$\mathcal{I}^2\mathcal{SDS}$

The Institute for Integrating Statistics in Decision Sciences

Technical Report TR-2009-10

August 10, 2009

Generalized Diagonal Band Copulae with Two-Sided Generating Densities

Samuel Kotz

Department of Engineering Management and Systems Engineering

The George Washington University

J. R. van Dorp

Department of Engineering Management and Systems Engineering

The George Washington University

Generalized Diagonal Band Copulae with Two-Sided Generating Densities

Samuel Kotz[a] and J. R. van Dorp[a,*],

[a] *Department of Engineering Management and Systems Engineering,*

The George Washington University.

Abstract: Maximum entropy copulae introduced by Bedford and Meeuwissen (1997) provide normative 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 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). Specifically, bivariate copulae density support shall cover the complete unit square when a full support generating density is selected from the two-sided (TS) framework of distributions introduced by Van Dorp and Kotz (2003). Even with unit square support, GDB copulae with TS generating densities allow for a complete rank correlation coverage by taking advantage of the TS framework's flexibility. GDB copulae distribution functions and properties shall be expressed in terms of the TS generating density. Instances shall be presented with closed form expressions utilizing only elementary functions. A straightforward expert judgment elicitation procedure for the GDB copula dependence parameter is suggested. Operationally, for rank correlations ranging from -0.4 to 0.4, GDB copulae with TS Slope generating densities closely approximate the minimal information measure of the maximum entropy copulae. For absolute rank correlation values larger than 0.4, a TS power generating density for GDB copulae performs very well in this regard. For demonstration purposes the sub-family of GDB copulae discussed herein is illustrated in a decision analysis example.

* 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. Any bivariate distribution of (X', Y') with known continuous marginal cumulative distribution functions $G(\cdot)$ and $H(\cdot)$ 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 distribution 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 $\varphi(1) = 0$. Another procedure for constructing copulae uses the geometric method. Nelsen (1999) discusses a variety of methods utilizing some information of a geometric nature.

Recent years have experienced a burst of applications utilizing the copulae approach in the fields of insurance (see, e.g., Frees et. al (1996, 1998, 2005) and finance (see, e.g. Härdle et. al (2002), Cherubini et al. (2004), McNeil et al. (2005), He and Gong (2009)) with Embrechts (2008) even referring to this attention as "the copula craze". For a sampling of other areas where the copulae approach was 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). A particular advantage of this approach 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 in part has resulted in the widespread applications of copulae constructs by "financial quants" whom may be characterized by a certain degree of mathematical sophistication. However as noted by Drouet and Kotz (2001); "*The concepts of dependence permeates the Earth and its inhabitants in a most profound manner. Examples of interdependent meteorological phenomena in nature and interdependence in the medical social, and political aspects of our existence, not to mention the economic structures are too numerous to be cited individually*" . The quote above expresses the need for continuous and increasing modeling and application of dependence between uncertain phenomena in a great variety of fields. Unfortunately, we have not seen a similar level of activity in fields other than finance or insurance.

In this paper, we shall further investigate the recently introduced bivariate family of Generalized Diagonal Band (GDB) copulae whose geometric copula construction method enjoys the ease of the algebraic generator method for copula construction of the Archimedean copulae. The diagonal band (DB) distribution introduced by Cooke and Waij (1986) and displayed in Figure 1 is the founding member of this copulae family and was geometrically constructed by limiting its support to a band of varying width around the unit diagonal. To facilitate the application of the family of GDB copulae, expressions for distribution functions and properties of several instances shall be derived in a closed form. We hope that their ease of use and geometric motivation facilitates the penetration of copulae 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.

Bojarski (2001) generalized the DB copula to a wider and more flexible family of copulae with the same diagonal band support as indicated in Figure 1A and by utilizing a symmetric generating density $f(\cdot)$ with support

$$[-(1-\theta),(1-\theta)]. \tag{1}$$

Setting $f(u)=\{2(1-\theta)\}^{-1}$ Bojarski's (2001) GDB copula reduce to the DB in Figure 1. To retain the sampling efficiency of the original DB copula, a closed form and preferably simple expression for the inverse cdf (or quantile function) of $f(\cdot)$ would be desirable. Bojarski (2001) considered symmetric beta distributions which do not meet that requirement. Lewandowski (2005) showed that Bojarski's generalization of DB copulae are equivalent to the family of copulae introduced by Ferguson (1995) with density

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

where $g(z)$ is a generating probability density function with support $[0,1]$. Ferguson (1995) demonstrated that copulas of the form (2) 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 $g(z)$.

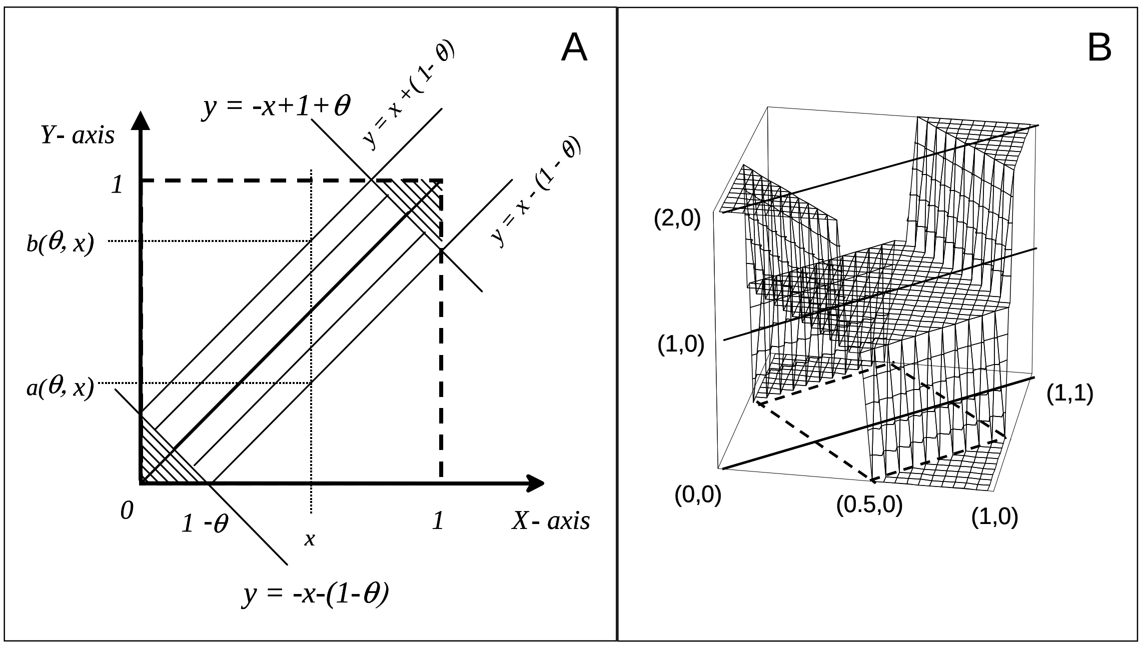

Figure 1: A: Diagonal Band Distribution ($DB(\theta)$) distribution seen from above B: Example of a $DB(0.5)$ distribution.

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 density $p(z)$ concept (but with support $[0,1]$) to define a sub-family of distributions within it. Members within the TS framework of distributions seem to provide a natural candidate for Bojarski's (2001) and Ferguson's (1995) generalizations of DB copulae. In Section 2, we shall follow

Bojarski's (2001) method to construct a sub-class of GDB copulae with a TS generating density and complete unit square copula density support. Its joint cumulative distribution function is derived in terms of the TS framework's generating cdf $P(z)$ and an efficient sampling algorithm for GDB copulae is presented, provided the TS framework's generating density $p(z)$ possesses a closed form quantile function $P^{-1}(z)$. Classical measures of positive dependence are given a decision analytic interpretation in Section 3 and are too expressed in terms of $P(z)$. 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 densities. Entropy measures have been used to design a variety of minimally informative constructs given partial information. Bedford and Meeuwissen (1997) specifically used it to construct maximum entropy copulae given a correlation constraint, which unfortunately are not available in a closed form. Abbas (2006) applied it to the construction of utility functions when only partial preference information is available. In Section 5, different TS generating densities are compared using a GDB's copula entropy in the context of matching an expert's elicited joint GDB probability. For demonstration purposes the use of the sub-family of GDB copulae discussed herein is exemplified in a decision analysis example in Section 6.

2. CONSTRUCTION

Consider the TS framework of symmetric distributions with support $[-1,1]$ and probability density function

$$f\{z|p(\cdot|\Psi)\} = \frac{1}{2} \times \begin{cases} p(z+1|\Psi), & \text{for } -1 < z \leq 0, \\ p(1-z|\Psi), & \text{for } \quad 0 < z < 1, \end{cases} \tag{3}$$

where $p(\cdot|\Psi)$ can be any *generating probability density function* with support $[0,1]$ and the parameters Ψ may in principle be vector-valued. The inverse cumulative distribution function (or quantile function) associated with (2) has the following form

$$F^{-1}\{u|p(\cdot|\Psi)\} = \begin{cases} P^{-1}(2u|\Psi) - 1, & \text{for } 0 < u \leq \frac{1}{2}, \\ 1 - P^{-1}(2-2u|\Psi), & \text{for } \frac{1}{2} < u < 1, \end{cases} \tag{4}$$

where $P^{-1}(\,\cdot\,|\psi)$ is the quantile function of $p(\,\cdot\,|\Psi)$. Hence, sampling from (3) is computationally efficient provided $P^{-1}(\,\cdot\,|\psi)$ is available in a closed form. For example, the Two-Sided Power (TSP) family of distributions follows from (3) by setting $p(z|n) = nz^{n-1}, n > 0$ and allows for efficient sampling utilizing (4) and the quantile function $P^{-1}(q|n) = q^{1/n},\ q \in [0,1]$. TSP distributions have been suggested as a flexible alternative to the classical beta distributions. They are discussed in more detail in Kotz and Van Dorp (2004).

Next, we construct a bivariate distribution $g(x,y)$ utilizing (3) for random variables X, Y, where X is uniformly distribution on $[0,1]$ and the conditional density function $g(y|x)$ has the following form :

$$g\{y|x, p(\,\cdot\,|\Psi)\} = f\{x - y|p(\,\cdot\,|\Psi)\}, x - 1 \leq y \leq x + 1. \tag{5}$$

From the uniformity of X, (5) and (3) it follows that

$$g\{x, y|p(\,\cdot\,|\Psi)\} = \frac{1}{2} \times \begin{cases} p(1 + x - y|\Psi), & -1 < x - y \leq 0, \\ p(1 - x + y|\Psi), & 0 < x - y < 1, \end{cases} \tag{6}$$

The construction of the bivariate density $g(x, y|n)$ (not to be confused with the univariate density $g(\,\cdot\,)$ in (2)) is demonstrated in Figure 2A for the case that $p(z) = 2z$. For $p(z) = 2z,\ z \in [0,1]$ the pdf (3) reduces to a symmetric triangular distribution with support $[-1,1]$.

From (6) we next construct a bivariate density distribution $c(x, y\,|\,p(\,\cdot\,|\Psi))$ 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 2A for a graphical depiction of this operation. Hence, we obtain for the relationship between $c\{x, y|p(\,\cdot\,|\Psi)\}$ and $g\{x, y|p(\,\cdot\,|\Psi)\}$

$$c\{x, y|p(\,\cdot\,|\Psi)\} = \begin{cases} g\{x, y|p(\,\cdot\,|\Psi)\} + g\{x, -y|p(\,\cdot\,|\Psi)\}, & 0 < x + y \leq 1, \\ g\{x, y|p(\,\cdot\,|\Psi)\} + g\{x, 2 - y|p(\,\cdot\,|\Psi)\}, & 1 < x + y \leq 2. \end{cases} \tag{7}$$

Combining (7) with (6) yields

$$c\{x, y|p(\,\cdot\,|\Psi)\} = \frac{1}{2} \times \begin{cases} p(1 - x - y|\Psi) + p(1 + x - y|\Psi), & (x, y) \in A_1, \\ p(1 - x - y|\Psi) + p(1 - x + y|\Psi), & (x, y) \in A_2, \\ p(x + y - 1|\Psi) + p(1 + x - y|\Psi), & (x, y) \in A_3, \\ p(x + y - 1|\Psi) + p(1 - x + y|\Psi), & (x, y) \in A_4. \end{cases} \tag{8}$$

where

$$\begin{aligned} A_1 &= \{(x,y) \in [0,1]^2 | \, 0 < x+y \le 1, \, -1 < x-y \le 0\}, \\ A_2 &= \{(x,y) \in [0,1]^2 | \, 0 < x+y \le 1, \quad 0 < x-y < 1\}, \\ A_3 &= \{(x,y) \in [0,1]^2 | \, 1 < x+y \le 2, \, -1 < x-y \le 0\}, \\ A_4 &= \{(x,y) \in [0,1]^2 | \, 1 < x+y \le 2, \quad 0 < x-y < 1\}. \end{aligned} \tag{9}$$

The areas $A_i, i = 1, \ldots, 4$ are depicted in Figure 2B. An example graph of the resulting bivariate distribution $c\{x, y | p(\cdot | \Psi)\}$ (8) using the generating density $p(z) = 2z$ for the TS framework (3) is provided in Figure 2C.

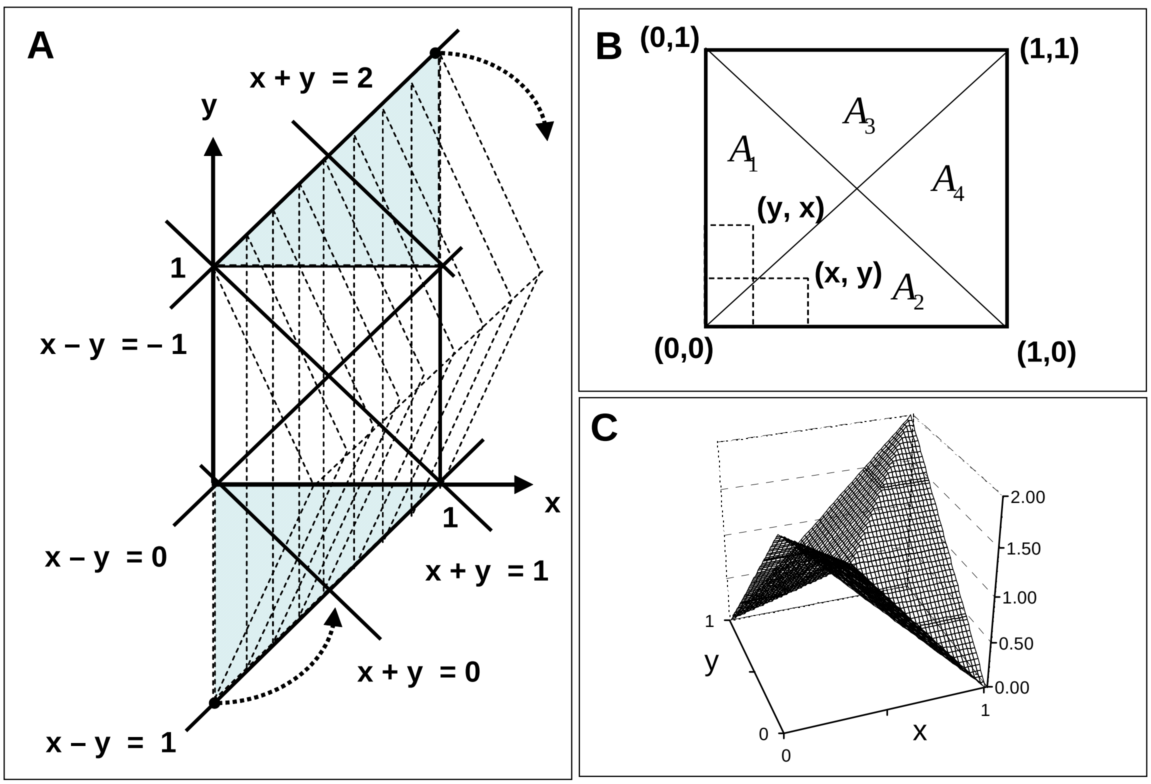

Figure 2. Construction of the TS copula; A: $g(x,y)$ given by (6); B: Areas A_i given by (9); C: $c\{x,y|p(\cdot|\Psi)\}$ given by (8) with $p(z) = 2z$ on $[0,1]$.

By design the random variable X in (6) is a uniform distribution on $[0,1]$. The operation exemplified in Figure 2A 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 (8) is uniformly distributed on $[0, 1]$ as well. From the density (8) it follows that

$$c\{x, y|p(\,\cdot\,|\Psi)\} = c\{y, x|p(\,\cdot\,|\Psi)\} \text{ for all } (x, y) \ \in [0, 1]^2. \tag{10}$$

Hence, following this symmetry argument (10), the random variable Y associated with (8) has to be uniform $[0, 1]$ distributed as well and one concludes that the bivariate distribution (8) 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 (3). 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 density (8). Since (3) reduces to a symmetric triangular density when $p(z) = 2z$, one could refer to the copula in Figure 2C as the *triangular* copula.

2.1. Cumulative distribution function

The joint cumulative distribution function follows directly from (8) as

$$C\{x, y|p(\,\cdot\,|\Psi)\} = \begin{cases} x - \frac{1}{2}\int_{1-x-y}^{1+x-y} P(z|\Psi)dz, & (x, y) \ \in A_1, \\ y - \frac{1}{2}\int_{1-x-y}^{1-x+y} P(z|\Psi)\}dz, & (x, y) \ \in A_2, \\ x - \frac{1}{2}\int_{x+y-1}^{1+x-y} P(z|\Psi)dz, & (x, y) \ \in A_3, \\ y - \frac{1}{2}\int_{x+y-1}^{1-x+y} P(z|\Psi)dz, & (x, y) \ \in A_4. \end{cases} \tag{11}$$

where $P(z|\Psi)$ is the cumulative distribution function of the generating density $p(z|\Psi)$ in (3) and the areas $A_i, i = 1, \ldots, 4$ are provided by (9). 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 (8). This apparent simple connection between (11) and (8), however, belies the effort in verifying (11). An example graph of the resulting bivariate cdf $C\{x, y|p(\,\cdot\,|\Psi)\}$ using the generating density $p(z) = 2z$ (and thus generating cdf $P(z) = z^2$) is provided in Figure 3.

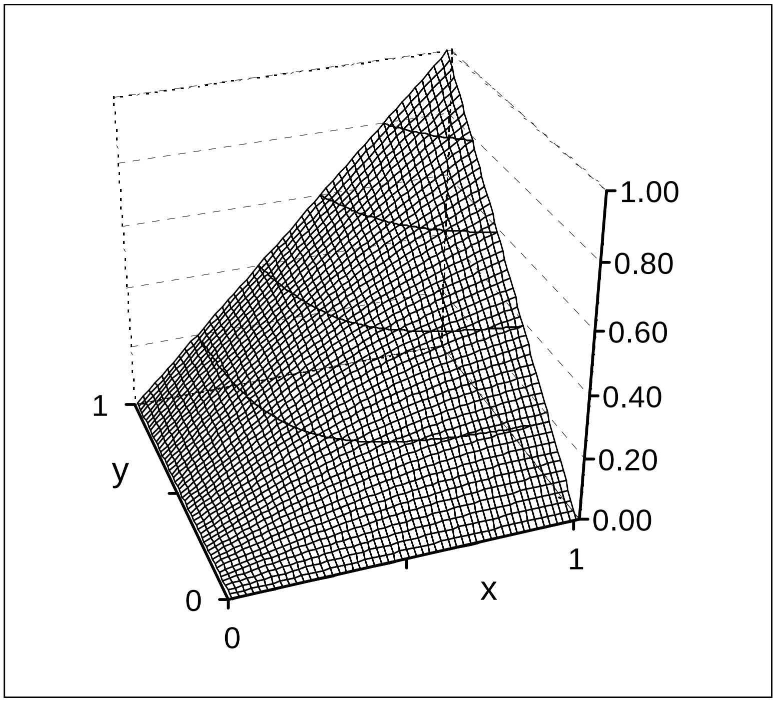

Figure 3. Graph of bivariate triangular copula cdf $C\{x, y|p(\cdot|\Psi)\}$ that follows using generating density $p(z) = 2z$ in (3) and generating cdf $P(z) = 2z$ in (11).

2.2. Dependence parameter elicitation

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

$$\{G(X'), H(Y')\} = (X, Y) \sim C\{x, y|p(\cdot|\Psi)\} \tag{12}$$

where $C\{x, y|p(\cdot|\Psi)\}$ is the bivariate cdf (11) and the generating cdf $P(z|\Psi)$ in (11) 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

$$Pr(Y \leq 0.5 | X \leq 0.5) = \pi \in [0, 1]. \tag{13}$$

From (13) and (11) we have

$$\frac{1}{2}\pi = Pr(Y \leq 0.5, X \leq 0.5) = C\{\frac{1}{2}, \frac{1}{2} | p(\,\cdot\,|\Psi)\} = \frac{1}{2}\int_0^1 \{1 - P(z|\Psi)\}dz \tag{14}$$

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

$$E[Z|\Psi] = \int_0^1 \{1 - P(z|\Psi)\}dz, \tag{15}$$

one arrives with (14) and (15) at the following simple expression

$$\pi = E[Z|\Psi] \tag{16}$$

from which to solve for the dependence parameter Ψ. Summarizing, solving for the dependence parameter ψ is equivalent to solving for the parameter ψ of the generating density $p(z\,|\,\psi)$ using the method of moments. Of course, one can only guarantee a solution to (16) 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 $c\{x, y|p(\,\cdot\,|\Psi)\}$ (8) follows its construction method and is efficient provided the quantile function $P^{-1}(z)$ of the TS framework's (3) generating density $p(z|\Psi)$ is available in a closed form. The following algorithm generates a bivariate sample (x, y) from $c\{x, y|p(\,\cdot\,|\Psi)\}$

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 $y = 1 - (y - 1)$

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 pay-offs of the probability trees in Figures 4A, B and C below, where in Figures 4A, B and C

$$(X_i, Y_i) = \{G(X_i'), H(Y_i')\}, i = 1, \dots, 3 \tag{17}$$

are three independent random bivariate samples from the distribution under consideration. Observe from (17) that $X_i' < Y_i'$ if and only if $X_i < Y_i, i = 1, \dots, 3$. 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:

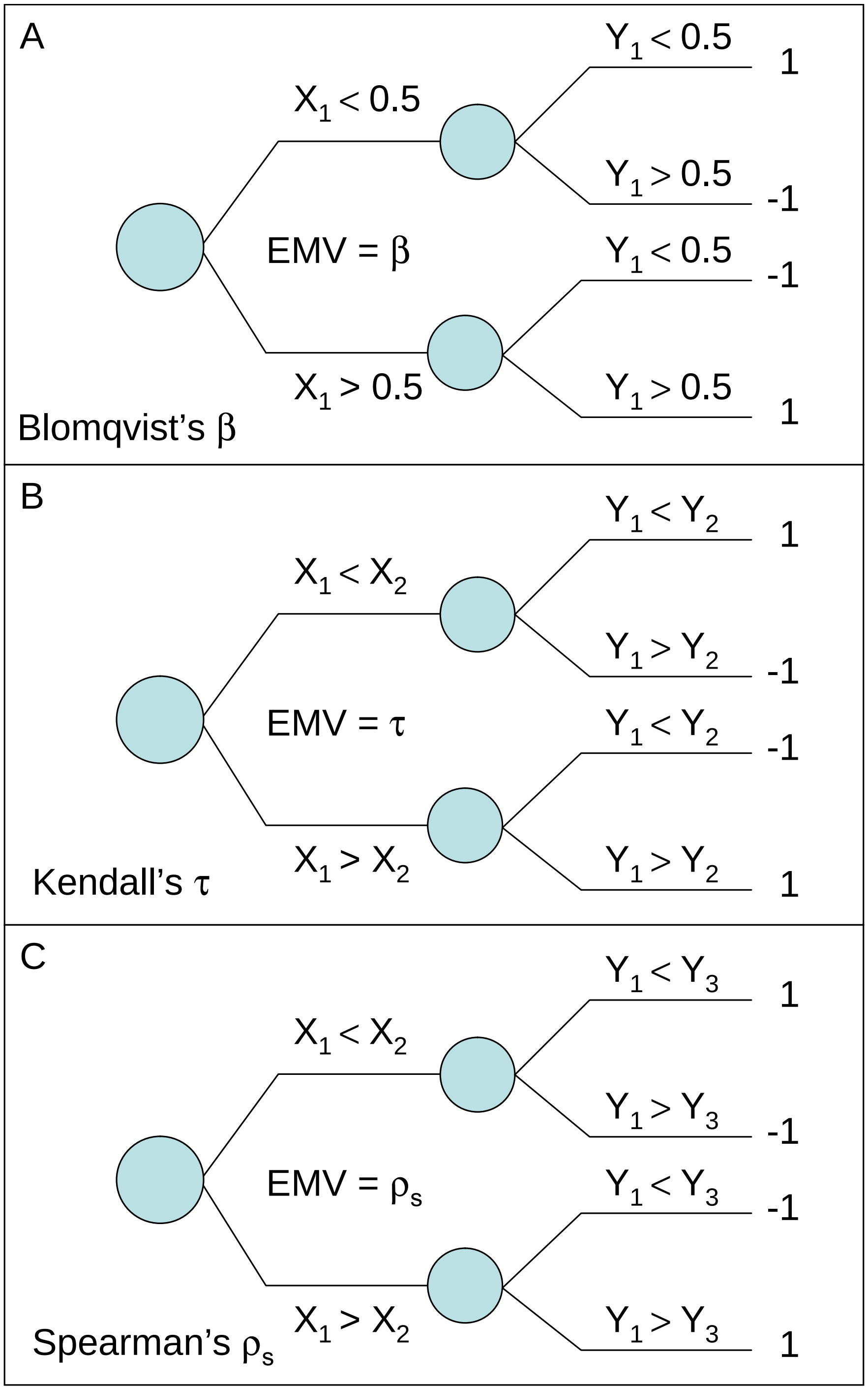

Figure 4. Operational interpretations of ordinal measures of association Blomquist's (1950) β (A), Kendall's (1938) τ (B) and Spearman's (1904) rank correlation ρ_s (C). Samples $(X_i, Y_i) = \{G(X_i'), H(Y_i')\}, i = 1, \ldots, 3$ are independent bivariate samples from a joint distribution under consideration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

$$\begin{cases} \beta(X,Y) = 4C(\frac{1}{2},\frac{1}{2}) - 1, \\ \tau(X,Y) = 4\int_0^1\int_0^1 C(x,y)c(x,y)dxdy - 1, \\ \rho_s(X,Y) = 12\int_0^1\int_0^1 xyc(x,y)dxdy - 3, \end{cases} \tag{18}$$

where $C(x,y)$ and $c(x,y)$ are the joint copula cumulative distribution function and density function of (X,Y), respectively. Hence, one concludes from (18) 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).

Substitution of $C\{x,y\,|\,p(\cdot\,|\Psi)\}$ and $c\{x,y\,|\,p(\cdot\,|\Psi)\}$ given by (11) and (8) in (18) we have after straightforward, but lengthy and tedious, algebraic manipulations that

$$\begin{cases} \beta\{X,Y|p(\cdot\,|\Psi)\} = 2E[Z|\Psi] - 1, \\ \tau\{X,Y|p(\cdot\,|\Psi)\} = 2E[Z^2] - 2\int_0^1 P^2(s|\Psi)ds + 4\int_0^1 sP^2(s|\Psi)ds - 1, \\ \rho_s\{X,Y|p(\cdot\,|\Psi)\} = -4E[Z^3|\Psi] + 6E[Z^2|\Psi] - 1, \end{cases} \tag{19}$$

where $Z \sim P(s|\Psi)$ and $P(s|\Psi)$ is the cumulative distribution functions of the generating probability density $p(\cdot\,|\Psi)$ of the TS framework of distributions (3). We have for $p(z) = 2z$, $E[Z] = \frac{2}{3}$, $E[Z^2] = 1/2$, $\int_0^1 P^2(s)ds = \frac{1}{5}$, $\int_0^1 sP^2(s)ds = \frac{1}{6}$ and $E[Z^3] = 2/5$. Thus utilizing (19) 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 2C.

We have from (19) and (16) 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 a GDB copula and the random variable $V \sim f(v)$:

$$\rho_s\{X,Y|\theta, f(\cdot)\} = 4E[|V|^3|\theta] - 6E[V^2|\theta] + 1, \tag{20}$$

where $f(v)$ is symmetric and has support $[-(1-\theta),(1-\theta)]$ (see (1)). The copula density $c\{x,y|p(\,\cdot\,|\Psi)\}$ (8) is a member within Bojarski's GDB class of copulae by setting $f(v)$ equal to (3) and thus $\theta=0$. Hence, the difference between Bojarski's (2001) larger class of GDB copulae and the sub-class (8) is that a full support $[0,1]$ generating density $p(\,\cdot\,|\Psi)$ yields a GDB copulae 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 copula in Figure 1. Utilizing (3) we have in the case that $\theta=0$

$$\begin{cases} E[V^2|\theta=0,f(\,\cdot\,)]=E[(1-Z)^2|p(\,\cdot\,|\Psi)] \\ E[|V|^3|\theta=0,f(\,\cdot\,)]=E[(1-Z)^3|p(\,\cdot\,|\Psi)], \end{cases} \tag{21}$$

and by substituting (21) into (20) the correlation coefficient $\rho_s\{X,Y|\theta,f(\,\cdot\,)\}$ reduces to expression for ρ_s in (19). Observe the similarity between the expressions for ρ_s in (19) and (20) (with (20) containing the absolute third moment).

3.1. Reflection Property

Let $q(z|\Psi)$ be the density function of $Z'=1-Z$, such that $Z\sim p(z|\Psi)$ and $p(z|\Psi)$ is the generating density of the GDB copula density $c\{x,y|p(\,\cdot\,|\Psi)\}$ given by (8). Hence,

$$q(z|\Psi)=p(1-z|\Psi),\ z\in[0,1]. \tag{22}$$

The density $q(z|\Psi)$ is referred to as the reflection density of the generating density $p(z|\Psi)$. 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 rotating. Figure 5 plots the density $c\{x,y|p(1-z|\Psi)\}$ where $p(z)=2z$ and observe it is a rotated version of the density depicted in Figure 2C.

From (22) and (19) it immediately follows that :

$$\begin{cases} \beta\{X,Y|q(z|\Psi)\}=\beta\{X,Y|p(1-z|\Psi)\}=-\beta\{X,Y|p(z|\Psi)\}, \\ \tau\{X,Y|q(z|\Psi)\}=\tau\{X,Y|p(1-z|\Psi)\}=-\tau\{X,Y|p(z|\Psi)\}, \\ \rho_s\{X,Y|q(z|\Psi)\}=\rho_s\{X,Y|p(1-z|\Psi)\}=-\rho_s\{X,Y|p(z|\Psi)\}, \end{cases} \tag{23}$$

The expression for $-\rho_s\{X,Y|p(z|\Psi)\}$, where $\rho_s\{X,Y|p(z|\Psi)\}$ is given by (19), is identical to the expression for $\rho_s\{X,Y|g(y)\}$ derived by Ferguson's (1995) for copula's (2). Hence, from (23)

it follows that Ferguson's generating density $g(z)$ in (2) enjoys the alternative interpretation of being a generating density of the TS framework of distributions (3).

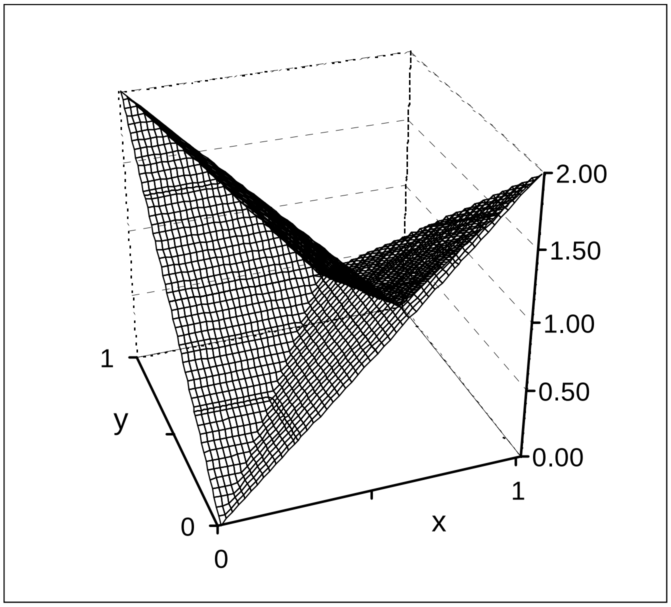

Figure 5. Graph of rotated triangular copula $c\{x, y|p(1-z|\Psi)\}$ using the reflection generating density $p(1-z|\Psi) = 2(1-z)$.

3.2. Lower and upper tail dependence

Recently, 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:

$$\begin{aligned}\lambda_L &= \lim_{x \downarrow 0} Pr\{Y' \leq H^{-1}(x)|X' \leq G^{-1}(x)\} \\ &= \lim_{x \downarrow 0} Pr(Y \leq x|X \leq x) = \lim_{x \downarrow 0} \frac{C(x,x)}{x},\end{aligned} \tag{24}$$

$$\lambda_U = \lim_{x \uparrow 1} Pr\{Y' > H^{-1}(x) | X' > G^{-1}(x)\} \tag{25}$$
$$= \lim_{x \uparrow 1} Pr(Y > x | X > x) = \lim_{x \uparrow 1} \frac{1 - 2x + C(x,x)}{1-x}.$$

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 (24), (25), $C\{x, y|p(\cdot|\Psi)\}$ given by (11) and by applying l'Hopitals rule, we have for GDB copula 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. 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 extremely negative press (and by association the copula approach) and some has 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 its lack of lower and upper tail dependence. We are puzzled by this assessment and we 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 in the time series of financial processes are typically not constant, which eventually lead to the introduction of, amongst others, the Auto-Regressive Conditional Heteroscedastic (ARCH) models by Nobel-Laureate Engle in 1982. Hence, it would seem extremely 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 these types of financial processes require more complex constructs than the use of a single bivariate copula.

4. GDB EXAMPLES WITH TS GENERATING DENSITIES

In this section we shall use the properties in Sections 2 and 3 to study GDB copulas with the following generating densities

$$p(z|\alpha) = 2 - \alpha + 2(\alpha - 1)z,\ 0 \leq \alpha \leq 2, \tag{26}$$

$$p(z|n) = nz^{n-1},\ n > 0, \tag{27}$$

$$p(z|m) = \frac{m+2}{3m+4}\{2(m+1)\sqrt{z^m} - mz^{m+1}\}, m > 0. \tag{28}$$

$$p(z|\theta) = \frac{1}{1-\theta} \times 1_{[\theta,1]}(z), 0 \leq \theta \leq 1, \tag{29}$$

$$p(z|a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\ x^{a-1}(1-x)^{b-1}, a > 0,\ b > 0, \tag{30}$$

Pdf's (26)-(28) 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 (all with full support $[0,1]$). Pdf's (30) is the classical beta distributions also with full support $[0,1]$ whereas pdf (29) may be recognized as a uniform distribution with limited support $[\theta,1]$.

Figure 6 displays GDB copulas and their generating densities (26)-(28) with parameter settings

$$\alpha = 1.5,\ n = 3, m = 4.916. \tag{31}$$

Figure 7 displays GDB copula and their generating densities (29)-(30) with parameter settings

$$\theta = 0.5; a = b = 5. \tag{32}$$

The generating densities in Figures 6A, C, E and 7A all have in common that $E[Z] = 0.75$. For the generating density in Figure 7C we have $E[Z] = 0.5$. Observe from Figure 7A that by reducing the support to $[\theta, 1]$ for the generating density of the TS framework, (8) leads via (29) to the original DB copula displayed in Figure 1. The generating densities $p(z|\psi)$ 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

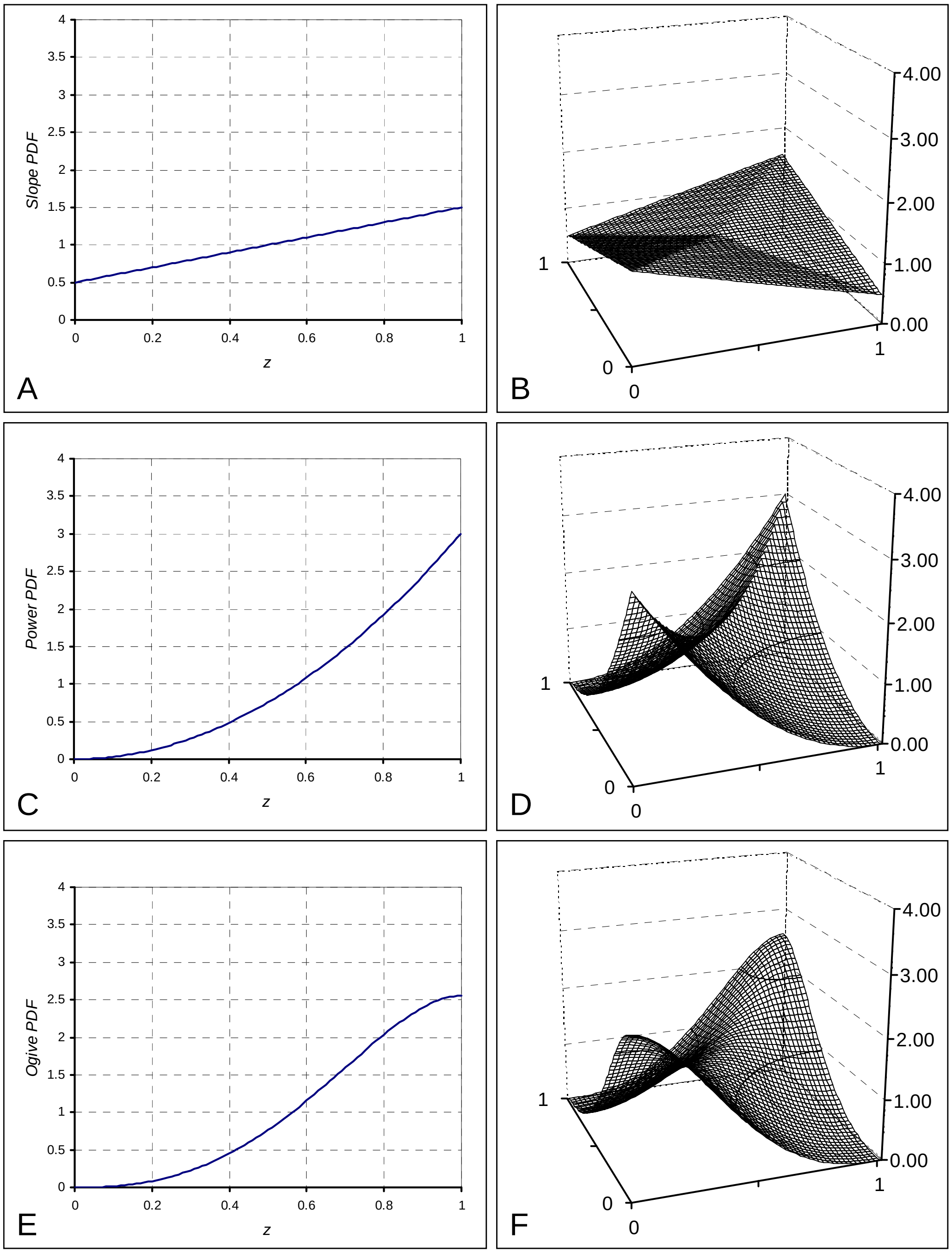

Figure 6. TS framework generating densities with associated GDB Copulae satisfying $E[Z] = 0.75$: A: Slope PDF; B: TS Slope - GDB Copula ($\alpha = 1.5$) C: Power PDF; D: TSP - GDB Copula ($n = 3$); E: Ogive PDF; F: TSO - GDB Copula ($m = 4.916$).

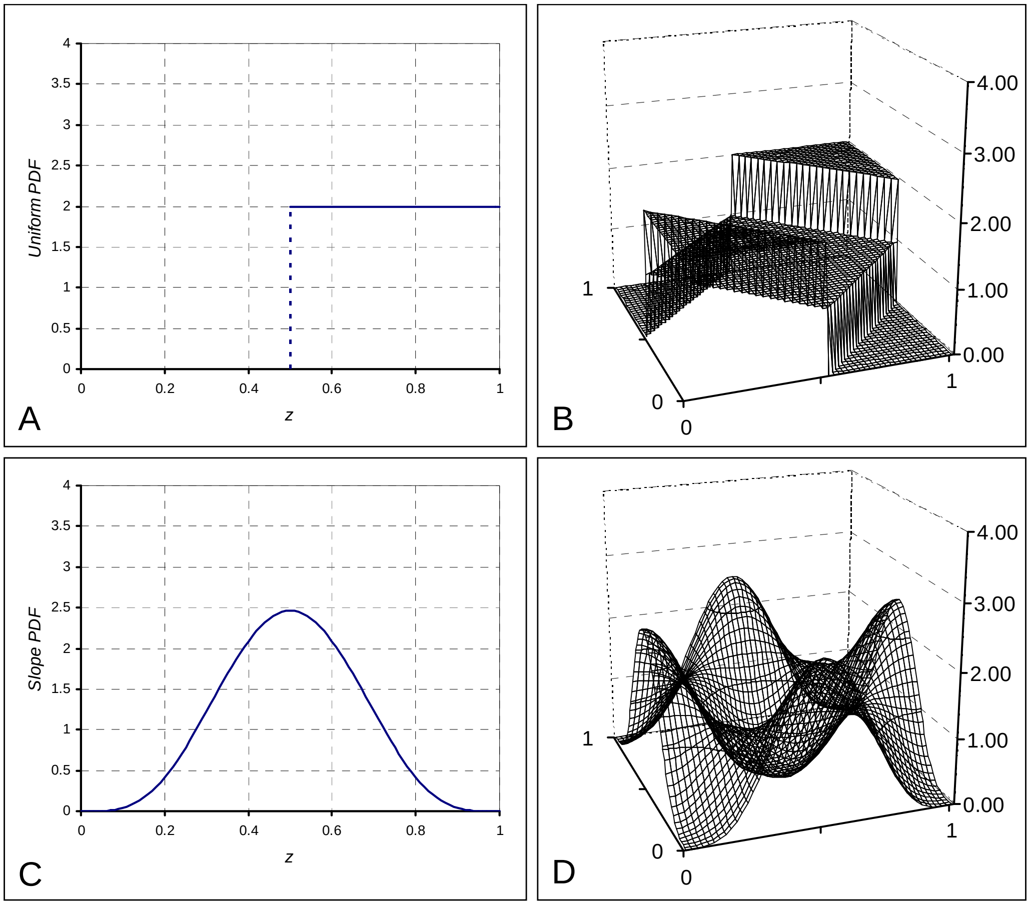

Figure 7. TS framework generating densities with associated GDB Copulae:

A: Uniform $[1-\theta, 1]$ PDF, $E[Z|\theta] = 0.75$; B: DB Copula ($\theta = 0.5$);

C: Beta PDF, $E[Z|a, b] = 0.5$; D: TS Beta - GDB Copula ($a = 5, b = 5$).

the generating densities 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 density 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 density 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 (23) it follows that Blomquist's β, Kendall's τ and Spearman's ρ_s all equal to 0. In other words, any symmetric generating density $p(z|\psi)$ for the TS framework (3) yields uncorrelatedness within a

GDB copula, whereas when $p(z|\psi)$ is non-uniform the variables (X, Y) are clearly not statistically independent (see Figure 7D).

GDB copulas with beta generating densities 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 (26) - (29) for an arbitrary quantile level $q \in (0,1)$:

$$P^{-1}(q|\psi) = \tag{33}$$
$$\begin{cases} \frac{-(2-\alpha)+\sqrt{(2-\alpha)^2+4(\alpha-1)q}}{2(\alpha-1)}, & p(z|\alpha), \alpha \neq 1, \quad \text{Eq. (26)}, \\ q^{1/n}, & p(z|n), \quad \text{Eq. (27)}, \\ \left[\frac{2(m+1)}{m} - \sqrt{\left\{\frac{2(m+1)}{m}\right\}^2 - q\frac{3m+4}{m}}\right]^{2/(m+2)}, & p(z|m), \quad \text{Eq. (28)}, \\ (1-\theta)q + \theta, & p(z|\theta), \quad \text{Eq. (29)}. \end{cases}$$

Thus, generating pdf's (26)-(29) allow for an efficient GDB copula sampling algorithm since their quantile functions are available in a closed form. Perhaps one could slightly favor densities (27) and (29) since their quantile functions require the least number of elementary operations for their evaluation.

We have from (23) for Blomquist β, Kendall's τ, Spearman's ρ_s for the generating densities (26) - (28), respectively:

$$\begin{cases} \beta\{X, Y|p(\cdot|\alpha)\} = -\frac{1}{3} + \frac{1}{3}\alpha, & \in [-\frac{1}{3}, \frac{1}{3}], \\ \tau\{X, Y|p(\cdot|\alpha)\} = -\frac{4}{15} + \frac{4}{15}\alpha, & \in [-\frac{4}{15}, \frac{4}{15}], \\ \rho_s\{X, Y|p(\cdot|\alpha) = -\frac{2}{5} + \frac{2}{5}\alpha, & \in [-\frac{2}{5}, \frac{2}{5}], \end{cases} \tag{34}$$

$$\begin{cases} \beta\{X, Y|p(\cdot|n)\} = \frac{n-1}{n+1}, & \in [-1, 1], \\ \tau\{X, Y|p(\cdot|n)\} = \frac{n-1}{n+2} + \frac{n-1}{(n+1)(n+2)(2n+1)}, & \in [-1, 1], \\ \rho_s\{X, Y|p(\cdot|n)\} = \frac{(n-1)(n+6)}{(n+2)(n+3)}, & \in [-1, 1], \end{cases} \tag{35}$$

$$\begin{cases} \beta\{X, Y|p(\cdot|m)\} = \frac{m(m+1)(3m+8)}{(m+3)(m+4)(3m+4)}, & \in [0, 1], \\ \tau\{X, Y|p(\cdot|m)\} = \frac{m(m+1)(162m^6+2643m^5+18132m^4+66108m^3+140032m+58880)}{(m+3)(m+4)(m+6)(2m+5)(3m+4)^2(3m+8)(3m+10)}, & \in [0, 1], \\ \rho_s\{X, Y|p(\cdot|m)\} = \frac{m(m+1)(3m^3+70m^2+424m+736)}{(m+4)(m+5)(m+6)(m+8)(3m+4)}, & \in [0, 1], \end{cases} \tag{36}$$

and

$$\begin{cases} \beta\{X,Y|p(\cdot|\theta)\} = \theta, & \in [0,1], \\ \tau\{X,Y|p(\cdot|\theta)\} = \theta(\theta+2)/3, & \in [0,1], \\ \rho_s\{X,Y|p(\cdot|\theta)\} = \theta(1+\theta-\theta^2), & \in [0,1]. \end{cases} \tag{37}$$

Figure 8A, B, C and D provides a comparison of β, τ and ρ_s for expressions (34) - (37).

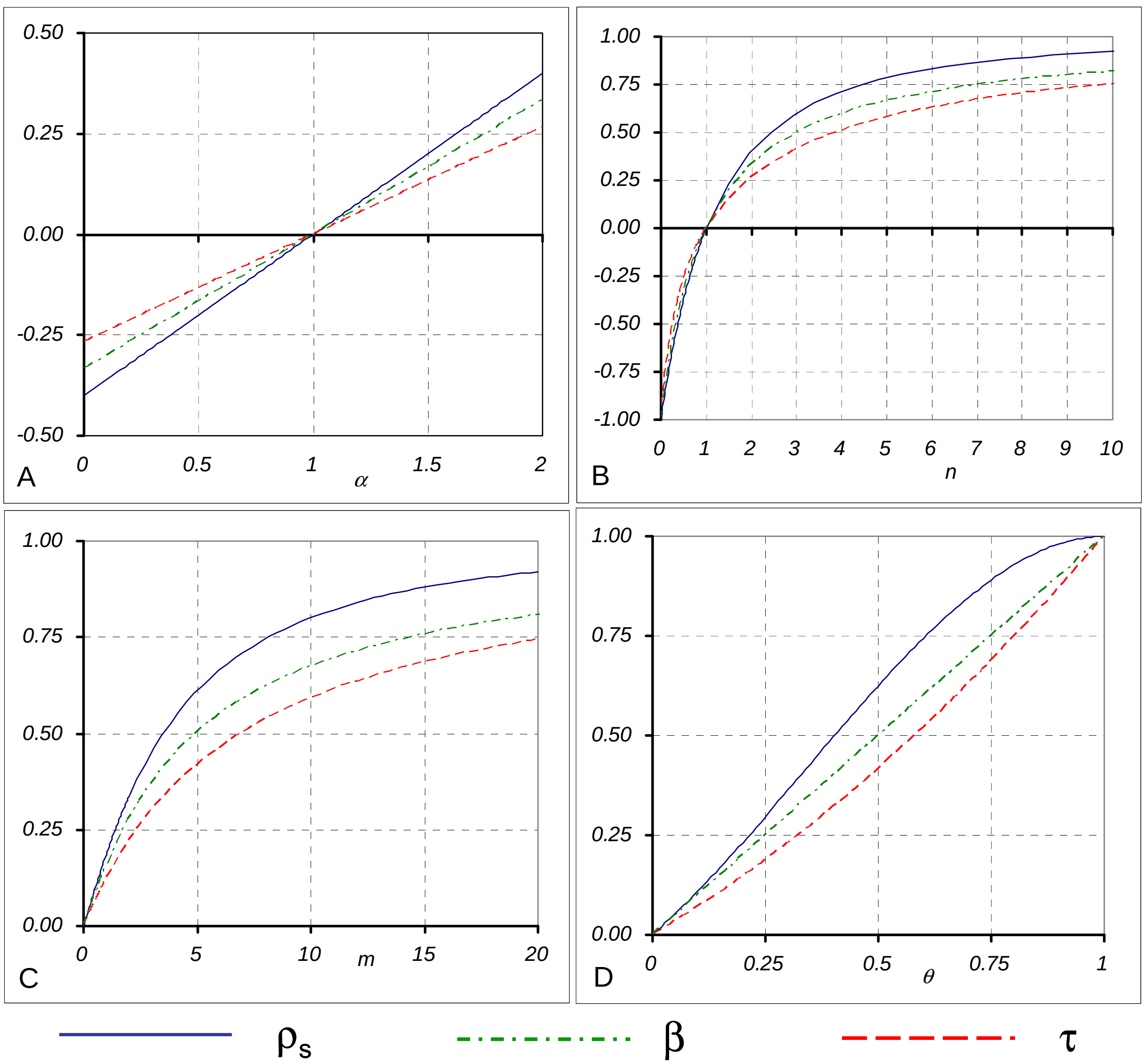

Figure 8. Comparison of ordinal measures of association (34) — (37) for GDB copulae with TS framework generating densities; A: Slope(α); B: Power(n); C: Uniform[θ, 1]; Ogive(m).

Observe from (35) and Figure 8B that GDB copulae with a TS power(n) generating density (27) allow for a complete coverage of β, τ and ρ_s. To achieve a full coverage for the generating densities (28) and (29) one would have to utilize the reflection property (see, Section 3.1) of GDB copulae. The slope generating density (26) only allows for a limited coverage of β, τ and ρ_s, but still larger than the coverage, for example, of $\rho_s \in [-\frac{1}{3}, \frac{1}{3}]$ for the Fairly-Gumbel-Morgenstern (FGM) copulae (see, Schucany et. al (1978)). Observe from Figure 8A, B, C and D that $\tau < \beta < \rho_s$ for all parameter values. Kruskal's (1958) paper, however, contains examples with a reversed orderings of these measures of association. We invite the reader to show that the ordering $\tau < \beta < \rho_s$ applies to all GDB copulae with TS generating densities, which is still an open question. Finally, observe from Figure 8A and expressions (34) 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 (34), (35) and (37) involves simple algebraic manipulations, whereas higher order expressions in (35), (36) and (37) 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 density that matches this constraint. We have from (26)-(29), (13) and (16) that

$$\pi\{X, Y|p(\cdot|\Psi)\} = \begin{cases} (2+\alpha)/6 \in [\frac{1}{3}, \frac{2}{3}], & p(z|\alpha), \quad \text{Eq. } (26), \\ n/(n+1) \in [0,1], & p(z|n), \quad \text{Eq. } (27), \\ \frac{(m+2)^2}{3m+4}\left[\frac{3m+6}{(m+4)(m+3)}\right] \in [0.5, 1], & p(z|\theta), \quad \text{Eq. } (28), \\ (\theta+1)/2 \in [0.5, 1], & p(z|m), \quad \text{Eq. } (29). \end{cases} \tag{38}$$

From (38) 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 densities (27) - (29), respectively. The corresponding copula densities are displayed in Figures 6D, 6F and 7B. For $m > 2$ it follows from the ogive pdf (28) 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 copulas in Figures 6D and 7B are not. Hence, if smoothness of the copula were a requirement one could favor the ogive generating density. For $m > 2$ one obtains for ogive pdf (28) 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 (35), (36) and (37) that for $n = 3, m = 4.916$, $\theta = 1/2$, respectively:

$$\rho(n) = \frac{3}{5},\ \rho(m) = 0.6059 \text{ and } \rho(\theta) = \frac{5}{8}. \tag{39}$$

Hence, this would favor the power(n) generating density, albeit ever so slightly. However, this raises the question if 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 density $p(z|\Psi)$. 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 copula with generating densities (27)-(29) the one that minimizes a distance measure between it and the copula with independent uniform marginals.

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

$$I(f|g) = \int\int f(x,y) \ln\{f(x,y)/g(x,y)\} dxdy. \tag{40}$$

The quantity (40) 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 (40), where $u(x,y)$ is the density on $[0,1]^2$ with independent uniform marginals, (40) reduces to:

$$I(c|u) = \int\int_{S_c} c(x,y)\ln\{c(x,y)\}dxdy, \tag{41}$$

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 density $c(x,y)$. The measure $-E = I(c|u) \geq 0$ measures the information imbedded within $c(x,y)$ relative to the uniform density $u(x,y)$.

Hence, given a particular constraint imposed on a sub-family of copulae, one could select that copula that is least informative by minimizing (41) or equivalently maximizing its entropy. Bedford and Meeuwissen (1997) specifically used the relative information (41) 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 copulas in Figures 6B, C and 7B and F and parameter settings (39), respectively,

$$I\{c(x,y)|p(\,\cdot\,|\psi)\} = \begin{cases} 0.2136 & p(z|n), n = 3, \\ 0.2222 & p(z|m), m = 4.916. \\ 0.3400 & p(z|\theta), \theta = 0.5, \end{cases} \tag{42}$$

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 TSP generating density with $n = 3$.

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 $0 \le \pi \le \frac{2}{3}$ ($0 \le \rho \le 0.4$) which coincides with the restricted range associated for the slope generating densities. Thus, a slope generating density analysis is not included in Figures 9C (9D) since it deals with $\frac{2}{3} \le \pi \le 0.95$ ($0.4 \le \rho \le 0.99$). Figures 9B and 9D also include combinations of correlations and relative information values for Bedford and Meeuwissen's (1997) minimal information copulas. 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.

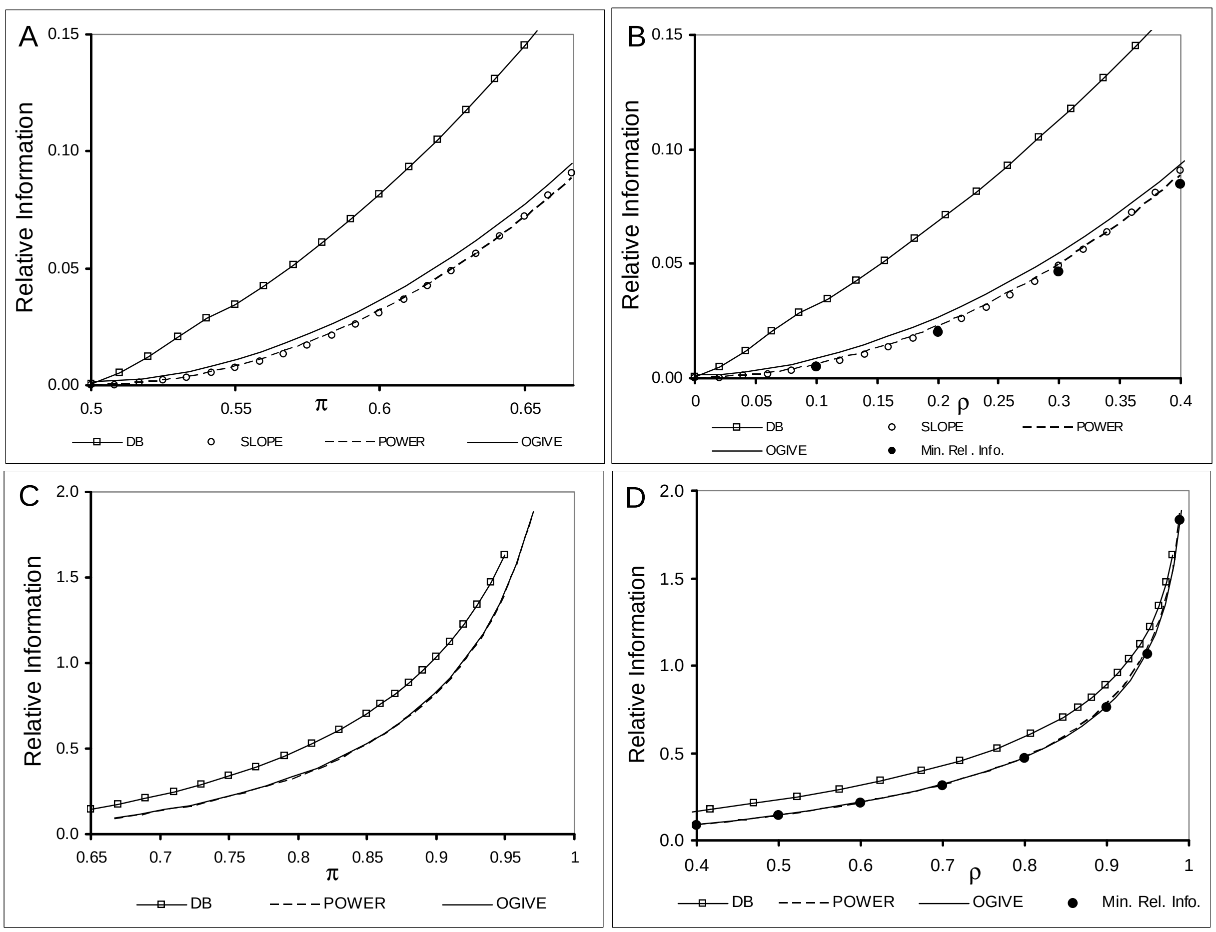

Figure 9. Behavior of relative information of GDB Copula as function of π and ρ

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 densities 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 densities are competitive. Observe from Figure 9B that the relative information values of GDB copulae with these generating densities 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 densities display similar results. Indeed, the relative information values for the power and ogive cases in Figure 9D are very close to the ones 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 loose 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 follows of $24,600.

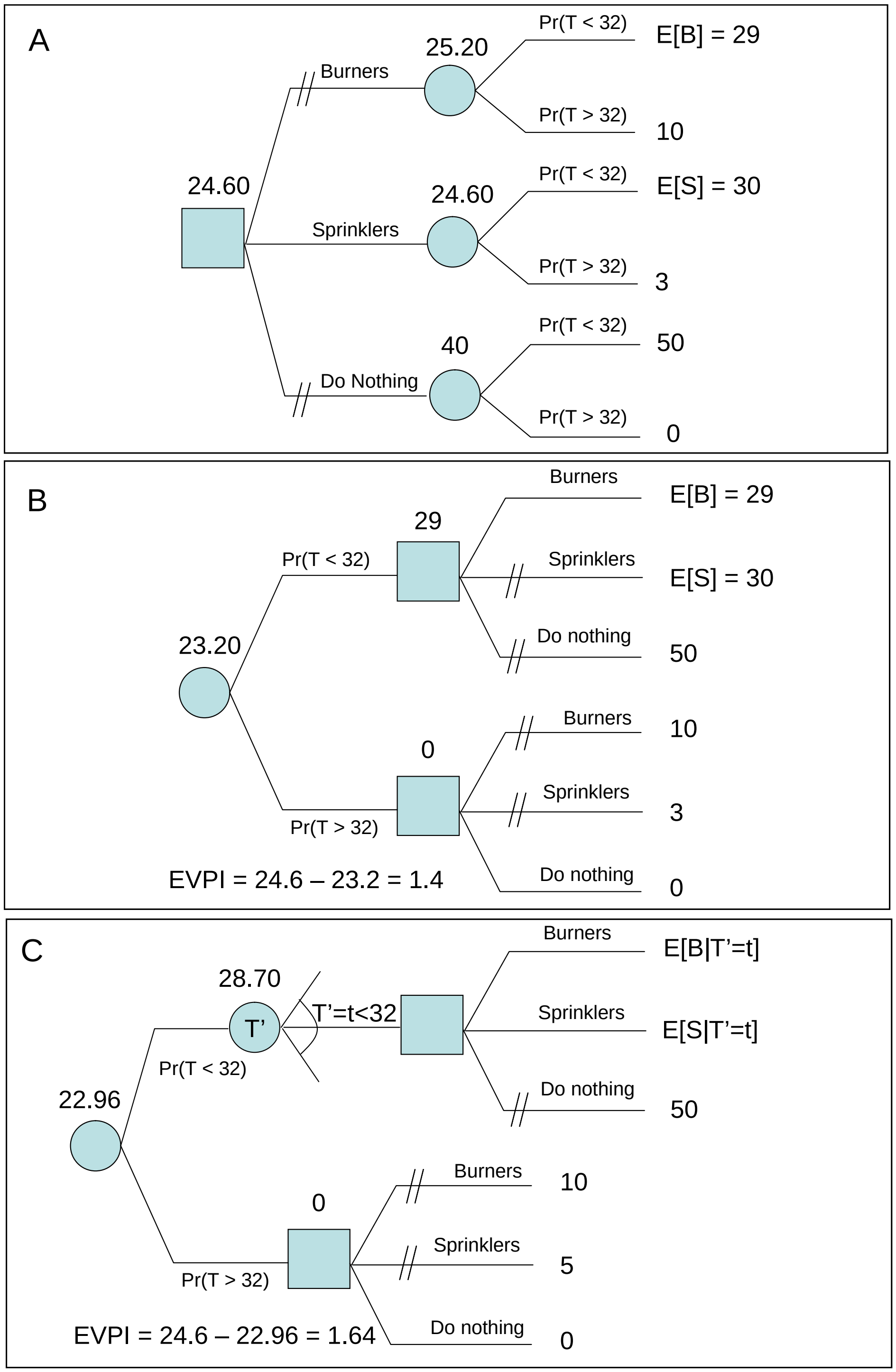

Figure 10. A: A farmer's DP; B: EVPI on "freezing"; C: EVPI on temperature *T*.

Effectiveness of both the burner and the sprinkler options depends 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 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:

$$Pr(B < b_{0.5} | T < 29) = 0.1, Pr(S < s_{0.5} | T < 29) = 0.4, \tag{43}$$

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 (38) we have $n = 1/11$ and $\alpha = 0.4$, respectively. Please note that since the probabilities in (43) 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 $E[B|t]$ and $E[S|t]$ 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 2.1, $n = 1/11$.

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 $slope(\alpha)$ generating density is used in Step 2 with $\alpha = 0.4$. Note that, Figure 10C contains a continuous fan node for

$$T' = (T|T < 32) \sim Uniform[24, 32], \tag{44}$$

since $T \sim Uniform[24, 34]$. Figure 11 plots the behavior of the functions $E[B|t]$ and $E[S|t]$ as a function of the temperature $t < 32$. The size of the hatched area in Figure 11 equals

$$E_{T'}(Min\{E[B|T'], E[S|T']\}) \approx \$28{,}700, \tag{44}$$

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).

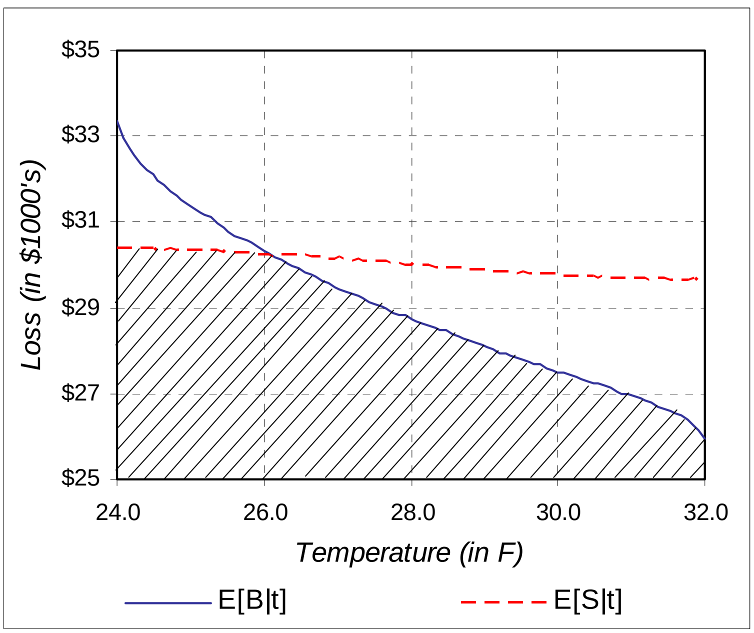

Figure 11. Graphical depiction of the evaluation $E_{T'}(Min\{E[B|T'], E[S|T']\})$.

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 an earlier version of this paper whose valuable comments improved the current contents and presentation.

REFERENCES

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

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.

Drouet, D. and Kotz, S. (2001). *Correlation and Dependence.* Imperial Press, Singapore.

Embrechts, P. (2008). Copulas: A personal view. To appear in *Journal of Risk and Insurance.*

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.