Copulas in Econometrics[*]

Yanqin Fan[†]
Department of Economics
University of Washington
Box 353330
Seattle, WA 98195
Email: fany88@u.washington.edu

Andrew J. Patton[‡]
Department of Economics
Duke University
Box 90097
Durham, NC 27708-0097
Email: andrew.patton@duke.edu

This version: September 16, 2013

Abstract

Copulas are functions that describe the dependence between two or more random variables. This article provides a brief review of copula theory and two areas of economics in which copulas have played important roles: multivariate modeling and partial identification of parameters that depend on the joint distribution of two random variables with *fixed* or *known* marginal distributions. We focus on bivariate copulas, but provide references on recent advances in constructing higher dimensional copulas.

Key words: Sklar's theorem, multivariate models, Fréchet-Hoeffding inequaity, bounds

1   Introduction

This article reviews the growing literature on the use of copulas in econometric research. We focus on two primary applications: multivariate models constructed using copulas and partial identification of parameters that depend on the joint distribution of two random variables with *fixed* or *known* marginal distributions. While copulas are used as a modeling tool in the first application, they are used as a mathematical technique in the second application.

As a first introduction to copulas, consider a pair of random variables X and Y, with (univariate) marginal cumulative distribution functions (cdfs) F and G, and joint cdf H. Assume that their corresponding probability density functions (pdfs) exist, and denote them as f, g and h. Copula theory (in particular, Sklar's theorem, see Nelsen, 2006, for example) enables one to "decompose" the joint pdf h into the product of the marginal densities and the "copula density" denoted as c:

$$h(x,y) = c(F(x), G(y)) f(x) g(y). \tag{1}$$

Recall that the joint density of a pair of independent random variables is equal simply to the product of the marginal densities; in this case the copula density function, c, is equal to unity across its whole support. When the variables are *dependent*, the copula density will differ from unity and can be thought of as "reweighting" the product of the marginal densities to produce a joint density for dependent random variables. Sklar's theorem applies to discrete as well as continuous random variables, and in its more general form it is used to map the marginal cdfs to the joint cdf:

$$H(x,y) = C(F(x), G(y)). \tag{2}$$

This review will focus on the bivariate case for simplicity, but Sklar's theorem applies to general d-dimensional distributions as well:

$$
\begin{aligned}
h(y_1, ..., y_d) &= c(G_1(y_1), ..., G_d(y_d)) \prod_{i=1}^{d} g_i(y_i) \text{ and} \\
H(y_1, ..., y_d) &= C(G_1(y_1), ..., G_d(y_d)),
\end{aligned}
\tag{3}
$$

where G_i (g_i) for $i = 1, 2, ..., d$ are the marginal cdfs (pdfs) of the joint cdf H.

The usefulness of Sklar's theorem for multivariate modeling stems from the fact that the marginal distributions and the copula need not belong to the same family of distributions; they can be symmetric or skewed, continuous or discrete, fat-tailed or thin-tailed, and the joint cdf formed by using equation (2) with any copula function C, any univariate cdf F, and any univariate cdf G will be a valid cdf.

The characterization of a joint cdf in terms of its marginal cdfs and the copula function in (2) suggests two-step procedures for the identification and estimation of the joint cdf H. In the first step, the identification and estimation of the marginals are investigated and in the second step,

the identification and estimation of the copula function are analyzed. When a bivariate sample from the joint distribution is available, both the marginals and the copula or equivalently the joint distribution are often identified. In this case, as we discuss in Section 2 below, the joint cdf H or generally copula-based models can often be estimated in stages (marginal distributions, then copula), which simplifies the computational burden. But when only univariate samples from the marginal distributions are available, such as in randomized experiments, the joint distribution may only be partially identified, as the sample information may not be sufficient to identify the copula function. As we demonstrate in Section 3, the identified sets of the joint cdf H and more generally parameters that depend on H may be characterized as solutions to the general Fréchet problem studied in the probability literature. We present three important applications there: bivariate option pricing, evaluation of the VaR of a linear portfolio, and evaluation of the distributional treatment effects of a binary treatment.

In the rest of this section, we provide a brief introduction to each of the two areas of applications of copulas that we will focus on in Sections 2 and 3: multivariate modeling and partial identification of parameters that depend on the joint distribution of two random variables with fixed or known marginal distributions. We finish this section by mentioning some of the other work reviewing copulas.

1.1 A brief introduction to Sklar's theorem and copulas

A copula is a multivariate distribution function with uniform marginal distributions on $[0, 1]$. Sklar's (1959) theorem states that if H is a bivariate distribution function with marginal distribution functions F and G, then there exists a copula $C : [0, 1]^2 \to [0, 1]$ such that

$$H(x, y) = C(F(x), G(y)) \text{ for all } (x, y) \in \mathcal{R}^2. \tag{4}$$

If F and G are continuous, then the copula C in equation (4) is unique,[1] else the copula is uniquely determined only on range of F and G. Conversely, for any marginal distributions F, G and any copula function C, the function $C(F(\cdot), G(\cdot))$ is a bivariate distribution function with marginal distributions F, G. This theorem provides the theoretical foundation for the widespread use of copulas in generating multivariate distributions from univariate distributions. The fact that the copula function C, the marginal distribution functions F, G in equation (4) are not necessarily of the same type allows the researcher a great deal of flexibility in specifying a multivariate distribution.

The copula is sometimes called the "dependence function," see Nelsen (2006) and Joe (1997), as it completely describes the dependence between two random variables: any measure of dependence that is scale invariant (i.e., is not affected by strictly increasing transformations of the underlying variables) can be expressed as a function of the copula alone. Such dependence measures include

---

[1]Genest and Nešlehova (2007) present an excellent discussion on copulas for discrete data.

Kendall's τ, Spearman's ρ, and tail dependence coefficients. Importantly, Pearson's *linear* correlation coefficient cannot be expressed in terms of the copula alone; it also depends on the marginal distributions. (This is revealed through the fact that linear correlation is known to change when a nonlinear transformation, e.g. the logarithm or exponential, of one or both of the variables is applied.)

Recall that the "probability integral transform" of a random variable X with distribution function F is defined as:

$$U = F(X). \tag{5}$$

If F is continuous, then U will have the $Unif(0,1)$ distribution, regardless of the original distribution F. If we define the probability integral transformation of Y as $V = G(Y)$, then the joint distribution of (U, V) is the copula of the original random variables (X, Y). That is, if $(X, Y) \sim H = C(F, G)$, then $(U, V) \sim C$.

Sklar's theorem can also be extended to apply to conditional distributions, see Patton (2006a), which is useful for forecasting and time series applications. In particular, such models can accommodate dynamics, such as time-varying conditional volatility (e.g., ARCH or stochastic volatility) and time-varying conditional dependence (correlation, or other measures). Let $\{(X_t, Y_t)\}_t$ denote a stochastic process and $\mathcal{F}_t$ denote an information set available at time t, and let the conditional distribution of $(X_t, Y_t)|\mathcal{F}_{t-1}$ be H_t, with conditional marginal distributions F_t and G_t. Then

$$H_t(x, y|\mathcal{F}_{t-1}) = C_t(F_t(x|\mathcal{F}_{t-1}), G_t(y|\mathcal{F}_{t-1})|\mathcal{F}_{t-1}). \tag{6}$$

The complication that arises when applying Sklar's theorem to conditional distributions is that the information set used for the margins and the copula must be the same; if different information sets are used then the resulting function H_t can no longer be interpreted as a multivariate conditional distribution. We refer interested readers to Fermanian and Wegkamp (2012) for analysis of the case when differing information sets are used.

Given the abundance of univariate time series models, it is natural to build multivariate time series models from existing univariate models. The copula approach turns out to be very convenient for this purpose thanks to Sklar's theorem or its conditional version (6). Indeed, recent work in empirical finance and insurance provide ample evidence on the success of multivariate time series models constructed by combining univariate time series models via the copula approach in risk management and in modelling the (nonlinear) dependence among different economic and financial series, see Section 2.4 for references. Commonly used parametric copulas in these applications include the Gaussian or Normal copula, the Student's t-copula, the Frank copula, the Gumbel copula, and the Clayton copula. We provide the Gaussian (Normal) copula and the Student's t-copula below and refer interested readers to Joe (1997) and Nelsen (1999) for properties of other parametric copulas.

Example 1 (Normal copula): the d-dimensional normal or Gaussian copula is derived from the d-dimensional Gaussian distribution. Let Φ denote the scalar standard normal cdf, and $\Phi_{\Sigma,d}$ the d-dimensional normal distribution with correlation matrix Σ. Then the d-dimensional normal copula with correlation matrix Σ is

$$C\left(\mathbf{u};\Sigma\right) = \Phi_{\Sigma,d}(\Phi^{-1}(u_1),...,\Phi^{-1}(u_d)),$$

whose copula pdf is

$$c\left(\mathbf{u};\Sigma\right) = \frac{1}{\sqrt{\det(\Sigma)}} \exp\left\{ -\frac{\left(\Phi^{-1}(u_1),...,\Phi^{-1}(u_d)\right)' \left(\Sigma^{-1} - I_d\right) \left(\Phi^{-1}(u_1),...,\Phi^{-1}(u_d)\right)}{2} \right\}.$$

Example 2 (Student's t-copula): the d-dimensional Student's t-copula is derived from the d-dimensional Student's t-distribution. Let T_ν be the scalar standard Student's t distribution with $\nu > 2$ degrees of freedom, and $T_{\Sigma,\nu}$ be the d-dimensional Student's t distribution with $\nu > 2$ degrees of freedom and a shape matrix Σ. Then the d-dimensional Student's t-copula with correlation matrix Σ is

$$C\left(\mathbf{u};\Sigma,\nu\right) = T_{\Sigma,\nu}(T_\nu^{-1}(u_1),...,T_\nu^{-1}(u_d)).$$

The Student's t copula density is:

$$c\left(\mathbf{u};\Sigma,\nu\right) = \frac{\Gamma(\frac{\nu+d}{2})[\Gamma(\frac{\nu}{2})]^{d-1}}{\sqrt{\det(\Sigma)}[\Gamma(\frac{\nu+1}{2})]^d} \left(1 + \frac{\mathbf{x}'\Sigma^{-1}\mathbf{x}}{\nu}\right)^{-\frac{\nu+d}{2}} \prod_{i=1}^{d}\left(1 + \frac{x_i^2}{\nu}\right)^{\frac{\nu+1}{2}},$$

where $\mathbf{x} = (x_1,...,x_d)'$ with $x_i = T_\nu^{-1}(u_i)$.

Just as the univariate Student's t distribution generalizes the normal distribution to allow for fat tails, the Student's t copula generalizes the normal copula to allow for *joint* fat tails, i.e., an increased probability of joint extreme events.

The above two copulas both impose that the joint upper tails of the distribution are identical to the joint lower tails, ruling out the *asymmetric* dependence often observed in asset return data. Asymmetric dependence may be modeled via certain Archimedean copulas such as the Gumbel copula, the Clayton copula (see Nelsen (2006) for example), the skewed t-copula of Demarta and McNeil (2005) and Christoffersen et al. (2012), or the factor copula model of Oh and Patton (2012).

1.2 An introduction to the Fréchet-Hoeffding inequality and correlation bounds

Besides its role in building multivariate econometric models from univariate models and in characterizing the dependence structure among multiple random variables, copula theory has also proven to be a useful mathematical tool in studying (partial) identification of parameters that depend on the joint distribution of several random variables with fixed or known marginal distributions.

Consider the bivariate case where the random variables denoted as X, Y have common support $\mathcal{R}$, the whole real line. For $(u,v) \in [0,1]^2$, let $M(u,v) = \max(u+v-1,0)$ and $W(u,v) = \min(u,v)$

denote the Fréchet-Hoeffding lower and upper bound copulas: $M(u,v) \leq C(u,v) \leq W(u,v)$. Then for any $(x,y) \in \mathcal{R}^2$ and any bivariate cdf H with marginal cdfs F, G, the Fréchet-Hoeffding inequality holds:

$$M(F(x), G(y)) \leq H(x,y) \leq W(F(x), G(y)). \tag{7}$$

The bivariate distribution functions $M(F(\cdot), G(\cdot))$ and $W(F(\cdot), G(\cdot))$ are referred to as the Fréchet-Hoeffding lower and upper bounds for bivariate distribution functions with fixed marginal distributions F and G. They are distributions of perfectly negatively dependent and perfectly positively dependent random variables respectively, see Joe (1997) and Nelsen (1999) for more discussions. Some of the applications of the Fréchet-Hoeffding inequality include Manski (1988) who uses it in decision theory; Manski (1997) on the mixing problem in program evaluation; Ridder and Moffitt (2007; section 3.1) who discuss the use of the Fréchet-Hoeffding inequality in data combination contexts; Hoderlein and Stoye (2009) who use it to bound violations of the revealed-preference axioms in a repeated cross-section context; and Tamer (2010) who presents an application of the Fréchet-Hoeffding inequality in the partial identification of the joint distribution of potential outcomes of a binary treatment where we observe one of the potential outcomes.

Oftentimes one may be interested in parameters that are functionals of the joint cdf H such as the correlation coefficient of X and Y. Suppose X and Y have finite means μ_X, μ_Y, finite variances σ_X^2, σ_Y^2, and correlation coefficient ρ. It is known that ρ lies between -1 and 1 and $|\rho| = 1$ if and only if X and Y are perfectly linearly dependent. Hoeffding (1940) (see also Mari and Kotz (2001)) shows that the covariance of X and Y has an alternative expression:

$$Cov(X,Y) = \int \int [H(x,y) - F(x)G(y)]dxdy. \tag{8}$$

Suppose the marginal distributions are *fixed*. Then applying the Fréchet-Hoeffding inequality to (8) implies sharp bounds on the correlation coefficient of X and Y with fixed marginals: $\rho_L \leq \rho \leq \rho_U$, where

$$
\begin{aligned}
\rho_L &= \frac{\int \int [M(F(x), G(y)) - F(x)G(y)]dxdy}{\sigma_X \sigma_Y} \text{ and} \\
\rho_U &= \frac{\int \int [W(F(x), G(y)) - F(x)G(y)]dxdy}{\sigma_X \sigma_Y}.
\end{aligned} \tag{9}
$$

It is known that ρ_L and ρ_U are sharp, see Embrechts et al. (2002) and McNeil et al. (2005) for detailed discussions on this result and other properties and pitfalls of the correlation coefficient as a dependence measure. It is interesting to point out that $\rho \in [-1,1]$ is implied by the Cauchy-Schwartz inequality and $|\rho| = 1$ if and only if the random variables X and Y with cdfs F and G are perfectly linearly dependent on each other. However when F and G belong to different families of distributions such as one continuous and one discrete, it is typically the case that $[\rho_L, \rho_U]$ is a

proper subset of $[-1, 1]$. In other words, for some marginal cdfs F and G, the random variables X and Y with these cdfs F and G can never be perfectly linearly dependent on each other.

As a simple application of the correlation bounds in (9), consider a randomized experiment on a binary treatment with two potential outcomes. The sample information contains two independent random samples, one on each potential outcome and thus identifies the marginal distributions. But it has no information on the copula of the two potential outcomes besides that contained in the marginal distributions. So $[\rho_L, \rho_U]$ is the identified set for the true correlation coefficient of the potential outcomes. Heckman et al. (1997) use ρ_L and ρ_U to bound the variance of the individual treatment effect and the correlation coefficient of the two potential outcomes for randomized experiments. Fan, Sherman, and Shum (2013) apply (9) to bounding counterfactual distributions and treatment effect parameters when outcomes and covariates are observed in separate datasets.

1.3 Other reviews of copula theory

Nelsen (2006) and Joe (1997) are two key textbooks on copula theory, providing clear and detailed introductions to copulas and dependence modelling, with an emphasis on statistical foundations. Cherubini et al. (2004) present an introduction to copulas using methods from mathematical finance, and McNeil et al. (2005) present an overview of copula methods in the context of risk management. Mikosch (2006) and the associated discussions and rejoinder contain a lively discussion on the value of copulas in multivariate modeling. Genest and Favre (2007) present a description of semiparametric inference methods for iid data with a detailed empirical illustration. This article draws on Fan (2010), Fan, Guerre, and Zhu (2013), and Patton (2013). We refer interested readers to Patton (2013) for a detailed review of copula-based methods for economic forecasting and for empirical examples illustrating some commonly-used methods, and to Fan, Guerre, and Zhu (2013) for a systematic treatment of partial identification and inference for parameters that depend on the joint distribution of two random variables with fixed or known marginal distributions.

2 Copulas and Multivariate Models

This section describes some of the key steps in the estimation of copula-based multivariate models. The majority of applications of copula-based multivariate models are in time series, and so we will make such models the focus of this section.[2] Models for iid data can essentially be treated as a special case of these. We will consider three key aspects of the problem: model specification, estimation and inference, and goodness-of-fit testing. A more detailed discussion of these topics is presented in Patton (2013), on which the discussion below is based.

---

[2]For discussion of copula-based *univariate* models, see Darsow, *et al.* (1992), Ibragimov (2009), Beare (2010), Chen and Fan (2006b) and Chen, *et al.* (2009) for example.

2.1 Model specification

A majority of applications of copula models for multivariate time series build the model in stages, starting with aspects of the marginal distributions, and then moving on to the copula. For example, it is common to assume some parametric models for the conditional means and variances of the individual variables:

$$
\begin{aligned}
E\left[X_t|\mathcal{F}_{t-1}\right] &\equiv \mu_x\left(\mathbf{Z}_{t-1}, \alpha_{xo}\right), \\
V\left[X_t|\mathcal{F}_{t-1}\right] &\equiv \sigma_x^2\left(\mathbf{Z}_{t-1}, \alpha_{xo}\right),
\end{aligned}
\tag{10}
$$

where $\mathbf{Z}_{t-1} \in \mathcal{F}_{t-1}$, $\mu_x\left(\cdot, \cdot\right)$ and $\sigma_x^2\left(\cdot, \cdot\right)$ are of known form, α_{xo} is the finite dimensional unknown parameter vector, and similarly for Y_t with $\mu_y\left(\cdot, \cdot\right)$, $\sigma_y^2\left(\cdot, \cdot\right)$, and α_{yo}. Models that can be used here include many common specifications: ARMA models, vector autoregressions, linear and non-linear regressions, and others. It also allows for a variety of models for the conditional variance: ARCH and any of its numerous parametric extensions (GARCH, EGARCH, GJR-GARCH, etc.), stochastic volatility models, and others. Given models for the conditional means and variances, the standardized residuals can be constructed:

$$
\varepsilon_{x,t} \equiv \frac{X_t - \mu_x\left(\mathbf{Z}_{t-1}, \alpha_{ox}\right)}{\sigma_x\left(\mathbf{Z}_{t-1}, \alpha_{ox}\right)} \text{ and } \varepsilon_{y,t} \equiv \frac{Y_t - \mu_y\left(\mathbf{Z}_{t-1}, \alpha_{oy}\right)}{\sigma_y\left(\mathbf{Z}_{t-1}, \alpha_{oy}\right)}.
\tag{11}
$$

The conditional distributions of $\varepsilon_{x,t}$ and $\varepsilon_{y,t}$ are treated in one of two ways, either parametrically or nonparametrically. In the former case, this distribution may vary through time as a (parametric) function of $\mathcal{F}_{t-1}$-measurable variables (e.g., the time-varying skewed t distribution of Hansen, 1994), or may be constant. In the nonparametric case, the majority of the literature assumes that the conditional distribution is constant, and estimates it using the empirical distribution function, see Chen and Fan (2006a). We discuss this choice further in the next subsection.

A bivariate time series model can be constructed from the univariate time series models as specified in (10) by coupling the conditional distributions of $\varepsilon_{x,t}$ and $\varepsilon_{y,t}$ using a conditional copula. The (conditional) copula is the (conditional) distribution of the probability integral transforms of the standardized residuals:

$$
U_{xt} = F_t\left(\varepsilon_{x,t}\right) \text{ and } U_{yt} = G_t\left(\varepsilon_{y,t}\right),
\tag{12}
$$

where F_t and G_t denote the cdfs of $\varepsilon_{x,t}$ and $\varepsilon_{y,t}$. The majority of the literature considers parametric copula models,[3] and the conditional copula can be assumed constant, or allowed to vary through time. Joe (1997) and Nelsen (2006) present numerous parametric copula functions which can be used in applied work; these include the Gaussian (or Normal) copula, the Student's t copula, and Archimedean copulas, such as the Clayton, Gumbel and Frank copulas.

---

[3]On nonparametric estimation of copulas, see Genest and Rivest (1993) and Capéraà, *et al.* (1997) for *iid* data, and Fermanian and Scaillet (2003), Fermanian, *et al.* (2004), Sancetta and Satchell (2004) and Ibragimov (2009) for time series data.

An illustration of the types of bivariate distributions that can be obtained with copula-based models is presented in Figure 1. These bivariate distributions all have standard Normal marginal distributions and the copula parameters are calibrated to imply linear correlation of one-half. We see from this figure that even after imposing such similarity across the bivariate distributions, there remains a great deal of flexibility, arising from the flexibility in the choice of copula. The variation across these distributions also provides an indication of the identification problems discussed in Section 3.

[INSERT FIGURE 1 ABOUT HERE]

Suppose the conditional copula model is parametric with unknown parameter γ_o. Let $\theta_o \equiv (\alpha_{ox}, \alpha_{oy}, \gamma_o)$ denote the parameter vector for the entire multivariate distribution model. Methods for estimating θ_o include the full maximum likelihood and multi-stage approaches, where the former is typically asymptotically efficient but the latter is computationally easier and thus more commonly used in empirical work.

In the next subsection, we review the basic idea underlying the multi-stage approach to estimating copula-based models and point to work that establishes asymptotic properties of such estimators.

2.2 Estimation and inference

To simplify the exposition of the idea underlying the multi-stage approach to estimating copula-based models, we first assume away the dynamics and use the log-likelihood function for an i.i.d. sample denoted as $\{(x_t, y_t)\}_{t=1}^{T}$. As in most existing applications, we assume that the copula is parametric, but the marginal distributions could be either parametric or nonparametric. The decomposition of a joint density into the product of the marginal densities and the copula density, as in equation (1), reveals a convenient decomposition of the joint log-likelihood:

$$\frac{1}{T} \sum_{t=1}^{T} \log h(x_t, y_t; \theta) = \frac{1}{T} \sum_{t=1}^{T} \log f(x_t) + \frac{1}{T} \sum_{t=1}^{T} \log g(y_t) + \frac{1}{T} \sum_{t=1}^{T} \log c(F(x_t), G(y_t); \gamma),$$
(13)

where $\theta = (\alpha_{xo}, \alpha_{yo}, \gamma)$ when the marginal distributions are parametric, say, $F(x_t) \equiv F(x_t, \alpha_{ox})$ and $G(y_t) \equiv G(y_t, \alpha_{oy})$ for some finite dimensional parameters α_{ox} and α_{oy}; and $\theta = (F, G, \gamma)$ when the marginal distributions are nonparametric.

If the marginal distributions are parametric, the most natural estimation method is the full (one-stage) maximum likelihood. Under regularity conditions, see White (1994) for example, standard results can be used to show that the MLE is consistent and asymptotically Normal, and an estimator of the asymptotic covariance matrix can also be obtained using standard methods. The drawback of this approach is that even for relatively simple bivariate models, the number of parameters to be

estimated simultaneously can be large, creating a computational burden. This burden is of course even greater in higher dimensions.

If the parameters of the marginal distributions are separable from those for the copula, as suggested by our use of the $(\alpha_x, \alpha_y, \gamma)$ notation, then we may estimate those parameters in a first stage, and estimate the copula parameters in a second stage. Specifically let

$$\widehat{\alpha}_x = \arg\min_{\alpha_x} \sum_{t=1}^{T} \log f(x_t, \alpha_x) \text{ and } \widehat{\alpha}_y = \arg\min_{\alpha_y} \sum_{t=1}^{T} \log f(x_t, \alpha_y).$$

Then the two-stage estimator of the copula parameter γ can be computed as follows:

$$\widehat{\gamma} = \arg\min_{\gamma} \sum_{t=1}^{T} \log c(F(x_t, \widehat{\alpha}_x), G(y_t, \widehat{\alpha}_y); \gamma).$$

This two-stage approach is sometimes called "inference functions for margins" in the copula literature, see Joe (1997) and Joe and Xu (1996), though more generally this is known as multi-stage maximum likelihood (MSML) estimation, see White (1994). Of course, two- or multi-stage estimation will yield parameters that are less efficient than one-stage maximum likelihood,[4] although simulation studies in Joe (2005) and Patton (2006b) indicate that the loss of efficiency is not great. As for one-stage MLE, under regularity conditions, see White (1994) or Patton (2006b), the multi-stage maximum likelihood estimator is asymptotically normal but the asymptotic covariance matrix now takes a non-standard form, see Patton (2013) for details on how to obtain this covariance matrix.

An attractive feature of the copula decomposition of a joint distribution is that it allows the marginal distributions and copula to be estimated separately, potentially via different methods. Semiparametric copula-based models exploit this feature and employ nonparametric models for the marginal distributions and a parametric model for the copula. In such cases, the estimation of the copula parameter is usually conducted in two steps: the first step estimates the marginal cdfs via rescaled empirical cdfs:

$$\widehat{F}_T(x) = \frac{1}{T+1} \sum_{t=1}^{T} I\{x_t \leq x\} \text{ and } \widehat{G}_T(y) = \frac{1}{T+1} \sum_{t=1}^{T} I\{y_t \leq y\}$$

and the second step estimates the copula parameter by maximizing the estimated log-likelihood function with the marginal cdfs replaced by the rescaled empirical cdfs:

$$\widehat{\gamma}_S = \arg\min_{\gamma} \sum_{t=1}^{T} \log c\left(\widehat{F}_T(x_t), \widehat{G}_T(y_t); \gamma\right).$$

This estimator is sometimes called "canonical maximum likelihood" in this literature. The asymptotic distribution of this estimator was studied by Genest et al. (1995) for *iid* data and by Chen

---

[4]Song et al. (2005) propose a maximization-by-parts algorithm for copula based models which upon convergence generates asymptotically efficient estimators from the two-stage estimator of the copula parameter, see also Fan et al. (2012).

and Fan (2006b) for (univariate) time series data. Chen et al. (2006) propose a (semiparametric) sieve ML method that achieves full efficiency.

For the copula-based multivariate time series models described in the previous subsection with parametric or nonparametric distributions for $\varepsilon_{x,t}$ and $\varepsilon_{y,t}$, an additional step of pre-filtering must be done before applying the two-step approaches discussed above for i.i.d. data, so the estimated standardized residuals would replace the raw data. The pre-filtering step can be done via any existing methods for estimating the univariate models specified in (10). Chan et al. (2009) and Chen and Fan (2006a) provide conditions under which an asymptotic normal distribution for the estimator of the copula parameter is obtained, and provide a method for estimating the asymptotic covariance matrix. Rémillard (2010) provides additional analyses of estimators of this type, and suggests a bootstrap approach for conducting inference. Oh and Patton (2013) consider SMM-type estimation of the parameters of semiparametric copula-based models. One surprising result in Chan et al. (2009) and Chen and Fan (2006a) when the marginal distributions are nonparametric is that under regularity conditions, the asymptotic distribution of the estimator of the copula parameter is not affected by the filtering step so in practice one can just proceed as though the parameters characterizing the dynamics were *known,* see Rémillard (2010) for some similar results.

It is important to note that the asymptotic theory for the estimation of semiparametric copula-based models, such as that in Chan et al. (2009), Chen and Fan (2006a), Rémillard (2010) and Oh and Patton (2013), only applies when the conditional copula is *constant.* The marginal distribution means and variances are allowed to vary (subject to regularity conditions) but the vector of standardized residuals are assumed to be *iid.* Theory for the case where the conditional copula is time varying is not available in the literature to date. Fully parametric models, i.e., those where the marginal distributions as well as the copula are parametric, can handle time-varying conditional copula specifications, although verifying the regularity conditions required for the asymptotic distribution theory can be difficult.

2.3 Goodness of fit testing and model selection

As with any parametric model, multivariate models constructed using a parametric copula are subject to model misspecification, thus motivating goodness-of-fit (GoF) testing. Two tests that are widely used for GoF tests of copula models are the Kolmogorov–Smirnov (KS) and the Cramér–von Mises (CvM) tests, see Rémillard (2010), both of which are based on comparing the fitted copula cdf to the empirical copula:

$$\hat{C}_T (u_x, u_y) \equiv \frac{1}{T} \sum_{t=1}^{T} 1(\hat{U}_{xt} \leq u_x, \hat{U}_{yt} \leq u_y), \tag{14}$$

where $\hat{U}_{xt}$ is the probability integral transform of the (estimated) standardized residual, based either on the fitted parametric marginal distribution or on a nonparametric estimate of this and similarly

for $\hat{U}_{yt}$. As for parameter estimation, inference or GoF tests differ depending on whether the model under analysis is parametric or semiparametric. The presence of estimation error, whether parametric or nonparametric in nature, means that standard critical values for the KS and CvM tests cannot be used. Simulation-based alternatives are available, see Rémillard (2010) and Patton (2013) for further discussion.

An alternative, though similar, GoF test is based on Rosenblatt's transform, which is a form of multivariate probability integral transformation; see Diebold et al. (1999) and Rémillard (2010). This approach is particularly useful when the copula is time varying. In this approach the data are first transformed so that, if the model is correct, the data are independent $Unif(0,1)$ random variables, and then KS and CvM tests are applied to the transformed data. GoF tests that use the empirical copula of the data rely on the assumption that the conditional copula is constant, and so are inappropriate for time-varying copula models.

Related to the problem of GoF testing is that of model selection. Rather than comparing a fitted copula model to the unknown true copula, model selection tests seek to identify the best model(s) from a given set of competing specifications. The problem of finding the model that is best, according to some criterion, among a set of competing models (i.e., the problem of "model selection") may be undertaken either using the full sample (in-sample) of data, or using an out-of-sample (OOS) period. The treatment of these two cases differs, as does the treatment of parametric and semiparametric models. Below we discuss pair-wise comparisons of models; for comparisons of large collections of models see White (2000), Romano and Wolf (2005) and Hansen et al. (2011) for general models and Chen and Fan (2005, 2006a, 2007) for copula models.

Full sample (or "in sample") comparisons of nested copula models can generally be accomplished via a likelihood ratio test or a Wald test,[5] with null being that the smaller model is correct, and the alternative that the larger model is correct. For example, a comparison of a Normal copula with a Student's t copula can be achieved via a test that the inverse degree of freedom parameter is equal to versus larger than zero. Full sample comparisons of non-nested, fully parametric, copula-based models can be conducted using the test of Vuong (1989) for iid data, and Rivers and Vuong (2002) for time series data. The latter paper allows for a variety of (parametric) estimation methods and a variety of evaluation metrics. For copula applications their results simplify greatly if the marginals and the copula are estimated by maximum likelihood (one-stage or multi-stage) and we compare models using their joint log-likelihood. In such cases a test of equal accuracy can be conducted as a simple t-test that the per-period difference in log-likelihood values is mean zero. The only complication is that a HAC estimator (e.g., Newey and West (1987)) of the variance of the difference in log-likelihoods will generally be needed. Chen and Fan (2006a) consider a similar

---

[5]The problem becomes more complicated if the smaller model lies on the boundary of the parameter space of the larger model, or if some of the parameters of the larger model are unidentified under the null that the smaller model is correct. See Andrews (2001) and Andrews and Ploberger (1994) for discussion of these issues.

case to Rivers and Vuong (2002) for semiparametric copula-based models, under the assumption that the conditional copula is constant. Chen and Fan (2006a) show that the likelihood ratio t test statistic is again Normally distributed under the null hypothesis, although the asymptotic variance is slightly more complicated, as the estimation error coming from the use of the nonparametric marginal distributions must be incorporated.

In forecasting applications, out-of-sample (OOS) comparisons of models are widely used, see West (2006) for a general review of this topic. Diks et al. (2010) propose comparing copula-based models via their OOS log-likelihoods, using the theory of Giacomini and White (2006) to conduct inference.

2.4 Applications of copula-based multivariate models in economics and finance

One of the main areas of applications of copula-based multivariate models has been in financial economics. For applications to risk management, see Hull and White (1998) and Embrechts et al. (2002) and Kaas et al. (2009) on Value-at-Risk, and Rosenberg and Schuermann (2006) and McNeil et al. (2005) on general risk management issues. For applications to the pricing of credit derivatives see Li (2000), Giesecke (2004), Hofert and Scherer (2011) and Duffie (2011). Applications of copulas in other derivatives markets include Rosenberg (2003), Cherubini et al. (2004), van den Goorbergh et al. (2005), and Zimmer (2012). Applications of copula-based models to consider portfolio decisions are presented in Patton (2004), Hong et al. (2007), Garcia and Tsafack (2011) and Christoffersen et al. (2012).

The uses of copula-based analyses in other economic applications are broad and growing. Bonhomme and Robin (2009) use copulas to model the dynamics in a panel of earnings data, and Smith et al. (2010) similarly use them to capture dynamics in a longitudinal analysis of electricity prices. In health economics, Cameron et al. (2004) use copulas to study differences between actual and self-reported visits to a physician, Zimmer and Trivedi (2006) use them to model the health care demand of married couples, and van Ophem (2011) models the determinants of visits to a physician. Brendstrup and Paarsch (2007) use a semiparametric copula-based model to identify and estimate the joint distributions of valuations in a multi-object English auction, and Hubbard et al. (2012) use a similar model to study first-price sealed bid auctions. Rothe (2012) uses copulas to study distributional policy effects.

3 Copulas and Partial Identification

The Fréchet-Hoeffding inequality in (7) and the correlation bounds in (9) provide sharp bounds on the joint distribution of X, Y and the correlation coefficient of X, Y respectively when their marginal distributions are fixed. Both are examples of the general Fréchet problem which is concerned with finding sharp bounds on functionals of the joint distribution of X, Y with fixed marginals.

Let $\theta_o \equiv E_o[\mu(X,Y)] \in \Theta \subset \mathcal{R}$ denote the parameter of interest, where $\mu(\cdot,\cdot)$ is a real-valued measurable function and $E_o[\mu(X,Y)]$ denotes the expectation of $\mu(X,Y)$ taken with respect to the true joint distribution of X, Y. To distinguish between the true joint cdf of X, Y and any bivariate cdf with marginals F, G, we denote the true joint cdf of X, Y as H_o.

In this section, we will present the general Fréchet problem as that of partial identification of θ_o and review explicit solutions to the general Fréchet problem for two important classes of functions in Subsection 3.1. In Subsection 3.2, we present three applications of the general Fréchet problem in economics and finance—bivariate option pricing, VaR evaluation of a linear portfolio, and evaluation of distributional treatment effects of a binary treatment. We conclude this section by a brief discussion on inference in Subsection 3.3. The presentation in this section draws heavily on Fan, Guerre, and Zhu (2013).

3.1 The general Fréchet problem

The general Fréchet problem can be regarded as the problem of partial identification of θ_o when the marginal cdfs F and G are *known* or *fixed*. Let $\mathcal{C}$ denote the class of bivariate copula functions. For a general function μ, the identified set for θ_o is given by

$$\Theta_I = \{\theta \in \Theta : \theta = E_H[\mu(X,Y)], \text{ where } H = C(F,G) \text{ for some } C \in \mathcal{C}\}, \tag{15}$$

where E_H denotes the expectation taken with respect to H.

The Fréchet-Hoeffding inequality and the correlation bounds presented in (7) and (9) provide explicit characterizations of Θ_I for two examples of the μ function: $\mu(X,Y) = I\{X \leq x, Y \leq y\}$ for a given $(x,y) \in \mathcal{R}^2$ and $\mu(X,Y) = (XY - \mu_X \mu_Y)/(\sigma_X \sigma_Y)$. Both functions belong to a general class of functions μ known as super-modular functions. Subsection 3.1.1 below reviews explicit characterizations of Θ_I for general super-modular functions extending the Fréchet-Hoeffding inequality and the correlation bounds. Subsection 3.1.2 reviews another well-known example of the Fréchet problem—finding sharp bounds on the distribution functions of the four basic arithmetic operations such as the sum and product of X, Y with fixed marginals. The corresponding functions μ are neither super-modular nor sub-modular.

3.1.1 Super-modular functions

We first provide the definition of a super-modular function.

Definition 3.1 *A function $\mu(\cdot,\cdot)$ is called super-modular if for all $x \leq x'$ and $y \leq y'$,*

$$\mu(x,y) + \mu(x',y') - \mu(x,y') - \mu(x',y) \geq 0, \tag{16}$$

and sub-modular if $-\mu(\cdot,\cdot)$ is super-modular.

If $\mu(\cdot,\cdot)$ is absolutely continuous, then it is super-modular if and only if $\partial^2\mu(x,y)/\partial x\partial y \geq 0$ a.e. Cambanis et al. (1976) provide many examples of super-modular or sub-modular functions, see also Tchen (1980). Many parameters of the joint distribution of potential outcomes of interest, including the correlation coefficient of the potential outcomes and many inequality measures of the distribution of treatment effects can be written as functions of θ_o corresponding to some super-modular (sub-modular) functions μ and parameters that depend on the marginal distributions of potential outcomes only. Also many payoff functions of bivariate options are either super-modular or sub-modular, see Subsection 3.2 below.

Suppose $\mu(\cdot,\cdot)$ is a super-modular and right continuous function. Sharp bounds on θ_o can be found in Cambanis et al. (1976), Tchen (1980), and Rachev and Ruschendorf (1998).[6] Let θ_L and θ_U denote the lower and upper bounds on θ_o respectively. They are

$$\theta_L \equiv E_{H^{(-)}}[\mu(X,Y)] = \int_0^1 \mu\left(F^{-1}(u),G^{-1}(1-u)\right)du \text{ and}$$

$$\theta_U \equiv E_{H^{(+)}}[\mu(X,Y)] = \int_0^1 \mu\left(F^{-1}(u),G^{-1}(u)\right)du,$$

where $F^{-1}(u) = \inf\{x: F(x) \geq u\}$ denotes the u-th quantile of F, $G^{-1}(u) = \inf\{y: G(y) \geq u\}$ denotes the u-th quantile of G, $H^{(-)}(x,y) \equiv M(F(x),G(y))$, and $H^{(+)}(x,y) \equiv W(F(x),G(y))$.

Below we restate Theorem 2 in Cambanis et al. (1976) which generalizes the Fréchet-Hoeffding inequality and the correlation bounds presented in (7) and (9).

Lemma 3.1 *Suppose $\mu(x,y)$ is super-modular and right continuous. If $E_{H^{(-)}}[\mu(X,Y)]$ and $E_{H^{(+)}}[\mu(X,Y)]$ exist (even if infinite valued), then the identified set for θ_o is $\Theta_I = [\theta_L,\theta_U]$ when either of the following conditions is satisfied: (a) $\mu(x,y)$ is symmetric and $E[\mu(X,X)]$ and $E[\mu(Y,Y)]$ are finite; (b) there are some fixed constants $\overline{x}$ and $\overline{y}$ such that $E[\mu(X,\overline{y})]$ and $E[\mu(\overline{x},Y)]$ are finite and at least one of $E_{H^{(-)}}[\mu(X,Y)]$ and $E_{H^{(+)}}[\mu(X,Y)]$ is finite.*

The idea underlying the proof of this lemma is not difficult to understand. By definition,

$$\theta_o = \int\int \mu(x,y)\,dH_o(x,y) = \int\int \mu(x,y)\,dC_o(F(x),G(y)).$$

Under the conditions stated in the lemma, one can express the right hand side expression for θ_o as a non-decreasing function of the copula C_o and the fixed marginal cdfs F and G. Since $M(u,v) \leq C_o(u,v) \leq W(u,v)$ for all $(u,v) \in [0,1]^2$, we obtain the Lemma.

3.1.2 The distribution function of $(X+Y)$

Let $Z = X + Y$ with cdf $F_Z(\cdot)$. For a given $z \in \mathcal{R}$, let $\theta_o = E_o[\mu(X,Y)] = E_o[I\{Z \leq z\}]$. Then $\theta_o = F_Z(z)$. The sharp bounds on $F_Z(z)$ can be found in Makarov (1981), Rüschendorf (1982), and

---

[6]Results for sub-modular functions follow straightforwardly from the corresponding results for super-modular functions. To save space, we will not present results for sub-modular functions here.

Frank et al. (1987), see Lemma 3.2 below. Frank et al. (1987) demonstrate that copula provides a useful tool in finding sharp bounds on the distribution function of the sum of two random variables with fixed marginals. In this section, we present sharp bounds for the distribution function of Z and refer the reader to Schweizer and Sklar (1983), Williamson and Downs (1990), and Fan, Guerre, and Zhu (2013) for sharp bounds for the distribution functions of other arithmetic operations on X, Y.

Frank et al. (1987) demonstrate that their proof based on copulas can be extended to more general functions than the sum, see Williamson and Downs (1990) and Embrechts et al. (2003) for details.

Lemma 3.2 *Let*

$$
\begin{aligned}
F_{\min,Z}(z) &= \sup_{x \in \mathcal{R}} \max(F(x) + G(z-x) - 1, 0) \text{ and} \\
F_{\max,Z}(z) &= 1 + \inf_{x \in \mathcal{R}} \min(F(x) + G(z-x) - 1, 0).
\end{aligned}
\tag{17}
$$

Then the identified set for $F_Z(z)$ is $\Theta_I = [F_{\min,Z}(z), F_{\max,Z}(z)]$. If either $F(\cdot)$ or $G(\cdot)$ is a degenerate distribution, then for all z, we have $F_{\min,Z}(z) = F_Z(z) = F_{\max,Z}(z)$, so $F_Z(\cdot)$ is point identified.

Unlike the sharp bounds for super-modular functions in Lemma 3.1 which are reached at the Fréchet-Hoeffding lower and upper bounds for the distribution of X, Y (when X and Y are perfectly negatively dependent or perfectly positive dependent), the sharp bounds for $F_Z(z)$ are not reached at the Fréchet-Hoeffding lower and upper bounds for the distribution of X, Y. Frank et al. (1987) provide explicit expressions for copulas that reach the bounds on $F_Z(z)$.

3.2 Three applications of the general Fréchet problem

In this section, we present three important applications of the results reviewed in Section 3.1: bivariate option pricing; evaluation of the VaR of a linear portfolio; and evaluation of the distributional treatment effects of a binary treatment.

3.2.1 Bivariate option pricing

Let X, Y denote the values of two individual assets or risks and θ_o denote the price of a European-style option on X, Y with the discounted payoff $\mu(X, Y)$. Following Rapuch and Roncalli (2001) and Tankov (2011), we assume that the econometrician observes random samples of prices on single-asset options on X, Y and that there is no arbitrage. It is known from option pricing theory that there exists a risk-neutral probability measure denoted as Q such that the option price is given by the discounted expectation of its payoff under Q. That is, $\theta_o = E_Q[\mu(X, Y)]$.

Let H_o denote the distribution function implied by Q with marginal cdfs F and G. Then $\theta_o = E_o\left[\mu\left(X, Y\right)\right]$ is the price of such an option. The sample information allows the identification of the marginal distributions of X, Y under Q. For example, if X is the price of an asset at time T and call options on this asset with prices $P_X\left(K\right) \equiv E_Q\left[\exp\left(-rT\right)\left(X - K\right)_+\right]$ are available, where r is the interest rate and K is the strike price, then the cdf of X is given by

$$F\left(K\right) = 1 - \exp\left(rT\right)\frac{\partial P_X\left(K\right)}{\partial K}. \tag{18}$$

Many options have payoff functions that are either super-modular or sub-modular. For example, the payoff function of a call on the minimum with strike K is super-modular given by $\mu\left(X, Y\right) = \left(\min\left(X, Y\right) - K\right)_+$ and the payoff function of a worst-off call option is also super-modular given by $\mu\left(X, Y\right) = \min\left\{\left(X - K_1\right)_+, \left(Y - K_0\right)_+\right\}$, where $\left(x\right)_+ = \max\left(x, 0\right)$ and K_1, K_0 are strike prices. A basket option with payoff function: $\left(X + Y - K\right)_+$ is yet another example. We refer interested readers to Table 1 in Rapuch and Roncalli (2001) or Tankov (2011) for more examples.

Applying Lemma 3.1 to the payoff functions of bivariate options with super-modular payoff functions yield the identified sets for the bivariate option prices. For example, the identified set for the price of a call on the minimum with strike K is given by

$$\left[E_{H^{(-)}}\left[\left(\min\left(X, Y\right) - K\right)_+\right], E_{H^{(+)}}\left[\left(\min\left(X, Y\right) - K\right)_+\right]\right]. \tag{19}$$

The above bounds are first obtained by Rapuch and Roncalli (2001). Tankov (2011) establishes improved bounds when additional information on the dependence between X and Y is available. Fan, Guerre, and Zhu (2013) provide closed-form expressions for the above bounds and bounds on the price of a worst-off call option.

3.2.2 The worst VaR of a linear portfolio

Let X, Y denote the values of two assets or two individual risks, $Z = X + Y$ denote a linear portfolio of X, Y, and θ_o denote the VaR of the linear portfolio Z. The VaR of Z at level $\alpha \in (0, 1)$ is defined as the α quantile of the distribution of Z denoted as $F_Z^{-1}\left(\alpha\right)$. Although $F_Z^{-1}\left(\alpha\right)$ cannot be written as $E_o\left[\mu\left(X, Y\right)\right]$ for some function μ, the distribution function of Z can and the bounds on the VaR of Z follow straightforwardly from the bounds on its distribution function. So with slight abuse of notation, we also refer to the VaR of a linear portfolio of X, Y as an example of θ_o.

The VaR of Z is of interest in risk management and is identified when a bivariate sample from the joint distribution of X, Y is available. However when only univariate samples on X, Y are available, we can not point identify the VaR of Z. To resolve this issue, researchers have adopted the independence assumption on X, Y. This assumption is often violated, see McNeil et al. (2005) and references therein for detailed discussion. Without the independence or any other specific assumption on the dependence of X, Y, one can find sharp bounds on the VaR of Z by inverting

the sharp bounds on $F_Z(z)$ in Lemma 3.2: for $\alpha \in (0,1)$, it holds that

$$F_Z^{-1}(\alpha) \in \left[\inf_{u \in (\alpha,1)} [F^{-1}(u) + G^{-1}(\alpha - u + 1)], \sup_{u \in (0,\alpha)} [F^{-1}(u) + G^{-1}(\alpha - u)] \right]. \tag{20}$$

These are first established in Makarov (1981). The upper bound on $F_Z^{-1}(\alpha)$ in (20) is known as the worst VaR of Z, see Embrechts et al. (2003) and Embrechts et al. (2005). Kaas et al. (2009) present the worst VaR of Z when additional information on the dependence between X, Y is available.

3.2.3 Distributional treatment effects

Let D be a binary treatment indicator such that an individual with $D = 1$ receives the treatment with a continuous outcome X and an individual with $D = 0$ does not receive the treatment with a continuous outcome Y. In addition to the average treatment effect parameters such as the average treatment effect and the treatment effect for the treated, distributional treatment effect parameters such as the proportion of people who benefit from the treatment and the quantile of the distribution of treatment effects may be of interest as well. Parameters in the latter category depend on the copula of the potential outcomes X, Y.

When the marginal distributions of X, Y are identified,[7] the bounds reviewed in Subsection 3.1 can be used to bound distributional treatment effect parameters. We review the recent work in this area for randomized experiments and threshold-crossing models below.

Randomized experiments. Data from randomized experiments contain two independent univariate random samples, one on each potential outcome, so they point identify the marginal distributions of X, Y and thus point identify average treatment effects. But they do not point identify the copula of X, Y, because they contain no information on the dependence of X, Y besides that in the marginal distributions. Since distributional treatment effects such as the proportion of people receiving treatment who benefit from the treatment and the median of the individual treatment effect depend on the copula of X, Y, they are not point identified from randomized experiments. Fan and Park (2009, 2010, 2011) provide a systematic study of partial identification and inference for these distributional treatment effect parameters using Lemma 3.2. Below we review some of their results.

Let $\Delta = X - Y$ denote the individual treatment effect with cdf $F_\Delta(\cdot)$. Given the marginals F and G, sharp bounds on $F_\Delta(\delta)$ for δ in the support of the distribution of Δ can be obtained from Lemma 3.2, see also Williamson and Downs (1990): $F_\Delta^L(\delta) \leq F_\Delta(\delta) \leq F_\Delta^U(\delta)$, where

$$F_\Delta^L(\delta) = \sup_y \max(F(y) - G(y - \delta), 0) \text{ and } F_\Delta^U(\delta) = 1 + \inf_y \min(F(y) - G(y - \delta), 0). \tag{21}$$

---

[7]See Manski (2003) and references therein for scenarios under which the marginal distributions of X, Y are partially identified.

Note that the proportion of people receiving treatment who benefit from it is given by

$$P(X > Y | D = 1) = P(\Delta > 0 | D = 1) = 1 - F_\Delta(0 | D = 1),$$

where $F_\Delta(\cdot | D = 1)$ denotes the conditional cdf of Δ given $D = 1$. For ideal randomized experiments, $P(X > Y | D = 1) = 1 - F_\Delta(0)$. Applying the bounds in (21) to $F_\Delta(0)$ leads to the identified set for $P(X > Y | D = 1)$.

Inverting the bounds on $F_\Delta(\delta)$ in (21), we get: $Q_\Delta^L(\alpha) \leq F_\Delta^{-1}(\alpha) \leq Q_\Delta^U(\alpha)$, where

$$
\begin{aligned}
Q_\Delta^L(\alpha) &= \sup_{u \in (0,\alpha)} [F^{-1}(u) - G^{-1}(u + 1 - \alpha)] \text{ and} \\
Q_\Delta^U(\alpha) &= \inf_{u \in (\alpha,1)} [F^{-1}(u) - G^{-1}(u - \alpha)].
\end{aligned}
$$

Fan and Park (2011) explore these bounds to construct inference procedures for $F_\Delta^{-1}(\alpha)$ for ideal randomized experiments.

Latent-threshold crossing model. The identification results for randomized experiments extend straightforwardly to the selection-on-observables framework—D is independent of (X, Y) conditional on observable covariates, see Fan, Guerre, and Zhu (2013). When selection into treatment is based on not only observable but also unobservable covariates, Heckman (1990) and Heckman and Vytlacil (2005) provide a systematic study of the identification of average treatment effects in latent threshold crossing models. Building on the results in Heckman (1990), Fan and Wu (2010) establish the identified sets for distributional treatment effect parameters in the latent threshold crossing model of Heckman (1990). We review some of the results in Fan and Wu (2010) below.

Consider the semiparametric threshold crossing model with continuous outcomes in Heckman (1990):

$$
\begin{aligned}
X &= g_1(X_1, X_c) + U_1, \ Y = g_0(X_0, X_c) + U_0, \text{ and} \\
D &= I\left\{(W, X_c)'\gamma + \epsilon > 0\right\},
\end{aligned}
\tag{22}
$$

where X_1, X_0, X_c, W are observable covariates, U_1, U_0, ϵ are unobservable covariates, $g_1(x_1, x_c)$, $g_0(x_0, x_c)$ and the distribution of $(U_1, U_0, \epsilon)'$ are completely unknown.

The sample information contains observations on the covariates (X_1, X_0, X_c, W) and the treatment indicator D for each individual in the sample, but only contains observations on X for individuals with $D = 1$ and observations on Y for individuals with $D = 0$. When conditional on the observable covariates (X_1, X_0, X_c, W), ϵ is independent of (U_1, U_0), the selection-on-observables assumption holds; otherwise, the unobservable error ϵ affects both the individual's decision to select into treatment and her potential outcomes. Suppose the unobservable covariates are independent of the observable covariates. Heckman (1990) provides conditions under which the distributions of $(U_1, \epsilon)'$ and $(U_0, \epsilon)'$, $g_1(x_1, x_c)$, $g_0(x_0, x_c)$, and γ are point identified from the sample information

alone. However, the joint distribution of $(U_1, U_0)'$ is only partially identified, see Fan and Wu (2010) for details.

Lemmas 3.1 and 3.2 allow us to establish sharp bounds for distributional treatment effect parameters that depend on the copula of X, Y or U_1, U_0. For example, the covariance of U_1, U_0 can be bounded as follows:

$$
\int \left[\int \int M(F_{1|\epsilon}(u), F_{0|\epsilon}(v)) du dv \right] dF_\epsilon(\epsilon) \leq Cov\,(U_1, U_0)
$$
$$
\leq \quad \int \left[\int \int W(F_{1|\epsilon}(u), F_{0|\epsilon}(v)) du dv \right] dF_\epsilon(\epsilon), \tag{23}
$$

where $F_{j|\epsilon}(u)$ denotes the conditional cdf of U_j on ϵ and $F_\epsilon(\epsilon)$ is the cdf of ϵ. The bounds in (23) may be used to infer the sign of $Cov\,(U_1, U_0)$ and are typically narrower than those based on the marginal cdfs of U_1, U_0 only.

Similarly consider the distribution of $\Delta = X - Y$. Let $ATE = g_1(x_1, x_c) - g_0(x_0, x_c)$. Then $F_\Delta(\delta) = E\left[P(U_1 - U_0 \leq \{\delta - ATE\}|\epsilon)\right]$. Applying Lemma 3.2 to $P(U_1 - U_0 \leq \{\delta - ATE\}|\epsilon)$, we obtain the sharp bounds on the distribution function of treatment effects: $F_\Delta^L(\delta) \leq F_\Delta(\delta) \leq F_\Delta^U(\delta)$, where

$$
\begin{aligned}
F_\Delta^L(\delta) &= \int_{-\infty}^{+\infty} \left[\sup_u \max \left\{ F_{X|\epsilon}(u) - F_{Y|\epsilon}(u - \{\delta - ATE\}), 0 \right\} \right] dF_\epsilon(\epsilon), \\
F_\Delta^U(\delta) &= \int_{-\infty}^{+\infty} \left[\inf_u \min \left\{ 1 - F_{Y|\epsilon}(u - \{\delta - ATE\}) + F_{X|\epsilon}(u), 1 \right\} \right] dF_\epsilon(\epsilon),
\end{aligned}
$$

where $F_{X|\epsilon}(u)$, $F_{Y|\epsilon}(u)$ are the conditional cdfs of X, Y on ϵ. For both examples, the bounds are point identified under the same conditions as Heckman (1990).

3.3 Inference

The expressions in (9), (21), and (20) share one common feature, i.e., they depend on the marginal distributions only. As a result, they can be consistently estimated without requiring any dependence information between X and Y, provided that univariate samples from F, G are available. For example, consider a linear portfolio Z of market risk X and credit risk Y. To estimate the VaR of Z requires a bivariate sample from the joint distribution of X, Y which may not always be available. In contrast, (20) implies that the smallest and worst VaR can be estimated with only univariate samples from F, G. Inference on ρ, $F_\Delta(\delta)$, and $F_Z^{-1}(\alpha)$ belongs to the recent, but fast-growing area in econometrics: inference for partially identified parameters pioneered by Imbens and Manski (2004), see also Stoye (2009), Andrew and Soares (2010), and Chernozhukov et al. (2007). We refer interested readers to Fan, Guerre, and Zhu (2013) for a detailed discussion and more references on recent developments on inference for partially identified parameters.

4   Conclusion

We conclude this survey by mentioning two active research areas on copulas in the current literature: the generalized Fréchet problem with additional constraints and the construction of higher dimensional copulas and the Fréchet problem in higher dimensions.

Section 3 reviewed the general Fréchet problem and its applications in bivariate option pricing, VaR evaluations, and distributional treatment effects. In each application, the available information is just enough to identify the marginal distributions and solutions to the general Fréchet problem provide the identified sets for parameters of interest. Oftentimes, additional information might be available that help restrict the class of copulas to which the true copula of X, Y belongs. For example, X and Y may be known to be non-negatively dependent; the value of a dependence measure such as Kendall's τ is known; or the values of the true copula at some specific points in $[0,1]^2$ may be known. Nelsen and Ubeda-Flores (2004) and Nelsen et al. (2001, 2004) establish improved Fréchet-Hoeffding bounds when such partial dependence information is available. Using the improved Fréchet-Hoeffding bounds, Tankov (2011) shows that the bounds in Cambanis et al. (1976) for super-modular functions can be tightened. Similar results for the distribution function of a sum of two random variables can be obtained using Williamson and Downs (1990) and Embrechts et al. (2003), see Fan and Park (2010) and Fan, Guerre, and Zhu (2013) for details.

Copulas have found most success in bivariate modeling, as there are numerous parametric families of bivariate copulas that the researcher can choose from, see Joe (1997) and Nelsen (1999). Compared with bivariate parametric copulas, the number of higher dimensional parametric copulas is rather limited. So far, most applications in higher dimensions have focused on Gaussian copulas and Student's t copulas; see Christoffersen et al. (2012) for a high dimension application based on a skew t copula. It is well known that construction of higher dimensional copulas is very difficult. Several methods have been proposed recently, including methods for constructing higher dimensional Archimedean copulas, see Hofert and Scherer (2011) and Hering et al. (2010), and the "pair-copula construction" approach known as vines, see Kurowicka and Cooke (2006), Aas et al. (2009) and Acar et al. (2012). Oh and Patton (2012) propose a new class of "factor copulas" and show that they have some desirable features in high dimension applications. The Fréchet problem in higher dimensions has also drawn attention recently; the reader is referred to Embrechts (2009) for a brief account and references. Finally, Kallsen and Tankov (2006) present a version of Sklar's theorem that allows to construct a general multivariate Lévy process from arbitrary univariate Lévy processes and an arbitrary Lévy copula.

References

[1] Aas, K., Czado, C., Frigessi, A., Bakken, H., 2009. Pair-copula constructions of multiple dependence. *Insurance, Mathematics, and Economics* 44: 182-198.

[2] Acar E. F., Genest, C., Nešlehová, J., 2012. Beyond simplified pair-copula constructions. *Journal of Multivariate Analysis*, in press.

[3] Andrews, D.W.K., 2001. Testing when a parameter is on the boundary of the maintained hypothesis. *Econometrica* 69: 683-734.

[4] Andrews, D.W.K., Ploberger, W., 1994. Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica* 62: 1383-1414.

[5] Andrews, D.W.K., Soares, G., 2010. Inference for parameters defined by moment inequalities using generalized moment selection. *Econometrica* 78: 119-157.

[6] Beare, B.K., 2010. Copulas and temporal dependence. *Econometrica* 78: 395-410.

[7] Bonhomme, S., Robin, J.-M., 2009. Assessing the equalizing force of mobility using short panels: France, 1990-2000. *Review of Economic Studies* 76(1): 63-92.

[8] Brendstrup, B., Paarsch, H. J., 2007. Semiparametric identification and estimation in multi-object, English auctions. *Journal of Econometrics* 141: 84-108.

[9] Cambanis, S., Simons, G. and Stout, W., 1976. Inequalities for $\mathcal{E}k(X,Y)$ when the marginals are fixed. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 36: 285-294.

[10] Cameron, A. C., Tong, L., Trivedi, P. K., Zimmer, D. M., 2004. Modelling the differences in counted outcomes using bivariate copula models with applications to mismeasured counts. *Econometrics Journal* 7: 566-584.

[11] Capéraà, P., Fourgères, A.-L., Genest, C., 1997. A nonparametric estimation procedure for bivariate extreme value copulas. *Biometrika* 84(3): 567-577.

[12] Chan, N.-H., Chen, J., Chen, X., Fan, Y., Peng, L., 2009. Statistical inference for multivariate residual copula of GARCH Models. *Statistica Sinica* 19: 53-70.

[13] Chen, X., Fan, Y., 2005. Pseudo-Likelihood Ratio Tests for Model Selection in Semiparametric Multivariate Copula Models. *The Canadian Journal of Statistics* 33: 389-414.

[14] —— 2006a. Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. *Journal of Econometrics* 135: 125-154.

[15] —— 2006b. Semiparametric estimation of copula-based time series models. *Journal of Econometrics* 130: 307-335.

[16] —— 2007. Model selection test for bivariate failure-time data. *Econometric Theory* 23: 414-439.

[17] Chen, X., Fan, Y., Tsyrennikov, V., 2006. Efficient estimation of semiparametric multivariate copula models. *Journal of the American Statistical Association* 101: 1228-1240.

[18] Chen, X., Wu, W.B., Yi, Y., 2009. Efficient estimation of copula-based semiparametric markov models. *Annals of Statistics* 37: 4214-4253.

[19] Chernozhukov, V., Hong, H., Tamer, E., 2007. Parameter set inference in a class of econometric models. *Econometrica* 75: 1243-1284.

[20] Cherubini, U., Luciano, E., Vecchiato, W., 2004. *Copula methods in finance*. John Wiley & Sons, England.

[21] Christoffersen, P., Errunza, V., Jacobs, K., Langlois, H., 2012. Is the potential for international diversification disappearing? A dynamic copula approach. *Review of Financial Studies* 25: 3711-3751.

[22] Darsow, W. F., Nguyen, B., Olsen, E.T., 1992. Copulas and markov processes. *Illinois Journal of Mathematics* 36: 600-642.

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

[24] Diebold, F.X., Hahn, J., Tay, A.S., 1999. Multivariate density forecast evaluation and calibration in financial risk management: high frequency returns on foreign exchange. *Review of Economics and Statistics* 81: 661-673.

[25] Diks, C., Panchenko, V., van Dijk, D., 2010. Out-of-sample comparison of copula specifications in multivariate density forecasts. *Journal of Economic Dynamics and Control* 34(9): 1596-1609.

[26] Duffie, D., 2011. *Measuring corporate default risk*. Clarendon Lectures in Finance, Oxford University Press, Oxford.

[27] Embrechts, P., 2009. Copulas: a personal view. *Journal of Risk and Insurance* 76: 639-650.

[28] Embrechts, P. , McNeil, A., Straumann, D., 2002. Correlation and dependence properties in risk management: properties and pitfalls. M. Dempster, ed., *Risk management: Value at Risk and beyond*. Cambridge University Press: 176-223.

[29] Embrechts, P., Hoeing, A., Juri, A., 2003. Using copulae to bound the Value-at-Risk for functions of dependent risks. *Finance & Stochastics* 7(2): 145-167.

[30] Embrechts, P., Hoeing, A., Puccetti,G., 2005. Worst VaR scenarios. *Insurance: Mathematics and Economics* 37: 115-134.

[31] Fan, Y., 2010. Copulas in econometrics. *Encyclopedia of Quantitative Finance*, Cont, R. (Ed.). John Wiley & Sons Ltd. Chichester, UK. pp. 375-379.

[32] Fan, Y., Park, S., 2009. Partial identification of the distribution of treatment effects and its confidence sets. In Thomas B. Fomby and R. Carter Hill (ed.) *Nonparametric Econometric Methods* (Advances in Econometrics, Volume 25), Emerald Group Publishing Limited, pp.3-70.

[33] ——2010. Sharp bounds on the distribution of treatment effects and their statistical inference. *Econometric Theory* 26: 931-951.

[34] ——2011. Confidence intervals for the quantile of treatment effects in randomized experiments. *Journal of Econometrics*, forthcoming.

[35] Fan, Y., Guerre, E., Zhu, D., 2013. *Partial identification and confidence sets for functionals of the joint distribution of "potential outcomes"*. Working paper.

[36] Fan, Y., Pastorello, S., Renault, E., 2012. *Maximization by Parts in Extremum Estimation*. Working paper.

[37] Fan, Y., Sherman, R., and Shum, M., 2013. *Identifying treatment effects under data combination*. Working paper.

[38] Fan, Y., Wu, J., 2010. Partial identification of the distribution of treatment effects in switching regime models and its confidence sets. *Review of Economic Studies* 77: 1002-1041.

[39] Fermanian, J.-D., 2005. Goodness of fit tests for copulas. *Journal of Multivariate Analysis* 95: 119-152.

[40] Fermanian, J.-D., Scaillet, O., 2003. Nonparametric estimation of copulas for time series. *Journal of Risk* 5(4): 25-54.

[41] Fermanian, J.-D., Wegkamp, M., 2012. Time dependent copulas. *Journal of Multivariate Analysis* 110: 19-29.

[42] Fermanian, J.-D., Radulović, D., Wegkamp, M., 2004. Weak convergence of empirical copula processes. *Bernoulli* 10(5): 847-860.

[43] Frank, M. J., Nelsen, R. B., Schweizer, B., 1987. Best-possible bounds on the distribution of a sum—a problem of Kolmogorov. *Probability Theory and Related Fields* 74: 199-211.

[44] Garcia, R., Tsafack, G., 2011. Dependence structure and extreme comovements in international equity and bond markets. *Journal of Banking & Finance* 35(8): 1954-1970.

[45] Genest, C., Favre, A.-C., 2007. Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering* 12: 347-368.

[46] Genest, C., Neslehova, J., 2007. A primer on copulas for count data. *Astin Bulletin* 37: 475-515.

[47] Genest, C., Rivest, L.-P., 1993. Statistical inference procedures for bivariate Archimedean copulas. *Journal of the American Statistical Association* 88(423): 1034-1043.

[48] Genest, C., Ghoudi, K., Rivest,L., 1995. A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* 82 (3): 543-552.

[49] Giacomini, R., White, H., 2006. Tests of conditional predictive ability. *Econometrica* 74(6): 1545–1578.

[50] Giesecke, K., 2004. Correlated default with incomplete information. *Journal of Banking and Finance* 28: 1521-1545.

[51] Hansen, B.E., 1994. Autoregressive conditional density estimation. *International Economic Review* 35(3): 705-730.

[52] Hansen, P.R., Lunde, A., Nason, J.M., 2011. Model confidence set for forecating models. *Econometrica*, forthcoming.

[53] Heckman, J. J., 1990. Varieties of Selection Bias. *American Economic Review, Papers and Proceedings* 80: 313-318.

[54] Heckman, J.J., Vytlacil, E., 2005. Structural equations, treatment effects, and econometric policy evaluation. *Econometrica* 73: 669-738.

[55] Heckman, J., Smith, J., Clements, N., 1997. Making the most out of programme evaluations and social experiments: accounting for heterogeneity in programme impacts. *Review of Economic Studies* 64: 487-535.

[56] Hering, C., Hofert, M., Mai, J.-F., Scherer, M., 2010. Constructing hierarchical Archimedean copulas with Lévy subordinators. *Journal of Multivariate Analysis* 101(6): 1428-1433.

[57] Hoderlein, S., Stoye, J., 2009. *Revealed preferences in a heterogeneous population.* Working paper, Boston College.

[58] Hofert, M., Scherer, M., 2011. CDO pricing with nested Archimedean copulas. *Quantitative Finance* 11(5): 775–787.

[59] Hoeffding, W., 1940. Maβstabvariante Korrelationstheorie. *Schriften Math. Inst. Univ. Berlin* 5: 181-233.

[60] Hong, Y., Tu, J., Zhou, G., 2007. Asymmetries in stock returns: Statistical tests and economic evaluation. *Review of Financial Studies* 20: 1547–1581.

[61] Hubbard, T., Li, T., Paarsch, H. J., 2012. Semiparametric estimation in models of first-price, sealed-bid auctions with affiliation. *Journal of Econometrics* 168, 4-16.

[62] Hull, J., White, A., 1998. Value at Risk when daily changes in market variables are not normally distributed. *Journal of Derivatives* 5, 9-19.

[63] Ibragimov, R., 2009. Copula-based characterizations for higher-order Markov processes. *Econometric Theory* 25: 819-846.

[64] Imbens, G. W., Manski, C.F., 2004. Confidence intervals for partially identified parameters. *Econometrica* 72: 1845–1857.

[65] Joe, H., 1997. *Multivariate models and dependence concepts.* Chapman & Hall/CRC.

[66] ——2005. Asymptotic efficiency of the two-stage estimation method for copula-based models. *Journal of Multivariate Analysis* 94: 401-419.

[67] Joe, H., Xu, J.J., 1996. *The estimation method of inference functions for margins for multivariate models.* Working paper, Department of Statistics, University of British Columbia.

[68] Kaas, R., Laeven, R.J.A., Nelsen, R.B., 2009. Worst VaR scenarios with given marginals and measures of association. Insurance: Mathematics and Economics 44: 146-158.

[69] Kallsen, J., Tankov, P., 2006. Characterization of dependence of multidimensional Lévy processes using Lévy copulas. *Journal of Multivariate Analysis* 97: 1551-1572.

[70] Kurowicka, D., Cooke, R., 2006. *Uncertain analysis with high dimensional dependence modelling.* Wiley Series in Probability and Statistics, John Wiley & Sons, Ltd.

[71] Li, D.X., 2000. On default correlation: a copula function approach. *Journal of Fixed Income* 9: 43-54.

[72] Makarov, G. D., 1981. Estimates for the distribution function of a sum of two random variables when the marginal distributions are fixed. *Theory of Probability and its Applications* 26: 803-806.

[73] Manski, C. F., 1988. Ordinal utility models of decision making under uncertainty. *Theory and Decision* 25(1): 79-104.

[74] ——1997. The mixing problem in programme evaluation. *Review of Economic Studies* 64: 537-553.

[75] ——2003. *Partial Identification of Probability Distributions.* Springer.

[76] Mari, D. D., Kotz, S., 2001. *Correlation and dependence.* Imperial College Press, London.

[77] McNeil, A.J., Frey, R., Embrechts, P., 2005. *Quantitative risk management: concepts, techniques and tools.* Princeton University Press, New Jersey.

[78] Mikosch, T., 2006. Copulas: tales and facts, with discussion and rejoinder. *Extremes* 9: 3-62.

[79] Nelsen, R.B., 2006. *An introduction to copulas.* Second Edition, Springer, U.S.A.

[80] Nelsen, R. B., Quesada-Molina, J. J., Rodriguez-Lallena, J. A., Ubeda-Flores, M., 2001. Bounds on bivariate distribution functions with given margins and measures of association. *Communications in Statistics-Theory and Methods* 30: 1155-1162.

[81] —— 2004. Best-possible bounds on sets of bivariate distribution functions. *Journal of Multivariate Analysis* 90: 348-358.

[82] Nelsen, R. B., Ubeda-Flores, M., 2004. A comparison of bounds on sets of joint distribution functions derived from various measures of association. *Communications in Statistics-Theory and Methods* 33: 2299-2305.

[83] Newey, W.K., West, K.D., 1987. A simple, positive semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55: 703-708.

[84] Oh, D.-H., Patton, A.J., 2012. *Modelling dependence in high dimensions with factor copulas.* Working paper, Duke University.

[85] ——2013. Simulated method of moments estimation for copula-based multivariate models. *Journal of the American Statistical Association* 108: 689-700.

[86] Patton, A.J., 2004. On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. *Journal of Financial Econometrics* 2(1): 130-168.

[87] ——2006a. Modelling asymmetric exchange rate dependence. *International Economic Review*. 47(2): 527-556.

[88] ——2006b. Estimation of multivariate models for time series of possibly different lengths. *Journal of Applied Econometrics* 21(2): 147-173.

[89] ——2013. Copula methods for forecasting multivariate time series. May 2012, in G. Elliott and A. Timmermann (eds.) *Handbook of Economic Forecasting* Volume 2, Springer Verlag, forthcoming.

[90] Rachev, S.T., Ruschendorf, L., 1994. Solution of some transportation problems with relaxed or additional constraints. *SIAM Journal of Control and Optimization* 32: 673-689.

[91] Rapuch, G., Roncalli, T., 2001. *Some remarks on two-asset options pricing and stochastic dependence of asset prices.* Groupe de Recherche Operationnelle, Credit Lyonnais, France.

[92] Rémillard, B., 2010. *Goodness-of-fit tests for copulas of multivariate time series.* Working paper, HEC Montreal.

[93] Ridder, G., Moffitt, R., 2007. Econometrics of data combination. In *The Handbook of Econometrics* 6B, chapter 75.

[94] Rivers, D., Vuong, Q., 2002. Model selection tests for nonlinear dynamic models. *The Econometrics Journal* 5(1): 1-39.

[95] Romano, J.P., Wolf, M., 2005. Stepwise multiple testing via formalized data snooping. *Econometrica* 73: 1237-1282.

[96] Rosenberg, J.V., 2003. Nonparametric pricing of multivariate contingent claims. *Journal of Derivatives* 10: 9-26.

[97] Rosenberg, J.V., Schuermann, T., 2006. A general approach to integrated risk management with skewed, fat-tailed risks. *Journal of Financial Economics* 79: 569-614.

[98] Rüschendorf, L., 1982. Random variables with maximum sums. *Advances in Applied Probability* 14: 623-632.

[99] Rothe, C., 2012. Partial distributional policy effects. *Econometrica* 80: 2269-2301.

[100] Sancetta, A., Satchell, S., 2004. The Bernstein copula and its applications to modeling and approximations of multivariate distributions. *Econometric Theory* 20: 535–562.

[101] Schweizer, B., Sklar, A., 1983. *Probabilistic metric spaces.* North-Holland, New York.

28

[102] Sklar, A., 1959. Fonctions de r'epartition 'a n dimensionset leurs marges. *Publ. Inst. Statis. Univ. Paris* 8: 229-231.

[103] Smith, M., Min, A., Almeida, C., Czado, C., 2010, Modeling longitudinal data using a pair-copula decomposition of serial dependence. *Journal of the American Statistical Association* 105 (492): 1467-1479.

[104] Song, P., Fan, Y., Kalbfleisch J., 2005. Maximization by parts in likelihood inference. *Journal of the American Statistical Association* 100: 1145-1158.

[105] Stoye, J., 2009. More on confidence intervals for partially identified parameters. *Econometrica* 77: 1299–1315.

[106] Tamer, E., 2010. Partial identification in econometrics. *Annu. Rev. Econ.* 2: 167-195.

[107] Tankov, P., 2011. Improved Fréchet bounds and model-free pricing of multi-asset options. *Journal of Applied Probability* 48: 389-403.

[108] Tchen, A. H., 1980. Inequalities for distributions with given marginals. *Annals of Probability* 8: 814-827.

[109] van den Goorbergh, R.W.J., Genest, C., Werker, B.J.M., 2005. Multivariate option pricing using dynamic copula models. *Insurance: Mathematics and Economics* 37: 101-114.

[110] van Ophem, H., 2011. The frequency of visiting a doctor: is the decision to go independent of the frequency?. *Journal of Applied Econometrics* 26: 872-879.

[111] Vuong, Q.H., 1989. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* 57(2): 307-333.

[112] West, K. D., 2006. Forecast Evaluation. *Handbook of Economic Forecasting, Volume 1*, G. Elliott, C.W.J. Granger and A. Timmermann eds., Elsevier, Oxford.

[113] White, H., 1994. *Estimation, inference and specification analysis*. Econometric Society Monographs No. 22, Cambridge University Press, Cambridge, U.K.

[114] —— 2000. A reality check for data snooping. *Econometrica* 68, 1097-1126.

[115] Williamson, R. C., Downs, T., 1990. Probabilistic arithmetic I: numerical methods for calculating convolutions and dependency bounds. *International Journal of Approximate Reasoning* 4: 89-158.

[116] Zimmer, D., 2012. The role of copulas in the housing crisis. *Review of Economics and Statistics* 94: 607-620.

[117] 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 & Economic Statistics* 24: 63-76.
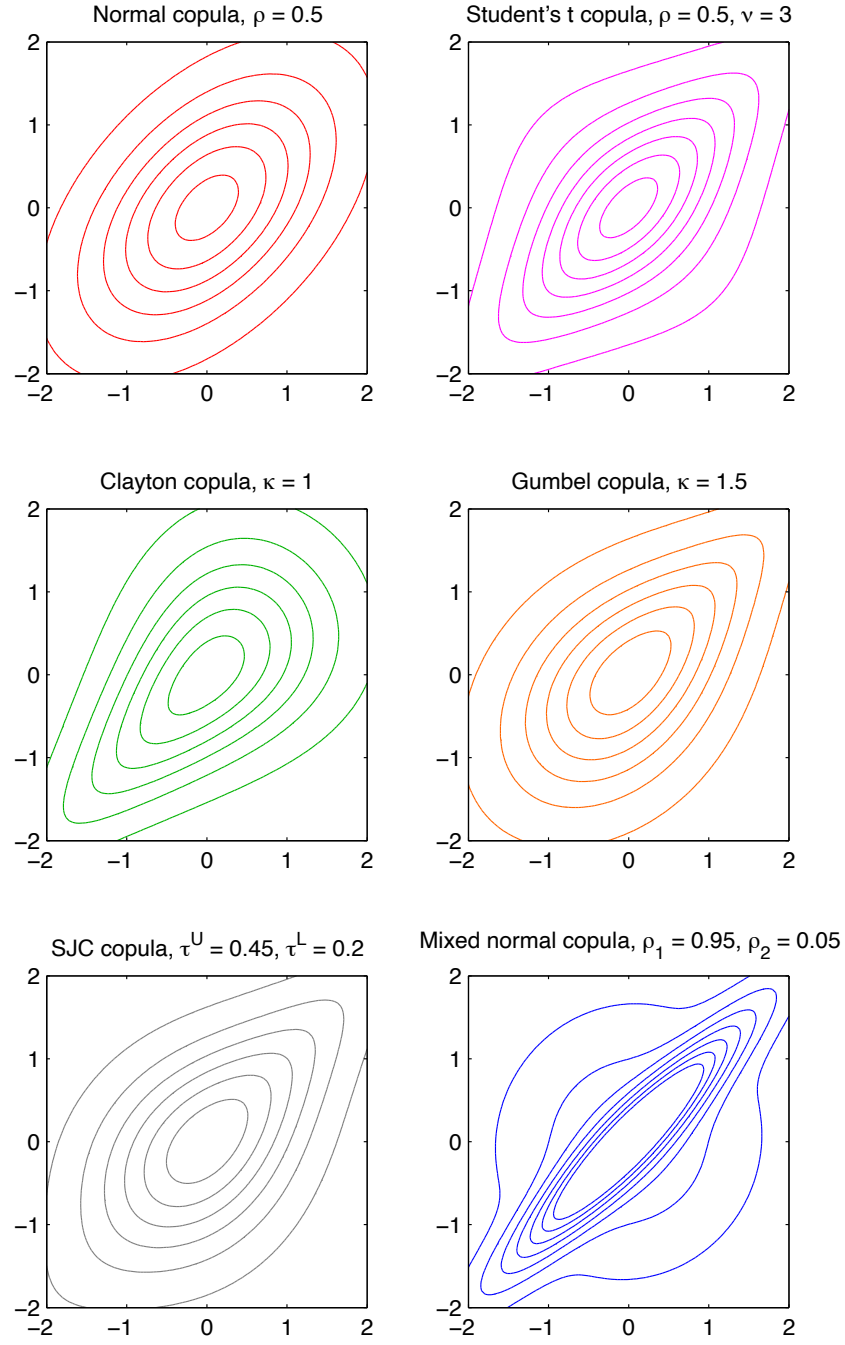
Figure 1: *This figure presents iso-probability contours from six bivariate densities, all with standard Normal margins and all implying linear correlation of one-half.*