HEDG
Health, Econometrics and Data Group

THE UNIVERSITY *of York*

HEDG Working Paper 07/22

The health-economic applications of copulas: methods in applied econometric research

Casey Quinn

October 2007

ISSN 1751-1976

The health-economic applications of copulas: methods in applied econometric research

Casey Quinn

University of York, York YO10 5DD, England

Lehigh University, Bethlehem, PA 18015, USA

October 22, 2007

Abstract

A copula is best described, as in Joe (1997), as a multivariate distribution function that is used to bind each marginal distribution function to form the joint. The copula parameterises the dependence between the margins, while the parameters of each marginal distribution function can be estimated separately.

This is a brief introduction to copulas and multivariate dependence issues within a health economics context. The research presented here will make its own contributions to the development of copulas as a methodology, but more importantly will make deliberate inroads into health economic applications of copulas. To do this, common analytic problems faced by health economists are considered. Some of the differences between the copula methodology and existing alternatives are discussed, and a generalisable, systematic approach to estimation is provided.

JEL classification: C1, C3, C5, I3, I10

1 Introduction

This is an introduction to copulas from the perspective of the needs of health-economic analysis. The presentation is intended to be non-technical, although copulas are, as constructed multivariate distribution functions, intrinsically technical. Important but non-necessary technical discussion is removed to footnotes, and can be overlooked. Technical discussion included in the body of the paper can also be overlooked: referencing is extensive, to enable more detailed research for a given item or characteristic of interest.

Essential references are Joe (1997) and Nelsen (1999, 2006). Frees and Valdez (1998) is particularly useful also, for its discussion and inclusion of an annotated bibliography. Since this is an exercise in applying copulas in a health economic context, only those characteristics most relevant to empirical applications will be discussed. The presentation will be stylistic rather than statistical, as much as is possible: the discussion of copulas is structured according to these characteristics, with statistical properties of copulas explained as they become relevant. Much of the statistical foundations of copulas will be overlooked, but can be found in the comprehensive studies by Dall'Aglio (1991), Nelsen (2006) and Joe (1997).

1.1 The motivation for copulas

Consider that economic and econometric analyses are, generally, based in assumptions. These can be explicit assumptions generated by prior information about the data being investigated, such as the use of models for discrete or continuous data, or only strictly positively defined distribution functions, for example. Others however are implicit assumptions of convenience. These can be assumptions made for analytic convenience, such as the normalisation to 1 of a standard deviation in a theoretical optimisation problem, to enable identification of the solution. Alternatively, it could be for computational convenience, such as the normalisation to 1 of standard deviation.

The primary motivation for introducing copulas is redressing the assumptions made in a specific context: the analysis of jointly-determined, or jointly-dependent, random variables. When analysing a single random variable there are many univariate distributions that can be used. There are fewer bivariate distributions however, for two correlated random variables, and fewer still in higher dimensions. When two or more dependent variables are jointly determined, the set of available distributions becomes even more restricted, leading to simplifying assumptions, such as the assumption of bivariate normality. However the analytic sequelae of assuming bivariate or multivariate normality can be extensive, though not usually acknowledged explicitly. This includes, for example, when manipulation of data is consistent only with symmetric distributions. This is not problematic if data is in fact joint normally distributed, or at least symmetric, but can become a problem otherwise. The fact that univariate normality may appear to be robust enough under reasonable mis-specification may not be sufficient to overcome errors when using non-normally, asymmetrically distributed data in higher dimensions, the central limit theorem notwithstanding (Keselman, *et al.* 2005).

A secondary motivation is mis-specification. Greene (2002) and Wooldrige (2002) discuss the use of pseudo-maximum likelihood when the marginal distributions are known but the joint is not. Pseudo-maximum likelihood techniques are a way of overcoming the need to specify a distribution correctly for methods of maximum likelihood, however in certain cases this will result in inconsistent estimates of some or all parameters (Greene 2002: specifying a Poisson incorrectly, in place of a Negative Binomial distribution, is one such case). Prokhorov and Schmidt (2006) discuss quasi-maximum likelihood estimation, the method used when independence is assumed, rather than some (potentially misspecified) joint distribution. They show that, when independence has been assumed incorrectly, estimates may be inconsistent also. In both cases corrections exist, but not in all; nor are they commonly undertaken in empirical analyses. The methods shown here obey rules in Prokhorov and Schmidt (2006) that, in the absence of these problems, the added moment conditions they include are only redundant, but in the presence of such problems provide consistent estimates and more efficient estimation overall.

A final motivation is redressing the other common concession made to analysing jointly-distributed random variables conveniently. That is, the association between them. There are many measures of dependence in statistics, some more robust than others. Usually, though, weaker measures of dependence are used in order to enable more straightforward estimation of parameters belonging to the margins themselves. This is related to the previous point: often these weaker measures of association are enabled by prior simplifying assumptions about the joint distribution, such as symmetry.

Simplifying assumptions generally are concessions to practicality, and they do not occur any more frequently in health economics than in economics or econometrics generally. Moreover other approaches exist, which either offset distributional asymmetry/dissimilarity, or avoid it altogether. Conditional likelihood is an example of the former, which is discussed here as it relates directly the motivation of this dissertation. The latter includes methods such as empirical likelihood and bootstrap/jackknife (Owen 2001; Shao and Tu 1995; Chernick 1999). These are not discussed because, although they overcome the need to specify any distribution, they are not generally practical for regression-based estimation, particularly in higher dimensions.[1]

1.2 Sklar's theorem

For univariate marginal distribution functions $F_1(x_1)$ and $F_2(x_2)$, a copula is a function that binds those margins precisely, to form the multivariate distribution function (Smith 2003). The copula parameterises the dependence between the margins, while the parameters of each marginal distribution function can be estimated separately. For the purposes of empirical analysis a copula is best described, as in Joe (1997), as a multivariate distribution function that separates each marginal distribution both from every other marginal distribution, and from the dependence between their associated random variables. Thus the two most important features of copulas: they exist as multivariate distribution functions which can feasibly contain any type and combination of marginal distributions; and each uniquely represents dependence. Depending on the functional form used, association of quite different types can either be assumed or tested, independently of the functional forms of the marginal distributions used.

By a theorem due to Sklar (1959) one can say that all multivariate distributions have a copula representation, in which each margin is invariant to transformations in every other margin, or independent of the choice of every other marginal distribution.[2] Consider two random variables X_1, X_2 with bivariate distribution function $H(x_1, x_2) = \Pr(X_1 \leq x_1, X_2 \leq x_2)$ and univariate marginal distributions $F_1(x_1)$ and $F_2(x_2)$ respectively. Then there exists a copula C that represents the joint distribution function in terms of the margins, such that

$$H(x_1, x_2) = C(F_1(x_1), F_2(x_2)) \tag{1}$$

for all real values of x_1, x_2 (or $(X_1, X_2) \in \mathbb{R}^2$). If F_1, F_2 are continuous, C is unique. Under discontinuity C is uniquely determined on its domain, the range of the margins $\mathrm{Ran}F_1 \times \mathrm{Ran}F_2$.[3] Moreover it can be seen using Sklar's theorem that, if C is a copula and F_1 and F_2 are distribution functions, then some function H as defined in Equation (1) is a joint distribution function (see Nelsen 2006 for this proof, as well as an explanation of quasi-inverses of non-strictly increasing margins, which can also be used to construct a copula). By taking the marginal distribution functions as explanators within which association is not contained, the copula separates the explanation of X_1 and X_2 from their association.

1.3 Background

Although copulas have appeared in the mathematical and statistical literature for several decades, in empirical studies they have only begun to be used relatively recently (Dall'Aglio 1991 contains an excellent analect of many of the developments in copulas). The earliest development of copulas empirically has been in finance. They have proven useful in non-parametric analyses of risk and asset returns, where covariate explanation is of little interest but precise estimation of the dependence is. Bouyé, *et al.* (2000) provide a comprehensive review of copulas for use in several methods of estimation beyond that of maximum-likelihood, exemplifying the approach with analysis of credit scoring, risk measurement and asset returns.

Sample selection is one more broadly economic area in which the majority of analyses begin with the normality assumption, even while acknowledging the limitations it imposes (Vella 1998). Smith (2003) uses a specific class of copulas to estimate self-selection, switching-regime and double-selection models, analysing labour supply and hospital lengths of stay. This is done to show the different properties of data, of which a copula approach can take advantage. Genius and Strazzera (2004) similarly used copulas to examine contingent value data with selection bias comparatively with standard Full Information Maximum Likelihood approaches, which assume bivariate normality.

Statistically, copulas are stronger with continuous random variables, although several studies now have applied them also to discrete data. Researchers at Université Catholique de Louvain's Institut de Statistique have successfully applied copulas to the analysis of categorical data from clinical trials (Vandenhende and Lambert 2000; Tajar, *et al.* 2001), as well as looking specifically at correlation, local and tail dependence (Almeida and Mouchart 2003; Cebrián, *et al.* 2003; Demarta and McNeil

2005; Vandenhende and Lambert 2004). Dardanoni and Lambert (2001) use copulas to analyse stochastic ordering. Quinn (2007c) applies them similarly to the analysis of income-related inequalities in health, where health is measured categorically. Advances also have been made in the analysis of measurement errors and mis-measured counts with respect to self-reported and actual physician visits (Cameron, *et al.* 2004). Joe (1997) showed several applications with a heavy focus on methodology: multivariate binary response, ordinal response and extreme value data, longitudinal binary and count data and serially correlated data, comparing models using their predictive power. Quinn (2005, 2007b) uses copulas to reinforce standard Bayesian economic evaluation techniques for clinical trial data. Also in a health economic context, Zimmer and Trivedi (2006) use a copula for three simultaneously-determined outcomes: health insurance status for couples and their individual health care demand.

The remainder of this paper discusses copulas in their own context as multivariate distribution functions, and in the health economics context according to various applications. Methods of constructing copulas of various types that may be of interest to health economics are presented, with copulas as multivariate distribution functions and models of multivariate dependence discussed in following sections. Finally estimation and goodness-of-fit is considered, before some concluding remarks.

2 Copulas

2.1 Symmetry and the bivariate normal distribution

Johnson and Kotz (1972) commented on the increasing recognition of the need for usable alternatives to the multivariate normal distribution, in instances when the marginal distributions were distinctly non-normal, or when the dependence between them deviated from the linear correlation associated with the bivariate normal. Such alternatives, while they exist, were not usually developed for data analysis *per se*, and so are not generalisable across marginal distributions or dependence (Cook and Johnson 1981). The multivariate normal is commonly selected for the convenience of its use and because the univariate normal distribution is robust under reasonable levels of non-normality, and so explain the margins of the joint distribution fairly well (Kowalski 1973).

The multivariate normal is employed also a result of the common practice of selecting a multivariate distribution according to identification of the margins; since the normal distribution is among the most robust, it is preferred to others such as the multivariate Pareto, Burr or Logistic, for example (Mardia 1962; Takahasi 1965, Satterthwaite and Hutchinson 1978; Cook and Johnson 1981 present a generalised model that nests each of these as special cases). Moreover, the normal distribution tends to be more easily extended to higher dimensions: the density or characteristic function of the normal distribution can be used, or a linear combination of normally-distributed random variables (Fang, *et al.* 1989). The multivariate normal can be preferred even when the joint density of the data being analysed appears not to be elliptically symmetric.

Alternatives, such as those found in Johnson, *et al.* (1997) for example, are typically less practicable than the multivariate normal under a given failure of the assumptions about the distributional form. Like most multivariate distribution functions they are extensions of the family of univariate distribution functions, based upon including some measure of association in the margins, which are combined to form the multivariate distribution. However this homogenisation of the margins in the joint distribution is commonly the problem: a different class of distribution may be needed for each margin, but different classes of univariate distributions generate different classes of multivariate distributions.

2.2 Conditional likelihood estimation

A related problem arises with the use of conditional distributions, which can allow different marginal distributions.[4] Conditional distributions are affected by the so-called Borel paradox (Kolmogorov 1950; Newey and Steigerwald 1997; Verhoeven and McAleer 2003). For practical purposes, this holds that the margins in a conditional distribution function are not 'swappable' - i.e. the conditional and conditioning distribution functions are not swappable, unless they are of the same, symmetric, family of distribution functions. Otherwise estimates of the parameters of the conditioning distribution function will not necessarily be consistent. Estimation of Conditional distributions in fact relies upon symmetry more than do pseudo - or quasi-maximum likelihood of bivariate distributions, which tend to overlook association implicitly during analysis.

These problems arise due to the inclusion, usually, of association as a parameter in the margins of the joint distribution function (consider Pearson's coefficient of correlation ρ in the bivariate normal as an example). Conditional distributions avoid this, but only by conditioning the moments of one marginal distribution on another, to the same effect: assumptions about the marginal distribution function direct the association between the random variables.

In order to estimate known-asymmetric data in a manner that includes no distributional restrictions, but which also provides estimates of conditionally-dependent parameters that are invariant to transformations or switching of margins, a measure of association is needed that is not contained in the margin itself (Frees and Valdez 1998). Measures of dependence such as rank correlation do this: these are invariant to non-linear increasing transformations of random variables.[5] They are therefore stable under switching, such that for a measure of dependence τ between random variables (X_1, X_2), $\tau_{X_1,X_2} = \tau_{X_2,X_1}$ regardless of the form (or skew, or kurtosis) of each marginal distribution $F_1(X_1)$ and $F_2(X_2)$.

2.3 Example 1: A bivariate Beta distribution with dissimilar margins

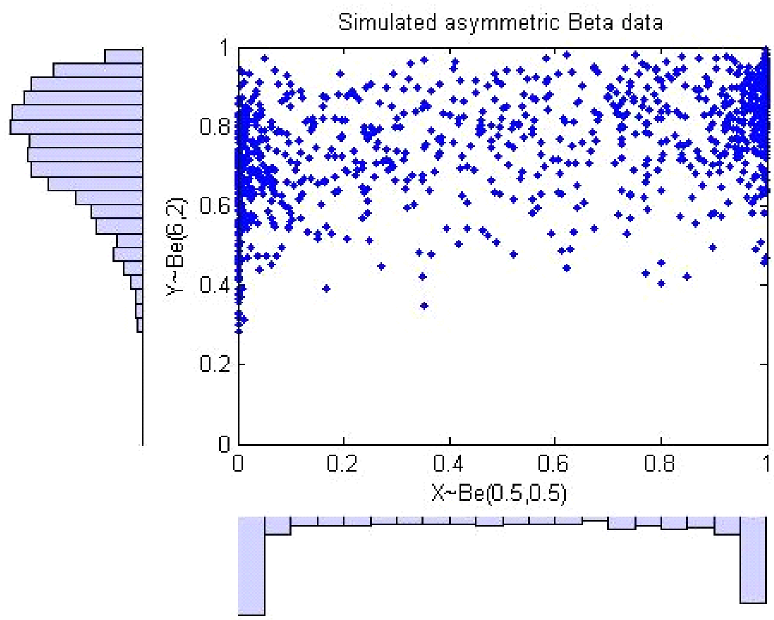
This example is useful to demonstrate the key contribution of copulas to econometrics and applied research: that assumptions about the joint and marginal distributions can be relaxed, when needed, thereby improving the performance of an economic model. The following examples do the same within the contexts of specific economic problems. This example however is motivated by a desire to demonstrate this key contribution.

Consider a joint (bivariate, for simplicity) distribution with non-normal, dissimilar and asymmetric margins, with two jointly-distributed random variables (X, Y), where $X \sim Be(.5, .5)$, $Y \sim Be(6, 2)$ and correlation $\rho = 0.6$. Figure 2.1 shows the marginal histograms of (X, Y), as well as the dispersion of jointly-observed data points.[6]

Note that although the margins are distributed within the same family, they are quite different. The distribution in Figure 2.1 is an example from a range of data-generating processes, including individual stock prices (Panas 2005), liquid asset ratios and education (Gordy 1998), income and the returns to research and development (De Castro and Goncalves 2002), entry and exit times in recreation use (Zarnoch, *et al.* 2004), dose response in toxicology studies and clinical trials (Calabrese and Baldwin 1998) and political partisanship (Box-Steffensmeier and Smith 1997).

This is an extreme example, because the nature of the marginal distributions make it clear that the bivariate normal distribution is not appropriate. Also, a known bivariate distribution, the Dirichlet, already exists for two Beta-distributed random variables. It is a useful example though, because the Beta itself is a flexibly-generalisable distribution, accommodating symmetry or asymmetry, nesting the Uniform and related to the Gamma, Exponential (and subsequently the Laplace, Chi-squared, Weibull and Raleigh), F and normal distributions (Aitchison 1986; Hutchinson and Lai 1990; Johnson, *et al.* 1997).

Figure 1: MARGINAL HISTOGRAMS AND BIVARIATE SCATTERPLOTS FOR JOINTLY-DEPENDENT BETA-DISTRIBUTED RANDOM VARIABLES X BE(0.5,0.5), Y BE(6,2), rho = 0.6



This does not however mean that the Dirichlet distribution will itself provide an appropriate fit. The Dirichlet is constructed from a trivariate form of reduction (where two dependent random variables are constructed from three independent random variables).[7] The parameters must conform across both distributions, and correlation is negative also (although $X' = (-X)$ can be used), and almost independent. Positive, or significantly negative, correlation is not accommodated well by this class. This is quite a strong property: generalisations that have attempted to find classes containing both the Dirichlet class and more flexible classes of distributions have not been very successful (many attempts involve lognormal classes of distributions: Aitchison 1986 contains the most comprehensive discussion of Dirichlet distributions). An important problem in this example is the asymmetry: the Dirichlet reduces only to margins where $X \sim Be(\alpha_1, \alpha_2 + \alpha_3)$ and $Y \sim Be(\alpha_2, \alpha_1 + \alpha_3)$, which cannot be reconciled with $F_1(X)$ and $F_2(Y)$ in this case (specifying α_1, α_2 accurately gives $\alpha_3 = -5.5$ for X, $\alpha_3 = 1.5$ for Y, but the parameters of the univariate Beta are strictly positive).[8]

These restrictions are emblematic of the problems caused by the usual constructions of multivariate distributions (Frees and Valdez 1998). This example is useful because it confronts both its own multivariate distribution and the most obvious alternative. Figure 2.2 contains the nearest approximation due to the bivariate normal, where the bi-modality problem can be seen more clearly. No approximation due to the Dirichlet can be found without altering one or both of the distributions of X and Y still further.

Figure 2.3, on the other hand, shows several classes of copulas, constructed purely from knowledge of the correlation and marginal distributions (such as, for example, an expectation of higher or lower association in one or both tails of the bivariate distribution). In each case the distribution has captured both the association between X and Y and the appropriate marginal distribution of each.

Figure 2: MARGINAL HISTOGRAMS AND BIVARIATE SCATTERPLOT FOR THE BIVARIATE NORMAL APPROXIMATION TO THE DATA IN FIGURE 2.1

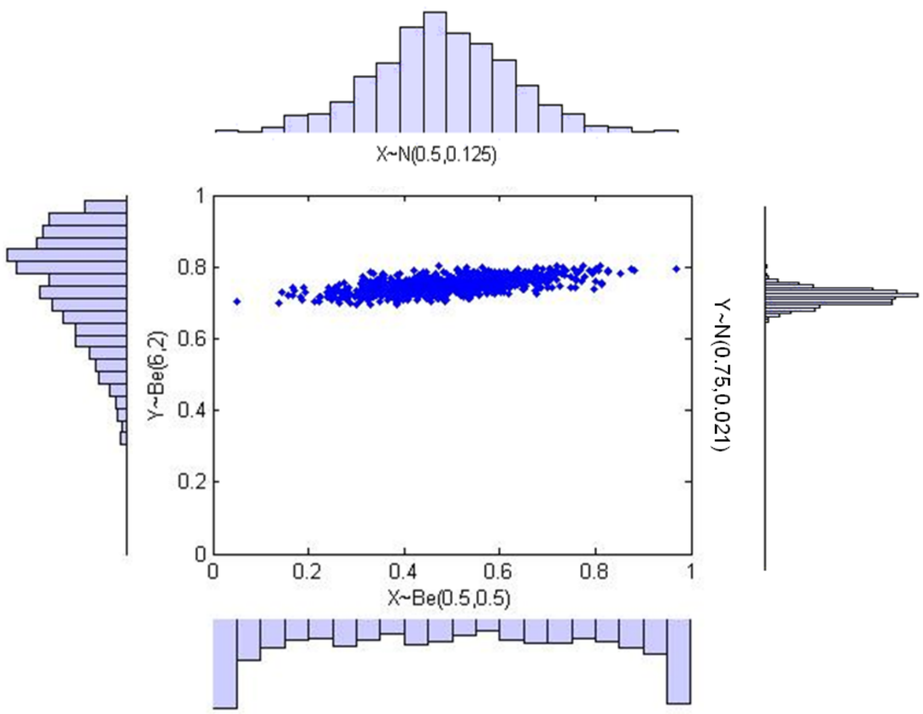
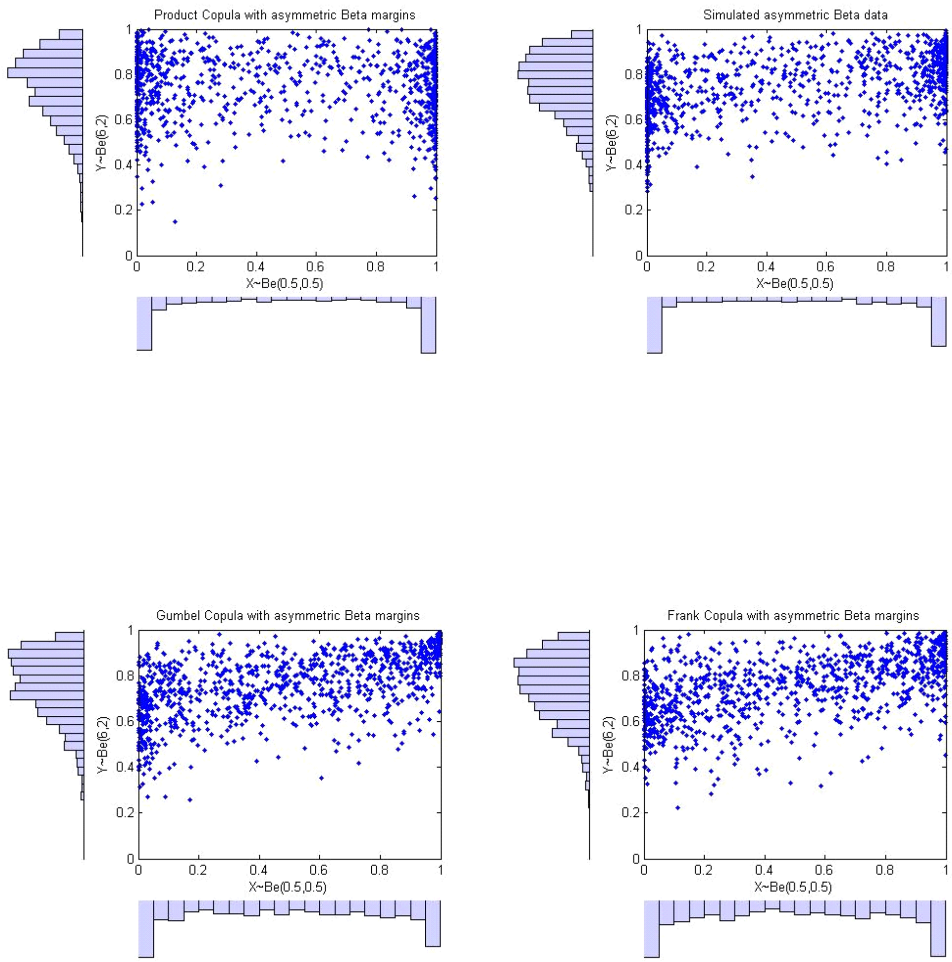
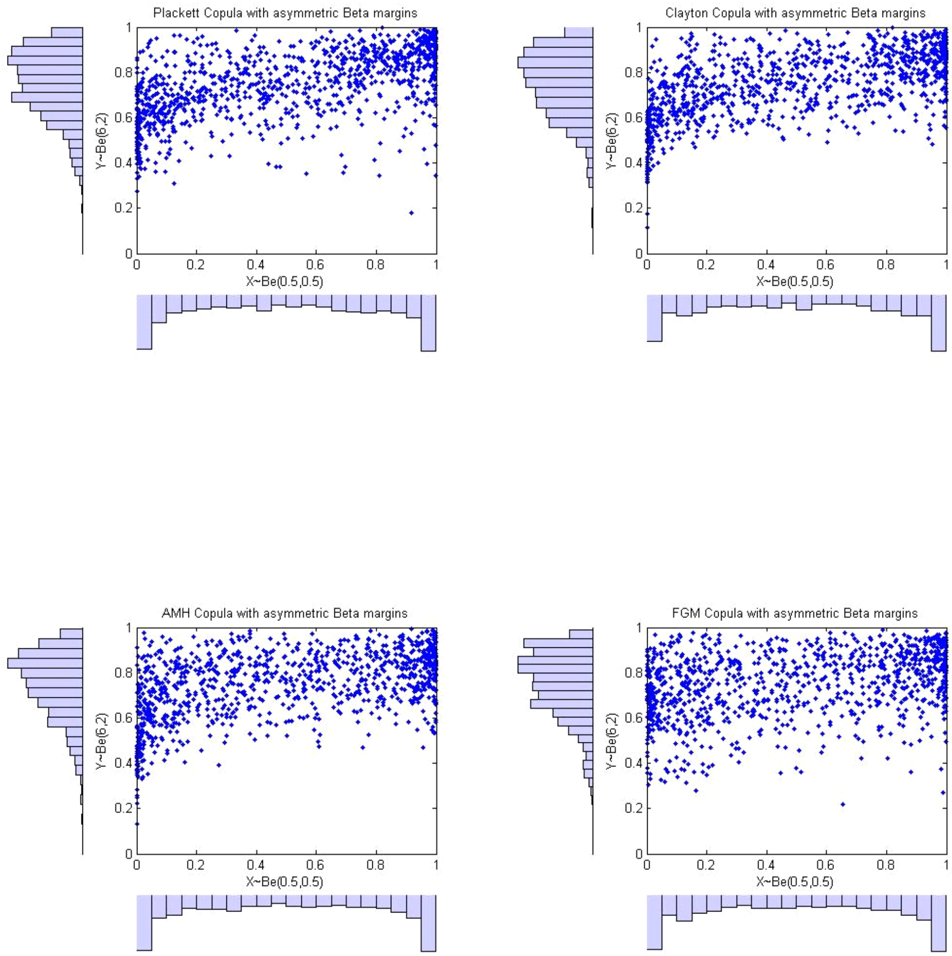


Figure 3: MARGINAL HISTOGRAMS AND BIVARIATE SCATTERPLOTS FOR X BE(0.5,0.5), Y BE(6,2), SIMULATED VIA INVERTING CONDITIONAL COPULAS (rho = 0.6)



MARGINAL HISTOGRAMS AND BIVARIATE SCATTERPLOTS FOR X BE(0.5,0.5), Y BE(6,2), SIMULATED VIA INVERTING CONDITIONAL COPULAS (rho = 0.6)



The notable exception is the so-called Product copula, which assumes $X \perp Y$. Looking at the marginal distribution of X, only the Product replicates that seen in purely simulated data, which is not surprising given that inversions from the Product copula draw upon no inherent dependence structure. The marginal distributions for Y show that the Farlie-Gumbel-Morgenstern (FGM), Gumbel and Plackett copulas do not quite contain the correlation in the data they generate due to their specific functional form. For the same reason the Gumbel and Clayton copulas show different association in the tails. None of the dependent copulas replicate the simulated distribution exactly: the marginal distribution of X shows this, and the scattered data in the dependent copula has more points in the centres than the empirical distribution. Comparison with Figure 2.2 though shows that they are more precise approximations than bivariate normality.

2.4 Example 2: Health insurance and health care utilisation

Zimmer and Trivedi (2006) used a copula for three simultaneously determined outcomes: health insurance status for couples and their individual health care demand. This involved discrete outcomes using a dichotomous selection equation for insurance coverage with negative binomial models for health care demand. The full model consisted of Archimedean copulas in a mixture to determine the trivariate distribution (an approach discussed later). Thus the dependence between insurance coverage and health care demand was retained, but split to allow for the fact that insured spouses will nevertheless demand health care in a trivariate framework: presumably based on need, but also partly on the needs of each other. This was done by taking advantage of the fact that the copula method offers closed-form multivariate distribution functions.

They showed primarily that the positive correlation, separately considered, indicates that policies to increase utilisation by women, for example, can 'spill over', increasing utilisation by their spouse also. This is relevant in terms of policy and concerns over excess utilisation. They also showed in their analysis the potential effect of the ordering of margins on the outcomes of analysis: they use a so-called mixture of powers, in which the placement of one margin in the distribution, relative to another, affects the dependence between them.

Pitt *et al.* (2006) discuss a similar problem: multi-dimensional measures of health care utilisation among the elderly. Their econometric problem is, as in Deb and Trivedi (1997), a multivariate count-data model. They use a Gaussian copula to model the dependence between these measures, and Markov simulation to establish the posterior multivariate distribution, and compare their results to the separate negative-binomial models fitted by Deb and Trivedi (1997).

2.5 Example 3: Cost-effectiveness analysis

Quinn (2005, 2007b) analyses the eVALuate hysterectomy trial. This was a multi-centre randomised trial comparing new laparoscopic procedures for hysterectomy with existing abdominal and vaginal procedures (Garry, *et al.* 2004). The trial collected data on total costs and the gain in Quality-adjusted Life Years (QALYs) per patient.

Standard methods for estimating Incremental Net Benefit (INB) consist of comparison of the means of some treatment group with those of the non-treated group (see for example Phelps and Mushlin 1991; Claxton and Posnett 1996; Gold, *et al.* 1996; Stinnet and Mullahy 1998; Drummond, *et al.* 2005), or Seemingly Unrelated Regression (SUR) analysis of individual cost and outcome (Willan, *et al.* 2004; Briggs, 2005; Vanness and Mullahy, 2005). Both assume bivariate normality (log-normality of cost can be accommodated).

The Frank and FGM copulas, seen in Figure 2.3, were used to estimate cost and QALYs gained via regression. Figure 2.4 contains the distribution of log-normal costs and Beta-distributed QALYs for the trial, overlaid with the density/spread of the models used.

Accommodating the Beta-distributed QALYs gained was most important in this application. The scatter-plots in each graph are observed pairs of Cost and QALYs gained for each individual; the histograms along the axes are empirical histograms for Cost and QALYs gained; finally the contours match the bivariate copulas that use those marginal distributions (log-normal and Beta for Cost and QALY gain, respectively). These contours are taken from the copula functions with specified margins, rather than the data itself.[9]

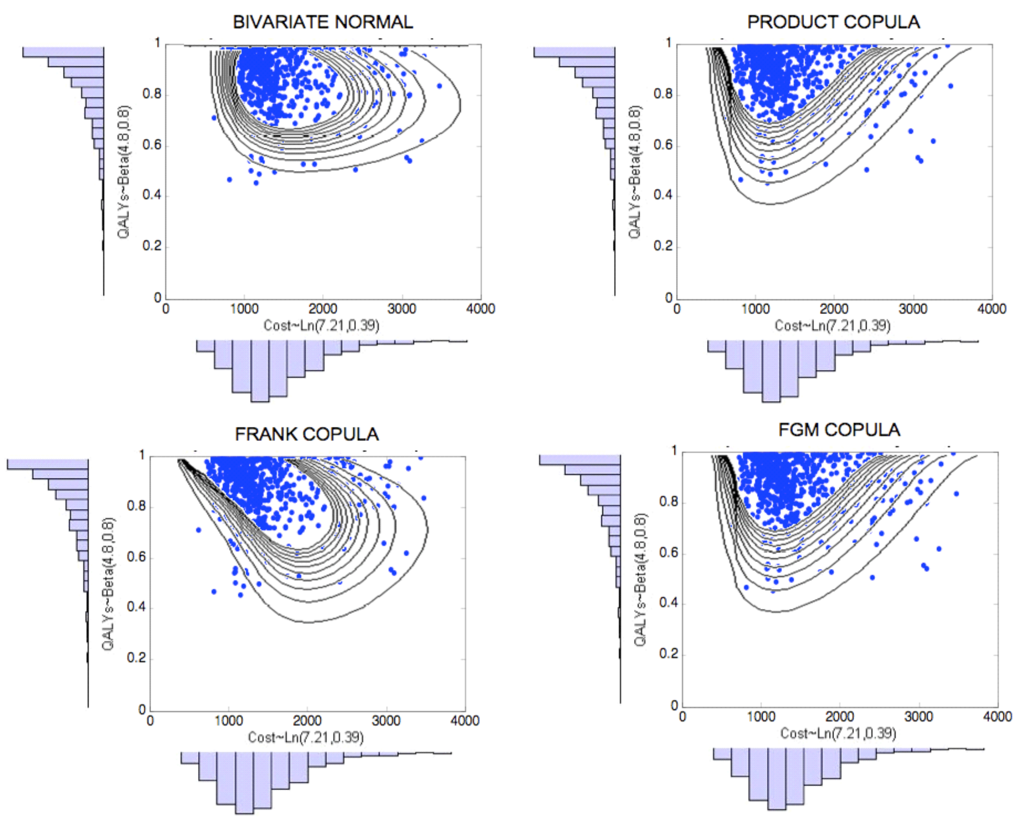
The graph labelled bivariate normal assumes log-normally-distributed Costs but normally-distributed QALYs gained, unlike the copulas. The FGM and Product copulas fit these data better than the Frank, and generated different cost-effectiveness thresholds for laparoscopic hysterectomy than the original study.[10]

2.6 Example 4: Sample selection

Smith's (2003) analysis of sample selection from the perspective of copulas was primarily methodological, however he did provide examples. The first was labour supply of females, using a study by Lee (1996), in which no association was found to exist between female participation and female labour supply - i.e. no selection bias. By employing copulas and comparing the results for the different dependence structures, Smith (2003) found varying degrees of association, but statistically significant selection bias in the same data.

In the same paper, Smith (2003) used Prieger's (2000) study of hospital length of stay. Smith (2003) employed copulas explicitly, and used a different class of copula (the Archimedean class, discussed later) and compared the results. Specifically, he included copulas capable of measuring negative dependence between hospitalisation (as the selecting event) and the length of stay. He determined that not only did the different measurement of dependence improve fit, but the estimated means lengths of

Figure 4: BIVARIATE COPULA DENSITIES, MARGINAL HISTOGRAMS AND BIVARIATE SCATTERPLOTS FOR OBSERVED INDIVIDUAL COSTS AND QALYS GAINED (QUINN 2007b).



stay varied accordingly.

Other examples are Genius and Strazzera (2003), who use copulas to examine the value of recreational forests, relaxing - like Smith (2003), the assumption of joint and marginal normality between participation in and valuation of such recreation. Trivedi and Zimmer (2005) also discuss sample selection in their general introduction to copulas.

2.7 Example 5: Stochastic frontier modelling

Smith (2005) used copulas to estimate stochastic frontier models, using US electric utilities and as a motivating example. In considering efficiency, which is typically a composite measure, independence is commonly assumed between technical inefficiency and (random) noise in the composite error of productivity modelling, although one would expect dependence between the two.

Smith (2005) employed copulas to overcome the independence assumption. He found in his results that, not only did the use of copulas re-rank companies within the industry, relative to standard stochastic frontier models, but the electric industry as a whole moved further below the production frontier when copulas were used. In his example, accommodating dependence showed that the US electric industry itself was less efficient than previously supposed, because the efficiency estimates of the utilities was quite sensitive to the dependence between components of the errors of the estimates.

In a health economics context, the analysis in Smith (2005) is directly comparable to the analysis of hospital efficiency, where the productivity or technical efficiency of hospitals is a function of several dimensions of hospital costs and outputs. The nature of copulas suggest also that they can be used in multilevel models, to a similar purpose.

3 Some statistical properties of copulas

3.1 Statistical foundations and methods of construction

To understand copulas fully it is worth considering them in the context within which they are generated, as different such methods create copulas with different characteristics and purposes in estimation. There are three general methods of construction: inversion, algebraic construction and geometric construction.

3.1.1 The inversion method

The most intuitive approach to constructing copulas is by using inverted distribution functions as arguments in known multivariate distributions. Thus, using inverses of the distributions in Equation (1) gives

$$C\left(G_1\left(x_1\right), G_2\left(x_2\right)\right) = F_{(1,2)}\left(F_1^{-1}\left(G_1\left(x_1\right)\right), F_2^{-1}\left(G_2\left(x_2\right)\right)\right) \tag{2}$$

for univariate CDFs $G_1\left(.\right)$ and $G_2\left(.\right)$. Note that $F_{(1,2)}$ is any joint distribution; different formulations of $F_{(1,2)}$ will generate different forms of C, one of which will be the either the closest approximation to, or the true, bivariate distribution H (the bivariate normal distribution $\Phi_{(1,2)}\left(x_1, x_2\right)$, for example, is another such approximation). Using the uniformly-distributed $F_1^{-1}\left(G_1\left(x_1\right)\right), F_2^{-1}\left(G_2\left(x_2\right)\right)$ allows, via inversion, the subsequent use of any type of distribution in the margin. Extending Equations (1) and (2) to higher dimensions, for $\left(X_1, ..., X_n\right) \in \mathbb{R}^n$ there exists the so-called n-copula

$$C\left(G_1\left(x_1\right), ..., G_n\left(x_n\right)\right) = F_{(1,...,n)}\left(F_1^{-1}\left(G_1\left(x_1\right)\right), ..., F_n^{-1}\left(G_n\left(x_n\right)\right)\right) \tag{3}$$

which also depend upon unique formulations of $F_{(1,...,n)}$ and where, according to Sklar's theorem, there exists one n-copula such that

Table 1: 2 x 2 CONTINGENCY TABLE

		X_1		
		Low	High	$\sum$
X_2	Low	a	b	a+b
	High	c	d	c+d
	$\sum$	a+c	b+d	

$$H\left(x_1, ..., x_n\right) = C\left(G_1\left(x_1\right), ..., G_n\left(x_n\right)\right) \tag{4}$$

The so-called Gaussian copula, given by

$$C\left(G_1\left(x_1\right), ..., G_n\left(x_n\right)\right) = \Phi_{(1,...,n)}\left(\Phi_1^{-1}\left(G_1\left(x_1\right)\right), ..., \Phi_n^{-1}\left(G_n\left(x_n\right)\right)\right) \tag{5}$$

is a well-known example of the inversion method, as is the t copula, which uses the multivariate student's t distribution. Nelsen (2006) discusses some more examples.

3.1.2 The algebraic method

The Plackett copula, seen in Figure 2.3, is the most well-known example of an algebraically-constructed copula. It provides the best example of a copula as a function of its univariate margins. Consider, in the bivariate case, two random variables (X_1, X_2), which can be represented by a 2×2 contingency table, as in Table 2.1.

Then the association between X_1 and X_2 in Equation (6)is given by the odds-ratio $\frac{ad}{bc}$ rather than by correlation: the Plackett measures the dominance of high-high and low-low values for (X_1, X_2) of- or by- high-low and low-high values.[11] In probabilistic terms this is given by (Nelsen 2006)

$$\theta = \frac{H\left(x_1, x_2\right)\left[1 - F_1\left(x_1\right) - F_2\left(x_2\right) + H\left(x_1, x_2\right)\right]}{\left[F_1\left(x_1\right) - H\left(x_1, x_2\right)\right]\left[F_2\left(x_2\right) - H\left(x_1, x_2\right)\right]} \tag{6}$$

since, from Table 2.1, $a = H(x_1, x_2)$, $d = 1 - F_1(x_1) - F_2(x_2) + H(x_1, x_2)$, $b = F_1(x_1) - H(x_1, x_2)$ and $c = F_2(x_2) - H(x_1, x_2)$. From Sklar's theorem $H(x_1, x_2) = C(u, v)|_{u=F_1(x_1), v=F_2(x_2)}$, thus

$$\theta = \frac{C(u, v)[1 - u - v + C(u, v)]}{[u - C(u, v)][v - C(u, v)]} \tag{7}$$

With some manipulation this gives

$$C(u, v; \theta) = \frac{[1 + (\theta - 1)(u + v)] - \sqrt{[1 + (\theta - 1)(u + v)]^2 - 4uv\theta(\theta - 1)}}{2(\theta - 1)} \tag{8}$$

Nelsen (2006) contains the proof that only the negative root of the term in Equation (8) is a copula (Mardia 1970). As well as an example of construction, Equations (6)-(8) illustrate how a copula becomes a function of only uniform univariate margins and some measure of association. Joe (1997) illustrates other examples of construction, as does Smith (2005) in an early section of his discussion of estimating efficiency with copulas compared to stochastic frontier models.

3.1.3 The geometric method

The limits, for minimum and maximum dependence accommodated by copulas (or their minimum and maximum values as joint CDFs), are defined by the so-called Fréchet-Hoeffding bounds W and M, where

$$\begin{aligned} W(u, v) &= \max\{u + v - 1, 0\} && \text{(9a)} \\ &\leq C(u, v) && \text{(9b)} \\ &\leq \min\{u, v\} = M(u, v) && \text{(9c)} \end{aligned}$$

considering copulas within a bivariate framework only (Lindskog 2000; this can be extended to higher dimensionality). The simplifying notation $u = F_1(x_1)$, $v = F_2(x_2)$ will be used hereon. Fréchet-Hoeffding bounds exist as universal bounds on any copula C, for all $u, v \in [0,1]$ as well as similarly grounding distribution functions generally. Consider u, v in the unit interval $\mathbf{I} \in [0,1]$. Then a bivariate copula $C(u,v)$ is a function $C : \mathbf{I}^2 \to \mathbf{I}$ such that

$$C(u,0) = C(0,v) = 0 \tag{10}$$

$$C(u,1) = u \text{ and } C(1,v) = v \tag{11}$$

$$C(u_2,v_2) - C(u_1,v_2) - C(u_2,v_1) + C(u_1,v_1) \geq 0 \tag{12}$$

where $u_1 \leq u_2,\ v_1 \leq v_2 \in [0,1]$. These establish that the copula is grounded, bound by the unit square and 2-increasing ($C(u,v)$ is increasing for increases in u, v). This is the same for any bivariate distribution. $W(u,v)$ and $M(u,v)$ in Equation (9a)-(9c) are so-called extreme-value copulas. A point not explored here explicitly is that for $n > 2$ only M would be a copula, not W: a 3-copula, for example, will not reach its lower bound. More accurately, its lower bound will be the best possible bound on the n-copula C_n but will not equal the lower bound W. For more on this, see Nelsen, *et al.* (2004) and Nelsen and Úbeda-Flores (2004). These are fundamental to the definition of a copula, and it is from this point that copulas can be constructed geometrically.

Since the process of geometric construction is decidedly more statistical in nature, it is not important to include it here. Suffice to say, this approach enables construction of copulas that accord with specific desired properties: linear, quadratic or cubic sections, for example, are copulas with prescribed horizontal or vertical supports. The FGM copula in Figure 2.3 is an example of such. Copulas of geometric construction can also be given specific diagonal sections (such as for random variables with a common marginal distribution function). Again, Nelsen's (2006) book provides a more technical explanation of geometric construction. The method is of interest here because of the importance of the Fréchet-Hoeffding bounds in copulas, particularly in higher dimensions.

3.2 Some classes and families of copulas

In keeping with the discussion so far, the families and classes of copulas discussed here will correspond to those most often seen in the literature, particularly single-parameter families of copulas, rather than higher-order families. The categories are stylistic, not statistical: a single copula could be characterised across several of the following sub-sectioned properties.

3.2.1 Linear

The FGM copula is the most commonly seen copula in exposition, since lower polynomials are more convenient for discussion (Genius and Strazzera 2003; Smith 2003; Zimmer and Trivedi 2006). It is also the first-order Taylor approximation of another single-parameter family of copulas, the Frank.[12,13] The FGM copula is C such that

$$C(u, v; \theta) = uv(1 + \theta(1 - u)(1 - v)) \tag{13}$$

where $\theta \in [-1, 1]$, and recalling that $F_1(x_1)$, $F_2(x_2)$ are (at least) monotonic. This contains positive and negative dependence for $\pm\theta$ respectively, and multivariate extensions. Symmetry of this type is not exclusive to simple linear forms for C, though, and can exist in higher polynomials. In practical applications this copula has been shown to be somewhat limited: for copula dependence parameter $\theta_{FGM} \in [-1, 1]$, Spearman's correlation $\rho \in \left[-\frac{1}{3}, \frac{1}{3}\right]$ and Kendall's $\tau \in \left[-\frac{2}{9}, \frac{2}{9}\right]$ (Trivedi and Zimmer 2006). Mari and Kotz (2001) provide several extensions of the FGM copula, which expand this range to different degrees. Prieger (2000) also shows this limitation. The FGM density is given by

$$
\begin{aligned}
C_{12}(u,v;\theta) &= \frac{\partial^2 C(u,v;\theta)}{\partial u \partial v} && \text{(14a)}\\
&= (1+\theta(1-2u)(1-2v)) && \text{(14b)}
\end{aligned}
$$

which again is straightforward. For estimation the joint probability density function h of distribution H is given by,

$$
h(x_1,x_2) = f_1(x_1) f_2(x_2) C_{12}(F_1(x_1), F_2(x_2); \theta) \tag{15}
$$

Thus the FGM lends itself well to methods of maximum likelihood. Being linear, the FGM is also extended into higher dimension with less difficulty than other families.

3.2.2 Elliptical copulas

The Gaussian copula defined in Equation (5) is an example of an elliptical copula. Despite its similarity to the bivariate normal distribution, notably to the extent that association is measured by Pearson's correlation ρ, the Gaussian copula is convenient in particular for simulation, which is not usually straight-forward in the case of jointly dependent data (see Perkins and Lane 2003 for more examples).

Another common elliptical copula is the student's t copula. This is, as discussed in the section on the inversion method, different from the Gaussian copula only in the use of the multivariate student's t, rather than the normal, distribution. With v degrees of freedom, the t copula is given by (Bouyé, *et al.* 2000)

$$
C(G_1(x_1), G_2(x_2)) = T_{(1,2),v}\left(t_v^{-1} G_1(x_1), t_v^{-1} G_2(x_2)\right) \tag{16}
$$

Although suited to simulation exercises, in practice these can be difficult to implement due to the form of the distribution. Copula density functions, as in Equation (15), use both the marginal density and distribution, which in the Gaussian case uses the bivariate distribution

$$\begin{aligned} C\left(F_1\left(x_1\right), F_2\left(x_2\right); \rho\right) &= \int_{-\infty}^{\Phi_1^{-1}\left(F_1\left(x_1\right)\right)} \int_{-\infty}^{\Phi_2^{-1}\left(F_2\left(x_2\right)\right)} \frac{1}{2\pi\sqrt{\left(1-\rho^2\right)}} \\ &\quad \times \exp\left\{\frac{-\left(s^2-2\rho st+t^2\right)}{2\left(1-\rho^2\right)}\right\} \mathrm{dsdt} \end{aligned} \tag{17}$$

Elliptical copulas can be found in other general classifications of copulas also. The FGM copula, above, and the Frank copula given by

$$C\left(u, v; \theta\right) = -\frac{1}{\theta} \ln\left(1+\frac{\left(e^{-\theta u}-1\right)\left(e^{-\theta v}-1\right)}{e^{-\theta}-1}\right) \tag{18}$$

where $\theta \in (-\infty, \infty) \setminus \{0\}$ are also examples of copulas with ellipsoid margins. More importantly for analysis, these also have closed-form CDFs.

3.2.3 Extreme Value copulas

Seger (2004) motivates his exposition of extreme value copulas with a discussion of the probability of flooding, given the height of two dykes and that of a river. In health economics a more appropriate hypothetical situation would be the carrying capacity of two hospitals and the occurrence of injuries, and the associated probability that hospitals would be overtaxed during peak periods (such as specific holidays when alcohol consumption is high or a lot of motorists are on the road simultaneously). In finance the returns of two or more correlated stocks during a given trading period is a common concern. In each instance the area of interest is the tails of the distribution, where little information can usually be found. The multivariate normal distribution, for example, treats extrema as independent events.

Galambos (1987), Resnick (1987), Joe (1997) and Abdous, *et al.* (1999) are good sources for more discussion of extreme value distributions and copulas from a statistical perspective. The focus here is primarily upon the families themselves. Extreme value distributions are due to the so-called three-types theorem: if there exists a non-degenerate limit distribution G for the random variable X, G is one of the following (Bouyé, *et al.* 2000)

$$\text{Fréchet} \quad G(x) = \begin{cases} 0 & x \leq 0 \\ \exp\left\{-x^{-\theta}\right\} & x > 0 \end{cases} \tag{19a}$$

$$\text{Weibull} \quad G(x) = \begin{cases} \exp\left\{-\left(-x\right)^{\theta}\right\} & x \leq 0 \\ 1 & x > 0 \end{cases} \tag{19b}$$

$$\text{Gumbel} \quad G(x) = \left\{ \exp\left\{-\exp\left\{-x\right\}\right\} \quad x \in \mathbb{R} \right. \tag{19c}$$

By symmetry the possible limits for minima of X are the reflections of $G(x)$. If one of these is the extreme of a given distribution function F, then F is said to be in the domain of attraction, thus $F^m\left(a_m x + b_m\right) \to G(x)$ for some constants a_m and b_m (Joe 1997; Galambos 1987; Resnick 1987).[14] Extreme value copulas can also be found for multivariate copulas.[15]

Some popular single-parameter families of extreme value copulas are

$$C\left(u,v\right) = uv \qquad \mathrm{C}^{\perp} \tag{20a}$$

$$C\left(u,v;\theta\right) = \exp\left\{-\left(-\ln\left(u\right)^{\theta} - \ln\left(v\right)^{\theta}\right)^{\frac{1}{\theta}}\right\} \qquad \text{Gumbel(I)} \tag{20b}$$

$$C\left(u,v;\theta\right) = uv\exp\left\{\theta\frac{\left(-\ln(u)\right)\left(-\ln(v)\right)}{-\ln(u)-\ln(v)}\right\} \qquad \text{Gumbel(II)} \tag{20c}$$

$$C\left(u,v;\theta\right) = uv\exp\left\{\left(-\ln\left(u\right)^{-\theta} - \ln\left(v\right)^{-\theta}\right)^{\frac{-1}{\theta}}\right\} \qquad \text{Galambos} \tag{20d}$$

$$C\left(u,v;\theta\right) = \min\left\{u,v\right\} \qquad \mathrm{C}^{+} \tag{20e}$$

3.2.4 Archimedean copulas

Archimedean copulas are a particular *class* of copula that includes several popular families. These are copulas whose form, in n dimensions, is reduced to a single function, called a generator. This is a strictly decreasing, convex and continuous function $\varphi : [0,1] \rightarrow [0,\infty]$ in a set Ω of the same, where $\varphi(0) = \infty$, $\varphi(1) = 0$ and with inverse $\varphi^{-1} : [0,\infty] \rightarrow [0,1]$, $\varphi^{-1}(0) = 1$ and $\varphi^{-1}(\infty) = 0$. Archimedean copulas are symmetric, associative such that $C(C(u,v),w) = C(u, C(v,w))\ \forall\ u,v,w$ in $\mathbf{I}$ and linearly transformable such that for some constant $c > 0$, $c\varphi$ is also a generator of C.

For some u, v in $\mathbf{I} \in [0,1]$ an Archimedean copula is C such that

$$C(u,v) = \varphi^{-1}(\varphi(u) + \varphi(v)) \tag{21}$$

This satisfies Equations (10)-(11) since

$$\begin{aligned} C(u,0) &= \varphi^{-1}(\varphi(u) + \varphi(0)) \\ &= 0 \end{aligned} \tag{22}$$

and

$$\begin{aligned} C(u,1) &= \varphi^{-1}(\varphi(u) + \varphi(1)) \\ &= \varphi^{-1}(\varphi(u)) \\ &= u \end{aligned} \tag{23}$$

Satisfaction of symmetry means $C(0,v) = 0$ and $C(1,v) = v$ also. Equation (21) provides the basic foundation for conceptualising Archimedean copulas. By way of exemplification, consider the Ali-Mikhail-Haq (AMH) copula of the odds in favor of failure against survival, also seen in Figure 2.3 (Ali, *et al.* 1978). This is the copula C such that

$$C(u, v; \theta) = \frac{uv}{1 - \theta(1-u)(1-v)} \tag{24}$$

where $\theta \in [-1, 1)$. This also is an algebraically-constructed copula such as the Plackett, using the odds in favour of failure against survival. Using a proof from Nelsen (2006) that, for almost all u, v in $\mathbf{I}$,

$$\frac{\varphi_\theta'(u)}{\varphi_\theta'(v)} = \frac{\partial C_\theta(u,v)/\partial u}{\partial C_\theta(u,v)/\partial v} \tag{25}$$

a generator for C can be found, using partial derivatives. For the AMH copula this provides, for some t in $\mathbf{I}$,

$$\varphi_\theta(t) = \ln\left(\frac{1-\theta(1-t)}{t}\right) \tag{26}$$

The Frank copula from Equation (18) is also a popular Archimedean class copula (Zimmer and Trivedi 2006; Smith 2003; Genius and Strazzera 2004; Smith 2005). Here the generator is given by

$$\varphi_\theta(t) = -\ln\left(\frac{e^{-\theta t}-1}{e^{-\theta}-1}\right) \tag{27}$$

This is a comprehensive family, such that association $\theta \in [-\infty, \infty] \setminus \{0\}$ corresponds to $\tau \in [-1, 1] \setminus \{0\}$.[16]

A third popular example is the Clayton copula, given by (Clayton 1978)

$$C(u, v; \theta) = \max\left[\left(u^{-\theta} + v^{-\theta} - 1\right)^{\frac{-1}{\theta}}, 0\right] \tag{28}$$

where $\theta \in [-1, \infty) \setminus \{0\}$. The Clayton copula is constructed using the following generator

$$\varphi_\theta(t) = \frac{1}{\theta}\left(t^{-\theta} - 1\right) \tag{29}$$

This also corresponds to association $\theta \rightarrow \tau \in [-1, 1] \setminus \{0\}$. This is one of several families of Archimedean copulas whose distributions are based on maximands (see Nelsen 2006).[17]

As an argument in only one margin, the generator φ can be used to extend Archimedean copulas into higher dimensions easily. For example, for u, v, w, z in $\mathbf{I}$,

$$C(u, v, w, z) = \varphi^{-1}(\varphi(u) + \varphi(v) + \varphi(w) + \varphi(z)) \tag{30}$$

All that is required to extend C is the addition of the generator function for a new margin. Note that φ_θ belongs to a single-parameter family of generators. Two-parameter generators also exist. These are $\varphi_{\theta_1\theta_2}$ such that $\varphi_{\theta_1\theta_2}$ also belongs to the set Ω of continuous decreasing convex functions as above. These generate so-called Rational Archimedean copulas $C(u, v) = P(u, v)/Q(u, v)$ where P, Q are polynomials, such as the AMH family. An example is a parametric extension of the AMH family of copulas, thus

$$C(u, v; \theta_1, \theta_2) = \frac{uv - \theta_2(1-u)(1-v)}{1 - \theta_1(1-u)(1-v)} \tag{31}$$

subject to appropriate values for θ_1, θ_2 (Nelsen 2006). Like φ_θ, $\varphi_{\theta_1\theta_2}$ can be used for straightforward additive extensions to generate Archimedean n-copulas, according to Equation (30). This will be discussed more explicitly in the following section.

4 The tractability of copulas as multivariate distribution functions

Equation (30) shows one of the useful properties of copulas. The copula can be made tractable in its margins, making estimation of dependent multivariate data much more practicable. Smith's (2005) examination of Fisher Information for copulas is particularly useful. He shows that the vector of parameters β_i in each margin of a closed-form copula are invariant to every other margin in the distribution.[18] Estimates of $\hat{\beta}_i$ are dependent only upon association parameter θ and the functional form of u_i.

4.1 Multivariate copula families

4.1.1 Example 6: Mortality, health and lifestyles

Consider the representative example of Balia and Jones (2007). They analyse the relationship between mortality, lifestyle and socioeconomic status, using longitudinal data from the British Health and Lifestyle Survey (HALS). Their model is a structure of 8 equations: mortality, self-assessed health and 6 health behaviours: smoking, alcohol consumption, exercise, sleep, consumption of breakfast and obesity. These are all binary, and with a reduced form of endogenous regressors. Assuming correlated errors to be normally distributed, this problem requires the method of simulated maximum likelihood to overcome an 8-dimensional probit, or 8 simultaneous integrals over the normal density with 28 pairwise correlations, with a mixture of signs.

Maximising likelihoods across 8 dimensions is time-consuming and computationally intensive, however a closed form solution can be found using the copula method (Muthén 1979, 1984 discusses the former issue in some detail). Quinn (2006, 2007c) used multivariate Gaussian and t copulas to analyse the same data, demonstrating the efficiency gains. While Balia and Jones (2007) analysed a recursive model in structural form, Quinn (2006, 2007c) used the system in reduced form.[19]

Although multivariate copulas must be constructed with significantly more care, they can provide advantages: as well as closed-form multivariate CDFs, they offer flexible multivariate dependence structures and autonomy over the functional form of each marginal CDF. Multivariate copulas also offer disadvantages compared to a multivariate normal framework, which is computationally more intensive but can require less interactive construction by the analyst. These disadvantages will also be discussed.

4.1.2 Multivariate FGM copulas

The FGM is an example of a simple closed-form CDF, where Equation (13) is extended into n dimensions. Using the notation of Joe (1997), the multivariate FGM copula for the problem in Balia and Jones (2007) can be given as

$$C(F_1(x_1), ..., F_8(x_8); \theta) = \prod_{i=1}^{8} u_i \left(1 + \sum_{1 \leq i < j}^{k=8} \theta_{ij} [1 - u_i] [1 - u_j] \right) \tag{32}$$

giving, like the 8-dimensional normal distribution, ${}^8C_2 = 28$ bivariate association parameters (since $\theta_{ij} = \theta_{ji} \ \forall \ i \neq j$).[20] Here $\theta_{ij} \in [-1, 1]$ as before, however more restrictions are introduced: θ_{ij} faces a limit also in sum, so that more margins means a narrower range of dependence for each non-zero θ_{ij}. Specifically

$$1 + \left| \sum_{1 \leq i < j < n}^{n-1} \theta_{ij} \right| \leq \theta_{1n} \leq 1 + \left| \theta_{12} \sum_{2 \leq i < j \leq n}^{n} \theta_{ij} \right| \tag{33}$$

so that $\lim_{n \to \infty} \theta_{ij} = 0$. In fact this limit is much narrower: in practice much fewer than 28 unique values for θ_{ij} would be preferred. Although the multivariate FGM offers a parameter for association in each bivariate margin, this is not usually feasible in practice.

4.1.3 Multivariate Archimedean copulas

Multivariate Archimedean class copulas are a popular alternative, as exemplified in Equations (21)-(30). Estimation in n dimensions however can be limited: for any $n > 2$-distributions to be a copula, the generator $\varphi_\theta^{-1}[0, \infty)$ must be completely monotonic. In Archimedean copulas that extend to negative dependence, $\varphi_\theta^{-1}[0, \infty)$ when $n > 2$ fails to be monotonic when $\theta \in \tau < 0$. In capturing positive multivariate dependence, Archimedean copulas are bound also by their parameterisation. Unlike the multivariate FGM, where θ_{jk} exists for each bivariate pair (u_j, u_k), φ_θ^{-1} is usually a function of a single parameter. Equations (26), (27) and (29) for example show that any bivariate pair will share a common association parameter. Consider for example $C(u, v, w)$ where C is the Frank copula. The generator in Equation (27) gives

$$C(u, v, w; \theta) = -\frac{1}{\theta} \ln \left(1 + \frac{\left(e^{-\theta u} - 1\right)\left(e^{-\theta v} - 1\right)\left(e^{-\theta w} - 1\right)}{\left(e^{-\theta} - 1\right)^{n-1}} \right) \tag{34}$$

where $\theta > 0$.

An alternative for multivariate Archimedean class copulas is generation according to inverse Laplace transforms and mixtures of powers (Joe 1997; Zimmer and Trivedi 2006). This is a transform $\phi(s)$ of some univariate CDF $M(\alpha)$ such that

$$\phi(s) = \int_0^\infty e^{-s\alpha} M(\alpha) \tag{35}$$

for $s > 0$. Joe (1993, 1997) calls the function M the mixing function. In the univariate case he shows that any arbitrary distribution function F will have a unique Laplace transform G, where

$$\begin{aligned} F(x) &= \int_0^\infty G^\alpha dM(\alpha) \\ &\equiv \phi(-\log G(x)) \end{aligned} \tag{36}$$

Zimmer and Trivedi (2006) present the parameter $\alpha > 0$ as a form of heterogeneity affecting the random variable X. Since copulas are distribution functions, like F, the bivariate case can be considered as

$$\begin{aligned} C(u, v; \theta) &= \int_0^\infty G_u^\alpha G_v^\alpha dM(\alpha) \\ &\equiv \phi(-\log G_u - \log G_v) \\ &\equiv \phi\left(\phi^{-1}(u) + \phi^{-1}(v)\right) \end{aligned} \tag{37}$$

where $G_u = \exp\left\{\phi^{-1}(u)\right\}, G_v = \exp\left\{\phi^{-1}(v)\right\}$. This can continue into any number of dimensions, using different Laplace transforms to overcome the singularity of the dependence structure. Only $n-1$ distinct transforms exist across $\frac{n(n-1)}{2}$ bivariate margins in an n-copula, though, so that distinct bivariate margins nevertheless share a common association. A trivariate mixture using two distinct transforms $\phi(s) \neq \psi(s)$ will give

$$C(u, v, w; \theta) \equiv \psi\left(\psi^{-1} \circ \phi\left(\phi^{-1}(u) + \phi^{-1}(v)\right) + \psi^{-1}(w)\right) \tag{38}$$

where $\psi^{-1} \circ \phi$ belongs to a class of infinitely differentiable increasing functions (Joe 1997).[21] Importantly, dependence is symmetric with respect to u and v, but not w now: this is an improvement upon, for example, Equation (34).

In terms of the motivating example in this section, this will produce 7 distinct measures of dependence for the 28 bivariate pairs of 8 distributions, but all positive. This is the Jouini and Clemen (1996) condition that $\theta > 0$ under Laplace transforms and multivariate Archimedean copulas. Correlations from the HALS data used by Balia and Jones (2007) show 9 of these 28 pairs are negatively associated, contra-indicating the use of Archimedean copulas.

Finally, consider mixtures of max-infinitely divisible distributions (Joe and Hu 1996; Joe 1997). A multivariate distribution H is called max-id if H^γ is a CDF for all $\gamma > 0$ and for all n dimensions.[22] In fact the mixture-of-powers approach just discussed is a mixture of powers of a max -or min-id multivariate distribution function. This approach can be extended to negative dependence for some bivariate margins, although such extensions are less common or straightforward.[23] If a copula is of the form in Equation (37), C can take the general form $C(u_1, ..., u_n) = \phi(-\ln H(u_1, ..., u_n))$, and C is a multivariate CDF if H is max-id and $-\ln\phi$ belongs to a class of infinitely differentiable increasing functions. This general form contains copulas of the form in Equation (38), however extensions to negative dependence (a sufficient condition for which is when $-\ln\phi$ is convex) do not have mixture representations. Moreover, such extensions tend to generate multivariate copulas whose margins are all RR_2, or negatively dependent: this is because the n-copula would be a mixture of min-id distributions, such that each bivariate margin is RR_2. This is the case even with general dependence such as the FGM in Equation (32) that allow unique bivariate association. Consider the copula C such that

$$C(u_1, ..., u_n) = \psi\left(-\sum_{i<j} \ln K_{ij}\left(e^{-p_i\psi^{-1}(u_i)}, e^{-p_j\psi^{-1}(u_j)}\right) + \sum_{i=1}^{n}(q_i + n - 2)\, p_i\psi^{-1}(u_i)\right) \tag{39}$$

where q_i is another max-id mixing parameter specific to each marginal CDF. Each K_{ij} in this expression is a bivariate margin; specifically a bivariate copula. Each K_{ij} then is max-id, giving $C(u_1, ..., u_n)$ positive orthant dependence.[24] Using the survival function in each case will instead give negative orthant dependence (Joe 1997; Belzunce and Semeraro 2004).

One solution is to use the Laplace transform $\phi(s) = \max\left\{(1+\theta s)^{\frac{-1}{\theta}}, 0\right\}$, which does permit negative association.[25] Using this, and another Laplace transform ψ, a multivariate copula can be constructed from Equation (38), such that each bivariate margin has the appropriate association (in terms of sign: the limit to the number of unique Laplace transforms that can be used still exists). Estimation of Balia and

Jones' (2007) lifestyle and mortality model would require this approach, and/or that of Equation (32) to a closed-form multivariate distribution function. Applications not facing negative correlation are much more straightforward: any or all of the multivariate copulas in Equations (32), (38) or (39) would be appropriate. A recent paper by Savu and Trede (2006) discuss this, considering hierarchical Archimedean copulas - the same as this mixing of multivariate Archimedean copulas - using 12 Euro-Stoxx-50 stocks in 3 sectors, all positively correlated. The explicit hierarchy means they are never modelling more than 4 stocks at any one level, and the absence of negatively-correlated returns allows them to bypass the critical flaw preventing the use of this procedure in the Balia and Jones (2007) case.

One remaining method appropriate to parametric estimation of dependent multivariate distribution functions is the method of Inference Functions for Models, discussed in a later section.

4.1.4 Multivariate elliptical copulas

The Gaussian copula in Equation (5), and the t copula in Equation (16) can be extended to higher dimensions simply by adding in more margins to the multivariate normal or multivariate t distributions, thus:

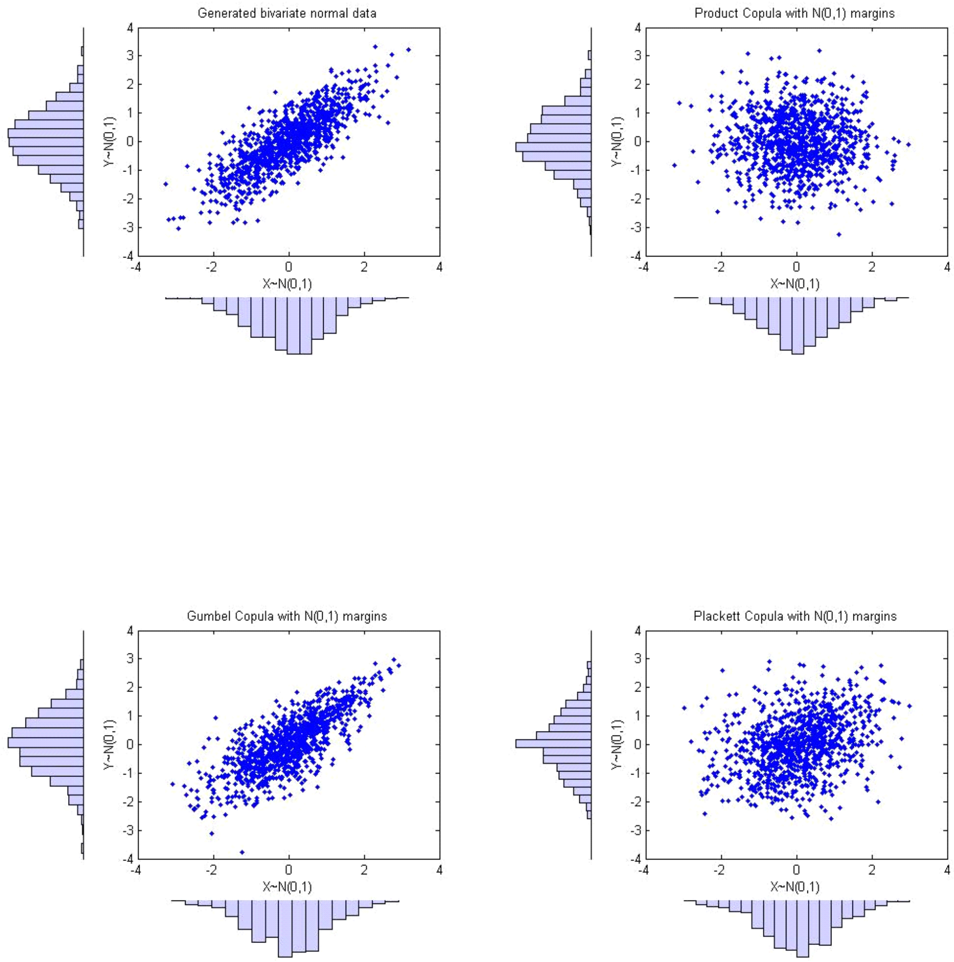
$$C\left(G_1\left(x_1\right),...,G_n\left(x_n\right)\right)=T_{(1,..,n),v}\left(t_v^{-1}G_1\left(x_1\right),..,t_v^{-1}G_n\left(x_n\right)\right) \tag{40}$$

Quinn (2006, 2007c) employed these, and variations due to Azzalini and Dalla Valle (1996) and Azzalini and Capitanio (2003) to consider skewness, in his analysis of the HALS data. Savu and Trede (2006), in their paper, criticise the unimodality of the multivariate Gaussian and t copulas, however they have the only robust and generalisable dependence structure. In terms of notation, derivation and computation, they are also significantly more straightforward to manage than the mixture-of-Max-ID approach in the preceding section.

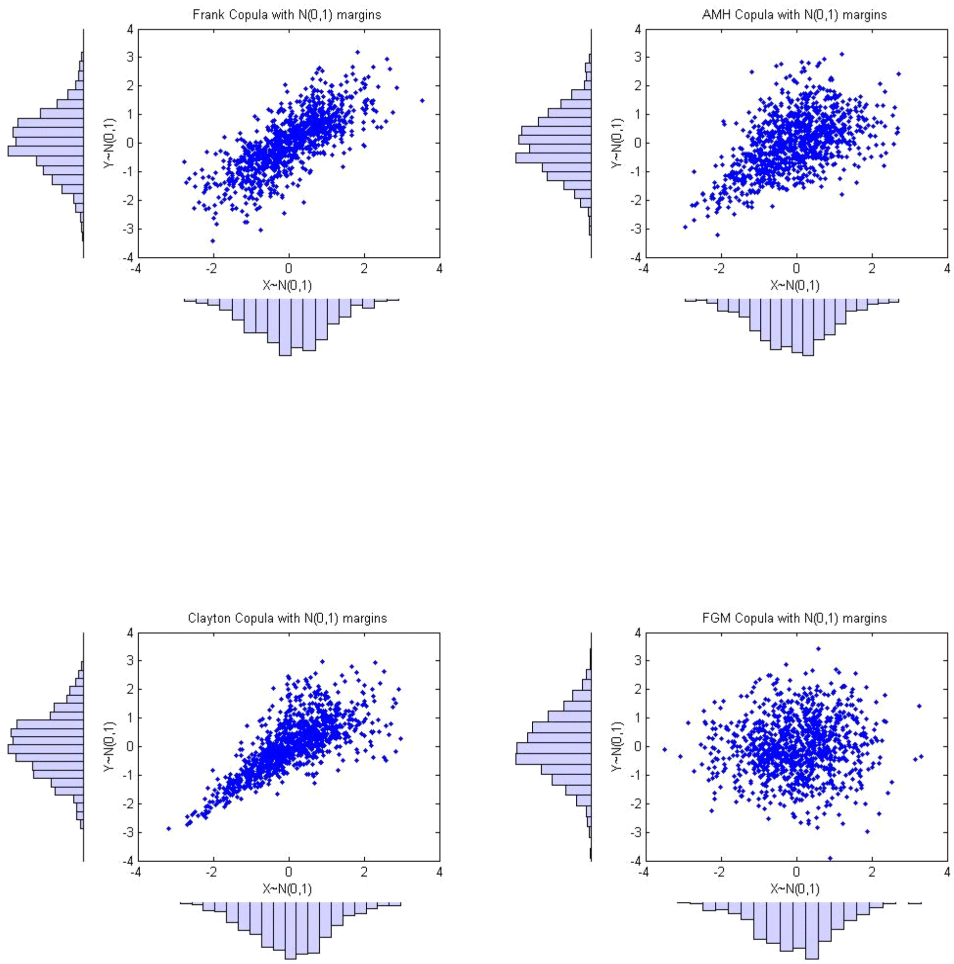
5 Copulas as functions of dependence

The central properties of copulas established thus far are that copulas are functions of two essential elements: dependence (bivariate or multivariate) and the distribution functions of each random variable. The selection of the univariate CDF F however is done independently of the selection of the functional form of the copula, which includes only the final uniformly-distributed F as its argument. Thus, copulas are functions of dependence only. This was seen to some extent in Figure 2.3. It is more obvious in Figure 2.5.

Figure 5: MARGINAL HISTOGRAMS AND BIVARIATE SCATTERPLOTS FOR X N(0,1), Y N(0,1), SIMULATED BY INVERTING CONDITIONAL COPULAS (tau = 0.6)



MARGINAL HISTOGRAMS AND BIVARIATE SCATTERPLOTS FOR X N(0,1), Y N(0,1), SIMULATED BY INVERTING CONDITIONAL COPULAS (tau = 0.6)



Smith (2003) and Trivedi and Zimmer (2005) also provide an illustration of copulas from this perspective, though slightly differently. Figure 2.5 contains simulated bivariate normal data, with correlation $\rho = 0.8$. Each contains its own representation of the dependence structure. The differences are more marked than in Figure 2.3 because the tails of the distribution can be seen more easily. The limited dependence of the FGM copula in particular stands out in this comparison. So too does that of the AMH copula, for which $\theta \in [-1, 1) \rightarrow \tau \in \left[-0.1817, \frac{1}{3}\right]$ (Nelsen 2006). In this case the upper limit is reached (for standard bivariate normality with $\rho = 0.8$, $\tau \approx 0.6$).

The Gumbel and Clayton copulas are also noticeable in their different and opposite capture of association in the tails of the bivariate distribution. The Gumbel copula exhibits a narrower spread in particular at the right-hand corner $(X \rightarrow 1, Y \rightarrow 1)$, reflecting its better capture of tail dependence. The Clayton on the other hand is known as a limiting lower tail copula, which can be seen in its particular drift at the left of the bivariate distribution $(X \rightarrow 0, Y \rightarrow 0)$.

The Plackett exhibits a similar trend in Figure 2.3, though not Figure 2.5, but without maintaining the correlation structure as strictly. This is a result the limits some copulas can place upon simulation: the Plackett copula is constructed according to an odds-ratio, such that it does not then naturally invert back into data: re-ranking is based on this odds-ratio, rather than re-ranking according to correlation.

5.1 Modelling tail dependence

A given distribution function can display stronger (positive or negative) dependence in its upper or lower quadrants, as was seen with the copulas illustrated in Figure 2.5. Positive dependence, for example, indicates that high-high and low-low combinations of two random variables are more likely than cross-combinations. Positive upper quadrant dependence indicates that high-high combinations are more likely than any other combination, including low-low combinations. The prices of tech stocks during the late 1990s would be a good example. Conversely many tech stocks in the early part of this decade would therefore be a good example of lower quadrant dependence.

This non-linearly structured dependence is the most useful property of copulas when estimating dependence. In the finance, insurance and risk fields this is the case: regression, covariates and similar methods are not of intrinsic interest. Only the structure of the dependence between two or more random variables is. Each univariate CDF F from the preceding discussion can be fitted non-parametrically and forgotten. Copulas with more parameters of association are important in this regard, such as the two-parameter AMH copula in Equation (31). To create these, the Laplace transformations considered in multivariate Archimedean copulas are again employed. A more general bivariate mixture of max-id distribution functions is given by (Joe 1997)

$$\begin{aligned} C\left(u,v\right) &= \psi\left(-\ln K\left(e^{-\psi^{-1}(u)}, e^{-\psi^{-1}(v)}\right)\right) &\text{(41a)} \\ &= \psi_\delta\left(-\ln K_\theta\left(K_\theta^{-1}\left(e^{-\psi_\delta^{-1}(u)}\right), K_\theta^{-1}\left(e^{-\psi_\delta^{-1}(v)}\right)\right)\right) &\text{(41b)} \end{aligned}$$

where again K is a bivariate copula: C becomes a two-parameter copula when Laplace transformation ψ_δ and the copula K_θ are parameterised separately.[26] Thus if K_θ is taken to be, for example, the Galambos copula from Equation (20d) and ψ_δ is the Gamma-form Laplace transformation used above, where $\psi\left(s\right) = \left(1+\delta s\right)^{\frac{-1}{\delta}}$, then C is given by

$$C\left(u,v;\delta,\theta\right) = \left(u^{-\delta} + v^{-\delta} - 1 - \left[\left(u^{-\theta}-1\right)^{-\theta} + \left(v^{-\theta}-1\right)^{-\theta}\right]^{\frac{1}{\theta}}\right)^{\frac{-1}{\delta}} \tag{42}$$

This is positive such that $\delta \geq 0$, $\theta > 0$, and can generalise into other families. Although it is not shown here, these methods generate representative parameters quantifying tail dependence: for this mixture of Galambos upper and lower tail dependence are given by $\left(\left(2-2^{\frac{1}{\theta}}\right)^{\frac{-1}{\delta}}\right)$ and $\left(2^{\frac{1}{\theta}}\right)$ respectively. Joe (1997) discusses some of this family's other properties, including extreme value limits and concordance increases. In particular the lower extreme value of this extension is itself an extreme value copula at the lower Fréchet-Hoeffding bound.

Cebrián, *et al.* (2003) and Demarta and McNeil (2004) are two recent applications considering tail dependence. de Matteis (2001) also considers elements of tail dependence, generally as well as specific to Archimedean-class copulas. In practice it is not necessary to understand the construction of tail-dependent copulas: Joe (1997) has pointed out that there is no theoretical model underlying selection of, for example, a given Laplace transform. The example above used one that resulted in a relatively straightforward two-parameter family of Galambos copula. Prepared forms can be found (most extensively in Joe 1997), and/or used according to upper and/or lower tail dependence parameters. Simulations such as those in Figures 2.3 and 2.5 can also be used to select the most likely best-fitting copula.

5.2 Simulating dependent multivariate data

Simulation is a very useful application of the unique dependence due to each copula. It is relatively straightforward, making use of known programmes or generalisable procedures (Nelsen 2006. Some applications can be found in, for example, Romano 2002). Dynamic simulation with Markov processes (or Joe's (1997) serial dependence) contain some of the most extensive use of copulas and simulation in this regard (Darsow, *et al.* 1992; Kulpa 1997; Bouyé, *et al.* 2000; Roncalli 2001).

Correlation is not invariant to non-linear transformations: generating multivariate normal data and manipulating each margin via inversion will result in a given desired marginal distribution function, but the correlation given to the generated data will not be upheld in the resulting non-normal data. Due to the transform invariance of measures of dependence, and hence copulas, this will be the case when copulas are used to simulate multivariate dependent data. Thus generating data is straightforward, and inverting margins via copulas will not only retain the desired dependence but include a dependence structure unique to each copula used. The inverted marginal distribution can be of any form, ergo any dependent data should feasibly be able to be generated. As well as simulation in a dynamic or Markov context, simulation of specific cross-sectional dependence structure can be used when necessary for purely theoretic or methodological work.

6 Estimation procedures for copulas

Copulas can be estimated both parametrically and non-parametrically, using either a single-step or two-step approach. Using one approach or the other typically involves a trade-off of potential misspecification for computational convenience, either of which might be more important. Although these approaches are presented in general terms, the full set of alternatives is appropriate only to Archimedean copulas, which should be identified first. It is assumed however that the procedure can be undertaken for more than one copula in estimation, and information on goodness-of-fit can be used subsequently, to select the most appropriate copula.

Bouyé, *et al.* (2000) provide discussion on estimating copulas for financial analyses with a focus upon likelihood estimation, which is assumed to be generally understood. The procedure outlined in de Matteis (2001) is more accessible in general estimation, as is the discussion in Trivedi and Zimmer (2006).

6.1 Non-parametric estimation

Two procedures for non-parametric identification of copulas exist that are free of any specification of the functional form of the margins. The first of these is the so-called empirical copula, which is loosely equivalent to the method of empirical likelihood. The second is the Genest and Rivest (1993) identification for Archimedean copulas.

6.1.1 Non-parametric estimation of empirical copulas

Step 1: According to Deheuvels (1979), for some sample $X \in \{(x_1^t, ..., x_N^t)\}_{t=1}^T$ the empirical copula is given by (Bouyé, *et al.* 2000)

$$\hat{C}\left(\frac{t_1}{T}, ..., \frac{t_N}{T}\right) = \frac{\sum_{t=1}^T 1\left(x_1^t \leq x_1^{\langle t_1 \rangle}, ..., x_N^t \leq x_N^{\langle t_N \rangle}\right)}{T} \tag{43}$$

where $x_n^{\langle t_1 \rangle}$ are the order statistics of X_i $\forall$ $1 \leq t_1, ...t_N \leq T$.[27]

Step 2: From this specification, measures of rank correlation can be found so that empirical copulas can function as measures of dependence. Empirical copulas however are of limited practical use for applied research, in much the same way as the method of empirical likelihood.

6.1.2 Non-parametric estimation of Archimedean copulas

Non-parametric identification of Archimedean copulas is a more useful consideration. Rather than the methods of construction discussed above, or a more systematic selection based on goodness-of-fit, Genest and Rivest (1993) provide an empirical method of identification. Consider that copulas can be represented in terms of the measures of rank correlation, such that

$$\text{Kendalls's } \tau_C = 4 \int \int_{I^2} C(u, v)\, dC(u, v) - 1 \tag{44}$$

and

$$\text{Spearman's } \rho_C = 12 \int \int_{I^2} C(u, v)\, dudv - 3 \tag{45}$$

for continuous margins u, v. For any Archimedean copula C from Equation (21), C is uniquely determined by the function

$$K(t) = t - \frac{\varphi(t)}{\varphi'(t)} \tag{46}$$

where $K(t) \in [0,1]$. Nelsen (2006) also shows that $K(t)$ is a distribution function on $(0,1)$. The differences between Spearman's ρ and Kendall's τ are discussed in Nelsen (2006) and Fredericks and Nelsen (2007). For absolutely continuous distributions u and v the use of either is equivalent. No general guideline exists, suggesting which circumstances are preferred for one method or another.[28] For applied research purposes, the use of Kendall's τ can be more convenient as the functional form of the relationship with the copula parameter θ is available (Genest and Rivest 1993; Nelsen 2006).

Using the relationship seen in Equations (44), (25) where the sample is further assumed to be generated from an Archimedean copula $C(x_1, x_2)$, Genest and Rivest (1993) show that Kendall's τ is such that

$$\begin{aligned} \tau &= 4\int\int_{\mathbf{I}^2} C(u,v)dC(u,v) - 1 \\ &= 1 + 4\int_0^1 \frac{\varphi(t)}{\varphi'(t)}dt \end{aligned} \tag{47}$$

Then for X_1, X_2

$$\hat{\tau} = \binom{n}{2}^{-1} \sum_{i<j} sign\left[(X_{1i} - X_{1j})(X_{2i} - X_{2j})\right] \tag{48}$$

Estimation proceeds as follows.

Step 1: Estimate rank correlation (Kendall's τ or Spearman's ρ). This can depend upon the data - in Quinn (2007a), for example, Kendall's τ_b is used specifically for a joint distribution with one ordinal, and one continuous, marginal distribution function.

Step 2: Select the functional form of the generator function $\varphi(t)$ (i.e. select the copula to be estimated). The copula generated by $\varphi(t)$ can be selected according to known information, or characteristics desired in the joint distribution function.

Step 3: Solve Equations (47) and (48) for $\hat{\tau} = \tau$ according to the generator function $\varphi(t;\theta)$, selected in Step 2. Since, following Equation (47), one can say that τ is a strictly increasing function of θ, the estimate $\hat{\theta}$ can be found for $\hat{\tau}$.[29]

For example, using Equation (26) and some observed $\hat{\tau}$ one would solve for θ in

$$\begin{aligned} \hat{\tau} &= 1 + 4\int_0^1 \frac{\ln\left(\frac{1-\hat{\theta}(1-t)}{t}\right)}{\left(\frac{\hat{\theta}-1}{t-\hat{\theta}t+t^2}\right)} dt \\ &= 1 - \frac{\hat{\theta} + 6\ln\left(1-\hat{\theta}\right)\left(1-\hat{\theta}\right)^2}{9\hat{\theta}^2} \end{aligned} \tag{49}$$

This has no analytical solution, however substitution of the limits shows that $\tau \in \left[-0.181726, \frac{1}{3}\right]$ for $\theta \in [-1, 1)$. This is not necessarily a procedure that needs to be undertaken every time: descriptions of the range of copulas (in terms of dependence) can be found in de Matteis (2001), Melchiori (2003) and Patton (2003). Quinn (2007a) uses procedures given in Perkins and Lane (2003) to apply this approach to his analysis of income-related inequalities in health, which he characterised using the dependence between Self-Assessed Health and income. Following work by Dardanoni and Lambert (2001), Genest and Rivest (1993) and Vandenhende and Lambert (2000, 2003), copulas were constructed for some countries in the European Community Household Panel, and used to rank-order them according to their association.

Although parametric estimation for covariates is required for regression analysis, non-parametric estimation of dependence and the parameters of each marginal distribution is sufficient for testing goodness-of-fit. If available, they can also be used as precise starting values in parametric estimation.

6.2 Parametric estimation

Parametric estimation of copulas (in a regression model, for example) can be undertaken in either a single-step, via Full-Information Maximum Likelihood (FIML), or over two steps, using a procedure known as Inference Functions for Margins (IFM). The latter is typically preferred when a copula cannot be constructed to match the desired joint distribution function (such as when dealing with a high order of dimensions, and/or negative dependence), or when the copula likelihood will be overly complex (such as when multiple dependence parameters are used in the copula).

6.2.1 Estimation via Full-Information Maximum Likelihood

FIML estimation follows the same procedure for copulas as for ordinary FIML estimation. Location and scale parameters are estimated in each marginal distribution function (the functional form of each of which is selected separately from the others) simultaneously with the copula parameters for dependence. Specifically, for some multivariate distribution function $H(X_1, .., X_n; \beta_1, .., \beta_n, \theta)$, consider the corresponding copula $C(F_1(X_1; \beta_1), .., F_n(X_n; \beta_n); \theta)$.

Step 1: Specify the functional forms of each marginal distribution $F_1(X_1; \beta_1), ...,$ $F_n(X_n; \beta_n)$. This can be done parametrically (by prior FIML estimation of each margin, for example), or non-parametrically (Matlab, for example, has some distribution-fitting tools). Selection can also be made visually, or according to any other prior information.

Step 2: Specify the functional form of the copula, $C(F_1(X_1; \beta_1), .., F_n(X_n; \beta_n); \theta)$. This can be done according to some knowledge of the dependence structure (such as with examination of the variance-covariance matrix) or any characteristics desired of the joint distribution.

Step 3: Construct the copula density $c\left(F_1\left(X_1;\beta_1\right),..,F_n\left(X_n;\beta_n\right);\theta\right)$ according to Equations (14a)-(15), as well as the likelihood and log-likelihood functions.

Step 4: The copula log-likelihood can be estimated according to any maximum-likelihood procedure. If point-estimates are available, for the parameters in either the copula or in the univariate marginal distribution functions, they should be given as precise starting values.

6.2.2 Estimation via Inference Functions for Margins

A two-step method due to Lee (1983), McLeish and Small (1988), Joe and Xu (1996), Xu (1996) and Joe (1997) is inferencing (IFM).[30] For some multivariate distribution function $H\left(X_1,..,X_n;\beta_1,..,\beta_n,\theta\right)$, consider the corresponding copula $C\left(F_1\left(X_1;\beta_1\right),..,F_n\left(X_n;\beta_n\right);\theta\right)$. The marginal parameter vectors $\beta_1,..,\beta_n$ can contain coefficients due to regression, and/or simple parameters for each distribution. The vector θ contains measures of assocation for the copula as a whole. The IFM method is a two-step procedure is as follows.

Step 1: Each marginal vector of coefficients $\beta_{i\in n}$ from marginal univariate distribution functions $F_1\left(X_1;\beta_1\right),...,F_n\left(X_n;\beta_n\right)$ is estimated first, and separately, to determine $\left\{\hat{\beta}_1,..,\hat{\beta}_n\right\}$ such that

$$\hat{\beta}_i = \arg\max_{\beta_i}\sum_{i=1}^{n}\ln f_i\left(x_i;\beta_i\right) \tag{50}$$

Step 2: The estimates $\hat{\beta}_i$ can be used to calculate the evaluated marginal distribution functions $\hat{F}_i\left(X_i;\beta_i\right)=F_i\left(X_i;\hat{\beta}_i\right)$. It is these, rather than $F_i\left(X_i;\beta_i\right)$, that are passed into the copula likelihood for estimation of θ.

Step 3: Using $\hat{F}_i\left(X_i;\beta_i\right)$, the copula likelihood $L\left(\hat{\beta}_1,..,\hat{\beta}_n,\theta\right)$ is maximised to find only $\hat{\theta}$ such that

$$\hat{\theta} = \arg\max_{\theta} \sum_{i=1}^{n} \ln c\left(\hat{F}_1(x_1;\beta_1),..,\hat{F}_n(x_n;\beta_n);\theta\right) \tag{51}$$

for some copula C with density $c\left(\hat{F}_1(x_1;\beta_1),..,\hat{F}_n(x_n;\beta_n);\theta\right)$.

Ordinarily, ML solves $(\partial L/\partial\beta_1,..,\partial L/\partial\beta_n,\partial L/\partial\theta) = 0$. Estimates from the method of IFM are such that $(\partial L_1/\partial\beta_1,..,\partial L_n/\partial\beta_n,\partial L/\partial\theta) = 0$ for univariate log-likelihoods $L_1,..,L_n$ as well as the joint likelihood L. This holds under regularity conditions, and Joe (1997, 2005) shows that the IFM method is efficient relative to the method of maximum likelihood, particularly for discrete marginal distributions with few categories. It is less so for more categories, and for continuous marginal distributions with strong dependence, although standard errors for the parameters in this approach can corrected post-estimation using jackknife methods.

6.3 Testing goodness of fit

Goodness-of-fit is a still-developing area of research related to copulas (see Fermanian 2005). They can be graphical or algebraic, representing in both cases a measure of distance between distributions (the copula C and the true - or empirical - distribution H).

6.3.1 Analytical goodness-of-fit

There are several approaches to testing goodness-of-fit analytically. Following de Matteis (2001), one can employ the Kolmogorov-Smirnoff test-statistic

$$T = \max_{x}\left\{\left|\hat{F}(x) - F(x)\right|\right\} \tag{52}$$

where $\hat{F}(x)$ and $F(x)$ are the empirical and theoretical CDFs, respectively. Although it has the advantage of being distribution-free, it suffers from a sensitivity of power to significant departures from the null (Janssen 2000). An alternative to this is the χ^2 statistic (Fermanian 2005)

$$T = \sum_{i=1}^{k} \frac{(f_i - np(x_i))^2}{np(x_i)} \tag{53}$$

where for a given class i, $np(x_i)$ is the supposed frequency and f_i the observed. Genest and Rivest (1993) and Frees and Valdez (1998) present a method for Archimedean-class copulas using the generator of the distribution function. Using Equation (46), the distance

$$\int (K_\theta(t) - K_n(t))^2 \, dK_n(t) \tag{54}$$

can be used analytically to measure goodness-of-fit, where $K_\theta(t)$ and $K_n(t)$ are the copula and empirical measures of $\Pr\left(C(u,v)\,|_{H(x_1,x_2)} \leq t\right)$ respectively. Fermanian (2005) also discusses multidimensional Chi-squared measures, which employ comparisons of the fit of copulas according to the dimensions in the multivariate distribution, rather than the value of the multivariate distribution function alone.

Following Joe (1997), two other approaches can be taken. The first is to use either the log-Likelihood directly, or information criteria such as the Akaike Information Criterion (AIC), given by $AIC = 2k - 2\ln(L)$ for log-likelihood L and k parameters, or Bayesian Information Criterion (BIC), given by $k\ln(n) - 2\ln(L)$ and where n is the sample size. Models do not need to be nested for comparison, and information criteria are particularly useful for copulas estimated via IFM, when the multivariate log-Likelihood is immediately available to be used. Empirical applications of copulas employ this method.

A second approach recommended by Joe (1997) is analysis of the predictive ability of the models estimated. That is, some comparison of the predicted summaries from the models with the observed summaries of the data itself. This is useful for IFM also, in particular for models whose purpose is predictive, rather than explanatory, as a check on the stability of the predictions under different approaches to inferencing.

6.3.2 Graphical goodness-of-fit

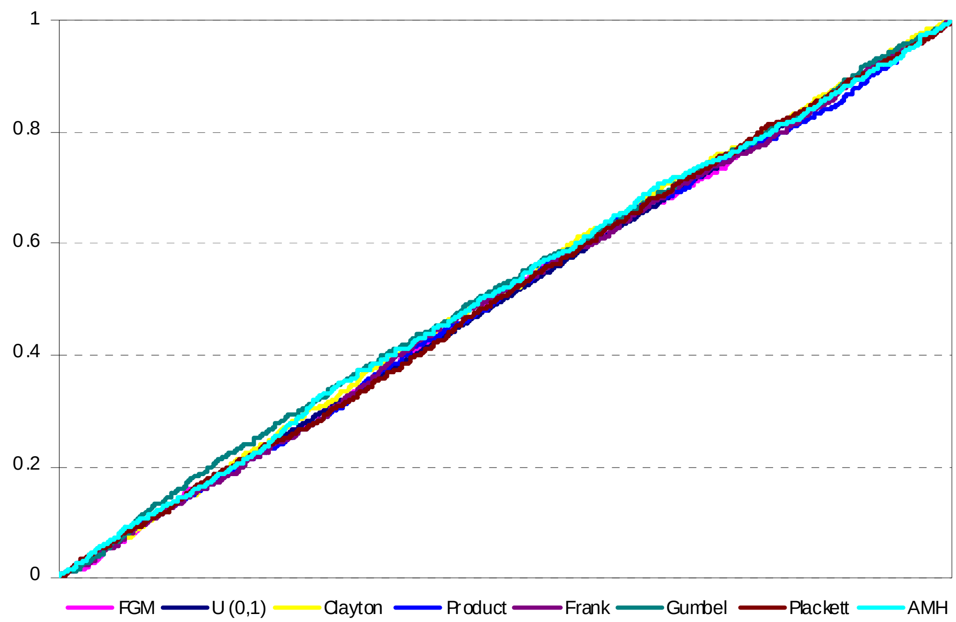
Goodness-of-fit can be tested graphically also: since both the copula C and empirical or true distribution function H are uniform, comparison can be made visually. The Genest and Rivest (1993) $K_\theta(t)$, for example, can be used in both analytical and graphical methods, since $K_\theta C(u,v)$ has a Uniform distribution. Therefore any evaluation of this distribution plotted against the standard uniform will qualify, giving so-called QQ-plots, or quantile-quantile plots, a graphical measure of the distance between a copula and some 'true' distribution. This method also allows in turn more straightforward numerical measures of dominance, where necessary for comparison.

Rather than the multivariate copula function, the conditional copula can be used. This is $C_u(u,v)$, where

$$\begin{aligned} C_u(u,v) &= \frac{\partial}{\partial u} C(u,v;\theta) \\ &\sim U(0,1) \end{aligned} \tag{55}$$

This holds when the arguments are uniformly-distributed random variables: i.e., after the transformation $x_1 \to u$, $x_2 \to v$.[31] This measures the fit of a distribution function, also compared to the standard uniform. Since any distribution function is a function mapping data from n-dimensions of Uniform distributions to the $[0,1]$ plane, or $[0,1]^n \to [0,1]$, the proximity of evaluations based upon each conditional distribution to the standard uniform will be an indicator of goodness-of-fit. Conditional copulas adequately accommodate the effect the dependence structure of each copula will have on the distribution, and is a fairly easily-implemented test (Durrleman, *et al.* 2000; Fermanian 2005). Applying the conditional copulas to the data simulated in Figure 2.4, for example, can be seen in Figure 2.6.

Figure 6: QUANTILE-QUANTILE PLOTS FOR COMPARISON OF CONDITIONAL COPULAS (APPLIED TO COPULAS FROM FIGURE 2.3)



Notice that the QQ-plots for each copula are indistinguishable, visually. As in this example, this approach to comparing models can be constrained by close proximity of all models, to themselves and to the diagonal. Moreover, the approach is not appropriate for the Product copula.[32]

Numerically, following Fermanian (2005), a fairly straightforward Chi-squared test-statistic comparing each copula (as a bivariate distribution function) to the bivariate standard uniform distribution can be calculated. The bivariate approach suggested by Fermanian (2005) partitions the bivariate distribution into a contingency table, and compares it to a similar partioning of a bivariate standard Uniform $U(0,1)^2$ distribution. A Chi-squared statistic will measure the absolute relative difference between the bivariate distributions and standard Uniform, but within a bivariate framework.

One benefit to this approach is that the Product copula can be tested along with the other distributions, since the products of the empirical margins will not be exactly Uniform. The method will still be biased in favour of the Product copula, though, forcing some trade-off between this test and any statistical significance of the dependence parameter. The approach also does not work for higher dimensions, although the test can be replicated across all bivariate pairs of margins within the copula.

7 Discussion

The discussion here has covered several potential health-economic applications of copulas, according to various of their properties that make them most useful. The most practical and versatile families of copulas have been shown, as well as their particular characteristics and various methods for estimation. In particular multivariate copulas have been described in a manner that should make them more easily understood by applied researchers seeking more flexible multivariate distributions while retaining general dependence.

Health-economic analyses using the copula method are, so far, relatively uncommon, as befitting a method only slowly expanding in general econometric analysis. The advantages to the method however are wide-ranging, as are the potential applications. These include not only applications comparing the results of copulas to standard methods, taking advantage of the freedom to construct any bivariate distribution according to the distributional behaviour of each random variable, but also new applications not heretofore considered feasible without very strong distributional and dependence assumptions.

Acknowledgements

Murray Smith, Martin Forster, Andrew Jones, Nigel Rice, Nicholas Biddle and Gabriel Picone were instrumental in improving this paper, and their input is appreciated. Advice from the Health, Econometrics and Data Group at the University of York was also useful, and financial support by the Centre for Health Economics is acknowledged gratefully.

Notes

[1]Although empirical likelihood methods can be used to obtain parameter estimates, they are mostly useful when estimating the parameters of a distribution, not covariates. Moreover, in higher dimensions the empirical likelihood assumes independent random variables: dependence is accomodated though blockwise observation of the data, and/or reductions to independence (Kitamura 1997, Owen 2001; El Ghouch, et al. 2005).

Non-likelihood methods such as kernel density are less informative still. The concern here is with simultaneous estimation of dependent random variables, explained by covariates, and able to be specifically or approximately parameterised according to some known class of distribution(s). To this end, empirical likelihood is not considered to be a useful alternative to approximating joint or conditional likelihoods.

[2]As with the discussion of conditional distributions, transform invariance in this case is, for two random variables (X_1, X_2), invariance of $F_1(X_1)$ to the use of $F_2(X_2)$ or $F_3(X_2)$, if $F_2(.)$ and $F_3(.)$ are both almost surely increasing functions. Consequently, if the dependence between joint-uniformly distributed (X_1, X_2) is defined by some copula $C\left(X_1, X_2\right)$, then $C\left(X_1, X_2\right) = C\left(X_1, F_2(X_2)\right) = C\left(X_1, F_3(X_2)\right).$

[3]This is not usually considered problematic since the region outside this is not usually of interest (Smith 2003).

[4]This discussion is not restricted to the use of conditional distributions for convenience: they can be necessary also because conditioning a distribution on past observations is more informative (analysis with Markov processes is an example; Fermanian and Wegkamp 2004). Statistically there is no difference.

[5]According to invariance to transformations, if $F_2(.)$ and $F_3(.)$ are both almost

surely increasing functions then the dependence between (X_1, X_2), as defined by $\theta(X_1, X_2)$, is such that $\theta(X_1, X_2) = \theta(X_1, F_2(X_2)) = \theta(X_1, F_3(X_2))$.

[6]Data were generated using Matlab, following the procedures in Perkins and Lane (2003). First, a bivariate CDF with correlation coefficient of 0.6 was generated, then each margin was inverted to create a joint distribution, still with correlation coefficent 0.6, but with Beta-distributed random variables in the margins.

[7]Trivariate reduction is such that $X \perp \frac{Y}{1-X}$ and, by symmetry, $Y \perp \frac{X}{1-Y}$.

[8]This is because correlation between X and Y is given by $\rho = -\sqrt{\frac{\alpha_1\alpha_2}{(\alpha_1+\alpha_3)(\alpha_2+\alpha_3)}} \in [-1, 0]$. This triangularity shows, as above, that X and Y must be similarly distributed for the joint Dirichlet $H(X, Y)$ to exist.

Devroye (1986) similarly criticises trivariate reduction. He uses a bivariate Gamma to show that, where marginal parameters are specified prior to reduction, there will be either precisely one solution for $\alpha_1, \alpha_2, \alpha_3$, or no solution.

[9]I.e. theoretical contours for the specifited copulas. They do not follow the observed data. Procedures for generating these contours in Maple or Mathematica are available from the author.

[10]The effect of laparoscopic surgery on QALYs gained (relative to standard treatments) was not statistically significant, though, in the regression models employed by Quinn (2005, 2007b).

It should be noted that the study by Quinn (2005, 2007b) was not intended to direct policy on hysterectomies; the illustrating example was used merely demonstrate the copula method.

[11]By comparison, Kendall's $\tau = \frac{(a+d)-(b+c)}{(a+d)+(b+c)}$

[12]The FGM copula is a first-order Taylor approximation of the more flexible Frank

copula. Its subsequent linearity in the margins has made it a popular exemplar (Smith 2003, Zimmer and Trivedi 2006).

[13]'Single-parameter' refers to the parameterisation of association: single-parameter families use only one parameter of association. Joe's (1997) presentation of single-and multiple-parameter copulas is particularly useful.

[14]This example would be the *maximum* domain of attraction, $G - F \in MDA$. By symmetry, the reflection of F would be the minimum domain of attraction if the lower extreme were a reflection of G (Joe 1997).

[15]In the multivariate case, from a proposition due to Resnick (1987), proven in Joe (1997) and Bouyé, Durrleman and Nikeghbali, *et al* (2000), it is known that multivariate extreme value copulas are copulas such that, for random variables $(X_1, ..., X_n)$, each with univariate extreme value distribution $\left(F_1^k(x_1), ..., F_n^k(x_n)\right)$ the extreme value copula $C(F_1(x_1), ..., F_n(x_n))$ is an extreme value distribution if $C^k(F_1(x_1), ..., F_n(x_n))$ and $C(F_1^k(x_1), ..., F_n^k(x_n))$ have the same limit distribution for any value of k. That is, C is in the domain of attraction if $C(F^m(a_{1m}x_1 + b_{1m}), ..., F^m(a_{nm}x_n + b_{nm})) \to G(x_1, ..., x_n)$.

[16]Algebraically, the Frank copula does not nest independence because of the term $\frac{1}{\theta}$. Nelsen (2006), however, demonstrates that $\lim\limits_{\theta \to 0} C_{Frank} = uv$, i.e. the Product Copula.

[17]The Clayton copula originally was given by Kimeldorf and Sampson (1975), and in Clayton 1978) as

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

where $\theta \in (0, \infty)$. This is also known as the Kimeldorf and Sampson (or, in Joe 1997, the B4 family) copula, as well as the Cook and Johnson or Pareto (Genest and

Mackay 1986 and Hutchison and Lai 1990, respectively). As a result of the extension to negative dependence, the Clayton copula no longer obeys Total Positivity of Order 2 (Nelsen 1998, Mari and Kotz 2001): it has a non-zero Lebesgue measure even with absolutely continuous u and v such that, in extended form, the Clayton copula must defined as the maximum of itself or zero.

[18]This will not necessarily mean that every copula has a closed form. Equation (17) illustrates this point: although tractable in terms of X_1 and X_2, the copula itself still requires integration, just like a standard multivariate normal distribution. In fact, this is the reason the Gaussian and t-copulas can be problematic in practice.

[19]The recursive model is essentially one of conditional distributions. This is not problematic when they are all normal, however a structural equation model does not permit the use of copulas, non-normal univariate distributions or flexible dependence structures, as discussed previously.

[20]Nelsen (2006) and Mari and Kotz (2001), whose presentation draws on that of Nelsen (2006), provide a different form for the multivariate FGM, giving

$$C(x_1, ..., x_8; \theta) = \prod_{i=1}^{8} F_i(x_i) \left(1 + \sum_{k=2}^{8} \sum_{1<j_1<...<j_k<n} \theta_{j_1,...j_k} [1 - F_{j_1}(x_{1j})] ... [1 - F_{j_k}(x_{j_k})]\right)$$

which contains not ${}^nC_2$ but $2^n - n - 1$, or $\sum_{i=2}^{8} {}^8C_i$. In the current problem this would mean 247 different $\theta_{j_1,...j_k}$ terms, which is not considered practicable. Estimation issues aside, the limits on θ in multivariate FGM copulas would render them all null.

[21]This is a condition assuring monotonicity of $\psi^{-1} \circ \phi$ mixtures, and hence the mixture-of-powers copula itself. Since known transforms are readily available in Joe's (1997) appendix, the requirement of infinite differentiability is not one the typical analyst will face.

[22]A univariate CDF F is such that F^{γ} is a CDF for all $\gamma > 0$, but this is not the case for multivariate distribution functions. In general the n-dimensional CDF H is such that H^{γ} is a CDF for all $\gamma > n-1$ (Joe 1997). Max-id is therefore a stronger dependence condition - it is equivalent to Total Positivity of Order 2 where, for $x_1 < x_2$ and $y_1 < y_2$, F is TP_2 if $F(x_1, y_1)F(x_2, y_2) > F(x_1, y_2)F(x_2, y_1)$.

[23]This is the reverse of TP_2, or what is called Reverse Rule of Order 2, or RR_2.

[24]Bivariate distributions are Positive Quadrant Dependent if higher values of one variable are correlated with higher values of the other, and vice versa (essentially $\tau > 0$). Positive Orthant Dependence is the multivariate equivalent. Note that these are weaker than (Multivariate) Total Positivity of Order 2, which implies positive (orthant) quadrant dependence (Mari and Kotz 2001 is a good reference for these dependence concepts).

[25]This is Joe's (1997) Laplace transform B, or Gamma-form LT, given by $\phi(s) = (1+\theta s)^{\frac{-1}{\theta}}$, where $\theta > 0$. The extension to negativity is, statistically, similar to that of the Clayton copula. That is, after extension the negative LTB is no longer strictly monotonic. This is also why Laplace-transformed multivariate copulas do not have mixture representations when extended to negative dependence.

[26]The previous discussion concerning min-id mixtures and RR_2 applies in this case also, with the exception that negative quadrant dependence is not a problem in the bivariate case, simplifying extensions to two-parameter families capturing negative association. However, co-movement is usually of more concern where tail dependence is of interest.

[27]Order statistics are $X_{(i)}$ such that, for some distribution of variables $\{x_{1,...,}x_n\}$

$$\begin{aligned} X_{(1)} &= \min\{x_{1,...,}x_n\} \\ X_{(2)} &= \min\left\{(x_{1,...,}x_n) - X_{(1)}\right\} \\ &\quad ... \\ X_{(n)} &= \max\{x_{1,...,}x_n\} \end{aligned}$$

[28]For some families of copulas, closed-form solutions for either Spearman's ρ or Kendall's τ may - or may not - be available. Packages such as Mathematica or Maple, however, can be employed to find numerical solutions.

[29]This is true for continuous margins. If the margins are discrete, $\hat{\tau}$ will not necessarily equal τ exactly. Quinn (2007c), following work by Vandenhende and Lambert (2000), examines this more closely.

[30]Lee (1983) does not refer to the method as IFM, though.

[31]This is because differentiation of the bivariate distribution function $H(x_1, x_2) = C(F_1(x_1), F_2(x_2))$ will, via the chain rule, leave a marginal density function in the conditional distribution function. In which case the conditional distribution function will uniformly distributed only if the random variables are uniformly distributed. By transforming the random variables first, only uniformly-distributed univariate distribution function are left as arguments in the copula.

[32]This is because, if the transformations $x_1 \rightarrow u$, $x_2 \rightarrow v$ are made empirically (such as with empirical CDFs or kernel distributions), the conditional Product copula will be an empircal marginal distribution function, which is standard Uniformly distributed.

REFERENCES

Abdous, B., Ghoudi, K., Khoudraji, A., 1999. Nonparametric estimation of the limit dependence function of multivariate extremes, Extremes, 22: 243-265.

Aitchison, J., 1986. The Statistical Analysis of Compositional Data. Chapman & Hall, New York.

Ali, M. M., Mikhail, N. N., Haq, M. S. 1978. A class of bivariate distributions including the bivariate logistic, Journal of Multivariate Analysis. 8: 405-412.

Almeida, C., Mouchart 2003. A note on a copula approach to polychoric correlations, Discussion Paper 03-22, Institute de Statistique, Université Catholique de Louvain, Louvain-la-Neuve.

Balia, S., Jones, A. M., 2007. Mortality, lifestyle and socio-economic status, Journal of Health Economics, forthcoming.

Belzunce, F., Semeraro, P., 2004. Preservation of some dependence concepts under mixtures and applications, Journal of Applied Probability, 414: 961-974

Bouyé, E., Durrleman, V., Nikeghbali, A., Riboulet, G., Roncalli, T. 2000. Copulas for finance: a reading guide and some applications. Groupe de Recherche Operationnelle, Credit Lyonnais.

Box-Steffensmeier, J. M., Smith, R. M., 1997. Heterogeneity and individual party identification, presented at the annual meetings of the Midwest Political Science Association, Chicago, 1996.

Briggs, A., 2005. Statistical methods for cost-effectiveness analysis alongside clinical trials, in: Jones, A. M., ed. Elgar Companion to Health Economics. Edward Elgar Publishing, Cheltenham.

Calabrese, E. J., Baldwin, L. A., 1998. A general classification of U - shaped dose - response relationships, Human and Experimental Toxicology, 17:353 - 364.

Cameron, A. C., Li, T., Trivedi, P. K., Zimmer, D. M., 2004. Modelling the differences in counted outcomes using bivariate copula models with an application to mismeasured counts. Econometrics Journal, 72: 566-584.

Cebrián, A. C., Denuit, M., Lambert, P., 2003. Analysis of bivariate tail dependence using extreme value copulas: an application to the SOA medical large claims database. Discussion Paper 03-02, Institute de Statistique, Université Catholique de Louvain, Louvain-la-Neuve.

Chernick, M. R., 1999. Bootstrap Methods: A Practitioner's Guide. Wiley, New York.

Claxton, K., Posnett, J., 1996. An economic approach to clinical trial design and research priority-setting. Health Economics, 5: 513-524.

Clayton, D. G. 1978. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence, Biometrika. 65:141-151.

Cook, R. D., Johnson, M. E., 1981. A family of distributions for modelling non-elliptically symmetric and multivariate data, Journal of the Royal Statistical Society B, 432: 210-218.

Dall'Aglio, G., Kotz, S., Salinetti, G., eds. 1991. Advances in Probability Distributions With Given Marginals: Beyond the Copulas. Kluwer Academic Publishers, Dordrecht.

Dardanoni, V., Lambert, P. J. 2001. Horizontal inequity comparisons, Social Choice. and Welfare. 18: 799-816.

Darsow, W., Nguyen, B., Olsen, E., 1992. Copulas and Markov processes, Illinois Journal of Mathematics, 36: 600-642.

De Castro, S., Goncalves, F., 2002. False contagion and false convergence clubs in stochastic growth theory, Discussion Paper 237, Departamento de Economia, Universidade de Brasilia, Brasilia.

Deb, P., Trivedi, P. K., 1997. Demand for medical care by the elderly: a finite mixture approach, Journal of Applied Econometrics, 12: 313-336.

Deheuvels, P., 1979. La fonction de dépendance empirique et ses propriétés, un test non paramétrique d'indépendance, Bulletin de la Classe des Sciences Académie Royale de Belgique, 65: 274–292.

Demarta, S., McNeil, A. J., 2005. The t-copula and related copulas, International Statistical Review 73: 111-129.

Drummond, M. F., Sculpher, M. J., Torrance, G. W., O'Brien, B. J., Stoddart, G. L., 2005. Methods for the economic evaluation of health care programmes. 3rd Edition. Oxford University Press, Oxford.

Durrleman, V., Nikeghbali, A., Roncalli, T., 2000. Which copula is the right one? Groupe de Recherche Operationnelle, Credit Lyonnais.

Fang, K. T., Kotz, S., Ng, K. W., 1989. Symmetric Multivariate and Related Distributions. Chapman and Hall, New York.

Fermanian, J-D., 2005. Goodness of fit tests for copulas, Journal of Multivariate Analysis, 951: 119–152.

Frank, M.J., 1979. On the simultaneous associativity of F(x,y). and x + y – F(x,y). Aequationes Mathematica, 19: 194–226.

Fredericks, G. A., Neslen, R. B., 20007. On the relationship between Spearman's rho and Kendall's tau for pairs of continuous random variables, Journal of Statistical Planning and Inference. 137: 2143-2150.

Frees, E. W., Valdez, E. A., 1998. Understanding relationships using copulas, North American Actuarial Journal 2: 1-25.

Galambos, J., 1987. The Asymptotic Theory of Extreme Order Statistics, 2nd ed. Kreiger Publishing Company, Malabar.

Garry, R., Fountain, J., Brown, J., Manca, A., Mason, S., Sculpher, M., Bridgman, S., Gray, J., Lilford, R., 2004. eVALuate hysterectomy trial: a multi-centre randomised trial comparing abdominal, vaginal and laparoscopic methods of hysterectomy. Health Technology Assessment, 826.: 1–154.

Garry, R., Fountain, J., Mason, S., Napp, V., , Hawe, J., Clayton, R., Abbott, J., Phillips, G., Whittaker, M., Lilford, R., Bridgman, S., Brown, J., 2004b. The eVALuate study: two parallel randomised trials, one comparing laparoscopic with abdominal hysterectomy, the other comparing laparoscopic with vaginal hysterectomy. British Medical Journal, 328: 129.

Genest, C., Rivest, L-P. 1993. Statistical inference procedures for bivariate Archim edean copulas, Journal of the American Statistical Association. 88 (423): 1034-1043.

Genius, M., Strazzera, E., 2004. The copula approach to sample selection modelling: an application to the recreational value of forests, Nota di Lavoro 73.2004. Fondazione Eni Enrico Matei, Milan.

Gold, M.R., Siegel, J.E., Russell, L.B., Weinstein, M.C., 1996. Cost-effectiveness in Health and Medicine. Oxford University Press, New York.

Gordy, M. B., 1998. A generalization of generalized beta distributions, Finance and Economics Discussion Series 1998-18, Board of Governors of the Federal Reserve System, Washington DC.

Greene, W. H., 2002. Econometric Analysis, 5th ed. Prentice-hall International Inc. New York.

Hutchinson, T. P., Lai, C. D., 1990. Continuous Bivariate Distributions, Emphasizing Applications. Rumsby Scientific Publishing, Adelaide.

Joe, H., 1997. Multivariate Models and Dependence Concepts. Chapman and Hall, London.

Joe, H., 2005. Asymptotic efficiency of the two-stage estimation method for copula-based models, Journal of Multivariate Analysis. 942: 401-419.

Joe, H., Hu, T., 1996. Multivariate distributions from mixtures max-infinitely divisible distributions, Journal of Multivariate Analysis, 572: 240-265.

Joe, H., Xu, J. J. 1996. The estimation method of inference functions for margins for multivariate models. Technical Report no. 166. Department of Statistics. University of British Columbia, Vancouver.

Johnson, N. L., Kotz, S., 1972. Distributions in Statistics: Continuous Multivariate Distributions. John Wiley & Sons, New York.

Johnson, N. L., Kotz, S., Balakrishnan, N., 1997. Discrete Multivariate Distributions. John Wiley & Sons, New York.

Jouini, M. N., Clemen, R. T., 1996. Copula models for aggregating expert opinions, Operations Research, 44: 444-457.

Keselman, H. J., Wilcox, R. R., Lix, L. M., 2005. A generally robust approach to hypothesis testing in independent and correlated groups designs, Psychophysiology, 404: 586-96.

Kolmogorov, A.N., 1950. Foundations of theory of Probability. Chelsea Publishing Company, New York.

Kowalski, C. J., 1973. Non-normal bivariate distributions with normal marginals, The American Statistician, 273: 103-106.

Kulpa, T., 1999. On approximation of copulas, International Journal of Mathematics and mathematical Science, 222: 259-269.

Lee, L-F., 1983. Generalized econometric models with selectivity, Econometrica, 512: 507-512.

Lee, M-J., 1996. Methods of Moments and Semiparametric Econometrics for Limited Dependent Variable Models. Springer, New York.

Lindskog, F., 2000. Modelling dependence with copulas and applications to risk management, Working Paper, RiskLab, ETH Zurich, Zurich.

Mardia, K. V., 1962. Multivariate Pareto distributions, Annals of Mathematical Statistics, 33: 1008-1015.

Mari, D.D., Kotz, S., 2001. Correlation and dependence. Imperial College Press, London.

de Matteis, R., 2001. Fitting copulas to data. Diploma thesis. University of Zurich, Zurich.

McLeish, D. L., Small, C. G., 1988. theory and Applications of Statistical Inference Functions. Lecture Notes in Statistics 44. Springer, New York.

Melchiori, M., 2003. Which Archimedean copula is the right one? http://www.defaultrisk.com/pp_corr_68.htm.

Nelsen, R. B., Quesada Molina, J. J., Rodríguez Lallena, J. A., Úbeda Flores, M., 2004. Best-possible bounds on sets of bivariate distribution functions, Journal Multivariate Analysis, 902: 348-358.

Nelsen, R. B., Úbeda Flores, M., 2004. A comparison of bounds on sets of joint distribution functions derived from various measures of association, Communications in Statistics - Theory & Methods, 3310: 2299-2305.

Nelsen, R. B. 1999. An Introduction To Copulas. Springer Verlag, New York.

Nelsen, R. B. 2006. An Introduction To Copulas. 2nd Ed. Springer Verlag, New York.

Newey, W., Steigerwald, D., 1997. Asymptotic bias for quasi maximum likelihood estimators in conditional heteroskedasticity models. Econometrica, 65: 587–599.

Owen, A., 2001. Empirical Likelihood. Chapman& Hall/CRC, New York.

Panas, E., 2005. Generalized beta distributions for stock market data: testing the U-shape pattern, Applied Economics, 37: 191-199.

Patton, A. J., 2003. Applications of Copula Theory in Financial Econometrics. PhD dissertation, University of California, San Diego.

Perkins, P., Lane, T., 2003. Monte-Carlo simulation in MATLAB using copulas. MATLAB News & Notes, November 2003.

Phelps, C. E., Mushlin, A., 1991. On the near equivalence of cost-effectiveness and cost-benefit analysis. International Journal of Technology Assessment in Health Care, 71: 12-21.

Pitt, M. D., Chan, D., Kohn, R., 2006. Efficient bayesian inference for Gaussian copula regression, Biometrika, 933: 537-554.

Prieger, J.E., 2000. A flexible parametric selection model for non-normal data with application to health care usage, Journal of Applied Econometrics 174: 367-392.

Prokhorov, A. B., Schmidt, P., 2006. Redundancy, Robustness, and Validity of Copulas in Likelihood Models, presented at the North American Winter Meeting of the Econometric Society, Boston, 2006.

Quinn, C., 2005. Generalisable regression methods for cost-effectiveness using copulas. Working Paper 05/13 of the Health, Econometrics and Data Group at the University of York, York.

Quinn, C., 2006. Alternative methods for estimating systems of health equations, Working Paper 06/05, Health, Econometrics and Data Group, University of York, York.

Quinn, C., 2007a. Estimating reduced-form systems of equations using copulas, mimeo.

Quinn, C., 2007b. Improved precision in cost-effectiveness analysis using copulas, mimeo.

Quinn, C., 2007c. Measuring income-related inequalities in health using copulas, mimeo.

Resnick, S. I., 1987. Extreme Values, Regular Variation, and Point Processes. Springer Verlag, New York.

Romano, C., 2002. Calibrating and simulating copula functions: an application to the Italian stock market, Working Paper 12/2002, Centro Interdipartimentale sul Diritto e l'Economia dei Mercati, Rome.

Roncalli, T., 2001. Modelling dependence in finance using copulas, Groupe de Recherche Operationnelle, Credit Lyonnais.

Satterthwaite, S. P., Hutchinson, T. P., 1978. A generalisation of Gumbel's bivariate logistic distribution, Metrika, 25: 163-170.

Savu, C., Trede, M. 2006. Hierarchical Archimedean copulas. Mimeo.

Seger, J., 2004. Extreme-value copulas, Nekst, 131: 38–41.

Shao, J., Tu, D., 1995, The Jackknife and Bootstrap. Springer-Verlag,New York.

Sklar, A., 1959. Fonctions de répartition à n dimensions et leur marges. Publications of the Institute of Statistics. University of Paris. 8: 229-231.

Smith, M. D., 2003. Modelling sample selection using archimedean copulas. Econometrics Journal, 6: 99–123.

Smith, M. D., 2005. Stochastic frontier models with dependent error components, Proceedings of the Econometric Society World Congress, London, 2005.

Stinnet, A., Mullahy, J., 1998. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. Medical Decision Making, 18: S68–S80.

Tajar, A., Denuit, M., Lambert, P., 2001. Copula-type representation for random couples with Bernoulli margins, Discussion paper 01-18, Institut de Statistique, Université catholique de Louvain, Louvain-la-Neuve.

Takahasi, K., 1965. Note on the multivariate Burr's distribution, Annals of the Institute of Statistical Mathematics, 17: 257-260.

Trivedi, P. K., Zimmer, D. M., 2006. Copula modeling: an introduction for practitioners, Foundations and Trends in Econometrics, 11: 1-110.

Vandenhende, F., Lambert, P., 2000. Modeling repeated ordered categorical data using copulas, Discussion Paper 00-25, Institut de Statistique, Université catholique de Louvain, Louvain-la-Neuve.

Vandenhende, F., Lambert, P. 2003. Improved rank-based dependence measures for categorical data. Statistics & Probability Letters. 63: 157-163.

Vandenhende, F., Lambert, P., 2004. Local dependence estimation using non-parametric archimedean copulas. Discussion Paper 04-02, Institut de Statistique, Université catholique de Louvain, Louvain-la-Neuve.

Vanness, D.J., Mullahy, J., 2005. Perspectives on mean-based evaluation of health care, Jones AM ed. Elgar Companion to Health Economics. Cheltenham, UK: Edward Elgar Publishing, forthcoming.

Vella, F., 1998. Estimating models with sample selection bias: a survey, Journal of Human Resources, 3: 127-144.

Verhoeven, P., McAleer, M., 2003. Fat tails and asymmetry in financial volatility models, Centre for International Research on the Japanese Economy Discussion Paper F-211, Centre for International Research on the Japanese Economy, The University of Tokyo, Tokyo.

Willan, A.R., Briggs, A.H., Hoch, J.S., 2004. Regression methods for covariate adjustment and subgroup analysis for non-censored cost-effectiveness data. Health Economics, 13: 461–475.

Wooldridge, J. M., 2002. Econometric Analysis of Cross Section and Panel Data. MIT Press, Cambridge.

Xu, J. J., 1996. Statistical modelling and inference for multivariate and lingitudinal discrete response data. PhD thesis. Department of Statistics. University of British Columbia.

Zarnoch, S. J., English, D. B. K., & Kocis, S. M. 2004. A model for evaluating dispersed outdoor recreation use estimation. In: Proceedings of the Second International Conference on Monitoring and Management of Visitor Flows in Recreational and Protected Areas, June 16-20, Rovaniemi, Finland.

Zimmer, D. M., Trivedi, P. K., 2006. Using trivariate copulas to model sample selection and treatment effects: application to family health care demand, Journal of Business and Economic Statistics, 24: 63-76.