in: *Decision Analysis*, published online before print November 25, DOI:10.1287/deca.1090.0162

Generalized Diagonal Band Copulae with Two-Sided Generating Densities

Samuel Kotz+ and J. René van Dorp*

Department of Engineering Management and Systems Engineering,

The George Washington University, Washington DC, 20052.

Abstract: Copulae are joint continuous distributions with uniform marginals and have been proposed to capture probabilistic dependence between random variables. Maximum entropy copulae introduced by Bedford and Meeuwissen (1997) provide experts the option of making minimally informative assumptions given a degree of dependence constraint between two random variables. Unfortunately, their distributions functions are not available in a closed form, and their application requires the use of numerical methods. In this paper we shall study a sub-family of generalized diagonal band (GDB) copulae, separately introduced by Ferguson (1995) and Bojarski (2001). Similar to Archimedean copulae, GDB copulae construction require a generator function. Bojarski's GDB copula generator functions are symmetric probability density functions. In this paper, members of a symmetric two-sided framework of distributions introduced by Van Dorp and Kotz (2003) shall be considered. This flexible set-up allows for derivations of GDB copula properties resulting in novel convenient expressions. A straightforward elicitation procedure for the GDB copula dependence parameter is proposed. Closed form expressions for specific examples in the sub-family of GDB copulae are presented, which enhance their transparency and facilitate their application. These examples closely approximate the entropy of maximum entropy copulae. Application of GDB copulae is illustrated via a value of information decision analysis example.

Keywords: Copula; probability distribution; probability assessment; expert judgement; elicitation; value of information.

+E-mail: kotz@gwu.edu, Tel.: +1 202 994 7187; Fax:+1 202 994 0245.
* Corresponding author. E-mail: dorpjr@gwu.edu, Tel.: +1 202 994 6638; Fax:+1 202 994 0245.

1. INTRODUCTION

Copulae are joint distributions with uniform marginals and were initially discovered by Sklar (1959), who was interested in pure mathematical aspects. When considering random variables X' and Y' with known continuous marginal cumulative distribution functions $G(\cdot)$ and $H(\cdot)$, any bivariate joint distribution of (X', Y') can be transformed to a bivariate copula $(X, Y) = (G(X'), H(Y'))$. The mapping $X' \to X = G(X')$, where X is uniformly distributed on the unit interval is commonly called the probability integral transformation (e.g. Nelsen (1999)). As such, many authors, mostly indirectly, studied copulae by considering bivariate distributions with known continuous marginals. Gaussian copulae and Student t-copulae, on the other hand, have been studied explicitly and are prime examples of this construction procedure. Both belong to the larger elliptical family of copulae which are characterized by elliptically contoured distributions (see, e.g. Clemen and Reilly (1999) and Lewandowski (2008)). Genest and Mackay (1986) and Nelsen (1999) studied an elegant framework for modeling a class of copulae in a direct manner, known as Archimedean copulae. Archimedean copulae are popular for their ease of construction via an algebraic method involving a convex decreasing function $\varphi : (0,1] \to [0, \infty)$, called a *generator*, such that $\varphi(1) = 0$. They possess joint cumulative distribution function (cdf)

$$C\{x, y|\varphi(\cdot)\} = \begin{cases} \varphi^{-1}\{\varphi(x) + \varphi(y)\}, & \varphi(x) + \varphi(y) \leq 0, \\ 0, & \text{elsewhere,} \end{cases} \tag{1}$$

and joint probability density function (pdf)

$$c\{x, y|\varphi(\cdot)\} = -\frac{\varphi''\{C(x,y)\}\varphi'(x)\varphi'(y)}{[\varphi'\{C(x,y)\}]^3}, \tag{2}$$

where $\varphi''(x) = d^2Q(x)/dx^2$ and $\varphi'(x) = dQ(x)/dx$. Another procedure for constructing copulae uses the geometric method. Nelsen (1999) discusses a variety of methods utilizing some information of a geometric nature.

In recent years, the fields of finance and insurance have experienced such a burst of applications utilizing the copulae approach that references are too numerous to be cited individually. We shall suffice with providing a sampling for insurance (see, e.g., Frees et al. (1996, 1998, 2005), Denuit et al. (2005), Embrechts (2009)) and finance (see, e.g. Härdle et al. (2002), Cherubini et al. (2004), McNeil et al. (2005), He and Gong (2009)). For examples of applications in other areas where the copulae approach was also suggested for statistical dependence modeling, see, e.g., Clemen and Reilly (1999), Van Dorp and Duffey (1999), Yi and Bier (1998), Kallen and Cooke (2002), De Michele et al. (2007), Genest and Favre (2007), Norris et al. (2008). Abbas and Howard (2005) and Abbas (2009) considered extensions of copula functions and applied them in the context of multi attribute utility theory. A particular advantage of the copula approach for statistical dependence modeling is that it utilizes a decomposition principle by separately describing uncertainty phenomena via marginal distributions and the dependence between these phenomena via a copula. Evidently, this has in part resulted in the widespread applications of copulae constructs of this type by "financial quants". Unfortunately, we have not seen a similar burst of applications in areas other than finance or insurance.

In this paper, we shall further investigate the bivariate family of generalized diagonal band (GDB) copulae separately introduced by Ferguson (1995) and Bojarski (2001). Their copula construction method is geometrically motivated, but enjoys an ease similar to that of the algebraic generator construction method of the Archimedean copulae. The diagonal band copula introduced by Cooke and Waij (1986) is the founding member of this copula family. It was originally geometrically constructed by firstly uniformly distributing a total probability mass of 1 over a band of varying width around the unit diagonal defined by four (x, y) points

$$(0, \theta - 1), (0, 1 - \theta),\ (1, \theta) \text{ and } (1, 2 - \theta), \tag{3}$$

depicted in Figure 1A. The density function in the y direction given a value for x in that case follows as a uniform distribution with bounds

$$[x - (1 - \theta), x + (1 - \theta)] \tag{4}$$

and interval width $2(1 - \theta)$. Next, by "folding back" the masses in the triangles outside the unit-square, indicated in Figure 1A, on top of the unit-square, the $DB(\theta)$ copula with dependence parameter $\theta \in [0, 1]$ was constructed. Its area with strictly positive density values (commonly referred to as density support) is limited to the gray areas $DB_i,\ i = 1, 2, 3$ in Figure 1A. Its density values are:

$$\begin{cases} \frac{1}{1-\theta}, & (x, y) \in\ DB_1 \cup DB_3, \\ \frac{1}{2(1-\theta)}, & (x, y) \in DB_2, \\ 0, & \text{elsewhere.} \end{cases} \tag{5}$$

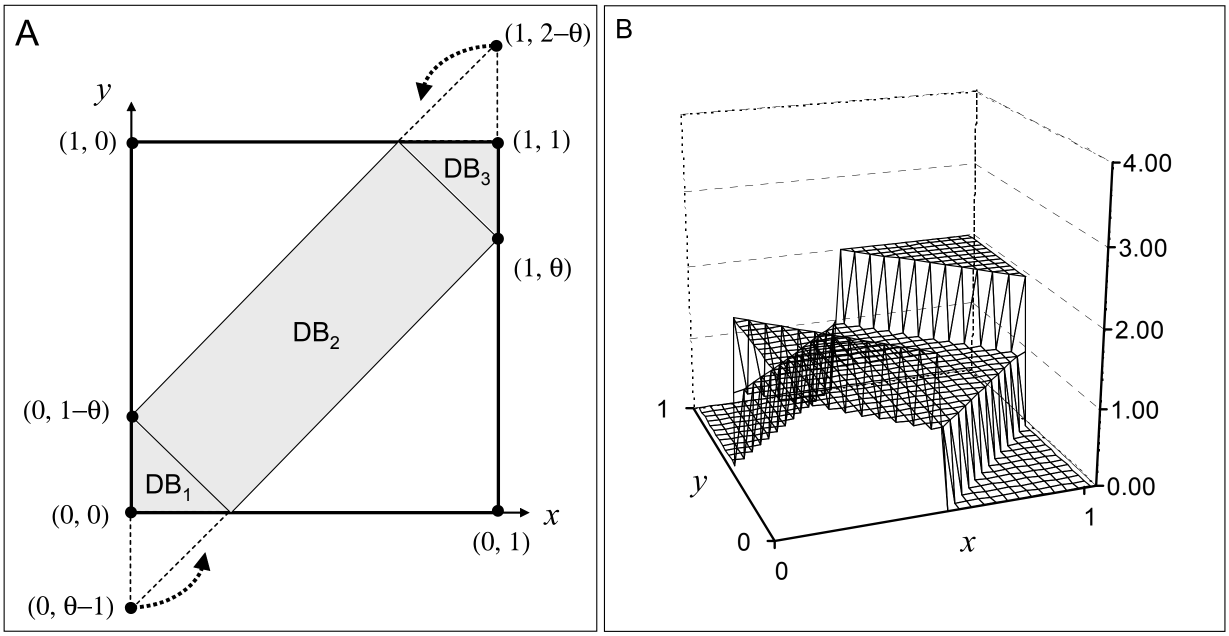

Figure 1: A: Gray area support of a $DB(\theta)$ copula comprised of sub-areas $DB_i, i = 1, 2, 3$; B: Example of a $DB(0.5)$ copula.

Observe that the probability mass in DB_2 ($DB_1 \cup DB_3$) equals the reciprocal of (twice) the width of the interval (4) (due to the "folding" operation depicted in Figure 1A). For $\theta = 0$, the area DB_2 in Figure 1A vanishes and (5) reduces to a density value of 1 across the complete unit-square implying independence between X and Y when $(X, Y) \sim\ DB(0)$. For $\theta = 1$, the interval (4) reduces to the singleton $\{x\}$ and the support of a $DB(1)$ copula in Figure 1A reduces to the positive unit-square diagonal implying complete dependence, i.e. $Y = X$. Figure 1B displays an example of a $DB(\theta)$ copula density with $\theta = \frac{1}{2}$.

Bojarski (2001) generalized the $DB(\theta)$ copula to a wider and more flexible family of copulae with the same diagonal band support in Figure 1A. Similar to Archimedean copulae, their construction requires a generator function $f(\cdot|\theta)$. However, here the generator function $f(\cdot|\theta)$ is a symmetric pdf with support $[-(1-\theta),(1-\theta)]$ (substitute $x=0$ in (4)). To retain the sampling efficiency of the original $DB(\theta)$ copula, a closed form and preferably simple expression for the inverse cdf of $f(\cdot|\theta)$ would be desirable. Bojarski (2001) considered symmetric beta distributions which do not meet that requirement. Lewandowski (2005) showed that Bojarski's GDB copulae are equivalent to the family of copulae introduced by Ferguson (1995) with density

$$c(x,y)=\frac{1}{2}\{g(|x-y|+g(1-|1-x-y|)\}, \tag{6}$$

where $g(z)$ is a generating pdf with support $[0,1]$. Ferguson (1995) demonstrated that copulae of the form (6) arise as a continuous mixture of bivariate uniform densities on rectangles with boundaries $(0,z)$, $(z,0)$, $(1,1-z)$ and $(1-z,1)$ with mixture density $g(z)$. To facilitate the application of the family of GDB copulae, expressions for distribution functions and properties of several instances shall be derived herein in a closed form. We hope that their ease of use combined with their geometric motivation facilitates a larger penetration of copula techniques in other application domains, such as, e.g., decision analysis and uncertainty analysis, to a level that triangular distributions (which too were geometrically motivated) have facilitated and contributed to a growth of uncertainty analysis applications.

Shortly after Bojarksi's (2001) generalization of DB copulae, Van Dorp and Kotz (2003) introduced a flexible Two-Sided (TS) framework of bounded distributions that too uses the generating pdf $p(z)$ concept to define a sub-family of distributions within it. Symmetric members within the TS framework of distributions seem to provide a natural candidate for the generator of Bojarski's (2001) GDB copulae. In Section 2, we shall introduce the TS framework of symmetric distributions with generating pdf $p(z)$ and express GDB copula properties in terms of this pdf $p(z)$. GDB copula density construction with a TS generating pdf following Bojarski's (2001) method is reserved for the appendix. Classical measures of dependence are given a decision analytic interpretation in Section 3 and are expressed utilizing the TS framework's generating cdf $P(z)$. In Section 4, we use the general property formulations from Section 3 to derive their closed form expressions for GDB copulae with specific TS generating pdf's. Entropy measures have been used to design a variety of minimally informative constructs given partial information. Bedford and Meeuwissen (1997) specifically used one of them to construct maximum entropy copulae given a correlation constraint, which unfortunately are not available in a closed form. Abbas (2006) applied the entropy concept to the construction of utility functions when only partial preference information is available. In Section 5, different TS generating pdf's are compared using a GDB's copula entropy in the context of matching an expert's elicited joint GDB probability. For illustration, the use of the sub-family of GDB copulae discussed herein is exemplified in a value of information decision analysis example in Section 6.

2. SOME PROPERTIES OF GDB COPULAE WITH TS GENERATING DENSITIES

Bojarski's (2001) generalizations of DB copulae with complete unit-square support require a symmetric generating pdf $f(\cdot)$ with support $[-1,1]$. A general method for generating symmetric distributions is provided by the following Two-Sided framework of distributions with support $[-1,1]$ and pdf

$$f\{z|p(\cdot|\Psi)\} = \frac{1}{2} \times \begin{cases} p(z+1|\Psi), & \text{for } -1 < z \leq 0, \\ p(1-z|\Psi), & \text{for } \quad 0 < z < 1, \end{cases} \tag{7}$$

where $p(\cdot|\Psi)$ can be any *generating pdf* with support $[0,1]$ and the parameters Ψ may in principle be vector-valued. Observe that the pdf (7) consists of two separate branches (hence the name "Two-Sided") and that the shape of each branch is described by the same generating pdf $p(\cdot|\Psi)$. From (7) it immediately follows that $f(-z|p(\cdot|\Psi)) = f\{z|p(\cdot|\Psi)\}$ for $z \in [0,1]$. Thus, similar to a triangular distribution with lower and upper bounds -1 and 1 and most-likely value 0, the pdf (7) is indeed symmetric. In fact, by substituting $p(z) = 2z$ in (7), the pdf (7) reduces to this Triang$(-1,0,1)$ pdf. The inverse cdf (or quantile function) associated with (7) has the following form

$$F^{-1}\{u|p(\cdot|\Psi)\} = \begin{cases} P^{-1}(2u|\Psi) - 1, & \text{for } 0 < u \leq \frac{1}{2}, \\ 1 - P^{-1}(2 - 2u|\Psi), & \text{for } \frac{1}{2} < u < 1, \end{cases} \tag{8}$$

where $P^{-1}(\cdot|\psi)$ is the quantile function of $p(\cdot|\Psi)$. Hence, sampling from (7) is computationally efficient provided $P^{-1}(\cdot|\psi)$ is available in a closed form. For example, the Two-Sided Power (TSP) family of distributions, discussed in more detail in Kotz and Van Dorp (2004), arise from their TS framework by setting $p(z|n) = nz^{n-1}, n > 0$. They allow for efficient sampling utilizing its closed form quantile function $P^{-1}(q|n) = q^{1/n}$, $q \in [0, 1]$.

Utilizing $f\{z|p(\cdot|\Psi)\}$ given by (7) as the generator function for a GDB copula (while setting $\theta = 0$), we derive in the appendix its copula pdf in terms of the TS framework's (7) generating pdf $p(\cdot|\Psi)$. The resulting copula pdf is:

$$c\{x,y|p(\cdot|\Psi)\} = \frac{1}{2} \times \begin{cases} p(1-x-y|\Psi) + p(1+x-y|\Psi), & (x,y) \in A_1, \\ p(1-x-y|\Psi) + p(1-x+y|\Psi), & (x,y) \in A_2, \\ p(x+y-1|\Psi) + p(1+x-y|\Psi), & (x,y) \in A_3, \\ p(x+y-1|\Psi) + p(1-x+y|\Psi), & (x,y) \in A_4, \end{cases} \tag{9}$$

where

$$\begin{aligned} A_1 &= \{(x,y) \in [0,1]^2 | \, 0 < x+y \leq 1, \, -1 < x-y \leq 0\}, \\ A_2 &= \{(x,y) \in [0,1]^2 | \, 0 < x+y \leq 1, \quad 0 < x-y < 1\}, \\ A_3 &= \{(x,y) \in [0,1]^2 | \, 1 < x+y \leq 2, \, -1 < x-y \leq 0\}, \\ A_4 &= \{(x,y) \in [0,1]^2 | \, 1 < x+y \leq 2, \quad 0 < x-y < 1\}. \end{aligned} \tag{10}$$

Substitution of the generating pdf $p(z) = 2z$ in (9) yields

$$c(x,y) = 2 \times \begin{cases} 1-y, & (x,y) \in A_1, \\ 1-x, & (x,y) \in A_2, \\ x, & (x,y) \in A_3, \\ y, & (x,y) \in A_4. \end{cases} \tag{11}$$

Figure 2A plots the joint copula pdf $c(x,y)$ given by (11). Since (7) reduces to a symmetric triangular pdf when $p(z) = 2z$, one could refer to the copula in Figure 2A as the *triangular* copula. Figure 2B displays iso-contours for the density in Figure 2A with their pdf values indicated along the unit-square border. The areas $A_i, i = 1, \ldots, 4$ defined by (10) are depicted in Figure 2 as well.

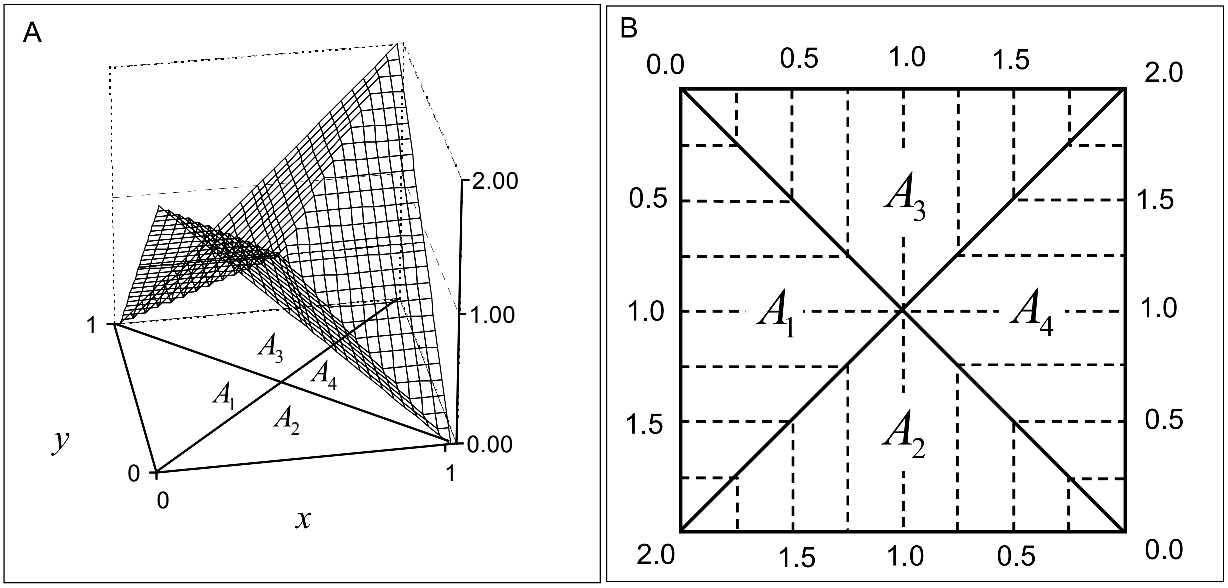

Figure 2. A: Copula density $c\{x,y\}$ given by (11) and (10); B: Density contour plot for Figure 2A. Density values of the iso-contours in Figure 2B are indicated along the unit-square border. Solid diagonal lines in Figures 2A and B are the unit-square diagonals.

2.1. Cumulative distribution function

The joint cumulative distribution function follows directly from (9) as

$$C\{x,y|p(\,\cdot\,|\Psi)\} = \begin{cases} x - \frac{1}{2}\int_{1-x-y}^{1+x-y} P(z|\Psi)dz, & (x,y)\in A_1, \\ y - \frac{1}{2}\int_{1-x-y}^{1-x+y} P(z|\Psi)\}dz, & (x,y)\in A_2, \\ x - \frac{1}{2}\int_{x+y-1}^{1+x-y} P(z|\Psi)dz, & (x,y)\in A_3, \\ y - \frac{1}{2}\int_{x+y-1}^{1-x+y} P(z|\Psi)dz, & (x,y)\in A_4. \end{cases} \quad (12)$$

where $P(z|\Psi)$ is the cumulative distribution function of the generating density $p(z|\Psi)$ in (7) and the areas $A_i, i=1,\ldots,4$ are provided by (10). Please note that the integration boundaries in $C\{x,y|p(\,\cdot\,|\Psi)\}$ coincide with the arguments in the definition of the joint density $c\{x,y|p(\,\cdot\,|\Psi)\}$ given by (9). This apparent simple connection between (12) and (9), however, belies the effort in verifying (12). Substitution of generating pdf $p(z)=2z$ in (12) (and thus generating cdf $P(z)=z^2$) yields

$$C\{x,y\} = \frac{1}{3} \times \begin{cases} -x^3 - 3xy^2 + 6xy\,, & (x,y) \in A_1, \\ -y^3 - 3x^2y + 6xy, & (x,y) \in A_2, \\ y^3 - 3y^2 + 3y(x^2+1) - 3x^2 + 3x - 1, & (x,y) \in A_3, \\ x^3 - 3x^2 + 3x(y^2+1) - 3y^2 + 3y - 1, & (x,y) \in A_4. \end{cases} \quad (13)$$

An example graph of the joint cdf $C\{x,y\}$ given by (13) is provided in Figure 3.

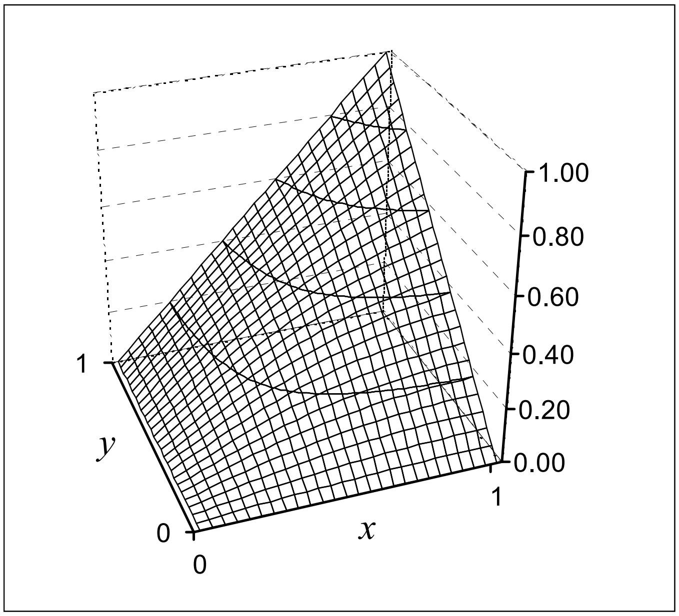

Figure 3. Graph of joint triangular copula cdf $C(x,y)$ given by (13).

2.2. Dependence parameter elicitation

Consider a pair or random variables (X', Y') with marginal cdf's G and H respectively. Assume further that the dependence between (X', Y') is described such that

$$\{G(X'), H(Y')\} = (X,Y) \sim C\{x,y|p(\,\cdot\,|\Psi)\} \quad (14)$$

where $C\{x,y|p(\,\cdot\,|\Psi)\}$ is the bivariate cdf (12) and the generating cdf $P(z|\Psi)$ in (12) is a member of a single parameter family of distributions with support $[0,1]$. Let $x'_{0.5}$ be the median of X' and $y'_{0.5}$ of Y'. To elicit the dependence parameter Ψ we suggest the elicitation of the conditional probability $Pr(Y' \leq y'_{0.5}|X' \leq x'_{0.5})$ (or vice versa). This elicitation procedure falls within the conditional fractile estimates method for eliciting correlations described in Clemen and Reilly (1999).

Should (X', Y') be a pair of independent variables, one has $Pr(Y' \leq y'_{0.5}|X' \leq x'_{0.5}) = 0.5$. If the expert judges that high values of X' tend to be associated with high (low) values of Y' he/she would provide a value larger (less) than 0.5. Suppose the expert answers

$$Pr(Y \leq 0.5|X \leq 0.5) = \pi \in [0, 1]. \tag{15}$$

From (15) and (12) we have

$$\frac{1}{2}\pi = Pr(Y \leq 0.5, X \leq 0.5) = C\{\frac{1}{2}, \frac{1}{2}|p(\cdot|\Psi)\} = \frac{1}{2}\int_0^1 \{1 - P(z|\Psi)\}dz \tag{16}$$

Utilizing the following relationship between $E[Z\,|\,\Psi]$ and the generating cdf $P(z\,|\,\Psi)$

$$E[Z|\Psi] = \int_0^1 \{1 - P(z|\Psi)\}dz, \tag{17}$$

one arrives with (16) and (17) at the following simple expression

$$\pi = E[Z|\Psi]. \tag{18}$$

Hence, substitution of the generating pdf $p(z) = 2z$ in (16) and (18) yields the joint probability value $Pr(Y \leq 0.5,\ X \leq 0.5) = \frac{1}{2} \times \frac{2}{3} = \frac{1}{3}$ for the cdf in Figure 3 which can be easily verified with (13).

Summarizing, having elicited π from an expert, solving for the dependence parameter ψ is equivalent to solving for the parameter ψ of the generating pdf $p(z\,|\,\psi)$ from (18) using the method of moments. Of course, one can only guarantee a solution to (18) if the range of $E[Z|\Psi]$ as a function ψ of equals $[0, 1]$. (Recall that the random variable Z has support $[0, 1]$).

2.3. Sampling procedure

Sampling from the copula $c\{x, y|p(\cdot|\Psi)\}$ (9) follows its construction method presented in the appendix and is efficient provided the quantile function $P^{-1}(z)$ of the TS framework's (7) generating pdf $p(z|\Psi)$ is available in a closed form. The following algorithm generates a bivariate sample (x, y) from the bivariate copula $c\{x, y|p(\cdot|\Psi)\}$ given by (9) :

Step 1: Sample x from a uniform random variable X on $[0,1]$.

Step 2: Sample u from a uniform random variable U on $[0,1]$.

Step 3: If $u \leq \frac{1}{2}$ then $z = P^{-1}(2u) - 1$ else $z = 1 - P^{-1}(2 - 2u)$

Step 4: $y = z + x$

Step 5: If $y < 0$ then $y = -y$

Step 6: If $y > 1$ then $y = 1 - (y - 1)$

Hence, using the generating pdf $p(z) = 2z$ in the algorithm above and given samples $x = \frac{3}{4}, u = \frac{3}{4}$ in Steps 1 and 2, we have, utilizing $P^{-1}(v) = \sqrt{v}$ in Step 3, from Step 4 that $y = 1\frac{3}{4} - \frac{1}{2}\sqrt{2} \approx 1.043$. Finally, Step 6 yields the bivariate sample (x, y) where $x = \frac{3}{4}$ and $y = \frac{1}{4} + \frac{1}{2}\sqrt{2} \approx 0.957$.

3. ORDINAL MEASURES OF ASSOCIATION

Positive (negative) dependence is present amongst two continuous random variables $X' \sim G(\cdot)$ and $Y' \sim H(\cdot)$ when large values of one are associated with large (small) values of the other. In case of positive (negative) dependence, X' and Y' are said to be concordant (disconcordant). Classical quantities that measure the degree of positive or negative dependence between X' and Y' are Blomquist's (1950) β (sometime also referred to as Blomquist's q), Kendall's (1938) τ and Spearman's (1904) rank correlation ρ_s. All three dependence measures attain values ranging from -1 to 1. They are ordinally invariant, which implies that the degree of dependence between the pair (X', Y') is the same as that between the pair $(X, Y) = \{G(X'), H(Y')\}$. Recall that by the probability integral transformation (e.g. Nelsen (1999)), X and Y are uniformly distributed random variables on $[0,1]$ and thus the bivariate distribution of (X, Y) is a copula.

An excellent exposition and comparison of these ordinal measures of association is provided by Kruskal (1958). Kruskal (1958) provides in his paper operational interpretations for all three measures which are equivalent to the expected monetary value (EMV) pay-offs of the probability trees in Figures 4A, B and C, where

$$(X_i, Y_i) = \{G(X_i'), H(Y_i')\}, i = 1, \ldots, 3 \quad (19)$$

are three independent random bivariate samples from the distribution under consideration. Observe from (19) that $X_i' < Y_i'$ if and only if $X_i < Y_i, i = 1, \ldots, 3$. We conclude from Figure 4 that the operational interpretations of Blomquist's β (Figure 4A), Kendall's τ (Figure 4B) and Spearman's rank correlation ρ_s (Figure 4C) involve 1, 2 and 3 independent random bivariate samples, respectively. In case of independence between X' and Y' (and thus X and Y) we have immediately from Figure 4 that $\beta = \tau = \rho_s = 0$. Moreover, in case of complete positive (negative) dependence, i.e. $Y_i = X_i$ ($Y_i = -X_i$), one observes from Figures 4A and B that $\beta = \tau = 1$ ($\beta = \tau = -1$). Kruskal (1958) (page 824) showed that the same applies to ρ_s in case of complete positive (negative) dependence. The equivalent population quantities for the expected pay-offs β, τ and ρ_s in Figures 4A, B and C are:

$$\begin{cases} \beta(X,Y) = 4C(\frac{1}{2}, \frac{1}{2}) - 1, \\ \tau(X,Y) = 4\int_0^1\int_0^1 C(x,y)c(x,y)dxdy - 1, \\ \rho_s(X,Y) = 12\int_0^1\int_0^1 xyc(x,y)dxdy - 3, \end{cases} \quad (20)$$

where $C(x,y)$ and $c(x,y)$ are the joint copula cdf and pdf of (X,Y), respectively. Hence, one concludes from (20) and $(X,Y) = \{G(X'), H(Y')\}$, where $X' \sim G(\cdot)$ and $Y' \sim H(\cdot)$, that the value for Spearman's ρ_s of (X', Y') equals that of the Pearson's (1920) product moment correlation for (X,Y).

Substituting $C\{x,y\,|\,p(\cdot\,|\Psi)\}$ and $c\{x,y\,|\,p(\cdot\,|\Psi)\}$ given by (12) and (9) in (20), we have after straightforward, but lengthy and tedious, algebraic manipulations that

$$\begin{cases} \beta\{X,Y|p(\cdot\,|\Psi)\} = 2E[Z|\Psi] - 1, \\ \tau\{X,Y|p(\cdot\,|\Psi)\} = 2E[Z^2|\Psi] - 2\int_0^1 P^2(s|\Psi)ds + 4\int_0^1 sP^2(s|\Psi)ds - 1, \\ \rho_s\{X,Y|p(\cdot\,|\Psi)\} = -4E[Z^3|\Psi] + 6E[Z^2|\Psi] - 1, \end{cases} \quad (21)$$

where $Z \sim P(s|\Psi)$ and $P(s|\Psi)$ is the cdf of the generating pdf $p(\cdot\,|\Psi)$ of the TS framework of distributions (7). We have for $p(z) = 2z$, $E[Z] = \frac{2}{3}, E[Z^2] = 1/2$, $\int_0^1 P^2(s)ds = \frac{1}{5}$,

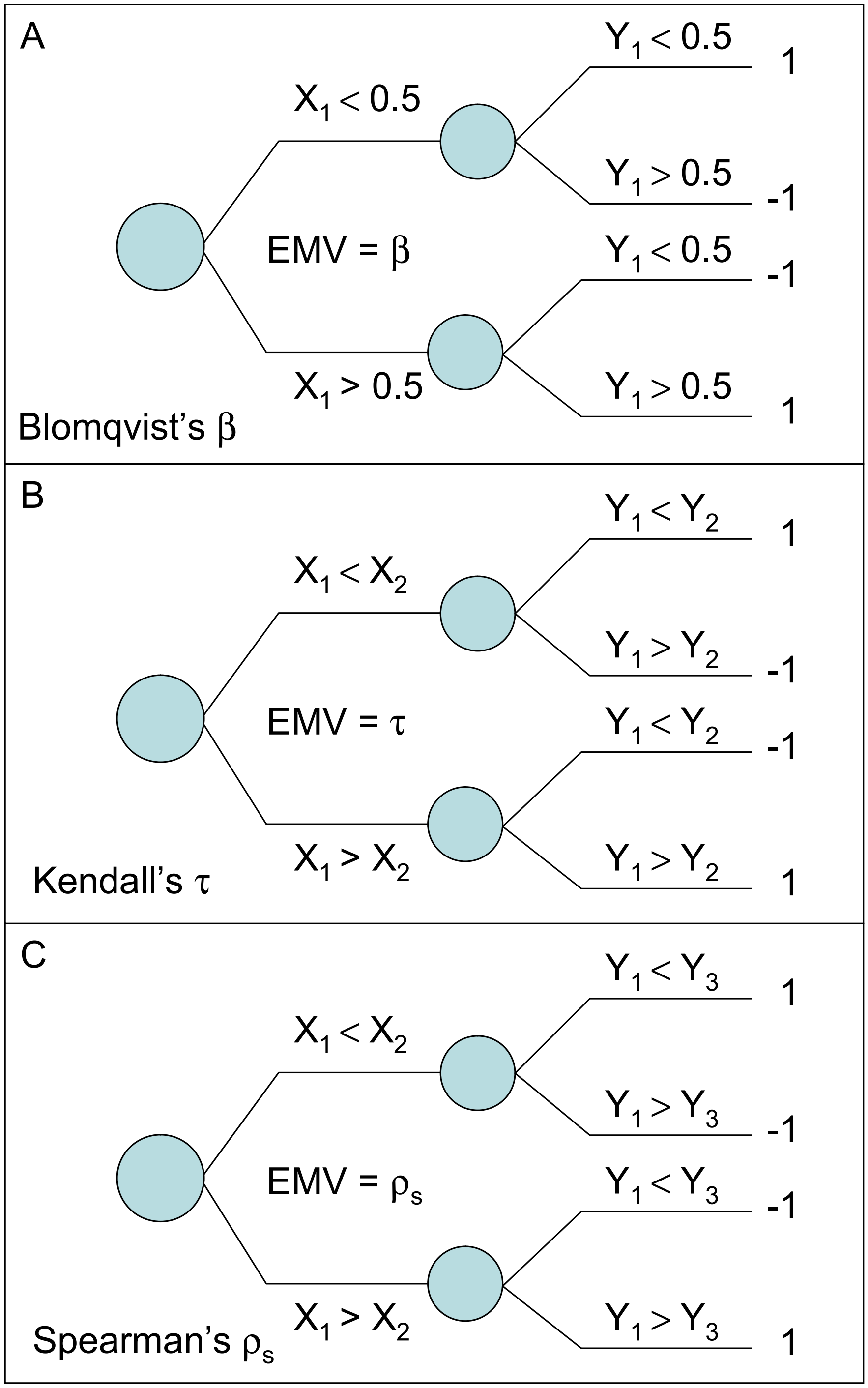

Figure 4. Operational interpretations of ordinal measures of association Blomquist's (1950) β (A), Kendall's (1938) τ (B) and Spearman's (1904) rank correlation ρ_s (C). Samples $(X_i, Y_i) = \{G(X_i'), H(Y_i')\}, i = 1, \ldots, 3$ are independent bivariate samples from a joint distribution under consideration with marginals $X' \sim G(\cdot)$ and $Y' \sim H(\cdot)$.

$\int_0^1 sP^2(s)ds = \frac{1}{6}$ and $E[Z^3] = 2/5$. Thus utilizing (21) we obtain $\tau(X,Y) = \frac{4}{15} < \beta(X,Y) = \frac{1}{3}$ $< \rho(X,Y) = \frac{2}{5}$ for the copula in Figure 2A.

We have from (21) and (18) in our case that $\beta\{X,Y|p(\cdot|\Psi)\} = 2\pi\{X,Y|p(\cdot|\Psi)\} - 1$, where $\pi\{X,Y|p(\cdot|\Psi)\} = Pr(Y < 0.5|X < 0.5)$. Hence, the elicitation of $\pi\{X,Y|p(\cdot|\Psi)\}$ suggested in Section 2.2 is equivalent to an indirect elicitation procedure for Blomquist's β. Moreover, observe from Figure 4 that Blomquist's β interpretation requires the least cognitive processing compared to τ and ρ_s, since it only involves one bivariate random sample, as opposed to two and three, respectively. This further supports the indirect elicitation of Blomquist's β over Kendall's τ or Spearman's ρ_s.

Bojarski (2001) derived the following relationship between the correlation coefficient $\rho(X,Y|\theta)$ of his GDB copula and the random variable $V \sim f(v|\theta)$:

$$\rho_s\{X,Y|f(\cdot|\theta)\} = 4E[|V|^3|\theta] - 6E[V^2|\theta] + 1, \tag{22}$$

where $f(v)$ is its symmetric generating density with support $[-(1-\theta),(1-\theta)]$. The copula pdf $c\{x,y|p(\cdot|\Psi)\}$ (9) is a member within Bojarski's GDB class of copulae by setting $f(v|\theta)$ equal to (7) and thus $\theta = 0$. Hence, the difference between Bojarski's (2001) larger class of GDB copulae and the sub-class (9) is that a full support $[0,1]$ generating density $p(\cdot|\Psi)$ yields a GDB copula with full unit-square $[0,1]^2$ support, whereas the former continues to have the more restricted diagonal band support shared also by the $DB(\theta)$ copula in Figure 1. Utilizing (7) we have in the case that $\theta = 0$

$$\begin{cases} E[V^2|\theta = 0, f(\cdot)] = E[(1-Z)^2|p(\cdot|\Psi)] \\ E[|V|^3|\theta = 0, f(\cdot)] = E[(1-Z)^3|p(\cdot|\Psi)], \end{cases} \tag{23}$$

and by substituting (23) into (22) the correlation coefficient $\rho_s\{X,Y|f(\cdot|\theta)\}$ reduces to expression for ρ_s in (21). Observe the similarity between the expressions for ρ_s in (21) and (22) (with (22) containing the absolute third moment).

3.1. Reflection Property

Consider a GDB copula density $c\{x, y|p(\,\cdot\,|\Psi)\}$ given by (9) such that $Z \sim p(z|\Psi)$ and $p(z|\Psi)$ is the generating density of (7). When $c\{x, y|p(\,\cdot\,|\Psi)\}$ exhibits positive (negative) dependence, a GDB copula $c\{x, y|q(\,\cdot\,|\Psi)\}$ may be constructed exhibiting the same degree of negative (positive) dependence, where $q(z|\Psi)$ is the density function of $Z' = 1 - Z$. Hence,

$$q(z|\Psi) = p(1 - z|\Psi),\ z \in [0, 1]. \tag{24}$$

The density $q(z|\Psi)$ is referred to as the *reflection* density of the generating density $p(z|\Psi)$. From (24) and (21) it immediately follows that :

$$\begin{cases} \beta\{X, Y|q(z|\Psi)\} = \beta\{X, Y|p(1 - z|\Psi)\} = -\beta\{X, Y|p(z|\Psi)\}, \\ \tau\{X, Y|q(z|\Psi)\} = \tau\{X, Y|p(1 - z|\Psi)\} = -\tau\{X, Y|p(z|\Psi)\}, \\ \rho_s\{X, Y|q(z|\Psi)\} = \rho_s\{X, Y|p(1 - z|\Psi)\} = -\rho_s\{X, Y|p(z|\Psi)\}. \end{cases} \tag{25}$$

The copula density $c\{x, y|q(z|\Psi)\} = c\{x, y|p(1 - z|\Psi)\}$ may be obtained from $c\{x, y|p(\,\cdot\,|\Psi)\}$ via a right angle rotation. Figure 5 plots the density $c\{x, y|p(1 - z|\Psi)\}$ where $p(z) = 2z$. Observe it is a rotated version of the density depicted in Figure 2A.

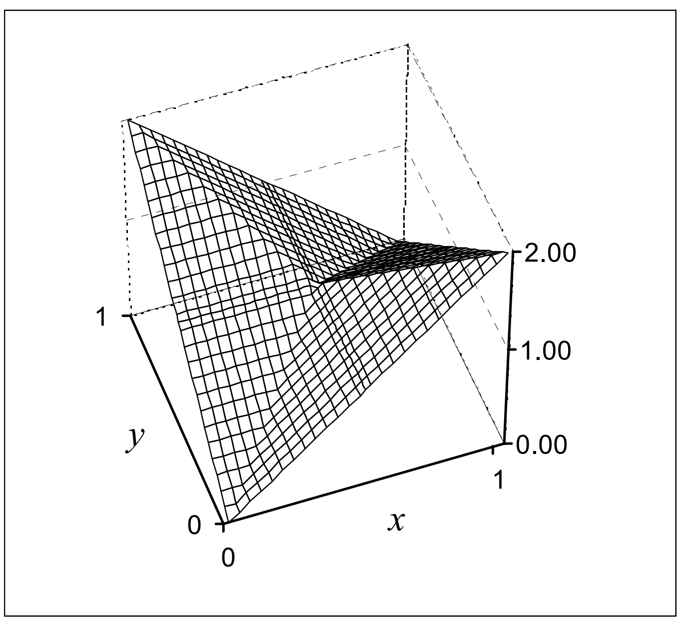

Figure 5. Graph of rotated triangular copula $c\{x, y|p(1 - z|\Psi)\}$ defined by (9) using the reflection generating density $p(1 - z|\Psi) = 2(1 - z)$.

The expression for $-\rho_s\{X, Y|p(z|\Psi)\}$, where $\rho_s\{X, Y|p(z|\Psi)\}$ is given by (21), is identical to the expression for $\rho_s\{X, Y|g(y)\}$ derived by Ferguson's (1995) for copula's (6). Hence, from (25) it follows that Ferguson's generating density $g(z)$ in (6) enjoys the alternative interpretation of being a generating density of the TS framework of distributions (7) for GDB copula construction.

3.2. Lower and upper tail dependence

The burst of applications and attention to the copula approach may be credited to the Gaussian copula which has been widely adopted by the "financial quants" in recent years. Unfortunately, it has also recently received negative press (and by association the copula approach) and some have gone as far (see, e.g., Salmon, 2009) as to blame the 2008 financial crash on the use of the Gaussian copulae, perhaps in part due to their lack of lower and upper tail dependence. Currently, lower and upper tail dependence measures are in vogue, particularly in problem contexts dealing with modeling the joint occurrence of extreme events, such as insurance and modeling of default risk in finance. These measures too are ordinal measures of association, although they focus primarily on modeling positive dependence and not negative dependence. Recalling the two continuous random variables $X' \sim G(\cdot)$ and $Y' \sim H(\cdot)$, the population expressions for lower tail dependence λ_L and upper tail dependence λ_U are, respectively:

$$\lambda_L = \lim_{x \downarrow 0} Pr\{Y' \leq H^{-1}(x)|X' \leq G^{-1}(x)\} \tag{26}$$
$$= \lim_{x \downarrow 0} Pr(Y \leq x|X \leq x) = \lim_{x \downarrow 0} \frac{C(x,x)}{x},$$

$$\lambda_U = \lim_{x \uparrow 1} Pr\{Y' > H^{-1}(x)|X' > G^{-1}(x)\} \tag{27}$$
$$= \lim_{x \uparrow 1} Pr(Y > x|X > x) = \lim_{x \uparrow 1} \frac{1 - 2x + C(x,x)}{1 - x}.$$

Copulae that do exhibit strictly lower or upper tail dependence (i.e. $\lambda_L > 0$ or $\lambda_U > 0$) are the Clayton, Frank and Gumbel copulae that all belong to the Archimedean class of copulae (see, e.g., Joe, 1997).

From (26), (27), $C\{x, y|p(\cdot|\Psi)\}$ given by (12) and by applying l'Hôpital's rule, we have for GDB copulae with TS generating densities $\lambda_L = \lambda_U = 0$, similar to the Gaussian copulae (Embrechts et al., 2002). In our opinion, traditional measures of dependence such as Blomquist's β, Kendall's τ and Spearman's ρ_S are more applicable in problem contexts not dealing with the modeling of joint extreme events per se, but dealing with the modeling of joint events in general. Indeed, these traditional ordinal measures pertain to the full support of a copula and not just to its asymptotic extreme values.

We are somewhat puzzled by the level of criticism that Gaussian copulae have been exposed to (see, e.g., Salmon, 2009) and would like to caution those who believe that the Clayton, Frank and Gumbel copulae could serve as the panacea. Indeed, it has long been recognized that the variances (volatility) in the time series of financial processes are typically not constant, which eventually led to the introduction of, amongst others, the Auto-Regressive Conditional Heteroscedastic (ARCH) models by Nobel-Laureate Engle in 1982. Hence, it would seem unlikely that a joint covariance process exists that cancels the volatility of two separate financial marginal processes leading to a single constant correlation value over time. Summarizing, it appears that dependence modeling between two dependent financial processes require more complex constructs than the use of a single bivariate copula, regardless of it displaying tail dependence or not.

4. GDB COPULA EXAMPLES WITH TS GENERATING DENSITIES

In this section we shall use the properties in Sections 2 and 3 to study GDB copulae with the following generating pdf:

Slope(α) pdf, $$p(z|\alpha) = 2 - \alpha + 2(\alpha - 1)z,\ 0 \leq \alpha \leq 2, \tag{28}$$

Power(n) pdf, $$p(z|n) = nz^{n-1},\ n > 0, \tag{29}$$

Ogive(m) pdf, $$p(z|m) = \frac{m+2}{3m+4}\{2(m+1)\sqrt{z^m} - mz^{m+1}\}, m > 0. \tag{30}$$

Uniform$[\theta, 1]$, $$p(z|\theta) = \frac{1}{1-\theta}, \, \theta \leq z \leq 1, \, 0 \leq \theta \leq 1, \tag{31}$$

Beta(a, b) pdf, $$p(z|a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \; x^{a-1}(1-x)^{b-1}, a > 0, \, b > 0, \tag{32}$$

where all pdf's, except (31), have support $[0, 1]$ and where in (32) $\Gamma(\,\cdot\,)$ is the gamma function with property $\Gamma(a+1) = a\Gamma(a)$. Pdf's (28)-(30) were utilized in Kotz and Van Dorp (2003) to introduce and exemplify their TS framework of distributions and are referred to as the slope, power and ogive distributions, respectively. Pdf (32) is the classical beta distribution, whereas the pdf (29) may be recognized as a uniform distribution with limited density support $[\theta, 1]$.

Figure 6 displays GDB copulae and their generating pdf's (28)-(30) with parameter settings

$$\alpha = 1.5, \, n = 3, m = 4.916. \tag{33}$$

Figure 7 displays GDB copula and their generating pdf's (31)-(32) with parameter settings

$$\theta = 0.5; a = b = 5. \tag{34}$$

The generating pdf's in Figures 6C, E and 7A all have in common that $E[Z] = 3/4$. For the generating pdf in Figure 7C (6A) we have $E[Z] = 1/2$ $(7/12)$. Observe from Figure 7A that by reducing the support to $[\theta, 1]$ for the generating pdf of the TS framework, GDB copula pdf (9) leads via (31) to the original $DB(\theta)$ copula displayed in Figure 1. The generating pdf's (28) - (32) $p(z|\psi)$ may also be observed in Figures 6B, D and F and 7B, D as conditional densities of $(Y|X=1)$ or $(X|Y=1)$. Reflected versions of the generating pdf's are observed in the same figures as conditional densities of $(Y|X=0)$ or $(X|Y=0)$. It may be somewhat surprising that the bivariate pdf in Figure 7D is a copula, i.e. that it possesses uniform marginals. Perhaps even more remarkable might be that the copula in Figure 7D has a product moment correlation of 0. However, for any symmetric pdf on $[0, 1]$ we have that $p(1-z) = p(z),$ for all $z \in [0, 1]$, and thus from the reflection property in Section 3.1 and (25) it follows that Blomquist's β, Kendall's τ and

in: *Decision Analysis*, published online before print November 25, DOI:10.1287/deca.1090.0162

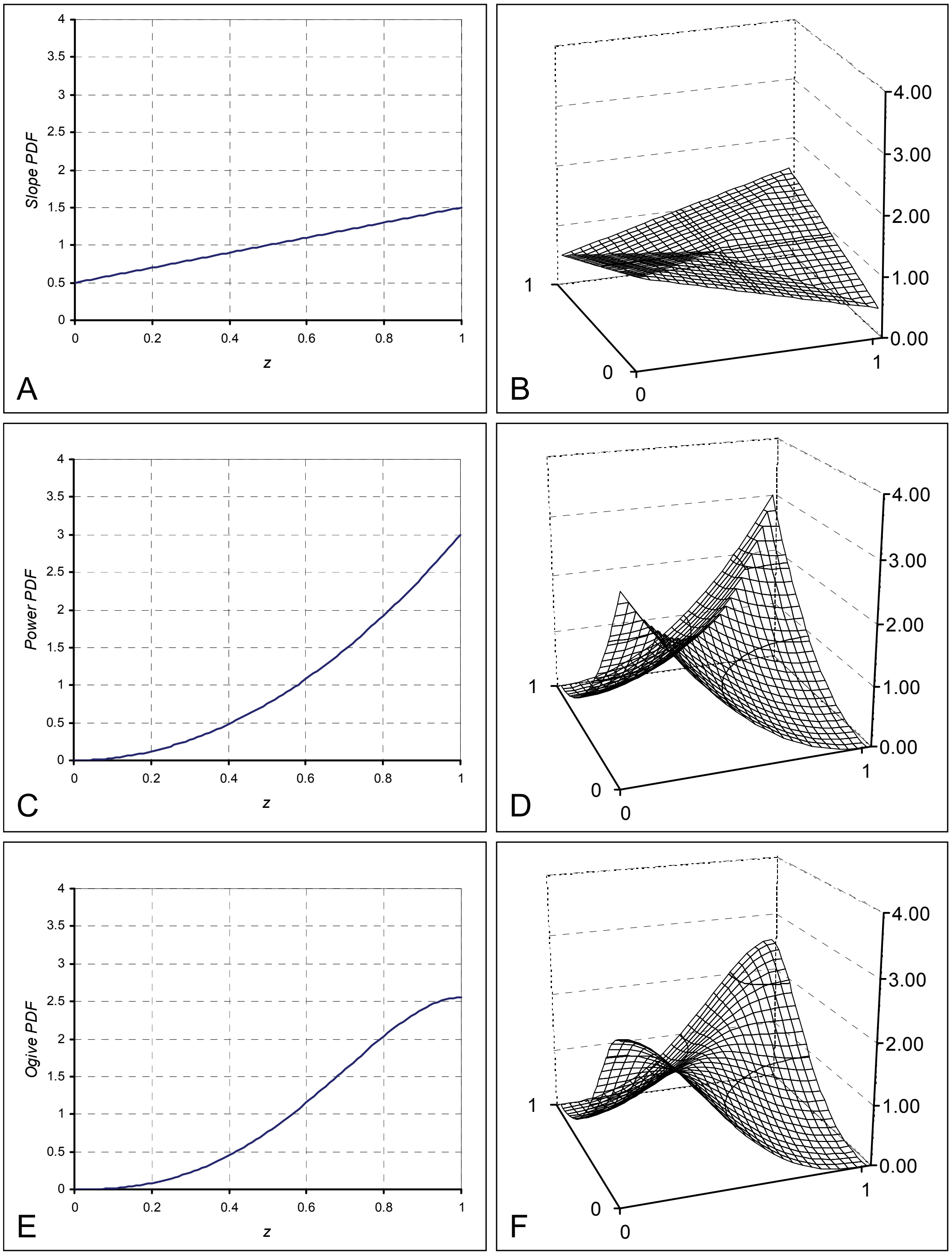

Figure 6. TS framework generating densities with associated GDB Copulae :
A: Slope PDF, $E[Z|\alpha] = 7/12$; B: TS Slope - GDB Copula ($\alpha = 1.5$)
C: Power PDF, $E[Z|n] = 3/4$; D: TSP - GDB Copula ($n = 3$);
E: Ogive PDF, $E[Z|m] = 3/4$; F: TSO - GDB Copula ($m = 4.916$).

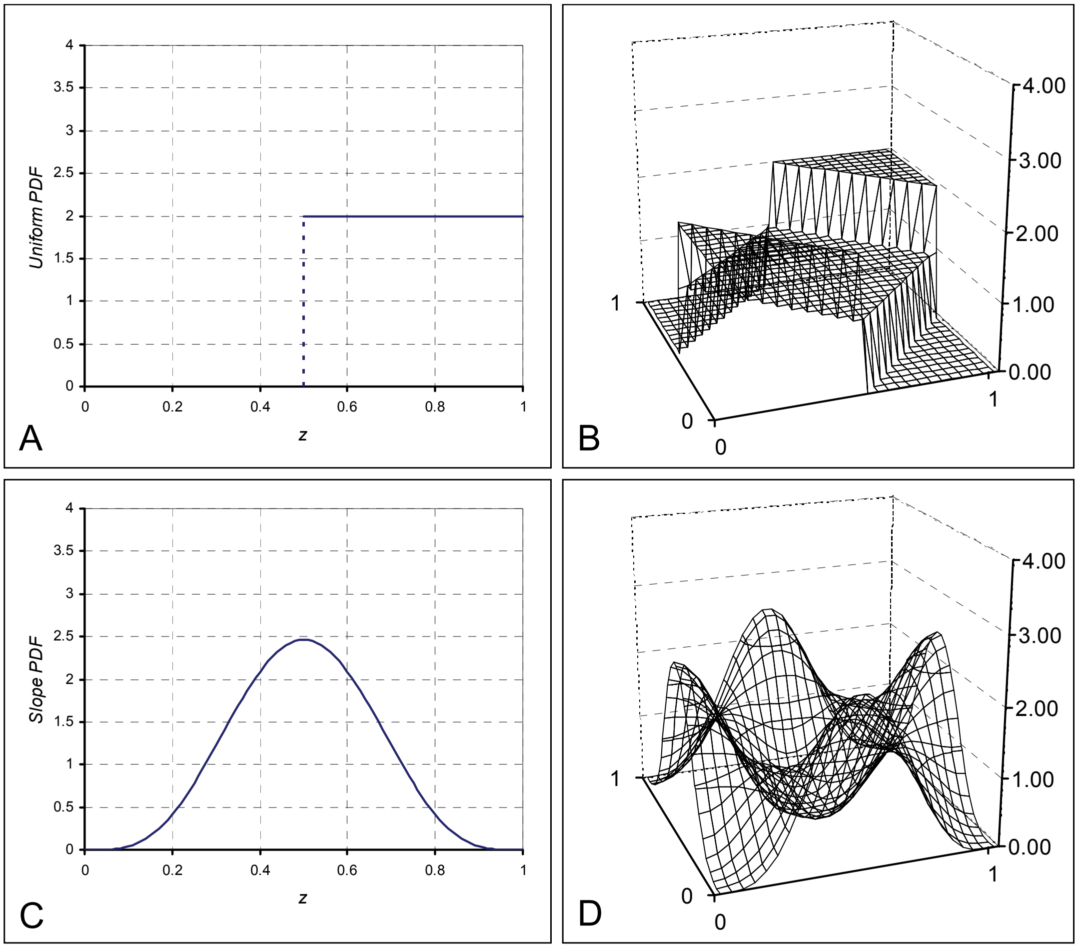

Figure 7. TS framework generating densities with associated GDB Copulae:
A: Uniform $[1-\theta, 1]$ PDF, $E[Z|\theta] = 3/4$; C: Beta PDF, $E[Z|a, b] = 1/2$;
B: DB Copula ($\theta = 0.5$); D: TS Beta - GDB Copula ($a = 5, b = 5$).

Spearman's ρ_s all equal to 0. In other words, any symmetric generating pdf $p(z|\psi)$ for the TS framework (7) yields uncorrelatedness within a GDB copula, whereas when $p(z|\psi)$ is non-uniform on $[0, 1]$ the variables (X, Y) are clearly not statistically independent (see Figure 7D).

GDB copulae with beta generating pdf's were studied by Bojarski (2001), but do not possess a closed form cdf nor a closed form quantile function. We have for the remaining generating densities (28) - (31) for an arbitrary quantile level $q \in (0, 1)$:

$$P^{-1}(q|\psi) = \begin{cases} \frac{-(2-\alpha)+\sqrt{(2-\alpha)^2+4(\alpha-1)q}}{2(\alpha-1)}, & p(z|\alpha), \alpha \neq 1, \text{Eq. (28)}, \\ q^{1/n}, & p(z|n), \text{Eq. (29)}, \\ \left[\frac{2(m+1)}{m} - \sqrt{\left\{\frac{2(m+1)}{m}\right\}^2 - q\frac{3m+4}{m}}\right]^{2/(m+2)}, & p(z|m), \text{Eq. (30)}, \\ (1-\theta)q+\theta, & p(z|\theta), \text{Eq. (31)}. \end{cases} \tag{35}$$

Thus, generating pdf's (28)-(31) allow for an efficient GDB copula sampling algorithm since their quantile functions are available in a closed form. Perhaps one could slightly favor densities (29) and (31) since their quantile functions require the least number of elementary operations for their evaluation.

We have from (21) for Blomquist β, Kendall's τ, Spearman's ρ_s for the generating densities (28) - (31), respectively:

$$\begin{cases} \beta\{X,Y|p(\,\cdot\,|\alpha)\} = -\frac{1}{3}+\frac{1}{3}\alpha, & \in [-\frac{1}{3},\frac{1}{3}], \\ \tau\{X,Y|p(\,\cdot\,|\alpha)\} = -\frac{4}{15}+\frac{4}{15}\alpha, & \in [-\frac{4}{15},\frac{4}{15}], \\ \rho_s\{X,Y|p(\,\cdot\,|\alpha) = -\frac{2}{5}+\frac{2}{5}\alpha, & \in [-\frac{2}{5},\frac{2}{5}], \end{cases} \tag{36}$$

$$\begin{cases} \beta\{X,Y|p(\,\cdot\,|n)\} = \frac{n-1}{n+1}, & \in [-1,1], \\ \tau\{X,Y|p(\,\cdot\,|n)\} = \frac{n-1}{n+2} + \frac{n-1}{(n+1)(n+2)(2n+1)}, & \in [-1,1], \\ \rho_s\{X,Y|p(\,\cdot\,|n)\} = \frac{(n-1)(n+6)}{(n+2)(n+3)}, & \in [-1,1], \end{cases} \tag{37}$$

$$\begin{cases} \beta\{X,Y|p(\,\cdot\,|m)\} = \frac{m(m+1)(3m+8)}{(m+3)(m+4)(3m+4)}, & \in [0,1], \\ \tau\{X,Y|p(\,\cdot\,|m)\} = \frac{m(m+1)(162m^6+2643m^5+18132m^4+66108m^3+140032m+58880)}{(m+3)(m+4)(m+6)(2m+5)(3m+4)^2(3m+8)(3m+10)}, & \in [0,1], \\ \rho_s\{X,Y|p(\,\cdot\,|m)\} = \frac{m(m+1)(3m^3+70m^2+424m+736)}{(m+4)(m+5)(m+6)(m+8)(3m+4)}, & \in [0,1], \end{cases} \tag{38}$$

and

$$\begin{cases} \beta\{X,Y|p(\,\cdot\,|\theta)\} = \theta, & \in [0,1], \\ \tau\{X,Y|p(\,\cdot\,|\theta)\} = \theta(\theta+2)/3, & \in [0,1], \\ \rho_s\{X,Y|p(\,\cdot\,|\theta)\} = \theta(1+\theta-\theta^2), & \in [0,1]. \end{cases} \tag{39}$$

Figures 8A, B, C and D provides a comparison of β, τ and ρ_s for expressions (36) - (39). Observe from (37) and Figure 8B that GDB copulae with a TS power(n) generating pdf (29) allow for a

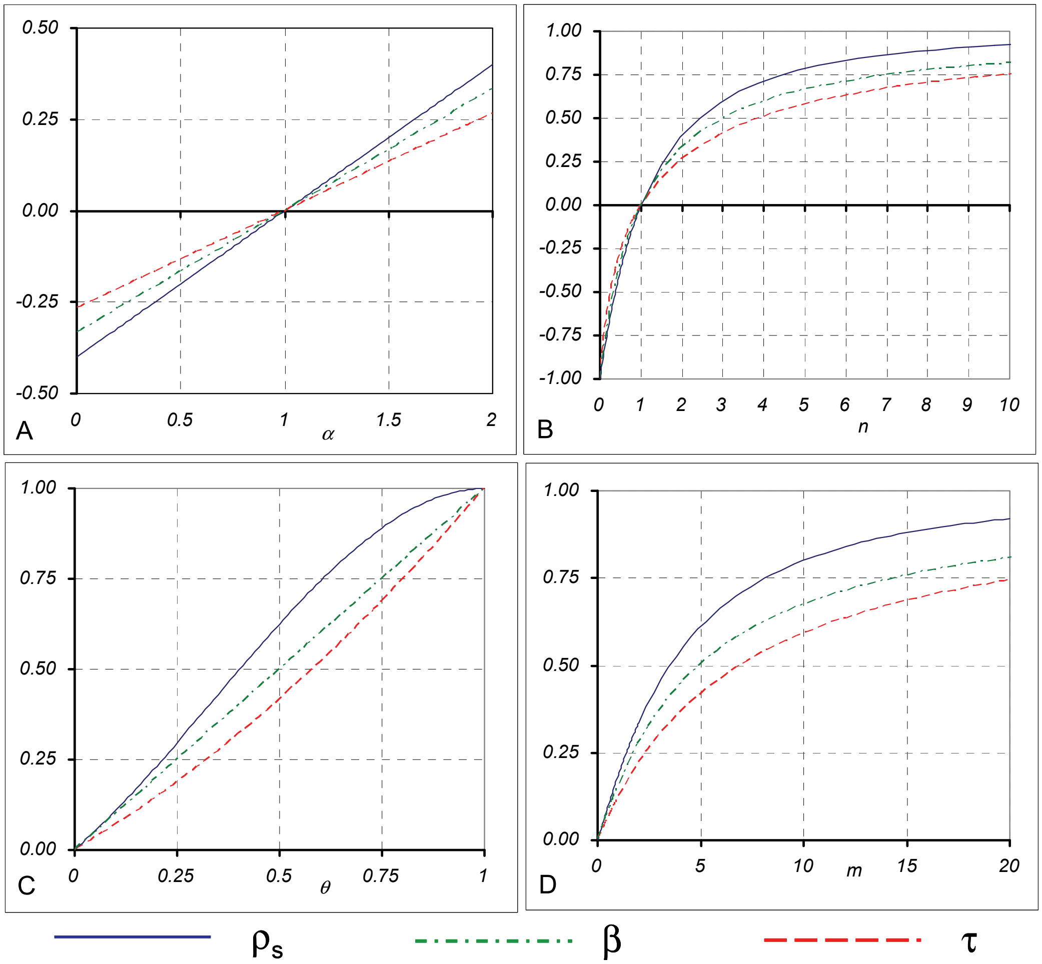

Figure 8. Comparison of ordinal measures of association (36) - (39) for GDB copulae with TS framework generating densities; A: Slope(α); B: Power(n); C: Uniform[θ, 1]; D: Ogive(m).

complete coverage of β, τ and ρ_s. To achieve a full coverage for the generating pdf's (30) and (31) one would have to utilize the reflection property (see, Section 3.1) of GDB copulae. The slope generating pdf (28) only allows for a limited coverage of β, τ and ρ_s, but still larger than the coverage of, for example, $\rho_s \in [-\frac{1}{3}, \frac{1}{3}]$ for the Fairly-Gumbel-Morgenstern (FGM) copulae (see, Schucany et al. (1978)). Observe from Figure 8A, B, C and D that in case of positive (negative) dependence $\tau < \beta < \rho_s$ $(\tau > \beta > \rho_s)$. Kruskal's (1958) paper, however, contains examples with a reversed order of these measures of association. We invite the reader to demonstrate that the ordering relationship above applies to all GDB copulae with TS generating pdf's, which is still an open question. Finally, observe from Figure 8A and expressions (36) that all measures β, τ and ρ_s are linear functions of the slope parameter α, which is remarkable. Solving for α, n and θ given a value for β, τ or ρ_s from second order or less equations in (37), (38) and (39) involves simple algebraic manipulations, whereas higher order expressions therein may require the use of root finding algorithms.

5. AN ELICITATION EXAMPLE

Assume that an expert has assessed a value $\pi = Pr(Y \leq 0.5 | X \leq 0.5) = 0.75$ and that we are tasked to develop a GDB copula with a TS generating pdf that matches this constraint. We have from (28)-(31), (15) and (18) that

$$\pi\{X, Y | p(\cdot|\Psi)\} = \begin{cases} (2+\alpha)/6 \in [\frac{1}{3}, \frac{2}{3}], & p(z|\alpha), \text{Eq. } (28), \\ n/(n+1) \in [0,1], & p(z|n), \text{Eq. } (29), \\ \frac{(m+2)^2}{3m+4}\left[\frac{3m+6}{(m+4)(m+3)}\right] \in [0.5,1], & p(z|m), \text{Eq. } (30), \\ (\theta+1)/2 \in [0.5,1], & p(z|\theta), \text{Eq. } (31). \end{cases} \quad (40)$$

From (40) it follows that for $n = 3, m = 4.916$ and $\theta = 1/2$, the above $\pi = 0.75$ constraint is met using the power(n), ogive (m) and uniform$[\theta, 1]$ generating pdf's (29) - (31), respectively. The slope pdf (28) does not meet this constraint since $0.75 \notin [1/3, 2/3]$. The corresponding copula pdf's are displayed in Figures 6D, 6F and 7B. For $m > 2$ it follows from the ogive pdf (30) that its derivative equals 0 at $z = 0$ and $z = 1$. As a result, the GDB copula in Figure 6F is smooth over its entire support whereas the other copulae in Figures 6D and 7B are not. Hence, if smoothness of the copula were a requirement one could favor the ogive generating pdf (30). For $m > 2$ one obtains for ogive(m) pdf (30) that $\pi(m) > \frac{16}{25} = 0.64$.

What if smoothness were not a requirement? How would one choose amongst these three generating densities? From an argument of being as uniform as possible, one could perhaps select that copula with the smallest correlation coefficient. We have from (37), (38) and (39) that for

$n = 3, m = 4.916, \theta = 1/2$, respectively:

$$\rho_s(n) = \frac{3}{5},\ \rho_s(m) = 0.6059 \text{ and } \rho_s(\theta) = \frac{5}{8}. \tag{41}$$

Hence, this would favor the power(n) generating pdf (26), albeit ever so slightly. However, this raises the question of whether the argument of selecting a copula with the smallest correlation coefficient from a family of copulae that matches a conditional probability constraint, is generally applicable. In the case that $\pi = Pr(Y \leq 0.5|X \leq 0.5) = 0.5$ between uniform$[0, 1]$ random variables X and Y, one would perhaps prefer the copula with independent uniform marginals. However, the copula in Figure 7D also matches the constraint $\pi = 0.5$ and too possesses a zero correlation. In fact, recall from Section 4 that the same holds for any symmetric generating pdf $p(z|\Psi)$. Thus, one arrives at the conclusion that the above question cannot be answered affirmatively. On the other hand, a procedure that selects a copula by minimizing the distance between it and the uniform copula with independent marginals would have selected the latter (and not the one in Figure 7D) given the $\pi = 0.5$ constraint. This suggests to select amongst the GDB copulae with generating pdf's (29)-(31) the one that minimizes a distance measure between it and the copula with independent uniform marginals.

5.1. *Using entropy to compare pdf's*

A well known distance measure between two pdf's $f(x, y)$ and $g(x, y)$ is the relative information of one candidate pdf $f(x, y)$ with respect to another specified pdf $g(x, y)$ given by

$$I(f|g) = \int\int f(x, y)\ln\{f(x, y)/g(x, y)\}dxdy. \tag{42}$$

The quantity (42) is known as the cross entropy or the Kullback-Liebler distance between two distributions f and g. The quantity $I(f|g)$ is non-negative and only equal to zero when $f(x, y) = g(x, y)$ everywhere. Soofi and Retzer (2002) provide a more general discussion on various information indices. Setting $f(x, y) = c\{x, y\}$ and $g(x, y) = u(x, y)$ in (42), where $u(x, y)$ is the density on $[0, 1]^2$ with independent uniform marginals, (42) reduces to:

$$I(c|u) = \int\int_{S_c} c(x,y)\ln\{c(x,y)\}dxdy, \tag{43}$$

where $c(x,y)$ is a copula with support $S_c \subset [0,1]^2$. The quantity $E = -I(c|u)$ is known as the entropy of the pdf $c(x,y)$. The measure $-E = I(c|u) \geq 0$ measures the information imbedded within $c(x,y)$ relative to the uniform pdf $u(x,y)$.

Hence, given a particular constraint imposed on a sub-family of copulae, one could select that copula that is least informative by minimizing (43) or equivalently by maximizing its entropy. Bedford and Meeuwissen (1997) specifically used the relative information (43) to construct maximum entropy copulae given a correlation constraint. Unfortunately, their maximum entropy copulae do not possess closed form distribution function expressions and require a discrete approximation on a fine grid on $[0,1]^2$ for their evaluation, which is not computationally efficient from a sampling perspective. Utilizing numerical integration over a 100 by 100 grid over $[0,1]^2$, we have for the copulae in Figures 6D, F and 7B, respectively,

$$I\{c(x,y)|p(\cdot|\psi)\} = \begin{cases} 0.2136, & p(z|n), n = 3, \\ 0.2222, & p(z|m), m = 4.916, \\ 0.3400, & p(z|\theta), \theta = 0.5. \end{cases} \tag{44}$$

Summarizing, given the constraint set by $\pi = Pr(Y \leq 0.5|X \leq 0.5) = 0.75$, the relative information approach above would suggest to use the GDB copula with the power(n) generating pdf (29) with $n = 3$.

5.2. Results of entropy comparison

Figure 9 provides a relative information analysis for DB copulae and GDB copulae with TS slope, power and ogive generating densities as a function of $\pi = Pr(Y \leq 0.5|X \leq 0.5)$ and as a function of the copula correlation coefficient ρ_s (since Bedford and Meeuwissen's maximum entropy copulae were constructed with ρ_s in mind). Figure 9 is split in four sub-panels. Panels 9A (9B) deals with the range $0 \leq \pi \leq \frac{2}{3}$ ($0 \leq \rho_s \leq 0.4$), which coincides with the restricted range associated for the slope generating densities. Thus, a slope generating pdf analysis is not included in

Figures 9C (9D) since it deals with $\frac{2}{3} \le \pi \le 0.95$ ($0.4 \le \rho_s \le 0.99$). Figures 9B and 9D also include combinations of correlations and relative information values for Bedford and Meeuwissen's (1997) minimal information copulae. These data points are indicated by the large solid bullets in Figures 9B and 9D and were provided by Lewandowski's (2005) Table 1, page 65.

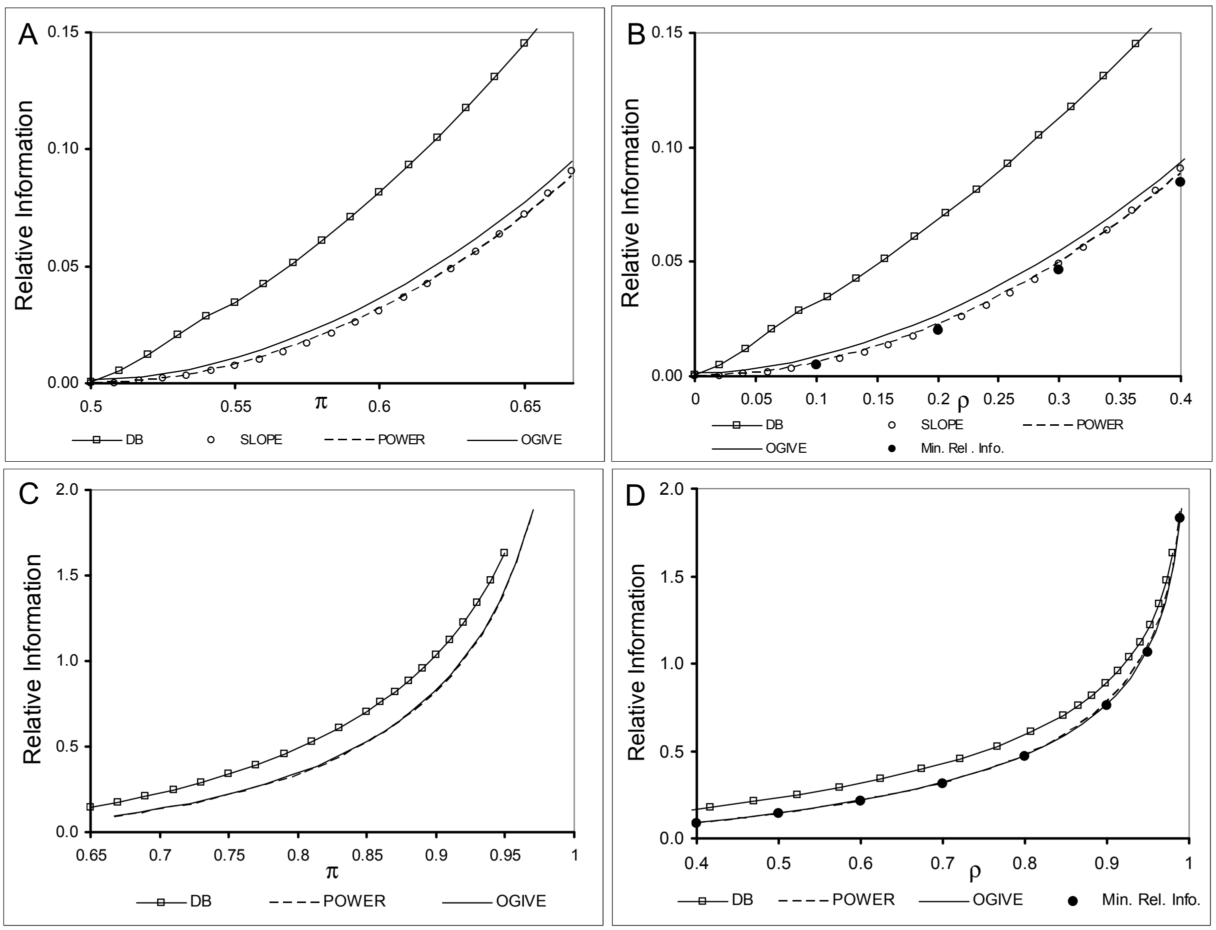

Figure 9. Behavior of relative information of GDB Copula as function of π and ρ

A: $0.5 \le \pi \le \frac{2}{3}$; B: $0 \le \rho \le 0.4$; C: $\frac{2}{3} \le \pi \le 0.95$; D: $0.4 \le \rho \le 0.99$.

From Figure 9 one immediately concludes that GDB copulae with TS slope, power, or ogive generating pdf's outperform the original DB copulae from a relative information perspective. Secondly, from Figures 9A and 9B it follows that for the lower ranges in these figures the slope and power generating pdf's are competitive. Observe from Figure 9B that the relative information values of GDB copulae with these generating pdf's are close to those of Bedford and Meeuwissen's (1997) minimum information copulae. From Figures 9C and 9D it follows that for the higher ranges in these figures the power and ogive generating pdf's display similar results. Indeed, the relative information values for the power and ogive cases in Figure 9D are very close to those obtained for Bedford and Meeuwissen's (1997) minimum information copulae.

6. A DECISION ANALYSIS EXAMPLE

As a matter of illustration, we apply GDB copulae with TS generating densities to a farmer's decision problem (DP) reminiscent of the one presented in Clemen and Reilly (2002) (Problems 5.9 and 12.13, pages 209, 521, respectively). The farmer is faced with protecting his/her crop of oranges (with a total worth of \$50,000) against freezing weather with the objective of minimizing his or her losses. In case the temperature drops below freezing (32 degrees Fahrenheit) he/she will lose the entire crop without protection. The farmer assesses the temperature T that evening to be between 24 and 34 degrees and uniformly distributed in between. Hence, the probability of freezing $Pr(T < 32)$ that evening is assessed at 80%. To protect the crop, the farmer has two alternatives: (1) to use burners with a fixed mobilization cost of \$10,000 or (2) sprinklers with a fixed mobilization cost of \$3,000. Effectiveness of the burning (sprinkler) option is uncertain and the all-in loss $B\,(S)$, including mobilization and crop loss, is assessed by the farmer to vary between $a =$ \$25000 (\$28,000) and $b =$ \$35,000 (\$33,000) with a most likely value of $m =$ \$27,000 (\$29,000). Both B and S are assumed to be triangular distributed with parameters a, m and b, respectively. Recalling that the mean of a triangular distribution equals the arithmetic mean of a, m and b (see, e.g., Kotz and van Dorp (2004), we have $E[B] =$ \$29,000 and $E[S] =$ \$30,000. Hence, due to its lower mobilization cost the sprinkler option follows from Figure 10A as the optimal decision with an expected loss of \$24,600.

Effectiveness of both the burner and the sprinkler options depend on the temperature T that evening. Since the protection of the sprinkler option is based on an insular layer of freezing water on the oranges and the burner option is based on gas usage, effectiveness of the burner option is more

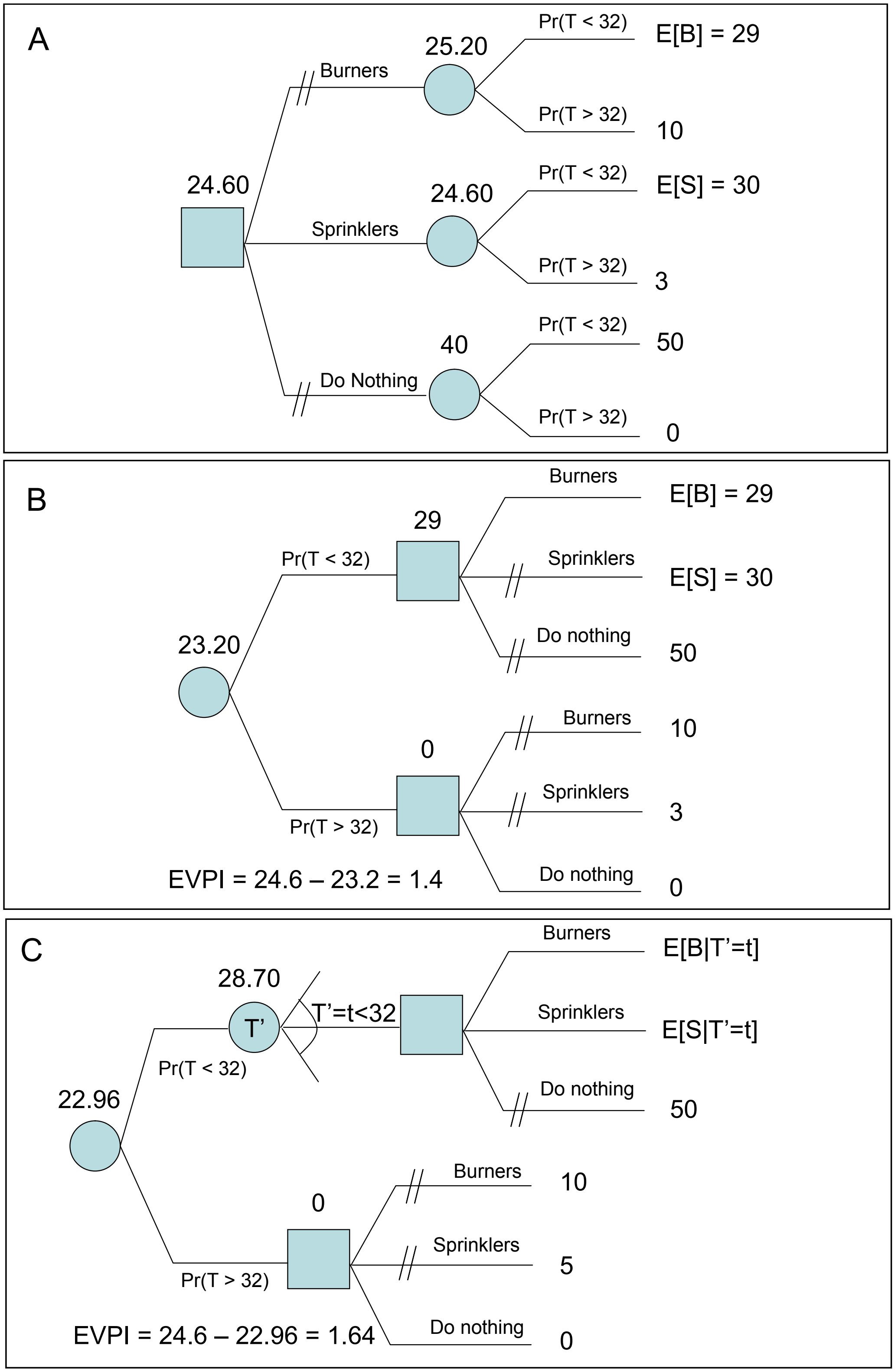

Figure 10. A: A farmer's DP; B: EVPI on "freezing"; C: EVPI on temperature T.

adversely affected by low temperatures than the sprinkler option. The farmer assesses a 90% chance (60% chance) that the burning loss B (sprinkler loss S) is above it median value $b_{0.5}$ ($s_{0.5}$) when the temperature T is below its median value $29F$. Hence, we have:

$$Pr(B < b_{0.5}| T < 29) = 0.1, Pr(S < s_{0.5}| T < 29) = 0.4, \tag{45}$$

where $b_{0.5} \approx$ \$28,675 and $s_{0.5} \approx$ \$29,838. Following the suggestions from Section 5 the dependence between B (S) and T is modeled using a GDB copula with a power (slope) generating density and utilizing (40) we have $n = 1/11$ ($\alpha = 0.4$). Please note that since the probabilities in (45) are less than $1/2$ negative dependence follows between (T, B) and (T, S) consistent with the notion that lower temperatures result in higher losses.

To reduce his losses further, the farmer considers consulting either a clairvoyant Expert A on "freezing" or a clairvoyant Expert B on the temperature T that evening. Recalling $E[B] =$ \$29,000, $E[S] =$ \$30,000 and $Pr(T < 32) = 0.8$, it immediately follows from Figure 10B that the expected value of perfect information (EVPI) for Expert A equals \$1,400. Observe from Figures 10A and 10B that the optimal decision switches to the burner option given the information that ($T < 32$).

The evaluation of the EVPI on the temperature T from Expert B is more complicated due to the dependence between T and B (T and S). The structure for its evaluation is depicted in Figure 10C. Firstly, given a value t for the temperature T, we evaluate $E[B|t]$ using $s = 2500$ realizations using the following steps:

Step 1: $x = \frac{t-24}{34-24}$ (Recall, $T \sim Uniform[24, 34]$)

Step 2: Sample quantile levels $y_i, i = 1, \ldots, s$ from GDB copula with power(n) generating density for B as per Section 2.3, $n = 1/11$.

Step 3: $E[B|t] = \frac{1}{s}\sum_{i=1}^{s} H^{-1}(y_i),\ B \sim$ Triang(\$25,000; \$27,000; \$35,000),

where $H^{-1}(\,\cdot\,)$ is the inverse cdf or quantile function of B. Evaluation of $E[S|t]$ is analogous realizing that $S \sim$ Triang(\$28,000; \$29,000; \$33,000) and a GDB copula with a slope(α) generating density is used in Step 2 with $\alpha = 0.4$. Please note that Figure 10C contains a continuous fan node for

$$T' = (T|T < 32) \sim \text{Uniform}[24, 32], \qquad (46)$$

since $T \sim \text{Uniform}[24, 34]$. Figure 11 plots the behavior of the functions $E[B|t]$ and $E[S|t]$ as a function over the temperature range $t \in [24, 32]$. The size of the hatched area in Figure 11 continuing downwards to a loss-value \$0 and $t \in [24, 32]$, equals

$$E_{T'}(Min\{E[B|T'], E[S|T']\}) \approx \$28{,}700, \qquad (47)$$

which was evaluated by averaging 101 equidistant values of $Min\{E[B|t], E[S|t]\}$ over the temperature range $[24, 32]$. Hence, we obtain from Figure 10C for the EVPI of Expert B \$1640 (\$240 dollars more than the EVPI for Expert A).

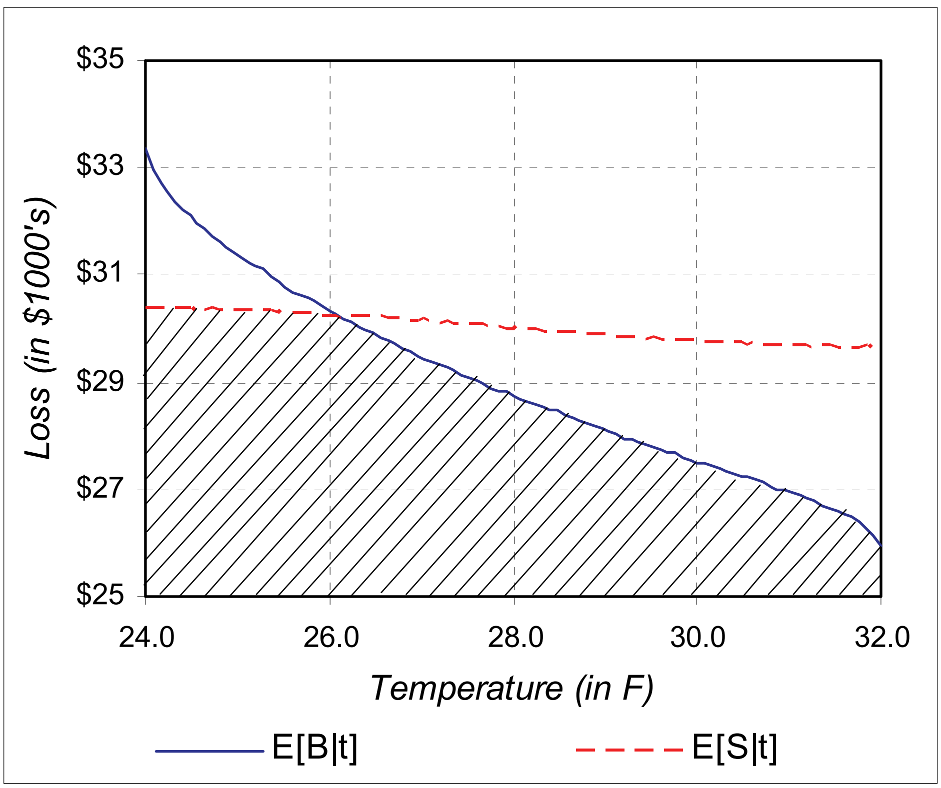

Figure 11. Graphical depiction of the evaluation $E_{T'}(Min\{E[B|T'], E[S|T']\})$.

Summarizing, the farmer is willing to pay \$240 dollars more for perfect information on the temperature T that evening than for the more limited (but perfect) information on whether it will freeze or not. Also observe from Figure 11 that given perfect information on T, operationally the optimal decision switches from the sprinkler option to the burner option at $T \approx 26F$. It is worthwhile to note that by the law of total expectation the total area underneath the solid line curve reduces to $E[B]$, while the total area underneath the dotted curve reduces to $E[S]$. Hence, we visually observe from Figure 11 that $E[B] < E[S]$ which further explains the optimal decision in Figure 10B given $T < 32$. Finally, it is illuminating that in the case of independence between (T, B) and (S, B) that the decision tree in Figure 10C reduces to the one in Figure 10B yielding the same EVPI value of \$1,400 for Expert A currently displayed in Figure 10B.

ACKNOWLEDGEMENTS

We are indebted to Thomas A. Mazzuchi who has been very gracious in donating his time to provide comments and suggestions in the development of this paper. The authors are thankful to the referees of earlier versions of this paper whose valuable comments improved the contents and presentation.

APPENDIX. GDB COPULA PDF CONSTRUCTION WITH A TS GENERATING PDF

In this appendix we shall derive a GDB copula's pdf (with complete unit-square support) in terms of the generating pdf $p(\cdot|\psi)$ of the TS framework of symmetric distributions defined by (4). The derivation follows Bojarski's (2001) construction method for GDB copulae.

Firstly, a bivariate pdf $g(x, y)$ for random variables X, Y is constructed, where X is uniformly distributed on $[0, 1]$ and the conditional pdf $g(y|x)$ has the following form :

$$g\{y|x, p(\cdot|\Psi)\} = f\{x - y|p(\cdot|\Psi)\}, x - 1 \leq y \leq x + 1, \tag{48}$$

where $f\{\cdot|p(\cdot|\Psi)\}$ is defined by (7). From the uniformity of X, (48) and (7) it follows that

$$g\{x,y|p(\,\cdot\,|\Psi)\} = \frac{1}{2} \times \begin{cases} p(1+x-y|\Psi), & -1 < x-y \leq 0, \\ p(1-x+y|\Psi), & 0 < x-y < 1, \end{cases} \tag{49}$$

The construction of the bivariate pdf $g(x,y|n)$ in (49) is demonstrated in Figure 12A for the case that $p(z) = 2z$. For $p(z) = 2z,\ z \in [0,1]$ the pdf (7) reduces to a symmetric triangular distribution with support $[-1,1]$ which may be observed as $g(y|x=0)$ in Figure 12A.

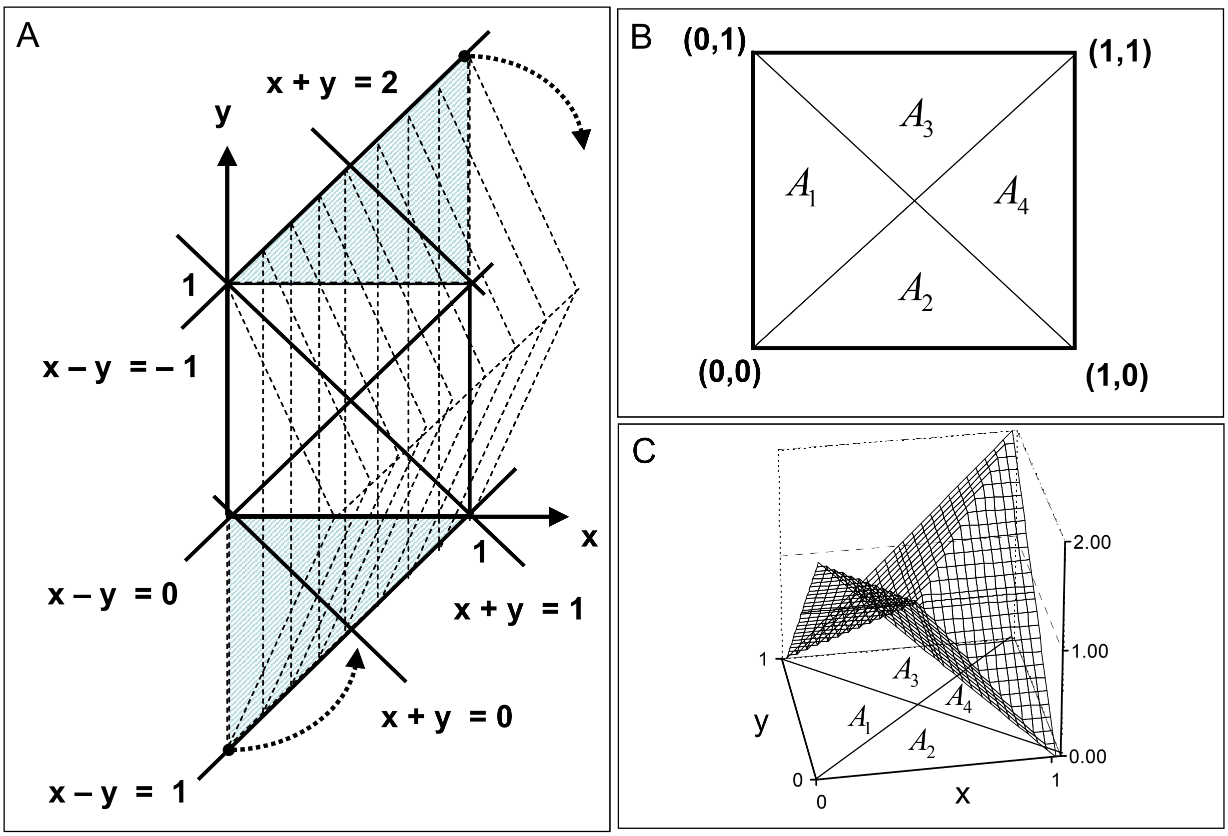

Figure 12. Construction of the GDB copula with TS generating pdf; A: $g(x,y)$ pdf (49); B: Areas A_i given by (10) in Sect. 2; C: $c\{x,y|p(\,\cdot\,|\Psi)\}$ pdf (9) with $p(z) = 2z$ on $[0,1]$.

From (49) we next construct a bivariate pdf $c(x,y\,|\,p(\,\cdot\,|\Psi))$ on the unit-square $[0,1]^2$ by folding back the probability masses of $g\{x,y|p(\,\cdot\,|\Psi)\}$ outside the unit-square $[0,1]^2$ onto it, using "folding" lines $y = 1$ and $y = 0$. See Figure 12A for a graphical depiction of this operation. Hence, we obtain for the relationship between $c\{x,y|p(\,\cdot\,|\Psi)\}$ and $g\{x,y|p(\,\cdot\,|\Psi)\}$ in (49) :

$$c\{x,y|p(\,\cdot\,|\Psi)\} = \begin{cases} g\{x,y|p(\,\cdot\,|\Psi)\} + g\{x,\,-y|p(\,\cdot\,|\Psi)\}, & 0 < x+y \leq 1, \\ g\{x,y|p(\,\cdot\,|\Psi)\} + g\{x,2-y|p(\,\cdot\,|\Psi)\}, & 1 < x+y \leq 2. \end{cases} \tag{50}$$

Combining (50) with (49) now yields the joint pdf $c\{x,y|p(\,\cdot\,|\Psi)\}$ defined by (9) in Section 2 with areas A_i defined by (10). The areas $A_i, i = 1,\ldots,4$ are depicted in Figure 12B and C. An example graph of the resulting bivariate distribution $c\{x,y|p(\,\cdot\,|\Psi)\}$ (11) using the generating pdf $p(z) = 2z$ for the TS framework (7) is provided in Figure 12C.

By design the random variable X in (49) is a uniform distribution on $[0,1]$. The operation exemplified in Figure 12A does not affect the marginal distribution of X and thus the random variable X associated with the bivariate distribution $c\{x,y|p(\,\cdot\,|\Psi)\}$ in (9) in Section 2 is uniformly distributed on $[0,1]$ as well. From the density (9) it follows however that

$$c\{x,y|p(\,\cdot\,|\Psi)\} = c\{y,x|p(\,\cdot\,|\Psi)\} \text{ for all } (x,y) \in [0,1]^2. \tag{51}$$

Hence, following the symmetry argument (51), the random variable Y associated with pdf (9) has to be uniformly distributed on $[0,1]$ as well and one concludes that the bivariate distribution $c\{x,y|p(\,\cdot\,|\Psi)\}$ given by (9) is in fact a copula. This holds regardless of the form of the generating density $p(\,\cdot\,|\Psi)$ of the TS framework of symmetric distributions (7). Moreover, the total probability mass in each of the four areas $A_i, i = 1,\ldots,4$ equals $\frac{1}{4}$ for all copulae with pdf $c\{x,y|p(\,\cdot\,|\Psi)\}$.

REFERENCES

Abbas, A.E. (2006). Maximum entropy utility. *Operations Research*, 54 (2), pp. 277-290.

Abbas, A. E. (2009). Multi attribute utility copulas, Forthcoming in *Operations Research*.

Abbas, A.E. and Howard, R.A. (2005). Attribute dominance utility, *Decision Analysis*, 2 (4), pp. 185 - 206.

Bedford, T. and Meeuwissen, A.M.H. (1997). Minimally informative distributions with given rank correlations for use in uncertainty analysis. *Journal of Statistical Computation and Simulation,* 57 (1-4), pp. 143-175.

Blomquist, N. (1950). On a measure of dependence between two random variables. *Annals of Mathematical Statistics*, 21, pp. 593–600.

Bojarski, J. (2001). A new class of band copulas - distributions with uniform marginals. *Technical Publication,* Institute of Mathematics, Technical University of Zielona Góra.

Cherubini, U., Luciano, E. and Vecchiato, W. (2004). *Copula Methods in Finance*. John Wiley & Sons, e-book.

Clemen, R. and Reilly, T. (1999). Correlations and copulas for decision and risk analysis. *Management Science*, 45, pp. 208-224.

Clemen, R.T. and Reilly, T. (2001). *Making hard decisions with decision tools.* Pacific Grove, CA: Duxbury.

Cooke, R.M. and Waij, R. (1986). Monte carlo sampling for generalized knowledge dependence, with application to human reliability. *Risk Analysis*, 6 (3), pp. 335-343.

De Michele, C., Salvadori, G., Passoni, G. and Vezzoli, R. (2007), A multivariate model of sea storms using copulas, *Coastal Engineering*, 54 (10), 734–751.

Denuit, M., Dhaene, J., Goovaerts, M. and Kaas, R. (2005) *Actuarial Theory for Dependent Risks. Measures, Orders and Models.* Wiley, Chichester.

Embrechts, P. (2009). Copulas: A personal view. *Journal of Risk and Insurance*, 76 (3), pp. 639-650.

Embrechts, P., McNeil, A. and Straumann, D. (2002). Correlation and dependence in risk management: properties and pitfalls. In: *Risk Management: Value at Risk and Beyond*, ed. M.A.H. Dempster, Cambridge University Press, Cambridge, pp. 176-223.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflations. *Econometrica* 50, pp. 987 − 1007.

Ferguson, T.F. (1995). A class of symmetric bivariate uniform distributions. *Statistical Papers*, 36 (1), pp. 31-40.

Frees, E.W., Carriere, J. and Valdez, E.A. (1996). Annuity valuation with dependent mortality. *Journal of Risk Insurance,* 63, pp. 229-261.

Frees, E.W. and Valdez, E.A. (1998). Understanding relationships using copula's. *North America Actuarial Journal,* 2, pp. 1-25.

Frees, E.W. and Wang P. (2005). Credibility using copulas. *North American Actuarial Journal*, 9(2), 31–48.

Genest, C. and Mackay, J. (1986). The joy of copulas, bivariate distributions with uniform marginals. *The American Statistician*, 40 (4), pp. 280-283.

Genest, C. and Favre A.C. (2007). Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering*, 12, pp. 347–368.

Härdle W., Kleinow, T. , Stahl, G. (2002). *Applied Quantitative Finance, Theory and Computational Tools*, Springer, e-book.

He, X. and Gong, P. (2009). Measuring the coupled risks: A copula-based CVAR model. *Journal of Computational and Applied Mathematics,* 223, pp. 1066–1080.

Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman & Hall, London.

Kallen, M.J. and Cooke, R.M. (2002). Expert aggregation with dependence. In, *Proceedings of the 6th International Conference on Probability Safety and Management,* E.J. Bonano, A.L. Camp, M.J. Majors and R.A. Thompson (Eds), pp. 1287-1294.

Kendall, M.G. (1938), A new measure of rank correlation. *Biometrika*, 30, pp. 81-93.

Kotz, S. and Van Dorp, J.R. (2004). *Beyond Beta, Other Continuous Families of Distributions with Bounded Support and Applications*. Singapore, World Scientific Publishing Co.

Kruskal, W. H. (1958). Ordinal measures of association. *Journal of the American Statistical Association*, 53 (284), pp. 814-861

Lewandowski, D. (2005). Generalized diagonal band copulas. *Insurance: Mathematics and Economics,* 37, pp. 49–67.

Lewandowski, D. (2008). *High Dimensional Dependence Copulae, Sensitivity, Sampling*. Dissertation, Delft University of Technology, Wörhman Print Service.

McNeil, A.J., Frey, R. and Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques, Tools,* Princeton University Press, Princeton.

Nelsen, R.B. (1999). *An Introduction to Copulas*. Springer, New York.

Norris, P. M., Oreopoulos, L. , Hou, A. Y., Tao, W.-K. and Zeng, X. (2008). Representation of 3D heterogeneous cloud fields using copulas: Theory for water clouds. *Quarterly Journal of the Royal Meteorological Society*, 134, pp. 1843–1864.

Pearson, K. (1920), Notes on the history of correlation. *Biometrika,* 13, pp. 25-45.

Salmon, F. (2009)"Recipe for disaster: The formula that killed wall street", *Wired Magazine* 17.03, http://www.wired.com/techbiz/it/magazine/17-03/wp_quant/, last accessed: 7/23/09.

Schucany, W.R, Parr, W.C., and Boyer, J.E. (1978). Correlation structure in Farlie-Gumbel-Morgenstern distributions. *Biometrika*, 65 (3), pp. 650-653.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris,* 8, pp. 229-231.

Soofi, E. S., and Retzer, J.J. (2002). Information indices: Unification and applications. *Journal of Econometrics,* 107, pp. 17-40.

Spearman, C. (1904), The proof an measurement of association between two things. *American Journal of Psychology*, 14, pp. 72-101.

Van Dorp, J.R. and M.R. Duffey (1999). Modeling statistical dependence in risk analysis for project networks. *International Journal of Production Economics*, 58., 17-29.

Van Dorp, J.R and Kotz, S. (2003). Generalizations of two sided power distributions and their convolution. *Communications in Statistics: Theory and Methods*, 32 (9), pp. 1703 - 1723.

Yi, W. and Bier, V. (1998). An application of copulas to accident precursor analysis. *Management Science*, 44, S257-S270.