Stübinger, Johannes; Mangold, Benedikt; Krauss, Christopher

Working Paper

Statistical arbitrage with vine copulas

FAU Discussion Papers in Economics, No. 11/2016

Provided in Cooperation with:
Friedrich-Alexander University Erlangen-Nuremberg, Institute for Economics

FAU Discussion Papers in Economics

No. 11/2016

Statistical arbitrage with vine copulas

Johannes Stübinger
University of Erlangen-Nürnberg

Benedikt Mangold
University of Erlangen-Nürnberg

Christopher Krauss
University of Erlangen-Nürnberg

ISSN 1867-6707

Friedrich-Alexander-Universität Erlangen-Nürnberg
Institute for Economics
https://www.iwf.rw.fau.de/research/iwf-discussion-paper-series/

Statistical arbitrage with vine copulas

Johannes Stübinger[a,1], Benedikt Mangold[a,1], Christopher Krauss[a,1]

[a] *University of Erlangen-Nürnberg, Department of Statistics and Econometrics, Lange Gasse 20, 90403 Nürnberg, Germany*

Tuesday 25th October, 2016

Abstract

We develop a multivariate statistical arbitrage strategy based on vine copulas - a highly flexible instrument for linear and nonlinear multivariate dependence modeling. In an empirical application on the S&P 500, we find statistically and economically significant returns of 9.25 percent p.a. and a Sharpe ratio of 1.12 after transaction costs for the period from 1992 until 2015. Tail risk is limited, with maximum drawdown at 6.57 percent. The high returns can only partially be explained by common sources of systematic risk. We benchmark the vine copula strategy against other variants relying on the multivariate Gaussian and t-distribution and we find its results to be superior in terms of risk and return characteristics. The multivariate dependence structure of the vine copulas is time-varying, and we see that the share of copulas capable of modeling upper and lower tail dependence increases well over 90 percent at times of high market turmoil.

Keywords: Finance, statistical arbitrage, pairs trading, quantitative strategies, copulas.

Email addresses: `johannes.stuebinger@fau.de` (Johannes Stübinger), `benedikt.mangold@fau.de` (Benedikt Mangold), `christopher.krauss@fau.de` (Christopher Krauss)

[1]The authors have benefited from many helpful discussions with Ingo Klein.

1. Introduction

Pairs trading is a relative-value arbitrage strategy, where an investor seeks to profit from mean-reversion properties of the price spread between two co-moving securities. Gatev et al. (2006) provide the first major academic study on this subject, with excess returns of up to 11 percent p.a. from 1962 until 2002 on the US stock universe. Ever since its publication, several pairs trading approaches have emerged, using different methodologies for pairs selection and pairs trading - Krauss (2016) provides a recent survey.

One stream of literature focuses on copula-based pairs trading. Key representatives are Liew and Wu (2013); Xie and Wu (2013); Stander et al. (2013); Xie et al. (2014); Krauss and Stübinger (2015); Rad et al. (2016). These studies use bivariate copulas to model the dependence structure between two stock return time series, and to identify mispricings that can potentially be exploited in a pairs trading application. The most comprehensive contribution is provided by Rad et al. (2016), which is briefly described following Krauss (2016); Rad et al. (2016). First, during a formation period, similar pairs are selected based on minimizing the sum of squared distances in normalized price space, as in Gatev et al. (2006). The top 20 pairs are retained. Second, the authors fit parametric marginal distribution functions to the return time series of each stock of the top 20 pairs. Third, the returns are transformed into their relative ranks. Then, several different copulas are fitted for each pair and the best-fitting one is selected based on information criteria. Fourth, conditional distributions are derived as first partial derivatives of the copula function, given as $C(u_1, u_2)$:

$$\begin{aligned} h_1\left(u_1|u_2\right) &= P\left(U_1 \leq u_1|U_2 = u_2\right) = \frac{\partial C\left(u_1, u_2\right)}{\partial u_2}, \\ h_2\left(u_2|u_1\right) &= P\left(U_2 \leq u_2|U_1 = u_1\right) = \frac{\partial C\left(u_1, u_2\right)}{\partial u_1}. \end{aligned} \tag{1}$$

In the next step, the conditional probabilities from equation (1) are transformed to daily mispricings $m_{1,t}$ and $m_{2,t}$ for a time t by subtracting a median value of 0.5:

$$m_{1,t} = h_1\left(u_{1,t}|u_{2,t}\right) - 0.5, \quad m_{2,t} = h_2\left(u_{2,t}|u_{1,t}\right) - 0.5, \quad t \in T. \tag{2}$$

Fifth, $m_{1,t}$ and $m_{2,t}$ are used to construct mispricing indices $M_{1,t}$ and $M_{2,t}$, given as

$$M_{1,t} = M_{1,t-1} + m_{1,t}, \quad M_{2,t} = M_{2,t-1} + m_{2,t}, \quad t \in T, \tag{3}$$

with $M_{1,0} = M_{2,0} = 0$. Following Rad et al. (2016), positive values of $M_{1,t}$ and negative values of $M_{2,t}$ indicate that stock 1 is overvalued compared to stock 2 and vice versa. Rad et al. (2016) open a pairs trade at time t if $M_{1,t} > 0.4$ and simultaneously $M_{2,t} < -0.4$ and vice versa. The trade is closed when both mispricing indices reach a level of zero again. Lau et al. (2016) perform an initial abstraction of this concept to three-dimensional space for Bernstein copulas, with a demonstration on three stocks.

We enhance the existing literature in several respects. First, instead of a two-dimensional pairs trading framework in the sense of Gatev et al. (2006); Rad et al. (2016), we construct a multivariate copula-based statistical arbitrage framework in the sense of Avellaneda and Lee (2010). Specifically, for each stock in our S&P 500 data base, we find the three most suitable partners by leveraging different selection criteria. As such, we operate in four-dimensional space (one target stock, three partner stocks) - one of the simplest show cases to benchmark the multivariate models we deploy. A generalization to higher dimensions is straightforward. Empirically, increasing the dimension of the partner portfolio usually leads to higher performance - see, for example, Perlin (2007); Avellaneda and Lee (2010); Chen et al. (2012). Second, we benchmark various multivariate copula models to capture the dependence structure of our quadruple, consisting of one target stock i and three partner stocks. We make use of the multivariate Gaussian and the multivariate t-distribution as baseline models for financial market data. These reference cases are compared against vine copulas, a novelty in high-dimensional dependence modeling and state-of-the-art in the copula literature due to their superior flexibility (Low et al. (2013); Weiß and Supper (2013)). Third, we perform a large-scale empirical study on the S&P 500 from January 1990 until October 2015. We find that our vine copula strategy produces statistically and economically significant returns of 9.25 percent p.a. after transaction costs. The results are far superior compared to the multivariate Student's t-copula (6.76 percent p.a.) or a naive strategy that neglects all partner stocks (0.57 percent p.a.). Similar to Gatev et al. (2006), returns of the vine strategy exhibit low exposure to systematic sources of risk - except for a short-term reversal factor. Monthly alpha after transaction costs still lies at 0.34 percent and tail risk is much lower compared to a simple buy-and-hold investment in the S&P 500. Especially surprising is the fact that the vine strategy does not suffer from consistently negative annualized returns in the recent part of our sample - an issue common among many pairs trading implementations (see, for example, Gatev et al. (2006); Do and Faff (2010); Clegg and Krauss (2016)). Fourth, we analyze the change in chosen copula families in the vine graph over time. We find that in recent years - and especially during financial turmoil - the demand for more flexible copulas increases, allowing for modeling both, upper and lower tail dependence.

The rest of this paper is organized as follows. Section 2 briefly describes our data and the software packages we use. Section 3 outlines the methodology, i.e., the partner selection procedure, the workings of the different copula models, the generation of trading signals, and the backtesting approach. In section 4, we present our results and discuss key findings in light of the relevant literature. Finally, section 5 concludes and provides suggestions for further research.

2. Data and Software

We run our empirical study on the S&P 500, a highly liquid subset of the U.S. stock market, covering 80 percent of available market capitalization (S&P Dow Jones Indices (2015)). Given intense analyst coverage and high investor attention, this market segment serves as a true acid test for any potential capital market anomaly. We follow Krauss and Stübinger (2015) in order to eliminate survivor bias from our data base. First, using Thomson Reuters Datastream, we obtain all month end constituent lists for the S&P 500 from December 1989 to September 2015. Then, we aggregate these lists into a binary matrix, where "1" indicates that a stock is a constituent of the S&P 500 in the subsequent month and "0" the opposite. For all these index constituents, we download the total return indices[2], covering the period from January 1990 until October 2015, equally from Thomson Reuters Datastream. By combining both data sets, we are able to replicate the S&P 500 index constituency and the respective prices over time.

All relevant analyses are conducted in the programming language `R`. Table 1 lists the additional packages for dependence modeling, data handling, and financial modeling.

[2]Return indices reflect prices including reinvested dividends and adjusted for all further corporate actions and stock splits.

Application	R package	Authors of the R package
Dependence modeling	condMVNorm	Varadhan (2015)
	copula	Hofert et al. (2015)
	fCopulae	Rmetrics Core Team et al. (2014)
	permute	Simpson (2015)
	Rcpp	Eddelbuettel et al. (2016)
	VineCopula	Schepsmeier et al. (2015)
	vines	Gonzalez-Fernandez and Soto (2015)
Data handling	dplyr	Wickham and Francois (2016)
	ReporteRs	Gohel (2016)
	xlsx	Dragulescu (2014)
	xts	Ryan and Ulrich (2014)
	zoo	Zeileis et al. (2015)
Financial modeling	fUnitRoots	Wuertz (2013)
	lmtest	Hothorn et al. (2015)
	PerformanceAnalytics	Peterson and Carl (2014)
	QRM	Pfaff and McNeil (2014)
	quantmod	Ryan (2015)
	sandwich	Lumley and Zeileis (2015)
	texreg	Leifeld (2015)
	timeSeries	Rmetrics Core Team et al. (2015)
	tseries	Trapletti and Hornik (2016)
	TTR	Ulrich (2015)

Table 1: R packages used in this paper.

3. Methodology

We slice our data set in 281 overlapping study periods. Each study period consists of a twelve-month initialization, a twelve-month formation, and a six-month out-of-sample trading period. Consequently, we have a total of 281 trading periods, of which six overlap and run in parallel. Their resulting returns are averaged in the sense of Gatev et al. (2006), thus consolidating the six portfolio returns to one final return time series. For each study period j, we consider a total of n_j stocks that are (i) an index constituent on the last day of the formation period and (ii) exhibit full historical price data, meaning no NA's.

The initialization period (subsection 3.1) is two-staged. The partner selection (subsection 3.1.1) deals with four different approaches for obtaining the most suitable partner stocks. Every approach is based on a different measure of association and emphasizes different aspects of the joint four dimensional dependence structure. The model fit (subsection 3.1.2) characterizes four different variants to adequately describe this multivariate dependence structure. At first, as a reference case, the naive E-model is created, only incorporating past returns of the target stock in the mispricing index and thus neglecting all partner stocks. Then, we construct the G-model and the T-model, relying on the multivariate Gaussian distribution and the multivariate t-distribution for identifying mispricings of the target stock relative to its partner stocks. Finally, we benchmark these implementations against the V-model, making use of highly flexible vine copulas for capturing multivariate mispricings.

The formation period (subsection 3.2) is used for creating one out-of-sample mispricing index per model for each target stock. Then, all mispricing indices per model type are ranked based on their augmented Dickey-Fuller (ADF) test statistics in ascending order.

Afterwards, all models are re-estimated based on the new return data of the formation period, to achieve an updated calibration for the out-of-sample trading period (subsection 3.3). The top r mispricing indices per model type are continued in the trading period and serve as trading signal for the corresponding top r target stocks. Specifically, a stock is bought (sold short) for each of the strategy variants, when its mispricing index falls below (exceeds) certain threshold levels. The models are called strategies in the trading process, i.e. the E-model corresponds to the E-strategy (E-strat), the G-model corresponds to the G-strategy (G-strat), the T-model corresponds to the T-strategy (T-strat), and the V-model corresponds to the V-strategy (V-strat).

3.1. Initialization period

3.1.1. Partner selection

The partner selection procedure aims at identifying a partner triple for each target stock, based on adequate measures of association. All four stocks together (one target stock and its three suitable partners) form the quadruple Q. Given that every stock of the S&P 500 is consecutively considered as target stock, we effectively create n_j such quadruples, which are logged in an $(n_j \times 4)$-output matrix.

We would like to make two preliminary remarks. First, all measures of association are calculated using the ranks of the daily discrete returns X of our samples. The rank transformation provides some robustness, since the impact of large values (outliers) is reduced by only considering the position within the ordered sample, not the value itself. Second, we only take into account the top 50 most highly correlated stocks (approximately 10 percent of available stocks n_j) for a given target as potential partner stocks, in order to limit the computational burden. This bivariate preselection speeds up the required calculation time by a factor of 1,000.

Traditional approach. A natural way of describing bivariate linear dependence between two variables is correlation. As baseline approach, the high dimensional relation between the four stocks is approximated by their pairwise bivariate correlations via Spearman's ρ. In addition to the robustness obtained by rank transformation, it allows to capture nonlinearities in the data to a certain degree. Also, we ensure consistency with the other three approaches, which are equally calculated on ranks.

The procedure itself is rather simple. First, we calculate the sum of all pairwise correlations for all possible quadruples, consisting of a fixed target stock and of one of the $\binom{50}{3}$ triples of partner stocks. Second, the quadruple with the largest sum of pairwise correlations is considered as Q and saved to the output matrix.

Extended approach. Schmid and Schmidt (2007) introduce multivariate rank based measures of association. We rely on a measure that generalizes Spearman's ρ to arbitrary dimensions - a natural extension of the traditional approach.

In contrast to the strictly bivariate case, this extended approach – and the two following approaches – directly reflect multivariate dependence instead of approximating it by pairwise measures only. We expect a more precise modeling of high dimensional association and thus a better performance in trading strategies.

Q for a given target stock is obtained by the following procedure: Build every quadruple out of the $\binom{50}{3}$ possible combinations containing the target stock, calculate the multivariate version of Spearman's ρ for each quadruple, and select for Q the one with the largest value.

Geometric approach. We introduce an intuitive geometric approach for measuring multivariate association in order to select Q. For the sake of clarity, we illustrate this measure in the bivariate case. A generalization to higher dimensions is straightforward.

Consider the relative ranks of a bivariate random sample, where every observation takes on values in the $[0, 1] \times [0, 1]$ square. If there exists a perfect linear relation among both the ranks of the components of the sample, a plot of the relative ranks would result in a perfect line of dots between the points $(0, 0)$ and $(1, 1)$ – the diagonal line. However, if this relation is not perfectly linear, at least one point differs from the diagonal. By dropping a perpendicular from that deviating point to the diagonal, one could calculate the Euclidean distance of the deviation. The more the relative ranks deviate from the diagonal, the larger the sum of all their respective deviations. This sum can be used as a measure of deviation from linearity, the diagonal measure.

Hence, we try to find the quadruple Q that leads to the minimal value of the sum of Euclidean distances from the relative ranks to the (hyper-)diagonal in four dimensional space for a given target stock. As such, we calculate the four dimensional diagonal measure for every of the $\binom{50}{3}$ combinations of partner stocks. The target stock together with the triple, that induces the lowest value of the diagonal measure, is saved as Q in the output matrix.

Extremal approach. Mangold (2015) proposes a nonparametric test for multivariate independence. The resulting χ^2 test statistic can be used to measure the degree of deviation from independence, so dependence. Main focus of this measure is the occurrence of joint extreme events. A disproportionately high or low occurrence of joint extreme events inflates the measure. Q is the combination of the target stock together with the triple of partner stocks that maximizes this extremal measure. With this approach, we focus more on the multivariate extremal regions of the unit cube, since those events are crucial for any kind of trading strategy.

Similar to the geometric approach, the partner selection operates as follows: for a given target stock, we calculate the extreme measure for every combination of the $\binom{50}{3}$ possible partner triples. The combination that leads to the largest value of the extremal measure is considered as Q and saved to the output matrix.

It is important to highlight the differences between the four approaches. The traditional, the extended, and the geometric approach share a common feature - they measure the deviation from linearity in ranks. All three aim at finding the quadruple that behaves as linearly as possible to ensure that there is an actual relation between its components to model. While it is true that this aspiration for linearity excludes quadruples with components that are not connected (say, independent), it also rules out nonlinear dependencies in ranks. On the other hand, the extremal approach tries to maximize the distance to independence with focus on the joint extreme observations. This includes both, linear and nonlinear relations among the components of Q. Since two of our introduced models (T-model and V-model) can implement nonlinear and tail dependencies, the extremal approach is promising and we expect a better preselection and thus better results compared to the other routines.

3.1.2. Model fit

In this subsection, we describe the E-, G-, T-, and V-model in detail, and how the fitting process works. At this point, we are facing four dimensional samples $\boldsymbol{X} = (X_1, X_2, X_3, X_4)$, where X_1 describes the return of the target stock and X_2, X_3, X_4 describe the returns of the partner stocks. Clearly, we are interested in the conditional distribution function $X_1|X_2, X_3, X_4$ in order to calculate the mispricing index. If necessary, we transform discrete return data X_i $(i = 1, 2, 3, 4)$ to relative ranks, henceforth denoted as U_i.

E-model. The baseline approach relies on the assumption that the partner stocks are independent of the target stock. Therefore, only the target stock's own history is used to value current returns. Conditioning on (X_2, X_3, X_4) leads to the empirical cumulative distribution function F_E based on the past daily discrete return data of the target stock. Therefore, $X_1|X_2, X_3, X_4$ follows a law with distribution function

$$h_E(x_1|x_2, x_3, x_4) = h(x_1) = F_E(x_1). \tag{4}$$

G-model. We assume our sample $\boldsymbol{X}$ to follow a Gaussian law with vector of expected values μ and covariance matrix Σ. The conditional distribution of the target stock given the three partners follows

$$X_1|X_2, X_3, X_4 \sim \mathcal{N}\left(\mu_1 + \Sigma_{1,2:4}\Sigma_{2:4,2:4}^{-1}\left(x_{2:4} - \mu_{2:4}\right), \Sigma_{1,1} - \Sigma_{1,2}\Sigma_{2:4,2:4}^{-1}\Sigma_{2:4,1}\right), \tag{5}$$

where 2:4 denotes the dimensions 2, 3, and 4 (see Eaton (1983)). Note, that $\Sigma_{1,1} - \Sigma_{1,2}\Sigma_{2:4,2:4}^{-1}\Sigma_{2:4,1}$ is called the Schur complement of Σ with respect to $\Sigma_{1,1}$. In this setup, our model is fitted by estimating μ and Σ from the sample. We refer to equation (5) as the function $h_G(x_1|x_2, x_3, x_4)$ in the following.

T-model. A 4-dimensional random vector $\boldsymbol{X} \sim t_4(\mu, \Sigma; \lambda, \nu)$ has a density function of

$$f_{\boldsymbol{X}}(x) = \frac{\Gamma(\nu/2 + 2)}{(\pi\lambda)^2\Gamma(\nu/2)|\Sigma|^{1/2}}\left(1 + \frac{1}{\lambda}(x-\mu)^\top\Sigma^{-1}(x-\mu)\right)^{-(\nu/2+2)}, \tag{6}$$

where $\mu \in \mathbb{R}^4$, $\Sigma \in \mathbb{R}^{4\times 4}$, $\lambda, \nu \in \mathbb{R}^+$. In particular, Kotz and Nadarajah (2004); Nadarajah and Kotz (2005) prove that under weak assumptions, the conditional distribution of a multivariate t-distribution follows a t-distribution as well. Kotz and Nadarajah (2004) show that

$$X_1|X_2, X_3, X_4 \sim t(\mu_1 + \Sigma_{1,2:4}\Sigma^{-1}_{2:4,2:4}(x_{2:4} - \mu_{2:4}), \Sigma_{1,1} - \Sigma_{1,2:4}\Sigma^{-1}_{2:4,2:4}\Sigma_{2:4,1}; \nu + 3, \nu + 3), \quad (7)$$

where 2:4 denotes the dimensions $2, 3$, and 4. For an intuitive proof, see Ding (2016). The parameters of the unconditional distribution of $\boldsymbol{X}$ are estimated using the 'Expectation/Conditional Maximization Either' algorithm, developed by Liu (1994). In the following, we refer to equation (7) as function $h_T(x_1|x_2, x_3, x_4)$.

V-model. A copula is a multivariate distribution function $C(F_1(x_1), F_2(x_2), F_3(x_3), F_4(x_4)) = C(u_1, u_2, u_3, u_4)$ with univariate distributions F_i, respective densities f_i, and uniformly distributed marginal distributions u_i ($i = 1, 2, 3, 4$). The corresponding density function is denoted as $c(u_1, u_2, u_3, u_4)$. According to Sklar (1959), any multivariate distribution function can be decomposed into its univariate marginal distributions and a copula function. This decomposition is unique if and only if the univariate marginals are continuous. For an excellent oeuvre on copulas see Nelsen (2007).

Any continuous four dimensional density function can be decomposed in multiple ways, as for example by

$$\begin{aligned} f(x_1, x_2, x_3, x_4) =& f_1(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \cdot f_4(x_4) \\ & \cdot c_{12}(F_1(x_1), F_2(x_2)) \cdot c_{23}(F_2(x_2), F_3(x_3)) \cdot c_{24}(F_2(x_2), F_4(x_4)) \\ & \cdot c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)) \cdot c_{34|2}(F_{3|2}(x_3|x_2), F_{4|2}(x_4|x_2)) \\ & \cdot c_{14|23}(F_{1|23}(x_1|x_2, x_3), F_{4|23}(x_4|x_2, x_3)), \end{aligned} \quad (8)$$

where c describes the density of the copulas (see Aas et al. (2009) for standard notation for vine copulas). Based on such decompositions, Joe (1994, 1996, 1997); Bedford and Cooke (2001, 2002); Whelan (2004) introduce the concept of vine copulas. The main idea is that multivariate parametric copulas are often very rigid and thus cannot specifically model partial dependence characteristics between some marginal distributions. As an example, the widely used class of Archimedean copulas only has one dependence parameter, even for high dimensional distributions. Vine copulas however, decompose the multivariate dependence into bivariate and conditional bivariate dependencies which are easy to model - see Czado (2010) for detailed information about vine copulas. A key feature of using vine copulas instead of elliptical or Archimedean copulas is the flexible modeling of combinations of tail dependencies. The downside however, is the lack of interpretability of the parameters and the cumbersome fitting process (see Brechmann et al. (2014)).

Figure 1 gives a showcase on the copula families used in the fitting process. The BB-copulas are mixture models of Archimedean copulas (see Joe (1997)) and allow for different combinations of upper and lower tail dependencies.

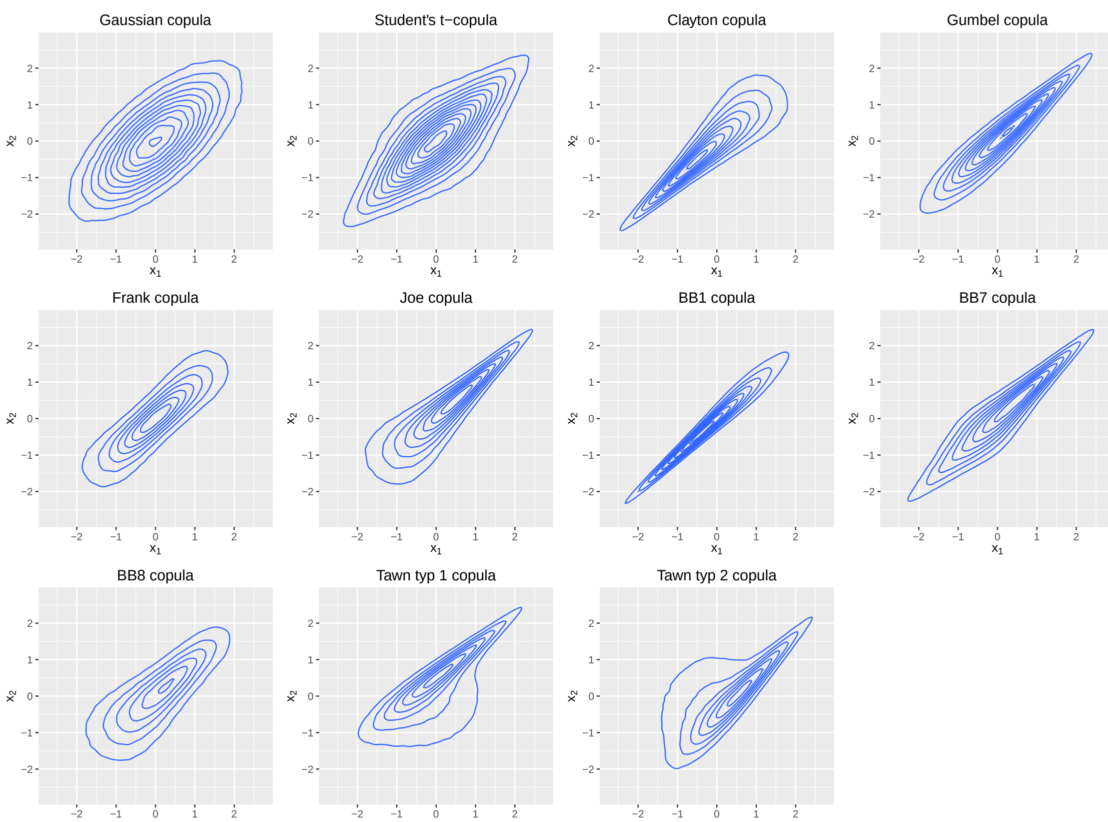


Figure 1: Two dimensional copula models that are used to fit the data. The univariate marginal distributions are standard Gaussian to illustrate deviations from the bivariate Gaussian distribution.

A way of describing the structure of a vine copula is with a node graph (see Bedford and Cooke (2001)). Depending on the structure of the graph, Aas et al. (2009) classify vine copulas into C-vine (star form) and D-vine (line form) copulas, among other classes. In the following, we will focus on C-vine copulas only, since we aim to sort the components rather by importance than by temporal aspects. Note that in our application, the structure of the vine tree is prespecified, hence no further selection heuristics are required.

In four dimensions, there are three trees of decomposition, each of which is represented by a graph (see figure 2). Each edge in the graph stands for a bivariate copula, modeling the dependence structure of the connected nodes. For the first tree of C-vines, one component is placed at the center of the graph (U_2 in the example) and connected to each of the other three nodes. Hence, three bivariate copulas need to be estimated. In total, there are four different constellations, since one could set every component at the center of the graph.

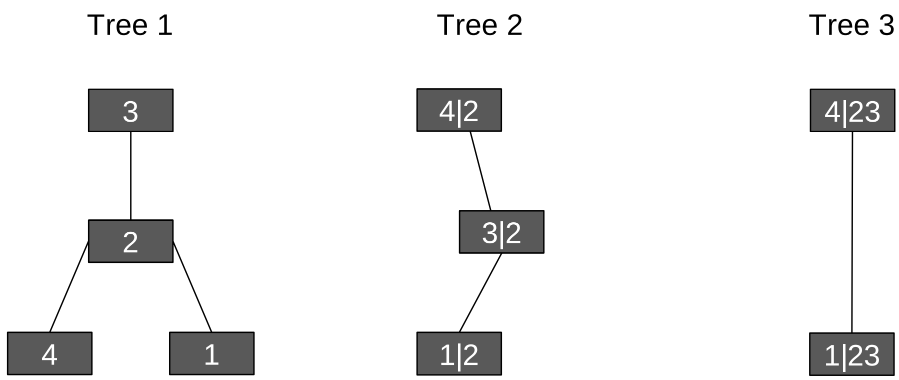


Figure 2: Example of a four dimensional C-vine copula (see equation (8)). Each edge models the bivariate dependence of the (un-)conditional data associated with the connected nodes.

The center of the first tree is the component on which the second tree is conditional on. The second tree connects the conditional distributions (conditional on U_2) of the first graph. The two bivariate copulas that need to be estimated on this level connect $U_1|U_2$, $U_3|U_2$, and $U_4|U_2$. Again, one component needs to be set as center ($U_3|U_2$ in the graph), so three constellations are possible. The third tree connects the two remaining components, both conditional on the centers of the prior trees ($U_1|U_2, U_3$, and $U_4|U_2, U_3$). Hence, one bivariate copula needs to be estimated.

In total, we have 12 possible constellations of C-vine graphs and six bivariate copulas that need to be estimated per constellation in the four dimensional setting. However, if we set the target stock to be the first component without loss of generality, not all constellations are conducive to our problem of generating the mispricing index using the conditional distribution of $U_1|U_2, U_3, U_4$. In fact, this conditioning is only possible, if U_1 is never at the center of a single graph, conditionally nor unconditionally. This restraint leaves us with six feasible constellations, see figure 3.

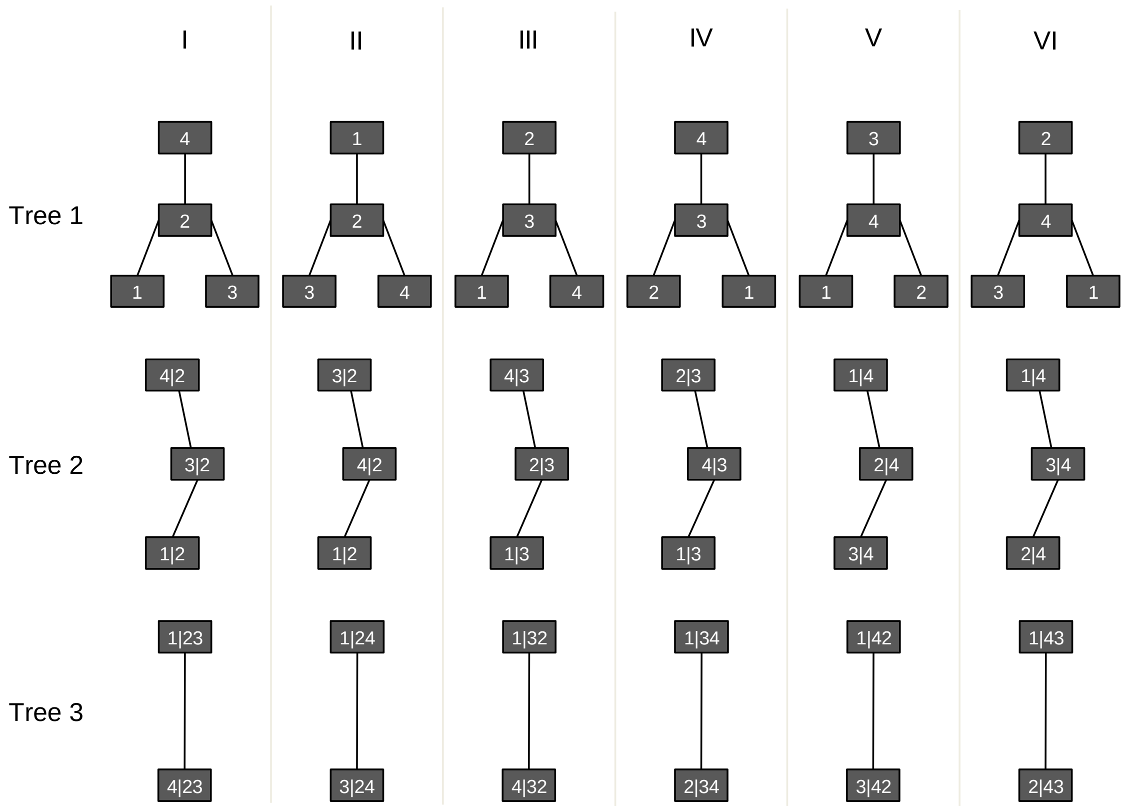


Figure 3: The three trees of all six constellations of four dimensional C-vine copulas that are suitable for calculating a mispricing index.

It is impossible to know up front which of the remaining six constellations is most suitable for our problem at hand. As such, we fit all of them at first and select the most suitable one at a later stage. In total, 39 different copula families[3] could potentially have been chosen by maximum likelihood (Brechmann et al. (2012); Brechmann and Czado (2013); Dissmann et al. (2013)). However, a pre-analysis show that barely one half of the 39 copula models are

[3]The R package VineCopula allows 39 different copula families.

actually selected in the fitting process. In fact, the top 19 most selected copula families were chosen in over 99 percent. This finding reassures us to constrain the set of potential copula families to the top 19, thus reducing the computational burden by a significant amount.

Once the six vine copulas are fitted, the conditional distribution of $U_1|U_2, U_3, U_4$ is of interest. To obtain $U_1|U_2, U_3, U_4$, we repeatedly condition on the partner stocks. In our example, the first step would be

$$h_C(u_1|u_2, u_3, u_4) = \frac{\partial C(u_1|u_2, u_3, u_4)}{\partial u_2}. \tag{9}$$

3.2. Formation period

In the formation period T_{for}, we aim to select the top r target stocks to be transferred to the trading period for actual trading. At first, we determine $\forall t \in T_{\text{for}}$ the daily mispricing m_t as

$$m_t = h_t - 0.5, \tag{10}$$

where h_t denotes the respective conditional distribution function of subsection 3.1 evaluated at time t. If $m_t > 0$ ($m_t < 0$), the target stock is considered to be overvalued (undervalued) on day t relative to its own past (E-model) or to its own past and its partner stocks (G-, T-, V-model). In case of fair pricing, the daily mispricing is close to 0. Since h_t specifies a probability, the daily mispricing m_t lies between -0.5 and 0.5. The mispricing index at time t ($t \in T_{\text{for}}$) is defined as

$$M_t = M_{t-1} + m_t, \qquad (M_0 = 0), \tag{11}$$

where $M_{for} = (M_t)_{t \in T_{\text{for}}}$. In other words, we accumulate the daily mispricings over time to a mispricing index. Xie et al. (2014) discusses stochastic properties of such a process. If the daily mispricing is relevant, the mispricing index should be mean-reverting, i.e., pricing errors are corrected over time. As such, we select the top r target stocks per model type based on the ADF test statistic of their corresponding mispricing index. Specifically, the top r target stocks with the lowest associated test statistics are transferred to the trading period - irrespective of the fact if the null hypothesis "unit root" is rejected or not. In other words, we use the ADF test statistic as heuristic to assess mispricing indices in terms of their propensity to mean-revert. We use one additional constraint for the V-model. From the six possible constellations elaborated in subsection 3.1 (and the six possible mispricing indices) per target stock, only the constellation with the lowest ADF test statistic is considered in the ranking, thus ensuring that each target stock may only be selected once.

3.3. Trading period

The top target stocks with lowest ADF test statistics of the mispricing index are transferred to the trading period and their corresponding models are re-estimated using historical data of the formation period. For every model and every newly arriving return on day t, with $t \in T_{trad}$, we update the mispricing index outlined in equation (11), using the conditional probabilities of the respective model.

Increasing deviations from equilibrium constitute larger mispricings of the target security relative to its own history or relative to its own history and its peers - see subsection 3.2. If our assumptions hold and mispricings are corrected over time, the mispricing index should revert to its equilibrium value. We aim to capture this potentially mean-reverting behavior with a simple trading strategy based on Bollinger bands of Bollinger (1992). For constructing the Bollinger bands, we calculate the running mean and standard deviation of the mispricing index of the past $d = 20$ days - corresponding to one trading month and according to Bollinger (1992) the most common parametrization. We obtain the upper (lower) band by adding (subtracting) k-times the running standard deviation to (from) the running mean. We set $k = 1$, a value similar to Avellaneda and Lee (2010); Clegg and Krauss (2016), and others, who aim at achieving a higher trading frequency compared to $k = 2$.

We go long (short) the target stock with 1 USD when the mispricing index crosses its lower (upper) Bollinger band, given that it is undervalued (overvalued) in this situation. We exit the trade when the mispricing index crosses its running 20-day mean - the reversion to equilibrium. Alternatively, all trades are exited at the end of the trading period, or upon delisting. Note that we only trade the target stock, and not the partner stocks, which are exclusively used for evaluating mispricings of the target stock. Since we still aspire a classic long-short investment strategy in the sense of Gatev et al. (2006), we follow Avellaneda and Lee (2010) and hedge market exposure day-by-day with corresponding investments in the

S&P 500 index, rendering the overall portfolio (consisting of the top target stocks, where each stock is either long, short, or flat) dollar-neutral.[4]

Return calculation follows Gatev et al. (2006). We scale the payoffs of the overall portfolio by the number of stocks that actually open during the six-month trading period, providing us with the return on actually employed capital - the more common metric in the pairs trading literature. Following Avellaneda and Lee (2010), we assume transaction costs of 5 bps per half-turn for the target stock, so 10 bps for the round-trip trade.

4. Results

4.1. General results

In table 2, we report mean returns per year as well as annualized Sharpe ratios for each of the four partner stock selection procedures (see subsection 3.1.1) and for each of the four strategy variants (see subsection 3.1.2) - before and after transaction costs. Following common practice in the pairs trading literature, we focus on a portfolio containing the top 20 target stocks. We see that the E-strat results in yearly returns of merely 0.57 percent after transaction costs and a negative Sharpe ratio - which is caused by subtracting the risk-free rate. Note that these results are identical across all four selection procedures, given that the E-strat only considers a stock's own history for constructing the mispricing index. This puristic approach does not seem to allow for investment results that beat the performance of the S&P 500. Taking into account a stock's own as well as the partner stocks' history leads to clear improvements for the G-strat and the T-strat. Mean returns range between 3.52 and 6.76 percent per year and Sharpe ratios between 0.15 and 0.73 after transaction costs - depending on the model and the selection algorithm. The V-strat results in yearly returns between 7.43 and 9.25 percent and Sharpe ratios between 0.75 and 1.12. We believe that this outperformance is driven by the higher flexibility of vine copulas. Instead of enforcing a multivariate elliptical dependence structure over the target and the partner stocks, vine copulas allow for a plethora of linear and non-linear dependencies - individually tailored to

[4] In contrast to Avellaneda and Lee (2010), we aim for a dollar-neutral portfolio for consistency reasons with Gatev et al. (2006) and the majority of the pairs trading literature.

all relevant interactions. Furthermore, Cherubini et al. (2004); Fischer et al. (2009) show that elliptical models have a suboptimal model performance, because of their symmetric tail dependence.

Regarding the selection algorithm, we observe that the traditional approach, the extended approach, and the geometric approach lead to similar results per strategy. As such, we can cautiously infer that our findings seem to be fairly robust to the partner portfolio that is selected - irrespective of the actual selection metric (traditional, i.e., purely bivariate, extended, i.e., multivariate Spearman's ρ, or geometric, i.e., based on distances). Slight improvements are achieved with the extremal selection algorithm, which emphasizes the degree of deviation from independence (see subsection 3.1).

Return	Before transaction costs				After transaction costs				
	E-strat	G-strat	T-strat	V-strat	E-strat	G-strat	T-strat	V-strat	S&P 500
Traditional	0.0303	0.0716	0.0757	0.1084	0.0057	0.0469	0.0497	0.0785	0.0581
Extended	0.0303	0.0594	0.0723	0.1035	0.0057	0.0352	0.0470	0.0763	0.0581
Geometric	0.0303	0.0712	0.0750	0.1021	0.0057	0.0460	0.0495	0.0743	0.0581
Extremal	0.0303	0.0756	0.0937	0.1210	0.0057	0.0507	0.0676	0.0925	0.0581
Sharpe ratio	Before transaction costs				After transaction costs				
	E-strat	G-strat	T-strat	V-strat	E-strat	G-strat	T-strat	V-strat	S&P 500
Traditional	0.0497	0.8212	0.8993	1.3234	-0.2740	0.3767	0.4304	0.8422	0.2119
Extended	0.0497	0.5767	0.8039	1.1743	-0.2740	0.1535	0.3660	0.7655	0.2119
Geometric	0.0497	0.7919	0.8472	1.1770	-0.2740	0.3507	0.4071	0.7476	0.2119
Extremal	0.0497	0.8375	1.1812	1.5642	-0.2740	0.4171	0.7315	1.1186	0.2119

Table 2: Yearly returns and Sharpe ratios for the top 20 target stocks of the E-, G-, T-, and V-strat compared to the S&P 500 from January 1992 until October 2015 for the four partner selection algorithms.

In the following subsections, we focus on the extremal selection algorithm, given that it produces the most favorable results with theoretical underpinning, i.e., the deviation from multivariate independence. In subsection 4.2, we evaluate the performance of all four strategy variants. Specifically, we analyze the return distributions, different Value at Risk metrics, risk-return characteristics, and exposure to common risk factors. The majority of selected performance metrics is discussed in detail in Bacon (2008). In subsection 4.3, we analyze strategy results over time. Finally, in subsection 4.4 we perform bootstrap trading, a robustness check, and conduct a deep dive on the identified dependence structures.

4.2. Strategy performance

Table 3 reports monthly return characteristics for the top 20 target stocks per strategy variant from January 1992 until October 2015. We find statistically significant returns for the G-strat, the T-strat, and the V-strat, with Newey-West (NW) t-statistics above 5.6 before transaction costs and above 3.9 after transaction costs for all multivariate strategies. From an economic perspective the returns are significant as well, ranging between 0.43 percent per month for the G-strat and 0.75 for the V-strat - even after transaction costs. As expected, the E-strat does not produce statistically and economically significant returns after tradings costs - the mere 0.07 percent per month are way behind the 0.56 of the S&P 500 benchmark.

All strategy variants exhibit positive skewness and follow a leptokurtic distribution. Also, results do not seem to be driven by strong outliers - compare the minimum and maximum values. Following the methodology of Mina and Xiao (2001), Value at Risk (VaR) levels are low, with a minimum of -2.74 percent for the 1%-VaR for the V-strat. Tail risk is much less expressed compared to an investment in the S&P 500 (1%-VaR at -11.65 percent) and at a similar level as classical pairs trading (1%-VaR at -1.94 percent for the top 20 pairs in Gatev et al. (2006)). Maximum drawdown after transaction costs is greatly reduced for the multivariate strategies with 16.03 percent for the G-strat, 15.21 percent for the T-strat, and only 6.57 percent for the V-strat - compared to 43.21 percent for the E-strat and 56.88 percent for the S&P 500. Introducing partner stocks in the creation of the mispricing index seems to have a strongly positive effect on the overall portfolio risk. Also, the hit rate is surprisingly high - more than 68 percent of monthly returns are positive for the V-strat after transaction costs - considerably higher than the 63 percent of the general market or the 52 percent of the naive E-strat.

	Before transaction costs				After transaction costs				
	E-strat	G-strat	T-strat	V-strat	E-strat	G-strat	T-strat	V-strat	S&P 500
Mean return	0.0027	0.0062	0.0076	0.0097	0.0007	0.0043	0.0056	0.0075	0.0056
Standard error (NW)	0.0011	0.0011	0.0012	0.0012	0.0011	0.0011	0.0011	0.0011	0.0028
t-Statistic (NW)	2.4405	5.5958	6.5979	8.3132	0.6292	3.8825	4.8993	6.6890	2.0073
Minimum	-0.0811	-0.0675	-0.0614	-0.0339	-0.0834	-0.0712	-0.0627	-0.0426	-0.1856
Quartile 1	-0.0083	-0.0025	-0.0013	-0.0006	-0.0101	-0.0040	-0.0032	-0.0029	-0.0180
Median	0.0028	0.0049	0.0072	0.0080	0.0009	0.0031	0.0050	0.0058	0.0105
Quartile 3	0.0124	0.0131	0.0145	0.0174	0.0104	0.0114	0.0128	0.0153	0.0321
Maximum	0.1106	0.0748	0.0862	0.0875	0.1079	0.0723	0.0831	0.0792	0.1023
Standard deviation	0.0218	0.0165	0.0160	0.0170	0.0214	0.0163	0.0158	0.0166	0.0420
Skewness	0.4773	0.4036	0.6532	1.3088	0.3672	0.3421	0.6415	1.1303	-0.8622
Kurtosis	4.6373	4.0878	4.1694	4.2611	4.5531	4.2552	4.1976	3.5849	1.8690
Historical VaR 1%	-0.0612	-0.0388	-0.0292	-0.0258	-0.0639	-0.0402	-0.0307	-0.0274	-0.1165
Historical CVaR 1%	-0.0702	-0.0571	-0.0469	-0.0315	-0.0721	-0.0594	-0.0483	-0.0354	-0.1533
Historical VaR 5%	-0.0272	-0.0154	-0.0134	-0.0110	-0.0295	-0.0176	-0.0150	-0.0127	-0.0724
Historical CVaR 5%	-0.0460	-0.0277	-0.0249	-0.0202	-0.0484	-0.0297	-0.0268	-0.0226	-0.1008
Maximum drawdown	0.3148	0.1016	0.1285	0.0491	0.4321	0.1603	0.1521	0.0657	0.5688
Share with return > 0	0.5594	0.6748	0.7133	0.7308	0.5210	0.6119	0.6713	0.6853	0.6294

Table 3: Monthly return characteristics for the top 20 target stocks of the E-, G-, T-, and V-strat compared to the S&P 500 from January 1992 until October 2015 for the extremal partner selection algorithm. NW denotes Newey-West standard errors with six-lag correction.

Table 4 describes the trading statistics, which are very similar across all strategies. The target stocks open in almost all cases, meaning that close to 20 out of 20 possible target stocks are actually traded. On average, 9 round-trip trades are executed per six-month period and trades are open for 0.6 months. The similarity is potentially driven by the same underlying logic based on Bollinger bands - however the resulting returns are vastly different, depending on the information level contained in the different mispricing indices.

	E-strat	G-strat	T-strat	V-strat
Average number of target stocks traded per six-month period	20.00	19.91	19.91	19.96
Average number of round-trip trades per target stock	9.25	9.12	9.29	9.54
Standard deviation of number of round-trip trades per target stock	2.32	2.66	2.70	2.62
Average time target stocks are open in months	0.59	0.58	0.57	0.56
Standard deviation of time open, per target stock, in months	0.21	0.22	0.21	0.19

Table 4: Trading statistics for the top 20 target stocks of the E-, G-, T-, and V-strat from January 1992 until October 2015 for the extremal partner selection algorithm, per six-month trading period.

Table 5 depicts annualized risk and return metrics. After transaction costs, the E-strat produces a mere 0.57 percent p.a. - clearly inferior to an investment in the naive S&P 500 buy-and-hold strategy. Multivariate dependence modeling pays off - with 5.07 percent p.a. for the G-strat and 6.76 percent p.a. for the T-strat. The V-strat is best in class with a mean return of 9.25 percent p.a. and a Sharpe ratio of 1.12 - almost twice the Sharpe ratio of the top 20 pairs of Gatev et al. (2006) and on a similar level as Avellaneda and Lee (2010); Clegg and Krauss (2016). The V-strat has also a low annualized standard deviation of 5.77 percent and a low downside deviation[5] of a mere 2.24 percent. The downside deviation only accounts for approximately 39 percent of total standard deviation for the V-strat, compared to 71 percent for the S&P 500 - a favorable effect for investors, given that volatility is largely driven by upside deviations. The lower partial moment risk leads to a high Sortino ratio[6] of the V-strat of 4.14, compared to 2.51 for the T-strat and only 0.56 for the S&P 500.

	Before transaction costs				After transaction costs				
	E-strat	G-strat	T-strat	V-strat	E-strat	G-strat	T-strat	V-strat	S&P 500
Mean return	0.0303	0.0756	0.0937	0.1210	0.0057	0.0507	0.0676	0.0925	0.0581
Mean excess return	0.0038	0.0479	0.0656	0.0923	-0.0204	0.0236	0.0401	0.0645	0.0308
Standard deviation	0.0754	0.0572	0.0555	0.0590	0.0743	0.0566	0.0548	0.0577	0.1455
Downside deviation	0.0466	0.0282	0.0246	0.0193	0.0498	0.0307	0.0269	0.0224	0.1037
Sharpe ratio	0.0497	0.8375	1.1812	1.5642	-0.2740	0.4171	0.7315	1.1186	0.2119
Sortino ratio	0.6508	2.6800	3.8127	6.2666	0.1136	1.6489	2.5123	4.1380	0.5603

Table 5: Annualized returns and risk measures for the top 20 target stocks of the E-, G-, T-, and V-strat compared to the S&P 500 from January 1992 until October 2015 for the extremal partner selection algorithm.

In table 6, we evaluate the exposure of the V-strat after transaction costs to systematic sources of risk. We employ the Fama-French 3-factor model (FF3) discussed in Fama and French (1996), the Fama-French 3+2-factor model (FF3+2) outlined in Gatev et al. (2006), and the Fama-French 5-factor model (FF5) suggested in Fama and French (2015). The first model captures exposure to the general market (MKT), small minus big capitalization stocks (SMB) and high minus low book-to-market stocks (HML). The second model enhances

[5] The downside deviation is a lower partial moment risk measure, which only takes into account returns lower than a minimum acceptable return - zero percent in our case.

[6] The Sortino ratio scales mean return by the downside deviation.

the first one by a momentum factor and a short-term reversal factor. The third model enhances the first model by a factor capturing robust minus weak (RMW) profitability and a factor capturing conservative minus aggressive (CMA) investment behavior. We observe statistically significant monthly alphas ranging between 0.34 percent and 0.55 percent. The FF3+2 model has highest explanatory content with an adjusted R^2 of 0.2317. We observe loadings very close to zero for MKT, SMB, HML, RMW, and CMA - driven by the long-short portfolio we are constructing. Loading on the momentum factor is small, but significant and it has the expected negative sign. Loading on the reversal factor is well expressed and highly significant, underlining the mean-reversion that the strategy captures.

	FF3	FF3+2	FF5
(Intercept)	0.0055***	0.0034***	0.0053***
	(0.0010)	(0.0010)	(0.0010)
MKT	0.0030	−0.0567*	0.0133
	(0.0234)	(0.0222)	(0.0282)
SMB	−0.0669*	−0.0441	
	(0.0314)	(0.0279)	
HML	−0.0361	−0.0399	
	(0.0328)	(0.0293)	
Momentum		−0.0572**	
		(0.0192)	
Reversal		0.1602***	
		(0.0206)	
SMB5			−0.0331
			(0.0356)
HML5			−0.0176
			(0.0453)
RMW5			0.0772
			(0.0480)
CMA5			−0.0615
			(0.0650)
R^2	0.0169	0.2451	0.0270
Adj. R^2	0.0065	0.2317	0.0096
Num. obs.	286	286	286
RMSE	0.0164	0.0144	0.0164

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 6: Exposure to systematic sources of risk after transaction costs for the monthly returns of the top 20 target stocks of the V-strat from January 1992 until October 2015 for the extremal partner selection algorithm. Standard errors are depicted in parentheses.

4.3. Sub-period analysis

In figure 4, we exhibit the development of an investment of 1 USD after transaction costs in all four strategy variants, from January 1992 until October 2015. We observe stagnation of the naive E-strat, and a growth to 3.38 and 4.94 for the G-strat and the T-strat. The difference between the latter two strategies may be driven by the higher flexibility of the multivariate t-distribution, allowing for capturing tail dependencies. This fact can typically be found in financial data. As expected, the V-strat is best in class with a final value of 8.64. Particularly surprising is the smooth an steady growth we can observe for all multivariate strategies. Neither the bust of the dot-com bubble, nor 9/11, nor the Irak war, nor the global financial crisis seem to have had any major impact on capital growth. To our knowledge, no other academic statistical arbitrage or pairs trading strategy exhibits such steady growth up until recent times - see, for example, Avellaneda and Lee (2010); Clegg and Krauss (2016); Rad et al. (2016). In stark contrast, we see the development of the S&P 500, which shows strong swings and large drawdowns - especially during the time of the financial crisis.

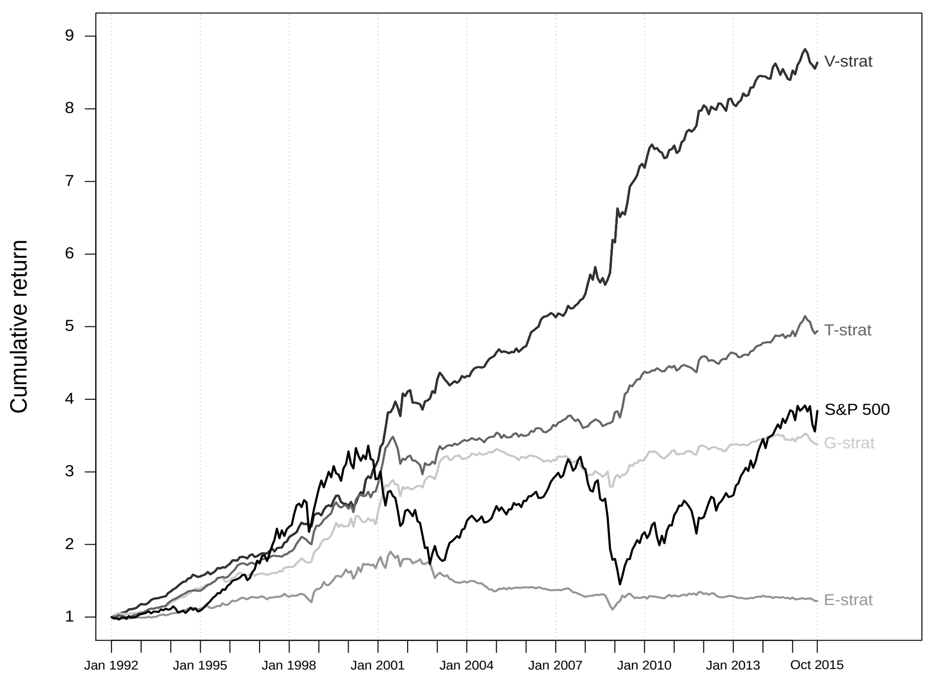
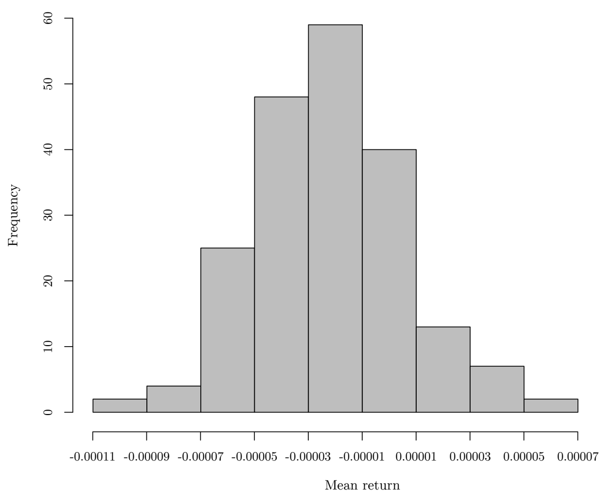


Figure 4: Development of an investment of 1 USD after transaction costs for the top 20 target stocks of the E-strat, G-strat, T-strat, and V-strat compared to the S&P 500 from January 1992 until October 2015 for the extremal partner selection algorithm.

4.4. Further analyses

4.4.1. Bootstrap trading

Furthermore, we compare the results of the V-strat with 200 bootstraps of random trading. Similar to Gatev et al. (2006), we use the entry signals of the selected pairs of the V-strat, but substitute the actual target stocks with random choices from the S&P 500 at that time. Table 5 reports the monthly return characteristics of the bootstrapped random tradings and the V-strat before transaction costs. As expected, the average monthly return of the random trading is very close to zero (-0.002 percent) - compared to the return of the V-strat of 12.10 percent before transaction costs. Hence, we may cautiously conclude that the constructed mispricing indices are truly idiosyncratic for the selected target stock and capture mispricings that may be profitably exploited.



	Bootstrap trading
Mean return	-0.00002250
Standard deviation	0.00002720
Median	-0.00002465

Figure 5: Monthly return characteristics before transaction costs for the top 20 target stocks of bootstrap trading from January 1992 until October 2015.

4.4.2. Robustness check

As stated previously, the trading threshold of one standard deviation ($k = 1$), the length of the moving average of 20 days ($d = 20$), and the number of 20 target stocks ($r = 20$) have been motivated based on the available literature - see section 3. Given that data snooping is a major issue across many applications, we would like to investigate the sensitivity of our

V-strat results in light of variations to these parameters. In table 7 we vary k, r, and d in two directions each and report annualized mean return and Sharpe ratio for the V-strat with extremal partner selection process.

Return		Before transaction costs			After transaction costs		
	k \ d	10	20	60	10	20	60
Top 10	0.5	0.1118	0.1040	0.1070	0.0681	0.0697	0.0881
	1	0.1175	0.1129	0.1054	0.0758	0.0840	0.0897
	2	0.1099	0.1243	0.1246	0.0764	0.1000	0.1114
Top 20	0.5	0.1106	0.1114	0.1025	0.0671	0.0785	0.0838
	1	0.1161	0.1210	0.1013	0.0749	0.0925	0.0865
	2	0.1184	0.1285	0.1201	0.0844	0.1045	0.1082
Top 30	0.5	0.1083	0.1071	0.1003	0.0648	0.0749	0.0814
	1	0.1148	0.1141	0.1017	0.0737	0.0862	0.0867
	2	0.1284	0.1284	0.1209	0.0935	0.1045	0.1085
Sharpe ratio		Before transaction costs			After transaction costs		
	k \ d	10	20	60	10	20	60
Top 10	0.5	1.1831	1.0576	1.1850	0.5890	0.6183	0.9247
	1	1.1387	1.0868	1.1122	0.6288	0.7357	0.9057
	2	0.6737	1.1606	1.1678	0.4059	0.8740	0.9993
Top 20	0.5	1.6265	1.5098	1.3774	0.8036	0.9639	1.0561
	1	1.5719	1.5642	1.3042	0.8688	1.1186	1.0523
	2	1.1365	1.4691	1.4665	0.7226	1.1378	1.2803
Top 30	0.5	1.7896	1.5992	1.4827	0.8590	0.9926	1.1285
	1	1.7788	1.6272	1.4449	0.9743	1.1294	1.1750
	2	1.4444	1.6125	1.6073	0.9546	1.2518	1.4124

Table 7: Yearly returns and Sharpe ratios for a varying number of target stocks (r), the number of days to use in the window (d), and the k-times of of the standard deviation of the V-strat to the S&P 500 from January 1992 until October 2015 for the extremal partner selection algorithm.

First of all, we see that our results are robust in light of parameter variations. After transaction costs, returns remain strictly positive. Lower values can generally be found at lower levels of d, given that the higher trading frequency generates higher transaction costs which are apparently not compensated for by higher returns. In terms of Sharpe ratio, we see that optimal strategies seem to be located in the lower right corner. A larger number of target stocks decreases portfolio standard deviation (i.e., the denominator of the Sharpe ratio) and higher k and a longer moving average d tend to improve returns (i.e., the numerator of the Sharpe ratio). Higher k decreases trading frequency (hence transaction costs) and a longer d leads to a more stable moving average - both of them seem to be beneficial to strategy returns. These findings reinforce our initial parameter choices. Clearly, we have not hit an optimum, but we find robust trading results irrespective of variations to our baseline setting.

4.4.3. Dependence structure

In this section, we investigate which copula families are chosen when fitting the V-model. Figure 6 summarizes the fitted bivariate copula families over time - in total and for each tree of the vine copula.

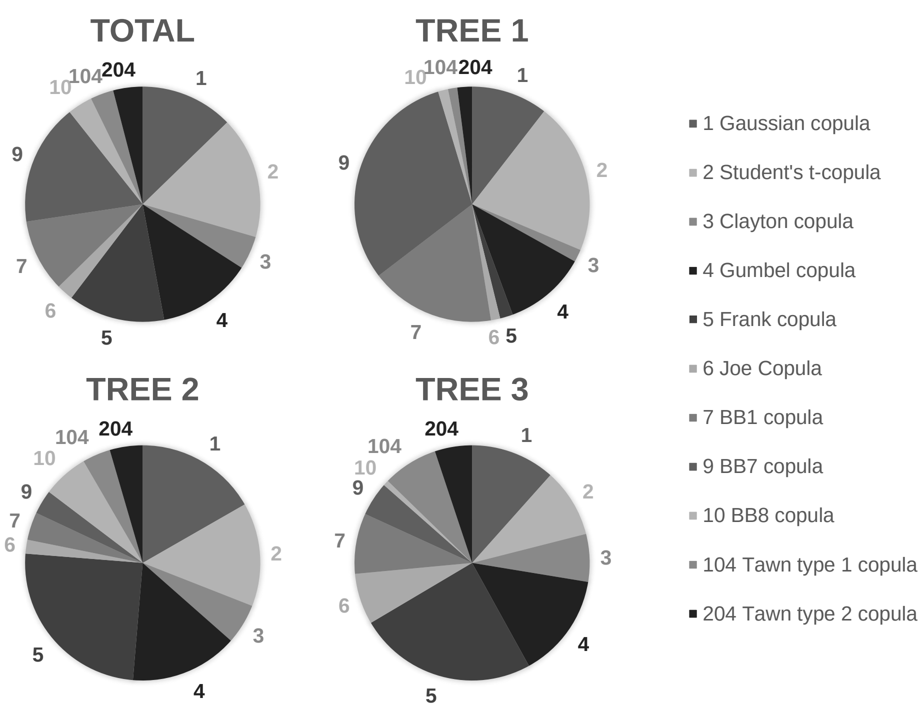


Figure 6: Usage of different copula families during the fitting process, for each tree of the vine copula and in total, from January 1990 until April 2015.

We find that the majority of copula families that are used in the first tree can model tail dependencies both in the upper and the lower tail (e.g., the t-, the BB1, and the BB7 copula in over 68 percent of the time). By contrast, the Gaussian copula - often a simple model in the context of financial data - is barely deployed at the first tree.

The more advanced the tree, the higher the diversity of families that are used in the modeling process. The focus shifts towards asymmetric copulas (only one kind of tail dependence, e.g., Clayton, Gumbel, Tawn copula, over 33 percent in the third tree) and simpler copulas (no tail dependence, e.g., Gaussian and Frank copula, over 36 percent in the third tree). We presume that this forking occurs because the first tree is directly fitted on rank-transformed financial market data, whereas the second and third tree is fitted on conditional values. The latter do not exhibit the usual properties of financial data anymore, but often simpler or more exotic, asymmetric dependence structures.

For comparing the families of copulas that are chosen for the first tree in the fitting step over time, we classify the discussed copulas by their tail behavior. We distinguish between copulas that can model tail dependence in both tails, the upper or lower tail only, and no tail dependence at all. Figure 7 shows the tree one occurrence of any of those four classes over the time from January 1990 to April 2015, where each point on the horizontal axis refers to the families that are used in the two years prior to the respective point.

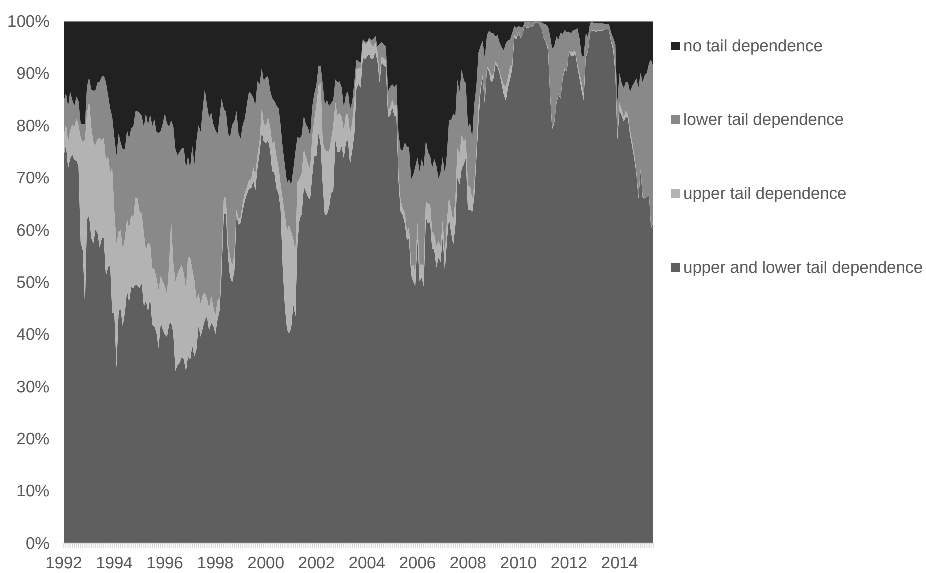


Figure 7: Share of fitted copula families of the first tree from January 1990 until April 2015 in two-year moving windows. A point on the horizontal axis refers to a time window two years prior to the respective point.

We can observe that in times of financial crises the families that can model tail dependence are more likely to be chosen. In fact, the resulting vine copulas of study periods whose data include returns of the financial crisis (2008 - 2012) almost exclusively use families that can model at least one kind of tail dependence. Furthermore, we find that lower tail dependence does occur more often than upper tail dependence, especially in recent times. This finding is in line with Fortin and Kuzmics (2002) who state that the general tail dependence of return pairs on stock indices displays a low tail dependence in the upper, and a high tail dependence in the lower tail.

5. Conclusions

We have developed several different statistical arbitrage strategies, where we trade portfolios of stocks against other portfolios of stocks. In this respect, we make several contributions to the literature.

At first, we theoretically generalize two-dimensional copula-based pairs trading, as initially suggested in Xie et al. (2014) and largely tested in Rad et al. (2016) to higher dimensions with a focus on vine copulas. In this context, we implement four different partner selection algorithms - a naive selection, based on all bivariate correlations, an extended selection relying on the multivariate Spearman's ρ of Schmid and Schmidt (2007), a geometric selection based on the sum of squared distances from the diagonal in a hypercube and, an extremal selection leveraging deviations from multivariate independence, as initially proposed in Mangold (2015).

Also, and more importantly, we implement different strategy variants, reflecting different degrees of flexibility in multivariate dependence modeling. The naive E-strat only takes into account a stock's own history in assessing mispricings. The G-strat and T-strat rely on the target stock and three partner stocks. Mispricings are determined by leveraging the multivariate Gaussian and the multivariate t-distribution. Finally, the V-strat allows for a multitude of linear and nonlinear dependence structures by using vine copulas as instrument of choice.

In a large-scale empirical application to the S&P 500, we find that the V-strat produces statistically and economically significant returns of 9.25 percent p.a. after transaction costs.

A portion of these returns can be explained by common sources of systematic risk, (mostly, a short-term reversal factor), but monthly alpha still remains at 0.34 percent. Especially surprising is the low (tail) risk associated with the V-strat - with a maximum drawdown of only 6.57 percent, Value at Risk at the 1% level of merely -2.74 percent, and annualized standard and downside deviations at 5.77 and 2.24 percent, respectively. The T-strat comes in second, just before the G-strat which has increasingly stronger restrictions on the multivariate dependence structure and fewer parameters. By contrast, the naive E-strat produces annualized returns close to zero - thus powerfully underlining the importance of the partner portfolio. The results are positive over time up until 2015 - even during times of severe market turmoil, i.e., the dot-com crisis or the global financial crisis.

Finally, we observe that the most-frequently chosen copula family in the vine graph varies over time. In the nineties, copulas allowing for (i) no tail dependence, (ii) lower or (iii) upper tail dependence, and (iv) upper and lower tail dependence can be observed. During the dot-com bubble, the share of copulas allowing for modeling upper and lower tail dependence severely increases to over 70 percent, only to slump back down to 40 percent shortly thereafter. In the following periods, this copula family increasingly dominates the dependence structure, with shares close to 100 percent during the global financial crisis. We may cautiously conclude that highly volatile market environments with increasing levels of associated insecurity demand more flexible dependence structures.

Bibliography

Aas, K., Czado, C., Frigessi, A., Bakken, H., 2009. Pair-copula constructions of multiple dependence. Insurance: Mathematics and Economics 44 (2), 182–198.

Avellaneda, M., Lee, J.-H., 2010. Statistical arbitrage in the US equities market. Quantitative Finance 10 (7), 761–782.

Bacon, C. R., 2008. Practical portfolio performance: Measurement and attribution, 2nd Edition. Wiley finance. Wiley, Chichester, England and Hoboken, NJ.

Bedford, T., Cooke, R. M., 2001. Probability density decomposition for conditionally dependent random variables modeled by vines. Annals of Mathematics and Artificial Intelligence 32 (1-4), 245–268.

Bedford, T., Cooke, R. M., 2002. Vines: A new graphical model for dependent random variables. Annals of Statistics, 1031–1068.

Bollinger, J., 1992. Using bollinger bands. Stocks & Commodities 10 (2), 47–51.

Brechmann, E. C., Czado, C., 2013. Risk management with high-dimensional vine copulas: An analysis of the Euro Stoxx 50. Statistics & Risk Modeling 30 (4), 307–342.

Brechmann, E. C., Czado, C., Aas, K., 2012. Truncated regular vines in high dimensions with application to financial data. Canadian Journal of Statistics 40 (1), 68–85.

Brechmann, E. C., Czado, C., Paterlini, S., 2014. Flexible dependence modeling of operational risk losses and its impact on total capital requirements. Journal of Banking & Finance 40, 271 – 285.

Chen, H., Chen, S. J., Li, F., 2012. Empirical investigation of an equity pairs trading strategy. Working paper, University of British Columbia, University of Michigan.

Cherubini, U., Luciano, E., Vecchiato, W., 2004. Copula methods in finance. John Wiley & Sons.

Clegg, M., Krauss, C., 2016. Pairs trading with partial cointegration. FAU Discussion Papers in Economics, University of Erlangen-Nürnberg.

Czado, C., 2010. Pair-copula constructions of multivariate copulas. In: Jaworski, P., Durante, F., Härdle, W. K., Rychlik, T. (Eds.), Copula theory and its applications: Proceedings of the workshop held in Warsaw, 25-26 September 2009. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 93–109.

Ding, P., 2016. On the conditional distribution of the multivariate t distribution. The American Statistician (just-accepted), 00–00.

Dissmann, J., Brechmann, E. C., Czado, C., Kurowicka, D., 2013. Selecting and estimating regular vine copulae and application to financial returns. Computational Statistics & Data Analysis 59, 52–69.

Do, B., Faff, R., 2010. Does simple pairs trading still work? Financial Analysts Journal 66 (4), 83–95.

Dragulescu, A. A., 2014. xlsx: Read, write, format Excel 2007 and Excel 97/2000/XP/2003 files. R package version 0.5.7.

Eaton, M. L., 1983. Multivariate statistics: A vector space approach. John Wiley & Sons, Inc.

Eddelbuettel, D., Francois, R., Allaire, J., Ushey, K., Kou, Q., Bates, D., Chambers, J., 2016. Rcpp: Seamless R and C++ integration. R package version 0.12.4.

Fama, E. F., French, K. R., 1996. Multifactor explanations of asset pricing anomalies. The Journal of Finance 51 (1), 55.

Fama, E. F., French, K. R., 2015. A five-factor asset pricing model. Journal of Financial Economics 116 (1), 1–22.

Fischer, M., Köck, C., Schlüter, S., Weigert, F., 2009. An empirical analysis of multivariate copula models. Quantitative Finance 9 (7), 839–854.

Fortin, I., Kuzmics, C., 2002. Tail-dependence in stock-return pairs. Intelligent Systems in Accounting, Finance and Management 11 (2), 89–107.

Gatev, E., Goetzmann, W. N., Rouwenhorst, K. G., 2006. Pairs trading: Performance of a relative-value arbitrage rule. Review of Financial Studies 19 (3), 797–827.

Gohel, D., 2016. ReporteRs: Microsoft Word, Microsoft PowerPoint and HTML documents generation. R package version 0.8.6.

Gonzalez-Fernandez, Y., Soto, M., 2015. vines: Multivariate dependence modeling with vines. R package version 1.1.3.

Hofert, M., Kojadinovic, I., Maechler, M., Yan, J., 2015. copula: Multivariate dependence with copulas. R package version 0.999-13.

Hothorn, T., Zeileis, A., Farebrother, R. W., Cummins, C., 2015. lmtest: Testing linear regression models. R package version 0.9-34.

Joe, H., 1994. Multivariate extreme-value distributions with applications to environmental data. Canadian Journal of Statistics 22 (1), 47–64.

Joe, H., 1996. Families of m-variate distributions with given margins and m (m-1)/2 bivariate dependence parameters. Lecture Notes-Monograph Series, 120–141.

Joe, H., 1997. Multivariate models and multivariate dependence concepts. CRC Press.

Kotz, S., Nadarajah, S., 2004. Multivariate t-distributions and their applications. Cambridge University Press.

Krauss, C., 2016. Statistical arbitrage pairs trading strategies: Review and outlook. Journal of Economic Surveys, forthcoming, http://onlinelibrary.wiley.com/doi/10.1111/joes.12153/full.

Krauss, C., Stübinger, J., 2015. Nonlinear dependence modeling with bivariate copulas: Statistical arbitrage pairs trading on the S&P 100. IWQW Discussion Paper Series, University of Erlangen-Nürnberg.

Lau, C. A., Xie, W., Wu, Y., 2016. Multi-dimensional pairs trading using copulas. Working paper, Nanyang Technological University.

Leifeld, P., 2015. texreg: Conversion of R regression output to LaTeX or HTML tables. R package version 1.35.

Liew, R. Q., Wu, Y., 2013. Pairs trading: A copula approach. Journal of Derivatives & Hedge Funds 19 (1), 12–30.

Liu, C., 1994. Statistical analysis using the multivariate t distribution. Ph.D. thesis, Harvard University.

Low, R. K. Y., Alcock, J., Faff, R., Brailsford, T., 2013. Canonical vine copulas in the context of modern portfolio management: Are they worth it? Journal of Banking & Finance 37 (8), 3085–3099.

Lumley, T., Zeileis, A., 2015. sandwich: Robust covariance matrix estimators. R package version 2.3-3.

Mangold, B., 2015. A multivariate linear rank test of independence based on a multiparametric copula with cubic sections. IWQW Discussion Paper Series, University of Erlangen-Nürnberg.

Mina, J., Xiao, J. Y., 2001. Return to RiskMetrics: The evolution of a standard. RiskMetrics Group.

Nadarajah, S., Kotz, S., 2005. Mathematical properties of the multivariate t distribution. Acta Applicandae Mathematica 89 (1-3), 53–84.

Nelsen, R. B., 2007. An introduction to copulas. Springer Science & Business Media.

Perlin, M. S., 2007. M of a kind: A multivariate approach at pairs trading. Working paper, ICMA/Reading University.

Peterson, B. G., Carl, P., 2014. PerformanceAnalytics: Econometric tools for performance and risk analysis. R package version 1.4.3541.

Pfaff, B., McNeil, A., 2014. QRM: Provides R-language code to examine quantitative risk management concepts. R package version 0.4-10.

Rad, H., Low, Rand Kwong Yew, Faff, R. W., 2016. The profitability of pairs trading strategies: Distance, cointegration and copula methods. Quantitative Finance, 1–18.

Rmetrics Core Team, Wuertz, D., Setz, T., 2014. fCopulae: Rmetrics - bivariate dependence structures with copulae. R package version 3011.81.

Rmetrics Core Team, Wuertz, D., Setz, T., Chalabi, Y., 2015. timeSeries: Rmetrics - financial time series objects. R package version 3012.99.

Ryan, J. A., 2015. quantmod: Quantitative financial modelling framework. R package version 0.4-5.

Ryan, J. A., Ulrich, J. M., 2014. xts: eXtensible time series. R package version 0.9-7.

Schepsmeier, U., Stoeber, J., Brechmann, E. C., Graeler, B., Nagler, T., Erhardt, T., 2015. VineCopula: Statistical inference of vine copulas. R package version 1.6.

Schmid, F., Schmidt, R., 2007. Multivariate extensions of Spearman's rho and related statistics. Statistics & Probability Letters 77 (4), 407–416.

Simpson, G. L., 2015. permute: Functions for generating restricted permutations of data. R package version 0.8-4.

Sklar, M., 1959. Fonctions de répartition à n dimensions et leurs marges. Publications de l'Institut de Statistique de l'Université de Paris (8), 229–231.

S&P Dow Jones Indices, 2015. Equity S&P 500. http://us.spindices.com/indices/equity/sp-500.

Stander, Y., Marais, D., Botha, I., 2013. Trading strategies with copulas. Journal of Economic and Financial Sciences 6 (1), 83–107.

Trapletti, A., Hornik, K., 2016. tseries: Time series analysis and computational finance. R package version 0.10-35.

Ulrich, J., 2015. TTR: Technical trading rules. R package version 0.23-0.

Varadhan, R., 2015. condMVNorm: Conditional multivariate normal distribution. R package version 2015.2-1.

Weiß, G. N. F., Supper, H., 2013. Forecasting liquidity-adjusted intraday Value-at-Risk with vine copulas. Journal of Banking & Finance 37 (9), 3334–3350.

Whelan, N., 2004. Sampling from Archimedean copulas. Quantitative Finance 4 (3), 339–352.

Wickham, H., Francois, R., 2016. dplyr: A grammar of data manipulation. R package version 0.5.0.

Wuertz, D., 2013. fUnitRoots: Trends and unit roots. R package version 3010.78.

Xie, W., Liew, Q. R., Wu, Y., Zou, X., 2014. Pairs trading with copulas. Working paper, Nanyang Technological University.

Xie, W., Wu, Y., 2013. Copula-based pairs trading strategy. In: Asian Finance Association (AsFA) 2013 Conference.

Zeileis, A., Grothendieck, G., Ryan, J. A., 2015. zoo: S3 infrastructure for regular and irregular time series (z's ordered observations). R package version 1.7-12.