.25$	Doubly naive s.e.	0.1567	0.0495	0.0355	0.0235	0.0186	0.0166
	Truly Naive s.e.	0.1881	0.0439	0.0345	0.0226	0.0182	0.0163
	MSML s.e.	0.1881	0.0439	0.0346	0.0226	0.0182	0.0163
	$\ln \mathcal{L}$	0.1620	6.8908	12.4649	17.6078	47.0480	61.4585
	$\hat{\rho}$	-0.0389	0.1247	0.1585	0.1443	0.1514	0.1495
	naive s.e.	0.1115	0.0493	0.0352	0.0247	0.0175	0.0156
$\rho = 0.1$	Doubly Naive s.e.	0.1152	0.0489	0.0349	0.0244	0.0174	0.0154
	Truly Naive s.e.	0.1297	0.0594	0.0375	0.0263	0.0195	0.0173
	MSML s.e.	0.1298	0.0594	0.0376	0.0264	0.0196	0.0173
	$\hat{\nu}^{-1}$	0.0411	0.1556	0.2131	0.1760	0.1866	0.1771
	naive s.e.	0.0975	0.0529	0.0406	0.0286	0.0203	0.0180
$\nu^{-1} = 0.25$	Doubly Naive s.e.	0.0746	0.0505	0.0404	0.0292	0.0209	0.0185
	Truly Naive s.e.	0.0777	0.0553	0.0395	0.0310	0.0243	0.0216
	MSML s.e.	0.0777	0.0555	0.0397	0.0312	0.0244	0.0217
	$\ln \mathcal{L}$	0.2494	10.7064	33.0923	51.0783	115.0035	133.9189

This table presents the results of a random simulation of n observations from Student's t copula with true copula parameters ρ and ν^{-1}, with a Gaussian marginal distribution. Corresponding standard errors and log likelihood are also reported.

6 Conclusion

In this paper we have investigated parameter estimation in semiparametric copula-based univariate time series models. We take a critical stand on the results obtained in Patton (2012), who models the S&P 100 returns as a first order Markov Chain using semiparametric copula models. The author estimates the marginal distributions nonparametrically with the Empirical Distribution Function, and combines this with 4 different copula models. The fact that the marginal distributions are estimated needs to be taken into account in the computation of standard errors, resulting in the so-called MSML standard errors. The author also defines 'naive' standard errors, which ignore the fact that the marginal distributions are estimated with the EDF.

First of all, we cast doubt on the standard errors on the rank autocorrelations obtained by Patton, by obtaining those standard errors in three different ways that lead to outcomes that are very similar each other, and different from the ones obtained by Patton. Secondly, we provide a derivation of the naive standard errors of copula parameter estimates that shows that there can also be issues in case of autocorrelation in observations or copula misspecification. We define the alternative Truly Naive standard error that is correct apart from ignoring the use of the EDF, such that any differences between the Truly Naive and MSML standard errors can only be attributed to this aspect. We also define the Doubly Naive standard error, which is naive both in ignoring the use of the EDF and the autocorrelation in observations. Hence, asymptotically, differences between the Doubly Naive and Truly Naive standard error can be attributed to autocorrelation and differences between the Doubly Naive and naive standard errors can be attributed to misspecification.

Thirdly, we compute the naive, Truly Naive, Doubly Naive and the MSML standard errors for the S&P 100 data set and find that the Truly Naive standard errors are equal to or almost equal to the MSML standard errors. This indicates that the difference between the naive standard errors and the MSML standard errors can, contrarily to the interpretation by Patton, not be attributed to estimation of the marginal distributions with the EDF. We also find that for those copulas that appear to be correctly specified, there is hardly a difference between the Doubly Naive and naive standard errors. This fits our claim that differences between those two standard errors are caused by misspecification. Moreover, we repeat our estimation for returns from the Nikkei 225 and the Bel 20 Indices and find very comparable results. Fifth, we perform a simulation study with 4 different copulas, with each 4 different parameter values and 6 different samples sizes. We again find that the Truly Naive standard errors are equal to the MSML standard errors in almost every case. However, we now also find that the Doubly Naive and naive standard errors are close to the MSML standard errors for most large samples, indicating that also autocorrelation and misspecification were not very influential. As we made use of the correctly specified copula, it is conform expectations that misspecification is not an issue. That also autocorrelation did not appear to play a big role can be explained by the fact that the data was simulated to only have a first order dependence, whereas the empirical autocorrelations showed multiple autocorrelations that were signicantly different from 0. We conclude that, for univariate time series, differences between the naive and the MSML standard error are not caused by estimating the marginal distribution, and that ignoring this use of the EDF has negligible impact on the size of the standard errors.

Our research is limited in various ways. First, there exists an abundance of different copulas and different marginal distributions, and this papers only considers a few of those. Furthermore, the simulation considers a limited number of samples, and a larger simulation study would allow more conclusive results. Thirdly, it could be investigated if heteroskedasticity in the scores is also a cause of the difference between the HAC estimator of the variance of the scores and the sum of outer products estimator of the variance. Besides, we only consider the univariate case, whereas it would also be of interest to consider Truly Naive (and Doubly Naive) standard errors in the multivariate case. Moreover, we can only model first order dependence with a bivariate copula; modelling a higher order dependence in a time series requires a multivariate copula. For further research, it would also be interesting to consider time-varying copulas. Volatility and autocorrelation in time series can arguably change over time, which can be modeled using such time-varying copulas.

References

Beare, B. K. (2010). Copulas and temporal dependence. *Econometrica*, 78(1):395–410.

Beare, B. K. (2012). Archimedean copulas and temporal dependence. *Econometric Theory*, 28(6):1165–1185.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327.

Bouyé, E. and Salmon, M. (2013). Dynamic copula quantile regressions and tail area dynamic dependence in forex markets. In *Copulae and Multivariate Probability Distributions in Finance*, pages 125–154. Routledge.

Chan, N.-H., Chen, J., Chen, X., Fan, Y., and Peng, L. (2009). Statistical inference for multivariate residual copula of garch models. *Statistica Sinica*, pages 53–70.

Chen, X. and Fan, Y. (2006a). Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. *Journal of econometrics*, 135(1-2):125–154.

Chen, X. and Fan, Y. (2006b). Estimation of copula-based semiparametric time series models. *Journal of Econometrics*, 130(2):307–335.

Choroś, B., Ibragimov, R., and Permiakova, E. (2010). Copula estimation. In *Copula theory and its applications*, pages 77–91. Springer.

Daul, S., De Giorgi, E. G., Lindskog, F., and McNeil, A. (2003). The grouped t-copula with an application to credit risk. *Available at SSRN 1358956*.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007.

Genest, C. and Favre, A.-C. (2007). Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of hydrologic engineering*, 12(4):347–368.

Genest, C., Nešlehová, J., and Ziegel, J. (2011). Inference in multivariate archimedean copula models. *Test*, 20(2):223.

Genest, C. and Rivest, L.-P. (1993). Statistical inference procedures for bivariate archimedean copulas. *Journal of the American statistical Association*, 88(423):1034–1043.

Geyer, C. J. (2003). 5601 notes: the sandwich estimator. *School of Statistics, University of Minnesota.*

Hartmann, P., Straetmans, S., and Vries, C. d. (2004). Asset market linkages in crisis periods. *Review of Economics and Statistics*, 86(1):313–326.

Jaworski, P., Durante, F., Hardle, W. K., and Rychlik, T. (2010). *Copula theory and its applications*, volume 198. Springer.

Joe, H. (1997). *Multivariate models and multivariate dependence concepts.* Chapman and Hall/CRC.

Joe, H. (2005). Asymptotic efficiency of the two-stage estimation method for copula-based models. *Journal of Multivariate Analysis*, 94(2):401–419.

Joe, H. and Xu, J. J. (1996). The estimation method of inference functions for margins for multivariate models.

Jones, S. (2009). The formula that felled wall st. *Financial Times*, 24:2009.

MacKenzie, D. and Spears, T. (2014). 'the formula that killed wall street': The gaussian copula and modelling practices in investment banking. *Social Studies of Science*, 44(3):393–417.

McNeil, A. J., Frey, R., Embrechts, P., et al. (2005). *Quantitative risk management: Concepts, techniques and tools*, volume 3. Princeton university press Princeton.

Nelsen, R. B. (2003). Properties and applications of copulas: A brief survey. In *Proceedings of the first brazilian conference on statistical modeling in insurance and finance*, pages 10–28. Citeseer.

Newey, W. K. and McFadden, D. (1994). Large sample estimation and hypothesis testing. *Handbook of econometrics*, 4:2111–2245.

Newey, W. K. and West, K. D. (1986). A simple, positive semi-definite, heteroskedasticity and autocorrelationconsistent covariance matrix.

Papadopoulos, A. (2014). The information matrix equality: proof, misspecification, and the quasi-maximum likelihood estimator. *Athens University of Economics and Business.*

Patton, A. (2013). Copula methods for forecasting multivariate time series. In *Handbook of economic forecasting*, volume 2, pages 899–960. Elsevier.

Patton, A. J. (2006). Estimation of multivariate models for time series of possibly different lengths. *Journal of applied econometrics*, 21(2):147–173.

Patton, A. J. (2012). A review of copula models for economic time series. *Journal of Multivariate Analysis*, 110:4–18.

Politis, D. N. and Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical association*, 89(428):1303–1313.

Rémillard, B. (2017). Goodness-of-fit tests for copulas of multivariate time series. *Econometrics*, 5(1):13.

Sklar, M. (1959). Fonctions de repartition an dimensions et leurs marges. *Publ. inst. statist. univ. Paris*, 8:229–231.

Smith, M. S., Gan, Q., and Kohn, R. J. (2012). Modelling dependence using skew t copulas: Bayesian inference and applications. *Journal of Applied Econometrics*, 27(3):500–522.

Wackerly, D., Mendenhall, W., and Scheaffer, R. L. (2014). *Mathematical statistics with applications*. Cengage Learning.

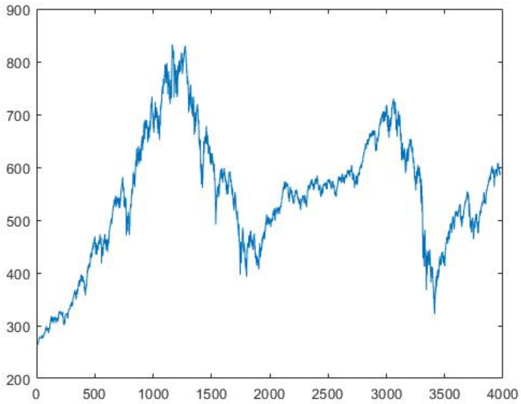
Yahoo!Finance (2019a). Bel 20. Retrieved from `https://finance.yahoo.com/quote/BEL2I.BR/history?p=BEL2I.BR&.tsrc=fin-srch` on May 15, 2019.

Yahoo!Finance (2019b). Nikkei 225. Retrieved from `https://finance.yahoo.com/quote/\%5EN225/history?p=^N225&.tsrc=fin-srch` on May 15, 2019.

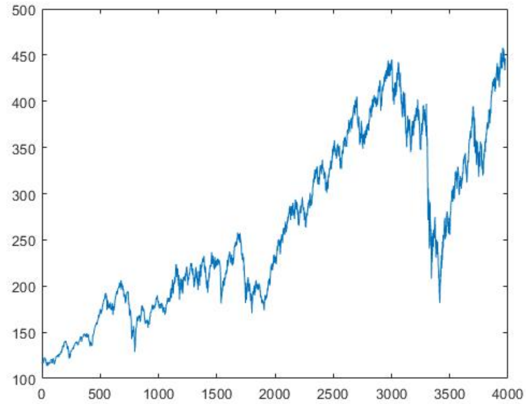
Yahoo!Finance (2019c). S&P 100. Retrieved from `https://finance.yahoo.com/quote/\%5EOEX/history?p=^OEX&.tsrc=fin-srch` on May 15, 2019.

A Appendix

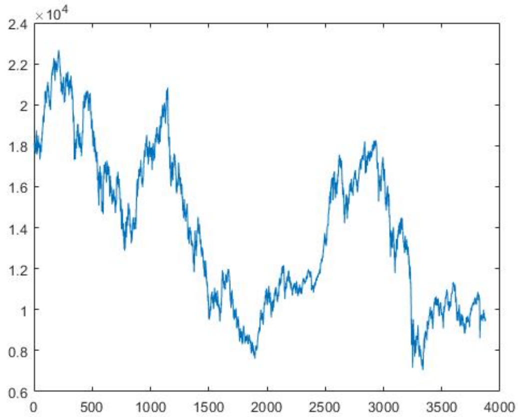
A.1 Plots of index values



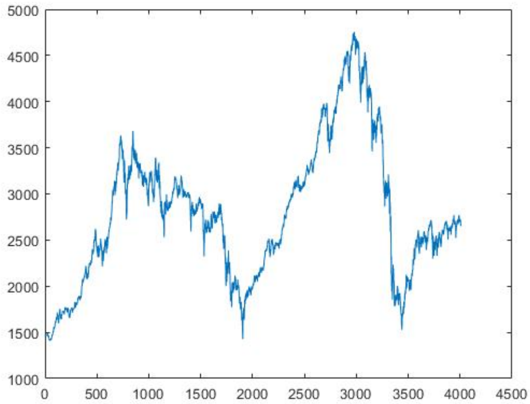
(a) S&P 100



(b) S&P 600



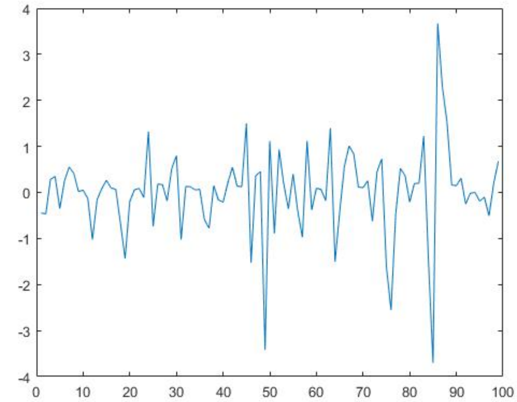
(c) Nikkei 225



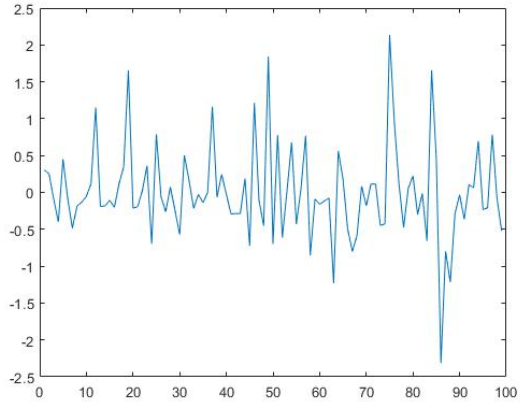
(d) Bel 20

Figure 2: Plots of the index values of the S&P 100, S&P 600, Nikkei 225 and Bel 20 for the period 17 August 1995 till 30 May 2011.

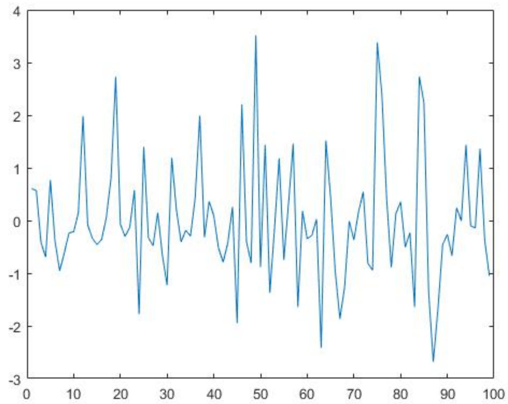
A.2 Plots of scores



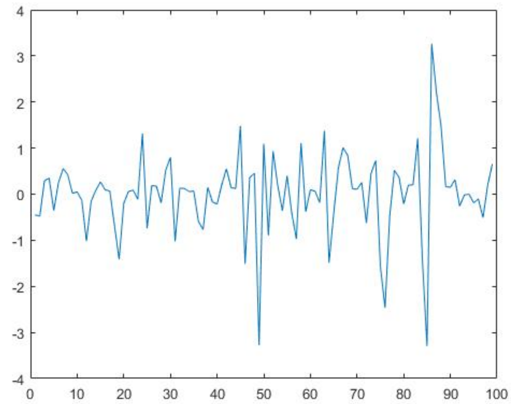
(a) Normal Copula



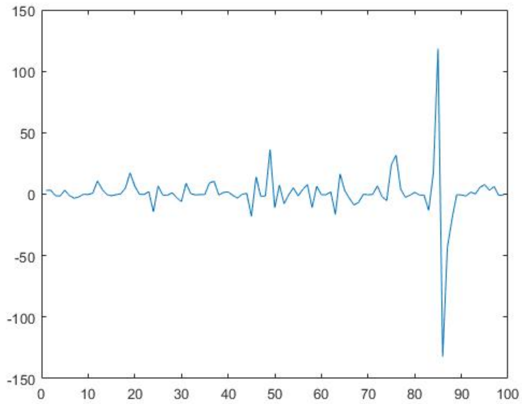
(b) Half Rotated Clayton Copula



(c) Half Rotated Gumbel Copula



(d) Student's t Copula (ρ)



(e) Student's t Copula (ν^{-1})

Figure 3: Plots of the scores $l_\alpha(U_{t-1}, U_t; \alpha^*)$ for the first 100 observations of the S&P 100 index for four copulas.

A.3 Matlab code

We have built on the code provided by Patton provided for his chapter in the Handbook of Economic Forecasting (Patton, 2013), which can be found at `https://public.econ.duke.edu/~ap172/code.html`. Furthermore, we make use of the Econometrics Toolbox of LeSage (`http://www.spatial-econometrics.com/`), Sheppard's MSFE Toolbox (`https://www.kevinsheppard.com/MFE_Toolbox`) and the stationary bootstrap code provided by MathWorks (`mathworks.com/matlabcentral/fileexchange/53701-bootstrapping-time-series`). Files that we have written or edited ourselves are:

- *belestimates* Provides the copula parameter estimates for the Bel 20, Table 6 in our paper.
- *claytoncop* Returns parameter estimate, naive, Doubly Naive, Truly Naive and MSML standard error and log likelihood of the Clayton copula.
- $copula\_deriv2\_chen\_fan\_new3$ Returns $\hat{\Sigma}^{Dnaive}$, $\hat{\Sigma}^{Tnaive}$, $\hat{\Sigma}$, $W_1(U_{t-1}$, $W_2(U_t)$ and the scores of a copula.
- *edf* Returns the EDF of the data.
- *gumbelcop* Returns parameter estimate, naive, Doubly Naive, Truly Naive and MSML standard error and log likelihood of the Gumbel copula.
- *niestimates* Provides the copula parameter estimates for the Nikkei 225, Table 5 in our paper.
- *normalcop* Returns parameter estimate, naive, Doubly Naive, Truly Naive and MSML standard error and log likelihood of the Normal copula.
- *rotclaytoncop* Returns parameter estimate, naive, Doubly Naive, Truly Naive and MSML standard error and log likelihood of the Half Rotated Clayton copula.
- *rotclaytoncop*2 Returns parameter estimate, naive, Doubly Naive, Truly Naive and MSML standard error and log likelihood of the Half Rotated Clayton copula II.
- *rotgumbelcop* Returns parameter estimate, naive, Doubly Naive, Truly Naive and MSML standard error and log likelihood of the Half Rotated Gumbel copula.
- *rotgumbelcop*2 Returns parameter estimate, naive, Doubly Naive, Truly Naive and MSML standard error and log likelihood of the Half Rotated Gumbel copula II.
- *simulation* Provides the simulation results in Tables 7, 8, 9 and 10.

- *sp100estimates* Provides the copula parameter estimates for the S&P 100, Table 4 in our paper.
- *stat_bootstrap_21* Patton's stationary bootstrap.
- *stationary_bootstrap* Sheppard's stationary bootstrap.
- *stationaryBB* Mathworks' stationary bootstrap.
- *Table1* Replicates Table 1 of Patton (2012), which is Table 2 in our paper.
- *Table1_sheppard_mathworks* Replicates Table 1 of Patton (2012) with bootstrap codes by Sheppard and Mathworks, which is Table 3 in our paper.
- *tcop* Returns parameter estimate, naive, Doubly Naive, Truly Naive and MSML standard error and log likelihood of Student's t copula.