Copulas in Econometrics

Yanqin Fan[1] and Andrew J. Patton[2]

[1]Department of Economics, University of Washington, Seattle, Washington 98195;
email: fany88@u.washington.edu

[2]Department of Economics, Duke University, Durham, North Carolina 27708-0097

Keywords

Sklar's theorem, multivariate models, Fréchet-Hoeffding inequality,
bounds

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 im-
portant roles: multivariate modeling and partial identification of
parameters that depend on the joint distribution of two random var-
iables with fixed or known marginal distributions. We focus on bivar-
iate copulas but provide references on recent advances in constructing
higher-dimensional copulas.

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. Whereas 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; e.g., see Nelsen 2006) 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 focuses on the bivariate case for simplicity, but Sklar's theorem applies to general d-dimensional distributions as well:

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

where G_i (g_i) for $i = 1, 2, \ldots, d$ are the marginal CDFs (PDFs) of the joint CDF H.

Sklar's theorem for multivariate modeling is useful because 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 Equation 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 discussed in Section 2, 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 Value at Risk (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 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 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 $C : [0, 1]^2 \rightarrow [0, 1]$ such that

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

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

The copula is sometimes called the dependence function (see Joe 1997, Nelsen 2006), 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 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. (Linear correlation is known to change when a nonlinear transformation, for example, 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 2006b), 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

---

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

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

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 & Wegkamp (2012) for an analysis of the case in which 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 (Equation 6). Indeed, recent work in empirical finance and insurance provides 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 modeling 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, Student's t copula, the Frank copula, the Gumbel copula, and the Clayton copula. We provide the Gaussian (normal) copula and Student's t copula below and refer interested readers to Joe (1997) and Nelsen (2006) 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

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

whose copula PDF is

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

> **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

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

The Student's t copula density is

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

where $\mathbf{x} = (x_1, \ldots, x_d)'$ 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, 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 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, e.g., Nelsen 2006), the skewed t copula of Demarta & McNeil (2005) and Christoffersen et al. (2012), and the factor copula model of Oh & 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 in which the random variables denoted as X and 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 and G, the Fréchet-Hoeffding inequality holds:

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

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 2006 for more discussion). Some applications of the Fréchet-Hoeffding inequality include those by Manski (1988), who uses it in decision theory; Manski (1997), who employs it on the mixing problem in program evaluation; Ridder & Moffitt (2007, section 3.1), who discuss its use in data-combination contexts; Hoderlein & Stoye (2014), who use it to bound violations of the revealed-preference axioms in a repeated cross-sectional context; and Tamer (2010), who employs it in the partial identification of the joint distribution of potential outcomes of a binary treatment in which one of the potential outcomes is observed.

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 and μ_Y, finite variances σ_X^2 and 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 & Kotz 2001) shows that the covariance of X and Y has an alternative expression:

$$\mathrm{cov}(X, Y) = \iint \big[H(x, y) - F(x)G(y)\big] \, dx \, dy. \tag{8}$$

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

$$\rho_L = \frac{\iint \big[M\big(F(x), G(y)\big) - F(x)G(y)\big] \, dx \, dy}{\sigma_X \sigma_Y},$$

$$\rho_U = \frac{\iint \big[W\big(F(x), G(y)\big) - F(x)G(y)\big] \, dx \, dy}{\sigma_X \sigma_Y}. \tag{9}$$

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). Interestingly, $\rho \in [-1, 1]$ 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 Equation 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 et al. (2014) apply Equation 9 to bounding counterfactual distributions and treatment effect parameters when outcomes and covariates are observed in separate data sets.

1.3. Other Reviews of Copula Theory

Nelsen (2006) and Joe (1997) provide two key textbooks on copula theory, with clear and detailed introductions to copulas and dependence modeling, emphasizing 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 discourse on the value of copulas in multivariate modeling. Genest & Favre (2007) present a description of semiparametric inference methods for independently and identically distributed (i.i.d.) data with a detailed empirical illustration. This article draws on Fan (2010), Fan et al. (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 et al. (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 key steps in the estimation of copula-based multivariate models. The majority of applications of copula-based multivariate models are in time series, so we make such models the focus of this section (for a discussion of copula-based univariate models, see, e.g., Darsow et al. 1992, Chen & Fan 2006b, Chen et al. 2009, Ibragimov 2009, Beare 2010). Models for i.i.d. data can essentially be treated as a special case of these. We consider three key aspects of the problem: model specification, estimation and inference, and goodness-of-fit (GoF) testing. Patton (2013) presents a more detailed discussion of these topics, on which the discussion below is based.

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:

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

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

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

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 & 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 Equation 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:

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

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 (for nonparametric estimation of copulas, see Genest & Rivest 1993 and Capéraà et al. 1997 for i.i.d. data and Fermanian & Scaillet 2003, Fermanian et al. 2004, Sancetta & Satchell 2004, and Ibragimov 2009 for time series data). The conditional copula can be assumed constant or allowed to vary through time. Joe (1997) and Nelsen (2006) present numerous parametric copula functions that can be used in applied work; these include the Gaussian (or normal) copula, Student's t copula, and Archimedean copulas, such as the Clayton, Gumbel, and Frank copulas.

Figure 1 illustrates the types of bivariate distributions that can be obtained with copula-based models. These bivariate distributions all have standard normal marginal distributions, and the copula parameters are calibrated to imply a linear correlation of one-half. 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.

Suppose the conditional copula model is parametric with unknown parameter γ_o. Let $\theta_o \equiv (\alpha_{xo}, \alpha_{yo}, \gamma_o)$ denote the parameter vector for the entire multivariate distribution model. Methods for estimating θ_o include the full maximum likelihood and multistage 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 multistage approach to estimating copula-based models and point to work that establishes asymptotic properties of such estimators.
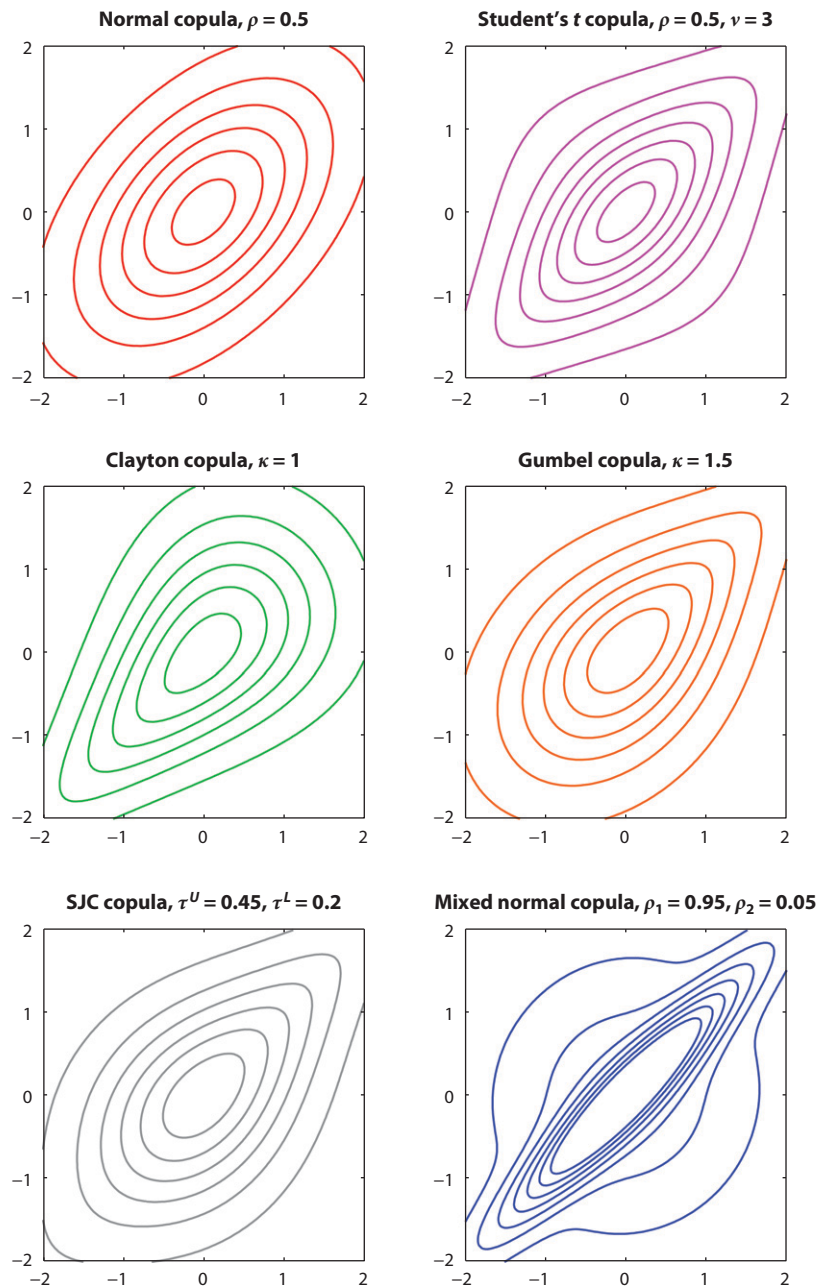
Figure 1

Isoprobability contours from six bivariate densities, all with standard normal margins and all implying a linear correlation of one-half. The parameters listed in the heading of each panel are those that describe the copula in that panel. Abbreviation: SJC, symmetrized Joe-Clayton copula.

2.2. Estimation and Inference

To simplify the exposition of the idea underlying the multistage 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 $\left\{(x_t, y_t)\right\}_{t=1}^{T}$. 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:

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

where $\theta = (\alpha_x, \alpha_y, \gamma)$ when the marginal distributions are parametric, for example, $F(x_t) \equiv F(x_t, \alpha_{xo})$ and $G(y_t) \equiv G(y_t, \alpha_{yo})$ for some finite-dimensional parameters α_{xo} and α_{yo}, 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, e.g., White 1994), standard results can be used to show that the maximum likelihood estimator 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 then estimate the copula parameters in a second stage. Specifically, let

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

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

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

This two-stage approach is sometimes called inference functions for margins in the copula literature (see Joe & Xu 1996, Joe 1997), although more generally this is known as multistage maximum likelihood estimation (see White 1994). Of course, two- or multistage estimation will yield parameters that are less efficient than one-stage maximum likelihood,[2] although simulation studies in Joe (2005) and Patton (2006a) indicate that the loss of efficiency is not great. As for one-stage maximum likelihood, under regularity conditions (see White 1994, Patton 2006a), the multistage maximum likelihood estimator is asymptotically normal, but the asymptotic covariance matrix now takes a nonstandard 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

---

[2]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).

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:

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

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

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

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 i.i.d. data and by Chen & Fan (2006b) for (univariate) time series data. Chen et al. (2006) propose a (semiparametric) sieve maximum likelihood 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 prefiltering 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 prefiltering step can be done via any existing methods for estimating the univariate models specified in Equation 10. Chan et al. (2009) and Chen & 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 & Patton (2013) consider simulated method of moments–type estimation of the parameters of semiparametric copula-based models. One surprising result in Chan et al. (2009) and Chen & Fan (2006a) in cases in which 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. Therefore, in practice, one can just proceed as though the parameters characterizing the dynamics were known (see Rémillard 2010 for some similar results).

Importantly, we note that the asymptotic theory for the estimation of semiparametric copula-based models, such as that in Chan et al. (2009), Chen & Fan (2006a), Rémillard (2010), and Oh & 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 is assumed to be i.i.d. Theory for the case in which the conditional copula is time varying is not available in the literature to date. Fully parametric models (i.e., those in which 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 GoF testing. Two widely used 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 to the empirical copula:

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

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, although similar, GoF test is based on Rosenblatt's transform, which is a form of multivariate probability integral transformation (see Diebold et al. 1999, 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 transformed 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 & Wolf 2005, and Hansen et al. 2011 for general models and Chen & 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,[3] 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 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 nonnested, fully parametric, copula-based models can be conducted using the test of Vuong (1989) for i.i.d. data and Rivers & 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 multistage) 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 heteroskedasticity and autocorrelation robust estimator (e.g., Newey & West 1987) of the variance of the difference in log-likelihoods will generally be needed. Chen & Fan (2006a) consider a similar case to Rivers & Vuong (2002) for semiparametric copula-based models, under the assumption that the conditional copula is constant. Chen & Fan (2006a) show that the likelihood ratio t test statistic is again normally distributed under the null hypothesis,

---

[3]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. Readers are referred to Andrews (2001) and Andrews & Ploberger (1994) for a discussion of these issues.

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, 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 & 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, readers are referred to Hull & White (1998), Embrechts et al. (2002), and Kaas et al. (2009) on VaR and to Rosenberg & Schuermann (2006) and McNeil et al. (2005) on general risk management issues. Li (2000), Giesecke (2004), Hofert & Scherer (2011), and Duffie (2011) provide applications to the pricing of credit derivatives. Applications of copulas in other derivative 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 by Patton (2004), Hong et al. (2007), Garcia & Tsafack (2011), and Christoffersen et al. (2012).

The use of copula-based analyses in other economic applications is broad and growing. Bonhomme & Robin (2009) utilize copulas to model the dynamics in a panel of earnings data, and Smith et al. (2010) similarly employ 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 & Trivedi (2006) employ them to model the health care demand of married couples, and van Ophem (2011) models the determinants of visits to a physician. Brendstrup & Paarsch (2007) work with a semiparametric copula-based model to identify and estimate the joint distributions of valuations in a multiobject English auction, and Hubbard et al. (2012) utilize 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 Equation 7 and the correlation bounds in Equation 9 provide sharp bounds on the joint distribution of X and Y and the correlation coefficient of X and 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 and 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 and Y and any bivariate CDF with marginals F and G, we denote the true joint CDF of X and Y as H_o.

In this section, we 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 Section 3.1. In Section 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 with a brief discussion on inference in Section 3.3. The presentation in this section draws heavily from Fan et al. (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

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

where E_H denotes the expectation taken with respect to H.

The Fréchet-Hoeffding inequality and the correlation bounds presented in Equations 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 supermodular functions. Section 3.1.1 reviews explicit characterizations of Θ_I for general supermodular functions extending the Fréchet-Hoeffding inequality and the correlation bounds. Section 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 and Y with fixed marginals. The corresponding functions μ are neither supermodular nor submodular.

3.1.1. Supermodular functions. We first provide the definition of a supermodular function.

Definition 1: A function $\mu(\cdot, \cdot)$ is called supermodular 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 is submodular if $-\mu(\cdot, \cdot)$ is supermodular.

If $\mu(\cdot, \cdot)$ is absolutely continuous, then it is supermodular if and only if $\partial^2\mu(x, y)/\partial x\partial y \geq 0$ almost everywhere. Cambanis et al. (1976) provide many examples of supermodular or submodular 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 supermodular (submodular) functions μ and parameters that depend on the marginal distributions of potential outcomes only. Many payoff functions of bivariate options also are either supermodular or submodular (see Section 3.2).

Suppose $\mu(\cdot, \cdot)$ is a supermodular and right continuous function. Sharp bounds on θ_o can be found in Cambanis et al. (1976), Tchen (1980), and Rachev & Ruschendorf (1994).[4] Let θ_L and θ_U denote the lower and upper bounds on θ_o, respectively. They are

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

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

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 of Cambanis et al. (1976), which generalizes the Fréchet-Hoeffding inequality and the correlation bounds presented in Equations 7 and 9.

---

[4]Results for submodular functions follow straightforwardly from the corresponding results for supermodular functions. To save space, we do not present results for submodular functions here.

Lemma 1: Suppose $\mu(x, y)$ is supermodular and right continuous. If $E_{H^{(-)}}\big[\mu(X, Y)\big]$ and $E_{H^{(+)}}\big[\mu(X, Y)\big]$ 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 $E\big[\mu(X, \overline{y})\big]$ and $E\big[\mu(\overline{x}, Y)\big]$ are finite, and at least one of $E_{H^{(-)}}\big[\mu(X, Y)\big]$ and $E_{H^{(+)}}\big[\mu(X, Y)\big]$ is finite.

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

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

Under the conditions stated in the lemma, one can express the right-hand side expression for θ_o as a nondecreasing function of the copula C_o and the fixed marginal CDFs F and G. Because $M(u, v) \leq C_o(u, v) \leq W(u, v)$ 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 we obtain $\theta_o = F_Z(z)$. The sharp bounds on $F_Z(z)$ can be found in Makarov (1981), Rüschendorf (1982), and Frank et al. (1987) (see Lemma 2 below). Frank et al. (1987) demonstrate that copulas provide useful tools for 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 readers to Schweizer & Sklar (1983), Williamson & Downs (1990), and Fan et al. (2013) for sharp bounds for the distribution functions of other arithmetic operations on X and Y.

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

Lemma 2: Let

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

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 supermodular functions in Lemma 1, which are reached at the Fréchet-Hoeffding lower and upper bounds for the distribution of X and 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 and 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 and Y denote the values of two individual assets or risks and θ_o denote the price of a European-style option on X and Y with the discounted payoff $\mu(X, Y)$. Following Rapuch & Roncalli (2001) and Tankov (2011), we assume that the econometrician observes random samples of prices on single-asset options on X and 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 $\theta_o = E_o[\mu(X, Y)]$ is the price of such an option. The sample information allows the identification of the marginal distributions of X and Y under Q. For example, if X is the price of an asset at time T and call options on this asset with prices $P_X(K) \equiv E_Q[\exp(-rT)(X - K)_+]$ are available, where r is the interest rate and K is the strike price, then the CDF of X is given by

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

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

Applying Lemma 1 to the payoff functions of bivariate options with supermodular payoff functions yields 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

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

The above bounds were first obtained by Rapuch & Roncalli (2001). Tankov (2011) establishes improved bounds when additional information on the dependence between X and Y is available. Fan et al. (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 and Y denote the values of two assets or two individual risks, $Z = X + Y$ denote a linear portfolio of X and 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 $F_Z^{-1}(\alpha)$. Although $F_Z^{-1}(\alpha)$ cannot be written as $E_o[\mu(X, Y)]$ 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 and 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 and Y is available. However, when only univariate samples on X and Y are available, we cannot point identify the VaR of Z. To resolve this issue, researchers have adopted the independence assumption on X and Y. This assumption is often violated (see McNeil et al. 2005, and references therein, for a detailed discussion). Without the independence or any other specific assumption on the dependence of X and Y, one can find sharp bounds on the VaR of Z by inverting the sharp bounds on $F_Z(z)$ in Lemma 2: For $\alpha \in (0, 1)$, it holds that

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

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

When the marginal distributions of X and Y are identified, the bounds reviewed in Section 3.1 can be used to bound distributional treatment effect parameters (see Manski 2003, and references therein, for scenarios under which the marginal distributions of X and Y are partially identified). 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 and Y and thus point identify average treatment effects. But they do not point identify the copula of X and Y because they contain no information on the dependence of X and Y besides that in the marginal distributions. Because 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 and Y, they are not point identified from randomized experiments. Fan & Park (2009, 2010, 2012) provide a systematic study of partial identification and inference for these distributional treatment effect parameters using Lemma 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 2 (see also Williamson & Downs 1990): $F_\Delta^L(\delta) \le F_\Delta(\delta) \le F_\Delta^U(\delta)$, where

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

(21)

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

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

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 Equation 21 to $F_\Delta(0)$ leads to the identified set for $P(X > Y | D = 1)$.

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

$$Q_\Delta^L(\alpha) = \sup_{u \in (0,\alpha)} \big[F^{-1}(u) - G^{-1}(u + 1 - \alpha)\big],$$

$$Q_\Delta^U(\alpha) = \inf_{u \in (\alpha,1)} \big[F^{-1}(u) - G^{-1}(u - \alpha)\big].$$

Fan & Park (2012) 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 et al. 2013). When selection into treatment is based not only on observable but also on unobservable covariates, Heckman (1990) and Heckman &

Vytlacil (2005) provide a systematic study of the identification of average treatment effects in latent threshold-crossing models. Building on the results of Heckman (1990), Fan & 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 of Fan & Wu (2010) below.

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

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

where X_1, X_0, X_c, and W are observable covariates, U_1, U_0, and ϵ are unobservable covariates, and $g_1(x_1, x_c)$, $g_0(x_0, x_c)$, 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), and the selection-on-observables assumption holds; otherwise, the unobservable error ϵ affects both the individual's decision to select into treatment and his or her potential outcomes. Suppose the unobservable covariates are independent of the observable covariates. Heckman (1990) provides conditions under which the distributions of $(U_1, \epsilon)'$, $(U_0, \epsilon)'$, $g_1(x_1, x_c)$, $g_0(x_0, x_c)$, 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 & Wu 2010 for details).

Lemmas 1 and 2 allow us to establish sharp bounds for distributional treatment effect parameters that depend on the copula of X and Y or U_1 and U_0. For example, the covariance of U_1 and U_0 can be bounded as follows:

$$
\begin{aligned}
&\int \left[\iint M\Big(F_{1|\epsilon}(u), F_{0|\epsilon}(v)\Big) du dv \right] dF_\epsilon(\epsilon) \le \operatorname{cov}(U_1, U_0) \\
&\le \int \left[\iint W\Big(F_{1|\epsilon}(u), F_{0|\epsilon}(v)\Big) du dv \right] dF_\epsilon(\epsilon),
\end{aligned}
\tag{23}
$$

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 Equation 23 may be used to infer the sign of $\operatorname{cov}(U_1, U_0)$ and are typically narrower than those based on the marginal CDFs of U_1 and 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 $F_\Delta(\delta) = E[P(U_1 - U_0 \le \{\delta - ATE\}|\epsilon)]$. Applying Lemma 2 to $P(U_1 - U_0 \le \{\delta - ATE\}|\epsilon)$, we obtain the sharp bounds on the distribution function of treatment effects: $F_\Delta^L(\delta) \le F_\Delta(\delta) \le F_\Delta^U(\delta)$, where

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

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

3.3. Inference

The expressions in Equations 9, 20, and 21 share one common feature; that is, 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 and G are available. For example, consider a linear portfolio Z of market risk X and credit risk Y. Estimating the VaR of Z requires a bivariate sample from the joint distribution of X and Y, which may not always be available. In contrast, Equation 20 implies that the smallest and worst VaR can be estimated with only univariate samples from F and G. Inference on ρ, $F_\Delta(\delta)$, and $F_Z^{-1}(\alpha)$ belongs to a recent, but fast-growing area in econometrics: inference for partially identified parameters pioneered by Imbens & Manski (2004) (see also Chernozhukov et al. 2007, Stoye 2009, Andrews & Soares 2010). We refer interested readers to Fan et al. (2013) for a detailed discussion and more references on recent developments on inference for partially identified parameters.

4. CONCLUSION

We conclude this article 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 reviews 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 helps restrict the class of copulas to which the true copula of X and Y belongs. For example, X and Y may be known to be nonnegatively dependent; the value of a dependence measure such as Kendall's τ may be known; or the values of the true copula at some specific points in $[0, 1]^2$ may be known. Nelsen & 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 supermodular functions can be tightened. Similar results for the distribution function of a sum of two random variables can be obtained using the methods of Williamson & Downs (1990) and Embrechts et al. (2003) (see Fan & Park 2010 and Fan et al. 2013 for details).

Copulas have found most success in bivariate modeling, as there are numerous parametric families of bivariate copulas from which the researcher can choose (see Joe 1997, Nelsen 2006). 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-dimensional 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 those for constructing higher-dimensional Archimedean copulas (see Hering et al. 2010, Hofert & Scherer 2011), and the pair-copula construction approach known as vines (see Kurowicka & Cooke 2006, Aas et al. 2009, Acar et al. 2012). Oh & Patton (2012) propose a new class of "factor copulas" and show that they have some desirable features in high-dimensional applications. The Fréchet problem in higher dimensions has also drawn attention recently; readers are referred to Embrechts (2009) for a brief account and references. Finally, Kallsen & Tankov (2006) present a version of Sklar's theorem that allows for the construction of a general multivariate Lévy process from arbitrary univariate Lévy processes and an arbitrary Lévy copula.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

LITERATURE CITED

Aas K, Czado C, Frigessi A, Bakken H. 2009. Pair-copula constructions of multiple dependence. *Insur. Math. Econ.* 44:182–98

Acar EF, Genest C, Nešlehová J. 2012. Beyond simplified pair-copula constructions. *J. Multivar. Anal.* 110:74–90

Andrews DWK. 2001. Testing when a parameter is on the boundary of the maintained hypothesis. *Econometrica* 69:683–734

Andrews DWK, Ploberger W. 1994. Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica* 62:1383–414

Andrews DWK, Soares G. 2010. Inference for parameters defined by moment inequalities using generalized moment selection. *Econometrica* 78:119–57

Beare BK. 2010. Copulas and temporal dependence. *Econometrica* 78:395–410

Bonhomme S, Robin J-M. 2009. Assessing the equalizing force of mobility using short panels: France, 1990–2000. *Rev. Econ. Stud.* 76:63–92

Brendstrup B, Paarsch HJ. 2007. Semiparametric identification and estimation in multi-object, English auctions. *J. Econom.* 141:84–108

Cambanis S, Simons G, Stout W. 1976. Inequalities for $\mathcal{E}k(X, Y)$ when the marginals are fixed. *Z. Wahrscheinlichkeitstheorie Verwandte Geb.* 36:285–94

Cameron AC, Tong L, Trivedi PK, Zimmer DM. 2004. Modeling the differences in counted outcomes using bivariate copula models with applications to mismeasured counts. *Econom. J.* 7:566–84

Capéraà P, Fourgères A-L, Genest C. 1997. A nonparametric estimation procedure for bivariate extreme value copulas. *Biometrika* 84:567–77

Chan N-H, Chen J, Chen X, Fan Y, Peng L. 2009. Statistical inference for multivariate residual copula of GARCH models. *Stat. Sinica* 19:53–70

Chen X, Fan Y. 2005. Pseudo-likelihood ratio tests for model selection in semiparametric multivariate copula models. *Can. J. Stat.* 33:389–414

Chen X, Fan Y. 2006a. Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. *J. Econom.* 135:125–54

Chen X, Fan Y. 2006b. Semiparametric estimation of copula-based time series models. *J. Econom.* 130:307–35

Chen X, Fan Y. 2007. Model selection test for bivariate failure-time data. *Econom. Theory* 23:414–39

Chen X, Fan Y, Tsyrennikov V. 2006. Efficient estimation of semiparametric multivariate copula models. *J. Am. Stat. Assoc.* 101:1228–40

Chen X, Wu WB, Yi Y. 2009. Efficient estimation of copula-based semiparametric Markov models. *Ann. Stat.* 37:4214–53

Chernozhukov V, Hong H, Tamer E. 2007. Parameter set inference in a class of econometric models. *Econometrica* 75:1243–84

Cherubini U, Luciano E, Vecchiato W. 2004. *Copula Methods in Finance*. New York: Wiley & Sons

Christoffersen P, Errunza V, Jacobs K, Langlois H. 2012. Is the potential for international diversification disappearing? A dynamic copula approach. *Rev. Financ. Stud.* 25:3711–51

Darsow WF, Nguyen B, Olsen ET. 1992. Copulas and Markov processes. *Ill. J. Math.* 36:600–42

Demarta S, McNeil AJ. 2005. The *t* copula and related copulas. *Int. Stat. Rev.* 73:111–29

Diebold FX, Hahn J, Tay AS. 1999. Multivariate density forecast evaluation and calibration in financial risk management: high frequency returns on foreign exchange. *Rev. Econ. Stat.* 81:661–73

Diks C, Panchenko V, van Dijk D. 2010. Out-of-sample comparison of copula specifications in multivariate density forecasts. *J. Econ. Dyn. Control* 34:1596–609

Duffie D. 2011. *Measuring Corporate Default Risk*. New York: Oxford Univ. Press

Embrechts P. 2009. Copulas: a personal view. *J. Risk Insur.* 76:639–50

Embrechts P, Hoeing A, Juri A. 2003. Using copulae to bound the Value-at-Risk for functions of dependent risks. *Finance Stoch.* 7:145–67

Embrechts P, Hoeing A, Puccetti G. 2005. Worst VaR scenarios. *Insur. Math. Econ.* 37:115–34

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

Fan Y. 2010. Copulas in econometrics. In *Encyclopedia of Quantitative Finance*, ed. R Cont, pp. 375–79. New York: Wiley & Sons

Fan Y, Guerre E, Zhu D. 2013. *Partial identification and confidence sets for functionals of the joint distribution of "potential outcomes."* Work. Pap., Univ. Wash., Seattle

Fan Y, Park S. 2009. Partial identification of the distribution of treatment effects and its confidence sets. In *Nonparametric Econometric Methods*, ed. TB Fomby, RC Hill, pp. 3–70. Bingley, UK: Emerald Group

Fan Y, Park S. 2010. Sharp bounds on the distribution of treatment effects and their statistical inference. *Econom. Theory* 26:931–51

Fan Y, Park S. 2012. Confidence intervals for the quantile of treatment effects in randomized experiments. *J. Econom.* 167:330–44

Fan Y, Pastorello S, Renault E. 2012. *Maximization by parts in extremum estimation*. Work. Pap., Univ. Wash., Seattle

Fan Y, Sherman R, Shum M. 2014. Identifying treatment effects under data combination. *Econometrica* 82:811–22

Fan Y, Wu J. 2010. Partial identification of the distribution of treatment effects in switching regime models and its confidence sets. *Rev. Econ. Stud.* 77:1002–41

Fermanian J-D, Radulović D, Wegkamp M. 2004. Weak convergence of empirical copula processes. *Bernoulli* 10:847–60

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

Fermanian J-D, Wegkamp M. 2012. Time dependent copulas. *J. Multivar. Anal.* 110:19–29

Frank MJ, Nelsen RB, Schweizer B. 1987. Best-possible bounds on the distribution of a sum: a problem of Kolmogorov. *Probab. Theory Relat. Fields* 74:199–211

Garcia R, Tsafack G. 2011. Dependence structure and extreme comovements in international equity and bond markets. *J. Bank. Finance* 35:1954–70

Genest C, Favre A-C. 2007. Everything you always wanted to know about copula modeling but were afraid to ask. *J. Hydrol. Eng.* 12:347–68

Genest C, Ghoudi K, Rivest L. 1995. A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* 82:543–52

Genest C, Nešlehová J. 2007. A primer on copulas for count data. *ASTIN Bull.* 37:475–515

Genest C, Rivest L-P. 1993. Statistical inference procedures for bivariate Archimedean copulas. *J. Am. Stat. Assoc.* 88:1034–43

Giacomini R, White H. 2006. Tests of conditional predictive ability. *Econometrica* 74:1545–78

Giesecke K. 2004. Correlated default with incomplete information. *J. Bank. Finance* 28:1521–45

Hansen BE. 1994. Autoregressive conditional density estimation. *Int. Econ. Rev.* 35:705–30

Hansen PR, Lunde A, Nason JM. 2011. Model confidence set for forecasting models. *Econometrica* 79:453–97

Heckman JJ. 1990. Varieties of selection bias. *Am. Econ. Rev.* 80:313–18

Heckman J, Smith J, Clements N. 1997. Making the most out of programme evaluations and social experiments: accounting for heterogeneity in programme impacts. *Rev. Econ. Stud.* 64:487–535

Heckman JJ, Vytlacil E. 2005. Structural equations, treatment effects, and econometric policy evaluation. *Econometrica* 73:669–738

Hering C, Hofert M, Mai J-F, Scherer M. 2010. Constructing hierarchical Archimedean copulas with Lévy subordinators. *J. Multivar. Anal.* 101:1428–33

Hoderlein S, Stoye J. 2014. Revealed preferences in a heterogeneous population. *Rev. Econ. Stat.* 96:197–213

Hoeffding W. 1940. Masstabvariante Korrelationstheorie. *Schr. Math. Inst. Univ. Berlin* 5:181–233

Hofert M, Scherer M. 2011. CDO pricing with nested Archimedean copulas. *Quant. Finance* 11:775–87

Hong Y, Tu J, Zhou G. 2007. Asymmetries in stock returns: statistical tests and economic evaluation. *Rev. Financ. Stud.* 20:1547–81

Hubbard T, Li T, Paarsch HJ. 2012. Semiparametric estimation in models of first-price, sealed-bid auctions with affiliation. *J. Econom.* 168:4–16

Hull J, White A. 1998. Value at Risk when daily changes in market variables are not normally distributed. *J. Deriv.* 5:9–19

Ibragimov R. 2009. Copula-based characterizations for higher-order Markov processes. *Econom. Theory* 25:819–46

Imbens GW, Manski CF. 2004. Confidence intervals for partially identified parameters. *Econometrica* 72:1845–57

Joe H. 1997. *Multivariate Models and Dependence Concepts*. Boca Raton, FL: Chapman & Hall/CRC

Joe H. 2005. Asymptotic efficiency of the two-stage estimation method for copula-based models. *J. Multivar. Anal.* 94:401–19

Joe H, Xu JJ. 1996. *The estimation method of inference functions for margins for multivariate models*. Work. Pap., Dep. Stat., Univ. Br. Columbia, Vancouver

Kaas R, Laeven RJA, Nelsen RB. 2009. Worst VaR scenarios with given marginals and measures of association. *Insur. Math. Econ.* 44:146–58

Kallsen J, Tankov P. 2006. Characterization of dependence of multidimensional Lévy processes using Lévy copulas. *J. Multivar. Anal.* 97:1551–72

Kurowicka D, Cooke R. 2006. *Uncertain Analysis with High Dimensional Dependence Modelling*. Chichester: Wiley & Sons

Li DX. 2000. On default correlation: a copula function approach. *J. Fixed Income* 9:43–54

Makarov GD. 1981. Estimates for the distribution function of a sum of two random variables when the marginal distributions are fixed. *Theory Probab. Appl.* 26:803–6

Manski CF. 1988. Ordinal utility models of decision making under uncertainty. *Theory Decis.* 25:79–104

Manski CF. 1997. The mixing problem in programme evaluation. *Rev. Econ. Stud.* 64:537–53

Manski CF. 2003. *Partial Identification of Probability Distributions*. New York: Springer

Mari DD, Kotz S. 2001. *Correlation and Dependence*. London: Imperial Coll. Press

McNeil AJ, Frey R, Embrechts P. 2005. *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton, NJ: Princeton Univ. Press

Mikosch T. 2006. Copulas: tales and facts, with discussion and rejoinder. *Extremes* 9:3–62

Nelsen RB. 2006. *An Introduction to Copulas*. New York: Springer. 2nd ed.

Nelsen RB, Quesada-Molina JJ, Rodriguez-Lallena JA, Úbeda-Flores M. 2001. Bounds on bivariate distribution functions with given margins and measures of association. *Commun. Stat. Theory Methods* 30:1155–62

Nelsen RB, Quesada-Molina JJ, Rodriguez-Lallena JA, Úbeda-Flores M. 2004. Best-possible bounds on sets of bivariate distribution functions. *J. Multivar. Anal.* 90:348–58

Nelsen RB, Úbeda-Flores M. 2004. A comparison of bounds on sets of joint distribution functions derived from various measures of association. *Commun. Statist. Theory Methods* 33:2299–305

Newey WK, West KD. 1987. A simple, positive semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55:703–8

Oh D-H, Patton AJ. 2012. *Modelling dependence in high dimensions with factor copulas*. Work. Pap., Duke Univ., Durham, NC

Oh D-H, Patton AJ. 2013. Simulated method of moments estimation for copula-based multivariate models. *J. Am. Stat. Assoc.* 108:689–700

Patton AJ. 2004. On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. *J. Financ. Econom.* 2:130–68

Patton AJ. 2006a. Estimation of multivariate models for time series of possibly different lengths. *J. Appl. Econ.* 21:147–73

Patton AJ. 2006b. Modelling asymmetric exchange rate dependence. *Int. Econ. Rev.* 47:527–56

Patton AJ. 2013. Copula methods for forecasting multivariate time series. In *Handbook of Economic Forecasting*, Vol. 2, ed. G Elliott, A Timmermann, pp. 899–960. Amsterdam: Elsevier

Rachev ST, Ruschendorf L. 1994. Solution of some transportation problems with relaxed or additional constraints. *SIAM J. Control Optim.* 32:673–89

Rapuch G, Roncalli T. 2001. *Some remarks on two-asset options pricing and stochastic dependence of asset prices*. Rep., Groupe Rech. Oper.,Credit Lyonnais, Paris

Rémillard B. 2010. *Goodness-of-fit tests for copulas of multivariate time series*. Work. Pap., HEC Montreal

Ridder G, Moffitt R. 2007. Econometrics of data combination. In *Handbook of Econometrics*, Vol. 6B, ed. JJ Heckman, EE Leamer, pp. 5469–547. Amsterdam: Elsevier

Rivers D, Vuong Q. 2002. Model selection tests for nonlinear dynamic models. *Econom. J.* 5:1–39

Romano JP, Wolf M. 2005. Stepwise multiple testing via formalized data snooping. *Econometrica* 73:1237–82

Rosenberg JV. 2003. Nonparametric pricing of multivariate contingent claims. *J. Deriv.* 10:9–26

Rosenberg JV, Schuermann T. 2006. A general approach to integrated risk management with skewed, fat-tailed risks. *J. Financ. Econ.* 79:569–614

Rothe C. 2012. Partial distributional policy effects. *Econometrica* 80:2269–301

Rüschendorf L. 1982. Random variables with maximum sums. *Adv. Appl. Probab.* 14:623–32

Sancetta A, Satchell S. 2004. The Bernstein copula and its applications to modeling and approximations of multivariate distributions. *Econom. Theory* 20:535–62

Schweizer B, Sklar A. 1983. *Probabilistic Metric Spaces*. Amsterdam: North-Holland

Sklar A. 1959. Fonctions de répartition à *n* dimensions et leurs marges. *Publ. Inst. Stat. Univ. Paris* 8:229–31

Smith M, Min A, Almeida C, Czado C. 2010. Modeling longitudinal data using a pair-copula decomposition of serial dependence. *J. Am. Stat. Assoc.* 105:1467–79

Song P, Fan Y, Kalbfleisch J. 2005. Maximization by parts in likelihood inference. *J. Am. Stat. Assoc.* 100:1145–58

Stoye J. 2009. More on confidence intervals for partially identified parameters. *Econometrica* 77:1299–315

Tamer E. 2010. Partial identification in econometrics. *Annu. Rev. Econ.* 2:167–95

Tankov P. 2011. Improved Fréchet bounds and model-free pricing of multi-asset options. *J. Appl. Probab.* 48:389–403

Tchen AH. 1980. Inequalities for distributions with given marginals. *Ann. Probab.* 8:814–27

van den Goorbergh RWJ, Genest C, Werker BJM. 2005. Multivariate option pricing using dynamic copula models. *Insur. Math. Econ.* 37:101–14

van Ophem H. 2011. The frequency of visiting a doctor: Is the decision to go independent of the frequency? *J. Appl. Econ.* 26:872–79

Vuong QH. 1989. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* 57:307–33

West KD. 2006. Forecast evaluation. In *Handbook of Economic Forecasting*, Vol. 1, ed. G Elliott, CWJ Granger, A Timmermann, pp. 100–32. Amsterdam: Elsevier

White H. 1994. *Estimation, Inference and Specification Analysis*. Cambridge, UK: Cambridge Univ. Press

White H. 2000. A reality check for data snooping. *Econometrica* 68:1097–126

Williamson RC, Downs T. 1990. Probabilistic arithmetic I: numerical methods for calculating convolutions and dependency bounds. *Int. J. Approx. Reason.* 4:89–158

Zimmer DM. 2012. The role of copulas in the housing crisis. *Rev. Econ. Stat.* 94:607–20

Zimmer DM, Trivedi PK. 2006. Using trivariate copulas to model sample selection and treatment effects: application to family health care demand. *J. Bus. Econ. Stat.* 24:63–76

Ⱥ

Contents

Symposium: The Institutional Underpinnings of Long-Run Income Differences

Indexes

Errata

An online log of corrections to *Annual Review of Economics* articles may be found at http://www.annualreviews.org/errata/economics

ANNUAL REVIEWS
It's about time. Your time. It's time well spent.

New From Annual Reviews:

Annual Review of Statistics and Its Application

Volume 1 • Online January 2014 • http://statistics.annualreviews.org

Editor: **Stephen E. Fienberg,** *Carnegie Mellon University*
Associate Editors: **Nancy Reid,** *University of Toronto*
Stephen M. Stigler, *University of Chicago*

The *Annual Review of Statistics and Its Application* aims to inform statisticians and quantitative methodologists, as well as all scientists and users of statistics about major methodological advances and the computational tools that allow for their implementation. It will include developments in the field of statistics, including theoretical statistical underpinnings of new methodology, as well as developments in specific application domains such as biostatistics and bioinformatics, economics, machine learning, psychology, sociology, and aspects of the physical sciences.

Complimentary online access to the first volume will be available until January 2015.

TABLE OF CONTENTS:

Access this and all other Annual Reviews journals via your institution at www.annualreviews.org.

ANNUAL REVIEWS | Connect With Our Experts

Tel: 800.523.8635 (US/CAN) | Tel: 650.493.4400 | Fax: 650.424.0910 | Email: service@annualreviews.org