A New Copula for Modeling Tail Dependence

November 10, 2010

J. Holman and G. Ritter

Abstract

We introduce a new copula which simultaneously allows fully-general correlation structures in the bulk of a multivariate distribution and an arbitrarily high degree of dependence in the left tails. This is ideally suited for modeling financial assets which may display moderate correlation in normal times, but which experience simultaneous left tail events, such as during a financial crisis. Our new copula is shown to be fully flexible in the sense that the user can specify a desired structure for a sequence of increasingly dire events in the left tail, while still retaining the same correlation structure in the bulk. We illustrate the use of this copula with an application to hedge fund returns.

1 Introduction

Modeling the bulk and the tails of a multivariate distribution simultaneously requires a balancing act that is difficult if not impossible to achieve with the well-known copulas. In this paper we construct a multivariate density which resembles an elliptical distribution in its bulk, but has the following two properties in its tails:

1. Association that is more extreme in the tails than in the bulk of the distribution
2. Association that is asymmetric between the upper and lower tails

We achieve this by introducing a new "Cube copula", that can accommodate arbitrarily precise modeling of the joint tails. We then describe two methods to break the link between the tails and the bulk of the distribution, copula mixing and copula nesting, and we illustrate with an application to hedge fund returns.

Copulas have become a key addition to the financial modeler's toolkit in the past decade. Formally, a copula is a multivariate cumulative distribution function defined on the unit cube such that its marginals are uniform. A key theorem, due to Sklar [34], lays out the use of such a function – for any set of marginal distributions defined on the reals, and any joint distribution function, there exists a copula that reproduces the joint distribution when applied to the marginal distribution functions. Modulo continuity conditions on the marginals, this copula is unique. The reason for the appeal to financial modeling is that researchers often have better information about marginal distributions than joint distributions, and a copula approach lets them fully use this information, and then choose from a small menu of well-known copulas to splice them into a joint distribution.

The set of well-known copulas leaves much to be desired when modeling financial returns. A brief, stylized history of univariate return modeling is instructive. Early modeling of financial (log) returns relied heavily on the Gaussian distribution, primarily for its analytical tractability. However, returns typically exhibit kurtosis much greater than that of the Gaussian, and so modelers expanded into fatter tailed distributions, such as the Student's-t and Generalized Extreme Value. In the multivariate case, the Gaussian copula played a similarly tractable role in early applications. The Gaussian copula is defined (via Sklar) as that copula which composes univariate Gaussians into a multivariate Gaussian. Just as in the univariate case, the Gaussian copula imposes a particular structure in the joint tails of multivariate distributions that is often empirically violated. Specifically, and we define this concept rigorously in section 2.3.1, the Gaussian copula requires that variables become asymptotically independent in the tails, while in practice, dependence even in extreme tail events often remains strong.

The "fix" has often been to simply move to a copula with a fatter joint-tail, such as the Student's-t copula, which is that copula that composes Student's-t marginals into a multivariate Student's-t. Other popular copulas with non-zero asymptotic tail dependence are the Archimedean copulas, which encompasses the Clayton, Frank, and Gumbel copulas. However, in making this move, one loses control of the ability to model both the bulk of the multivariate distribution and the joint tails. For example, the bivariate Student's-t copula has two parameters, η, the degrees of freedom, and correlation ρ. The amount of left tail dependence is a decreasing function of η and an increasing function of ρ. Thus ρ and η can serve as tuning parameters for tail dependence. However, neither of these parameters changes solely tail dependence. The ρ parameter is in fact the correlation of the bivariate distribution in the case of Student's-t marginals, so that a side effect of increasing tail dependence via increasing ρ is a stronger dependence in the bulk of the distribution. Frequently in applications we will have an estimate of correlations in the bulk of the distribution, and want to increase tail dependence while holding fixed the correlation in the bulk. The η parameter is less intuitive, but has equally unattractive properties upon scaling. First, η has limited ability to generate extreme tail dependence. Figure 1 plots the tail dependence as a function of η for several values of ρ.

For example, with $\rho = 0.3$ the maximum achievable tail dependence is 0.41. Furthermore, very low values of η generate behavior that most would consider odd in the upper left and lower right quadrants of the multivariate.

Fig. 2 plots a Monte Carlo simulation of the t copula with $\rho = 0.3$ for $\eta \in \{10, 5, 2\}$. Note the extreme right tail outliers in univariate Y that are associated with left tail outliers in univariate X with increasing frequency as η decreases. The intuition for this behavior is that the t copula needs to enforce correlation ρ, and so must balance out what we call the "double-tail" observations with "anti-double-tail" points.

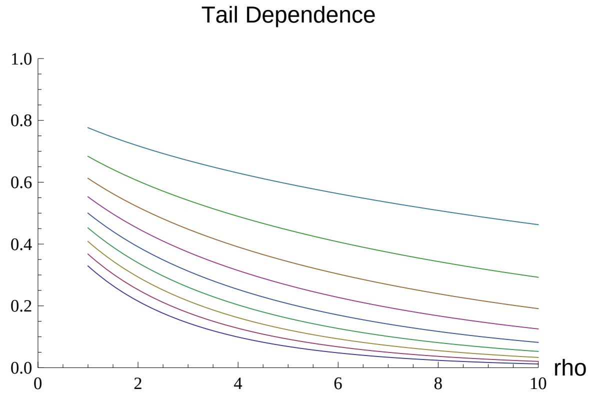

Figure 1: Tail dependence λ_L as a function of η for several $\rho = 0.1 \ldots 0.9$.

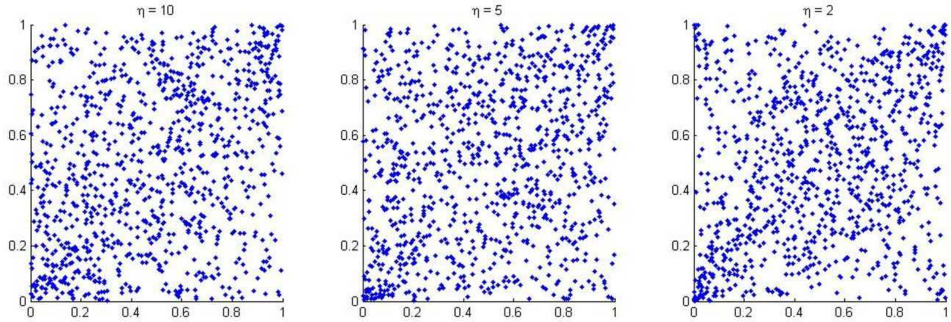

Figure 2: the t copula with $\rho = 0.3$ for $\eta \in \{10, 5, 2\}$

2 The Cube Copula

2.1 Construction in n-dimensions

Definition 1. *Let X be a random variable on the sample space $\Omega_n = [0,1]^n$. Let $\mathbf{a}$ be an element of Ω_n. If $X_i \leq a_i$ then we say that X* experiences an i-th tail event *(with respect to $\mathbf{a}$).*

For an arbitrary set of indices $I \subseteq \{1, \ldots, n\}$, define the set

$$T_I = \{x \in \Omega_n \mid x_j \leq a_j \text{ if and only if } j \in I\}.$$

Furthermore when I has cardinality $|I| = k$, we refer to T_I as a k-tail region.

In other words, a k-tail region is a subset of the sample space with exactly k of the variables experiencing tail events simultaneously. For any fixed k, the number of k-tail regions equals $\binom{n}{k}$.

Let $\tau_{k,n}$ denote the union of all k-tail regions in Ω_n. The sets $\tau_{k,n}$ for various k form a partition of Ω_n; that is,

$$\Omega_n = \bigcup_{k=0}^{n} \tau_{k,n}, \quad \tau_{k,n} \cap \tau_{\ell,n} = \emptyset \text{ if } k \neq \ell. \tag{1}$$

Figure 3 illustrates the 2-tail and 3-tail regions for $n = 3$ and $\mathbf{a} = (0.1, 0.1, 0.1)$

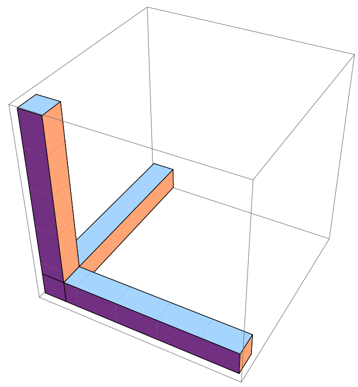

Figure 3: The 2-tail and 3-tail regions for $n = 3$

Theorem 1. *Consider a real vector $q = (q_0, \ldots, q_n) \in \mathbb{R}^{n+1}$ and the corresponding density on Ω_n:*

$$p_c(x) = q_k \text{ when } x \in \tau_{k,n}. \tag{2}$$

Then p_c is a copula density if and only if conditions (a) and (b) below are met:

(a) Total Density Condition

Defining $v_{k,n}[\mathbf{a}] := \mathrm{vol}(\tau_{k,n})$, one has

$$1 = \sum_k q_k \, v_{k,n}[\mathbf{a}], \tag{3}$$

(b) Unit Marginals Condition

Defining $\hat{\mathbf{a}}^{(j)} = (a_1, a_2, \ldots, a_{j-1}, a_{j+1}, \ldots, a_n)$, one has for all j:

$$M_{j,1} := \sum_{i=1}^{n} q_i v_{i-1,n-1}[\hat{\mathbf{a}}^{(j)}] \quad = \quad 1 \tag{4}$$

$$M_{j,2} := \sum_{i=1}^{n} q_{i-1} v_{i-1,n-1}[\hat{\mathbf{a}}^{(j)}] \quad = \quad 1. \tag{5}$$

Furthermore, one has

$$v_{k,n}[\mathbf{a}] = \sum_{|I|=k} \mathrm{vol}(T_I) = \sum_{|I|=k} \Big[\prod_{i\in I} a_i \prod_{j\in I^c} (1-a_j) \Big]. \tag{6}$$

where the sum is over the $\binom{n}{k}$ *subsets* $I \subseteq \{1,2,\ldots,n\}$ *with length* k*, and* I^c *is the complement of* I*.*

Proof. In order that (2) be a probability density, we must have

$$\int_{\Omega_n} p_c(x)\,dx = 1 = \sum_k v_{k,n}[\mathbf{a}] q_k, \quad \text{where} \quad v_{k,n}[\mathbf{a}] := \mathrm{vol}(\tau_{k,n}). \tag{7}$$

This is condition (a) in the theorem. To prove (6), and thus derive an explicit formula for (3), we must calculate $v_{k,n}[\mathbf{a}]$. Note that $\tau_{k,n}$ is the union of $\binom{n}{k}$ connected components, each a rectangular prism. Let $I \subseteq \{1,2,\ldots,n\}$; then for $k = |I|$ the region T_I defined above is one of the $\binom{n}{k}$ components of $\tau_{k,n}$. The volume of such a region is the product of its side lengths, and the regions are disjoint as noted above in eq. (1). This establishes (6).[1]

In order that $p_c(x)$ forms a copula density, we require the condition of uniform marginals:

$$m_j(x_j) = \int_{[0,1]^{n-1}} p_c(x) \prod_{i\neq j} dx_i = 1. \tag{8}$$

We claim that each m_j is a step function on $[0,1]$, and more specifically can be written in the form

$$m_j(x_j) = \begin{cases} M_{j,1} & \text{if } x_j \le a_j, \\ M_{j,2} & \text{if } x_j > a_j \end{cases}. \tag{9}$$

where M is an $n \times 2$ matrix that is a polynomial function of $\mathbf{q}$ and $\hat{\mathbf{a}}^{(j)}$. We now focus on establishing the representation (9) by calculating the required integrals explicitly.

The marginal $m_j(x_j)$ is given by the integral of p_c over the lower-dimensional "slice"

$$S(x_j) \;=\; \{y \in \Omega_n : y_j = x_j\}$$

with respect to $n-1$ dimensional volume, i.e.

$$m_j(x_j) = \int_{S(x_j)} p_c(z) \prod_{i\neq j} dz_i. \tag{10}$$

Since $p_c(x)$ is a step function, (10) can be written as a finite sum of density times volume; it remains to determine the explicit form of this sum.

[1] In the special case $\mathbf{a} = (a,a,\ldots,a)$, this simplifies to $v_{k,n} = \binom{n}{k} a^k (1-a)^{n-k}$.

The slice $S(x_j)$ is isomorphic Ω_{n-1} and hence has the same combinatorial structure as the original problem in one lower dimension. For each $k = 0, 1, \ldots, n-1$ the slice $S(x_j)$ has $\binom{n-1}{k}$ connected components which play the role of k-tails in Ω_{n-1}. The side lengths for the resulting partition of Ω_{n-1} are determined by the truncated vector $\hat{\mathbf{a}}^{(j)}$ defined in part (b) of the theorem. Hence our strategy is to use a form of induction on n.

First consider the case $x_j \le a_j$; then the $\mathbf{q}$-vector relevant for calculating the $(n-1)$ dimensional density on the slice is $(q_1, q_2, \ldots, q_n)$, because $S(x_j)$ doesn't intersect the 0-tail in Ω_n. The set of points in $S(x_j)$ where $p_c = q_i$ has the structure of an $(i-1)$ tail region in Ω_{n-1}, and is a subset of an i-tail in Ω_n. Hence

$$M_{j,1} = \sum_{i=1}^{n} q_k v_{i-1,n-1}[\hat{\mathbf{a}}^{(j)}]. \tag{11}$$

For the case $x_j > a_j$ the logic is the same, but instead of working with $(q_1, q_2, \ldots, q_n)$ we have to work with $(q_0, q_1, \ldots, q_{n-1})$ because $S(x_j)$ intersects the 0-tail but not the n-tail in Ω_n. In this case the set of points in $S(x_j)$ where $p_c = q_i$ has the structure is a subset of an $(i-1)$-tail in Ω_n. Hence we shift the index on q in (11) to yield

$$M_{j,2} = \sum_{i=1}^{n} q_{i-1} v_{i-1,n-1}[\hat{\mathbf{a}}^{(j)}]. \tag{12}$$

With this, we establish that formulas (4) and (5) are correct, and complete the proof. □

Note also that for any $\mathbf{a}$ and for any n there is always at least one $\mathbf{q}$ that trivially defines a Cube copula, namely $\mathbf{q} = (1, 1, \ldots, 1)$. In fact we conjecture that there are always infinitely many such consistent $\mathbf{q}$; the argument involves the degrees of freedom allowed in the Unit Marginals Condition.

For the remainder of the paper we assume that $\mathbf{a} = (a, a, \ldots, a)$; our results below hold more generally but this assumption simplifies notation considerably. For instance, with this assumption, the Total Density Condition and Unit Marginals Conditions of Theorem 1 simplify to just three equations:

(a) *Total Density Condition*

$$1 = \sum_k v_{k,n} q_k \text{ where } v_{k,n} = \binom{n}{k} a^k (1-a)^{n-k} \tag{13}$$

(b) *Unit Marginals Condition*

$$1 = \sum_{i=1}^{n} q_i v_{i-1,n-1} = \sum_{i=1}^{n} q_{i-1} v_{i-1,n-1}. \tag{14}$$

2.2 Examples: 2 and 3 dimensions

In $n = 2$ dimensions eq. (6) implies

$$v = (v_{0,2}, v_{1,2}, v_{2,2}) = ((1-a)^2, 2a(1-a), a^2).$$

We then have the copula conditions as in Theorem 1:

$$\begin{aligned}
(1-a)^2 q_0 + 2a(1-a)q_1 + a^2 q_2 &= 1 \\
(1-a)q_1 + aq_2 &= 1 \\
(1-a)q_0 + aq_1 &= 1
\end{aligned}$$

These equations can of course be solved explicitly; we choose to express the solution in terms of q_2, the density in the double-tail region:

$$q_0 = \frac{1 - 2a + a^2 q_2}{(a-1)^2}, \qquad q_1 = \frac{aq_2 - 1}{a-1}.$$

Positivity of q_0, q_1 give the constraint that $(2a-1)a^{-2} \le q_2 \le a^{-1}$. This implies $0 \le q_0 \le (1-a)^{-1}$ and $0 \le q_1 \le a^{-1}$. These constraints are useful in maximum-likelihood optimization to guide the optimizer and ensure that it doesn't go outside the valid parameter space.

In $n = 3$ dimensions eq. (6) implies

$$v = (v_{0,3}, \ldots, v_{3,3}) = ((1-a)^3, 3a(1-a)^2, 3a^2(1-a), a^3).$$

Again as in Theorem 1, the conditions for a copula are:

$$\begin{aligned}
(1-a)^3 q_0 + 3a(1-a)^2 q_1 + 3a^2(1-a)q_2 + a^3 q_3 &= 1 \\
(1-a)^2 q_1 + 2a(1-a)q_2 + a^2 q_3 &= 1 \\
(1-a)^2 q_0 + 2a(1-a)q_1 + a^2 q_2 &= 1
\end{aligned}$$

In any number n of dimensions, we can represent this system as $Aq = \mathbf{1}$ where A is a $3 \times n+1$ matrix. Hence in n dimensions there is a space of solutions (copulas) which is of dimension $n - 2 + p$, where p is the dimension of the null space of A. We showed above that for $n = 2$, we have $p = 1$ and the solution space is one-dimensional.

2.3 Properties of the Cube Copula

Above we referred to "tail dependence" informally, but there is a natural definition that is standard within the copula literature. We provide this definition, and compute the tail dependence of our Cube copula. This tail dependence, and its finite analog, serve as parameters that we can estimate/calibrate when applying the Cube copula to data. In addition, we consider three broader measures of association – Pearson's product-moment correlation, Spearman's rank correlation, and Kendall's τ – that we will use later to parameterize the relationships that exist within the bulk of a multivariate distribution.

2.3.1 Tail Dependence

Tail dependence between a pair of distributions is typically formalized via a conditional tail probability called the coefficients of tail dependence:

Definition 2. *Let X, Y be random variables with cdfs F_X and F_Y, with H as their bivariate cdf. The **lower-u tail dependence** of H is*

$$\begin{aligned}\lambda_{L-u} &= P\left(Y < F_Y^{-1}(u)\,|X < F_X^{-1}(u)\right)\\ &= \frac{H(F_X^{-1}(u), F_Y^{-1}(u))}{u}\end{aligned}$$

Note that the bivariate cdf H is symmetric by its definition, and so lower-u tail dependence is a symmetric property.

*Similarly define the **upper-u tail dependence** of H as*

$$\begin{aligned}\lambda_{U-u} &= P\left(Y > F_Y^{-1}(u)\,|X > F_X^{-1}(u)\right)\\ &= \frac{1-2u-H(F_X^{-1}(u), F_Y^{-1}(u))}{1-u}\end{aligned}$$

*The limits of these quantities as are the **lower tail dependence** and **upper tail dependence**, respectively (provided that the limits exist):*

$$\lim_{u\searrow 0}\lambda_{L-u} = \lambda_L, \quad \lim_{u\nearrow 1}\lambda_{U-u} = \lambda_U.$$

We can use Sklar's theorem to establish tail dependence as a property of the copula, independent of the marginals.

Theorem 2. *Sklar (1959) Let X, Y be random variables with cdfs F_X and F_Y, with H as their bivariate cdf. There exists a copula C such that*

$$H(x, y) = C(F_X(x), F_Y(y)).$$

If F_X and F_Y are continuous, then C is unique.

If C is H's corresponding copula in the definition of λ_{L-u} and λ_{U-u}, then by Sklar's Theorem, $\lambda_{L-u} = \frac{C(u,u)}{u}$ and $\lambda_{U-u} = \frac{1-2u-C(u,u)}{1-u}$.

Note that the CDF of the Cube copula, F, in the bivariate case is given by

$$
\begin{aligned}
F(u,v) &= \int_0^v \int_0^u f(x,y)dxdy \\
&= \begin{cases} q_2uv & \text{if} \quad u \le a, v \le a \\ u[q_2a + q_1(v-a)] & \text{if} \quad u \le a < v \\ v[q_2a + q_1(u-a)] & \text{if} \quad v \le a < u \\ q_2a^2 + q_1a[(v-a)+(u-a)] + q_0(u-a)(v-a) & \text{if} \quad u > a, v > a \end{cases}
\end{aligned}
$$

So,

$$
\lambda_{L-u} = \begin{cases} q_2u & \text{if } u \le a \\ [q_2a^2 + 2q_1a(u-a) + q_0(u-a)^2]/u & \text{if } u > a \end{cases}
$$

and

$$
\lambda_{U-u} = \begin{cases} \frac{1-2+q_2u}{1-u} & \text{if} \quad u \le a \\ \frac{1-2u+q_2a^2+2q_1a(u-a)+q_0(u-a)^2}{1-u} & \text{if} \quad u > a \end{cases}
$$

Note in particular, λ_L and λ_U, the limiting values, are zero. So, the Cube copula allows for precise modeling of λ_{L-u}, but because its density is bounded, the lower tail dependence vanishes at arbitrarily small percentiles. In section 4 we propose a method for modeling not just a single λ_{L-u} but a countably infinite set of lower tail dependencies.

2.3.2 Measures of Association

Many authors have commented on the shortcomings of Pearson's product-moment correlation, ρ_P, for measuring associations in copula-based models. The most obvious is that ρ_P is not defined when marginals have infinite second moments, for example t_η with $\eta \le 2$, and many copula applications use such fat-tailed distributions.

A second shortcoming is that ρ_P is not a copula property. That is, ρ_P depends on both the copula and the marginal distributions. This contrasts with Spearman's rank correlation and Kendall's tau, ρ_S and ρ_τ respectively, both of which are copula properties. Here we focus on ρ_S rather than ρ_τ, for reasons described in Sec. 3 on copula mixtures.

For a copula with cdf F, we have $\rho_S = 12\int_0^1\int_0^1 F(x,y)dxdy - 3$. For the Cube copula then we have

$$
\rho_S(F) = 3(a-1)^4q_0 + 3(a-2)^2a^2q_2 - 12a(a-1)^3(a+1)q_1 - 3 \tag{15}
$$

Our explicit formula (15) for the Spearman correlation of a Cube copula will prove useful in Sec. 3, when we consider mixing the Cube copula with another copula in order to achieve a desired correlation in the bulk of the distribution.

3 Mixing the Cube Copula

Above we constructed a copula with unusually high tail-dependence; indeed the tail dependence arising from this copula is maximally high within the space of copulas that have the particularly simple structure we have laid out. However, the correlation with the bulk is zero by construction. How can we incorporate a non-zero bulk correlation structure together with tail dependence? We use a simple result that any mixture of two copulas is again a copula. We can then take a convex combination of a Cube copula with another copula that exhibits bulk correlation, and the resulting copula will exhibit both left tail dependence and bulk correlation.

3.1 Properties of Copula Mixtures

Definition 3. *Let V be a real vector space, and let $K \subseteq V$ be any convex subset. A function $f : K \to \mathbb{R}$ is said to be* convex-linear *if*

$$f(tx + (1-t)y) = tf(x) + (1-t)f(y) \quad \textit{for all} \quad t \in [0,1], \ \ x, y \in K.$$

Note that convex-linearity extends naturally to compositions with affine maps in either order. Specifically if $f : V \to \mathbb{R}$ is convex-linear and $\phi : V \to V$ is an affine map defined by $\phi(x) = Ax + b$, then $f \circ \phi$ is also convex-linear. Similarly if $\phi' : \mathbb{R} \to \mathbb{R}$ is affine, then $\phi' \circ f$ is again convex-linear. These statements remain true when the target space $\mathbb{R}$ is replaced by an arbitrary vector space, but we will only use the real-valued case.

3.1.1 Tail Dependence

Let $\mathfrak{C}_k$ denote the set of k-variate copulas; note that $\mathfrak{C}_k$ is a convex subset of the vector space of all functions from $\Omega_2 \to \mathbb{R}$, hence def. 3 applies. Consider λ_{L-u}, the lower-u tail dependence of the previous section, as a real-valued function defined on $\mathfrak{C}_2$.

Let C_1 and C_2 be bivariate copulas, and $t \in [0,1]$. Note that

$$\begin{aligned}\lambda_{L-u}(tC_1 + (1-t)C_2) &= \frac{1}{u}\left(tC_1 + (1-t)C_2\right)(u,u) \\ &= t\frac{C_1(u,u)}{u} + (1-t)\frac{C_2(u,u)}{u} \\ &= t\lambda_{L-u}(C_1) + (1-t)\lambda_{L-u}(C_2).\end{aligned}$$

So λ_{L-u} is convex-linear, and the lower-u tail dependence of a mixture copula is the mixture of the component copulas' lower-u tail dependencies. Also convex-linear is $\lambda_L = \lim_{u \searrow 0} \lambda_{L-u}$, provided that the limits of the components' exist.

3.1.2 Correlations

Let (X_1, Y_1) and (X_2, Y_2) be independent continuous random variables with common margins F (of X_1 and X_2) and G (of Y_1 and Y_2). Let C_i denote the copula of (X_i, Y_i). Define

$$Q = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0].$$

as in [9]. Then

$$Q = Q(C_1, C_2) = 4\int C_2 dC_1 - 1 = Q(C_2, C_1). \tag{16}$$

It is immediate from (16) that Q is convex-linear in each of its two arguments, if the other is held fixed. Let Π denote the independence copula (constant density).

Two popular non-parametric measures of concordance are *Kendall's tau* defined as $\rho_\tau(C) = Q(C, C)$, and *Spearman's rho* given by $\rho_S(C) = 3Q(C, \Pi)$. From the observation that Q is convex-linear in each argument, it follows that $\rho_S(C)$ is convex-linear in C, while

$$\rho_\tau(tA + (1-t)B)$$

is a polynomial of degree 2 in t, hence cannot be convex-linear. Similarly, Gini's coefficient

$$\gamma = 2\int (|u + v - 1| - |u - v|) dC(u, v)$$

and Blomqvist's "medial correlation" $\beta = 4C(\frac{1}{2}, \frac{1}{2}) - 1$ are also convex-linear functions of C.

Finally, we remark that the n-dimensional Spearman's ρ_S, given by

$$\rho_{S,n} = \frac{n+1}{2^n - (n+1)}\Big[2^{n-1}\Big(\int C\, d\Pi^n + \int \Pi^n dC\Big) - 1\Big]$$

is convex-linear on the convex cone of n-copulas.

These considerations imply that when dealing with mixture copulas, all of the usual measures of concordance *except Kendall's* τ are convex-linear and can be summed across the components of the mixture.

3.2 The Cube-Gaussian Mixture Copula

Let ρ be an $n \times n$ correlation matrix. The *Gaussian copula* with correlation ρ is defined by its PDF:

$$p_g(\mathbf{u}) = |\rho|^{-1/2} \exp[-\frac{1}{2}\zeta'(\rho^{-1} - I)\zeta] \quad \text{where} \quad \zeta = \Phi^{-1}(\mathbf{u}).$$

Here Φ is the normal CDF applied componentwise to vectors, and $\mathbf{u} \in [0,1]^n$. Building on this we define the *Cube-Gaussian mixture copula* by

$$p_{gt}(\mathbf{u}) = \lambda p_c(\mathbf{u}) + (1-\lambda) p_g(\mathbf{u}). \tag{17}$$

Here $\lambda \in [0,1]$ is the mixture probability. One can view this as a hierarchical model, where a mixing random variable defined on $\{0,1\}$ determines which copula that X will be drawn from.

Eq. (17) is our first example of a copula which has desirable properties for modeling portfolios of risky assets. In the bulk of the distribution, i.e. the region in which assets in the portfolio are not experiencing VaR events individually, the assets behave as though they have correlation matrix ρ, but if one or more assets is experiencing a VaR event, the conditional probability that others are also seeing their VaR events is much higher than it could be with a Gaussian copula or a t-copula.

One advantage of any copula of the form (17) is that, due to the result in sec. 3.1.2 that many of the standard measures of association are convex-linear on the space of copulas, we see that these measures will be no more difficult to compute for the Cube-Gaussian copula (17) than for either of its components, and we have already shown how to compute spearman's correlation for the cube in sec. 2.3.2. The more advanced copulas we will introduce in Sec. 4 are also mixtures, and also benefit from the results in sec. 3.1.2.

Eq. (17) also lends itself well to Monte Carlo simulation, since each of the components p_c and p_g is easy to simulate. Given a simulation, one can proceed to a full portfolio-level analysis of Value-at-Risk (VaR). The simulation suffices to compute each asset's marginal contribution to portfolio VaR as a numerical derivative. Fig. 4 illustrates the behavior of the Cube-Gaussian mixture via a simulation histogram with $\rho = 0.3$ bulk correlation, $a = 0.015$ and maximal tail dependence for these parameters.

From Fig. 4 it's intuitively clear that this copula satisfies the two desirable properties laid out in Sec. 1. In particular note that there is a large probability density

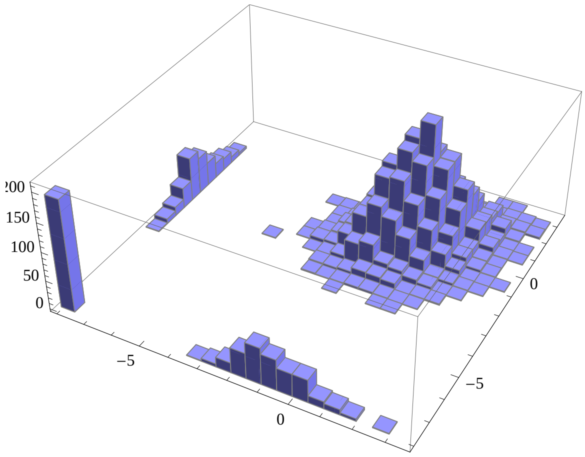

Figure 4: Histogram of a simulation from the Cube-Gaussian mixture with $\rho = 0.3$ bulk correlation, $a = (0.015, 0.015)$ and maximal tail dependence.

in the simultaneous *lower left* tail, but no corresponding density in the simultaneous *upper right* tail; this is typical of portfolios of financial assets, and completely impossible to achieve with the t-Copula.

Although surely more realistic for portfolio risk modeling than either the normal copula alone, or the t-Copula, even Eq. (17) has an important shortcoming for the intended application. Fortunately the correction for this shortcoming is known, and leads to interesting further mathematics. This will be the subject of the next section; for the moment, we simply expose the indicated shortcoming.

Note that the probabilities of simultaneous tails beyond the a-th percentile die off quickly. Suppose for illustration a Cube with $a = 0.05$, $n = 2$, and $q_2 = 16$. This is equivalent to assuming that the chance of a double tail is twice the chance of a single tail. Then conditional on the first asset having a 95%-VaR event, the probability that both simultaneously have 95%-VaR events (that is, $\lambda_{L-0.05}$) is 0.8. The corresponding probability for a Gaussian copula with $\rho = 0.9$ is around 0.64. However what if we consider 99%-VaR? Under the Cube, $\lambda_{L-0.01}$ falls to 0.43, while for the Gaussian is 0.54. Given the Cube's zero asymptotic tail dependence we know this conditional probability converges to zero, but it does so sufficiently quickly to cause concern in some applications, such as in cases where one wants to forecast both 95% and 99% VaRs. The next section introduces a modeling technique, which we call copula nesting, that allows for modeling arbitrarily many points along the tail dependence surface with a sequence of nested Cube copulas.

4 The Copula Nesting Theorem

In Sec. 3, we defined the Cube copula and noted that via selecting $\mathbf{q}$ and a appropriately one can precisely specify λ_{L-a}. However, the tail dependence beyond a degrades, and asymptotes to zero. Suppose that one wants to specify tail dependence at a set of quantile points A. Figure 5 illustrates the method we propose for doing so. The key observation is that one can nest a second Cube copula within the lower left region of an initial Cube copula. We prove below that the resulting function remains a copula. One can repeat this nesting arbitrarily many times, and in doing so, precisely model λ_{L-a_n} for $a_n \in A$.

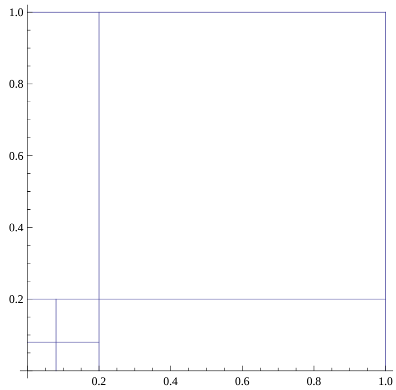

Figure 5: Cell complex underlying a doubly-nested Cube copula in two dimensions. The figure illustrates a cube with $a = 0.4$ nested inside a cube with $a = 0.2$.

4.1 The Nesting Theorem and Proof

Let f be the Cube copula density on $\Omega = [0,1]^n$ with tail parameter a and values q_k on the k-tail for each $k = 0,\ldots,n$. Let s be any copula on Ω, which we extend to $\mathbb{R}^n$ by specifying that $s = 0$ outside Ω. Let $\phi : \Omega \to \tau_{n,n}$ be the bijective affine map between the indicated regions given by rescaling each coordinate.

Consider the density

$$\hat{s}(\mathbf{u}) = a^{-n}\, s(\phi^{-1}(\mathbf{u})) \tag{18}$$

Then by our convention $\hat{s}$ vanishes outside of $\tau_{n,n}$. Since we have multiplied by the inverse of the Jacobian, the overall normalization is preserved: $\int \hat{s}(\mathbf{u})\, d\mathbf{u} = 1$.

In Theorem 3 we construct a copula $\hat{f}$ which, intuitively, consists of modifying f by replacing its values in $\tau_{n,n}$ with a scaled version of s.

Theorem 3. *The multivariate probability density defined by*

$$\hat{f}(\mathbf{x}) = \begin{cases} q_n a^n \hat{s}(\mathbf{x}), & \text{if } \mathbf{x} \in \tau_{n,n} \\ f(\mathbf{x}) & \text{otherwise} \end{cases}$$

with $\hat{s}$ defined as in Eq. (18)*, is a copula density.*

Proof. The scaling is such that the integral over $\tau_{n,n}$ is unchanged. It follows that $\hat{f}$ is a probability density. We also claim that $\hat{f}$ has uniform marginals. We need to show that the marginal in the x_j direction is uniform for each $j = 1, \dots, n$. For notational simplicity we show this for $j = n$; the same proof holds in each direction. Then we may write $\mathbf{x} = (\mathbf{y}, x_n)$ where $\mathbf{y} \in [0,1]^{n-1}$ and $x_n \in [0,1]$. The marginal function is then

$$m(x_n) = \int_{[0,1]^{n-1}} \hat{f}(\mathbf{y}, x_n)\, d\mathbf{y}.$$

Note that if $x_n > a$, then $\hat{f}(\mathbf{x}) = f(\mathbf{x})$ and hence $m(x_n) = 1$ since $m(x_n)$ is a marginal of the copula f. Therefore suppose $x_n \leq a$ and split the integral as follows:

$$\begin{aligned} \int_{[0,1]^{n-1}} \hat{f}(\mathbf{y}, x_n)\, d\mathbf{y} &= \int_{\tau_{n-1,n-1}} \hat{f}(\mathbf{y}, x_n)\, d\mathbf{y} + \int_{\tau_{n-1,n-1}{}^c} \hat{f}(\mathbf{y}, x_n)\, d\mathbf{y} \qquad (19) \\ &= q_n a^n \int_{\tau_{n-1,n-1}} \hat{s}(\mathbf{y}, x_n)\, d\mathbf{y} + \int_{\tau_{n-1,n-1}{}^c} f(\mathbf{y}, x_n)\, d\mathbf{y} \qquad (20) \end{aligned}$$

Also in the region $x_n \leq a$ one has

$$\begin{aligned} q_n a^n \int_{\tau_{n-1,n-1}} \hat{s}(\mathbf{y}, x_n)\, d\mathbf{y} &= q_n a^n \int_{\tau_{n-1,n-1}} \hat{s}(a\mathbf{z}, x_n)\, d(a\mathbf{z}) = q_n a^{n-1} \int_{[0,1]^{n-1}} s(\mathbf{z}, x_n/a)\, d\mathbf{z} \\ &= q_n a^{n-1} = \int_{\tau_{n-1,n-1}} f(\mathbf{y}, x_n)\, d\mathbf{y} \end{aligned}$$

since $\int_{[0,1]^{n-1}} s(\mathbf{z}, x_n/a)\, d\mathbf{z}$ is a marginal of s.

Plugging this back into the expression (20), we see that the sum (20) collapses into an expression for the marginal of f at x_n, which we know to be 1. This completes the proof. □

Remark 1. The same argument also shows that an arbitrary copula can be nested within the 0-tail region $\tau_{0,n}$. The Cube copula is essentially the only copula that admits a nesting theorem of this form.

If the nested copula s is itself a Cube copula, then further copulas can be nested within the inner copula s. One can in fact do this infinitely-many times, leading to a fractal structure, though for applications in finance or engineering one would typically stop when the tails being modeled are so low-probability that one has no further view on tail dependence or need to model it in those regions.

4.2 Improving The Cube-Gaussian Mixture

The multiple-nesting property allows the practitioner to customize the copula's tail-dependence properties by specifying as ingredients not only the probability of an a-quantile VaR event, but also the probabilities of the $a/10$-quantile, $a/100$-quantile, etc.

In this way we can resolve the fundamental difficulty which plagued the simple form of the Cube-Gaussian mixture discussed in Sec. 3.2. With nested copulas, it need not be the case that the conditional probability of an n-tail a-quantile event, conditioned on the occurrence of an $(n-1)$-tail a-quantile event, goes to zero as $a \to 0$. By suitably choosing the $\mathbf{q}$-vectors for the inner nested copulas, one can ensure that these probabilities remain bounded away from zero and so that the full copula has a non-zero tail dependence coefficient.

Suppose that in n-dimensions we have the Cube copula $p_c(\mathbf{x})$ with parameters $a, \mathbf{q}$ and we define an *inner Cube copula* $\hat{p}(\mathbf{x})$ which has the same structure, but different parameters $\hat{a}, \hat{\mathbf{q}}$ and a normalizing constant set according to eq. (21) below. Note that there are no constraints on $\hat{a}, \hat{\mathbf{q}}$ aside from the general constraints set by Theorem 1 which apply to all Cube copulas.

As before, we set $\tau_{k,n}$ to be the k-tail region of the outer copula. We will use $\hat{\tau}_{k,n}$ to denote the corresponding regions for the inner copula. Then the nested copula is:

$$p(\mathbf{x}) = \begin{cases} \hat{p}(\mathbf{x}), & \mathbf{x} \in \tau_{n,n} \\ p_c(\mathbf{x}) & \text{otherwise} \end{cases}.$$

The normalizing constant for $\hat{p}$ is set so that

$$\int_{\tau_{n,n}} \hat{p}(\mathbf{x}) d^n\mathbf{x} = q_n a^n. \tag{21}$$

Suppose, as is common in financial risk modeling, we are interested in 95% and 99% VaR, and we wish to build a model with higher-than-normal probabilities of joint tail events occurring at these quantiles. The doubly-nested Cube achieves this: parameters $a = 0.05$, $\hat{a} = 0.2$ imply

$$\begin{aligned} \text{Prob(joint 95\% quantile)} &= (0.05)^2 q_n \\ \text{Prob(joint 99\% quantile)} &= (0.01)^2 \hat{q}_n \,. \end{aligned}$$

This illustrates that the nested copula allows us to tailor the probabilities of these events representing joint observations of extreme outliers. After mixing with the normal, of course the necessary integrals become more difficult to do, but even these can easily be handled numerically.

5 Literature Review

Sklar introduced the mathematical structure of copulas into the probability and statistics literature in 1959, coining the phrase “copula” with Schweizer in their 1983 textbook [33]. The topic received much attention in decades following its introduction, summarized nicely by Schweizer’s “Thirty Years of Copulas” [32]. This research spawned several introductory papers and textbooks meant to introduce the advanced undergraduate or graduate student to the topic, see for example the appropriately titled “Joy of Copulas” [16], and the excellent texts by Joe [21] and Nelsen [27]. The concept found applications within the fields of engineering and biology, but only recently have researchers applied copulas to economic data. The earliest instances came in the insurance and operations research literature insurance during the mid 1990s. Frees et all [12] in the Journal of Risk and Insurance consider the problem of pricing an annuity promised on two lives, and apply Frank’s copula [11], a special case of the Archimedean family of copulas. Jouini and Clemen [22] investigate aggregating expert opinions, also with Frank’s copula. The first mention of copulas within *Management Science* arrives in [37], who study accident “precursors” or “near-misses”, where the joint distribution modeled is that of the failure probability of some safety system under two states of the world depending on whether some other safety system has or has not failed. The first appearance of copulas in an economics journal is also via an investigation into an insurance problem, in the context of a principal agent problem with adverse selection [23]. Overall, the use of copulas in the economics literature has been sparse and very recent. The journals of the American Economic Association record four articles mentioning copulas, all between 2007 and present; *Econometrica* records three mentions, all published in 2010; the *Journal of Political Economy* records one mention [19]; while the *Quarterly Journal of Economics* records none.

The use of copulas in the financial literature was also recent, but has grown explosively in the last ten years. In their widely circulated 1999 working paper, Embrechts, McNeil, and Strauman [10] introduce copulas into modeling financial asset returns. They focus on correcting what they perceive as commonly held views on correlations that “arise from the naive assumption that dependence properties of the elliptical world also hold in the non-elliptical world” and they propose copulas and rank correlations as a remedy. With this background, it is perhaps not surprising that the highly non-elliptical world of credit derivatives emerged as fertile ground for copulas. Li [24] was the earliest published instance, although he cites technical documents from the industry that predate his research (although not explicitly using the copula terminology).[2] Soon after Li’s article, examples of copulas in credit modeling rapidly proliferated; key references are [13] and [31], both of which unify Li’s approach with the latent variable approach of older industry research (KMV and CreditMetrics). Bouye et al [6] provide a reading guide that both introduces the mathematics of copulas and illustrates with applications to

[2] Li’s use of the Gaussian copula was pilloried in an article in Wired magazine dubbed “Recipe for Disaster: The Formula That Killed Wall Street”, [30].

credit scoring, asset returns, and risk measurement. Longin and Solnik [26] provide the first published example of copulas used in modeling returns from different equity markets, as well as the first mention of copulas in a top finance journal. They use Gumbel's copula [18], although interestingly they neither cite the seminal statistical references nor use the phrase "copula" in their paper. Other highly regarded finance journals follow suit: [7] in the *Journal of Financial and Quantitative Analysis*; [5] in the *Journal of Business*; [28] in the *Review of Financial Studies*. Interestingly, the earliest mention of copulas in the *Journal of Financial Economics* is in a footnote to [4], which states that "Embrechts et al. (1999) have recently advocated the use of copulas and rank statistics when measuring dependence in non-normally distributed financial data. However, because the unconditional distributions that we explore . . . are all approximately Gaussian, the linear correlation affords the most natural measure in the present context." Unfortunately, the evidence presented that the financial data they study are normal concerns only the marginal distributions, and not the joint distribution, which is what is relevant for determining whether to use copulas. Finally, several textbooks provide very thorough introductions to copulas in finance, namely [8] and [36].

One interesting application comes from Rosenberg and Schuermann [29]. They attempt to model the various risks that a complicated financial institution faces (market, credit, and operational) via flexible modeling of marginal distributions each of which is allowed to have a very different shape. The authors conclude that the VaRs of the individual components and the weights that aggregate these components into a portfolio play a larger role in determining portfolio VaR than the choice of the copula. However, the authors consider only Gaussian and Student's-t copulas, which we suspect drives this conclusion.

In general, many of the existing financial applications in the literature seem to view the primary benefit of copulas as simply allowing for arbitrary marginals, without much attention given to the implications of the copula for modeling tail dependence. The emphasis, then, becomes sophisticated modeling of the marginals, with the copula chosen as an afterthought. Rosenberg and Schuermann clearly fits in this category, as do most of the early credit modeling references provided above. A notable exception, and the approach most similar to ours, is Hu's [20], which estimates mixtures of Gaussian, Gumbel, and Gumbel survival copulas using monthly returns from the S&P 500, FTSE 100, Nikkei 225, and Hang Seng. Like our application below, Hu estimates marginals non-parameterically, focusing on the dependence structure rather than the marginals.

6 An Application to Hedge Fund Returns

6.1 Data

Hedge fund returns provide a natural setting to apply our copula mixture. We think of multivariate hedge fund return distributions as operating under two regimes. In normal times, hedge funds strategies operate with whatever correlations arise naturally from their common exposures to risk factors and correlated trading strategies. However, during stress scenarios, strategies correlate to a much higher degree, as industry-wide balance sheet reductions beget negative returns via market impact, which beget further balance sheet reductions. The causes of these correlated portfolio liquidations can be acute, such as when sudden losses by a large fund become common knowledge (eg Long Term Capital Management in 1998), or more diffuse, such as a contraction in banks' willingness to finance transactions or elevated redemption requests by investors (which are two commonly-cited causes of large hedge fund losses in the fall of 2008). Regardless of cause, the presence of simultaneous deleveragings creates a left tail dependency that can be much more extreme that what one would expect from observing returns during normal times.

To illustrate, we use the Hedge Fund Research indexes (HFRI), which HFR describes as "a series of benchmarks designed to reflect hedge fund industry performance by constructing equally weighted composites of constituent funds, as reported by the hedge fund managers listed within HFR Database." While HFRI returns suffer some serious biases in their construction[3], they are generally considered the industry standard and have been used in many of the seminal studies of hedge fund returns (Ackerman, McEnally, and Ravenscraft [1]; Liang [25]; Agarwal and Naik [2], [3]; Getmansky, Lo, Makarov [17]; Fung and Hsieh [14]). Specifically, we investigate the joint distribution of the Event Driven (ED) and Relative Value (RV) strategy indexes.[4]

Our data consist of the monthly returns for Event Driven and Relative Value from February 1990 through August 2010, measured in excess of the US 3-month Treasury bill rate. The cumulative returns are shown in Fig. 6, and the scatterplots in Fig. 7. The excess returns are highly correlated, with a ρ_S of 0.67 over the full sample, despite strategy descriptions that would not suggest such high correlations.

Much of this correlation is due to persistent exposure to common risk factors. We attempted to control for these exposures by OLS regression of each strategy against the excess total returns of several market indexes:

[3] Most seriously, returns are self-reported and funds are free to self-censor.

[4] Event Driven includes as sub-categories: Activist, Credit Arbitrage, Distressed/Restructuring, Merger Arbitrage, Private Issue/Regulation D, Special Situations, and Multistrategy. Relative Value includes as sub-categories: Fixed Income-Asset Backed, Fixed Income-Convertible Arbitrage, Fixed Income-Corporate, Fixed Income-Sovereign, Volatility, Yield Alternatives:Energy Infrastructure, Alternatives:Real Estate, and Multistrategy.

(a) S&P 500,

(b) Barclays Capital Aggregate Total Treasury,

(c) Barclays Capital US Corporate High Yield, and

(d) S&P Goldman Sachs Commodity.

For each HFR index, all four market indexes (a)–(d) had t-statistics above 2 (with average absolute t-statistics of 6.0 for ED and 4.8 for ED), and adjusted R^2 were 65% for ED and 52% for RV. The Spearman correlation of the residuals to these factors drops to 0.49, but clearly, left-tail correlation remains present even in the residualized data.

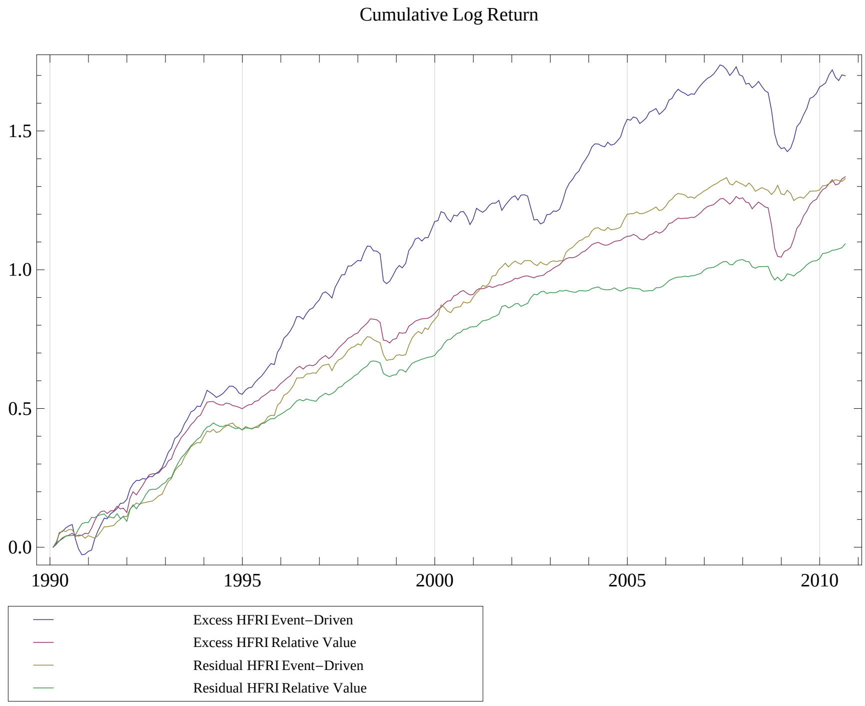

Figure 6: Cumulative Excess Log Returns and Residual Log Returns for HFRI Event Driven and Relative Value Indexes, Monthly from Jan 1990–Aug 2010.

6.2 Methodology

We estimate the parameters of our Cube-Gaussian Mixture copula, described in Section 3.2, via a two-step, pseudo-maximum likelihood estimation procedure. First, we estimate marginals via the empirical CDF, and apply an inverse empirical CDF

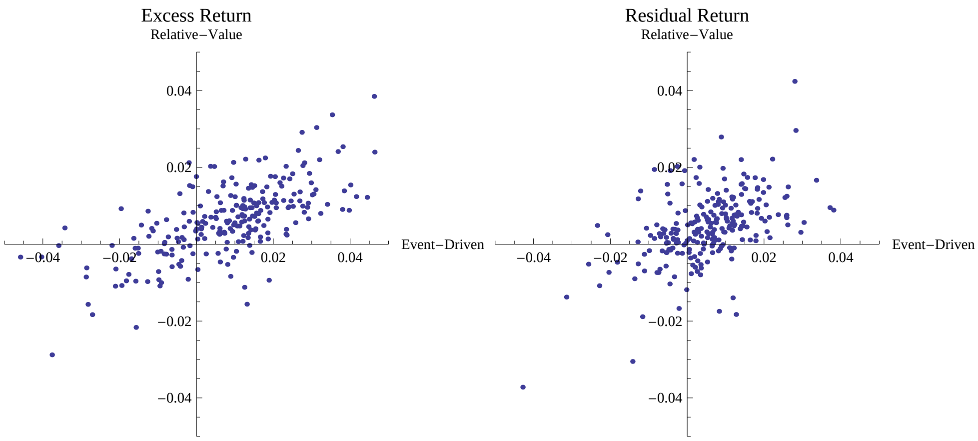

Figure 7: Cumulative Monthly Excess and Residual Log Returns for HFRI Event Driven (horizontal) and Relative Value (vertical) Indexes, Jan 1990-Aug 2010.

to each variable to transform it into a uniform. In the second step, we estimate the parameters of the copula on these transformed data via maximum likelihood. Alternatively, we could estimate a full information maximum likelihood by specifying marginals, and then maximizing a joint likelihood function that is both a function of the parameters of the marginals and the parameters of the copula. The benefit of using the two-step procedure and non-parametrically estimating marginals is that if parametric marginals are mis-specified and included in a joint likelihood, they will interfere with the copula estimates, which are our focus. Note that the standard errors that arise from this two-step procedure are larger than those that would be naively computed by assuming that the transformed data were the actual data. Intuitively, the inverse empirical CDF does not equal the true inverse CDF, and this source of estimation error much be accounted for in the copula's parameter estimates. The corrected standard errors come via a "sandwich estimator" of the asymptotic covariance matrix; [15] provide a derivation.

We chose to use a single breakpoint a, fixed in advance, which defines double-tail region in the Cube, and we estimated the following three parameters: q_2, which determines the likelihood of the double-tail,[5] λ, the mixing parameter, and ρ, the correlation of the Gaussian copula. Provided that a is chosen not to coincide with one of the values of the sample, the likelihood function is differentiable in each parameter, and the maximization was easily solved via Matlab's fmincon function.[6]

[5]Note from Section 2.2 that q_1 and q_0 are simple functions of q_2.

[6]In preliminary research, we also experimented with estimating a, but this proved difficult. The typical solution to this problem involves an a that coincides exactly with a sample value, which creates a non-differentiability at the maximum likelihood estimate. In fact, not only is the likelihood function not differentiable in a at such a point, it is not differentiable in q_2 either. This resulted in instability in our numerical routines.

We estimated with a equal to 0.01 and 0.05, representing 95% and 99% VaR, two commonly used breakpoints in financial risk modeling. Neither 0.01 nor 0.05 was exactly equal to an observation of our sample after it had been mapped to uniform via the inverse empirical CDF, since our sample contained 247 observations.[7]

We also evaluated the fit of the Mixture copula against three competing copulas: pure Gaussian, Student's-t, and Clayton. We estimated the parameters of each of these three copulas via maximum likelihood: ρ for the Gaussian, ν and ρ for the Student's-t, and θ for the Clayton. We evaluate each copula's fit via the value of the log-likelihood and two information criteria that penalize for over-fitting: Akaike's Information Criterion (AIC) and the Bayesian Information Criterion (BIC).

Note that our sample consists of 247 monthly observations. With a sample this size, estimating the 95% and 99% VaRs is imprecise. To see this, note that one can compute a confidence interval for a quantile using the sample order statistics via the binomial distribution. Letting $X_{(j)}$ denote the jth largest of a collection of n iid continuously distributed random variables, and q_α the α-quantile, we have $P(X_{(j)} < q_\alpha < X_{(j+1)}) = \binom{n}{j}\alpha^j(1-\alpha)^{j-1}$. Starting from the sample α-quantile, one can expand to successively larger intervals with order statistics as endpoints to obtain confidence intervals with increasing rates of coverage. See Stark [35] for details. In particular, with a sample of 247 observations, a 95% confidence interval for $q_{0.05}$ centered at the sample 0.05-quantile is contained within $[x_{(6)}, x_{(19)}]$, that is, the sample 2.4%-ile to the sample 7.7%-ile. A 95% confidence intervals for $q_{0.01}$ is contained within $[-\infty, x_{(5)}]$, that is, from $-\infty$ to the sample 2%-ile. If one prefers confidence intervals of finite length, a 90% confidence intervals for $q_{0.01}$ is contained within $[x_{(1)}, x_{(6)}]$; however there is no interval between the minimum sample observation and maximum sample observation that contains a 95% confidence interval for $q_{0.01}$.

However, we don't view this as an impediment to estimating the Cube copula on a dataset of this size. The Cube copula is concerned with the dependence between extreme observations, but not with the specific shape of the tails of a particular marginal distribution. To see this another way, note that our maximum likelihood procedure transforms the data into ranks before the Cube even sees them, so that taking the most extreme observation and scaling it by a factor of 10 along one dimension would have no impact on the Cube's estimates. What is a problem however, is that with a small number of observations, one may not observe extreme dependence. Thus estimating the Cube requires confidence that one has indeed observed extreme dependence within a sample. The uneasiness one should feel making statements about having already observed extreme dependence suggests that in any financial application (and perhaps any application more generally), stress-testing the Cube's parameters (for example for their impact on forecast VaR or contribution to VaR) is critical.

[7] Had we instead had 100 observations, so that 0.01 and 0.05 would have been elements of our U(0,1) transformed sample, we could have chosen for example an a of 0.011 and 0.051 to avoid numerical difficulties. However, choosing 0.011 rather than 0.009 could have a sizable impact on the estimates of the other parameters.

6.3 Results

First we report the results of the OLS regression of the HFR indexes on the market indexes:

	HFRI Event Driven	HFRI Relative Value
Intercept	0.005377	0.004426
	(0.000762)	(0.000592)
S&P 500	0.1857	0.0426
	(0.0209)	(0.0162)
BarCap Agg Total Treasury	-0.1728	-0.0920
	(0.0561)	(0.0436)
BarCap US Corp High Yield	0.3172	0.2549
	(0.0332)	(0.0258)
S&P GS Commodity	0.0301	0.04336
	(0.0121)	(0.0094)
n	247	247
R^2	0.655	0.523
Adj-R^2	0.650	0.515

Standard errors in parentheses

Next we report the estimated parameters of the Cube-Gaussian Mixture, both for excess and residualized HFRI returns:

	$a = 0.05$		$a = 0.01$	
	HFRI Excess Returns	HFRI Residual Returns	HFRI Excess Returns	HFRI Residual Returns
q_2	15.45	5.87	100.0	100.0
	(37.05)	(5.01)	(198.1)	(149.5)
λ	0.047	0.393	0.130	0.394
	(0.097)	(0.103)	(0.090)	(0.101)
ρ_S	0.726	0.751	0.768	0.758
	(0.046)	(0.039)	(0.037)	(0.039)
-log L	-82.22	-37.09	-83.07	-38.52

Approximate standard errors in parentheses

Note that with $a = 0.01$, the corner solution of $q_2 = a^{-1}$ maximizes the likelihood, and at that estimate the derivative of the likelihood function is undefined, so the approximate standard error is not valid. Given that only 2 observations per variable lie below the 0.01 percentile, we should expect the estimator to be somewhat unstable. The fact that the most extreme observation for HFRI ED residual returns coincides precisely with the most extreme observation for HFRI RV's residual returns [8] means that fitting this highly unlikely (from the perspective of the Gaussian) point perfectly brings a large likelihood gain. This instability would prevent most practitioners from taking the $a = 0.01$ estimates very seriously, and so we have included them solely to illustrate the mechanics of the model.

Finally, we report the values of the likelihood, AIC, and BIC evaluated at the maximum likelihood estimates for the Cube-Gaussian (at $a = 0.05$ and $a = 0.01$), Gaussian, Clayton, and Student's-t copulas:

[8] These occurred during the LTCM crisis of August 1998. The most extreme HFRI ED excess return, also August 1998, occurred with HFRI RV's 2nd most extreme excess return, and vice versa during October 2008.

	HFRI Excess Returns	HFRI Residual Returns
-log L		
Cube-Gaussian ($a = 0.05$)	-82.22	-37.09
Cube-Gaussian ($a = 0.01$)	-83.07	-38.52
Gaussian	-82.09	-33.43
Clayton	-78.61	-33.42
Student's-t	-83.92	-36.11
AIC		
Cube-Gaussian ($a = 0.05$)	-158.45	-68.19
Cube-Gaussian ($a = 0.01$)	-160.13	-71.04
Gaussian	-162.18	-64.87
Clayton	-155.21	-64.83
Student's-t	-163.84	-68.22
BIC		
Cube-Gaussian ($a = 0.05$)	-147.92	-57.66
Cube-Gaussian ($a = 0.01$)	-149.60	-60.51
Gaussian	-158.67	-61.36
Clayton	-151.70	-61.33
Student's-t	-156.82	-61.20

Cube-Gaussian estimates 3 parameters; Gaussian, 1; Clayton, 1; and Student's-t, 2

On the HFRI Excess returns, Student's-t performs best in terms of the log-likelihood and AIC, but the heavier penalty that BIC enacts on extra estimated parameters causes the pure Gaussian to fit the best of the five models. The Cube-Gaussians both outperform the Clayton on AIC, but are last on BIC due to their extra parameter.

On the HFRI Residual returns, the Cube-Gaussians perform better. BIC still penalizes them sufficiently heavily that they are last, but both are considerably better than the three competing models under AIC's penalty. In terms of intuition for the very different ranking achieved by AIC and BIC, note that AIC's penalty term is a function solely of the number of parameters, namely $2k$ where k is the number of estimated parameters, while BIC's penalty term grows with the sample size, via $klog(n)$ where n is the sample size.

Our view of these results is that they at minimum establish that parameter estimation is straightforward for the simplest version of the mixed Cube copula, and that it produces results about as good as common alternative approaches to modeling tail dependence. However, we also view the Cube-Gaussian as a modeling tool, rather than just a statistical tool, in the sense that its parameters will generally require both intuition and experience in the arena in which the tool is applied. To make this point obvious, the practitioner who put no thought into the liquidation scenarios that could lead to simultaneous left tails across hedge fund strategies, but instead blindly applied our method, would not have been much better off going into August 1998 than a practitioner who instead used a Gaussian copula.

7 Conclusions

Motivated by applications to portfolio risk modeling, we searched for a copula which is flexible enough to accommodate a fully-general correlation matrix in the bulk, as well as a very high conditional probability of left-tail events, with no corresponding implication for right-tail events. Our conclusion was that none of the well-known copulas in the literature are quite so flexible. All of these seem to have the property that one can introduce high left tail dependence, but only at the cost of influencing the copula in other ways which make it inadmissible for this sort of risk modeling.

To address this problem, we created a new family of copulas which is more "flexible" in several important ways. It allows one in particular to separately specify the probabilities for a-quantile events, for a sequence of increasing values of a, while retaining a fixed correlation structure in the bulk and without requiring the introduction of artificial right-tail dependence.

The new copula we propose takes the form of a nested Cube mixed with a normal copula. The normal copula takes into account the correlation matrix, while the nesting and the cube structure creates a sequence of increasing probabilities of simultaneity in the left tails. The parameter vector $\mathbf{q}$ can be tuned to make these conditional probabilities as large as mathematically possible. Since the q_i's are subject only to simple linear constraints, it is not difficult to tune the $\mathbf{q}$-vector while maintaining consistency.

Theorem 1 and Theorem 3 ensure that the resulting mixtures are indeed copulas. This structure has several desirable properties, including the fact that certain important and widely-used measures of concordance, such as Spearman's rho, distribute over mixtures and can be computed explicitly for the components.

Applications in financial engineering include a better estimation of VaR and contributions to VaR for portfolios of assets which have moderate correlation in normal times, but which tend to experience highly correlated drawdowns in crisis periods.

References

[1] 1999. Ackermann, C.; McEnally, R.; Ravenscraft, D.: *The Performance of Hedge Funds: Rsik, Return, and Incentives*. Journal of Finance, Volume 54, Issue 3, Jun, pp. 833-874.

[2] 2000. Agarwal, C.; Naik, N.: *Multi-Period Performance Persistence Analysis of Hedge Funds*. Journal of Financial and Quantitative Analysis, Volume 35, Issue 3, pp. 327-342.

[3] 2000. Agarwal, C.; Naik, N.: *Generalized Style Analysis of Hedge Funds*. Journal of Asset Management, Volume 1, Issue 1, Jul, pp. 93-109.

[4] 2001. Andersen, T.; Bollerslev, T.; Diebold, F.; Ebens, H.: *The Distribution of Realized Stock Return Volatility*.

[5] 2003. Ane, T.; Kharoubi, C.: *Dependence Structure and Risk Measure*. Journal of Business, Volume 76, pp 411-438. Journal of Financial Economics, Volume 61, pp 43-76.

[6] 2000. Bouye, E.; Durrelman, V.; Nikeghbali, A.; Riboulet, G.; Roncalli, T.: *Copulas for Finance: A Reading Guide and Some Applications*. Unpublished working paper.

[7] 2003. Chen, R.; Sopranzetti, B.: *The Valuation of Default-Triggered Credit Derivatives*. Journal of Financial and Quantitative Analysis, Volume 38, pp 359-382.

[8] 2004. Cherubini, U.; Luciano, E.; Vecchiato, W.: *Copula Methods in Finance*, Wiley.

[9] 2003. Embrechts, P.; Lindskog, F.; McNeil, A.: *Modelling Dependence with Copulas and Applications to Risk Management*. In: Handbook of Heavy Tailed Distributions in Finance, ed. S. Rachev, Elsevier, Chapter 8, pp. 329-384.

[10] 1999. Embrechts, P.; McNeil, A.; Straumann, D.: *Correlation and Dependence in Risk Management: Properties and Pitfalls*. Unpublished working paper.

[11] 1979. Frank, M.: *On the simultaneous associativity of F(x, y) and x + y - F(x, y)*. Aequationes Mathematicae, Volume 19, 194-226.

[12] 1996. Frees, E.; Carriere, J.; Valdez, E.: *Annuity Valuation with Dependent Mortality*. Journal of Risk and Insurance, Volume 63, pp. 229-261.

[13] 2001. Frey, R.; McNeil, A.; Nyfeler, M.: *Copulas and Credit Models*. RISK, October, pp 111-114.

[14] 1999. Fung, W.; Hsieh, D.: *A Primer on Hedge Funds*. Journal of Empirical Finance, Volume 6, pp. 309-331.

[15] 1995. Genest, C.; Ghoudi, K.; Rivest, L.-P.: *A Semiparametric Estimation Procedure of Dependence Parameters in Multivariate Families of Distributions*. Biometrika, Volume 82, pp 543-552.

[16] 1986. Genest, L.; MacKay, J.: *The Joy of Copulas: Bivariate Distributions with Uniform Marginals*. American Statistician, Volume 40, pp. 280-283.

[17] 2004. Getmansky, L.; Lo, A.; Makarov, I.: *An Econometric Analysis of Serial Correlation and Illiquidity in Hedge-Fund Returns*. Journal of Financial Economics, Volume 74, pp. 529-609.

[18] 1961. Gumbel, E.: *Bivariate Logistic Distributions*. Journal of the American Statistical Association, Volume 56, pp 335-349.

[19] 2003. Haile, P.; Tamer, E.: *Inference with an Incomplete Model of English Auctions*. Journal of Political Economy, Volume 111, pp. 1-51

[20] 2006. Hu, L.: *Dependence Patterns Across Financial Markets: A Mixed Copula Approach*. Applied Financial Economics, Volume 16, pp 717-729.

[21] 1997. Joe, H.: *Multivariate Models and Dependence Concepts*, CRC Press.

[22] 1996. Jouini, N.; Clemen, R.: *Copula Models for Aggregating Expert Opinions*. Operations Research, Volume 44, pp. 444-457.

[23] 1999. Landsberger, M.; Meilijson, I.: *A General Model of Insurance under Adverse Selection*. Economic Theory, Volume 14, pp. 331-352.

[24] 2000. Li, D.: *On Default Correlation: A Copula Function Approach*. Journal of Fixed Income, Volume 9, pp 43-54.

[25] 1999. Liang, B.: *On the Performance of Hedge Funds*. Financial Analysts Journal, Volume 55, Issue 4, Jul-Aug, pp. 72-85.

[26] 2001. Longin, F.; Solnik, B.: *Extreme Correlation of International Equity Markets*. Journal of Finance, Volume 56, pp 649-676.

[27] 1998. Nelsen, R.: *An Introduction to Copulas*, Springer.

[28] 2004. Poon, S.; Rockinger, M.; Tawn, J.: *Extreme Value Dependence in Financial Markets: Diagnostics, Models, and Financial Implications*. Review of Financial Studies, Volume 17, pp 581-610.

[29] 2006. Rosenberg, J.; Schuermann, T.: *A General Approach to Integrated Risk Management with Skewed, Fat-Tailed Risks*. Journal of Financial Economics, Volume 79, pp 569-614.

[30] 2009. Salmon, F.: *Recipe for Disaster: The Formula that Killed Wall Street*. Wired Magazine, February 23.

[31] 2001. Schonbucher, P.; Schubert, D.: *Copula-Dependent Default Risk in Intensity Models*. Working paper, Department of Statistics, Bonn University.

[32] 1991. Schweizer, B.: *Thirty Years of Copulas*, in ed: Dall'Aglio, G.; Kotz, S.; Salinetti, G.: *Advances in Probability Distributions with Given Marginals: Beyond the Copulas*, Springer.

[33] 1983. Schweizer, B.; Sklar, A.: *Probabilistic Metric Spaces*. North-Holland, Amsterdam.

[34] 1959. Sklar, A.: *Fonctions de repartition a n dimensions et leurs marges*. Publications de l'Institut de Statistique de l'Universite de Paris, Volume 8, pp. 229-231.

[35] 2010. Stark, P.; *SticiGui*, http://statistics.berkeley.edu/ stark/SticiGui/index.html.

[36] 2007. Trivedi, P.; Zimmer, D.: *Copula Modeling: An Introduction for Practitioners*, Now Publishers.

[37] 1998. Yi, W.; Bier, V.: *An Application of Copulas to Accident Precursor Analysis*. Management Science, Volume 44, pp. S257-S270.