....

Spline approximations to conditional Archimedean copula

Philippe Lambert[a,b]

We propose a flexible copula model to describe changes with a covariate in the dependence structure of (conditionally exchangeable) random variables. The starting point is a spline approximation to the generator of an Archimedean copula. Changes in the dependence structure with a covariate x are modelled by flexible regression of the spline coefficients on x. The performances and properties of the spline estimate of the reference generator and the abilities of these conditional models to approximate conditional copulas are studied through extensive simulations. Inference is made using Bayesian arguments with posterior distributions explored using importance sampling or adaptive MCMC algorithms. The modelling strategy is illustrated with the analysis of bivariate growth curve data. Copyright © 2014 John Wiley & Sons, Ltd.

....

1. Introduction

Sklar (1959) has proven that any distribution $H(y_1, \ldots, y_p)$ with marginal distributions $F_j(y_j)$ $(j = 1, \ldots, p)$ can be written as

$$H(y_1, \ldots, y_p) = C(F_1(y_1), \ldots, F_p(y_p)), \tag{1.1}$$

where C denotes a distribution function (named *copula*) on $(0, 1)^p$ with uniform margins. If the margins are continuous, then C is unique. Conversely, if C is a copula and $F_j(\cdot)$ are distribution functions, then (1.1) defines a multivariate distribution with marginal distributions F_j $(j = 1, \ldots, p)$.

In most practical applications where copula are used, the marginal distributions and their potential link with covariates $\mathbf{x}$ are investigated in a first step, yielding marginal fitted quantiles, $\hat{u}_{j|x} = \hat{F}_j(y_j|\mathbf{x})$ $(j = 1, \ldots, p)$. A parametric copula is then selected to describe the dependence structure of the fitted quantiles. That copula is usually assumed to be

....

[a]Institut des sciences humaines et sociales, Méthodes quantitatives en sciences sociales, Université de Liège, Boulevard du Rectorat 7 (B31), B-4000 Liège, Belgium.
[b]Institut de statistique, biostatistique et sciences actuarielles (ISBA), Université catholique de Louvain, Louvain-la-Neuve, Belgium.
*Email: p.lambert@ulg.ac.be

....

independent of the covariates as if the strength of association between the margins did not change with the unit characteristics. Then, (1.1) becomes

$$H(y_1, \ldots, y_p | \mathbf{x}) = C(F_1(y_1 | \mathbf{x}), \ldots, F_p(y_p | \mathbf{x})), \tag{1.2}$$

Unfortunately, the previous modelling assumption is not always realistic, as shown on Fig. 1 where the scatterplot of weight (in kg) and height (in cm) of young boys is given for different age classes. Indeed, one probably notices that the link between the marginal quantiles is getting looser as age increases. When the selected copula C_θ is parametric, one can let the copula parameter (and, hence, the strength of association) change with covariates, yielding

$$H(y_1, \ldots, y_p | \mathbf{x}) = C_{\theta(\mathbf{x})}(F_1(y_1 | \mathbf{x}), \ldots, F_p(y_p | \mathbf{x})), \tag{1.3}$$

An early example of that can be found in Lambert & Vandenhende (2002) where the effect of an antidepressant on blood pressures and heart rate were studied in a longitudinal setting. Besides the effects of covariates on the marginal distributions of these three responses, their strengths of association were also allowed to change with sex and the presence of drug in the plasma. The same idea was used in a financial context by Patton (2006) where the name *conditional copula* for $C_{\theta(\mathbf{x})}$ was coined.

Nonparametric versions are desirable to suggest or to validate parametric specifications, or even as a substitute for these models. This is a growing topic of research in the copula literature. Hafner & Reznikova (2010) and Acar & Craiu (2011) use local likelihood to estimate changes with a covariate in the dependence parameter of a parametric copula, while Craiu & Sabeti (2012) explore the use of cubic splines. Nonparametric estimates for the conditional copula were also proposed and studied by Gijbels et al. (2011), Veraverbeke et al. (2011) and Abegaz et al. (2012) where a local kernel-weighted pseudo likelihood is used to estimate $\theta(\mathbf{x})$ in a local polynomial approach. Extensions to multivariate or even functional covariates can be found in Gijbels et al. (2012).

Here, we explore the possibility to use B-splines to specify a copula 'non-parametrically' and to model its evolution with a covariate. The plan of the paper is as follows. In Section 2, we extend the work of Lambert (2007) and Vandenhende & Lambert (2005) to provide a smooth estimate for the generator of an Archimedean copula. Its properties are revealed by a simulation study in Section 2.3 with excellent results already available with small sample sizes. Flexible power transforms of the copula generator are introduced in Section 3.2 to model smooth changes of the copula with covariates. The abilities of the *flex-power* and of the *additive conditional spline Archimedean copula families* to model such changes are studied in Sections 3.1 and 3.2. We conclude the paper with an application in Section 4 followed by a discussion.

2. Spline estimation of an Archimedean copula

We shall restrict our attention to the estimation of bivariate Archimedean copulas. A (bivariate) copula is said *Archimedean* (Genest & MacKay, 1986) if it can be written as

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

where $\varphi(\cdot)$ is a decreasing and convex function (named the *generator*) taking values on $(0, 1)$ and such that $\varphi(0^+) = +\infty$ and $\varphi(1) = 0$. It is symmetric in u and v, and characterized by that univariate function. The strength of association between the two variates can be quantified using e.g. Spearman's rho or Kendall's tau. They can be computed from the generator with the latter given by

$$\tau = 1 + 4 \int_0^1 \lambda(s) ds, \tag{2.1}$$

....

where $\lambda(\cdot) = \varphi(\cdot)/\varphi'(\cdot)$.

Various parametric proposals $\varphi = \varphi_\theta$ have been made for the generator in the literature, see e.g. Table 1, where Kendall's tau is expressed as a function of θ. For a given choice of the copula family and independent pairs $\{(u_i, v_i) : i = 1, \ldots, n\}$ of data with uniform margins, one can estimate θ using e.g. the maximum likelihood principle. One can show that the likelihood is

$$L(\theta|\mathcal{D}) = \prod_{i=1}^{n} \frac{\partial^2}{\partial u \partial v} C_\theta(u_i, v_i) = -\prod_{i=1}^{n} \frac{\varphi_\theta''(C_i)\varphi_\theta'(u_i)\varphi_\theta'(v_i)}{\left(\varphi_\theta'(C_i)\right)^3}, \tag{2.2}$$

where $C_i = \varphi_\theta^{-1}\left(\varphi_\theta(u_i) + \varphi_\theta(v_i)\right)$ and $\mathcal{D}$ generically denotes the available data.

2.1. The spline approximation

A (spline based) nonparametric estimate of $\varphi(\cdot)$ was first proposed in Lambert (2007). Here, we present another formulation with superior properties and embedding the proposal made by Vandenhende & Lambert (2005) where the function $g(\cdot)$ in (2.3) was piecewise linear. It can be written as

$$\varphi_{\boldsymbol{\theta}}(u) = \exp\{-g_{\boldsymbol{\theta}}(S(u))\} \tag{2.3}$$

where $S(u) = -\log(-\log(u))$ is the quantile function of the extreme value distribution and

$$\frac{d}{ds}g_{\boldsymbol{\theta}}(s) = \sum_{k=1}^{K} b_k(s)(1 + \theta_k^2) = 1 + \sum_{k=1}^{K} b_k(s)\theta_k^2 \tag{2.4}$$

is a linear combination of K cubic B-splines associated to equidistant knots on $(S(\epsilon), S(1-\epsilon))$ for a small quantity ϵ ($= 10^{-6}$, say). One can show that such a formulation ensures that (2.3) provides a valid generator for any $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_K)'$ in $\mathbb{R}^K$. In the special case where $\theta_k = \theta$ for all k, the generator is $\varphi(\cdot) = (-\log(\cdot))^\zeta$, i.e. that of a Gumbel copula with dependence parameter $\zeta = 1 + \theta^2$. When, in addition, $\theta = 0$, one obtains the independence copula.

2.2. Inference

Given K, the inferential problem ends up to the selection of the parameters $\boldsymbol{\theta}$ defining the spline coefficients. We suggest to follow the proposal made by Eilers & Marx (1996) by taking a large number of equidistant knots and to counterbalance the introduced flexibility by penalizing changes in rth order differences of the spline coefficients. Then, the penalized log-likelihood (see the Appendix for computational details) is

$$l_{pen}(\boldsymbol{\theta}|\mathcal{D}) = \log L(\boldsymbol{\theta}|\mathcal{D}) - \frac{\kappa}{2}\boldsymbol{\theta}'P\boldsymbol{\theta},$$

where $P = D_r'D_r$ and D_r is the $(K - r) \times K$ matrix yielding rth order differences of the splines coefficients when applied on $\boldsymbol{\theta}$.

The penalty parameter κ can be selected using cross validation or an information criterion. A possible translation in Bayesian terms (Lang & Brezger, 2004) takes as prior for $\boldsymbol{\theta}$

$$p(\boldsymbol{\theta}|\kappa) \propto \kappa^{\rho(P)/2} \exp\left(-\frac{\kappa}{2}\boldsymbol{\theta}'P\boldsymbol{\theta}\right),$$

where $\rho(P)$ denotes the rank of P. A gamma distribution is a possible choice for the prior of the penalty parameter, $\kappa \sim \mathcal{G}(a, b)$ (with mean a/b and variance a/b^2). The marginal posterior distribution for θ (Jullion & Lambert, 2007) is obtained by integrating out κ from the joint posterior for $(\boldsymbol{\theta}, \kappa)$:

$$p(\boldsymbol{\theta}|\mathcal{D}) \propto \int_0^{+\infty} L(\boldsymbol{\theta}|\mathcal{D}) \, p(\boldsymbol{\theta}|\kappa) \, p(\kappa) \, d\kappa \propto \frac{L(\boldsymbol{\theta}|\mathcal{D})}{\left(b + \frac{1}{2}\boldsymbol{\theta}' P \boldsymbol{\theta}\right)^{a + \frac{\rho(P)}{2}}} \tag{2.5}$$

That expression reveals that it is equivalent to assuming the following independent Student priors for the differences $D_r\boldsymbol{\theta}$ of the splines coefficients:

$$D_r\boldsymbol{\theta} \sim t_{\nu=2a}\left(0, \frac{b}{a}I_{\rho(P)}\right), \tag{2.6}$$

where $t_\nu(\mu, \Sigma)$ is the multivariate Student-t distribution with ν degrees of freedom, mean μ and variance-covariance matrix $\frac{\nu}{\nu-2}\Sigma$ when these two moments exist. Therefore, taking $a = 1$ and a 'large' value for b (1, say) leaves a priori a lot of freedom to the spline coefficients with a regularization of their rth order differences towards 0. The pertinence of that recommendation will be confirmed by the simulation study in Section 2.3.

A possible point estimate for $\boldsymbol{\theta}$, and, hence, for the copula generator, can be obtained by maximizing (2.5), yielding the maximum posterior probability (MAP) estimate $\hat{\theta}$. A sample $\left\{\boldsymbol{\theta}^{(m)} : m = 1, \ldots, M\right\}$ from the joint posterior for $\boldsymbol{\theta}$ can also be obtained using an importance sampler with vectors generated using the multivariate Student distribution $t_\upsilon\left(\hat{\boldsymbol{\theta}}, (-H_{\hat{\boldsymbol{\theta}}})^{-1}\right)$ where $H_{\hat{\boldsymbol{\theta}}}$ denotes the Hessian matrix evaluated at $\hat{\boldsymbol{\theta}}$ and υ the degrees of freedom.

2.3. Simulation study

A first simulation study was set up to evaluate the merits of the spline model to estimate the generator of an Archimedean copula. $S = 500$ datasets $\{(u_{1i}, u_{2i}) : i = 1, \ldots, n\}$ were generated with $n = 100$, 250, 500 or 2000 pairs with uniform margins and an association structure characterized by a Clayton, a Frank or a Gumbel copula with a dependence parameter corresponding to a Kendall's tau equal to 0.15, 0.30 or 0.45. The copula generator was approximated using (2.3-2.4) with $K = 11$ knots, a 3rd order penalty, and a gamma prior for the penalty coefficient, $\kappa \sim \mathcal{G}(a, b)$ with $a = b = 1$.

The posterior mode (MAP) $\hat{\boldsymbol{\theta}}$ for the spline coefficients was first estimated by maximizing (2.5) w.r.t. $\boldsymbol{\theta}$. Then, a sample $\left\{\boldsymbol{\theta}^{(m)} : m = 1, \ldots, M\right\}$ from the joint posterior for $\boldsymbol{\theta}$ was obtained using an importance sampler, see Section 2.2 for details. The posterior mean $\tilde{\lambda}(u)$ and pointwise credible intervals for $\lambda(u) = \varphi(u)/\varphi'(u)$ were estimated using the induced sample $\{\lambda(u|\boldsymbol{\theta}^{(m)}) : m = 1, \ldots, M\}$ and the associated importance weights (Chen & Shao, 1999). The bias and the RMSE of $\tilde{\lambda}(u)$ for a grid of values for u in (0.05, .95) are reported in Table 2 when $\tau = 0.30$ and available upon request when $\tau = 0.15$ or 0.45. It suggests that the biases (if any) are very small, even with small sample sizes. Root mean squared errors (RMSE) decrease with sample size, with empirical results suggesting that they are proportional to $n^{-1/2}$. The root mean integrated squared errors (RMISE), $RMISE = \left(\int_0^1 (\lambda(u) - \tilde{\lambda}(u))^2 du\right)^{1/2}$, were also computed as a summary value of the quality of the generator estimator, see Table 3.

The (mean) coverages of the pointwise credible intervals (computed from a grid of 19 equidistant values for u between 0.05 and 0.95) for $\lambda(u)$ are reported in Table 4. They tend to be slightly larger than their nominal values with better performances obtained for larger underlying Kendall's tau.

The number K of B-splines in the basis should be large enough to ensure sufficient flexibility to the approximation of the copula generator. That flexibility is counter-balanced by the penalty part, resulting in the Student prior for rth

order differences of the spline coefficients, see (2.6). The simulation results suggest that $K = 11$ is indeed sufficient to have an estimator with low bias for the generator.

Besides the excellent statistical properties of the generator estimator (including a low bias and a good agreement between the nominal and the effective coverages of the computed credible regions), the estimated functions turn to be smooth whatever the sample size.

3. Flexible conditional Archimedean copula families

The flexible form presented in Section 2 for the generator can be generalized by letting the spline coefficients change smoothly with a continuous covariate X. Extending the preceding definitions, we propose to take $\varphi_\Theta(u|x)$ as conditional generator for the Archimedean copula when $X = x$, where

$$\varphi_\Theta(u|x) = \exp\left\{-g_\Theta(S(u)|x)\right\},$$

$$\frac{d}{ds}g_\Theta(s|x) = \sum_{k=1}^{K} b_k(s)\left(1 + \theta_k^2(x)\right) = \sum_{k=1}^{K} b_k(s)\left\{1 + \left(\sum_{\ell=1}^{K^*} b_\ell^*(x)\theta_{k\ell}\right)^2\right\},\tag{3.1}$$

for a B-spline basis $\{b_\ell^*(\cdot)\}_{\ell=1}^{K^*}$ on the domain $\mathcal{X}$ of the covariate values, a $K \times K^*$ matrix $\Theta = (\theta_{kl})$ of spline coefficients and $S(u) = -\log(-\log(u))$. It directly affects Kendall's tau and relates it to the covariate. Indeed, one can show that

$$\tau_\Theta(x) = 1 + 4\int_0^1 \frac{\varphi_\Theta(u|x)}{\varphi_\Theta'(u|x)}\,du = 1 - 4\int_0^1 \left\{g_\Theta'(S(u)|x)\,S'(u)\right\}^{-1}\,du.$$

This approximation to the conditional copula generator shares the qualities of the spline approximation to the generator proposed in Section 2. The price to pay is in the number of spline parameters: $K \times K^*$ instead of K in the unconditional case. It motivates the two extra approximations proposed below in Sections 3.1 and 3.2.

Smoothness can be forced using penalties on the r_1th order differences in the θ_{kl}'s for a given ℓ and on their r_2th order differences for a given k, yielding the following quadratic form for the penalty

$$\text{pen}(\Theta|\kappa_1, \kappa_2) = \text{vec}(\Theta)'\left(\kappa_1 I_{K^*} \otimes P_1 + \kappa_2 P_2 \otimes I_K\right)\text{vec}(\Theta)$$

where $P_1 = D_{r_1}'D_{r_1}$ and $P_2 = D_{r_2}'D_{r_2}$ denote the penalty matrices of sizes K and K^*, respectively, and I_r the identity matrix of size r. The first term in the penalty induces smoothness on $g_\Theta(s|x)$ in the $s-$dimension, while in the second case, smoothness is obtained in the $x-$direction.

3.1. The additive conditional spline Archimedean copula family

One can reduce the number of spline parameters to $(K + K^* - 1)$ free parameters by using the additive form (3.2) for $\theta_k(x)$ in (3.1):

$$\varphi_{\gamma,\beta}(u|x) = \exp\left\{-g_{\gamma,\beta}(S(u)|x)\right\}$$

$$\frac{d}{ds}g_{\gamma,\beta}(s|x) = \sum_{k=1}^{K} b_k(s)\left(1 + \theta_k^2(x)\right) \quad;\quad \theta_k(x) = \gamma_k + \beta(x) = \gamma_k + \sum_{\ell=1}^{K^*} b_\ell^*(x)\beta_\ell,\tag{3.2}$$

for a B-spline basis $\{b_\ell^*(\cdot)\}_{\ell=1}^{K^*}$ on the domain $\mathcal{X}$ of the covariate values, an identification constraint $\sum_\ell \beta_\ell = 0$ (say) and $S(u) = -\log(-\log(u))$. Like in the general case, one can show that the conditional Kendall's tau is given by

$$\tau_{\gamma,\beta}(x) = 1 + 4 \int_0^1 \frac{\varphi_{\gamma,\beta}(u|x)}{\varphi'_{\gamma,\beta}(u|x)} du = 1 - 4 \int_0^1 \left\{ g'_{\gamma,\beta}(S(u)|x) S'(u) \right\}^{-1} du. \tag{3.3}$$

It can be further extended to settings with multiple covariates. If $\boldsymbol{x} = (x_1, \ldots, x_p)'$ denote p continuous covariates, we suggest to generalize (3.2) to

$$\frac{d}{ds} g_{\gamma,\beta}(s|\boldsymbol{x}) = \sum_{k=1}^{K} b_k(s) \left(1 + \theta_k^2(\boldsymbol{x})\right) \quad ; \quad \theta_k(\boldsymbol{x}) = \gamma_k + \sum_{j=1}^{p} \beta_j(x_j) = \gamma_k + \sum_{j=1}^{p} \sum_{\ell=1}^{K^*} b_\ell^*(x_j)\beta_{j\ell}, \tag{3.4}$$

with identification constraints $\sum_\ell \beta_{j\ell} = 0$ for all j. That model contains $K + p(K^* - 1)$ free parameters for the conditional copula.

3.1.1. Inference We shall focus on the single covariate case, the extension to the additive setting being straightforward. The likelihood is given by

$$L(\boldsymbol{\gamma},\boldsymbol{\beta}|\mathcal{D}) = \prod_{i=1}^{n} \frac{\partial^2}{\partial u \partial v} C_{\boldsymbol{\gamma},\boldsymbol{\beta}}(u_i, v_i|x_i) = -\prod_{i=1}^{n} \frac{\varphi''_{\boldsymbol{\gamma},\boldsymbol{\beta}}(C_i|x_i) \, \varphi'_{\boldsymbol{\gamma},\boldsymbol{\beta}}(u_i|x_i) \, \varphi'_{\boldsymbol{\gamma},\boldsymbol{\beta}}(v_i|x_i)}{\left(\varphi'_{\boldsymbol{\gamma},\boldsymbol{\beta}}(C_i|x_i)\right)^3}, \tag{3.5}$$

see the Appendix for computational details. Smoothness is encouraged on $\varphi_{\gamma,\beta}(u|x)$ in the $u-$ and $x-$scales by penalizing changes in 2nd or 3rd order differences, $D_\gamma \boldsymbol{\gamma}$ and $D_\beta \boldsymbol{\beta}$, of the spline coefficients $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ (Eilers & Marx, 1996). Letting $P_\gamma = D'_\gamma D_\gamma$ and $P_\beta = D'_\beta D_\beta + \epsilon I_{K^*}$, the following penalties

$$\text{pen}(\boldsymbol{\gamma}|\kappa_\gamma) = -\frac{\kappa_\gamma}{2} \boldsymbol{\gamma}' P_\gamma \boldsymbol{\gamma} \quad ; \quad \text{pen}(\boldsymbol{\beta}|\kappa_\beta) = -\frac{\kappa_\beta}{2} \boldsymbol{\beta}' P_\beta \boldsymbol{\beta},$$

are added to the log-likelihood to define an estimation criterion for the spline parameters. A small ridge penalty is added in the definition of P_β to include the identifiability constraint on $\boldsymbol{\beta}$. In Bayesian terms, it translates into priors on the spline coefficients:

$$p(\boldsymbol{\gamma}|\kappa_\gamma) \propto \kappa_\gamma^{\rho(P_\gamma)/2} \exp\left(-\frac{\kappa_\gamma}{2} \boldsymbol{\gamma}' P_\gamma \boldsymbol{\gamma}\right) \quad ; \quad p(\boldsymbol{\beta}|\kappa_\beta) \propto \kappa_\beta^{K^*/2} \exp\left(-\frac{\kappa_\beta}{2} \boldsymbol{\beta}' P_\beta \boldsymbol{\beta}\right).$$

With gamma priors on the penalty coefficients, $\kappa_\gamma \sim \mathcal{G}(a_\gamma, b_\gamma), \kappa_\beta \sim \mathcal{G}(a_\beta, b_\beta)$, one can show that the marginal posterior for $(\boldsymbol{\gamma}, \boldsymbol{\beta})$ is

$$p(\boldsymbol{\gamma},\boldsymbol{\beta}|\mathcal{D}) \propto \frac{L(\boldsymbol{\gamma},\boldsymbol{\beta}|\mathcal{D})}{\left(b_\gamma + \frac{1}{2}\boldsymbol{\gamma}' P_\gamma \boldsymbol{\gamma}\right)^{a_\gamma + \frac{\rho(P_\gamma)}{2}} \left(b_\beta + \frac{1}{2}\boldsymbol{\beta}' P_\beta \boldsymbol{\beta}\right)^{a_\beta + \frac{\rho(P_\beta)}{2}}} \tag{3.6}$$

MAP estimates $(\hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{\beta}})$ for the splines coefficients can be obtained by maximizing the last expression. Again, an importance sampler based on a normal approximation could be set up to explore the joint posterior of the spline parameters. Given the increase in the number of parameters, we prefer to use a block-Metropolis algorithm where vectorial proposals are made sequentially for $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$. Our experience shows that the following algorithm is very efficient:

-- Obtain an estimation $(\hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{\beta}})$ of the posterior mode and an approximation $(\hat{\boldsymbol{\sigma}}_\gamma, \hat{\boldsymbol{\sigma}}_\beta)$ to their marginal standard errors from (the diagonal of minus the inverse of) the Hessian matrix at the mode. Let $\hat{\Sigma}_\gamma = \text{diag}(\hat{\boldsymbol{\sigma}}_\gamma^2), \hat{\Sigma}_\beta = \text{diag}(\hat{\boldsymbol{\sigma}}_\beta^2)$.

- At iteration m,

   1. Generate a proposal $\tilde{\boldsymbol{\gamma}}$ from the multivariate normal distribution
   $\mathcal{N}_K \left(\boldsymbol{\gamma}^{(m-1)}, \varsigma_\gamma \hat{\Sigma}_\gamma\right)$ and u from a uniform distribution $\mathcal{U}_{(0,1)}$. Let $\varrho = \min \left\{1, \frac{p(\tilde{\boldsymbol{\gamma}}, \boldsymbol{\beta}^{(m-1)} | \mathcal{D})}{p(\boldsymbol{\gamma}^{(m-1)}, \boldsymbol{\beta}^{(m-1)})}\right\}$. Set $\boldsymbol{\gamma}^{(m)} = \tilde{\boldsymbol{\gamma}}$
   if $u \leq \varrho$ and $\boldsymbol{\gamma}^{(m)} = \boldsymbol{\gamma}^{(m-1)}$ otherwise.
   2. Generate a proposal $\tilde{\boldsymbol{\beta}}$ from the multivariate normal distribution
   $\mathcal{N}_{K^*} \left(\boldsymbol{\beta}^{(m-1)}, \varsigma_\beta \hat{\Sigma}_\beta\right)$ and u from a uniform distribution $\mathcal{U}_{(0,1)}$. Let $\varrho = \min \left\{1, \frac{p(\boldsymbol{\gamma}^{(m)}, \tilde{\boldsymbol{\beta}} | \mathcal{D})}{p(\boldsymbol{\gamma}^{(m)}, \boldsymbol{\beta}^{(m-1)} | \mathcal{D})}\right\}$. Set $\boldsymbol{\beta}^{(m)} = \tilde{\boldsymbol{\beta}}$ if $u \leq \varrho$ and $\boldsymbol{\beta}^{(m)} = \boldsymbol{\beta}^{(m-1)}$ otherwise.

- The algorithm starts with $(\boldsymbol{\gamma}^{(0)}, \boldsymbol{\beta}^{(0)}) = (\hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{\beta}})$. The values of ς_γ and ς_β are tuned automatically during the burn-in to have acceptance probabilities around 0.20 (Haario et al., 2001). The matrices $\hat{\Sigma}_\gamma$ and $\hat{\Sigma}_\beta$ are updated half-way during the burnin using the empirical variance-covariance of the corresponding chains.

The generated chain $\left\{(\boldsymbol{\gamma}^{(m)}, \boldsymbol{\beta}^{(m)}) : m = 1, \ldots, M\right\}$ can be used to estimate a (simultaneous) credible region for any function of the conditional copula generator. For example, if one is interested in such a region for the conditional Kendall's tau, compute

$$\tau^{(m)}(x_j) = 1 + 4 \int_0^1 \lambda_{\boldsymbol{\gamma}^{(m)}, \boldsymbol{\beta}^{(m)}}(s|x_j) \, ds \tag{3.7}$$

on a fine grid of values $\{x_j\}_{j=1}^J$ on $(0, 1)$. The so-obtained M trajectories, $\left\{\left(x_j, \tau^{(m)}(x_j)\right) : j = 1, \ldots, J\right\}_{m=1}^M$, can be used to estimate pointwise or simultaneous (Held, 2004) credible regions for $\tau(\cdot)$.

3.1.2. Simulation study We have chosen to report the results of a study where the data are simulated from a conditional copula outside the additive conditional spline Archimedean copula family to assess the ability of that tool to model the dynamics of specific dependence structures. More specifically, for given covariate values $\{x_i : i = 1, \ldots, n\}$ uniformly distributed on $(0, 1)$, we have simulated a data pair (u_i, v_i) from a Clayton or from a Frank copula with a dependence parameter such that the corresponding Kendall's tau is

$$\tau(x_i) = .5 + .3 \sin\left(1.6\pi x_i^{1.5}\right),$$

yielding values for tau oscillating between .2 and .8, a particularly challenging set-up. The considered samples sizes are $n = 250, 500$ and 2000. $S = 500$ replicates were considered.

Given that the data were simulated outside the assumed model family and the demanding simulation set-up, we did not expect a perfect reconstruction of the underlying dependence structure, but hoped that reasonable estimations for the changing strength of dependence would be obtained by forcing that model. This is indeed the conclusion that can be drawn by inspecting Fig. 2 where the mean (over the S replicates) of the estimated conditional Kendall's tau (see (3.7)) corresponding to the fitted additive conditional spline Archimedean copula model can be compared to the function $\tau(x)$ used to simulate the data. Whatever the underlying copula or sample size, the global evolution of Kendall's tau with the covariate is correctly captured, with results improving with sample size. The bias becomes negligible for large n, except perhaps for the largest values of x when the data are generated using a Clayton copula with a small Kendall's tau.

3.2. The flex-power Archimedean copula family

Other approaches were investigated. One is based on (what we suggest to name) the *power Archimedean copula family*. It relies on the following result (see e.g. Nelsen, 1999; Hofert, 2011): if $\varphi(\cdot)$ is an Archimedean copula generator, then

1. *Interior power transform*: $\varphi_{\alpha,1}(t) = \varphi(t^\alpha)$ is also a generator if $\alpha \in (0,1]$;
2. *Exterior power transform*: $\varphi_{1,\beta}(t) = (\varphi(t))^\beta$ is also a generator if $\beta \geq 1$.

The effect of α and β on a reference generator $\varphi(t)$ (with Kendall's tau τ_0) is illustrated on Fig. 3 when starting from a Clayton($\theta = 1$) generator. Remembering the relationship between Kendall's tau and the area decribed by the lambda function and the horizontal axis, see (2.1), one can see that Kendall's tau increases with α (with a maximal value τ_0 when $\alpha = 1$, corresponding to the reference generator) and with β (with a minimal value τ_0 when $\beta = 1$).

We suggest to consider both transforms to model changes of the copula generator with covariates, i.e. to take $\varphi_{\alpha,\beta}(t|x) = \left[\varphi(t^{\alpha(x)})\right]^{\beta(x)}$ with flexible forms for the reference generator $\varphi(\cdot)$ (see (2.3)-(2.4)), the interior power α and the exterior power β such that

$$\alpha = \alpha(x) = \left[1 + \left(\sum_{k=1}^{K^*} b_k^*(x)\alpha_k\right)^2\right]^{-1} \;\; ; \;\; \beta = \beta(x) = 1 + \left(\sum_k^{K^*} b_k^*(x)\beta_k\right)^2 ,$$

where $\{b_k^*(\cdot) : k = 1, \ldots, K^*\}$ denotes a cubic B-spline basis on the covariate space (relocated and rescaled to take values in $(0,1)$). That model includes $(K + 2K^*)$ parameters. Like for the reference generator, penalties can be introduced to favor smooth changes with x for $\alpha(x)$ and $\beta(x)$, and, hence, for the conditional copula generator $\varphi_{\alpha,\beta}(t|x)$. If κ_α (resp. κ_β) denote the penalty coefficients for the spline parameters $\boldsymbol{\alpha}$ (resp. $\boldsymbol{\beta}$) in $\alpha(x)$ (resp. $\beta(x)$) and P_α (resp. P_β) the corresponding penalty matrix, then one can show, using gamma priors $\kappa_\alpha \sim \mathcal{G}(a_\alpha, b_\alpha)$, $\kappa_\beta \sim \mathcal{G}(a_\beta, b_\beta)$ and the same reasoning as in Section 2, that the marginal posterior for the spline coefficients is

$$p(\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta}|\mathcal{D}) \propto \int_0^{+\infty} \int_0^{+\infty} \int_0^{+\infty} L(\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta}|D) \, p(\boldsymbol{\theta}|\kappa) \, p(\boldsymbol{\alpha}|\kappa_\alpha) \, p(\boldsymbol{\beta}|\kappa_\beta) \, p(\kappa) \, p(\kappa_\alpha) \, p(\kappa_\beta) \, d\kappa \, d\kappa_\alpha \, d\kappa_\beta$$

$$\propto \frac{L(\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta}|\mathcal{D})}{\left(b + \frac{1}{2}\boldsymbol{\theta}'P\boldsymbol{\theta}\right)^{a + \frac{\rho(P)}{2}} \left(b_\alpha + \frac{1}{2}\boldsymbol{\alpha}'P_\alpha\boldsymbol{\alpha}\right)^{a_\alpha + \frac{\rho(P_\alpha)}{2}} \left(b_\beta + \frac{1}{2}\boldsymbol{\beta}'P_\beta\boldsymbol{\beta}\right)^{a_\beta + \frac{\rho(P_\beta)}{2}}} \tag{3.8}$$

MAP estimates $(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}})$ for the splines coefficients can be obtained by maximizing the last expression. An adaptive block-Metropolis algorithm can be set up by mimicking the approach in Section 3.1.1 to sample the joint posterior of the spline parameters.

3.2.1. Simulation study When the data are generated from the power Archimedean family where the interior and exterior powers are functions of a covariate x, one can show, using simulations (not reported to save space) that the flex-power family is able to estimate the underlying conditional copula in a precise and nearly unbiased way for large sample sizes. This is not surprising given the results in Section 2.3.

Instead, we report the results obtained when fitting the flex-power Archimedean family to the same datasets as in Section 3.1.2. Given that the data were simulated outside the flex-power family, we did not expect a perfect reconstruction of the underlying dependence structure, but like for the additive model, hoped that reasonable estimations for the changing strength of dependence would be obtained by forcing that model. This is indeed the conclusion that can be drawn by inspecting Fig. 4 where the mean (over the S replicates) of the estimated conditional Kendall's tau corresponding to the fitted flex-power Archimedean copula can be compared to the function $\tau(x)$ used to simulate the data. Whatever the underlying copula or sample size, the bias for a given x is small and decreases with n. A comparison of Figs. 2 and 4 suggests that, in the framework of the simulation study, the flex-power family is a bit more performant than the additive one, at the cost of K^* extra (spline) parameter (per covariate).

4. Application

The additive conditional spline (Model 1) and flex-power (Model 2) Archimedean copula models were also applied on the growth data mentioned in Section 1 (see also Fig. 1). It provides the height (Y_1) and weight (Y_2) of $n = 441$ young Dutch boys aged between 3 and 21. The margins were first modelled using the nonparametric additive location-scale model described in Lambert (2013),

$$Y_j = \mu_j(x) + \sigma_j(x)\epsilon_j,$$

where $\mu_j(x)$ and $\sigma_j(x)$ denotes the conditional median and inter-quartile range of Y_j given Age ($= X$). The smooth evolution of these quantities, as well as the pivotal distributions of the ϵ_j's were described using penalized B-splines. The resulting fitted conditional deciles for Y_1 and Y_2 are displayed on Fig. 5. It reveals the nonlinear relationship between the responses and Age as well as their increasing dispersion with the latter variate.

The unconditional and the two conditional copula models were adjusted on the fitted conditional quantiles,

$$\left\{ \left(x_i, \hat{u}_{1|x_i} = \hat{F}_1(y_{1i}|x_i), \hat{u}_{2|x_i} = \hat{F}_2(y_{2i}|x_i) \right) : i = 1, \dots, n \right\},$$

using the methodology described in Sections 2 (Model 0), 3.1 (Model 1) and 3.2 (Model 2), respectively, with $K = 11$, $K^* = 5$ and $b = b_\alpha = b_\gamma = b_\beta = 1$. Therefore, Model 0, Model 1 and Model 2 include $K = 11$, $K + K^* = 16$ and $K + 2K^* = 21$ parameters, respectively (with identifiability constraints). Adaptive block-Metropolis algorithms with chains of length $30,000$ (and a burnin of $1,000$) and initial values set at the MAP estimates were used to explore the joint posterior of the spline parameters in the two models. The deviance information criterions (DIC, Spiegelhalter et al. (2002)) are much smaller in Models 1 & 2 than in Model 0, suggesting a significant change of the strength of association between the two responses with the covariate, see Table 5. Smaller values were obtained for the DIC and for the effective number of parameters in Model 1 (compared to Model 2). The posterior mean of the conditional Kendall's tau, also estimated from the MCMC chains (see e.g. (3.7)), is plotted with a 95% credible region in Fig. 6 and 7. It confirms and quantifies the decreasing association between weight and height suspected from Fig. 1. It is very large at early childhood and decreases afterwards with an apparent acceleration at the start of puberty.

The fitted joint distribution can also be visualized for any value of the covariate, see Fig. 8, where its contours for three values of Age are superposed to the scatterplot of the whole dataset.

5. Discussion

A performant spline estimator for the generator of an Archimedean copula has been proposed. A large simulation study has shown that it has nearly negligible bias and credible regions with coverage probabilities closed to their nominal values even for moderate sample sizes.

It has been extended to change smoothly with a covariate using either an additive model for the spline coefficients or flexible (spline based) power transforms of the generator and of its argument. Again, simulation studies suggest that these models are quite flexible and useful to describe changes in conditional association structures.

It can be extended to handle multiple categorical or continuous covariates using additive forms such as in (3.4). One might also be interested by an extension to more than two responses using conditional copulas of the form

$$C(u_1, \dots, u_J | x) = \varphi^{-1} \left(\varphi(u_1 | x) + \dots + \varphi(u_J | x) | x \right),$$

but the validity of an exchangeable dependence structure should be questioned in the context of the modelling exercise. Instead, we suggest to use one of the proposed bivariate conditional copula models as building blocks in pair copula constructions such as in regular vines (see e.g. Aas et al., 2009).

A joint estimation of the parameters involved in the marginal models and in the conditional copula structure is also possible and straightforward to manage in a Bayesian framework. Unless at least one of the margins is discrete, it is not obvious that one would gain anything substantial by merging these modelling efforts in a single inferential procedure.

Besides additive and vine copula models, an extension to hierarchical settings involving repeated measurements (such as with clustered or longitudinal data) is certainly worth considering, but out of scope for this paper.

Acknowledgements

Appendix: Likelihood computation in the spline copula model

From a computational point of view, we found convenient to rewrite the likelihood as

$$L(\boldsymbol{\theta}) = -\prod_{i=1}^{n} (1 - \lambda'_{\boldsymbol{\theta}}(C_i)) \frac{\lambda_{\boldsymbol{\theta}}(C_i)}{\lambda_{\boldsymbol{\theta}}(u_i)\lambda_{\boldsymbol{\theta}}(v_i)} \frac{\varphi_{\boldsymbol{\theta}}(u_i)\,\varphi_{\boldsymbol{\theta}}(v_i)}{(\varphi_{\boldsymbol{\theta}}(u_i) + \varphi_{\boldsymbol{\theta}}(v_i))^2}$$

where

$$\lambda_{\boldsymbol{\theta}}(u) = \frac{\varphi_{\boldsymbol{\theta}}(u)}{\varphi'_{\boldsymbol{\theta}}(u)} = \frac{-1}{g'_{\boldsymbol{\theta}}(S(u))S'(u)}.$$

The value of C_i is given by

$$C_i = \varphi_{\boldsymbol{\theta}}^{-1}(\varphi_{\boldsymbol{\theta}}(u_i) + \varphi_{\boldsymbol{\theta}}(v_i)).$$

The function $\varphi_{\boldsymbol{\theta}}^{-1}(\cdot)$ can be computed numerically using an iterative method. The following algorithm turns to be very efficient. Assume that one looks for u such that $\varphi_{\boldsymbol{\theta}}^{-1}(x) = u$. This is equivalent to finding the root of $f(\cdot)$ with $f(u) = g_{\boldsymbol{\theta}}(S(u)) + \log(x)$. Starting from $u = u^{(0)}$, the guess $u^{(t+1)}$ at iteration $t+1$ (resulting from a Newton step) is such that $S(u^{(t+1)}) = S(u^{(t)}) + \delta^{(t)}$ with $\delta^{(t)} = -f(u^{(t)})/g'(S(u^{(t)}))$. Starting from $C_i^{(0)} = u_i v_i$ and letting $x_i = \varphi_{\boldsymbol{\theta}}(u_i) + \varphi_{\boldsymbol{\theta}}(v_i)$, that procedure can be vectorized to compute C_i $(i = 1, \dots, n)$ with a single loop in three or four iterations. Finally, note that the quantities $\lambda'_{\boldsymbol{\theta}}(C_i)$ in the likelihood can be approximated using finite differences of the lambda function.

References

Aas, K, Czado, C, Frigessi, A & Bakken, H (2009), 'Pair-copula constructions of multiple dependence,' *Insurance Mathematics and Economics*, **44**(2), pp. 182–198.

Abegaz, F, Gijbels, I & Veraverbeke, N (2012), 'Semiparametric estimation of conditional copulas,' *Journal of Multivariate Analysis*, **110**, pp. 43–73.

Acar, E & Craiu, R (2011), 'Dependence calibration in conditional copulas: A nonparametric approach,' *Biometrics*, **67**, pp. 445–453.

Chen, MH & Shao, QM (1999), 'Monte Carlo estimation of Bayesian credible and HPD intervals,' *Journal of the American Statistical Association*, **8**(1), pp. 69–92.

....

10
Stat **2014**, 00 1–11
Prepared using staauth.cls

Craiu, R & Sabeti, A (2012), 'In mixed company: Bayesian inference for conditional copulas models with discrete and continuous outcomes,' *Journal of Multivariate Analysis*, **110**, pp. 106–120.

Eilers, PHC & Marx, BD (1996), 'Flexible smoothing with B-splines and penalties,' *Statistical Science*, **11**, pp. 89–121.

Genest, C & MacKay, RJ (1986), 'The joy of copulas: Bivariate distributions with uniform marginals,' *The American Statistician*, **40**, pp. 280–283.

Gijbels, I, Omelka, M & Veraverbeke, N (2012), 'Multivariate and functional covariates and conditional copulas,' *Electronic Journal of Statistics*, **6**, pp. 1273–1306.

Gijbels, I, Veraverbeke, N & Omelka, M (2011), 'Conditional copulas, association measures and their applications,' *Computational Statistics and Data Analysis*, **55**, pp. 1919–1932.

Haario, H, Saksman, E & Tamminen, J (2001), 'An adaptive Metropolis algorithm,' *Bernoulli*, **7**, pp. 223–242.

Hafner, CM & Reznikova, O (2010), 'Efficient estimation of a semiparametric dynamic copula model,' *Computational Statistics and Data Analysis*, **54**(11), pp. 2609–2627.

Held, L (2004), 'Simultaneous posterior probability statements from Monte Carlo output,' *Journal of Computational and Graphical Statistics*, **13**, pp. 20–35.

Hofert, M (2011), 'Efficiently sampling nested Archimedean copulas,' *Computational Statistics and Data Analysis*, **55**, pp. 57–70.

Jullion, A & Lambert, P (2007), 'Robust specification of the roughness penalty prior distribution in spatially adaptive Bayesian P-splines models,' *Computational Statistics and Data Analysis*, **51**(5), pp. 2542–2558.

Lambert, P (2007), 'Archimedean copula estimation using Bayesian splines smoothing techniques,' *Computational Statistics and Data Analysis*, **51**(12), pp. 6307–6320.

Lambert, P (2013), 'Nonparametric additive location-scale models for interval censored data,' *Statistics and Computing*, **23**, pp. 75–90.

Lambert, P & Vandenhende, F (2002), 'A copula based model for multivariate non normal longitudinal data: analysis of a dose titration safety study on a new antidepressant,' *Statistics in Medicine*, **21**, pp. 3197–3217.

Lang, S & Brezger, A (2004), 'Bayesian P-splines,' *Journal of Computational and Graphical Statistics*, **13**, pp. 183–212.

Nelsen, RB (1999), *An Introduction to Copulas*, Springer-Verlag.

Patton, A (2006), 'Modelling asymmetric exchange rate dependence,' *International Economic Review*, **47**(2), pp. 527–556.

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

Spiegelhalter, D, Best, N, Carlin, B & van der Linde, A (2002), 'Bayesian measures of model complexity and fit,' *Journal of the Royal Statistical Society, Series B*, **64**, pp. 583–639.

Vandenhende, F & Lambert, P (2005), 'Local dependence estimation using semiparametric Archimedean copulas,' *The Canadian Journal of Statistics*, **33**, pp. 377–388.

Veraverbeke, N, Omelka, M & Gijbels, I (2011), 'Estimation of a conditional copula and association measures,' *Scandinavian Journal of Statistics*, **38**, pp. 766–780.
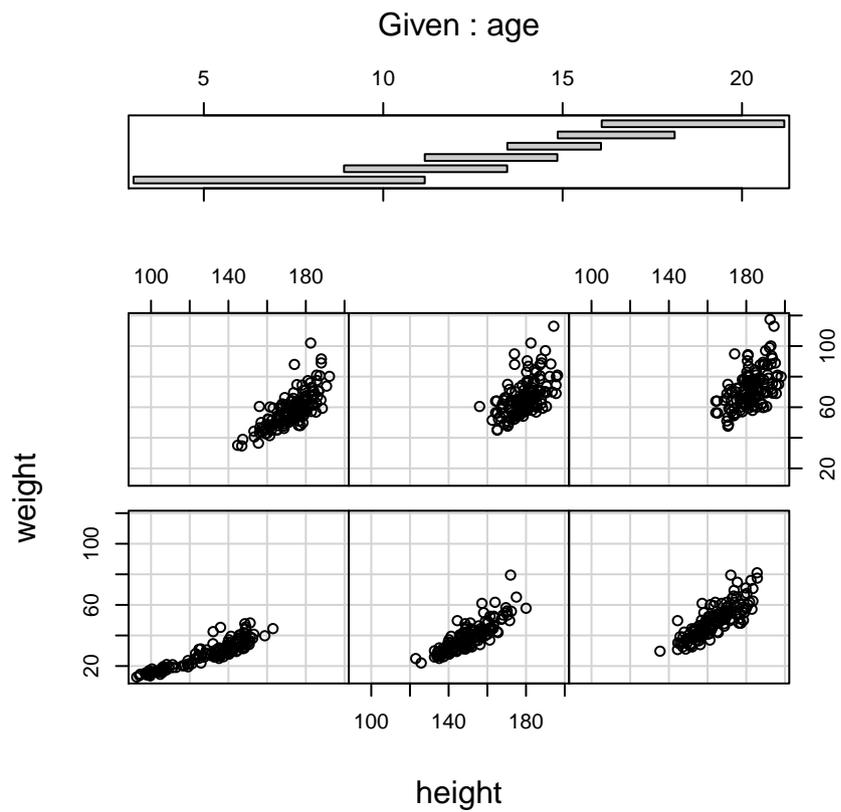
Figure 1. Growth dataset (for 441 Dutch boys): conditional scatterplot of weight versus height for different age classes.
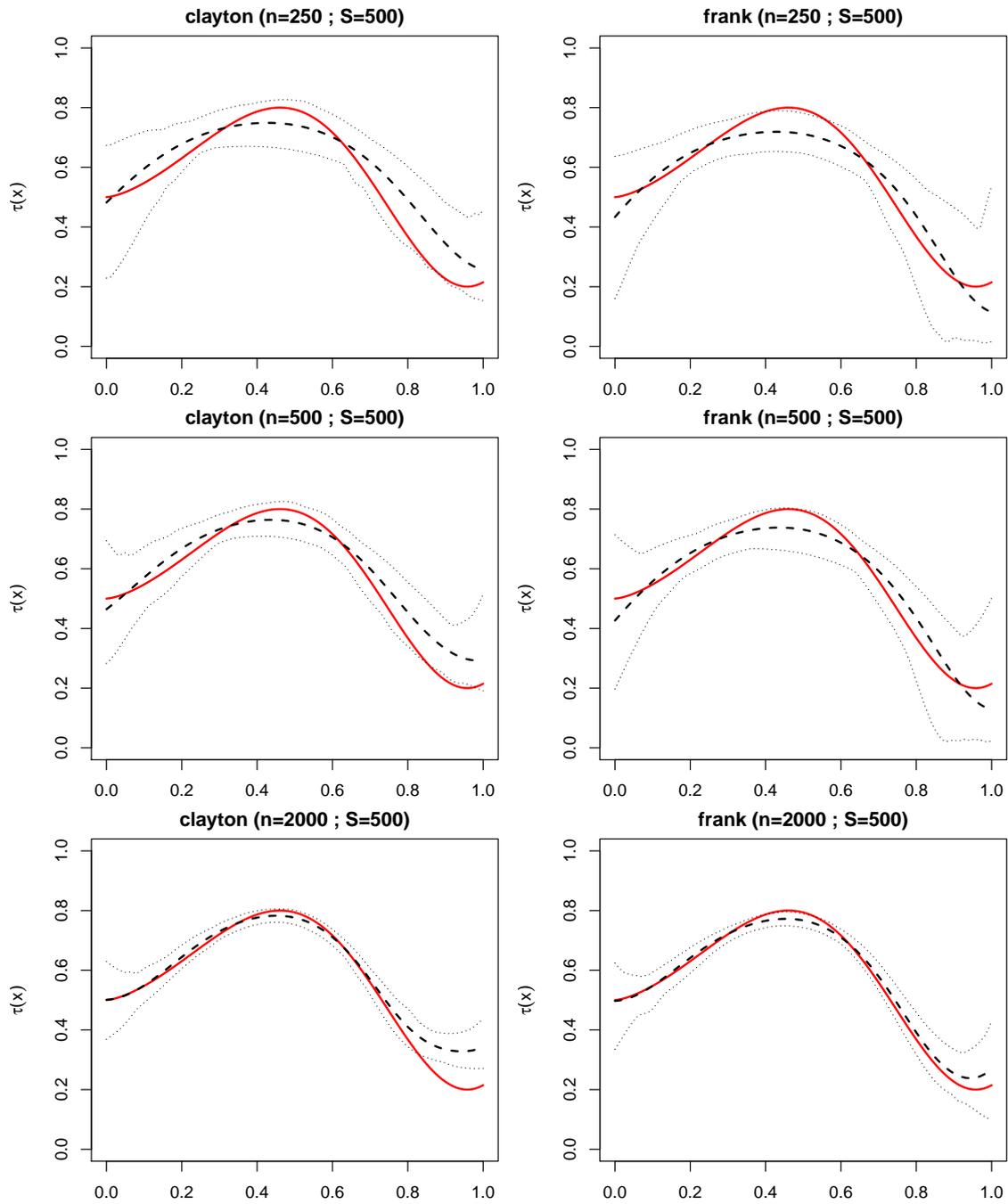
Figure 2. Estimate of the conditional Kendall's tau corresponding to the additive conditional spline Archimedean copula family fitted to n data generated from a Clayton or a Frank copula with varying Kendall's tau: true Kendall's $\tau(x)$ (solid), mean (dashed line) of the S estimated $\tau(x)$ and envelope (dotted lines) containing 95% of the $\tau(x)$ estimates over the $S = 500$ replicates.
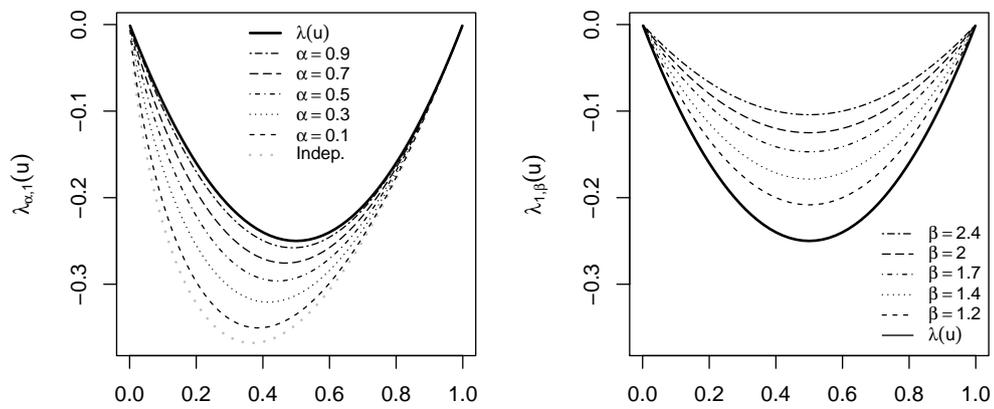
Figure 3. Effect of α and β on the lambda function in the power Archimedean family (for a Clayton($\theta = 1$) reference generator).
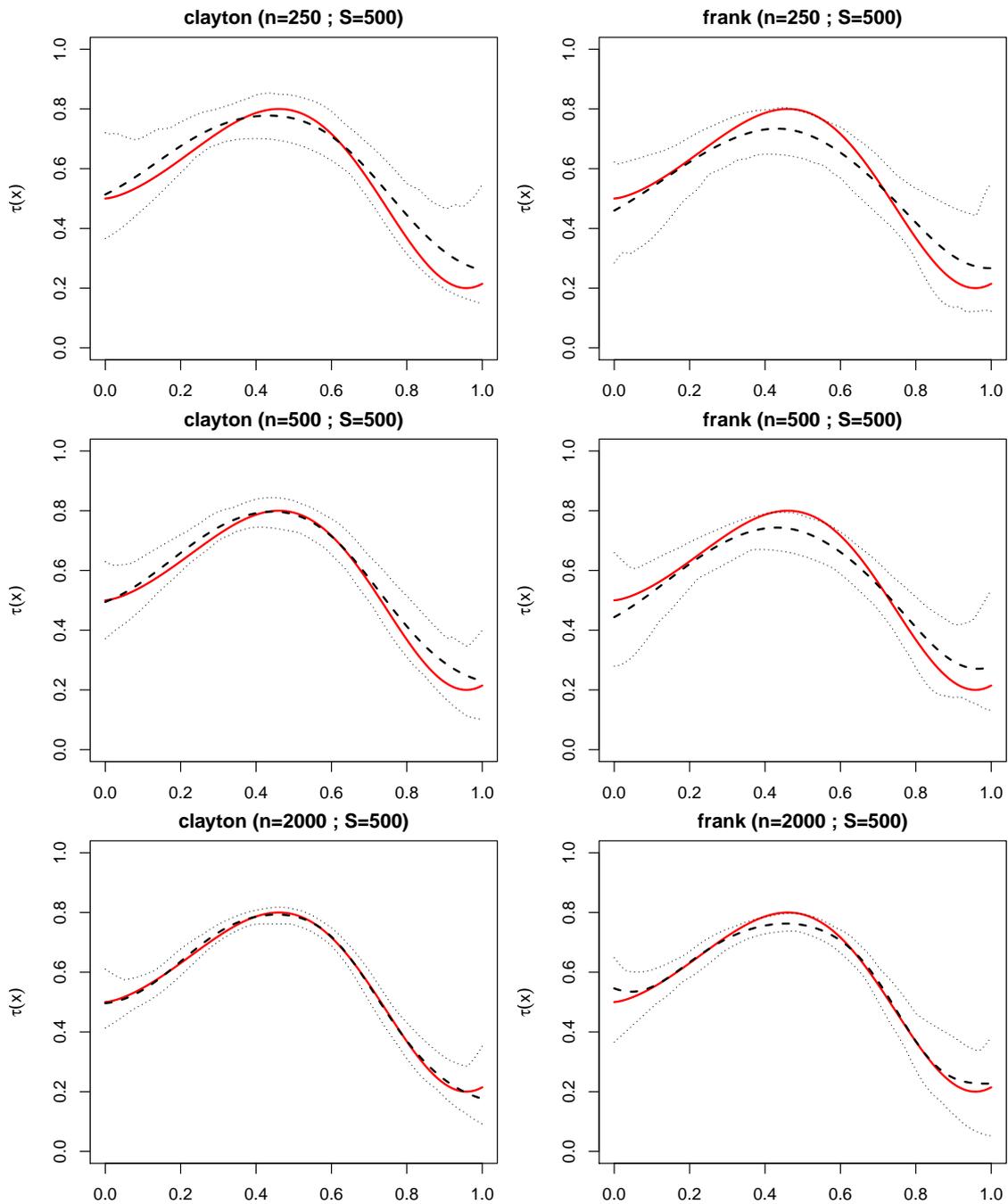
Figure 4. Estimate of the conditional Kendall's tau corresponding to a flex-power Archimedean copula fitted to n data generated from a Clayton or a Frank copula with varying Kendall's tau: true Kendall's $\tau(x)$ (solid), mean (dashed line) of the S estimated $\tau(x)$ and envelope (dotted lines) containing 95% of the $\tau(x)$ estimates over the $S = 500$ replicates.
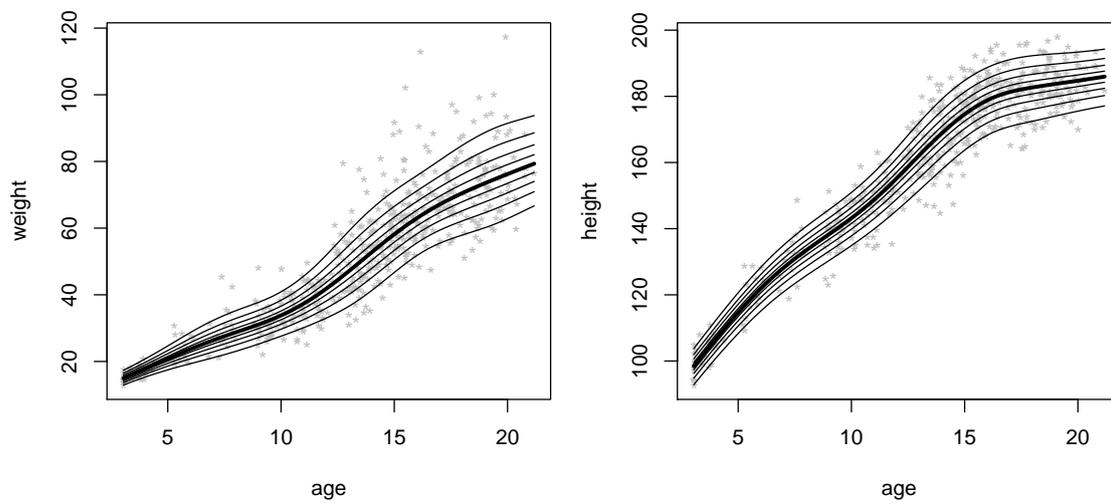
Figure 5. Growth dataset: fitted conditional deciles for `Weight` and `Height` given `Age` using the flexible location-scale model in Lambert (2013).
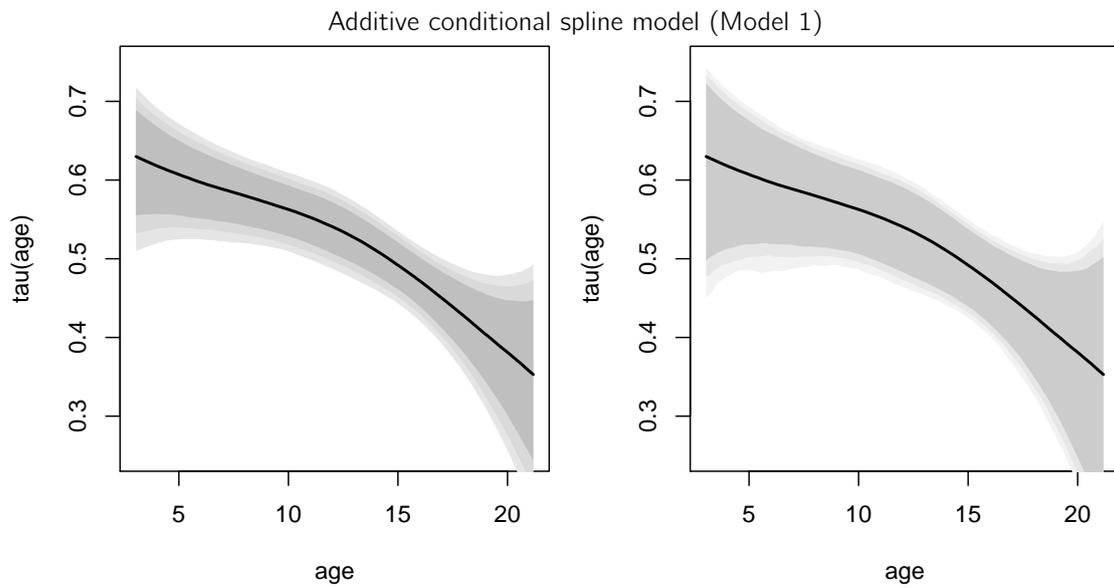
Figure 6. Growth dataset: fitted (solid line) conditional Kendall's tau under the additive conditional spline model (Model 1). Grey areas corresponds to .80, .90 and .95 pointwise (left panel) or simultaneous (right panel) credible regions for $\tau(\text{age})$.
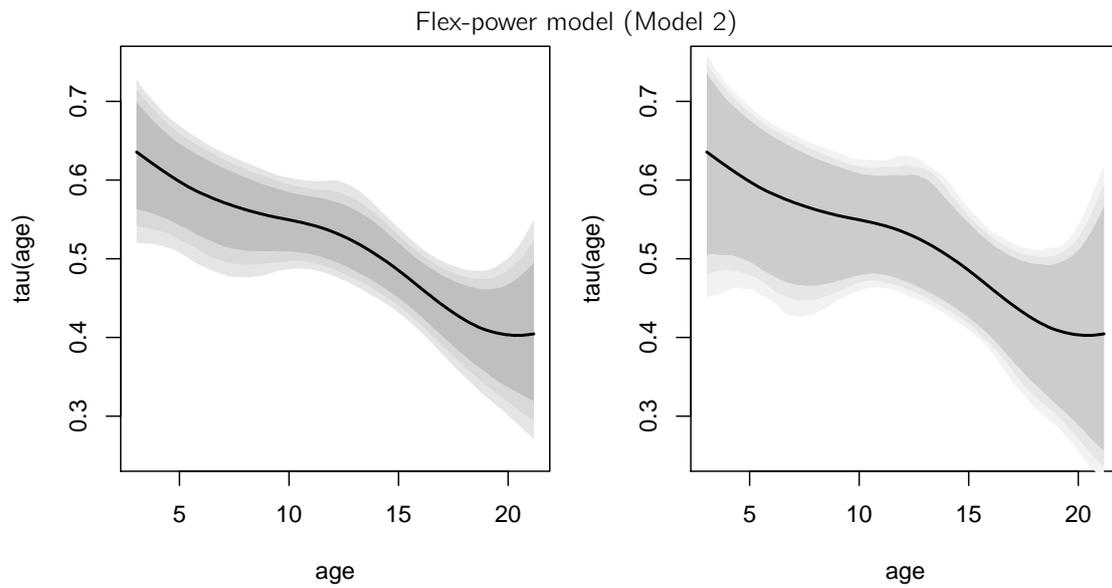
Figure 7. Growth dataset: fitted (solid line) conditional Kendall's tau under the flex-power Archimedean copula model (Model 2). Grey areas corresponds to .80, .90 and .95 pointwise (left panel) or simultaneous (right panel) credible regions for $\tau(\text{age})$.
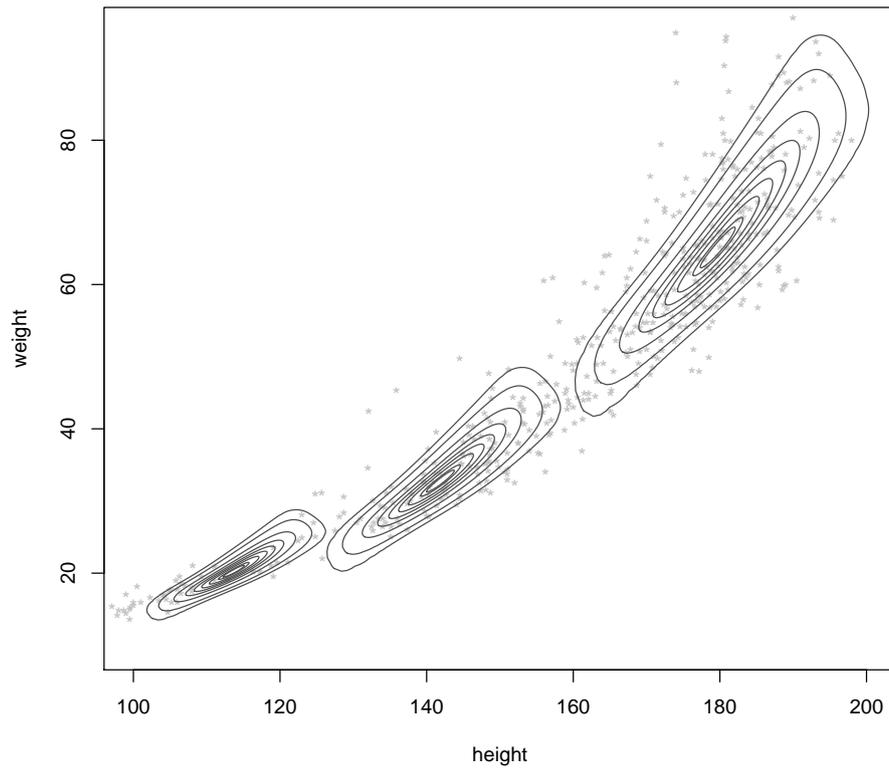
Figure 8. Growth dataset: contours of the fitted joint distribution for `Height` and `Weight` at age 5, 10 and 17.

Table 1. Examples of parametric Archimedean copulas.

Family	θ range	Generator	Kendall's tau
Frank	$(-\infty, +\infty)/\{0\}$	$-\log \frac{e^{-\theta u}-1}{e^{-\theta}-1}$	$1 - \frac{4}{\theta}\left(1 - \frac{1}{\theta}\int_0^\theta \frac{t}{e^t-1}dt\right)$
Clayton	$[-1, +\infty)/\{0\}$	$(u^{-\theta} - 1)/\theta$	$\theta/(\theta + 2)$
Gumbel	$[-1, +\infty)$	$(-\log u)^\theta$	$(\theta - 1)/\theta$

Table 2. Estimation of $\lambda(u) = \varphi(u)/\varphi'(u)$ on a grid of values for u using the spline estimator in (2.3). Biases and RMSE's were estimated using $S = 500$ randomly generated datasets of size n with Kendall's $\tau = .30$.

		Sample size							
		$n = 100$		$n = 250$		$n = 500$		$n = 2000$	
u	$\lambda(u)$	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
		Clayton copula ($\tau = 0.30$)							
0.05	-0.054	-0.007	0.016	-0.004	0.009	-0.002	0.006	-0.000	0.003
0.10	-0.100	-0.007	0.020	-0.002	0.011	-0.001	0.008	-0.001	0.005
0.20	-0.175	-0.003	0.022	0.003	0.014	0.000	0.010	-0.000	0.005
0.30	-0.225	0.000	0.024	0.005	0.018	0.001	0.013	-0.000	0.006
0.40	-0.254	0.003	0.024	0.005	0.019	0.001	0.014	-0.000	0.007
0.50	-0.261	0.005	0.023	0.003	0.018	0.001	0.013	0.000	0.006
0.60	-0.248	0.005	0.021	0.001	0.016	0.001	0.012	0.001	0.006
0.70	-0.215	0.005	0.017	-0.000	0.011	0.001	0.009	0.001	0.004
0.80	-0.162	0.005	0.012	0.000	0.006	0.001	0.006	0.000	0.003
0.90	-0.091	0.005	0.008	0.001	0.003	0.002	0.003	0.000	0.002
0.95	-0.048	0.004	0.006	0.002	0.003	0.001	0.002	0.000	0.001
		Frank copula ($\tau = 0.30$)							
0.05	-0.105	0.009	0.017	0.007	0.012	0.004	0.008	-0.000	0.004
0.10	-0.153	0.004	0.021	0.003	0.013	0.001	0.009	-0.000	0.005
0.20	-0.207	-0.008	0.026	-0.005	0.016	-0.005	0.012	-0.000	0.006
0.30	-0.232	-0.014	0.030	-0.009	0.017	-0.006	0.013	-0.000	0.006
0.40	-0.241	-0.015	0.029	-0.008	0.017	-0.005	0.013	-0.000	0.006
0.50	-0.237	-0.012	0.027	-0.005	0.016	-0.002	0.012	-0.000	0.006
0.60	-0.220	-0.007	0.022	-0.002	0.014	0.001	0.010	-0.000	0.006
0.70	-0.192	-0.001	0.018	0.001	0.012	0.002	0.009	0.000	0.004
0.80	-0.148	0.002	0.012	0.002	0.009	0.001	0.006	-0.000	0.004
0.90	-0.086	0.003	0.007	0.002	0.005	0.001	0.003	0.000	0.002
0.95	-0.046	0.003	0.005	0.001	0.003	0.001	0.002	0.000	0.001
		Gumbel copula ($\tau = 0.30$)							
0.05	-0.105	0.001	0.017	0.000	0.011	-0.001	0.008	-0.001	0.005
0.10	-0.161	-0.002	0.022	-0.001	0.015	-0.002	0.011	-0.001	0.006
0.20	-0.225	-0.006	0.027	-0.002	0.018	-0.002	0.013	-0.001	0.007
0.30	-0.253	-0.007	0.029	-0.002	0.018	-0.001	0.013	-0.000	0.007
0.40	-0.257	-0.007	0.029	-0.001	0.017	-0.000	0.012	0.000	0.006
0.50	-0.243	-0.007	0.028	-0.001	0.015	0.000	0.011	-0.000	0.006
0.60	-0.215	-0.006	0.026	-0.001	0.013	0.000	0.010	-0.000	0.005
0.70	-0.175	-0.005	0.022	-0.001	0.011	0.000	0.008	-0.000	0.004
0.80	-0.125	-0.005	0.017	-0.002	0.010	-0.000	0.007	-0.000	0.004
0.90	-0.066	-0.003	0.010	-0.000	0.006	-0.001	0.005	-0.000	0.003
0.95	-0.034	-0.000	0.006	-0.000	0.004	-0.001	0.003	0.000	0.002

Prepared using staauth.cls

Table 3. Estimation of $\lambda(u) = \varphi(u)/\varphi'(u)$ over $(0, 1)$: RMISE of the spline estimator in (2.3) estimated using $S = 500$ randomly generated datasets of size n.

Copula	τ	Sample size n			
		100	250	500	2000
Clayton	0.15	0.0147	0.0110	0.0081	0.0044
	0.30	0.0153	0.0122	0.0087	0.0043
	0.45	0.0149	0.0106	0.0075	0.0037
Frank	0.15	0.0158	0.0105	0.0075	0.0043
	0.30	0.0172	0.0107	0.0081	0.0042
	0.45	0.0187	0.0116	0.0078	0.0038
Gumbel	0.15	0.0153	0.0118	0.0078	0.0046
	0.30	0.0186	0.0109	0.0083	0.0046
	0.45	0.0182	0.0073	0.0079	0.0040

Table 4. Mean coverages of pointwise credible intervals for $\lambda(u)$ estimated using $S = 500$ randomly generated datasets of size n under the Clayton, Frank or Gumbel copula with different Kendall's τ.

		$\tau = 0.45$			$\tau = 0.30$			$\tau = 0.15$		
		Nominal coverage								
	n	0.80	0.90	0.95	0.80	0.90	0.95	0.80	0.90	0.95
Clayton	100	0.83	0.91	0.94	0.83	0.92	0.95	0.87	0.95	0.97
	250	0.84	0.93	0.97	0.84	0.92	0.96	0.86	0.94	0.97
	500	0.82	0.91	0.96	0.82	0.91	0.96	0.87	0.95	0.98
	2000	0.81	0.91	0.96	0.82	0.92	0.97	0.84	0.93	0.96
Frank	100	0.80	0.91	0.95	0.85	0.93	0.96	0.87	0.96	0.99
	250	0.82	0.91	0.96	0.86	0.94	0.97	0.86	0.94	0.97
	500	0.82	0.91	0.96	0.86	0.94	0.97	0.87	0.95	0.98
	2000	0.82	0.91	0.96	0.84	0.92	0.96	0.85	0.94	0.97
Gumbel	100	0.84	0.93	0.96	0.82	0.90	0.94	0.89	0.95	0.98
	250	0.90	0.96	0.98	0.86	0.95	0.98	0.83	0.93	0.97
	500	0.83	0.92	0.96	0.86	0.94	0.97	0.84	0.92	0.96
	2000	0.81	0.91	0.96	0.82	0.92	0.96	0.83	0.92	0.97

Table 5. Growth dataset: effective number of parameters (e.d.) and DIC in the fitted copula models.

Copula model	# par.	e.d.	DIC
Model 0 (Unconditional)	11	3.2	-310.4
Model 1 (additive)	16	6.0	-319.0
Model 2 (flex-power)	21	7.0	-318.4

Prepared using staauth.cls