Bernstein Approximations to the Copula Function and Portfolio Optimization*

Alessio Sancetta and Stephen Satchell
Faculty of Economics and Politics
Trinity College, Cambridge, UK

June 4, 2001

Abstract

The Copula function is considered within the context of financial multivariate data sets that are not normally distributed. The Bernstein polynomial approximation to copulae is given and motivated by its desirable properties. The multivariate convergence properties are analyzed. The concept of Bernstein copula is introduced as a generalization of some bivariate and higher dimensional families of copulae. Statistical properties of the Bernstein copula are studied together with implementation issues related to portfolio theory and expected utility optimization.

Key words: Copulae; Bernstein polynomials; approximation theory; portfolio optimization.

JEL: G11, C49.

*We would like to thank Mark Salmon for interesting us both independently in this topic.

1 Introduction

Multivariate normality is often assumed in portfolio theory in order to draw simple results. However, the normality assumption is violated in practice. Financial returns are found to be leptokurtic, *inter alia* Mandelbrot (1963, 1997) and Fama (1965). Therefore, a different kind of approach is needed in modeling multivariate returns when constructing a portfolio. It is desirable to find a general form for a distribution that would allow us both to describe the body of the distribution and the tails. This is particularly true for multivariate data where the degree of dispersion is high and the tails of a distribution become more important. In general it is desirable to have a distribution that could be used to model both normal and extreme events, a situation relevant to financial economics. Moreover, cross dependence and association among assets cannot be fully captured by such linear measures as correlation. Therefore, it is necessary to consider more general kinds of dependences to model the joint distribution of financial returns. However, results should allow for further inference and practical implementation. The purpose of this paper is to address the above problems allowing for solutions that can be used for inference purposes. Our approach can be seen as a competitive alternative to multivariate normality, or elliptic distributions in general when modeling the construction of a portfolio.

We consider the use of copulae and their potential application to portfolio construction. However, this is not a survey of copulae's families. We present an approximation to copulae that in some instances is exact for certain families up to an additive term. This approximation is obtained through Bernstein polynomials. It has some desirable properties like preserving convexity of all orders. Bernstein polynomials are connected with the theory of singular integrals, with probability (Feller, (1966) p. 218-230), with the sum of divergent series and with other branches of mathematics, Lorentz (1953).

We will discuss the use of Bernstein polynomials to approximate functions on the unit hyper-cube. We consider their speed of convergence to the approximated function. Moreover, we establish the link between the copula function and Bernstein polynomials, subject to parameter restrictions.

Having discussed the necessary theoretical results, we concentrate on its practical use in financial economics. We theoretically derive the density and characteristic function of an arbitrary k dimensional portfolio with fixed marginals. For our purpose the marginals are Weibull distributions. These find theoretical justification in the theory of multiplicative processes and are particularly suited for fat tails and extreme value theory; see Frisch and Sornette (1997), and Laherrère and Sornette (1998). Regardless of the dimension and dependency complexity of the portfolio, its characteristic function is easily derived. We show that once a result is established for the univariate case its extension to k dimensions is trivial by virtue of the use of the Bernstein approximation. Moreover, the characteristic function of the portfolio can be transformed via the use of a constant risk aversion utility function and optimized without discarding information as in mean-variance optimization. A worked example shows the potential of our approach.

The plan for the paper is as follows: section 2 discusses the reasons for using copulae in portfolio optimization; section 3 considers the approximation of continuous functions in the k dimensional hyper-cube; section 4 focuses on some technical issues of copulae and defines parametric conditions on the Bernstein polynomial to be a copula; section 5 uses the Bernstein approximation to derive results related to portfolio optimization and presents a simple worked example; section 6 concludes the paper. The appendix contains some proofs and give an alternative representation of the Bernstein copula through transcendental functions like the incomplete beta function and Gauss' hypergeometric series.

2 Portfolio Theory and Distributions

Results based on normality might not be appropriate for optimal portfolio construction and management. Optimal portfolios occur in all areas of intertemporal economics. It may be necessary to derive quantitative results that do not rely on normality.

For example, the distribution of returns is of fundamental importance in the capital asset pricing model (CAPM). As Hagerman (1978) puts it, "if the distribution of security returns is not stable under addition, it is hard to envision a theory of how to combine securities optimally into portfolios." Tobin (1958) showed that if asset returns are normally distributed then the variance is the proper measure of risk. Capital asset pricing models typically assume normality or ellipticity in the distribution of returns. If this assumption is not validated, then the model might deliver results that are not satisfactory.

The study of financial returns dates back to the beginning of the last century. Bachelier (1900) was the first to rigorously study the behavior of speculative prices. The results of his thesis implied that price changes are independent and identically distributed. That is, returns are white noise. This does not imply that they are

Gaussian, i.e. strict white noise.

Mandelbrot (1963) started to revive interest in the time series properties of assets prices. He observed that the unconditional distribution of many economic and financial variables have thick tails. Further, he noticed that variances are not constant and that they are correlated: large changes are followed by large changes of either sign.

Serious examination of the normality assumption in the CAPM by Fama (1965) confirmed Mandelbrot's findings. Both of them proposed that security returns follow a stable symmetric distribution with an infinite variance. Other two competing hypotheses were proposed. One assumes that a mixture of normal underlying distributions is the result of the empirical distribution of returns. The second is that daily security returns follow a student's t-distribution with more than two degrees of freedom. For further discussion and empirical study see Hagerman (1978) and Kon (1984).

The implication for the stability on the tail of the distribution is discussed next. For example, if W is an α-stable random variable, then for $\alpha < 2$, $\Pr(W > w) \simeq w^{-\alpha} L(w)$, where $L(...)$ is a slowly varying function; see Feller (1966) p. 268. However, if $\alpha = 2$ we have Brownian motion and the tail behavior differs from the above case. Therefore, for $\alpha < 2$, the returns can be closely approximated by a line with slope equal to α in log-log plot. The finding of Mandelbrot for cotton price changes showed $\alpha = 1.7$ (see Mandelbrot (1997), p. 34). The interesting implication is that if, for example, in the case of $\alpha = 2$ the probability of ruin is 10^{-20}, for $\alpha = 1.7$ the probability is approximately 10^{-1}. This should make one reflect before assuming normality.

Furthermore, it is necessary to use a measure that allows us to derive optimal weights for a portfolio. In order to do so, it is fundamental to construct families of multivariate distributions with a wide range of dependance properties that are not restricted to have normal marginals. In fact, if marginal densities are not normal, it is impossible to assume multivariate normality for their multidimensional extension. However, returns that are not normally distributed will not have a straightforward multivariate density extension. Therefore, it is important to find alternative ways to characterize the joint distribution of financial returns.

Moreover, the concept of dependency must be completely revised once the normality assumption is dropped. While the covariance matrix captures the essence of dependency for normally distributed assets, this does not hold for variables that are not normally distributed. Basic probability shows that no dependency implies zero covariance, but the other way does not follow. The reason lies in the particular kind of dependency measured by the covariance, namely linear dependency. See Embrechts et al. (1999) for further discussion on this.

One way to tackle the problem is just to use some non parametric technique. Usually, non parametric techniques are not scale invariant, a desirable property when dealing with a complex range of dependency for several assets. Nonparametric estimation is fully general, but because of this important information is discarded. An improvement on this can be achieved by semiparametric estimation. We think in particular of the case of known functional form for the marginal distributions. Most important, multidimensional nonparametric estimation results in problems due to the sparsity of data. Hence, the sample size necessary to obtain acceptable mean square errors must be too large even for financial time series. Estimation of a ten dimensional normal density through a normal kernel would require a sample size of 842,000 to ensure a mean square error less than 0.1 at zero; see Silverman (1986).

This is the reason for looking to some other alternatives, one of interest being copulae. However, before explicitly looking at copulae and their definition, we consider some results that will be important in the sequel and in the discussion of the copula function.

3 Approximations

There are several ways to estimate multivariate distribution functions. The choice should be based on how much one is willing to assume and how much efficiency one is willing to give up. As it will be shown in a worked simulation, a parametric copula function not only guaranties the efficiency of parametric estimation, but it also allows for a complex range of dependance in the data.

In many cases the copula function might have a very complex form and further calculations could be difficult to perform, for example portfolio optimization. Therefore, it is necessary to approximate the estimated function. Here we suggest the use of an operator with shape preserving property, i.e. that preserves convexity of all orders. Such a property is indeed desirable, obviously not necessary. Within the class of linear operators, we look at Bernstein polynomials. Bernstein polynomials have a slower rate of convergence as compared to other polynomial approximations.[1] However, they have the best rate of convergence within the class

[1] The simple Bernstein approximation can be improved by taking linear combitnations; see Butzer (1952b). Let $f^{(2l)} \in Lip\gamma$ be the $2l$ derivative of f, then, Buzter (1952b) shows that his

of all operators with the same shape preserving property; see Berens and DeVore (1980). In the last two decades there has been some interest for approximation theory in econometrics; *inter alia* Phillips (1982, 1983). Phillips advocate the use of extended rational approximates, though for different purposes. However, in the setting of portfolio optimization and copulae, we find the Bernstein approximation more tractable and adequate than, for example, rational polynomials. We remark on this later when we prove some convergence results. Moreover, simple restrictions on the coefficients of the polynomial allow us to derive a generic family of copulae. It is necessary to notice that some approximations to copulae including the Bernstein approximation have been recently considered in the mathematics literature by Kulpa (1997) and Li et al. (1998) in order to define convergence notions that would lead to joint continuity of the *-product defined by Darsow et al. (1992).

3.1 K Dimensional Bernstein Polynomials

The treatment in this paper is based on a condition that we make explicit.

Condition 1. *$f : A \to C$, $A \subset \Re^K$, $C \subset \Re$, where f is continuous, A and C are compact, i.e. f is defined on the Banach space $C_{[A]}$, the set of all continuous bounded functions in A.*

The following is a useful theorem on linear monotone operators. It is followed by the Weierstrass approximation theorem for functions of k variables. These theorems are fundamental to the understanding of our approximation strategy.

Theorem 1. *Linear Monotone Operators for Functions of k Variables. Under Condition 1, for any sequence of monotone linear operators G_n on f the following conditions are equivalent:*

i $G_n f \to f$ (uniformly) for any $f \in C_{[A]}$

ii $G_n f \to f$ for the following functions $f = 1, x_i, x_i^2$

iii $G_n 1 \to 1$ and $G_n \phi_T(T) \to 0$ uniformly in $T \in \Re^K$, where $\phi_T(X) \equiv \|T - X\|^2$.

Proof. See the Appendix. ■

Theorem 2. *Weierstrass Approximation for Functions of k Variables (Stone). Under Condition 1, to each $\epsilon > 0$ there corresponds a polynomial P such that $|f(X) - P(X)| < \epsilon$, for any $X \in A$.*[2]

---

liner combination of one dimensional Bernstein polynomials (equation (10)) has error $O\left(n^{-l-\gamma}\right)$ compared to $O\left(n^{-2l-\gamma}\right)$ for the best polynomials of order n.

[2] This is a particular instance of the Stone Weierstrass theorem. The proof given by Stone relies on the properties of the space where f is defined. It proves the existence of such polynomials, i.e. that the set of polynomials is dense in the set of continuous functions in a compact space. The

Proof. See Appendix ■

The proof we give in the appendix is through k dimensional Bernstein polynomials, which are defined next, inductively from the two dimensional definition of Butzer (1953).

Definition. *Let f be as in Condition 1, B_n be a monotone linear operator, and $P_{v_j,n_j}(x_j) \equiv \binom{n_j}{v_j} x_j^{v_j}(1-x_j)^{n_j-v_j}$, then*

$$(B_n f)(X) \equiv \sum_{v_1=0}^{n_1} ... \sum_{v_k=0}^{n_k} f\left(\frac{v_1}{n_1}, ..., \frac{v_k}{n_k}\right) P_{v_1,n_1}(x_1) \cdots P_{v_k,n_k}(x_k) \tag{1}$$

is a k dimensional Bernstein polynomial.

Before concluding this section we notice the following representation in terms of a Riemann Stieltjes integral of a one dimensional Bernstein polynomial (the k dimensional extension trivially follows).

$$\begin{aligned}(B_n f)(x) &\equiv \sum_{v}^{n} f\left(\frac{v}{n}\right) P_{v,n}(x) \\ &= \int_0^1 f(t) d_t Kn(x,t), \end{aligned} \tag{2}$$

where

$$\begin{aligned} Kn(x,t) &\equiv \sum_{v \le nt} \binom{n}{v} x^v (1-x)^{n-v}, \\ Kn(x,0) &\equiv 0 \end{aligned}$$

is the kernel function that is constant for $\frac{v}{n} \le t < \frac{v+1}{n}$ and has a jump of $\binom{n}{v} x^v (1-x)^{n-v}$ at $t = \frac{v}{n}$. This representation establishes some clear parallels to kernel density estimation in statistics. However, here the term kernel function is used in the language of singular integrals. We now turn to some convergence issues.

3.2 Convergence

By the convergence of the Bernstein polynomial it follows (e.g. see Feller (1966) p. 481) that $\phi_n(t) \to \phi(t)$, where $\phi_n(t)$ and $\phi(t)$ are the characteristic functions of $(B_n f)(X)$ and $f(X)$. Therefore, convergence of all moments is guaranteed. Obviously, the necessary and sufficient condition is that $\phi(t)$ exists.

---

proof given in the appendix is directly relevant and more concrete for our purposes.

Moreover, it is of interest to know the speed of convergence. Here, we present some results on convergence of the Bernstein polynomial approximating functions of k variables under the Tchebysheff's norm. The L_∞ norm is the most common choice of norm in this approximation context.

$$\begin{aligned}(B_n f)(X) - f(X) &= \sum_{v_1}^{n_1} \dots \sum_{v_k}^{n_k} P_{v_1,n_1}(x_1) \cdots P_{v_k,n_k}(x_k) \\ &\quad \times \left[f\left(\frac{v_1}{n_1}, \dots, \frac{v_k}{nk}\right) - f(x_1, \dots, x_k) \right] \\ &= \sum_{v_1}^{n_1} \dots \sum_{v_k}^{n_k} P_{v_1,n_1}(x_1) \cdots P_{v_K,n_K}(x_k) \int_{(x_1,\dots,x_k)}^{\left(\frac{v_1}{n_1},\dots,\frac{v_k}{n_k}\right)} \nabla f dr\end{aligned}$$

$\nabla f \equiv [f'^{1}(s_1, \dots, s_k), \dots, f'^{k}(s_1, \dots, s_k)]$, where $f'^{j}(s_1, \dots, s_k) \equiv \frac{\partial f(s_1,\dots,s_k)}{\partial s_j}$, and r is a vector valued function that defines the path between the end points of the integral. By definition, ∇f is a conservative vector field, so the path of integration is irrelevant. The above line integral can be split into k integrals along any paths parallel to the axis and perpendicular to each other. For example, we can write

$$\begin{aligned}\int_{(x_1,\dots,x_k)}^{\left(\frac{v_1}{n_1},\dots,\frac{v_k}{n_k}\right)} \nabla f dr &= \int_{x_1}^{\frac{v_1}{n_1}} f'^{1}(s_1, x_2, x_3, \dots, x_k)\, ds_1 + \dots \\ &\quad + \int_{x_j}^{\frac{v_j}{n_j}} f'^{j}\left(\frac{v_1}{n_1}, \frac{v_2}{n_2}, \dots, s_j, \dots, x_k\right) ds_j + \dots \\ &\quad + \int_{x_k}^{\frac{v_k}{n_k}} f'^{k}\left(\frac{v_1}{n_1}, \frac{v_2}{n_2}, \dots, \frac{v_{k-1}}{n_{k-1}}, s_k\right) ds_k\end{aligned}$$

Now,

$$\int_{x_j}^{\frac{v_j}{n_j}} f'^{j}\left(\frac{v_1}{n_1}, \frac{v_2}{n_2}, ..., s_j, ..., x_k\right) ds_j = f'^{j}\left(\frac{v_1}{n_1}, ..., \frac{v_j}{n_j}, x_{j+1}, ..., x_k\right)\left(\frac{v_j}{n_j} - x_j\right)$$
$$- \int_{x_j}^{\frac{v_j}{n_j}} (s_j - x_j)\, df'^{j}\left(\frac{v_1}{n_1}, \frac{v_2}{n_2}, ..., s_j, ..., x_k\right). \qquad (3)$$

From here a crude result can be obtained by assuming that $f'^{j} \in Lip_{M_j}1$, i.e. f'^{j} satisfies the Lipschitz condition with constant M_j and exponent 1: $|f'^{1}(s_1, ..., s_j + h_j, ..., s_k) - f'^{1}($ $M_j |h_j|$. It follows that the last integral in (3) does not exceed $M_j \int_{x_j}^{\frac{v_j}{n_j}} (s_j - x_j)\, ds_j = \frac{1}{2} M_j \left(\frac{v_j}{n_j} - x_j\right)^2$.

Therefore,

$$\begin{aligned} |(B_n f)(X) - f(X)| &\leq \sum_{v_1}^{n_1} ... \sum_{v_K}^{n_K} P_{v_1, n_1}(x_1) \cdots P_{v_K, n_K}(x_k) \\ &\quad \times \frac{1}{2} \sum_{j=1}^{k} M_j \left(\frac{v_j}{n_j} - x_j\right)^2 \\ &= \frac{1}{2}\left[M_1 \frac{x_1(1 - x_1)}{n_1} + ... + M_k \frac{x_k(1 - x_k)}{n_k}\right] \end{aligned}$$

for any X. This is just the multivariate analog of the standard univariate result. Therefore, we have just proved the following.

Theorem 3. *Let f be as in Condition 1, and $f'^{j} \in Lip_{M_j}1$, then*

$$|(B_n f)(X) - f(X)| \leq \sum_{j=1}^{k} M_j \frac{x_j(1 - x_j)}{2n_j},$$

where B_n is the Bernstein operator and the M_j's are constants.

Remark. *We only required f to be defined on a Banach space without restriction to the k dimensional unit hyper-cube. This can be achieved by a simple transformation: $x \in [a, b] \Longrightarrow t \in [0, 1]$, $t \equiv \frac{x-a}{b-a}$. In general, we can define a transformation that makes the real line isomorphic to the unit interval: $x \in \Re \Longrightarrow t \in [0, 1]$, $t \equiv \frac{x}{1-x} - \frac{1-x}{x}$.*

Remark. *The reader should note that we are interested in applying this approximation to the copula function. The copula is a special case of a t-norm as studied in probabilistic metric spaces (see Schweizer (1991) for a short review). Therefore, modeling the copula confine our attention to modeling the distance between the marginal distributions. It is the case that sharp changes in the gradient of a copula density are located in the tail area. This is exactly where, unlike other kind of approximations, the Bernstein approximation is relatively superior.*

A better result can be achieved by the use of *Theorem 3.1* in Schurer and Steutel (1977) for the univariate case:

$$\sup \frac{|(B_n f)(x) - f(x)|}{w_1(f;\delta)} = \Delta_n\left(\tilde{f};\delta\right),$$

where

$$w_1(f;\delta) = \sup\left\{\left|f'(x+h) - f'(x)\right| : |h| \le \delta\right\}$$

is the modulus of continuity of f',

$$\begin{aligned}
\Delta_n(f;\delta) &\equiv (B_n f)(x) - f(x) \\
&= \sum_{v=0}^{n} P_{k,n}(x) \int_{x}^{\frac{v}{n}} f'(t)dt,
\end{aligned}$$

and $\tilde{f}$ is defined for $x \in \Re$ by

$$\begin{aligned}
\tilde{f}(x_0) &= 0, \\
\tilde{f}'(x) &= l + \frac{1}{2}, \\
(l\delta &< x - x_0 \le (l+1)\delta,\ l = 0, \pm 1, \pm 2, ...).
\end{aligned}$$

Then, $\tilde{f}(x) = \frac{1}{2}|x - x_0| + \sum_{l=1}^{\infty}(|x - x_0| - l\delta)_+$, $g_+ \equiv \max(g, 0)$. Heuristically, this function is an extremal function such that $\tilde{f}'(x)$ is equal to $\sup \frac{f'}{w_1(f;\delta)}$; see Schurer and Steutel (1977) for details.

Let

$$w_{1_j}(f;\delta_j) = \sup\left\{\left|f'^j(x_1, ..., x_j + h_j, ..., x_k) - f'^j(x_1, ..., x_j, ..., x_k)\right| : |h_j| \le \delta_j\right\}.$$

Therefore,

$$\begin{aligned}
\tilde{f}'^j(s_j) &= l_j + \frac{1}{2}, \\
(l_j\delta &< s_j - x_j \le (l_j+1)\delta, l_j = 0, \pm 1, \pm 2, ...).
\end{aligned}$$

Now,

$$\int_{(x_1,\ldots,x_k)}^{\left(\frac{v_1}{n_1},\ldots,\frac{v_K}{n_K}\right)} \nabla f dr \leq \sum_{j=1}^{k} w_{1_j}(f;\delta_j) \int_{x_j}^{\frac{v_j}{n_j}} \tilde{f}'^{j}(s_j) ds_j$$

and the result for the k dimensional case follows. See Schurer and Steutel (1980) for a detailed treatment of the two dimensional case. Their approach is slightly different than ours. However, the actual calculations become prohibitive very soon.

An alternative proof that is instrumental in showing the properties of the Bernstein operator is through its Riemann Stieltjes integral representation. Its extension can be used for showing results related to nonparametric estimation (Sancetta and Satchell, 2001). We show it for the univariate case and then extend it to the k dimensional space. The error in the approximation can be written as follows,

$$(B_n f)(x) - f(x) = \int_0^1 f(t) d_t Kn(x,t) - f(x).$$

Taking a Taylor expansion of $f(t)$ around x and using the fact that $d_t Kn(x,t)$ integrates to one, we write

$$\begin{aligned}
(B_n f)(x) - f(x) &= \int_0^1 [f(x) + f'(x)(t-x) + f''(x)(t-x)^2 \\
&\quad + f'''(h)(t-x)^3 - f(x)] d_t Kn(x,t) \\
&= \int_0^1 [f'(x)(t-x) + f''(x)(t-x)^2 \\
&\quad + f'''(h)(t-x)^3] d_t Kn(x,t),
\end{aligned}$$

where $h = \rho x + (1-\rho) t$, $\rho \in [0,1]$. Then we just use a recurrence formula as in Lorentz (1953), p. 14. We state a version of it that is relevant to our calculations.

Lemma 1. *Let* $T_{n,s}(x) \equiv \sum_{v}^{n} (v - nx)^s \binom{n}{v} u^v (1-u)^{n-v}$, *n=1,2,... and s=0,1,...,* *then* $T_{n,0}(x) = 1$, $T_{n,1}(x) = 0$, $T_{n,2}(x) = nx(1-x)$, $T_{n,3}(x) = nx(1-x)(1-2x)$, *and in general* $T_{n,s+1}(x) = x(1-x)\left[T'_{n,s}(x) + nsT_{n,s-1}(x)\right]$.

As $n \to \infty$ $|t - x| \leq \frac{1}{n}$, *then clearly the third term in the expansion is smaller then the second. Therefore,* $(B_n f)(x) - f(x) = \frac{x(1-x)}{n} f''(x) + \frac{\epsilon}{n}$.

The multivariate extension follows by noticing the following.

$$(B_n f)(X) = \int_0^1 \cdots \int_0^1 f(t_1, .., t_k) d_{t_1} Kn(x_1, t_1) \cdots d_{t_k} Kn(x_k, t_k),$$

i.e. the operator can be applied recursively. Therefore, the multivariate Bernstein operator is a k iterate of the Bernstein operator. Writing the multivariate Taylor series expansion, the multivariate extension can be found to be $o\left(\frac{k}{n}\right)$, for fixed k, as in the theorem above when $n_j = n_i$ for all i and j.

As we increase the dimension of the operator, it is apparent that its performance at the end points is relatively superior. Moreover, to establish the approximation it is only required to evaluate the function at different points and take a sum. Computationally, this is an easy task to perform even for high dimensions. This is not the case for other approximations that have a faster rate of convergence.

4 The Copula Function

We are now in the position of considering the Bernstein approximation in conjunction with the copula function as outlined in Section 2. We start by giving a brief overview of copulae.

Copula functions were first used in the study of metric spaces. A classical article on their use in statistics is Sklar (1973). All the relevant basic theory is explained there. Therefore, the reader is referred to it for a brief self-contained discussion. Details and further references can be found in Joe (1993, 1997) and Nelsen (1997, 1998). Some articles of interest are in Dall'Aglio et al. (1991).

Theorem 4. *Sklar (1973). Let H be an n dimensional distribution function with 1-dimensional margins $F_1, F_2, ..., F_n$, then there exists a function C from the unit n-cube to the unit cube such that*

$$H(x_1, x_1, ..., x_1) = C\left(F_1(x_1), F_2(x_2), ..., F_n(x_n)\right);$$

C is referred to as the n-Copula. If each F_J is continuous, the copula is unique.

Proof. See Sklar (1973) ∎

Some properties will be discussed as we consider the properties of the Bernstein approximation. There are many families of copulae. In general a parametric family can be constructed by mixtures. To make the statement more clear we state the following theorem.

Theorem 5. *Marshall and Olkin (1988). Let $H_1, ..., H_n$ be univariate distribution functions, and let G be an n-variate distribution function such that $\overline{G}(0, ..., 0) = 1$, with univariate marginals G_i $(i = 1, ..., n)$. Denote the Laplace transform of G and G_i, respectively, by ϕ and ϕ_i $(i = 1, ..., n)$. Let K be an n-variate distribution function with all univariate marginals uniform on [0,1]. If $F_i(x) = \exp[-\phi_i^{-1} H_i(x)]$ $(i = 1, ..., n)$, then*

$$H(x_1, ..., x_n) = \int \cdots \int K\left(F_1^{\theta_1}(x_1), ..., F_n^{\theta_n}(x_n)\right) dG(\theta_1, ..., \theta_n)$$

is an n-variate distribution function with marginals $H_1, ..., H_n$.

Proof. See Marshall and Olkin (1988). ■

As a consequence of the above theorem,

$$H(x_1, ..., x_n) = \phi\left(\phi_1^{-1} H_1(x_1), ..., \phi_n^{-1} H_n(x_n)\right)$$

when $K(x_1, ..., x_n) = \prod_{i=1}^{n} x_i$. Clearly $C(u_1, ..., u_k) = \phi\left(\phi_1^{-1} u_1, ..., \phi_n^{-1} u_1\right)$ is a copula, where $u_j \equiv H_j(x_j)$.

4.1 The Bernstein Approximation to the Copula Function

By the Weierstrass theorem it is possible to approximate the above k dimensional copula by a Bernstein polynomial, i.e.

$$(B_n C)(u_1, ..., u_k) = \sum_{v_1} ... \sum_{v_k} \phi\left(\phi_1^{-1} \frac{v_1}{n_1}, ..., \phi_k^{-1} \frac{v_k}{n_k}\right) P_{v_1, n_1}(u_1) \cdots P_{v_K, n_k}(u_k),$$

where $P_{v_j, n_j}(u_j)$ is defined as in the previous section. However, it is interesting to note that some simple families of copulae have a particular structure that very much resemble a Bernstein polynomial. These are families with quadratic, cubic and hyper-cubic sections, i.e. multivariate polynomials of quadratic or cubic order. A simple example of a copula with polynomial structure is the Farlie-Gumbel-Morgenstern copula:

$$C(u_1, u_2) = u_1 u_2 + \theta u_1 u_2 (1 - u_1)(1 - u_2), \quad (4)$$

with $|\theta| \leq 1$ being a measure of either positive or negative dependance. However, the Farlie-Gumbel-Morgenstern copula has very little application to real problems. It is only a simple perturbation of the independence copula $u_1 u_2$. Its Spearman's rho is bounded in absolute value by 1/3.[3] Subject to specific constraints the above copula can be generalized to a polynomial of arbitrary order.

Copulae with polynomial structure have a nice interpretation in terms of conditional probabilities. In fact, the conditional copula is basically the same copula but with a lower dimensional section in the variable we condition on. For the copula in (4) this implies

$$\frac{\partial C(u_1,u_2)}{\partial u_2} = u_1 + \theta u_1(1-u_1)(1-2u_2). \tag{5}$$

As it can be seen, (5) has linear sections in u_2. For a treatment of copulae with polynomial structure and their properties see Nelsen (1998).

The following generalization can be proved to be a copula with hyper-cubic section in u_j:

$$C(u_1,...,u_k) = u_1\cdots u_k + \sum_{v_1}...\sum_{v_k}\alpha_{v_1,...,v_k}P_{v_1,n_1}(u_1)\cdots P_{v_k,n_k}(u_k) \tag{6}$$

Clearly, its resemblance to a Bernstein polynomial up to the additive term $u_1\cdots u_k$ is apparent. From section 3.1 we know that for finite n_j's $f\left(\frac{v_1}{n_1},...,\frac{v_k}{n_k}\right)$ in (1) is a step function in virtue of its arguments. If we legitimately let $\alpha_{v_1,...,v_k} = \alpha\left(\frac{v_1}{n_1},...,\frac{v_k}{n_k}\right)$ we just have a Bernstein polynomial. Therefore, we can refer to it as the Bernstein copula. However, its implementation as a copula function would require an estimation procedure that is more in the spirit of non semiparametric than parametric estimation.

It follows from the properties of the Bernstein linear operator that $\alpha(t_1,t_2,...,t_k)$ plays a crucial role in determining the dependence structure of the vector of uniform random variables $U = (u_1,\cdots,u_k)'$. For the above to be a copula, restrictions have to be imposed on $\alpha(t_1,t_2,...,t_k)$. Then, the parallel with a copula is direct. By an application of set theory to the probability of events, copulae satisfy the following inequality, see Joe (1997) or Nelsen (1998),

$$\max\{0, u_1+\cdots+u_k-(k-1)\} \leq C(u_1,...,u_k) \leq \min_{u_j}(u_1,...,u_k),$$

[3]Spearman's rho is a bivariate measure of dependence. For a copula $C(u_1,u_2)$, it is defined as

$$\begin{aligned}\rho_S &\equiv 12\int_0^1 u_1u_2dC(u_1,u_2)-3\\ &= 12\int_0^1 \bar{C}(u_1,u_2)du_1du_2-3\end{aligned}$$

where $\bar{C}(u_1,u_2) = 1-u_1-u_2+C(u_1,u_2)$ is the survival copula.

which are the Fréchet bounds. Moreover, $C(1,...,u_j,...,1) = u_j$ ought to be satisfied. The latter condition just says that integrating out all the other variables, we are just left with one marginal, that marginal being the distribution function of a uniform $[0,1]$ random variable. Therefore, $\alpha(t_1,...,t_k) = 0$ in (6) if at least $k-1$ of its k arguments are one. Notice that the Fréchet bounds apply if any of the arguments in $C(u_1,...,u_k)$ is zero, then $\alpha(t_1,...,t_k) = 0$ if any of its arguments is equal to zero.

It follows from section 3.1 that (6) can be regarded as an approximation to a copula with the following structure:

$$C(u_1,...,u_k) = u_1 \cdots u_k + \alpha(u_1,...,u_k).$$

Notice that for $C(u_1,...,u_k) = u_1 \cdots u_k$ the approximation is exact (see the proof of theorem 2 in the appendix), this is why we can just subtract $u_1 \cdots u_k$ from $C(u_1,...,u_k)$ and add it outside the Bernstein operator. Then, what we obtain is (6). The reason for explicitly extracting the term $u_1 \cdots u_k$ is to give an intuitive representation of a copula as the sum of the independence copula and a perturbation factor that can be very complex in nature. We can state this in a lemma.

Lemma 2. *Any copula $C(u_1,...,u_k)$ can be written as $u_1 \cdots u_k + G(u_1,...,u_k)$, where $u_1 \cdots u_k$ represents the case of independence and $G(u_1,...,u_k)$ is a perturbation factor containing all information about the dependence of $(u_1,...,u_k)$.*

Proof. It follows by the uniform convergence of the Bernstein operator. ■

Remark. *Notice that $G(u_1,...,u_k)$ is the distance of the copula from the independent copula. This is bounded above and below by the Fréchet bounds. For a 2-copula, the Fréchet bounds define a skewed quadrilateral where the product copula is the paraboloid inside it.*

For the Bernstein approximation, some particular care is needed for the perturbation factor to allow for a valid copula. Without loss of generality, we focus on the bivariate case. Write $\alpha\left(\frac{v_1}{n_1}, \frac{v_2}{n_2}\right) \equiv \gamma\left(\frac{v_1}{n_1}, \frac{v_2}{n_2}\right) - \frac{v_1}{n_1}\frac{v_2}{n_2}$, where for all practical matters we know that $\gamma(...)$ is very much related to some copula $C(...)$. However, for the time being, we want to treat the Bernstein approximation as a copula itself and therefore we use this notation. The following conditions must hold:

$$\gamma\left(\frac{v_1}{n_1}, \frac{v_2}{n_2}\right) \leq \min\left(\frac{v_1}{n_1}, \frac{v_2}{n_2}\right)$$

$$\gamma\left(\frac{v_1+m_1}{n_1},\frac{v_2+m_2}{n_2}\right)+\gamma\left(\frac{v_1}{n_1},\frac{v_2}{n_2}\right)\geq\gamma\left(\frac{v_1+m_1}{n_1},\frac{v_2}{n_2}\right)+\gamma\left(\frac{v_1}{n_1},\frac{v_2+m_2}{n_2}\right),$$

for any $0\leq m_j\leq n_j-v_j$;

$$\lim_{v_j\rightarrow n_j}\gamma\left(\frac{v_1}{n_1},\frac{v_2}{n_2}\right)=\frac{v_{-j}}{n_{-j}}$$

and

$$\lim_{v_j\rightarrow 0}\gamma\left(\frac{v_1}{n_1},\frac{v_2}{n_2}\right)=0,$$

where if $j=1$ then $-j=2$. The first condition is a consequence of the upper Frechét bound. The second follows by cumulative distribution functions satisfying the rectangle inequality. The third is necessary for having fixed marginals. The fourth is trivial.

Moreover,

$$\alpha\left(\frac{v_1}{n_1},\frac{v_2}{n_2}\right)=\frac{v_j}{n_j}-\frac{v_1}{n_1}\frac{v_2}{n_2},\quad j=\min_j(u_1,u_2)$$

is the Frechét upper bound. For no simple restriction on $\alpha\left(\frac{v_1}{n_1},\frac{v_2}{n_2}\right)$ is the Frechét lower bound attained. A few other characteristics are as follows: permutation symmetry if $\alpha(t_1,t_2,...,t_k)$ is symmetric in its arguments; symmetry about medians if $\alpha\left(\frac{v_1}{n},\frac{v_2}{n},...,\frac{v_k}{n}\right)=\alpha\left(\frac{n-v_1}{n},\frac{n-v_2}{n},...,\frac{n-v_k}{n}\right)$, $n_i=n_j$ for any i and j.

Dependance properties can be checked for the Bernstein copula. The terminology is standard in multivariate analysis. For a review see Joe (1997) or Nelsen (1998). A bivariate copula is positive quadrant dependent (PQD) if

$$\Pr(u_1>a_1,u_2>a_2)\geq\Pr(u_1>a_1)\Pr(u_2>a_2),$$

i.e. $G(u_1,u)\geq 0$. It is negative quadrant dependent (NQD) if the inequality is reversed. PQD implies that large (small) values of one variable are likely to occur with large (small) values of the other. NQD implies the opposite, then it is a measure of discordance. It is stochastically increasing (SI) in say u_2, if

$$\Pr(u_1>a_1|u_2=a_2)=1-C(a_1|a_2)\tag{7}$$

is decreasing in a_2 for any a_1. Stochastically decreasing is obtained when (7) is decreasing in a_2 for any a_1. Therefore, the Bernstein copula is SI if

$$\sum_{v_1}\sum_{v_2}\alpha\left(\frac{v_1}{n_1},\frac{v_2}{n_2}\right)\binom{n_1}{v_1}\binom{n_2}{v_2}a_1^{v_1}(1-a_1)^{n_1-v_1}a_2^{v_2-1}(1-a_2)^{n_2-v_2}$$

is decreasing in a_2 for any a_1. These properties and a few other important ones can be established once the parameters of the polynomial are known or some functional form to generate the parameters is specified. For simplicity we just mentioned some properties for the bivariate case. Indeed, dependence properties are easier to handle for the bivariate case, but they also provide a better intuition behind the concepts for the multivariate extension. However, some generality is lost and not all definitions will directly extend to the multivariate case. The reason for being interested in these measure of dependence is that they provide a link with stochastic orders. Stochastic orders can often be defined for variables having common copula. The fact that the copula is invariant under increasing transformations of the marginals allows us to apply it not only to intertemporal optimization, but also to areas of economics such as income distribution and applications of stochastic orders. We conclude this subsection and turn our interest to the density.

4.2 The Bernstein Density Copula

Differentiating (6) with respect to each variable and rearranging we can easily see the close link in structure with the density,

$$\begin{aligned}\frac{\partial^k C(u_1,...,u_k)}{\partial u_1 \cdots \partial u_k} &= 1 + n_1 \cdots n_k \sum_{v_1=0}^{n_1-1} ... \sum_{v_k=0}^{n_k-1} \Delta\alpha_{1,...,k}\left(\frac{v_1}{n_1},...,\frac{v_k}{n_k}\right) \\ &\quad \times P_{v_1,n_1-1}(u_1) \cdots P_{v_K,n_k-1}(u_k).\end{aligned}$$

Differentiating, a term in the summation is lost and the coefficients of the polynomial are written as a difference form which is directly linked to the generalization of the rectangle inequality,

$$\Delta_{1,...,k}\alpha\left(\frac{v_1}{n_1},...,\frac{v_k}{n_k}\right) \equiv \sum_{m_1=0}^{1} ... \sum_{m_k=0}^{1} (-1)^{k+m_1+...+m_k}\, \alpha\left(\frac{v_1+m_1}{n_1},...,\frac{v_k+m_k}{n_k}\right).$$

The above result follows by applying the difference operator Δ recursively. Therefore, we can easily find any derivative with respect to any of the variables. See Lorentz (1953) for the univariate case.

However, if interest lies on the density, it is more convenient to use the following definition for the density,

$$
\begin{aligned}
c(u_1,...,u_k) &= 1+\sum_{v_1=0}^{n_1}...\sum_{v_k=0}^{n_k}\beta\left(\frac{v_1}{n_1},...,\frac{v_k}{n_k}\right) \\
&\quad \times\prod_{j=1}^{k}\binom{n_j}{v_j}u_j^{v_j}(1-u_j)^{n_j-v_j}. \qquad (8)
\end{aligned}
$$

Here, we keep the one outside the operator for expositional convenience in what follows. The necessary and sufficient conditions for it to be a copula density are $c(u_1,...,u_k)\geq 0$ everywhere, and that integrating out all variables we are left with one marginal. The second condition is worth exploring. We proceed in an heuristic way. Assume that (8) is a valid copula density with univariate marginals $u_1,\cdots,u_k$. Integrating all the variables (see Appendix B for some details) and taking limits over their support for $k-1$ of them,

$$
\begin{aligned}
u_l &= u_l+\sum_{v_1=0}^{n_1}...\sum_{v_k=0}^{n_k}\beta\left(\frac{v_1}{n_1},...,\frac{v_k}{n_k}\right)\prod_{j\neq l}\binom{n_j}{v_j}B(v_j+1,n_j-v_j+1) \\
&\quad \times\binom{n_l}{v_l}B_{u_l}(v_l+1,n_l-v_l+1),
\end{aligned}
$$

where $B(a,b)$ and $B_c(a,b)$ are the beta and incomplete beta function. Though it is written in a complex way, the above expression is just a polynomial of order $n+1$. It can be written (see the appendix) as

$$
\begin{aligned}
u_l &= u_l+\sum_{v_1=0}^{n_1}...\sum_{v_k=0}^{n_k}\beta\left(\frac{v_1}{n_1},...,\frac{v_k}{n_k}\right) \\
&\quad \times\prod_{j\neq l}\binom{n_j}{v_j}B(v_j+1,n_j-v_j+1) \\
&\quad \times\binom{n_l}{v_l}\frac{u_l^{v_l+1}}{v_l+1}{}_2F_1(v_l+1,v_l-n_l;v_l+2;u_l) \\
&= u_l+\sum_{p=1}^{n+1}a_p u_l^p,
\end{aligned}
$$

where

$$
\begin{aligned}
a_p &\equiv \sum_{v_1=0}^{n_1}...\sum_{v_l\leq p-1}...\sum_{v_k=0}^{n_k}\beta\left(\frac{v_1}{n_1},...,\frac{v_l}{n_l},...,\frac{v_k}{n_k}\right) \\
&\quad \times\prod_{j\neq l}\binom{n_j}{v_j}B(v_j+1,n_j-v_j+1) \\
&\quad \times\sum_{q=0}^{v_l}\binom{n_l}{q}\frac{(q+1)_{p-q-1}(q-n)_{p-q-1}}{q+1\,(q+2)_{p-q-1}},
\end{aligned}
$$

which is a scalar obtained by collecting terms of the same order, where $(\xi)_s$ is Pochhammer's symbol and is defined in the appendix. Since the set of monomials $\{u \to u^v | v \leq n\}$ is linearly independent, then it must be true that $a_p = 0$ for all p's. However, it is not clear if there exist a solution to $a_p = 0$. By the Binomial theorem, which can be shown to hold for k dimensions (for example see Cheney and Ward, 2000), we know that the integral of a k dimensional Bernstein operator is equal to u_j when we integrate out all other variables. In fact, the sum of all monomials is equal to zero except for x^0 which is equal to one. Therefore, it can easily be shown that there exists a solution to the above problem which guaranties the density to be positive everywhere. One might argue that imposing the above restriction does not allow the density to be exactly zero at the limits of its compact support. However, we are concerned with a copula density and its tails differ substantially from usual densities. For an example, the reader is referred to the appendix for the three dimensional graph of the Kimeldorf and Sampson copula density. The Kimeldorf and Sampson copula is defined in the next subsection. Its density is unbounded at the origin. Therefore the Bernstein approximation cannot be directly applied to it, but just to the cumulative distribution. In a worked example below, we will use a copula with bounded density.

While very general, attempting to estimate a copula directly from a Bernstein polynomial might be computationaly infeasible and not efficient. It is probably more reasonable to use it as an approximation after having estimated a copula through maximum likelihood or some maximum entropy approach to guarantee fixed marginals; see Joe (1987). Indeed the choice is quite vast and should depend on several factors like data and computing power. Moreover, in multidimensional density estimation, it is known that data are much more dispersed and large sample data are required even for semiparametric or parametric estimation. Techniques and specific methodologies for estimating a k dimensional copula will be provided in future work.

4.3 Spearman's Rho and the Moment Generating Function of the Bernstein Copula

At this point some light can be shed on the dependency properties of the Bernstein copula. The copula is

$$\begin{aligned} C(u_1,...,u_k) &= u_1 \cdots u_k + \sum_{v_1=0}^{n_1} ... \sum_{v_k=0}^{n_k} \alpha\left(\frac{v_1}{n_1},...,\frac{v_k}{n_k}\right) \\ &\times \prod_{j=1}^{k} \binom{n_j}{v_j} u_j^{v_j}(1-u_j)^{n_j-v_j}, \end{aligned}$$

and its bivariate marginal distribution, say for u_1 and u_2, is

$$\begin{aligned} C(u_1,u_2,1,...,1) &= u_1u_2 + \sum_{v_1=0}^{n_1}\sum_{v_2=0}^{n_2} \alpha\left(\frac{v_1}{n_1},\frac{v_2}{n_2},1,...,1\right) \\ &\times \prod_{j=1,2} \binom{n_j}{v_j} u_j^{v_j}(1-u_j)^{n_j-v_j}. \end{aligned}$$

We now calculate Spearman's rho. Spearman's rho is a bivariate non-linear measure of dependence. It is the covariance of the distribution of two random variables under their joint probability measure. Therefore, assets which have zero covariance, could have positive Spearman's rho. Its use is advocated on the basis of the documented non-liniarities in finance and its easy practical calculation. For the Bernstein copula it is equal to,

$$\begin{aligned} \rho_S &\equiv 12\int_0^1\int_0^1 [1-u_1-u_2+C(u_1,u_2,1,...,1)]\,du_1du_2 - 3 \\ &= 12\sum_{v_1=0}^{n_1}\sum_{v_2=0}^{n_2} \alpha\left(\frac{v_1}{n_1},\frac{v_2}{n_2},1,...,1\right) \\ &\times \prod_{j=1,2}\binom{n_j}{v_j}\int_0^1\int_0^1 u_j^{v_j}(1-u_j)^{n_j-v_j}du_1du_2 \\ &= 12\sum_{v_1=0}^{n_1}\sum_{v_2=0}^{n_2} \alpha\left(\frac{v_1}{n_1},\frac{v_2}{n_2},1,...,1\right) \\ &\times \prod_{j=1,2}\binom{n_j}{v_j} B\left(v_j+1,n_j+1-v_j\right). \end{aligned}$$

The first equality follows by the definition of a Bernstein copula as the sum of the case of independence and a complex perturbation term, i.e.

$$12\int_0^1\int_0^1 (1-u_1-u_2+u_1u_2)\,du_1du_2 = 3.$$

All dependency information is contained in the perturbation term. Even when the Bernstein copula is used as an approximation, the above Spearman's rho can be used as an approximation to the true Spearman's rho of any copula. If enough terms are included, Spearman's rho can be easily found to any degree of accuracy without the need of evaluating complicated integrals. However, care has to be used in defining

$$\alpha\left(u_1,...,u_k\right) \equiv C(u_1,...,u_k) - u_1\cdots u_k.$$

For the sake of completeness the moment generating function of the density in (8) is found. We do it for the one variable case. Then we just extend it to the k dimensional case.

$$\begin{aligned} M_u(t) &= \int_0^1 \exp\{tu\}\, c(u)du \\ &= \sum_{v=0}^{n} \tilde{\beta}\left(\frac{v}{n}\right)\binom{n}{v}\int_0^1 \exp\{tu\}\, u^v(1-u)^{n-v}du, \end{aligned}$$

where without loss of generality we have omitted the first term and absorbed it into $\tilde{\beta} \equiv \beta - 1$. Before proceeding any further, we notice the following (see Marichev (1983), p. 87),

$${}_1F_1\left(a;c;z\right)B\left(a,c\right) = \int_0^1 \exp\{z\tau\}\,\tau^{a-1}(1-\tau)^{c-a-1}d\tau,$$

Re $c >$Re $a > 0$, where ${}_1F_1(a;c;z)$ is Kummer's confluent hypergeometric function and $\Gamma\left(c\right)$ is the gamma function. For $a \equiv v+1$, $c \equiv n+2$, and $z \equiv t$ this implies

$$\int_0^1 \exp\{tu\}\, u^v(1-u)^{n-v}du = {}_1F_1\left(v+1;n+2;t\right)B\left(v+1,n-v+1\right).$$

Therefore,

$$M_u(t) = \sum_{v=0}^{n} \tilde{\beta}\left(\frac{v}{n}\right)\binom{n}{v}{}_1F_1\left(v+1;n+2;t\right)B\left(v+1,n-v+1\right)$$

To obtain the moment generating function for the k dimensional Bernstein approximation just replace the univariate result in the multivariate definition,

$$\begin{aligned} M_u(t) &= \int_0^1 \cdots \int_0^1 \exp\{t(u_1 + ... + u_k)\} \sum_{v_1=0}^{n_1} ... \sum_{v_k=0}^{n_k} \tilde{\beta}\left(\frac{v_1}{n_1}, ..., \frac{v_k}{n_k}\right) \\ &\quad \times \prod_{j=1}^{k} \binom{n_j}{v_j} u^{v_j} (1 - u_j)^{n_j - v_j} du_1 \cdots du_k \\ &= \sum_{v_1=0}^{n_1} ... \sum_{v_k=0}^{n_k} \tilde{\beta}\left(\frac{v_1}{n_1}, ..., \frac{v_k}{n_k}\right) \\ &\quad \times \prod_{j=1}^{k} {}_1F_1(v_j + 1; n_j + 2; t) B(v_j + 1, n_j - v_j + 1). \end{aligned}$$

These results can be used to further investigate the properties of the Bernstein copula and its approximations. Notice that in the case of approximations the results are qualitatively good but quantitatively misleading if a large order of polynomials is not used. However, approximation by higher order of polynomials is trivial and numerically feasible. Deriving results on the joint moments of the Bernstein copula is quite easy in virtue of its incomplete Beta function representation. The joint moments are important to study the scale free dependence properties of the variables.

To give a simple example of the viability of the Bernstein approximation and its range of dependance we approximate the Kimeldorf and Sampson copula (see e.g. Joe (1997) p. 141), which is equal to

$$C(u, v) \equiv \left(u^{-\theta} + v^{-\theta} - 1\right)^{-\frac{1}{\theta}}.$$

Figure I shows the 3 dimensional graph of the Kimeldorf and Sampson copula density.[4] We report the values of Spearman's rho as a function of the dependance parameter θ in the approximation for $n_1 = n_2$ of order 10, 30, 50 and the corresponding ones for the Kimeldorf and Sampson copula (KS). Figure II and III show the contourplot of the two copulae when $\theta = 1.06$ and $n = 30$. In Table I, values for Spearman's rho in KS are from Joe (1997), values for the Bernstein copula (B_n) were calculated on a Pentium 150 MHz. Because of computational difficulties the limit of the dependence parameter to infinity was not calculated for the approximation. Differences are found as a result of polynomials being fairly

[4] Figures and Tables are in the Appendix.

slow in adjusting at turning points. In this case, improvements can be achieved by increasing the order of the polynomial. Though the polynomial might be of large order, its actual calculation is straightforward. Indeed, all computations required between a fraction of a second and 20 seconds. The evaluation of the integral for the computation of Sperman's rho for the Kimeldorf and Sampson copula could not be performed on the same computer using Maple. By contrast, for the Bernstein copula, the calculations on the same computer are straightforward. A polynomial of order fifty increases the computational time to about twenty seconds, and the error is not too great.

Even though we restricted ourself to Spearman's rho, other estimators can be employed to investigate non linear dependence properties. See Embrechts et al. (1999), Joe (1997) and Nelsen (1998) for a description of other measures of dependence.

5 Implementing an Extreme Portfolio Density: the Case of Marginal Weibull Distributions

In this section we consider a $k \times n$ matrix $X \equiv [X_1, X_2, ..., X_k]$, where X_j is the j^{th} column. We assume that the variables in each column $j = 1, ..., k$ are *iid* copies of some random variable x_j. However, no restriction is imposed on the dependance among column variables. Notice that this assumption is realistic. We can usually find a linear or non linear time series model such that the innovations are *iid*. GARCH models are an example. On the other hand, a multivariate standardization to generate a matrix of *iid* copies is often not feasible to implement and in any case not adequate for our purposes. We think about X_j as being a vector of relative prices. Let the distribution for each random variable have the following functional form as $x_j \to 0$:

$$F(x_j) \sim a_j x_j^{b_j}, \tag{9}$$

defined on the positive real axis, where $a_j, b_j \in \Re_{++}$, $1 \leq j \leq k$. We choose this limiting functional form because of its fat tails properties. The above (9) is a power function with $E(x_j^r) = a_j^{\frac{r}{b_j}} \left(\frac{b_j}{b_j+r}\right)$. We are interested in the limiting distribution of $Z_j \equiv \min(X_{1j}, X_{2j}, ..., X_{nj})$ as $n \to \infty$. In order to find the limiting distribution of X_j it is necessary to find a suitable standardization. Standardizing by mean and standard deviation is not adequate in this case since interest lies in the local behavior in the neighborhood of zero. Choose $Z_j^* = Z_j n^{c_j}$. Now,

$$\begin{aligned}\Pr(Z_j &\leq z) = 1 - \Pr(X_{1j} > z, X_{2j} > z, ..., X_{nj} > z) \\ &= 1 - [1 - F(z)]^n .\end{aligned}$$

Then,

$$\begin{aligned}\Pr(Z_j^* &\leq z) = \Pr(Z_j \leq \frac{z}{n^{c_j}}) \\ &= 1 - \left[1 - F(\frac{z}{n^{c_j}})\right]^n \\ &= 1 - \left[1 - \frac{a_j x_j^{b_j}}{n^{b_j c_j}}\right]^n ,\end{aligned}$$

for $c_j = \frac{1}{b_j}$, $\Pr(Z_j^* \leq z) \simeq 1 - \exp\left\{-a_j x_j^{b_j}\right\}$, as $n \rightarrow \infty$, which is a Weibull distribution. Such a distribution has theoretical justification in the description of multiplicative processes in nature and economics. The exponent b_j is the reciprocal of the number of multiplicative processes. See Frisch and Sornette (1997) for theoretical justifications, and Laherrère and Sornette (1998) for empirical evidence.[5]

However, a characterization for the returns that allows for a great deal of generality can be obtained by the use of the following modified Weibull density,

$$\frac{a_j b_j \left(\left|\varkappa_j - \mu_j\right| + m\right)^{b_j - 1}}{2} \exp\left\{-a_j \left(\left|\varkappa_j - \mu_j\right| + m\right)^{b_j}\right\}, \tag{10}$$

$\varkappa_j \equiv \ln(x_j)$, $E(\varkappa_j) = \mu$, where $\varkappa_j \in \Re$ is assumed to be ergodic, and $m \equiv \left(\frac{b-1}{ab}\right)^{\frac{1}{b}}$ if $b \geq 1$, $m \equiv 0$ otherwise, is equal to the maximum of a Weibull.[6] The parameter is such that the two sided distribution is unimodal. Further generalizations can be achieved allowing for asymmetry and bimodality. In that case further care is needed in defining the shifting parameter. For simplicity, we will not pursue these generalizations. However, even if a modified Weibull is used the calculations and results are basically the same. Therefore, no generality is loss in the sequel.

[5] In the terminology of these authors a Weibull distribution is called stretched exponential.

[6] If relative prices have a Weibull distribution, then the returns (i.e. the log of relative prices) have a Fisher-Tippet density. Nevertheless, a modified Weibull is defined on the real numbers and all results presented directly apply to it. Fisher-Tippet density can be used in this framework as well. Notice that Laherrère and Sornette (1998) used a Weibull distribution in log rank plot for the returns (i.e. they divided the sample into positive and negative data) and showed a very good fit. Therefore, the use of a modified Weibull is consistent with their results.

Restricting ourself to the Weibull distribution, let $u_j \equiv 1 - \exp\left\{-a_j x_j^{b_j}\right\}$, then we can write the multivariate density of X as (15),

$$\begin{aligned} c(u_1,...,u_k) &= \sum_{v_1=0}^{n_1} ... \sum_{v_k=0}^{n_k} \tilde{\beta}\left(\frac{v_1}{n_1},...,\frac{v_k}{n_k}\right) \\ &\times \prod_{j=1}^{k} \binom{n_j}{v_j} u_j^{v_j} {}_1F_0\left(v_j - n_j; u_j\right). \end{aligned}$$

Then,

$$\begin{aligned} c(x_1,...,x_k) &= \sum_{v_1=0}^{n_1} ... \sum_{v_k=0}^{n_k} \tilde{\beta}\left(\frac{v_1}{n_1},...,\frac{v_k}{n_k}\right) \\ &\times \prod_{j=1}^{k} \binom{n_j}{v_j} {}_1F_0\left(v_j - n_j; 1 - \exp\left\{-a_j x_j^{b_j}\right\}\right) \\ &\times \left(a_j b_j x_j^{b_j - 1}\right) \exp\left\{-a_j x_j^{b_j}\right\}^{v_j+1}, \end{aligned}$$

where $\prod_{j=1}^{k}\left(a_j b_j x_j^{b_j-1}\right)\exp\left\{-a_j x_j^{b_j}\right\}$ is the Jacobian of $u_j \to x_j$. The expression above has the disadvantage of being written in powers of $1 - \exp\left\{-a_j x_j^{b_j}\right\}$. Therefore, directly from (8) we write,

$$\begin{aligned} c(x_1,...,x_k) &= \sum_{v_1=0}^{n_1-1} ... \sum_{v_k=0}^{n_k-1} \tilde{\beta}\left(\frac{v_1}{n_1},...,\frac{v_k}{n_k}\right) \\ &\times \prod_{j=1}^{k} \sum_{s_j=0}^{v_j} \binom{n_j}{v_j} \frac{(-v_j)_{s_j}}{\Gamma(s_j)} \\ &\times \left(a_j b_j x_j^{b_j-1}\right) \left(\exp\left\{-a_j x_j^{b_j}\right\}\right)^{n_j - v_j + 1 + s_j}. \end{aligned} \qquad (11)$$

However, interest does not directly lie in the joint density of the assets, but in the following random variable: $S \equiv \sum_{j=1}^{k} w_j x_j$, the portfolio return, where $w_j \in \Re$. Formally, the general solution for the density of such a random variable is

$$pdf(S) = \int \cdots \int c(x_1,...,x_k)\delta\left(S - \sum_{j=1}^{k} w_j x_j\right) dx_1 \cdots dx_k$$

where $\delta(...)$ is the Dirac delta function. We now turn to finding its characteristic function.

5.1 The Characteristic Function

Here, we find the characteristic function of $S \equiv \sum_{j=1}^{k} w_j x_j$ where $w_j \in \Re$. We shall find the Fourier transform of $\left(a_j b_j x_j^{b_j-1}\right) \exp\left\{-A_{v_j s_j} a_j x_j^{b_j}\right\}$, $A_{v_j s_j} \equiv (n_j - v_j + s_j + 1)$:

$$\phi_{x_j}(t) = \int_0^\infty \left(a_j b_j x_j^{b_j-1}\right) \exp\left\{-A_{v_j s_j} a_j x_j^{b_j}\right\} \exp\{itx_j\}\, dx_j$$

We do not consider the moment generating function because this does not exist for $b_j < 1$. For our purposes b_j will often be less than 1; see Frisch and Sornette (1997). We note that the above integral always converges. Let $y = A_{v_j s_j} a_j x_j^{b_j}$, $\frac{dy}{dx_j} = A_{v_j s_j} a_j b_j x_j^{b_j-1}$, and expanding $\exp\{itx_j\}$ around zero,

$$\begin{aligned}
\phi_{x_j}(t) &= \int_0^\infty \frac{\exp\{-y\}}{A_{v_j s_j}} \sum_{l=0}^{\infty} \left(\frac{y}{A_{v_j s_j} a_j}\right)^{\frac{l}{b_j}} \frac{(it)^l}{l!} dy \\
&= \sum_{l=0}^{\infty} \left(\frac{1}{A_{v_j s_j}}\right)^{\frac{l+b_j}{b_j}} \left(\frac{1}{a_j}\right)^{\frac{l}{b_j}} \Gamma\left(\frac{l}{b_j}+1\right) \frac{(it)^l}{l!}
\end{aligned}$$

Therefore, the characteristic function of S is

$$\begin{aligned}
\phi_S(t) &= \sum_{v_1=0}^{n_1} \cdots \sum_{v_k=0}^{n_k} \tilde{\beta}\left(\frac{v_1}{n_1}, \ldots, \frac{v_k}{n_k}\right) \\
&\quad \times \prod_{j=1}^{k} \sum_{s_j=0}^{v_j} \binom{n_j}{v_j} \frac{(-v_j)_{s_j}}{\Gamma(s_j)} \sum_{l=0}^{\infty} \left(\frac{1}{A_{v_j s_j}}\right)^{\frac{l+b_j}{b_j}} \\
&\quad \times \left(\frac{1}{a_j}\right)^{\frac{l}{b_j}} \Gamma\left(\frac{l}{b_j}+1\right) \frac{(iw_j t)^l}{l!} \\
&= \sum_{v_1=0}^{n_1} \cdots \sum_{v_k=0}^{n_k} \tilde{\beta}\left(\frac{v_1}{n_1}, \ldots, \frac{v_k}{n_k}\right) \\
&\quad \times \prod_{j=1}^{k} \sum_{l=0}^{\infty} \left[\sum_{s_j=0}^{v_j} \binom{n_j}{v_j} \frac{(-v_j)}{\Gamma(s_j)} \left(\frac{1}{A_{v_j s_j}}\right)^{\frac{l+b_j}{b_j}}\right] \\
&\quad \times \left(\frac{1}{a_j}\right)^{\frac{l}{b_j}} \Gamma\left(\frac{l}{b_j}+1\right) \frac{(iw_j t)^l}{l!}. \qquad (12)
\end{aligned}$$

The above expression can be easily differentiated with respect to t evaluated at zero, and then with respect to w_j if some kind of index notation is introduced. The procedure is tedious, but straightforward. In the appendix we calculate the first two moments.

5.2 Maximizing the Negative Exponential Expected Utility Function

The characteristic function of S allows us to find moments and minimize them with respect to the weights given a first moment constraint. However, such a procedure is highly arbitrary and irrational from the point of view of a utility maximizer. Assume an agent having the following utility function,

$$U = -\exp\{-\gamma S\}$$

that is a negative exponential utility function. This function is characterized by a constant Arrow-Pratt coefficient of risk aversion equal to γ. It could be argued that this function is superior to quadratic utility functions in representing preferences. Quadratic utility functions exhibit increasing absolute risk aversion implying satiation and moreover that risky assets are inferior goods. However, they are often used to derive tractable results whose risk characteristics are questionable.

The expected negative exponential utility function is equal to

$$EU = -E\exp\{-\gamma S\} \tag{13}$$

which is just equal to minus the Laplace transform of the wealth's probability density function. Notice that the Laplace transform is directly related to the Fourier transform by a change of variables. Let $-\gamma \equiv it$ and the Laplace transform, say $\mathfrak{L}_S(t)$, can be minimized. Since $\phi_S(t)$ is written as a linear combination of an infinite non convergent series, it is impossible to find a solution to this problem

$$\sup_{w_j,\ 1\leq j\leq k} -E\exp\left\{-\gamma\sum_{j=1}^{k} w_j x_j\right\}.$$

Moreover, $\phi_S(t)$ is characterized by oscillations of hyper-exponential order, i.e. it is not a nicely behaved function. However, in the appendix we prove the following result.

Proposition.

$$\begin{aligned}\phi_{x_j}(t) &= \int_0^\infty \left(a_j b_j x_j^{b_j-1}\right)\exp\left\{-A_{v_j s_j} a_j x_j^{b_j}\right\}\exp\{itx_j\}\,dx_j \\ &= \frac{B_j}{A_{v_j s_j}^{1/b_j}}\frac{1}{r_j} q_j^{\frac{3}{2}} r_j^{\frac{1}{2}} (2\pi)^{\frac{2-\left(q_j+r_j\right)}{2}} \Sigma_{q_j}\left\{\left(q_j^{q_j} r_j^{r_j}\right)\left(-\frac{A_{v_j s_j}^{1/b_j}}{it}\right)^{r_j}\right\},\end{aligned}$$

where

$$\Sigma_{q_j}(\frac{1}{\tau}) = \sum_{i=1}^{q_j} \tau^{-b_i}\Gamma\left[(b)' - b_i, (a)' + b_i\right]$$
$$\times_{\mathrm{rj}} F_{\mathrm{qj\text{-}1}}\left((a)' + b_i; 1 + b_i - (b)'; \frac{(-1)^{-q_j}}{\tau}\right),$$

$(a) = \left(0, \frac{1}{r_j}, \frac{2}{r_j}, ..., \frac{r_j - 1}{r_j}\right)$, $(b) = \left(\frac{1}{q_j}, \frac{2}{q_j}, ..., 1\right)$ *,* a_i *is the* i^{th} *entry in vector* (a)*, similarly for* b_i*, and the empty product is replaced by* 1.

Proof. See the Appendix ■

The above expression is convergent for $\tau \neq 0$ when $q_j > r_j$, which is the value we are interested in: $0 < b < 1$. Nevertheless, for $q_j < r_j$ an alternative form can be easily found. Therefore, a numerical package could be used to minimize

$$\phi_S(t) = \sum_{v_1=0}^{n_1} ... \sum_{v_k=0}^{n_k} \tilde{\beta}\left(\frac{v_1}{n_1}, ..., \frac{v_k}{n_k}\right) \prod_{j=1}^{k} \sum_{s_j=0}^{v_j} \binom{n_j}{v_j} \frac{(-v_j)_{s_j}}{\Gamma(s_j)} \phi_{x_j}(w_j t)$$

with respect to the weights subject to $\sum_{j=1}^{k} w_j = 1$. Notice that there is no general closed form solution to this problem. A solution can only be found numerically. To see that the solution is a minimum just recall that $\frac{\partial^2 \exp\{-\gamma w_j x_j\}}{\partial w_j^2} = (\gamma x_j)^2 \exp\{-\gamma w_j x_j\}$, which is one standard result of Laplace transforms of densities with semicompact support being convex completely monotone functions. By the properties of Laplace transform we can further argue that there exist a real $\tilde{w}_j$ such that $\lim_{w_j \to \tilde{w}_j} \frac{\partial E \exp\{-\gamma w_j x_j\}}{\partial w_j} \leq 0$ subject to $\sum_{j=1}^{k} w_j = 1$. The result holding for the multivariate case establishes the existence of such a $\tilde{w}_j$.

The procedure advocated above has firmer economic ground than the simple mean-variance optimization. Moreover, the value of γ has direct economic interpretation. We feel that the above optimization has not been feasible because of difficulties caused by multivariate non-normal densities. However, the use of the copula function and the Bernstein operator makes this possible.

If we wish to simultaneously compute value at risk or utility free optimal weights,[7] this can be done by inverting the characteristic function to find the den-

[7] We use the term utility free optimal weights to indicate the solution to the following problem,

$$\min_{w_j, 1 \leq j \leq k} \int_A pdf(S)\, dS,$$

$$s.t. \sum_{j=1}^{k} w_j = 1, \; w_j \geq 0, \; 1 \leq j \leq k$$

sity of the portfolio,

$$pdf\left(S\right)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\left\{-iSt\right\}\phi_S\left(it\right)dt.$$

The calculation is tedious but, in this case, straightforward because of the polynomial structure of the characteristic function.

5.3 A Simple Illustrative Example

Optimal portfolio weights for the simple two dimensional case are calculated from a simulated random series. The results derived through our approach and the use of normality are compared to the optimal weights derived from minimizing the empirical Laplace transform. Under suitable conditions, the empirical Laplace transform is a consistent estimator of the true Laplace transform: $\text{plim}\frac{1}{N}\sum_{t=1}^{N}\exp\left\{-\gamma S_t\right\}\rightarrow E\left(\exp\left\{-\gamma S_t\right\}\right)$. In fact, allowing for several simulated observations that are ergodic and epoch uncorrelated, this seems to be the best available criterion for assessing the performance of our methodology versus normality. Two thousands observations for two data series were generated using the following data generating process:

$$z_i=\lambda_i+t_i+f_j\left(t_j\right)$$

for $i\neq j$ where $f_l\left(...\right)$ is some function and t_i is a t-distributed random variable. The Appendix contains details about it, plus summary statistics and a few details. The above specification will guaranty non ellipticity for a non trivial choice of $f_l\left(...\right)$. The correlation for the generated series is equal to 0.2516, where the first series has lower mean with lower variance. The data are defined over the real line and can be thought as log differences in prices. Therefore, a modified Weibull is used. As mentioned before, a modified Weibull allows for much generality. In fact it can capture fat tails and approximate densities that are both strictly convex or bell shaped.

Although in portfolio problems it is appropriate to use arithmetic returns, we use geometric returns. Arithmetic returns have the irritating prospect of being bounded below. Since one of our competing models is normal returns, we have chosen to use geometric returns so that both models have the same range. The economic rationale for not bounding returns below is to jettison the free disposal

where $A=\left\{S:S\leq c\right\}$.

assumption. Thus we model an investor who considers optimizing a portfolio of forward contracts, ownership of which confers a liability on the holder.

For simplicity, the parameters were estimated in a two step procedure. The likelihood for the univariate marginals were separately maximized. Using the estimated parameters from the univariate likelihoods, the likelihood for the copula was optimized with respect to the dependence parameter; see Joe (1997) p. 299-301 for details. The best fit, under the constraint $b \geq 1$, gives $b = 1$ for both series, and $a_1 = 1.1695$ and $a_2 = 1.0839$ for the first and the second, respectively. We use $b \geq 1$ as constraint in order to simplify the calculations. However, the fit still remains good. Figure IV shows the graph of the first series; clearly, it is highly peaked. Actually more complex features emerge, but for simplicity we do not try to model them. Therefore, we just assume symmetry. The copula function is estimated parametrically and a Plackett copula is used,

$$\begin{aligned} C\left(u, v; \theta\right) \equiv & \frac{1}{2}\{(\theta - 1)\, 1 + (\theta - 1)\,(u + v) \\ & - \left[(1 + (\theta - 1)\,(u + v))^2 - 4\theta\,(\theta - 1)\, uv\right]^{\frac{1}{2}}\}. \end{aligned}$$

Figure V shows the 3 dimensional graph of the Plackett copula density. The dependence parameter estimated for this copula is θ =2.1459. Then, we approximate the above copula by a Bernstein polynomial. The order of the Bernstein polynomial is $n = 10$. Because of the low dependence of the two series, this order of polynomial is adequate. Indeed, similar result up to 3 decimal point were found using a larger polynomial.[8] The Bernstein approximation to the copula density is the following,

$$
\begin{aligned}
c(\varkappa_1, \varkappa_2) &\simeq \sum_{v_1=0}^{n} \sum_{v_2=0}^{n} c\left(\frac{v_1}{n}, \frac{v_2}{n}; \theta\right) \\
&\quad \times \prod_{j=1,2} \sum_{s_j=0}^{v_j} \binom{n}{v_j} \frac{(-v_j)_{s_j}}{\Gamma(s_j)} \\
&\quad \times \frac{a_j}{2} \left(\frac{\exp\left\{-a_j\left(\varkappa_j - \mu_j\right)\right\}}{2} \right)^{n-v_j+1+s_j}, \\
\text{if } \varkappa_j &\geq \mu_j, \\
&\simeq \sum_{v_1=0}^{n} \sum_{v_2=0}^{n} c\left(\frac{v_1}{n}, \frac{v_2}{n}; \theta\right) \\
&\quad \times \prod_{j=1,2} \sum_{s_j=0}^{v_j} \binom{n}{v_j} \frac{(v_j - n)_{s_j}}{\Gamma(s_j)} \\
&\quad \times \frac{a_j}{2} \left(\frac{\exp\left\{a_j\left(\varkappa_j - \varkappa_j\right)\right\}}{2} \right)^{v_j+1+s_j}, \\
\text{if } \varkappa_j &\leq \mu_j,
\end{aligned}
$$

where $c(u, v; \theta) = \frac{\partial^2 C(u,v;\theta)}{\partial u \partial v}$. Therefore, the problem is reduced to the evaluation of the Laplace transform of a simple exponential function. However, care has to be used in evaluating the integral since the function is not centered at the origin and it is not differentiable at its maximum. Therefore, the integral was split into two parts. Details of the calculations are available upon request. The function was maximized with respect to the weight w_1 which is associated with the first asset. The same calculation was carried out assuming normality, and optimizing the empirical Laplace transform. The coefficient chosen for the expected utility function as defined in (13) was $\gamma = 0.5$. The results are shown in Table II.[9] The results show that under the normality assumption the risk is underestimated. Normality chooses a lower weight for the variable that is safer according to the empirical Laplace transform, but which has lower mean. By construction we can expect the first series to be representative of a less risky asset. In fact, this was generated using a t-distribution with 9 degrees of freedom plus a perturbation term, versus the 7 degrees of freedom of the second one. The reader is referred to the appendix

[8] Often, the order of polynomial improves the fit in terms of curvature. In simple cases like this one, the stationary point of the function would almost be the same as we increase the order of the polynomial.

[9] For $n = 20$, w was found to be 0.5255.

for details on the data generating process. As mentioned before, the risk of ruin is much higher, but normality does not capture this. Further, not looking at all higher order moments does not allow us to effectively assess risk. Indeed the second series has a lower kurtosis, but higher variance. These two alone do not allow for rigorous probabilistic assessment of risk.

Better results could be obtained, but the purpose of this simple exercise was just to show that results might be extremely different under the two approaches and that our method gives a quite close answer to the true even when just a small order of polynomial is used. We recall that the empirical Laplace transform does not gives the exact answer, but the estimated weight converges asymptotically to the true value by the continuity of the utility function and the epoch uncorrelation of the data series. This can be easily shown by delta method. Empirical applications will be provided in the future.

6 Concluding Thoughts

A specific technique to estimate dependencies and portfolio densities was discussed. This can be achieved either through semiparametric or parametric estimation. In the former case an empirical copula would be estimated, in the latter a parametric one.

Tractable results were derived by the introduction of the Bernstein operator. Indeed, parametric families of copulae allow for very complex dependency structure. However, when the dimension increases and symmetry in the dependence structure is ruled out, then big parametric models have to be estimated. The problem with expressions having complex structure is that it might be impossible to derive any further results. On the other hand the use of Bernstein polynomials allows results to be tractable. Moreover, some parametric families of copulae have a structure that is identical to a Bernstein polynomial plus a known additional term. As described in the paper, feasible restrictions can be imposed to the parameters in order to define a copula function. This allows to define generalized families of multivariate distribution.

The convergence properties of the Bernstein approximation were investigated for the arbitrary k dimensional case. There is a vast literature on convergence of Bernstein polynomials. However we did not find specific results for the k dimensional case. For the sake of completeness, we preferred to introduce them explicitly.

As mentioned in the introduction, Bernstein polynomials are objects that arise in many branches of mathematics and probability. The representation of Bernstein polynomials through transcendental functions is convenient to derive some results. For example we easily found the covariance of marginals using Bernstein polynomials. By the definition of copula, the joint moments with respect to the marginals are of interest, but do not tell the full story about dependence.

We also provided an application to the construction of a general portfolio. Fixed Weibull marginals were chosen because of their documented adequacy to capture the distribution of financial returns. It was shown that the problem is a univariate problem in virtue of the use of the Bernstein operator. We established the existence of optimal real weights for the portfolio according to economic theory without the restrictive use of quadratic utility. Finally, we worked a simple simulation example to show that the optimal weights derived under normality and our approach can give very different answers. Indeed, the answer under Normality might be misleading, while the results for the Bernstein copula were pleasingly close to the empirical Lapace transform which is known to converge to the true value under the stationary conditions of our experiment. Moreover, we remark that only a small order of polynomial was used in order to derive our results. The intuition is that while slow to adjust, the Bernstein polynomials capture fairly well the turning points of the function.

Several issues were only briefly considered. We just mentioned their parallel with non parametric estimation as a consequence of their singular integral representation. In fact, an empirical copula could be estimated using a Bernstein polynomial as smoother. We did that in a companion paper whose results are very promising in comparison to kernel estimation. In a way, this is a new idea in density estimation in the econometrics literature. Copulae open a door for the use of multivariate semiparametric estimation, i.e. if the marginals are known, the copula estimation can be carried out nonparametrically partially avoiding the curse of dimensionality, (Sancetta and Satchell, 2001). There is a large choice of estimation methods, however the best should be chosen on the basis of its feasibility and consistency with the data.

It is also necessary to remark on the limits of our approach. The direct construction of a copula through Bernstein polynomials might be difficult requiring the estimation of many parameters if strongly dependent variables are considered. Nevertheless, if Bernstein polynomials are used as an approximation after copula estimation, it is possible to let the order of the polynomial go to a large number still allowing for tractable results. Other better approximations would require several computations as the numerical evaluation of some complicated integrals and moreover they could not provide a clear parallel with the copula function itself.

It is felt that there is scope for applications of Bernstein copula to many problems concerning the aggregation of asset returns. We leave all these problems to further research. In particular, further empirical assessment of derived results is required as the principal test of effectiveness of the suggested techniques.

References

[1] Abadir, K.M. (1999) An Introduction to Hypergeometric Functions for Economists. *Econometric Reviews* 18, 287-330.

[2] Bachelier, L. (1900) *Theorie de la Speculation*, PhD thesis. Ecole Normale Superieure, Paris. Reprinted in P. Cootner (1964) *The Random Character of Stocks*. Cambridge: MIT Press.

[3] Berens, H. & R. DeVore (1980) A Characterization of Bernstein Polynomials. *Approximation Theory III.*

[4] Butzer, P.L. (1953) On Two Dimensional Bernstein Polynomials. *Canadian Journal of Mathematics* 5, 107-113.

[5] Cheney, E.W. (1981) *An Introduction to Approximation Theory.* Providence, Rhode Island: AMS Chelsea Publishing.

[6] Cheney, E.W. and E. Ward (2000) *A Course in Approximation Theory.* Pacific Grove, CA: Brooks/Cole Publishing Company.

[7] Darsow, W.F., B. Nguyen and E.T. Olsen (1992) Copulas and Markov Processes. *Illinois Journal of Mathematics* 36, 600-642.

[8] Embrechts, P., A. McNeil & D. Straumann (1999) Correlation: Pitfalls and Alternatives. *Risk* 5, 69-71.

[9] Fama, E. F. (1965) The Behavior of Stock Prices, *Journal of Business* 37, 34-105.

[10] Feller, W. (1966) *An Introduction to Probability Theory and its Applications*, vol. II. New York: John Wiley & Sons Inc.

[11] Frisch, U. & D. Sornette (1997) Extreme Deviations and Applications. *Journal de Physique* I France 7, 1155-1171.

[12] Hagerman, R. (1978) More Evidence on the Distribution of Security Returns. *Journal of Finance* 33, 1213-1221.

[13] Joe, H. (1987) Majorization, Ramdomness and Dependence for Multivariate Distributions. *The Annals of Probability* 15, 1217-1225.

[14] Joe, H. (1993) Parametric Families of Multivariate Distributions with Given Margins. *Journal of Multivariate Analysis* 46, 262-282.

[15] Joe, H. (1997) *Multivariate Models and Dependence Concepts.* London: Chapman & Hall Ltd.

[16] Kon, S. (1984) Models of Stock Returns- a Comparison. *Journal of Finance* 39, 147-165.

[17] Kulpa, T. (1999) On Approximation of Copulas. *International Journal of Mathematics and Mathematical Science* 22, 259-269.

[18] Laherrère, J. & D. Sornette (1998) Stretched Exponential Distributions in Nature and Economy: "Fat Tails" with Characteristic Scales. *The European Physical Journal* B 2, 525-539.

[19] Li, X., P. Mikusiński and M.D. Taylor (1998) Strong Approximation of Copulas. *Journal of Mathematical Analysis and Applications* 225, 608-623.

[20] Lorentz, G.G. (1953) *Bernstein Polynomials.* Toronto: University of Toronto Press.

[21] Mandelbrot, B. (1963) The Variation of Certain Speculative Prices. *Journal of Business* 36, 394-419.

[22] Mandelbrot, B.B. (1997) *Fractals and Scaling in Finance.* New York: Springer-Verlag.

[23] Marichev, O.I. (1983) *Handbook of Integral Transforms of Higher Transcendental Functions.* Chichester, England: Ellis Horwood Ltd.

[24] Marshall, A.W. & I. Olkin (1988) Families of Multivariate Distributions. *Journal of the American Statistical Association* 83, 834-841.

[25] Nelsen, R.B. (1997) Dependence and Order in Families of Archimedean Copulas. *Journal of Multivariate Analysis* 60, 111-122.

[26] Nelsen, R.B. (1998) *An Introduction to Copulas.* Lecture Notes in Statistics 139. New York: Springer Verlag.

[27] Phillips, P.C.B. (1982) Best Uniform and Modified Padé Approximations of Probability Densities in Econometrics. In W. Hildenbrand (ed.), *Advances in Econometrics*, 123-167. Cambridge: Cambridge University Press.

[28] Phillips, P.C.B. (1983) ERA's: a New Approach to Small Sample Theory. *Econometrica* 51, 1505-1525.

[29] Sancetta, A. and S. Satchell (2001) Empirical Bernstein Copula Estimation. *Preprint.*

[30] Schurer, F. & F.W. Steutel (1977) The Degree of Approximation of Functions in $C_1[0,1]$ by Bernstein Polynomials. *Journal of Approximation Theory* 19, 69-82.

[31] Schurer, F. & F.W. Steutel (1980) Exact Degree of Approximation for Two-Dimensional Bernstein Operators on C^1. In E.W. Cheney and E. Ward (ed.), *Approximation Theory III, 823-828.* New York: Academic Press.

[32] Schweizer, B. (1991) Thirty Years of Copulas. In G. Dall'Aglio, S. Kotz and G. Salinetti (ed.),.*Advances in Probability Distributions with Given Marginals, 13-50.* Dordrecht, The Netherlands: Kluwer Academic Publishers.

[33] Silverman, B.W. (1986) *Density Estimation for Statistics and Data Analysis.* London: Chapman and Hall Ltd.

[34] Sklar, A. (1973) Random Variables, Joint Distribution Functions, and Copulas. *Kybernetika* 9, 449-460.

[35] Tobin, J. (1958) Liquidity Preference as Behavior Towards Risk. *Review of Economic Studies* 26, 65-86.

A Proof of Theorems and Proposition

Proof. Theorem 1. $i \Longrightarrow ii$: by the condition given in i.
$ii \Longrightarrow iii$:write

$$\phi_T(X) = \sum_{i=1}^{k} t_i^2 f_0 - 2\sum_{i=1}^{k} t_i f_1^i + \sum_{i=1}^{k} f_2^i,$$

where$f_j^i \equiv x_i^j$; then

$$(G_n\phi_T) = \sum_{i=1}^{k} t_i^2(G_n f_0) - 2\sum_{i=1}^{k} t_i(G_n f_1^i) + \sum_{i=1}^{k}(G_n f_2^i).$$

Now,

$$\begin{aligned}(G_n\phi_T) &= \sum_{i=1}^{k} t_i^2\left[(G_n f_0)(T) - 1\right] - 2\sum_{i=1}^{k} t_i\left[(G_n f_1^i) - t_i\right] + \sum_{i=1}^{k}\left[(G_n f_2^i) - t_i^2\right] \\ &\leq \sum_{i=1}^{k} t_i^2\left\|(G_n f_0)(T) - 1\right\| + 2\sum_{i=1}^{k}|t_i|\left\|(G_n f_1^i) - t_i\right\| + \sum_{i=1}^{k}\left\|(G_n f_2^i) - t_i^2\right\|.\end{aligned}$$

Therefore, by the conditions stated in ii, $(G_n\phi_T)(T) \to 0$ uniformly.
$iii \Longrightarrow i$: by the continuity of f, $\exists\delta > 0$ such that

$$\|T - X\| < \delta \Longrightarrow |f(T) - f(X)| < \epsilon,$$

for any $\epsilon > 0$. Let T be an arbitrary fixed point in A.

$$\|T - X\| \geq \delta \Longrightarrow |f(T) - f(X)| \leq 2\|f\| \leq 2\|f\|\|T - X\|^2/\delta^2,$$

which for any X in A satisfies the following inequality:

$$|f(T) - f(X)| < \epsilon + \alpha\phi_T(X),$$

where $\alpha \equiv 2\|f\|/\delta^2$. Multiply both sides by $f_0 \equiv 1$, apply the linear operator G_n and then let $X \to T$. Hence,

$$|f(T)(G_n f_0) - (G_n f)(T)| < \epsilon(G_n f_0) + \alpha(G_n\phi_T)(T),$$

$(G_n\phi_T)(T) \to 0$ uniformly, by the conditions stated in iii; so the following inequality is satisfied:

$$|f(T) - (G_n f)(T)| < \epsilon,$$

for any arbitrary T, which proves the last part of the theorem. ■

Proof. Theorem 2. Without loss of generality, attention is restricted to the k^{th} dimensional unit hyper-cube. Consider the following k dimensional Bernstein Polynomial:

$$(B_n f)(X) \equiv \sum_{v_1=0}^{n_1}\cdots\sum_{v_K=0}^{n_k} f\left(\frac{v_1}{n_1},...,\frac{v_k}{n_k}\right)P_{v_1,n_1}(x_1)\cdots P_{v_K,n_K}(x_k),$$

where $P_{v_j,n_j}(x_j) \equiv \binom{n_j}{v_j} x_j^{v_j}(1-x_j)^{n_j-v_j}$. By the theorem on *Linear Monotone Operators* it is sufficient to show uniform convergence to f for the following cases: $f(X) = 1, x_j, x_j^2$, $1 \leq j \leq k$.

$$(B_n 1)(X) \equiv \sum_{v_1}^{n_1} \ldots \sum_{v_K}^{n_K} P_{v_1,n_1}(x_1) \cdots P_{v_K,n_K}(x_k) = 1,$$

by the binomial theorem.

$$\begin{aligned}
(B_n x_j)(X) &\equiv \sum_{v_1}^{n_1} \ldots \sum_{v_j}^{n_j} \ldots \sum_{v_K}^{n_K} \left(\frac{v_j}{n_j}\right) P_{v_1,n_1}(x_1) \cdots P_{v_j,n_j}(x_j) \cdots P_{v_K,n_K}(x_k) \\
&= \sum_{v_j}^{n_j} \left(\frac{v_j}{n_j}\right) \binom{n_j}{v_j} x_j^{v_j}(1-x_j)^{n_j-v_j} \\
&= x_j \sum_{v_j=1}^{n_j} \frac{n_j - 1!}{(n_j - v_j)!(v_j - 1)!} x_j^{v_j - 1}(1-x_j)^{n_j - v_j} \\
&= x_j \sum_{v_j=0}^{n_j-1} \binom{n_j - 1}{v_j} x_j^{v_j-1}(1-x_j)^{n_j-v_j-1} \\
&= x_j[x_j + (1-x_j)]^{n-1} = x_j,
\end{aligned}$$

where the first equality follows by the binomial theorem.

$$\begin{aligned}
(B_n x_j^2)(X) &\equiv \sum_{v_1}^{n_1} \ldots \sum_{v_j}^{n_j} \ldots \sum_{v_K}^{n_K} \left(\frac{v_j}{n_j}\right)^2 P_{v_1,n_1}(x_1) \cdots P_{v_j,n_j}(x_j) \cdots P_{v_K,n_K}(x_k) \\
&= \sum_{v_j}^{n_j} \left(\frac{v_j}{n_j}\right)^2 \binom{n_j}{v_j} x_j^{v_j}(1-x_j)^{n_j-v_j} \\
&= \sum_{v_j=1}^{n_j} \left(\frac{v_j}{n_j}\right) \binom{n_j - 1}{v_j - 1} x_j^{v_j}(1-x_j)^{n_j-v_j} \\
&= \frac{n_j - 1}{n_j} \sum_{v_j=1}^{n_j} \left(\frac{v_j - 1}{n_j - 1}\right) \binom{n_j - 1}{v_j - 1} x_j^{v_j}(1-x_j)^{n_j-v_j} \\
&\quad + \frac{1}{n_j} \sum_{v_j=1}^{n_j} \binom{n_j - 1}{v_j - 1} x_j^{v_j}(1-x_j)^{n_j-v_j} \\
&= \frac{n_j - 1}{n_j} x_j^2 + \frac{1}{n_j} x_j \to x_j^2.
\end{aligned}$$

■

Proof. Proposition. Let $-itx_j = y_j \Longrightarrow x_j = \frac{iy_j}{t}$, $\frac{dx_j}{dy_j} = \left|\frac{1}{-it}\right| = \left|\frac{i}{t}\right|$, $v = \frac{t}{iA_{v_js_j}^{1/b_j}}$, then

$$\begin{aligned}\phi_{x_j}(v) &= \int_0^\infty \frac{B_j}{A_{v_js_j}^{1/b_j}} \left(\frac{iy_jA_{v_js_j}^{1/b_j}}{t}\right)^{b_j} \exp\left\{-\left(\frac{iy_jA_{v_js_j}^{1/b_j}}{t}\right)^{b_j}\right\} \exp\{-y_j\} \frac{idy_j}{iy_j} \\ & \int_0^\infty \frac{B_j}{A_{v_js_j}^{1/b_j}} \left(\frac{v}{y_j}\right)^{-b_j} \exp\left\{-\left(\frac{v}{y_j}\right)^{-b_j}\right\} \exp\{-y_j\} \frac{dy_j}{y_j} \\ &= \frac{B_j}{A_{v_js_j}^{1/b_j}} \mathfrak{M}^{-1}\{\mathfrak{H}_1^*(l)\mathfrak{H}_2^*(l); v\},\end{aligned}$$

which is recognized as a Mellin convolution type of integral, where $\mathfrak{H}_1^*(s)$ and $\mathfrak{H}_2^*(s)$ are respectively the Mellin transforms of the functions with arguments $\frac{v}{y_j}$ and y_j. Make the following change of variables, $b_j{=}\frac{r_j}{q_j}$, $\lambda = v^{r_j}$, $z_j = y_j^{r_j}$, $\frac{dz_j}{dy_j} = r_jy_j^{r_j-1}$,

$$\begin{aligned}\phi_{x_j}\left(\lambda^{\frac{1}{r_j}}\right) &= \frac{B_j}{A_{v_js_j}^{1/b_j}} \frac{1}{r_j} \int_0^\infty \left(\frac{\lambda}{z_j}\right)^{-\frac{1}{q_j}} \exp\left\{-\left(\frac{\lambda}{z_j}\right)^{-\frac{1}{q_j}}\right\} \exp\left\{-z_j^{\frac{1}{r_j}}\right\} \frac{dz_j}{z_j} \\ &= \frac{B_j}{A_{v_js_j}^{1/b_j}} \frac{1}{r_j} \mathfrak{M}^{-1}\left\{\mathfrak{H}^*(l); \lambda^{\frac{1}{r_j}}\right\},\end{aligned}$$

Now,

$$\mathfrak{H}_1^*(l) = \int_0^\infty \left(\frac{\lambda}{z_j}\right)^{-\frac{1}{q_j}} \exp\left\{-\left(\frac{\lambda}{z_j}\right)^{-\frac{1}{q_j}}\right\} \left(\frac{\lambda}{z_j}\right)^{l-1} d\left(\frac{\lambda}{z_j}\right),$$

let $\tau = \left(\frac{\lambda}{z_j}\right)^{-\frac{1}{q_j}} \Longrightarrow \left(\frac{\lambda}{z_j}\right) = \tau^{-q_j}$, $\frac{d\tau}{d\left(\frac{\lambda}{z_j}\right)} = \left|-\frac{1}{q_j}\left(\frac{\lambda}{z_j}\right)^{-\frac{1}{q_j}-1}\right|$, therefore

$$\mathfrak{H}_1^*(l) = q_j \int_0^\infty \exp\{-\tau\} \tau^{-q_jl} d\tau = \Gamma(-q_jl+1).$$

Now,

$$\mathfrak{H}_2^*(l) = \int_0^\infty \exp\left\{-z_j^{\frac{1}{r_j}}\right\} z_j^{l-1} dz_j.$$

Let $\tau = z_j^{\frac{1}{r_j}} \Longrightarrow z_j = \tau^{r_j}$, $\frac{d\tau}{dz_j} = \frac{1}{r_j} z_j^{\frac{1}{r_j}-1}$, therefore

$$\mathfrak{H}_2^*(l) = r_j \int_0^\infty \exp\{-\tau\} \tau^{r_jl-1} d\tau = \Gamma(r_jl).$$

Then

$$\mathfrak{H}^*(l) = q_j r_j \Gamma(1 - q_j l)\,\Gamma(r_j l),$$

and using the Gauss multiplication formula for $\Gamma(1 - q_j l)$ and $\Gamma(r_j l)$ (see Marichev (1983), p. 110) we can write,

$$\begin{aligned}\mathfrak{H}^*(l) &= q_j^{\frac{3}{2}} r_j^{\frac{1}{2}} (2\pi)^{\frac{2-(q_j+r_j)}{2}} \left(q_j^{q_j} r_j^{r_j}\right)^s \\ &\times\Gamma\left[\frac{1}{q_j} - l, \frac{2}{q_j} - l, ..., 1 - l\right]\Gamma\left[l, \frac{1}{r_j} + l, \frac{2}{r_j} + l, ..., \frac{r_j - 1}{r_j} + l\right].\end{aligned}$$

We need to find the following inverse Mellin transform,

$$\mathfrak{H}(l) = \int_{\gamma - i\infty}^{\gamma + i\infty} \mathfrak{H}^*(l)\lambda^{-l} dl.$$

Notice that if $Q(\lambda)$ is the inverse Mellin transform of $\Gamma[...]$, then $Q(\frac{\lambda}{N})$ is the inverse Mellin transform of $N^s\Gamma[...]$. Now choose $\gamma \in \left(0, \frac{1}{q_j}\right)$ and deform the path of integration of the above integral to be the close contour comprising of the line γ parallel to the imaginary axis and a loop surrounding either all left or right singularities of the integrand. The integral along the arc will tend to zero as its radius goes to infinity; therefore,

$$\int_{\gamma - i\infty}^{\gamma + i\infty} \mathfrak{H}^*(l)\lambda^{-l} dl = \oint_c \mathfrak{H}^*(l)\lambda^{-l} dl$$

It follows that the value of the above integral is equal to the sum of the residues surrounded by the loop. To avoid convergence issues, we just apply Slater's theorem (see Marichev (1983)). By Slater's theorem,

$$Q(\tau) = \int_0^\infty \Sigma_{q_j}(\frac{1}{\tau})\tau^{l-1} d\tau$$

$$\begin{aligned}\Sigma_{q_j}(\frac{1}{\tau}) &= \sum_{i=1}^{q_j} \tau^{-b_i}\Gamma\left[(b)' - b_i, (a)' + b_i\right] \\ &\times {}_{\mathrm{r_j}}F_{\mathrm{q_j\text{-}1}}\left((a)' + b_i; 1 + b_i - (b)'; \frac{(-1)^{-q_j}}{\tau}\right),\end{aligned}$$

where in our case $(a) = \left(0, \frac{1}{r_j}, \frac{2}{r_j}, ..., \frac{r_j - 1}{r_j}\right)$, $(b) = \left(\frac{1}{q_j}, \frac{2}{q_j}, ..., 1\right)$, and $\tau = \frac{\lambda^{\frac{1}{r_j}}}{\left(q_j^{q_j} r_j^{r_j}\right)} = \frac{1}{\left(q_j^{q_j} r_j^{r_j}\right)} \frac{t}{iA_{v_j s_j}^{1/b_j}}$. ■

B Transcendental Functions and the Bernstein Operator

An alternative representation of the Bernstein copula is through incomplete beta functions,

$$
\begin{aligned}
B_x(p,q) &= \int_0^x t^{p-1}(1-t)^{q-1}dt \\
&= \frac{x^p}{p} {}_2F_1(p, 1-q; p+1; x)
\end{aligned}
$$

where $2F1$ is Gauss' hypergeometric series, or simply the hypergeometric function, i.e. the probability generating function of an hypergeometric distribution; see Abadir (1999) for an introduction for economists and references therein. Here we recall that

$$
pFq(a_1, ..., a_p; c_1, ..., c_q; x) = \sum_{i=0}^{\infty} \frac{(a_1)_i \cdots (a_p)_i}{(c_1)_i \cdots (c_q)_i} \frac{x^i}{i!},
$$

where $(a)_i \equiv \frac{\Gamma(a+i)}{\Gamma(a)}$ and similarly for $(c)_i$. Note that throughout the paper the following convention is used, $(0)_0 \equiv 0$.

Integrating (8) we can write the Bernstein copula as

$$
\begin{aligned}
C(u_1, ..., u_k) &= u_1 \cdots u_k + \sum_{v_1=0}^{n_1} ... \sum_{v_k=0}^{n_k} \beta\left(\frac{v_1}{n_1}, ..., \frac{v_k}{n_k}\right) \\
&\quad \times \prod_{j=1}^{k} \binom{n_j}{v_j} B_{u_j}(v_j+1, n_j - v_j + 1) \\
&= u_1 \cdots u_k + \sum_{v_1=0}^{n_1} ... \sum_{v_k=0}^{n_k} \beta\left(\frac{v_1}{n_1}, ..., \frac{v_k}{n_k}\right) \\
&\quad \times \prod_{j=1}^{k} \binom{n_j}{v_j} \frac{u_j^{v_j+1}}{v_j+1} {}_2F_1(v_j+1, v_j - n_j; v_j+2; u_j). \qquad (14)
\end{aligned}
$$

It should be noticed that for $v_j \leq n_j$, which is the case of the Bernstein copula, ${}_2F_1(v_j+1, v_j-n_j; v_j+2; u_j)$ is a polynomial of order $n_j - v_j$ The value of this representation through transcendental functions should not be underestimated. In fact, this always allows one to directly write the Bernstein copula in powers of the marginal distributions only. Therefore, it is possible to work with distributions that can be regarded as multivariate distribution functions of *iid* random variables. Moreover, in this case the Gauss hypergeometric series is a finite polynomial, then any convergence issue is avoided. A similar result can be obtained by writing $(1-u)^v = \sum_{s=0}^{v}(-1)^s\binom{v}{s}u^s$ which is the familiar binomial expansion, and substituting it in the copula distribution. Remember that when we differentiate we loose one factor in the summation.

Differentiating (14) we can find a useful expression for the Bernstein density copula

$$\begin{aligned}\frac{\partial^k C(u_1,...,u_k)}{\partial u_1 \partial u_2 \cdots \partial u_k} &= 1+\sum_{v_1=0}^{n_1}...\sum_{v_k=0}^{n_k}\beta\left(\frac{v_1}{n_1},...,\frac{v_k}{n_k}\right) \\ &\quad \times \prod_{j=1}^{k}\binom{n_j}{v_j}\left(\frac{\partial}{\partial u_j}\right)\frac{u_j^{v_j+1}}{v_j+1}{}_2F_1(v_j+1, v_j-n_j; v_j+2; u_j) \\ &= 1+\sum_{v_1=0}^{n_1}...\sum_{v_k=0}^{n_k}\beta\left(\frac{v_1}{n_1},...,\frac{v_k}{n_k}\right) \\ &\quad \times \prod_{j=1}^{k}\binom{n_j}{v_j}u_j^{v_j}{}_1F_0(v_j-n_j; u_j). \end{aligned} \tag{15}$$

C The First Two Moments of the Portfolio

The first two moments can be easily calculated from (12). Remember that $E(S^n)$ is the coefficient corresponding to t^n. Therefore, instead of multiplying all polynomials in (12) and find the corresponding coefficient, it is easier to take a subset of the above polynomials, i.e. the polynomials of order n at most and multiply them. This considerably reduces the calculations. Moreover, we can write

$$\begin{aligned}\prod_{j=1}^{k}\left[\sum_{l=0}^{n}g_{l_j}\frac{(iw_jt)^l}{l!}\right] &\equiv \prod_{j=1}^{k}\sum_{l=0}^{n}\left[\sum_{s_j=0}^{v_j}\binom{n_j}{v_j}(-v_j)_{s_j}\left(\frac{1}{A_{v_js_j}}\right)^{\frac{l+b_j}{b_j}}\right] \\ &\quad \times\left(\frac{1}{a_j}\right)^{\frac{l}{b_j}}\Gamma\left(\frac{l}{b_j}+1\right)\frac{(iw_jt)^l}{l!},\end{aligned}$$

where the summation goes to n and not to infinity, i.e. it is a subset of the equivalent term in (12). Since the dummy suffix l in g_{l_j} uniquely identifies the order of t, we can say that the coefficient in g_{l_j} of t^n is the sum of all partitions of $\frac{g_{l_1}\cdots g_{l_k}}{l_1!\cdots l_k!}$ such that $\sum_{j}^{k} l_j = n$.

To give a manageable example we only consider the first two moments. Usually interest lies just on these two. The first moment is given by the different ways of choosing one object out of k, i.e. k. This means that the first moment is given by the sum of the sum of the following k terms, $g_{1_1} g_{0_2} \cdots g_{0_k} + g_{0_1} g_{1_2} \cdots g_{0_k} + ... + g_{0_1} g_{0_2} \cdots g_{1_k}$:

$$\begin{aligned} E(S) &= \sum_{v_1=0}^{n_1} ... \sum_{v_k=0}^{n_k} \tilde{\beta}\left(\frac{v_1}{n_1}, ..., \frac{v_k}{n_k}\right) \left(g_{1_1} g_{0_2} \cdots g_{0_k}(w_1) + g_{0_1} g_{1_2} \cdots g_{0_k}(w_2) + ...\right. \\ &\left. + g_{0_1} g_{0_2} \cdots g_{1_k}(w_k)\right), \end{aligned}$$

where g_{l_j} is as defined above. The second moment is given by the sum of the sum of the following $k + \binom{k}{2}$ terms, $\frac{\left(g_{2_1} g_{0_2} \cdots g_{0_k} + g_{0_1} g_{2_2} \cdots g_{0_k} + ... + g_{0_1} g_{0_2} \cdots g_{2_k}\right)}{2} + (g_{1_1} g_{1_2} \cdots g_{0_k} + g_{1_1} g_{0_2} g_{1_3} \cdots g_{0_k} + ... + g_{0_1} g_{0_2} \cdots g_{1_{k-1}} g_{1_k})$:

$$\begin{aligned} E\left(S^2\right) &= \frac{1}{2} \sum_{v_1=0}^{n_1} ... \sum_{v_k=0}^{n_k} \tilde{\beta}\left(\frac{v_1}{n_1}, ..., \frac{v_k}{n_k}\right) \\ &\times \left(g_{2_1} g_{0_2} \cdots g_{0_k}(w_1)^2 + g_{0_1} g_{2_2} \cdots g_{0_k}(w_2)^2 + ...\right. \\ &\left. + g_{0_1} g_{0_2} \cdots g_{2_k}(w_k)^2\right) \\ &+ \sum_{v_1=0}^{n_1} ... \sum_{v_k=0}^{n_k} \tilde{\beta}\left(\frac{v_1}{n_1}, ..., \frac{v_k}{n_k}\right) \\ &\times \left(g_{1_1} g_{1_2} \cdots g_{0_k}(w_1 w_2) + g_{1_1} g_{0_2} g_{1_3} \cdots g_{0_k}(w_1 w_3) + ...\right. \\ &\left. + g_{0_1} g_{0_2} \cdots g_{1_{k-1}} g_{1_k}(w_{k-1} w_k)\right). \end{aligned}$$

For higher order the notation is lengthy and use of tensor notation is necessary. However, these first two moments are sufficient if mean variance optimization is performed.

D Estimation results and Graphs

The parameters for the data generating process were chosen as follows. Let $z_i = \lambda_i + t_i + f_j(t_j)$. Then, $t_1 \sim t(9)$, $t_2 \sim t(7)$, $\lambda_1 = 50/N$, $\lambda_2 = 70/N$, $f_2(t_2) = 1/10(|t_2| + 2t_2)$, $f_1(t_1) = 1/30\left(|t_1|^{\frac{1}{3}} + \frac{3}{2}t_1\right)$, where $t(n)$ is a randomly generated variable from a t-distribution with n degrees of freedom.

In order to give an idea of the distribution of the data generated, Table III contains descriptive statistics. Parameter estimates for the marginal distributions and the copula function are in Table IV.

Table I. Spearman's rho for different values of the dependance parameter θ

θ	0	.14	.31	.51	.76	1.06	1.51	2.14	3.19	5.56	∞
$\rho_S(KS)$	0	.1	.2	.3	.4	.5	.6	.7	.8	.9	1
$\rho_S(B_{10})$	0	.79	.16	.24	.32	.4	.48	.57	.65	.73	*
$\rho_S(B_{30})$	0	.09	.19	.28	.37	.46	.56	.65	.75	.84	*
$\rho_S(B_{50})$	0	.09	.19	.29	.38	.48	.58	.67	.77	.86	*

Table II. Optimal portfolio weight for a negative exponential utility function

	Empirical Laplace	Multivariate Normal	Plackett Copula (Bernstein Approximation)
w_1	0.5157	0.4485	0.5258

Table III. Descriptive Statistics

	Min	1st Qu.	Mean	Median	3rd Qu	Max
Z_1	-5.282	-0.607	0.061	0.057	0.722	6.543
Z_2	-5.323	-0.621	0.104	0.131	0.815	6.419

Variance	Kurtosis	N
1.286	4.8506	2000
1.486	4.675027	2000

	Variance	Kurtosis	N
Z_1	1.286	4.8506	2000
Z_2	1.486	4.675027	2000

Table IV. Parameter Estimates

a_1	b_1	a_2	b_2	θ
1.1695	1	1.0839	1	2.1459

Figure I. Kimeldorf and Sampson (KS) copula density (θ=1.06)

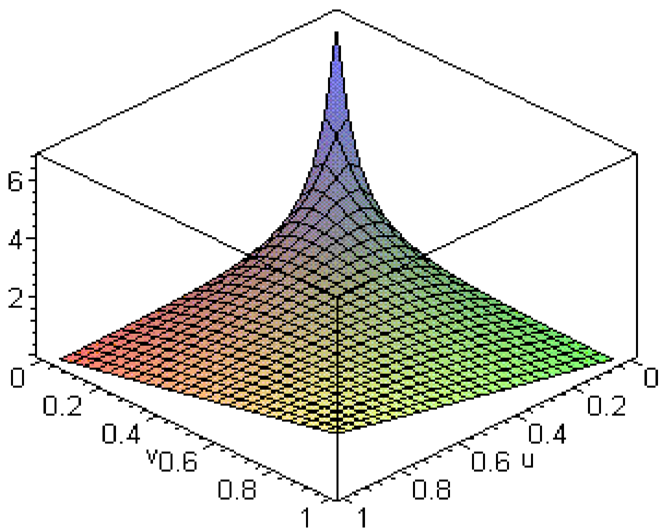

Figure II. KS copula (θ=1.06), contour plot

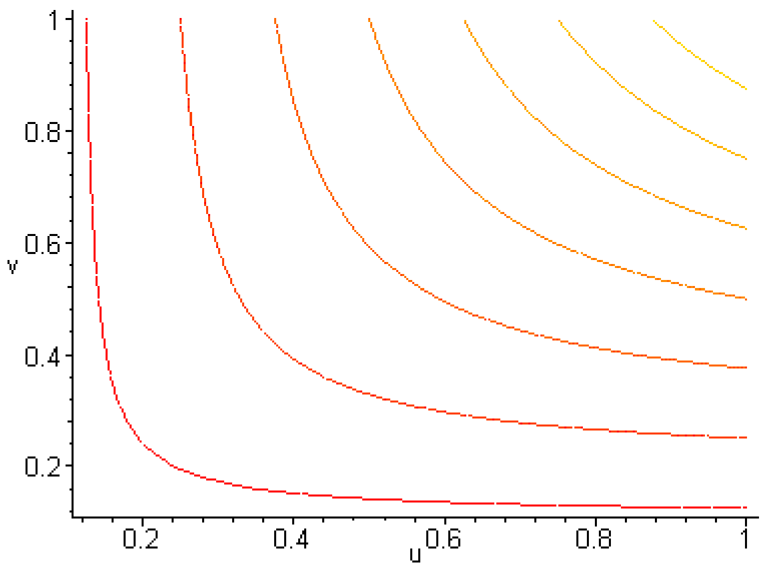

Figure III. Bernstein approximation to the KS copula (θ=1.06, n=30), contour plot

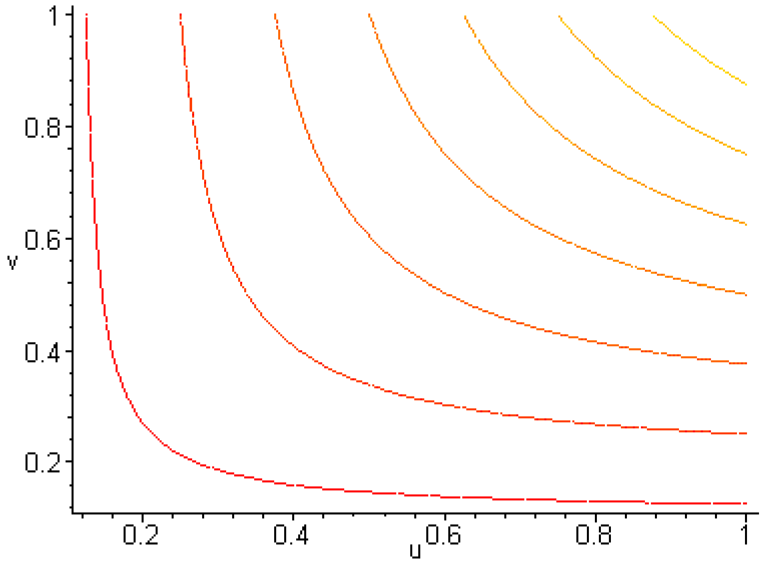

Figure IV. First series of simulated random data

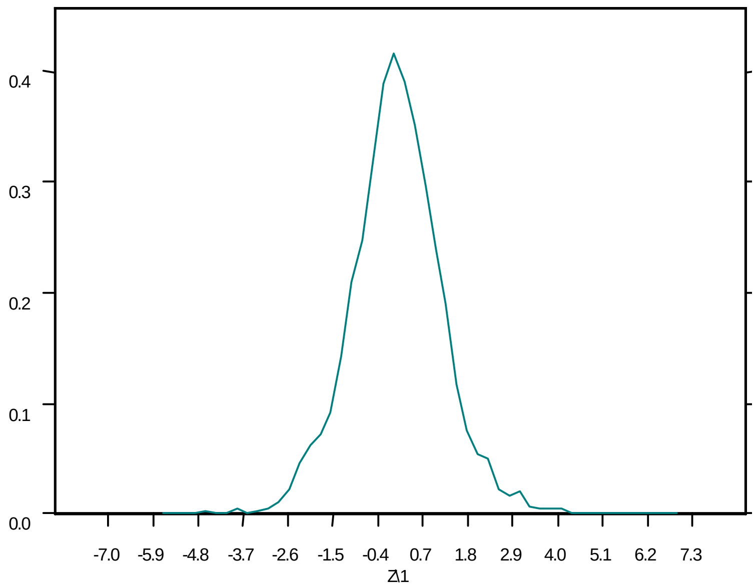

Figure V. Plackett copula (θ=2.14)