Semiparametric copula quantile regression for complete or censored data

Mickaël De Backer and Anouar El Ghouch

Institut de Statistique, Biostatistique et Sciences Actuarielles
Université catholique de Louvain, Voie du Roman Pays, 20, B-1348 Louvain-la-Neuve,
Belgium
e-mail: mickael.debacker@uclouvain.be*;* anouar.elghouch@uclouvain.be

Ingrid Van Keilegom

Institut de Statistique, Biostatistique et Sciences Actuarielles
Université catholique de Louvain, Voie du Roman Pays, 20, B-1348, Louvain-la-Neuve,
Belgium

KU Leuven, Research Centre for Operations Research and Business Statistics,
Naamsestraat 69, 3000 Leuven, Belgium
e-mail: ingrid.vankeilegom@kuleuven.be

Abstract: When facing multivariate covariates, general semiparametric regression techniques come at hand to propose flexible models that are unexposed to the curse of dimensionality. In this work a semiparametric copula-based estimator for conditional quantiles is investigated for both complete or right-censored data. In spirit, the methodology is extending the recent work of Noh, El Ghouch and Bouezmarni [34] and Noh, El Ghouch and Van Keilegom [35], as the main idea consists in appropriately defining the quantile regression in terms of a multivariate copula and marginal distributions. Prior estimation of the latter and simple plug-in lead to an easily implementable estimator expressed, for both contexts with or without censoring, as a weighted quantile of the observed response variable. In addition, and contrary to the initial suggestion in the literature, a semiparametric estimation scheme for the multivariate copula density is studied, motivated by the possible shortcomings of a purely parametric approach and driven by the regression context. The resulting quantile regression estimator has the valuable property of being automatically monotonic across quantile levels. Additionally, the copula-based approach allows the analyst to spontaneously take account of common regression concerns such as interactions between covariates or possible transformations of the latter. From a theoretical prospect, asymptotic normality for both complete and censored data is obtained under classical regularity conditions. Finally, numerical examples as well as a real data application are used to illustrate the validity and finite sample performance of the proposed procedure.

Keywords and phrases: Semiparametric regression, censored quantile regression, multidimensional copula modelling, semiparametric vine copulas, kernel smoothing, polynomial local-likelihood, probit transformation.

1. Introduction

Quantile regression is a prevailing method when it comes to investigating the possible relationships between a d-dimensional covariate $\boldsymbol{X}$ and a response variable T. Since the seminal work of Koenker and Basset [23], quantile regression has received notable interest in the literature on theoretical and applied statistics as a very attractive alternative to the classical mean regression model based on quadratic loss. As the latter only captures the central tendency of the data, there are many cases and nice examples where mean regression is uninformative with respect to studying the conditional upper or lower quantiles. For an interesting application, see for example Elsner, Kossin and Jagger [11]. A comprehensive review of quantile regression as a robust (to outliers) and flexible (to error distribution) method can be found in Koenker [22].

The first main concern of this paper is related to the estimation of a quantile regression function where the response variable T is completely observed. A wide literature on the subject includes fully parametric, semiparametric and nonparametric methodologies. When several covariates are to be taken into account, fully parametric methodologies are known to be highly sensitive to model misspecification and may lack the flexibility needed for an adequate modelling. On the other hand, in spite of their great flexibility, fully nonparametric methods such local quantile regression as proposed by Spokoiny, Wang and Härdle [42] are typically affected by the curse of dimensionality. In light of these restrictions, semiparametric estimation procedures such as a single-index regression (Wu, Yu and Yu [46], Zhu, Huang and Li [48]) come at hand when the dimension of the covariate is high.

In this context, Noh, El Ghouch and Bouezmarni [34] and Noh, El Ghouch and Van Keilegom [35] suggested to estimate a regression function based on the copula that defines the dependence structure between the variables of interest. The central idea of the methodology is to express the conditional quantile function in terms of a copula density and marginal distributions. Such an approach allows the analyst to take profit of copula modelling and straightforwardly avoid common regression issues such as the consideration of transformed covariates and the possible inclusion of interactions among the latter. In their original paper, Noh, El Ghouch and Bouezmarni [34] suggested subsequently to leave the marginal distributions unspecified while assuming a parametric model for the copula. Overall, their suggested approach results in a semiparametric regression estimator that is not exposed to the curse of dimensionality. However, Dette, Van Hecke and Volgushev [8] highlighted that this proposed methodology may suffer as such from possible shortcomings that are induced by the misspecification of the parametric copula. The first contribution of this paper is to circumvent such issues by proposing an alternative semiparametric estimation strategy for the copula itself that is motivated by the regression context. The resulting regression estimator is flexible for multidimensional data, easy to implement and does not require any iterative procedure in opposition to existing semiparametric alternatives.

The second objective of this paper is to propose a copula-based methodology in the context of survival analysis, where right censoring of T may arise. In this situation, instead of fully observing the variable of interest, one only observes the minimum of it and a censoring variable. For instance, in clinical studies, censoring may occur because of the withdrawal of patients from the study, the end of the follow-up period, etc. In this context, quantile regression becomes attractive as an alternative to popular regression techniques like the Cox proportional hazards model or the accelerated failure time model, as is argued in Koenker and Bilias [24], Koenker and Geling [25] and Portnoy [37]. Additional appealing properties of the method include the fact that it allows for modelling heterogeneity of the variance and it does not necessarily impose a proportional effect of the covariates on the hazard over the duration time as opposed to the popular Cox model.

As is the case for the uncensored situation, existing literature on censored quantile regression includes fully parametric, semiparametric and nonparametric methodologies such as local linear smoothing proposed by El Ghouch and Van Keilegom [10]. The introduction of censored quantile regression goes back to Powell [39] for linear models and fixed censoring, that is, presuming that the censoring times are known for all observations. For random censoring, a wide literature may be found, among others, in Ying, Jung and Wei [47], Portnoy [37], Wang and Wang [45] and Leng and Tong [27]. Still, similarly to regression models with complete observations, a linear approach may be too restrictive for real data applications.

For semiparametric models in the context of censored quantile regression, an interesting approach was proposed by Bücher, El Ghouch and Van Keilegom [2] where a single-index model for the conditional quantile function is studied under the assumption of independence between the covariates and the censoring variable. The single-index structure assumes that the objective function depends linearly on the covariates through an unknown link function, making the proposed model (i.e. under the aforementioned assumption) insensitive to the curse of dimensionality since the nonparametric part is of dimension one. However, besides the latter single-index model, existing literature on flexible multidimensional methodologies is rather sparse.

Hence, the second main contribution of this paper is to extend this literature on multivariate quantile regression estimation in the possible presence of censored data by providing a rich, flexible and robust alternative based on the copula function. In essence, the proposed methodology mimics the work of Noh, El Ghouch and Bouezmarni [34] and Noh, El Ghouch and Van Keilegom [35], as the central idea is here again to rewrite the conditional quantile function in terms of marginal distributions and an appropriate copula density. Employing the proposed multivariate copula density estimation strategy, the resulting regression estimator enjoys the same qualities as for complete observations. Consequently, by taking advantage of copula modelling in regression models, the proposed methodology provides a new class of estimators that allow practitioners to flexibly analyse multidimensional survival data.

The rest of this paper is organised as follows. Developing a semiparametric copula-based estimation procedure for quantile regression with complete observations is the topic of Section 2. Section 3 develops the methodology for the copula-based quantile regression estimator when confronted to possible censoring of the responses. The asymptotic properties of the proposed estimators for both complete and censored data are obtained in Section 4 and the finite sample performance is illustrated by means of Monte Carlo simulations in Section 5, where both the semiparametric copula estimation strategy and the overall performance of our estimator for complete and censored data are investigated. Section 6 provides a brief application to real censored data. Lastly, the proofs of our asymptotic properties are deferred to the Appendix.

2. Copula-based estimator for complete data

2.1. Background for copula-based quantile regression

Let $\boldsymbol{X} = (X_1, \ldots, X_d)^{\mathsf{T}}$ be a covariate vector of dimension $d \geq 1$ and T be a (time-to-event) response variable with marginal continuous cumulative distribution functions (c.d.f.) $F_1, \ldots, F_d$ and F_T, respectively. Throughout this paper, we denote by f_j and f_T the density of $X_j, j = 1, \ldots, d$, and T, respectively. From the pioneering work of Sklar [41], for a given $\boldsymbol{x} = (x_1, \ldots, x_d)^{\mathsf{T}}$, the c.d.f. of $(T, \boldsymbol{X})$ evaluated at $(t, \boldsymbol{x})$ can be expressed as $C_{T\boldsymbol{X}}(F_T(t), \boldsymbol{F}(\boldsymbol{x}))$, where $\boldsymbol{F}(\boldsymbol{x}) = (F_1(x_1), \ldots, F_d(x_d))^{\mathsf{T}}$ and $C_{T\boldsymbol{X}}$ is the unique copula distribution of $(T, \boldsymbol{X})$ defined by $C_{T\boldsymbol{X}}(u_0, u_1, \ldots, u_d) = \mathbb{P}(U_0 \leq u_0, U_1 \leq u_1, \ldots, U_d \leq u_d)$, with $U_0 = F_T(T)$ and $U_j = F_j(X_j)$, $j = 1, \ldots, d$. From Sklar's theorem, it is clear that the copula $C_{T\boldsymbol{X}}$ disjoints the marginal behaviors of T and $\boldsymbol{X}$ from their dependence structure, hence allowing a great modelling flexibility. For a book length treatment of copulas, see Nelsen [33] and Joe [21].

The object of interest of this paper, the τ-th conditional quantile function of the dependent variable T given $\boldsymbol{X} = \boldsymbol{x}$, denoted by $m_\tau(\boldsymbol{x})$, is defined for any $\tau \in (0, 1)$ as $m_\tau(\boldsymbol{x}) = \inf\{t : F_{T|\boldsymbol{X}}(t|\boldsymbol{x}) \geq \tau\}$ where $F_{T|\boldsymbol{X}}$ is the conditional c.d.f. of T given $\boldsymbol{X}$, or, equivalently,

$$m_\tau(\boldsymbol{x}) = \arg\min_a \mathbb{E}\big(\rho_\tau(T - a)|\boldsymbol{X} = \boldsymbol{x}\big), \tag{1}$$

where $\rho_\tau(u) = u(\tau - \mathbb{1}(u \leq 0))$ is the so-called "check" function, and $\mathbb{1}(\cdot)$ is the indicator function.

To motivate our approach, we suppose in this section that there is no censoring and that we observe an i.i.d. sample $(T_i, \boldsymbol{X}_i), i = 1, \ldots, n$, from $(T, \boldsymbol{X})$. In this context, following the definition of a copula function, Noh, El Ghouch and Van Keilegom [35] noted that the conditional quantile function of T given $\boldsymbol{X} = \boldsymbol{x}$ may be expressed as

$$m_\tau(\boldsymbol{x}) = \arg\min_a \mathbb{E}\Big[\rho_\tau(T - a)\, c_{T\boldsymbol{X}}(F_T(T), \boldsymbol{F}(\boldsymbol{x}))\Big], \tag{2}$$

where $c_{TX}(u_0, \boldsymbol{u}) \equiv c_{TX}(u_0, u_1, \ldots, u_d) = \partial^{d+1} C_{TX}(u_0, u_1, \ldots, u_d)/\partial u_0 \partial u_1 \ldots \partial u_d$ is the copula density corresponding to C_{TX}. Consequently, any given estimators $\widehat{c}_{TX}$, $\widehat{F}_T$ and $\widehat{F}_j$ of c_{TX}, F_T and F_j, $j = 1, \ldots, d$, respectively, automatically yield an estimator of $m_\tau(\boldsymbol{x})$ given by

$$\widehat{m}_\tau(\boldsymbol{x}) = \arg\min_a \sum_{i=1}^n \rho_\tau(T_i - a)\, \widehat{c}_{TX}(\widehat{F}_T(T_i), \widehat{\boldsymbol{F}}(\boldsymbol{x})), \tag{3}$$

with $\widehat{\boldsymbol{F}}(\boldsymbol{x}) = (\widehat{F}_1(x_1), \ldots, \widehat{F}_d(x_d))^{\mathsf{T}}$. As indicated earlier, Noh, El Ghouch and Van Keilegom [35] suggest to estimate the marginals nonparametrically and to consider a parametrization of the copula density, that is, assume that the latter belongs to a certain parametric family of copula densities $\mathcal{C} = \{c(u_0, \boldsymbol{u}; \boldsymbol{\theta}), \boldsymbol{\theta} \in \Theta \subset \mathbb{R}^p\}$.

2.2. A motivational one-dimensional example

In this paper however, we consider an alternative estimation approach for the copula density, motivated by the issues related to the possible misspecification of the parametric approach. To highlight this shortcoming and illustrate how one may circumvent it, we consider in this section the simplistic example reported by Dette, Van Hecke and Volgushev [8] with a single covariate, where $(T_i, X_{1i}), i = 1, \ldots, n$, are i.i.d. random variables with $T_i = (X_{1i} - 0.5)^2 + \sigma\epsilon_i$, $X_{1i} \sim U[0, 1]$, $\sigma = 0.025$ and $\epsilon_i, i = 1, \ldots, n$, are i.i.d. standard normal random variables. In this situation, where the true quantile regression function is non-monotonic in the covariate, it is found that most of the common parametric copula families still yield a monotone estimation of the regression function, thereby providing a rather poor fit of the latter. This is illustrated in Figure 1a, where the estimation is carried out for $\tau = 0.5$, $n = 500$ and using three common parametric copulas.

As the roots of the above-mentioned limitation are not intrinsic to a copula-based approach, but rather to be attributed to the limited set of parametric copula families existing in the literature, a natural alternative for low dimensional covariates would be to consider a fully nonparametric estimation of the copula density itself. The resulting, and adequate, quantile regression estimation is depicted in Figure 1b.

For this approach to be appropriate, a suitable nonparametric estimation of the copula density is required. In the literature, several specific kernel-based methodologies have been proposed in order to correct for well-known bias issues at the boundaries, as the density of interest is only supported on the unit square. These include for instance the mirror reflection method (Gijbels and Mielniczuk [16]) or the boundary kernel method (Chen and Huang [5]). In this paper however, we will adopt the technique proposed by Charpentier, Fermanian and Scaillet [3] and Geenens, Charpentier and Paindaveine [14] which is summarized as follows: to estimate for instance the copula density c_{TX_1} in our motivational example, given a bivariate sample $(U_{0i}, U_{1i}), i = 1, \ldots, n$, from $(F_T(T), F_1(X_1))$, the main idea is to appropriately project the initial
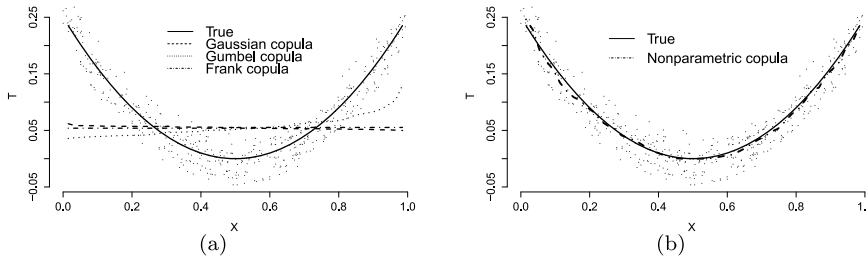
FIG 1. *Copula-based quantile regression estimates of the one-dimensional minimalistic example. The Gaussian, Gumbel and Frank copulas are used for parametric copula estimation in (a), while (b) depicts the regression fit resulting from a nonparametric estimation of the copula density using the procedure of Geenens, Charpentier and Paindaveine [14] (more details given below).*

data on an unbounded support with the purpose of then estimating the obtained bivariate transformed density by means of standard techniques (standard kernel (Charpentier, Fermanian and Scaillet) or polynomial local-likelihood (Geenens, Charpentier and Paindaveine)). Using the invariance property of copulas to increasing transformations of their margins, the estimation of the copula density is then obtained by back-transformation on the unit square. That is, using the example of a probit transformation, one may estimate the copula density c_{TX_1} at $(u_0, u_1) \in [0, 1]^2$ by

$$\frac{\widehat{f}_{01}\big(\Phi^{-1}(u_0), \Phi^{-1}(u_1)\big)}{\phi\big(\Phi^{-1}(u_0)\big)\phi\big(\Phi^{-1}(u_1)\big)},$$

where ϕ and Φ stand for the standard normal density and c.d.f., respectively, and where $\widehat{f}_{01}$ is a bivariate density estimator of the projected data $(\Phi^{-1}(U_{0i}), \Phi^{-1}(U_{1i}))$, $i = 1, \ldots, n$. This transformation technique, coupled with polynomial local-likelihood estimation for f_{01} in order to allow for possible unbounded copula density estimates, is shown to outperform its competitors in most scenarios in a detailed simulation study in Geenens, Charpentier and Paindaveine. An interested reader may find an exhaustive comparison of existing methodologies for bivariate copula density estimation may be found in Nagler [31]. Furthermore, fully nonparametric multidimensional copulas are studied in Hobæk Haff and Segers [19] and Nagler and Czado [32].

2.3. A semiparametric copula estimator for multivariate covariates

Recalling that we intend to handle multivariate covariates in this paper, we will not adopt a purely nonparametric approach as in the previously-described motivational example. Instead, we will prefer a copula estimation strategy that provides sufficient flexibility to the multidimensional estimator while avoiding dimension related constraints. More specifically, we note that any multivariate

copula density can be decomposed into two parts as follows:

$$c_{T\boldsymbol{X}}\big(F_T(t), \boldsymbol{F}(\boldsymbol{x})\big) = c_{TX_1}\big(F_T(t), F_1(x_1)\big) \times \ldots \times c_{TX_d}\big(F_T(t), F_d(x_d)\big) \qquad (4)$$

$$\times\, c_{X_1\ldots X_d | T}\big(F_{1|T}(x_1|t), \ldots, F_{d|T}(x_d|t)|t\big), \qquad (5)$$

where $F_{j|T}$, $j = 1, \ldots, d$, denotes the conditional c.d.f. of X_j given T. The first part of the decomposition contains the product of d bivariate copula densities related to the dependence of T with every covariate, whereas the second part captures the conditional dependence of $\boldsymbol{X}$ given $T = t$. In the general regression context, part (4) may then be interpreted as the dependence of actual interest since it focuses on the relationship between the response variable with every covariate. On the contrary, part (5) may be viewed in such framework as a 'noisy' dependence, or, more precisely, a correction parameter for possible (conditional) dependence among covariates. Consequently, a natural reasoning suggests to provide as much flexibility as possible to the modelling of part (4), while keeping the estimation of part (5) uncomplicated. We therefore advocate to estimate nonparametrically the d bivariate copulas of interest and, subsequently, exploit standard parametric techniques for the second part of the multivariate copula density. The latter involve, among others, nested Archimedean copulas (see e.g. Hofert and Pham [20] and Joe [21]), factor copulas (see e.g. Oh and Patton [36]) and, arguably the most popular, vine copulas (see e.g. Czado [7], Joe [21] and references therein). Note however that, for the estimation of (5) in practice, one is first advised to adopt the so-called *simplifying assumption* which stipulates that the conditioning on $T = t$ is fully captured by the conditional marginals. In other words, in (5), the conditional copula itself is not affected by the conditioning on T. This assumption turns out to be the cornerstone of vine copula models as it keeps them tractable for inference and model selection. For more details about this and its implications, see Hobæk Haff, Aas and Frigessi [18] and Stöber, Joe and Czado [43].

Conclusively, in this article we propose to adopt the following detailed procedure for the modelling and estimation of the multivariate copula density:

(1) Based on original observations $(T_i, X_{1i}, \ldots, X_{di}), i = 1, \ldots, n$, construct 'pseudo-observations', needed for the estimation of (4), using rescaled versions of empirical distributions:

$$\widehat{U}_{0i} = \frac{1}{n+1} \sum_{k=1}^n \mathbb{1}(T_k \le T_i) \qquad \widehat{U}_{ji} = \frac{1}{n+1} \sum_{k=1}^n \mathbb{1}(X_{jk} \le X_{ji}),$$

$$i = 1, \ldots, n, \quad j = 1, \ldots, d,$$

where the factor $1/(n+1)$, commonly adopted in the copula literature, aims at keeping the constructed observations in the interior of $[0, 1]$.

(2) Based on $(\widehat{U}_{0i}, \widehat{U}_{ji}), i = 1, \ldots, n$, estimate each bivariate copula density c_{TX_j}, $j = 1, \ldots, d$, in (4) using a bivariate kernel density estimator. This can be achieved via the local-polynomial probit methodology of Geenens, Charpentier and Paindaveine, or any other estimator satisfying assumption (C7) given below.

(3) Compute the pseudo-observations needed for the estimation of (5) as:

$$\widehat{F}_{j|T}(X_{ji}|T_i) = \int_0^{\widehat{U}_{ji}} \widehat{c}_{TX_j}(\widehat{U}_{0i}, s) \, \mathrm{d}s.$$

This relationship is at the origin of the sequential nature of the vine copula estimation scheme (see e.g. Czado [7]).

(4) Lastly, for the estimation of $c_{X_1\ldots X_d|T}$, adopt the simplifying assumption and use standard parametric vine techniques on the dataset $(\widehat{F}_{1|T}(X_{1i}|T_i),$ $\ldots, \widehat{F}_{d|T}(X_{di}|T_i))$, $i = 1, \ldots, n$.

Finally, to graphically illustrate this procedure in a multivariate example, we consider a simple extension to the previously-described setting with two independent covariates, where (T_i, X_{1i}, X_{2i}), $i = 1, \ldots, n$, are i.i.d. random variables with $T_i = (X_{1i}-0.5)^2 + (X_{2i}-0.5)^2 + \sigma\epsilon_i$, $X_{1i} \sim U[0,1]$, $X_{2i} \sim U[0,1]$, $\sigma = 0.025$ and $\epsilon_i, i = 1, \ldots, n$, are i.i.d. standard normal random variables. In this situation, a fully parametric vine copula estimation again yields an inappropriate quantile regression estimation, as depicted in Figure 2a for $\tau = 0.5$ and $n = 500$. In opposition, the proposed semiparametric copula density estimation allows the regression estimator to suitably capture the non-monotonic features of both covariates. This can be observed in Figure 2b. A more exhaustive analysis of the performance of the proposed estimator with respect to existing competitors will be provided in Section 5.
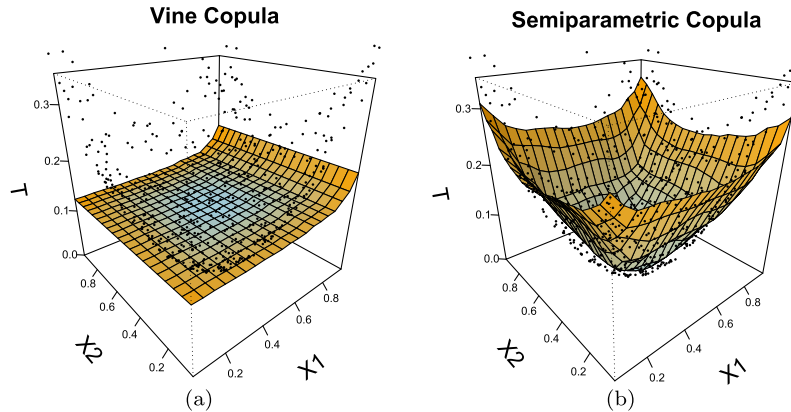

FIG 2. *Copula-based quantile regression estimates of the two-dimensional minimalistic example. A parametric vine copula estimation is at the basis of the quantile estimation in (a), while (b) illustrates the semiparametric copula estimation approach proposed in this paper.*

3. Copula-based estimator for censored data

In the presence of censoring, the estimation equation (3) becomes inappropriate as we do not fully observe the response variables T_i. Instead, we only observe a sequence of i.i.d. triplets $(Y_i, \Delta_i, \boldsymbol{X}_i)$, $i = 1, \ldots, n$, from $(Y, \Delta, \boldsymbol{X})$, where $Y = \min(T, C)$, $\Delta = \mathbb{1}(T \leq C)$ and C denotes the censoring variable, assumed to be independent of T given $\boldsymbol{X}$. In order to take censoring into account in the estimation procedure, the first step is to note that, for any measurable function $\varphi : \mathbb{R} \to \mathbb{R}$,

$$\mathbb{E}\big(\varphi(T)|\boldsymbol{X} = \boldsymbol{x}\big) = \mathbb{E}\left(\varphi(Y)\frac{\Delta}{1 - G_C(Y - |\boldsymbol{x})}\Big|\boldsymbol{X} = \boldsymbol{x}\right), \tag{6}$$

where $G_C(c|\boldsymbol{x}) = \mathbb{P}(C \leq c|\boldsymbol{X} = \boldsymbol{x})$ denotes the conditional distribution of C given $\boldsymbol{X} = \boldsymbol{x}$. This, along with (1), suggests a natural way to handle censoring for quantile regression by replacing the function φ with the check function.

At this stage, in the objective of proposing a similar methodology to the one developed in Section 2, we propose to work on the obtained conditional expectation in (6) before considering the introduction of copulas. The underlying rationale is that the latter conditional expectation is in fact the joint conditional expectation of (Y, Δ) given $\boldsymbol{X} = \boldsymbol{x}$. Adopting an analogous reasoning as the one presented by Noh, El Ghouch and Van Keilegom for the uncensored case at this point would therefore result in the insertion of the joint copula of $(Y, \Delta, \boldsymbol{X})$, hence exposing the estimation procedure to the lack of uniqueness of the copula given that Δ is a discrete (binary) variable. An interesting opinion on several pitfalls appearing for copulas with discrete variables can be found in Embrechts [12]. Additional details may also be found in Genest and Nešlehová [15].

Instead, the idea is to work on the joint conditional expectation so as to bypass the issues related to the copula of $(Y, \Delta, \boldsymbol{X})$. In short, our intention is to discard the problem by obtaining the copula of $(Y, \boldsymbol{X})$ conditionally on $\Delta = 1$, for which no specific technical difficulties are involved. To that end, using the notations H^u (resp. h^u) as a shorthand for a given distribution (resp. density) conditionally on $\Delta = 1$, first note that

$$\mathbb{E}\left(\rho_\tau(Y - a)\frac{\Delta}{1 - G_C(Y - |\boldsymbol{x})}\Big|\boldsymbol{X} = \boldsymbol{x}\right)$$
$$= \int_{\mathbb{R}^+} \rho_\tau(y - a)\,\frac{1}{1 - G_C(y - |\boldsymbol{x})}\,\mathrm{d}F_{Y,\Delta|\boldsymbol{X}}(y, 1|\boldsymbol{x}), \quad (7)$$

where $F_{Y,\Delta|\boldsymbol{X}}(y, 1|\boldsymbol{x}) = \mathbb{P}(Y \leq y, \Delta = 1|\boldsymbol{X} = \boldsymbol{x}) = p(\boldsymbol{x})\int_{-\infty}^{y} f^u_{Y\boldsymbol{X}}(z, \boldsymbol{x})/f^u(\boldsymbol{x})\,\mathrm{d}z$, with $p(\boldsymbol{x}) = \mathbb{P}(\Delta = 1|\boldsymbol{X} = \boldsymbol{x})$ and where $f^u_{Y\boldsymbol{X}}$ and f^u denote the conditional densities of $(Y, \boldsymbol{X})$ and $\boldsymbol{X}$ given $\Delta = 1$, respectively. Hence, using the definition of a copula function in a similar spirit as Noh, El Ghouch and Bouezmarni [34], one obtains

$$\mathrm{d}F_{Y,\Delta|\boldsymbol{X}}(y, 1|\boldsymbol{x}) = p(\boldsymbol{x})\,\frac{c^u_{Y\boldsymbol{X}}\big(F^u_Y(y), \boldsymbol{F}^u(\boldsymbol{x})\big)}{c^u_{\boldsymbol{X}}\big(\boldsymbol{F}^u(\boldsymbol{x})\big)}\,f^u_Y(y)\,\mathrm{d}y,$$

where $c_{Y\boldsymbol{X}}^{u}(u_0, \boldsymbol{u}) = \partial^{d+1} C_{Y\boldsymbol{X}}^{u}(u_0, u_1, \ldots, u_d)/\partial u_0 \partial u_1 \ldots \partial u_d$ is the copula density corresponding to the copula $C_{Y\boldsymbol{X}}^{u}$ of $(Y, \boldsymbol{X}|\Delta = 1)$, $c_{\boldsymbol{X}}^{u}(\boldsymbol{u}) = \partial^{d} C_{\boldsymbol{X}}^{u}(u_1, \ldots, u_d)/\partial u_1 \ldots \partial u_d$ is the copula density of $(\boldsymbol{X}|\Delta = 1)$, and $\boldsymbol{F}^{u}(\boldsymbol{x}) = (F_1^{u}(x_1), \ldots, F_d^{u}(x_d))^{\mathsf{T}}$. Inserting this last expression in (7), we may write

$$\mathbb{E}\left[\rho_\tau(T - a)|\boldsymbol{X} = \boldsymbol{x}\right]$$
$$= \mathbb{E}\left[\rho_\tau(Y - a)\frac{\Delta}{1 - G_C(Y - |\boldsymbol{x})}\frac{p(\boldsymbol{x})}{\mathbb{P}(\Delta = 1)}\frac{c_{Y\boldsymbol{X}}^{u}\left(F_Y^{u}(Y), \boldsymbol{F}^{u}(\boldsymbol{x})\right)}{c_{\boldsymbol{X}}^{u}\left(\boldsymbol{F}^{u}(\boldsymbol{x})\right)}\right].$$

Applying this equality in the context of quantile regression, interestingly, one eventually retrieves an expression analogous to (2):

$$m_\tau(\boldsymbol{x}) = \arg\min_a \mathbb{E}\left[\rho_\tau(Y - a)\,W(\boldsymbol{x})\,c_{Y\boldsymbol{X}}^{u}\left(F_Y^{u}(Y), \boldsymbol{F}^{u}(\boldsymbol{x})\right)\right], \qquad (8)$$

where $W(\boldsymbol{x}) \equiv \Delta/(1 - G_C(Y - |\boldsymbol{x}))$. Note that the copula in question in (8) is determined by strictly fully observed data. Hence, standard literature on copulas can be manipulated without any censoring related constraints. Given estimators $\widehat{G}_C(\cdot|\boldsymbol{x})$, $\widehat{F}_Y^{u}$ and $\widehat{F}_j^{u}$ of $G_C(\cdot|\boldsymbol{x})$, F_Y^{u} and F_j^{u}, $j = 1, \ldots, d$, satisfying certain high-level conditions which will be given in Section 4, this suggests to estimate the quantile regression in the presence of censoring by the empirical analogue of (8), that is

$$\widehat{m}_\tau(\boldsymbol{x}) = \arg\min_a \sum_{i=1}^{n}\left[\rho_\tau(Y_i - a)\,\widehat{W}_i(\boldsymbol{x})\,\widehat{c}_{Y\boldsymbol{X}}^{u}\left(\widehat{F}_Y^{u}(Y_i), \widehat{\boldsymbol{F}}^{u}(\boldsymbol{x})\right)\right], \qquad (9)$$

where $\widehat{W}_i(\boldsymbol{x}) = \Delta_i/(1 - \widehat{G}_C(Y_i - |\boldsymbol{x}))$, and where $\widehat{c}_{Y\boldsymbol{X}}^{u}$ denotes an estimator of $c_{Y\boldsymbol{X}}^{u}$ based on the four-step procedure described in Section 2.3. Explicitly,

$$\widehat{c}_{Y\boldsymbol{X}}^{u}\left(\widehat{F}_Y^{u}(y), \widehat{\boldsymbol{F}}^{u}(\boldsymbol{x})\right) = \widehat{c}_{YX_1}^{u}\left(\widehat{F}_Y^{u}(y), \widehat{F}_1^{u}(x_1)\right) \times \ldots \times \widehat{c}_{YX_d}^{u}\left(\widehat{F}_Y^{u}(y), \widehat{F}_d^{u}(x_d)\right)$$
$$\times \widehat{c}_{X_1 \ldots X_d|Y}^{u}\left(\widehat{F}_{1|Y}^{u}(x_1|y), \ldots, \widehat{F}_{d|Y}^{u}(x_d|y)\right),$$

where any two-dimensional kernel density estimator may be used for each bivariate copula density $\widehat{c}_{YX_j}^{u}$, $j = 1, \ldots, d$, such that condition (C7) of Section 4 holds, and where $\widehat{c}_{X_1 \ldots X_d|Y}^{u}$ is estimated by standard parametric vine procedures.

The resulting quantile regression estimator in (9) may then be viewed as a simple weighted quantile of the observed response variable, and is therefore easy to implement in practice using the efficient quantile regression code developed by Portnoy and Koenker [38] and Koenker [22]. Nonetheless, in the context of multivariate covariates, the estimation of $G_C(\cdot|\boldsymbol{x})$ requires further assumptions to overcome dimension related issues. Popular choices in the literature include, among others, independence between C and $\boldsymbol{X}$, the Cox model or the single-index model on $C|\boldsymbol{X} = \boldsymbol{x}$. Illustrations of such assumptions are treated in our simulation study.

As an interesting property, and similarly to the case without censoring, we note that the obtained regression function estimator is automatically monotonic

across quantile levels. Applying analogous arguments to the ones adopted in the proof of Theorem 2.5 of Koenker [22], one can indeed determine that

$$(\tau_2 - \tau_1)(\widehat{m}_{\tau_2}(\boldsymbol{x}) - \widehat{m}_{\tau_1}(\boldsymbol{x})) \sum_{i=1}^{n} \widehat{W}_i(\boldsymbol{x}) \, \widehat{c}^u_{Y\boldsymbol{X}} \big(\widehat{F}^u_Y(Y_i), \widehat{\boldsymbol{F}}^u(\boldsymbol{x}) \big) \geq 0. \qquad (10)$$

Given that $\widehat{W}_i(\boldsymbol{x}) \, \widehat{c}^u_{Y\boldsymbol{X}} \big(\widehat{F}^u_Y(Y_i), \widehat{\boldsymbol{F}}^u(\boldsymbol{x}) \big) \geq 0$ for all $i = 1, \ldots, n$, this signifies that $\widehat{m}_{\tau_2}(\boldsymbol{x}) \geq \widehat{m}_{\tau_1}(\boldsymbol{x})$ for $\tau_2 \geq \tau_1$.

Conclusively, in parallel to what has been stated for the uncensored case, the resulting estimator $\widehat{m}_\tau(\boldsymbol{x})$ defines a rich class of estimators built on the many different existing methods available in the literature for estimating copula densities and marginal distributions of both complete and censored data.

4. Asymptotic properties

We establish in this section the asymptotic normality of the proposed estimator $\widehat{m}_\tau(\boldsymbol{x})$. To that end, we first report the set of regularity conditions as well as the required high-level conditions on all estimators involved in the expression of $\widehat{m}_\tau(\boldsymbol{x})$. We then develop an asymptotic representation of our estimator for a general d-variate covariate. As the latter will result in a somewhat unpleasant expression for the asymptotic bias and variance for a general multivariate covariate, and given that the analytical reasoning is similar in spirit, we eventually restrict ourselves to the detailed asymptotic expression for the case $d = 2$.

For a fixed but arbitrary point of interest $\boldsymbol{x}$ in the support of $\boldsymbol{X}$, denoted by $\mathrm{supp}(\boldsymbol{X})$, let us suppose that there exists a neighborhood $\mathcal{V}_{\boldsymbol{F}^u(\boldsymbol{x})}$ of $\boldsymbol{F}^u(\boldsymbol{x})$ such that the following regularity conditions hold:

(C1) The conditional distribution $F_{T|\boldsymbol{X}}$ of T given $\boldsymbol{X}$ admits a conditional density $f_{T|\boldsymbol{X}}$ that is continuous, strictly positive and bounded uniformly on $\mathbb{R} \times \mathrm{supp}(\boldsymbol{X})$.

(C2) The point of interest $\boldsymbol{x}$ is such that $\boldsymbol{F}^u(\boldsymbol{x}) \in (0,1)^d$ and $\mathbb{P}(\Delta = 1|\boldsymbol{x}) > 0$. Furthermore, $0 < c^u_{\boldsymbol{X}}(\boldsymbol{F}^u(\boldsymbol{x})) < \infty$, $\sup_{t \in \mathbb{R}} c^u_{Y\boldsymbol{X}}(F^u_Y(t), \boldsymbol{F}^u(\boldsymbol{x})) < \infty$ and $\inf_{t \in \mathbb{R}} c^u_{Y\boldsymbol{X}}(F^u_Y(t), \boldsymbol{F}^u(\boldsymbol{x})) > 0$.

(C3) The point $m_\tau(\boldsymbol{x}) \in \mathbb{R}$ satisfies $G_C(m_\tau(\boldsymbol{x}) + \delta|\boldsymbol{x}) < 1$, for some $\delta > 0$.

(C4) Denote $\epsilon \equiv \epsilon(\boldsymbol{x}, \tau) = Y - m_\tau(\boldsymbol{x})$ and define $\psi_\tau(u) = \tau - I(u \leq 0)$. Then,

   (i) $\mathbb{E}(|\psi_\tau(\epsilon)| \, W(\boldsymbol{x})) < \infty$.

   (ii) $\mathbb{E}\left[\psi_\tau(\epsilon) \, W(\boldsymbol{x}) \, c^u_{Y\boldsymbol{X}}\big(F^u_Y(Y), \boldsymbol{F}^u(\boldsymbol{x})\big)\right]^2 < \infty$.

Concerning the high-level conditions, it is assumed that the multivariate copula density $c^u_{Y\boldsymbol{X}}$ is estimated using the proposed four-step strategy of Section 2.3, and that, for simplicity, the d bivariate kernel copula estimators of step (2) are based on the same bandwidth $\boldsymbol{H} = h^2 \boldsymbol{I}$ for a certain $h > 0$. The following conditions are then assumed to hold:

(C5) The marginal c.d.f. estimators are such that:

(i) $\sup_{t \in \mathbb{R}} \left| \widehat{F}_Y^u(t) - F_Y^u(t) \right| = O_{\mathbb{P}}(n^{-1/2})$.

(ii) $\widehat{\boldsymbol{F}}^u(\boldsymbol{x}) - \boldsymbol{F}^u(\boldsymbol{x}) = O_{\mathbb{P}}(n^{-1/2})$, where $\widehat{\boldsymbol{F}}^u(\boldsymbol{x}) = (\widehat{F}_1^u(x_1), \ldots, \widehat{F}_d^u(x_d))^{\mathsf{T}}$ and $\widehat{F}_j^u$ is an estimator of F_j^u.

(C6) $\sup_{t \le \tau_{F_Y}} |\widehat{G}_C(t|\boldsymbol{x}) - G_C(t|\boldsymbol{x})| = o_{\mathbb{P}}((nh^2)^{-1/2})$, and $\tau_{F_Y} < \tau_{G_C}$, where $\tau_{F_Y} = \inf\{t : F_Y(t) = 1\}$ and $\tau_{G_C} = \inf\{t : G_C(t) = 1\}$.

(C7) The multivariate copula estimator is such that:

(i) $\sup_{t \in \mathbb{R}} \left| \widehat{c}_{YX_j}^u (F_Y^u(t), F_j^u(x_j)) - c_{YX_j}^u (F_Y^u(t), F_j^u(x_j)) \right| = o_{\mathbb{P}}(1)$, $j = 1, \ldots, d$, where x_j is the j-th coordinate of $\boldsymbol{x}$.

(ii) $\sup_{u_0 \in (0,1)} \sup_{\boldsymbol{u} \in \mathcal{V}_{\boldsymbol{F}^u(\boldsymbol{x})}} |\partial_j \widehat{c}_{Y\boldsymbol{X}}^u(u_0, \boldsymbol{u})| = O_{\mathbb{P}}(1)$, $j = 1, \ldots, d+1$, where ∂_j denotes the partial derivative with respect to the j-th argument.

Assumption (C1) is standard in the context of quantile regression estimation. As for condition (C2), this is similar to assumption (C3)-(i) in Noh, El Ghouch and Van Keilegom [35] for the simplified case with no censoring, with an additional requirement on the conditional censoring probability that is resulting from the initial transformation of synthetic observations. Assumption (C3) is likewise emanating from the handling of censoring through these observations, and is rather usual in survival analysis. Note that, in the quantile regression framework, the latter assumption amounts to defining a natural upper bound for the quantile of interest that can be studied. Assumption (C4) reports a set of technical conditions to be met.

As regards conditions (C5)-(C7), (C5) is routinely made in the copula framework. For instance, it is readily satisfied for the empirical distributions when only uncensored observations are taken into account, and their rescaled versions which are prominent in the copula literature. Assumption (C6) imposes restrictions on the estimator one may consider for the conditional distribution of the censoring variable and is, for instance, fulfilled for a simple Kaplan-Meier estimator for G_C (see e.g. Theorem 2.1 in Chen and Lo [6] for sufficient and necessary conditions for (C6)). Lastly, the uniform consistency of the kernel density estimator required by assumption (C7) is, for instance, alluded to in Geenens, Charpentier and Paindaveine [14] for the probit-transformed copula estimator.

We now state the main result of this section that holds for a general d-dimensional covariate vector and for all bivariate kernel copula estimators based on the same bandwidth h. In practice, however, it may be recommended to adopt an unconstrained and non-diagonal bandwidth matrix, as is detailed in Section 4 of Geenens, Charpentier and Paindaveine. Nevertheless, when considering this general situation, the theoretical results become less tractable while equivalent in nature to the simplified situation considered here.

Theorem 4.1. *Let $h \equiv h_n \to 0$ be the common bandwidth of the d bivariate kernel copula density estimators. For h satisfying $nh^2 \to \infty$ as $n \to \infty$, and*

under assumptions (C1)-(C7), we have

$$\left(nh^2\right)^{1/2}\left(\hat{m}_\tau(\boldsymbol{x}) - m_\tau(\boldsymbol{x})\right) = \frac{w(\boldsymbol{x})}{f_{T|\boldsymbol{X}}(m_\tau(\boldsymbol{x})|\boldsymbol{x})}\frac{\left(nh^2\right)^{1/2}}{n}\sum_{i=1}^{n}\psi_\tau(\epsilon_i)W_i(\boldsymbol{x})$$
$$\times\left[\widehat{c}^u_{Y\boldsymbol{X}}\left(F^u_Y(Y_i),\boldsymbol{F}^u(\boldsymbol{x})\right) - c^u_{Y\boldsymbol{X}}\left(F^u_Y(Y_i),\boldsymbol{F}^u(\boldsymbol{x})\right)\right]$$
$$+ o_{\mathbb{P}}(1),$$

where $f_{T|\boldsymbol{X}}$ is the conditional density of T given $\boldsymbol{X}$ and $w(\boldsymbol{x}) = p(\boldsymbol{x})/\left[\mathbb{P}(\Delta = 1)c^u_{\boldsymbol{X}}(\boldsymbol{F}^u(\boldsymbol{x}))\right]$.

Theorem 4.1 implies, quite naturally, that the asymptotic behavior of $\hat{m}_\tau(\boldsymbol{x})$ will be characterized by the properties of the copula estimator, specifically through its nonparametric feature, provided that the estimation of $\widehat{G}_C(\cdot|\boldsymbol{x})$ is 'reasonable' when confronted to a multidimensional covariate vector (assumption (C6)). In particular, this suggests that the detailed discussion of Geenens, Charpentier and Paindaveine about the asymptotic bias and variance of their distinctive bivariate copula estimators may be transcribed in our context.

Additionally, Theorem 4.1 also covers an asymptotic representation of the copula-based quantile regression estimator when all responses are fully observed. In this situation, one would indeed obtain a similar result for the proposed semiparametric procedure, with the removal of all censoring related terms, that is $w(\boldsymbol{x})$, $W_i(\boldsymbol{x})$, $i = 1, \ldots, n$, and the superfluous conditioning on $\Delta = 1$ for the copula densities and marginal distributions. This results in the following:

$$\left(nh^2\right)^{1/2}\left(\hat{m}_\tau(\boldsymbol{x}) - m_\tau(\boldsymbol{x})\right) = \frac{1}{f_{T|\boldsymbol{X}}(m_\tau(\boldsymbol{x})|\boldsymbol{x})}\frac{\left(nh^2\right)^{1/2}}{n}\sum_{i=1}^{n}\psi_\tau(\epsilon_i)$$
$$\times\left[\widehat{c}_{Y\boldsymbol{X}}\left(F_Y(Y_i),\boldsymbol{F}(\boldsymbol{x})\right) - c_{Y\boldsymbol{X}}\left(F_Y(Y_i),\boldsymbol{F}(\boldsymbol{x})\right)\right]$$
$$+ o_{\mathbb{P}}(1).$$

We now consider a detailed asymptotic representation of our estimator for the simplified case where $d = 2$, and for a general nonparametric estimator of the bivariate copula densities. For convenience, we use the notation c^u_k as a shorthand for $c^u_{YX_k}$, and similarly for other functions depending on (Y, Δ, X_k), $k = 1, 2$.

Corollary 4.2. *Suppose that the assumptions of Theorem 4.1 hold for the case $d = 2$. Furthermore, suppose that the bivariate nonparametric copula estimators of c^u_1 and c^u_2 are such that*

$$\left(nh^2\right)^{1/2}\left(\widehat{c}^u_k(u_0, u_k) - c^u_k(u_0, u_k) - h^2 b_k(u_0, u_k)\right) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}Z^{nj}_k(u_0, u_k) + o_{\mathbb{P}}(1),$$

$$\forall u_k \in (0, 1), \text{ uniformly in } u_0 \in (0, 1), \text{ for } k = 1, 2,$$

for some some deterministic function $b_k(u_0, u_k)$, and for some function $Z_k^{nj}(u_0, u_k)$ depending on (Y_j, Δ_j, X_{kj}) and possibly on n, satisfying $\mathbb{E}\left(Z_k^{nj}(u_0, u_k)\right) = 0$, for all $u_0, u_k \in (0, 1)$.
Define

$$\widetilde{Z}^{ni}(u_0, \boldsymbol{u}) = \left[Z_1^{nj}(u_0, u_1)c_2^u(u_0, u_2) + Z_2^{nj}(u_0, u_2)c_1^u(u_0, u_1)\right]$$
$$\times c_{X_1 X_2 | Y}^u(u_1, u_2 | u_0),$$
$$b_{Y\boldsymbol{X}}(u_0, \boldsymbol{u}) = \left[b_1(u_0, u_1)c_2^u(u_0, u_2) + b_2(u_0, u_2)c_1^u(u_0, u_1)\right]$$
$$\times c_{X_1 X_2 | Y}^u(u_1, u_2 | u_0),$$
$$\lambda_n(Y_i, \Delta_i, \boldsymbol{X}_i, \boldsymbol{x}) = \mathbb{E}\left[\psi_\tau(\epsilon)W(\boldsymbol{x})\widetilde{Z}^{ni}(F^u(Y), \boldsymbol{F}^u(\boldsymbol{x})) | Y_i, \Delta_i, \boldsymbol{X}_i\right], i = 1, \ldots, n.$$

Suppose furthermore that the following technical conditions hold:

(C8) $\mathbb{E}\left[\psi_\tau(\epsilon)W(\boldsymbol{x})b_{Y\boldsymbol{X}}(F^u(Y), \boldsymbol{F}^u(\boldsymbol{x}))\right]^2 < \infty.$
(C9) $\mathbb{E}\left[\psi_\tau(\epsilon_i)W_i(\boldsymbol{x})\widetilde{Z}^{nj}(F^u(Y_i), \boldsymbol{F}^u(\boldsymbol{x}))\right]^2 = o(n)$ *for all* $i, j = 1, \ldots, n$, *where the expectation is taken with respect to* (Y_i, Δ_i) *and* $(Y_j, \Delta_j, \boldsymbol{X}_j)$.

Then, the copula-based quantile regression estimator at any point of interest $\boldsymbol{x}$ satisfying (C1)-(C9) is such that

$$\left(nh^2\right)^{1/2}\left(\widehat{m}_\tau(\boldsymbol{x}) - m_\tau(\boldsymbol{x}) - h^2 B(\boldsymbol{x})\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \sigma^2(\boldsymbol{x})\right),$$

where

$$B(\boldsymbol{x}) = \frac{w(\boldsymbol{x})}{f_{T|\boldsymbol{X}}(m_\tau(\boldsymbol{x})|\boldsymbol{x})}\mathbb{E}\left(\psi_\tau(\epsilon)W(\boldsymbol{x})b_{Y\boldsymbol{X}}\left(F^u(Y), \boldsymbol{F}^u(\boldsymbol{x})\right)\right)$$
$$and \ \sigma^2(\boldsymbol{x}) = \frac{w^2(\boldsymbol{x})}{f_{T|\boldsymbol{X}}^2(m_\tau(\boldsymbol{x})|\boldsymbol{x})} \lim_{n\to\infty} \mathbb{E}\left(\lambda_n(Y, \Delta, \boldsymbol{X}, \boldsymbol{x})^2\right).$$

Corollary 4.2 reports the asymptotic normality of our estimator at the expected convergence rate, implied by the nonparametric estimation of the bivariate copula densities. Depending on the choice in step (2) of the kernel density estimator fulfilling the conditions of Corollary 4.2, simple plug-in of the expression of Z_k^{nj}, $k = 1, 2$, in all quantities built upon the latter may then lead to the detailed, although arduous, expressions of the asymptotic bias and variance of the proposed estimator.

Furthermore, as this had yet to be covered, it is worth stressing out that Corollary 4.2 also encompasses the asymptotic normality of the suggested estimator based on semiparametric vine copulas with strictly complete data. Similarly to what has been stated for Theorem 4.1, one is indeed only required to withdraw all censoring related elements from Corollary 4.2 to obtain the expressions of the asymptotic bias and variance of the proposed semiparametric quantile regression estimator for complete observations.

5. Simulation study

In this section, we assess the practical finite-sample performance of the proposed methodology by means of Monte Carlo simulations. For this purpose, we first present a brief numerical study to further motivate the semiparametric copula strategy we intend to adopt for multivariate problems for both complete and censored observations. Secondly, we assess the numerical performance of the copula-based regression estimator for complete observations using the developed copula estimation strategy and compare it with established semiparametric and nonparametric techniques in the literature. Lastly, focusing on survival data, we illustrate the performance of our estimator in (9) by also showing promising results with respect to competitors in the domain, including when the generated scenario is to the advantage of the latter. All the simulations are carried out using the statistical computing environment R (R Core Team [44]) and its freely accessible packages.

5.1. Assessing the semiparametric copula estimation

This first section aims at numerically illustrating the choice of our semiparametric copula estimation strategy. For this purpose, we will consider here both complete and censored responses as this will result in an interesting analysis of our proposed strategy resulting from the fact that the censoring percentage will have an influence on the number of observations actually entering the copula estimation process. Overall, we consider two distinctive data generating processes (DGP) and compare our methodology with fully parametric and nonparametric procedures one might consider for the estimation of a multivariate copula density. For the general simulation settings, we consider $B = 500$ repetitions of each DGP; three (average) levels of censoring (0%, 30% and 50%), three sample sizes ($n \in \{100, 200, 400\}$) and the quantile level of interest $\tau = 0.3$. As the object of interest here is the copula modelling, when censoring is introduced, we only consider the simple case of independence between the censoring variable and the covariate vector in order to keep the estimation of $\widehat{G}_C$ needed for (9) uncomplicated, that is, using the Kaplan-Meier estimator. The detailed DGPs are as follows:

- **DGP A**: $(F_T(T), F_1(X_1), F_2(X_2)) \sim$ Gaussian copula with parameters $(\rho_{T1}, \rho_{T2}, \rho_{12}) = (0.3, 0.9, 0.5)$. Given standard uniform marginal distributions for all three variables, the resulting true quantile regression may be calculated as $m_\tau(\boldsymbol{x}) = \Phi\big(-0.2\Phi^{-1}(x_1) + \Phi^{-1}(x_2) + 0.4\Phi^{-1}(\tau)\big)$ (see Noh, El Ghouch and Van Keilegom [35]). To include censoring, we introduce the variable $C \sim U[0, M]$, where the parameter M is computed in order to obtain the desired average censoring proportion ($M = 5/3$ for 30% and $M = 1$ for 50%).
- **DGP B**: $(F_T(T), F_1(X_1), F_2(X_2), F_3(X_3)) \sim$ Gaussian copula with parameters $(\rho_{T1}, \rho_{T2}, \rho_{T3}, \ \rho_{12}, \rho_{13}, \rho_{23}) = (0.3, 0.9, 0.7, 0.5, 0.25, 0.5)$. The

resulting true quantile regression for standard uniform marginal distributions is determined as $m_\tau(\boldsymbol{x}) = \Phi\big(-0.2\Phi^{-1}(x_1) + 0.83\Phi^{-1}(x_2) + 0.33\Phi^{-1}(x_2) + 0.27\Phi^{-1}(\tau)\big)$. The censoring variable is $C \sim U[0, M]$ ($M = 5/3$ for 30% and $M = 1$ for 50% censoring).

For any general copula-based regression estimator, the marginal distribution estimations are performed, as suggested in Section 2.3, using rescaled versions of the empirical distributions:

$$\widehat{F}_Y^u(y) = \frac{1}{n^u + 1} \sum_{i=1}^n \Delta_i \mathbb{1}(Y_i \le y)$$

$$\widehat{F}_j^u(x_j) = \frac{1}{n^u + 1} \sum_{i=1}^n \Delta_i \mathbb{1}(X_{ij} \le x_j), j = 1, \dots, d,$$

where $n^u = \sum_{i=1}^n \Delta_i$ is the number of uncensored observations.

For the distinctive copula-based estimators, we consider the following procedures:

$\widehat{m}_{cop,\tau}^{SP}$: semiparametric estimation strategy detailed in Section 2.3. That is, we first estimate the d bivariate copulas of interest employing the probit transformation technique of Geenens, Charpentier and Paindaveine [14] coupled with local likelihood estimation based on quadratic polynomials. To that end, we follow their proposed nearest-neighbor bandwidth selection procedure. Concerning the estimation of the d-dimensional 'noisy' copula density (5), as mentioned above, we apply standard vine techniques using the R package `VineCopula`. Specifically, we adopt one automatically selected tree structure for the simplified decomposition of the copula density among many R-vine candidate structures (see Dißmann et al. [9]), and subsequently determine the appropriate pair-copula family to be selected and parametrically estimated. The selection criterion for bivariate copulas is chosen to be the Akaike information criterion (AIC), which revealed to be adequate in the R-vine context (see Brechmann [1], chap. 5), and ten potential family candidates, together with their rotations, are considered: eight of them are Archimedian (Clayton, Gumbel, Frank, Joe, Clayton-Gumbel, Joe-Gumbel, Joe-Clayton and Joe-Frank), and the last two are elliptical (Gaussian and Student t).

$\widehat{m}_{cop,\tau}^{NP}$: fully nonparametric estimation of the d-dimensional copula using vine techniques. Specifically, while the vine structure is kept identical, here all bivariate building blocks are estimated using the local likelihood technique based on probit-projected data with the bandwidth selection procedure of Geenens, Charpentier and Paindaveine (as is studied in Nagler and Czado [32]). Given its fully nonparametric nature, it should be mentioned that this estimator is not covered by the theoretical results of Section 4.

TABLE 1

Simulation results expressed in terms of $IMSE \times 1000$ for the estimation of $m_\tau(x)$ in DGP
A *and* **B**. *The number of repetitions operated is $B = 500$ for sample sizes*
$n \in \{100, 200, 400\}$, *average censoring proportions $p_c \in \{0, 0.3, 0.5\}$ and quantile level*
$\tau = 0.3$.

		DGP **A**					DGP **B**		
n	p_c	$\widehat{m}^P_{cop,\tau}$	$\widehat{m}^{SP}_{cop,\tau}$	$\widehat{m}^{NP}_{cop,\tau}$	n	p_c	$\widehat{m}^P_{cop,\tau}$	$\widehat{m}^{SP}_{cop,\tau}$	$\widehat{m}^{NP}_{cop,\tau}$
	0	1.442	1.486	2.369		0	1.192	1.416	3.307
100	0.3	2.203	2.173	3.167	100	0.3	2.036	2.459	3.972
	0.5	3.537	3.548	4.135		0.5	5.055	6.552	7.038
	0	0.689	0.737	1.462		0	0.560	0.658	2.609
200	0.3	0.990	1.016	1.905	200	0.3	0.850	0.945	2.857
	0.5	1.887	1.664	2.581		0.5	1.686	2.108	3.647
	0	0.337	0.371	0.987		0	0.259	0.354	2.140
400	0.3	0.503	0.525	1.243	400	0.3	0.403	0.487	2.266
	0.5	0.915	0.863	1.600		0.5	0.738	0.916	2.497

$\widehat{m}^P_{cop,\tau}$: fully parametric estimation of the d-dimensional copula density, where all bivariate copulas are estimated using the previously mentioned candidate families and selection criteria. However, unlike the above-mentioned estimators, we do not force here any structure for the vine decomposition. As a consequence, no explicit distinction is imposed between dependence of interest and noisy dependence. Instead, one data-driven selected structure is adopted, regardless of the arguments of Section 2.3. This will allow us to analyse the impact of such dependence distinction in our regression context, as is discussed below. Finally, as it is the case for $\widehat{m}^{NP}_{cop,\tau}$, this estimator is not covered by the asymptotic theory of Section 4.

Both DGPs concentrate on the situation where the dependence structure between the response variable and the covariate vector is characterized by a parametric copula. In such circumstances, $\widehat{m}^P_{cop,\tau}$ will have a critical advantage and may serve in order to evaluate the impact of the nonparametric part of the estimation scheme, especially when the dimension of the covariate vector increases. As a performance criterion, we consider here the empirical integrated mean squared error (IMSE), defined as

$$IMSE(\widehat{m}_\tau(\boldsymbol{x})) = \frac{1}{N} \sum_{i=1}^{N} \left(\frac{1}{B} \sum_{b=1}^{B} \left(\widehat{m}^{(b)}_\tau(\boldsymbol{x}_i) - m_\tau(\boldsymbol{x}_i) \right)^2 \right),$$

where $\{\boldsymbol{x}_i, i = 1, \ldots, N\}$ is a generated random sample of size $N = 10$ serving as an evaluation set spread on the domain of $\boldsymbol{X}$, and $\widehat{m}^{(b)}_\tau(\cdot)$ denotes the regression estimation for the b-th simulated sample.

The results of our simulation study are summarized in Table 1. Based on these, we detail our analysis in two parts, as the outcomes of our study offer relevant information on both the copula decomposition choice and the type of bivariate estimators one may adopt in the multivariate setting. Note that, for both DGPs, as the dependence structure is specified by a Gaussian copula, the

simplifying assumption intrinsic to the vine decomposition is here applicable (see Theorem 4 in Stöber, Joe and Czado [43]). In our context, this means that any observed difference between copula strategies is not to be attributed to a possible violation of the underlying simplifying assumption.

Focusing first on our decomposition strategy, we note as expected that $\widehat{m}_{cop,\tau}^{P}$ globally outperforms $\widehat{m}_{cop,\tau}^{SP}$ and $\widehat{m}_{cop,\tau}^{NP}$. However, strikingly enough, this is not observed for DGP **A** for different censoring proportions where $\widehat{m}_{cop,\tau}^{SP}$ details better results. This is interpreted here as evidence for the validity of our arguments regarding the decomposition choice: as the censoring proportion grows, the number of observations actually entering the copula estimation becomes more moderate, hereby implying two opposite effects in this context. First, the propagation of estimation approximations tends to be more important, signifying that the further we decompose, the more sensitive becomes the estimation of the involved bivariate copulas as these are tributary of the quality of previously estimated bivariate blocks. Using a purely data-driven decomposition may then result in a poor fit of the (conditional) copula of the response variable with one of the covariates, as it is not required that the latter would be primarily treated. This is interpreted as the reason why $\widehat{m}_{cop,\tau}^{SP}$ is able to outperform $\widehat{m}_{cop,\tau}^{P}$, admittedly by a small amount, when censoring increases for a fixed sample size $n \in \{200, 400\}$, even though the simulated scenario is issued from a purely parametric copula. However, on the other hand, when observations are more scarce, it is well-known that nonparametric estimations become more sensitive than parametric counterparts. This explains why the estimation results for $\widehat{m}_{cop,\tau}^{SP}$ are not superior to those of $\widehat{m}_{cop,\tau}^{P}$ for $n = 100$ with 50% censoring, as the former requires the nonparametric estimation of two bivariate copulas, whose complexity compared to $\widehat{m}_{cop,\tau}^{P}$ seems to override the positive effects of our decomposition choice. Overall, these noteworthy results for DGP **A** illustrate the effectiveness of our proposed copula decomposition in the regression context. When augmenting the covariate vector dimension, the price of estimating now three nonparametric bivariate copulas quite logically exceeds the potential gain of concentrating efforts on the dependence of interest. This is identified in DGP **B**.

Concentrating now on the modelling choice for the noisy dependence, the comparison between $\widehat{m}_{cop,\tau}^{SP}$ and $\widehat{m}_{cop,\tau}^{NP}$ offers valuable information, particularly in DGP **A**, as the only distinctness here is the estimation of a unique bivariate copula density $c_{X_1 X_2 | Y}^{u}$. Visibly, the implication of keeping a nonparametric approach for this part seems to be rather severe. This finding clearly also applies when the dimension of the covariate grows.

Conclusively, the recommended copula modelling strategy seems to propose an adequate trade-off between preventing the serious effects of a purely nonparametric estimation and providing the flexibility needed to overcome possible shortcomings associated to a purely parametric alternative.

5.2. Comparison with other estimation methods for complete responses

In this section, we briefly assess the global performance of the semiparametric copula-based regression estimator for complete observations with several established methodologies in the literature. To that end, we consider the following competing methodologies:

$\widehat{m}_{LL,\tau}$: Locally polynomial quantile regression estimator studied in Chaudhuri [4]. In this paper, we consider a local linear estimator which is implemented with Gaussian kernels in the function `lpqr` of the R package `quantreg`. The bandwidth is selected via leave-one-out cross validation in a set of candidates ranging from 0.05 to 2 by 0.05 increments.

$\widehat{m}_{inv,\tau}$: Nonparametric regression estimator studied in Li, Lin and Racine [28] where the idea is to numerically invert a kernel estimated conditional distribution function. This procedure is implemented in the R package `np` via the function `npreg`. The bandwidth selection is automatically implemented in the function `npcdistbw` of the same library.

$\widehat{m}_{si,\tau}$: Single-index regression estimator based on a two stage estimation method as proposed by Zhu, Huang and Li [48]. The univariate non-parametric part of the methodology is estimated with a Gaussian kernel and the bandwidth is selected by leave-one-out cross validation in a set of candidates ranging from 0.05 to 0.5 by 0.05 increments.

$\widehat{m}_{cop,\tau}^{P}$: Semiparametric copula quantile regression estimator of Noh, El Ghouch and Van Keilegom [35] where the copula density is estimated with purely parametric vine techniques as in Section 5.1.

$\widehat{m}_{cop,\tau}$: Proposed semiparametric copula quantile regression estimator with the same implementation as in Section 5.1.

These methodologies are compared in the following two DGPs:

- **DGP C**: $(F_T(T), F_1(X_1), F_2(X_2)) \sim$ Gaussian copula with parameters $(\rho_{T1}, \rho_{T2}, \rho_{12}) = (0.9, 0.8, -0.8)$. The marginal distributions are chosen to be standard exponential for T and standard normal for the covariates. The resulting quantile regression may be determined as $m_\tau(\boldsymbol{x}) = F_T^{-1}\left[\Phi\left(\beta^\mathsf{T}\boldsymbol{x} + 0.41\Phi^{-1}(\tau)\right)\right]$, where $\beta = (-13/18, 2/9)^\mathsf{T}$.
- **DGP D**: Data simulated from the model $T = |2X_1 - X_2^2 + 0.5|^{0.5} + 0.1X_4^3 X_5(0.5|X_3|+1)+0.1\epsilon$, where $\epsilon \sim \mathcal{N}(0,1)$ and the five-dimensional co-variate $(X_1, \ldots, X_5)$ is simulated from a Gaussian distribution with correlations $(\rho_{12}, \rho_{13}, \ldots, \rho_{45}) = (0.3, 0.4, 0.5, 0.6, 0.7, 0.3, 0.4, 0.5, 0.6, 0.7)$ and standard marginals.

The first DGP is a low-dimensional setting where we expect the single-index and both copula-based estimators to perform well given that the true quantile regression is in fact a single-index model and that the data is simulated from a known copula. In opposition, the second DGP explores a higher dimension of the covariates and considers a setting with non-monotonic dependencies between the response and the covariates, for which no parametric copula in the literature

TABLE 2

Simulation results expressed in terms of $IMSE \times 10$ for the estimation of $m_\tau(\boldsymbol{x})$ in DGPs
C *and* **D** *with respectively two and five covariates. The number of repetitions operated is*
$B = 500$ *for sample sizes $n \in \{100, 200, 400\}$ and with quantiles of interest $\tau \in \{0.3, 0.5, 0.7\}$.*

DGP	n	τ	$\widehat{m}_{LL,\tau}$	$\widehat{m}_{inv,\tau}$	$\widehat{m}_{si,\tau}$	$\widehat{m}_{cop,\tau}^{P}$	$\widehat{m}_{cop,\tau}$
		0.3	0.121	0.122	0.104	0.096	0.090
	100	0.5	0.152	0.157	0.126	0.126	0.121
C		0.7	0.212	0.241	0.193	0.180	0.176
		0.3	0.070	0.068	0.058	0.041	0.043
	200	0.5	0.090	0.083	0.073	0.053	0.057
		0.7	0.128	0.128	0.108	0.074	0.082
		0.3	1.507	1.599	6.227	3.207	1.264
	100	0.5	1.482	1.506	2.587	2.074	0.983
D		0.7	1.887	1.676	2.624	2.261	1.198
		0.3	0.822	0.674	2.020	3.110	0.852
	400	0.5	0.726	0.626	1.624	1.890	0.631
		0.7	1.187	0.765	2.170	1.986	0.830

is suited as illustrated in Section 2. Hence, in this setting $\widehat{m}_{cop,\tau}^{P}$ is expected to perform poorly and both $\widehat{m}_{cop,\tau}$ and $\widehat{m}_{si,\tau}$ are a priori no longer favored above the other methodologies.

To compare the performance of the different procedures, we consider 500 repetitions of each DGP, quantile levels $\tau \in \{0.3, 0.5, 0.7\}$ and sample sizes $n \in \{100, 200, 400\}$. Similarly to the previous section, we then consider the IMSE as a performance criterion calculated on a testing sample of 20 evaluation points $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{20}\}$.

The simulation results of this brief comparative study are presented in Table 2. As expected, for the two-dimensional setting we observe the copula-based and the single-index estimators to perform better than the remaining methodologies as the latter do not take profit of any underlying model assumption here. Furthermore, both copula-based estimators tend to outperform in this setting the single-index estimator for all considered sample sizes and quantile levels. Similarly to Section 5.1, we also note that $\widehat{m}_{cop,\tau}$ is here again able to outperform by a slight margin $\widehat{m}_{cop,\tau}^{P}$ for small sample sizes by focusing primarily on the estimation of the dependence of interest. Overall, this suggests that an appropriate semiparametric estimation of the copula density results in a performance of the regression estimator relatively similar to the estimator of Noh, El Ghouch and Van Keilegom when the latter is considered in its most advantageous situation.

Next, the results of DGP **D** numerically illustrate the gain in flexibility of the proposed methodology over $\widehat{m}_{cop,\tau}^{P}$. In this situation, the latter evidently misspecifies the regression model as can be observed from the fact that the IMSE does not tend to diminish much with an increasing sample size. In opposition, $\widehat{m}_{cop,\tau}$ still performs very competitively with respect to the literature, especially for smaller samples. This suggests here again that the copula estimation strategy proposed in this paper offers the required flexibility to a copula-based approach for quantile regression. To conclude, this results in a very competitive estimation procedure with respect to the established semiparametric and nonparametric literature.

5.3. Comparison with other estimation methods for censored responses

The objective of this last section is to provide a comparison study between the proposed copula-based methodology and existing competitors for survival data. On a general note, given that for multidimensional covariates an appropriate estimation of the conditional distribution $G_C(\cdot|\boldsymbol{x})$ for the weights $W(\boldsymbol{x})$ in (9) may be of crucial influence, we consider distinctive scenarios contrasting in the impact on $\widehat{G}_C(\cdot|\boldsymbol{x})$ in order to provide a sufficiently broad view on the performance of our methodology.

We examine three general simulation models, with $B = 500$ repetitions of each; two (average) levels of censoring (30% and 50%), two sample sizes ($n = 200$ and $n = 400$) and four values for the quantile level of interest ($\tau \in \{0.1, 0.3, 0.5, 0.7\}$). Specifically, we consider general data arising from a Cox regression model. Based on the hazard function (defined as $h(t) = f_T(t)/(1 - F_T(t))$), this prominent model for analysing survival times specifies that $h(t|\boldsymbol{X}_i) = h_0(t) \exp(\beta^{\mathsf{T}}\boldsymbol{X}_i), i = 1, \ldots, n$, where $h_0(t)$ is the so-called baseline hazard function. In this instance, given a nonnegative time-to-event variable, simple algebraic manipulations show that the general conditional quantile regression can be written as $m_\tau(\boldsymbol{x}) = H_0^{-1}\big(-\log(1 - \tau)\exp(-\beta^{\mathsf{T}}\boldsymbol{x})\big)$, where $H_0(t) \equiv \int_0^t h_0(s)\mathrm{d}s$. For every proposed setting, the baseline distribution is set to be standard exponential and consequently, for a given vector $\boldsymbol{x}$ the $\tau-$th conditional quantile function is given by

$$m_\tau(\boldsymbol{x}) = -\log(1 - \tau)\exp(-\beta^{\mathsf{T}}\boldsymbol{x}).$$

The distinction between our simulation scenarios is to be found in the three covariate vectors and the dimension of the latter that are taken into account. The first two settings are chosen to be of moderate dimension with $d = 5$ while the third setting exhibits an example with $d = 8$. Focusing first on the case $d = 5$, we consider $\beta = (1, -3/4, 1/2, 1/4, -3/5)^{\mathsf{T}}$ and 5 covariates $(X_1, \ldots, X_5)$ simulated form a Gaussian copula with parameters $(\rho_{12}, \rho_{13}, \ldots, \rho_{45}) = (0.3, 0.4, 0.5, 0.6, 0.7, 0.3, 0.4, 0.5, 0.6, 0.7)$. To distinguish our scenarios when $d = 5$, we consider two covariate vectors given by $\boldsymbol{X}^{(1)} = (X_1, X_2, \ldots, X_5)$ and $\boldsymbol{X}^{(2)} = (X_1, \exp(X_2), X_3, X_4, X_5)$, for which the resulting quantile regressions are both single-index models in $\boldsymbol{X}^{(1)}$ and $\boldsymbol{X}^{(2)}$, respectively. This will allow to highlight the performance of our procedure when its competing estimators are both in an ideal situation and a slightly altered version of it.

Lastly, for $d = 8$ we consider again a Gaussian copula for the initial covariate vector with parameters $\rho_{ij} = (-0.8)^{|i-j|}, i, j = 1, \ldots, 8$. The parameter vector of the Cox regression is $\beta = (1, -3/4, 1/2, 1/4, -3/5, 4/5, 3/5, -2/5)^{\mathsf{T}}$ and the covariate vector is $\boldsymbol{X}^{(3)} = (X_1, \exp(X_2), X_3, \ldots, X_7, X_8^2)$. The resulting model is hence a single-index and a Cox regression with respect to $\boldsymbol{X}^{(3)}$.

5.3.1. Model 1: Cox model and single-index model with $d = 5$

We first consider the simple case of data following a Cox regression model issued from covariate vector $\boldsymbol{X}^{(1)}$. The distributions and parameter values of the

censoring variables define the following scenarios:

- **DGP E**: $C \sim \mathrm{Exp}(\lambda_C)$, with λ_C independent of $\boldsymbol{X}^{(1)}$. To attain the desired average censoring proportions, we fix $\lambda_C = 0.464$ and $\lambda_C = 1.083$ (corresponding to approximately 30% and 50% censoring in average, respectively). The exact conditional censoring probability given $\boldsymbol{X}^{(1)} = \boldsymbol{x}$ is calculated as $\lambda_C/(\lambda_C + \exp(\beta^{\mathsf{T}}\boldsymbol{x}))$ and is hence a decreasing function of $\beta^{\mathsf{T}}\boldsymbol{x}$, making us conjecture better results for higher values of $\beta^{\mathsf{T}}\boldsymbol{x}$. Note that in this scenario, $G_C(\cdot|\boldsymbol{x})$ boils down to $G_C(\cdot)$, for which the Kaplan-Meier estimator is suited.
- **DGP F**: $C \sim \mathrm{Exp}(\lambda_C(\boldsymbol{x}))$, with $\lambda_C(\boldsymbol{x}) = 3/7 \times \exp(\beta^{\mathsf{T}}\boldsymbol{x})$ and $\lambda_C(\boldsymbol{x}) = \exp(\beta^{\mathsf{T}}\boldsymbol{x})$ corresponding to 30% and 50% censoring, respectively. In this scenario, the conditional censoring probability is independent of $\boldsymbol{x}$, and an adequate estimation of $G_C(\cdot|\boldsymbol{x})$ is fulfilled with a Cox model hypothesis for the relationship between the covariates and the censoring variable.

For comparison purposes, we consider the four following estimators:

$\widehat{m}_{cox,\tau}$: parametric estimator exploiting the information related to the parametric Cox model setting. This estimator will serve as a reference for the ideal, yet unknown in practice, situation. Specifically, $\widehat{m}_{cox,\tau}(\boldsymbol{x}) = -\log(1-\tau)\exp(-\widehat{\beta}^{\mathsf{T}}\boldsymbol{x})/\widehat{\lambda}_{T_0}$, where $\widehat{\beta}$ is estimated by maximum partial likelihood, and $\widehat{\lambda}_{T_0}$ is the maximum likelihood estimator of the exponential baseline distribution.

$\widehat{m}_{si,\tau}$: Single-index regression estimator studied in Bücher, El Ghouch and Van Keilegom [2], where the censoring distribution is supposed to be independent of $\boldsymbol{x}$. For the univariate nonparametric part of the estimation process, 10 different bandwidths are selected ($h \in \{0.05, 0.1, \ldots, 0.5\}$), and the optimal choice is performed using the described leave-one-out cross-validation procedure. Note that the quantile regression of interest $m_\tau(\boldsymbol{x})$ here is indeed a single-index model in $(X_1, X_2, \ldots, X_5)$. This should provide a critical advantage to the performance of $\widehat{m}_{si,\tau}$.

$\widehat{m}_{cop,\tau}^{(\perp\!\!\!\perp)}$: our copula-based estimator from estimation equation (9), where the conditional distribution $G_C(\cdot|\boldsymbol{x})$ is supposed to be independent of $\boldsymbol{x}$ and thereby estimated by the classical (unconditional) Kaplan-Meier estimator.

$\widehat{m}_{cop,\tau}^{(cox)}$: our copula-based estimator, where the relationship between the covariates and the censoring variable is supposed to follow a Cox regression model. Namely, $\widehat{G}_C(c|\boldsymbol{x})$ is estimated by $1 - \exp\left(-c\exp(\widehat{\beta}_C^{\mathsf{T}}\boldsymbol{x})/\widehat{\lambda}_{C_0}\right)$.

Following the arguments of Section 2.3 and the results of the previous sections, both copula-based procedures are implemented employing a semiparametric estimation for the d-variate copula built on the aforementioned candidate families and selection criterions.

Simulation results expressed in terms of IMAE and the dispersion measure in brackets for the estimation of $m_\tau(\boldsymbol{x})$ with 30% of censoring where the true model is a Cox and single-index model. The number of repetitions operated is $B = 500$ for sample sizes $n \in \{200, 400\}$ and with quantiles of interest $\tau \in \{0.1, 0.3, 0.5, 0.7\}$.

			Censoring = 30%			
DGP	n	τ	$\widehat{m}_{cox,\tau}$	$\widehat{m}_{si,\tau}$	$\widehat{m}_{cop,\tau}^{(\perp)}$	$\widehat{m}_{cop,\tau}^{(cox)}$
E	200	0.1	0.016 (0.019)	0.039 (0.053)	0.043 (0.059)	0.043 (0.057)
		0.3	0.054 (0.063)	0.083 (0.103)	0.101 (0.124)	0.102 (0.123)
		0.5	0.105 (0.123)	0.149 (0.185)	0.174 (0.199)	0.175 (0.194)
		0.7	0.183 (0.213)	0.280 (0.373)	0.273 (0.299)	0.279 (0.305)
	400	0.1	0.011 (0.013)	0.030 (0.038)	0.033 (0.044)	0.033 (0.043)
		0.3	0.036 (0.043)	0.070 (0.078)	0.079 (0.097)	0.081 (0.098)
		0.5	0.069 (0.084)	0.115 (0.140)	0.130 (0.160)	0.133 (0.159)
		0.7	0.120 (0.146)	0.221 (0.273)	0.212 (0.251)	0.212 (0.244)
F	200	0.1	0.016 (0.019)	0.037 (0.052)	0.043 (0.059)	0.043 (0.059)
		0.3	0.055 (0.064)	0.082 (0.105)	0.102 (0.129)	0.103 (0.124)
		0.5	0.108 (0.125)	0.152 (0.190)	0.174 (0.197)	0.174 (0.191)
		0.7	0.187 (0.218)	0.301 (0.474)	0.283 (0.310)	0.278 (0.298)
	400	0.1	0.011 (0.014)	0.029 (0.037)	0.035 (0.044)	0.034 (0.045)
		0.3	0.036 (0.046)	0.066 (0.077)	0.082 (0.102)	0.083 (0.098)
		0.5	0.071 (0.089)	0.118 (0.141)	0.138 (0.165)	0.135 (0.159)
		0.7	0.123 (0.155)	0.233 (0.287)	0.222 (0.259)	0.218 (0.241)

In order to compare the studied estimators' performance, we consider in this section an integrated version of the median absolute estimation error, that is

$$IMAE(\widehat{m}_\tau(\boldsymbol{x})) = \frac{1}{N} \sum_{i=1}^{N} \mathrm{med}^{(B)}\big(|\widehat{m}_\tau(\boldsymbol{x}_i) - m_\tau(\boldsymbol{x}_i)|\big),$$

where $\widehat{m}_\tau$ is a generic estimator of m_τ, $\{\boldsymbol{x}_i, i = 1, \ldots, N\}$ is an evaluation set corresponding to a generated random sample of size $N = 10$, spread on the domain of $\boldsymbol{X}^{(1)}$, and $\mathrm{med}^{(B)}$ denotes the median taken over all $B = 500$ simulations. The choice for this robust L_1-type of measure is motivated by the fact that the optimization routines involved in the single-index procedure may yield very unlikely results with a small probability, hereby strongly disadvantaging the estimator when considering a L_2-type of error measure. The same reasoning is underlying the determination of a robust dispersion measure on estimation errors, taken as the averaged interquartile range. More precisely, we choose $N^{-1} \sum_{i=1}^{N} \Big(Q_3^{(B)}\big(|\widehat{m}_\tau(\boldsymbol{x}_i) - m_\tau(\boldsymbol{x}_i)|\big) - Q_1^{(B)}\big(|\widehat{m}_\tau(\boldsymbol{x}_i) - m_\tau(\boldsymbol{x}_i)|\big)\Big)$, where $Q_3^{(B)}$ and $Q_1^{(B)}$ stand, respectively, for the third and first quartiles taken over all simulations.

The results of our simulation study for this model are reported in Tables 3 and 4 for 30% and 50% of censoring, respectively. As expected, the Cox regression estimator $\widehat{m}_{cox,\tau}$, serving as a reference case here, outperforms the other estimators for both scenarios and every sample size, quantile level and censoring percentage. More interestingly, while the single-index estimator $\widehat{m}_{si,\tau}$ quite logically displays better overall results than our copula-based estimators, it is worth noticing that the difference seems to fade away when moving to higher quantile

Table 4

Simulation results expressed in terms of IMAE and the dispersion measure in brackets for the estimation of $m_\tau(\boldsymbol{x})$ with 50% of censoring where the true model is a Cox and single-index model. The number of repetitions operated is $B = 500$ for sample sizes $n \in \{200, 400\}$ and with quantiles of interest $\tau \in \{0.1, 0.3, 0.5, 0.7\}$.

			Censoring = 50%			
DGP	n	τ	$\widehat{m}_{cox,\tau}$	$\widehat{m}_{si,\tau}$	$\widehat{m}_{cop,\tau}^{(\perp)}$	$\widehat{m}_{cop,\tau}^{(cox)}$
E	200	0.1	0.019 (0.022)	0.040 (0.056)	0.044 (0.056)	0.044 (0.054)
		0.3	0.063 (0.074)	0.089 (0.111)	0.113 (0.127)	0.115 (0.127)
		0.5	0.123 (0.144)	0.176 (0.242)	0.201 (0.209)	0.205 (0.213)
		0.7	0.213 (0.251)	0.411 (0.402)	0.354 (0.326)	0.362 (0.326)
	400	0.1	0.012 (0.015)	0.031 (0.040)	0.035 (0.042)	0.035 (0.042)
		0.3	0.042 (0.051)	0.071 (0.082)	0.089 (0.106)	0.089 (0.106)
		0.5	0.081 (0.099)	0.136 (0.170)	0.161 (0.179)	0.163 (0.178)
		0.7	0.140 (0.171)	0.299 (0.315)	0.284 (0.290)	0.281 (0.279)
F	200	0.1	0.019 (0.022)	0.040 (0.061)	0.046 (0.059)	0.044 (0.053)
		0.3	0.064 (0.074)	0.090 (0.106)	0.117 (0.129)	0.115 (0.123)
		0.5	0.125 (0.143)	0.202 (0.374)	0.212 (0.218)	0.210 (0.206)
		0.7	0.217 (0.249)	0.526 (0.854)	0.380 (0.341)	0.380 (0.324)
	400	0.1	0.013 (0.015)	0.029 (0.038)	0.038 (0.048)	0.036 (0.045)
		0.3	0.043 (0.052)	0.069 (0.082)	0.099 (0.113)	0.095 (0.105)
		0.5	0.084 (0.102)	0.140 (0.173)	0.177 (0.201)	0.172 (0.174)
		0.7	0.145 (0.176)	0.346 (0.632)	0.306 (0.322)	0.298 (0.283)

levels, especially for higher levels of censoring percentage. This is considered as a first encouraging result for the copula-based estimators as the simulated model is very strongly to the advantage of $\widehat{m}_{si,\tau}$. This indicates that our proposed procedure is flexible enough to compete with $\widehat{m}_{si,\tau}$ in its ideal setup when the number of observations actually entering the estimation scheme becomes moderate, that is when censoring percentage is important and the quantile level of interest is high. Of course, if this is not the case, there is no a priori reason to believe that the copula-based estimators could outperform $\widehat{m}_{si,\tau}$ when the latter is considered in its optimal setting.

Figure 3 serves to illustrate these results for the particular case of DGP **E** with $n = 400$ and for high quantile levels ($\tau \in \{0.5, 0.7\}$). For the sake of brevity, we only report here a graphical comparison between the copula-based estimators and the single-index estimator, as $\widehat{m}_{cox,\tau}$ clearly outperforms the latter. As can be observed from Figure 3, the difference between $\widehat{m}_{si,\tau}$ and both copula-based estimators is graphically modest, especially for the highest quantile of interest $\tau = 0.7$, hereby reinforcing the previously discussed results of Tables 3 and 4. Furthermore, note that all estimators tend to present worse performances for low levels of $\beta^\mathsf{T}\boldsymbol{x}$. This is expected in this simulation setting as these levels correspond to the region of $\beta^\mathsf{T}\boldsymbol{x}$ where the exact conditional censoring probability is the highest.

Focusing again on Tables 3 and 4, we further note that, while there is logically a difference between the performances of $\widehat{m}_{cop,\tau}^{(\perp)}$ and $\widehat{m}_{cop,\tau}^{(cox)}$ depending on the simulated scenario, the effect of an appropriate modelling of $\widehat{G}_C(\cdot|\boldsymbol{x})$ seems to be rather limited in this simulation setup for our estimators. Additionally, while both $\widehat{m}_{si,\tau}$ and $\widehat{m}_{cop,\tau}^{(\perp)}$ rely here on the assumption of in-
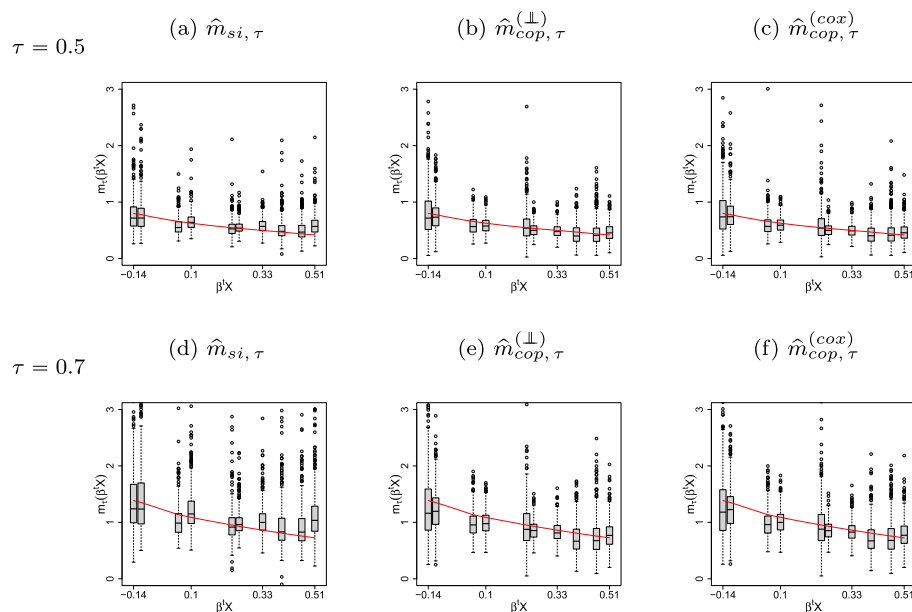
(a) $\widehat{m}_{si,\tau}$ (b) $\widehat{m}_{cop,\tau}^{(\perp)}$ (c) $\widehat{m}_{cop,\tau}^{(cox)}$

$\tau = 0.5$

(d) $\widehat{m}_{si,\tau}$ (e) $\widehat{m}_{cop,\tau}^{(\perp)}$ (f) $\widehat{m}_{cop,\tau}^{(cox)}$

$\tau = 0.7$

FIG 3. *Boxplots corresponding to the results in Tables 3 and 4 of three estimators considered for $m_\tau(\boldsymbol{x})$ for $N = 10$ points spread on the domain of $\beta^{\mathsf{T}}\boldsymbol{X}^{(1)}$ in DGP $\boldsymbol{E}$, and for $\tau \in \{0.5, 0.7\}$. In this setting, the true model is a single-index model and hence favors $\widehat{m}_{si,\tau}$ over both copula-based estimators. All pictures are based on 500 repetitions for $n = 400$ and 30% censoring. The red lines represent the true value of $m_{0.5}(\boldsymbol{x})$ and $m_{0.7}(\boldsymbol{x})$.*

dependence between the censoring variable and the covariate vector, the latter seems to numerically behave better when confronted to the violation of this hypothesis. This can be observed from the comparison between DGP **E** and **F** of the dispersion measures of both estimators, once again especially for high levels of quantile values and censoring percentages. Of course, one could argue that, given the results for the dispersion measure of $\widehat{m}_{si,\tau}$ for high quantiles, this can partly be due to a poor smoothing parameter choice. However, the latter was implemented using the proposed methodology of Bücher, El Ghouch and Van Keilegom [2], just as a practitioner would have resolved to act.

5.3.2. Models 2 & 3: Alteration of a Cox model and single-index model with $d = 5$ and $d = 8$

To pursue our simulation study, we now consider in this section the second and third models where the time-to-event data is simulated using covariate vectors $\boldsymbol{X}^{(2)}$ and $\boldsymbol{X}^{(3)}$, respectively. Consequently, the resulting quantile regressions are no longer a single-index nor a Cox regression model in $(X_1, X_2, \ldots, X_5)$ for model 2 and $(X_1, X_2, \ldots, X_8)$ for model 3. For comparison purposes, we

TABLE 5

Simulation results expressed in terms of $IMAE$ and the dispersion measure in brackets for the estimation of $m_\tau(\boldsymbol{x})$ for model 2. The number of repetitions operated is $B = 500$ for sample sizes $n \in \{200, 400\}$, censoring proportion $p_c \in \{0.3, 0.5\}$ and with four levels of quantile of interest $\tau \in \{0.1, 0.3, 0.5, 0.7\}$.

Model 2

n	p_c	τ	$\widehat{m}_{cox,\tau}$	$\widehat{m}_{si,\tau}$	$\widehat{m}_{cop,\tau}^{(\perp)}$	$\widehat{m}_{cop,\tau}^{(cox)}$
200	0.3	0.1	0.115 (0.046)	0.105 (0.142)	0.115 (0.160)	0.115 (0.159)
		0.3	0.391 (0.156)	0.279 (0.273)	0.249 (0.285)	0.250 (0.286)
		0.5	0.760 (0.302)	0.489 (0.538)	0.452 (0.532)	0.454 (0.538)
		0.7	1.319 (0.525)	0.988 (1.023)	0.781 (0.828)	0.783 (0.842)
	0.5	0.1	0.116 (0.054)	0.107 (0.156)	0.115 (0.139)	0.116 (0.143)
		0.3	0.391 (0.182)	0.334 (0.324)	0.303 (0.335)	0.311 (0.338)
		0.5	0.760 (0.359)	0.611 (0.742)	0.574 (0.552)	0.583 (0.558)
		0.7	1.321 (0.615)	1.441 (1.142)	1.053 (0.845)	1.074 (0.865)
400	0.3	0.1	0.115 (0.029)	0.084 (0.101)	0.092 (0.118)	0.092 (0.120)
		0.3	0.388 (0.099)	0.219 (0.222)	0.211 (0.264)	0.212 (0.265)
		0.5	0.753 (0.192)	0.372 (0.401)	0.358 (0.427)	0.359 (0.430)
		0.7	1.308 (0.333)	0.685 (0.844)	0.598 (0.682)	0.604 (0.683)
	0.5	0.1	0.114 (0.035)	0.084 (0.105)	0.096 (0.109)	0.096 (0.112)
		0.3	0.387 (0.118)	0.254 (0.231)	0.221 (0.283)	0.226 (0.285)
		0.5	0.751 (0.228)	0.479 (0.490)	0.431 (0.486)	0.438 (0.488)
		0.7	1.305 (0.397)	0.933 (0.878)	0.850 (0.781)	0.857 (0.793)

consider the four estimation procedures described in model 1, given covariate vectors $(X_1, X_2, \ldots, X_5)$ and $(X_1, X_2, \ldots, X_5)$, respectively, hereby implying a slight model bias for both $\widehat{m}_{si,\tau}$ and $\widehat{m}_{cox,\tau}$ as the latter would require transformations of the covariates. In opposition, both copula-based procedures, which are here constructed using the same previously-described semiparametric modelling strategy, automatically take account of such concerns.

For brevity, we only consider here the situation where the censoring variable is independent from the covariate vector, as the impact of a dependent scheme has been treated in model 1. Specifically, for model 2 we simulate the censoring variable from an exponential distribution with parameter values 0.208 and 0.486 for approximately 30% and 50% censoring. Concerning model 3, we only report the results for 30% censoring which are attained by simulating here a censoring variable following an exponential distribution with parameter 0.903.

The resulting performance of the considered estimators are depicted in Tables 5 and 6, once again in terms of $IMAE$. In Table 5 reporting the case $d = 5$, we observe that both copula-based estimators tend to outperform their competitors, with the exception of very low quantiles of interest where the effect of a ('small') misspecification of the underlying model seems to be moderate for both $\widehat{m}_{cox,\tau}$ and $\widehat{m}_{si,\tau}$. As we move away from these low quantile levels, the consequences of misspecifying the model become more severe for the competing estimators, notably for a purely parametric approach ($\widehat{m}_{cox,\tau}$). In contrast, the copula-based approach presents satisfactory results for varying censoring proportions and sample sizes, especially when keeping in mind that the simulated scenario is 'only' a slightly altered version of the ideal scenario for its competitors.

TABLE 6

Simulation results expressed in terms of IMAE and the dispersion measure in brackets for the estimation of $m_\tau(\boldsymbol{x})$ for model 3. The number of repetitions operated is $B = 500$ for sample sizes $n \in \{200, 400\}$, censoring proportion $p_c = 0.3$ and with four levels of quantile of interest $\tau \in \{0.1, 0.3, 0.5, 0.7\}$.

Model 3

p_c	n	τ	$\widehat{m}_{cox,\tau}$	$\widehat{m}_{si,\tau}$	$\widehat{m}_{cop,\tau}^{(\perp)}$	$\widehat{m}_{cop,\tau}^{(cox)}$
		0.1	0.031 (0.054)	0.037 (0.049)	0.029 (0.037)	0.031 (0.044)
	200	0.3	0.084 (0.182)	0.124 (0.174)	0.075 (0.080)	0.078 (0.091)
		0.5	0.149 (0.354)	0.241 (0.451)	0.139 (0.139)	0.142 (0.151)
0.3		0.7	0.273 (0.616)	0.419 (0.822)	0.240 (0.222)	0.247 (0.236)
		0.1	0.116 (0.035)	0.028 (0.038)	0.023 (0.028)	0.024 (0.030)
	400	0.3	0.067 (0.118)	0.094 (0.099)	0.062 (0.066)	0.062 (0.069)
		0.5	0.104 (0.230)	0.183 (0.344)	0.114 (0.116)	0.115 (0.122)
		0.7	0.220 (0.400)	0.317 (0.652)	0.194 (0.190)	0.194 (0.197)

Similar findings apply when analyzing the results of Table 6 for $d = 8$. In this situation, both copula-based estimator still globally outperform their competitors, although the margin of improvement is relatively smaller than for $d = 5$. This results from the fact that we have here to model a nine-dimensional copula density with no information on the underlying model, while both $\widehat{m}_{si,\tau}$ and $\widehat{m}_{si,\tau}$ rely on model assumptions that are still relatively close to the simulated model. Conclusively, this advocates, here again, for the flexibility of our procedure and its withstanding to misspecification of the underlying model for multidimensional problems when comparing with other semiparametric or fully parametric modelling techniques.

6. Real data application

We present in this section a brief application of our procedure for censored data by analysing the Colorado Plateau uranium miners cohort data (see e.g. Lubin et al. [29], Langholz and Goldstein [26]). The object of the study, for which 3347 Caucasian male miners having worked at least a month in the uranium mines of the Colorado Plateau were followed, is to investigate the risk of lung cancer related to smoking and radon exposure. Hence, the event of interest is defined as the time till lung cancer death (expressed as the logarithm of number of years), which affected a total of 258 miners. Besides failure time, the study also includes information about age at entry to the study, cumulative smoking (in number of packs) and radon exposure (in working level month (WLM)).

As the original data set is prone to heavy censoring (92.3%), and given the illustrative nature of this section, we first define a subsample on which the analysis will be performed. To that end, in order to preserve as best as possible the nature of the population at risk, we decide to define a threshold on the radon exposure above which observations will enter the subsample. The value of the threshold is practically chosen as a trade-off between censoring proportion and actual number of observations that are to be selected. Specifically, we find that by defining a threshold of 2831 WLM on radon exposure, a subsample of 176
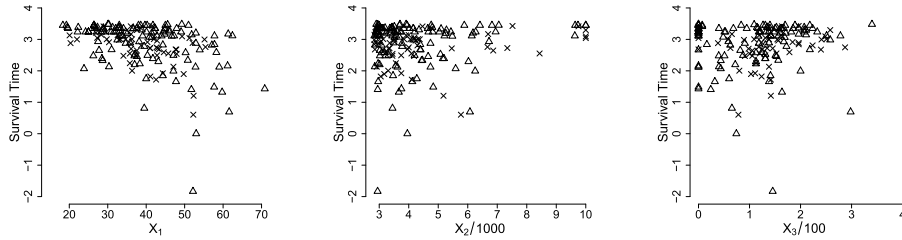
FIG 4. *Colorado Plateau uranium miners cohort data. Scatter plots of the survival time versus each covariate. Uncensored data points are given by ×, while censored observations are represented by △.*

observations is constituted, 55 of which were subject to the event of interest. Scatterplots of the selected data are represented in Figure 4, where X_1 is the age at entry into the study, X_2 is the cumulative radon exposure and X_3 is the cumulative smoking.

Regarding the data analysis, endorsing the role of a practitioner, we are faced with the choice of an appropriate estimator for the application of quantile regression in this context. We therefore consider the following distinctive candidates:

$\widehat{m}_{cop,\tau}^{(cox)}$: Copula-based quantile regression estimator, with Cox regression modelling for $G_C(\cdot|\cdot)$.

$\widehat{m}_{si,\tau}$: Single-Index methodology of Bücher, El Ghouch and Van Keilegom [2].

$\widehat{m}_{cox,\tau}$: Semiparametric estimator based on the Cox proportional hazards model. Specifically, $\widehat{m}_{cox,\tau}(\boldsymbol{x}) = \widehat{H}_0^{-1}(-\log(1-\tau)\exp(-\widehat{\beta}^\mathsf{T}\boldsymbol{x}))$, where $\widehat{\beta}$ is estimated by maximum partial likelihood, and $\widehat{H}_0$ is the Nelson-Aalen-type estimator of the cumulative baseline hazard.

As a general evaluation measure to compare models for the present data set, we consider the median quantile loss from predicting the τ-th conditional quantile of T for the uncensored observations. In other words, we use the following cross-validated prediction error criterion:

$$\mathrm{PE}\left(\widehat{m}_\tau\right) = \underset{\substack{1 \le i \le n \\ \Delta_i = 1}}{\mathrm{med}}\, \rho_\tau\left(Y_i - \widehat{m}_\tau^{-i}(\boldsymbol{X}_i)\right),$$

where $\widehat{m}_\tau^{-i}$ denotes any estimator of m_τ based on all observations except the i-th one.

For the implementation of the copula-based regression estimator, we adopt the methodology of Section 2.3 and our simulation study by opting for a semiparametric modelling of the four-variate copula density, where the bivariate copulas of the response variable with each covariate are estimated using the procedure of Geenens, Charpentier and Paindaveine [14] with quadratic polynomials along with the proposed data-driven bandwidth selection scheme. The remaining trivariate noisy copula is, afterwards, estimated using vine techniques with the same candidate families and selection criterion as in Section 5. Additionally, a general appropriate hypothesis has to be made for the modelling of

Colorado Plateau uranium miners cohort data. Prediction error multiplied by 10 for each censored quantile regression estimator for quantile levels $\tau \in \{0.1, 0.3, 0.5, 0.7\}$.

τ	PE $\left(\widehat{m}_{cop,\tau}^{(cox)}\right)$	PE $(\widehat{m}_{si,\tau})$	PE $(\widehat{m}_{cox,\tau})$
0.1	0.348	0.848	0.946
0.3	0.774	1.471	1.180
0.5	0.785	1.901	1.169
0.7	0.470	2.838	0.967

the conditional distribution $G_C(\cdot|\cdot)$. As it is shown in the literature that independence is unsuitable for the data of the study (see e.g. Leng and Tong [27]), we propose to model the conditional distribution through a semiparametric Cox regression for the censoring time with respect to the covariates. Therefore, as is common in practice, the parameter of the regression is estimated by maximum partial likelihood while the Breslow estimator is used for computing the baseline survivor function.

Concerning $\widehat{m}_{si,\tau}$, the bandwidth choice is performed using the proposed leave-one-out cross-validation procedure on normalized covariates with 15 candidates ($h \in \{0.1, 0.2, \ldots, 1.5\}$). Note that, as the latter procedure is computationally costly, it is here assumed that the selected bandwidth is constant for all datasets entering the estimation of $\text{PE}(\widehat{m}_{si,\tau})$, which is of little impact in this context given that only one observation at the time is to be removed from each dataset.

The results of the evaluation measure are reported in Table 7 for four quantile levels of interest $\tau \in \{0.1, 0.3, 0.5, 0.7\}$. It is observed that, in terms of cross-validated prediction error, our copula-based approach depicts quite confidently the best performance for every considered quantile level of interest. By comparison, the single-index structure seems to be inappropriate as such for the studied data set, and should possibly require further work and attention on, for instance, transformations of covariates. Furthermore, its sensitivity to higher quantile levels is, here again, highlighted. Lastly, despite being part of every practitioner's toolbox for survival analysis, the proportional hazards regression estimator is also relatively largely outperformed by the increased flexibility of the copula-based approach. Hence, this example clearly illustrates the ability of our estimator to adapt to the underlying regression structure of the data.

Conclusively, recalling that, from equation (10), our procedure enjoys the additional valuable property of being automatically monotonic across quantile levels, this real data application plainly highlights the relevance of our estimator for flexible analyses of multivariate censored data.

7. Conclusion

In this work we have proposed a semiparametric copula-based quantile regression estimator in the context of potentially right-censored responses. On a general note, for data with or without censoring, and motivated by the regression con-

text, a novel semiparametric estimation approach for the implied copula density was studied. Furthermore, in parallel to the procedure of Noh, El Ghouch and Van Keilegom [35], the proposed regression estimator in this work is obtained as a weighted quantile of the observed response variable, hereby opening the door to the practical use of the quantile regression code developed by Portnoy and Koenker [38] and Koenker [22]. Asymptotic normality of the resulting estimator for both complete and censored data was obtained with convergence rate determined by the nonparametric estimation of bivariate copula densities. For inferential purposes however, bootstrap procedures are to be preferred as the applicability of the obtained asymptotic normality is restrained by numerous unknown quantities to be estimated in practice. This bootstrap procedure could be inspired by the wild bootstrap procedure proposed by Feng, He and Hu [13]. Finally, supporting the objective of this work, an extensive simulation study and a real data application have been carried out to illustrate both the validity of the semiparametric copula estimation and the increased flexibility of our procedure in comparison with existing alternatives, especially when no a priori knowledge about the functional form of the quantile regression is available in practice.

Appendix

We develop in this appendix the proofs of Theorem 4.1 and Corollary 4.2.

Proof of Theorem 4.1

Define

$$\widehat{A}_n(s) = \sum_{i=1}^{n} \left[\rho_\tau(\epsilon_i - s/a_n) - \rho_\tau(\epsilon_i) \right] \widehat{W}_i(\boldsymbol{x}) \, \widehat{c}_{Y\boldsymbol{X}}^u \left(\widehat{F}_Y^u(Y_i), \widehat{\boldsymbol{F}}^u(\boldsymbol{x}) \right),$$

with $\epsilon_i \equiv \epsilon_i(\boldsymbol{x}, \tau) = Y_i - m_\tau(\boldsymbol{x})$, and $a_n = \sqrt{nh^2}$. Observe first that, by definition of $\hat{m}_\tau(\boldsymbol{x})$,

$$a_n(\hat{m}_\tau(\boldsymbol{x}) - m_\tau(\boldsymbol{x})) = \arg\min_s \widehat{A}_n(s).$$

Furthermore, given that ρ_τ is a convex function and $\widehat{W}_i(\boldsymbol{x}) \, \widehat{c}_{Y\boldsymbol{X}}^u \left(\widehat{F}_Y^u(Y_i), \widehat{\boldsymbol{F}}^u(\boldsymbol{x}) \right) \geq 0$ for all $i = 1, \ldots, n$, we have that $s \mapsto \widehat{A}_n(s)$ is convex. The idea is then to develop an expression of $\widehat{A}_n(s)$ leading to the application of the quadratic approximation Lemma of convex functions (Basic Corollary in Hjort and Pollard [17]). To that end, we have that (Knight's (1998) identity)

$$\rho_\tau(u - v) - \rho_\tau(u) = -v\psi_\tau(u) + R(u, v),$$

with $\psi_\tau(u) = \tau - \mathbb{1}(u \leq 0)$,

$$R(u, v) = \int_0^v (\mathbb{1}(u \leq s) - \mathbb{1}(u \leq 0)) \mathrm{d}s = (u - v)(\mathbb{1}(u \leq 0) - \mathbb{1}(u \leq v)),$$

and $0 \le R(u, v) \le |v|$. Hence, we may write

$$\widehat{A}_n(s) = -s\widehat{A}_{1n} + \hat{A}_{2n}(s),$$

with

$$\widehat{A}_{1n} = a_n^{-1} \sum_{i=1}^n \psi_\tau(\epsilon_i) \, \widehat{W}_i(\boldsymbol{x}) \, \widehat{c}^u_{Y\boldsymbol{X}} \big(\widehat{F}^u_Y(Y_i), \widehat{\boldsymbol{F}}^u(\boldsymbol{x}) \big),$$

and

$$\widehat{A}_{2n}(s) = \sum_{i=1}^n R(\epsilon_i, s/a_n) \, \widehat{W}_i(\boldsymbol{x}) \, \widehat{c}^u_{Y\boldsymbol{X}} \big(\widehat{F}^u_Y(Y_i), \widehat{\boldsymbol{F}}^u(\boldsymbol{x}) \big).$$

Focusing first on $\widehat{A}_{2n}(s)$, we will show that

$$\widehat{A}_{2n}(s) = \frac{s^2}{2} w^{-1}(\boldsymbol{x}) f_{T|\boldsymbol{X}}(m_\tau(\boldsymbol{x})|\boldsymbol{x}) \frac{n}{a_n^2} + o_{\mathbb{P}}\left(\frac{n}{a_n^2}\right), \tag{A.11}$$

where $f_{T|\boldsymbol{X}}$ is the conditional density of T given $\boldsymbol{X}$ and $w(\boldsymbol{x}) = p(\boldsymbol{x})/\big[\mathbb{P}(\Delta = 1)c^u_{\boldsymbol{X}}(\boldsymbol{F}^u(\boldsymbol{x}))\big]$. To that end, we write

$$\widehat{A}_{2n}(s) = A_{2n}(s) + \Delta_{1n}(s) + \Delta_{2n}(s),$$

where

$$A_{2n}(s) = \sum_{i=1}^n R(\epsilon_i, s/a_n) \, W_i(\boldsymbol{x}) \, c^u_{Y\boldsymbol{X}} \big(F^u_Y(Y_i), \boldsymbol{F}^u(\boldsymbol{x}) \big)$$

$$\Delta_{1n}(s) = \sum_{i=1}^n R(\epsilon_i, s/a_n) \, (\widehat{W}_i(\boldsymbol{x}) - W_i(\boldsymbol{x})) \, c^u_{Y\boldsymbol{X}} \big(F^u_Y(Y_i), \boldsymbol{F}^u(\boldsymbol{x}) \big)$$

$$\Delta_{2n}(s) = \sum_{i=1}^n R(\epsilon_i, s/a_n) \, \widehat{W}_i(\boldsymbol{x}) \big[\widehat{c}^u_{Y\boldsymbol{X}} \big(\widehat{F}^u_Y(Y_i), \widehat{\boldsymbol{F}}^u(\boldsymbol{x}) \big) - c^u_{Y\boldsymbol{X}} \big(F^u_Y(Y_i), \boldsymbol{F}^u(\boldsymbol{x}) \big) \big].$$

Concentrating on each term, we will show that

$$A_{2n}(s) = \frac{s^2}{2} w^{-1}(\boldsymbol{x}) f_{T|\boldsymbol{X}}(m_\tau(\boldsymbol{x})|\boldsymbol{x}) \frac{n}{a_n^2} + o_{\mathbb{P}}\left(\frac{n}{a_n^2}\right) \tag{A.12}$$

$$\Delta_{1n}(s) = o_{\mathbb{P}}\left(\frac{n}{a_n^2}\right) \tag{A.13}$$

$$\Delta_{2n}(s) = o_{\mathbb{P}}\left(\frac{n}{a_n^2}\right). \tag{A.14}$$

For the proof of (A.12), we first establish that

$$\mathbb{E}(A_{2n}(s)) = n \, \mathbb{E}\big(R(\epsilon_1, s/a_n) \, W_1(\boldsymbol{x}) \, c^u_{Y\boldsymbol{X}} \big(F^u_Y(Y_1), \boldsymbol{F}^u(\boldsymbol{x}) \big) \big)$$

$$= n \, w^{-1}(\boldsymbol{x}) \int_0^{s/a_n} \big(F_{T|\boldsymbol{X}}(m_\tau(\boldsymbol{x}) + t|\boldsymbol{x}) - F_{T|\boldsymbol{X}}(m_\tau(\boldsymbol{x})|\boldsymbol{x}) \big) \mathrm{d}t,$$

where $F_{T|\boldsymbol{X}}$ denotes, as in Section 2, the conditional c.d.f. of T given $\boldsymbol{X}$, and where, for the second equality, we used (see Section 3) the fact that, for any measurable function $\varphi : \mathbb{R} \to \mathbb{R}$, we have $\mathbb{E}(\Delta\varphi(Y) c^u_{Y\boldsymbol{X}}(F^u_Y(Y), \boldsymbol{F}^u(\boldsymbol{x}))) = w^{-1}(\boldsymbol{x})\mathbb{E}(\Delta\varphi(Y)|\boldsymbol{X} = \boldsymbol{x})$. Next, we write

$$F_{T|\boldsymbol{X}}(m_\tau(\boldsymbol{x}) + t|\boldsymbol{x}) - F_{T|\boldsymbol{X}}(m_\tau(\boldsymbol{x})|\boldsymbol{x}) = t\, f_{T|\boldsymbol{X}}(m_\tau(\boldsymbol{x}) + \theta t|\boldsymbol{x})$$
$$= t\, f_{T|\boldsymbol{X}}(m_\tau(\boldsymbol{x})|\boldsymbol{x}) + t\, R(t),$$

for some $\theta \in (0, 1)$, with $R(t) = f_{T|\boldsymbol{X}}(m_\tau(\boldsymbol{x}) + \theta t|\boldsymbol{x}) - f_{T|\boldsymbol{X}}(m_\tau(\boldsymbol{x})|\boldsymbol{x})$. This yields

$$\mathbb{E}(A_{2n}(s)) = \frac{s^2}{2} w^{-1}(\boldsymbol{x}) f_{T|\boldsymbol{X}}(m_\tau(\boldsymbol{x})|\boldsymbol{x})\frac{n}{a_n^2} + n\, w^{-1}(\boldsymbol{x})\int_0^{s/a_n} t\, R(t)\mathrm{d}t.$$

From the fact that, under the required bandwidth condition and assumption (C1),

$$\left|\int_0^{s/a_n} t\, R(t)\mathrm{d}t\right| \leq \frac{s^2}{2\,a_n^2}\sup_{|t|\leq |s|/a_n}|R(t)| = o(1/a_n^2),$$

we get that

$$\mathbb{E}(A_{2n}(s)) = \frac{s^2}{2} w^{-1}(\boldsymbol{x}) f_{T|\boldsymbol{X}}(m_\tau(\boldsymbol{x})|\boldsymbol{x})\frac{n}{a_n^2} + o\left(\frac{n}{a_n^2}\right),$$

provided assumption (C2) is satisfied. To conclude the proof of (A.12), it is then sufficient to show that

$$\mathbb{V}\mathrm{ar}(A_{2n}(s)) = o\left(\frac{n}{a_n^2}\right)^2.$$

To that end, observe that, for n sufficiently large, and under assumptions (C1), (C2) and (C3),

$$\mathbb{V}\mathrm{ar}(A_{2n}(s))$$
$$\leq n\,\mathbb{E}\left(R(\epsilon_1, s/a_n)\, W_1(\boldsymbol{x})\, c^u_{Y\boldsymbol{X}}(F^u_Y(Y_1), \boldsymbol{F}^u(\boldsymbol{x}))\right)^2$$
$$\leq n\,\mathbb{E}\left(R(\epsilon_1, s/a_n)\, W_1(\boldsymbol{x})\, c^u_{Y\boldsymbol{X}}(F^u_Y(Y_1), \boldsymbol{F}^u(\boldsymbol{x}))\right)\frac{|s|}{a_n}$$
$$\times \left(1 - G_C(m_\tau(\boldsymbol{x}) + \delta)|\boldsymbol{x})\right)^{-1}\sup_{t\in\mathbb{R}} c^u_{Y\boldsymbol{X}}(F^u_Y(t), \boldsymbol{F}^u(\boldsymbol{x})), \quad \text{for some } \delta > 0$$
$$= O\left(\frac{n}{a_n^3}\right) = o\left(\frac{n}{a_n^2}\right)^2,$$

as a_n/n converges to 0.

Concentrating now on the proof of (A.13), observe that, for n sufficiently large

$$|\Delta_{1n}(s)|$$

$$= \left| \sum_{i=1}^{n} R(\epsilon_i, s/a_n) \, W_i(\boldsymbol{x}) \, c_{Y\boldsymbol{X}}^u \big(F_Y^u(Y_i), \boldsymbol{F}^u(\boldsymbol{x}) \big) \frac{\widehat{G}_C(Y_i - |\boldsymbol{x}) - G_C(Y_i - |\boldsymbol{x})}{1 - \widehat{G}_C(Y_i - |\boldsymbol{x})} \right|$$

$$\leq A_{2n}(s) \sup_{\substack{t \leq \max_{1 \leq i \leq n} Y_i \\ \Delta_i = 1}} \frac{|\widehat{G}_C(t|\boldsymbol{x}) - G_C(t|\boldsymbol{x})|}{1 - \widehat{G}_C(t|\boldsymbol{x})}$$

$$= o_{\mathbb{P}}\left(\frac{n}{a_n^2} \right),$$

provided that $\widehat{G}_C(\cdot|\boldsymbol{x})$ satisfies assumption (C6).

Lastly, for the proof of (A.14), note that, for n sufficiently large,

$$|\Delta_{2n}(s)| = \left| \sum_{i=1}^{n} R(\epsilon_i, s/a_n) \, \widehat{W}_i(\boldsymbol{x}) \, c_{Y\boldsymbol{X}}^u \big(F_Y^u(Y_i), \boldsymbol{F}^u(\boldsymbol{x}) \big) \right.$$

$$\left. \times \frac{\widehat{c}_{Y\boldsymbol{X}}^u \big(\widehat{F}_Y^u(Y_i), \widehat{\boldsymbol{F}}^u(\boldsymbol{x}) \big) - c_{Y\boldsymbol{X}}^u \big(F_Y^u(Y_i), \boldsymbol{F}^u(\boldsymbol{x}) \big)}{c_{Y\boldsymbol{X}}^u \big(F_Y^u(Y_i), \boldsymbol{F}^u(\boldsymbol{x}) \big)} \right|$$

$$\leq (A_{2n}(s) + \Delta_{1n}(s)) \sup_{t \in \mathbb{R}} \frac{\left| \widehat{c}_{Y\boldsymbol{X}}^u \big(\widehat{F}_Y^u(t), \widehat{\boldsymbol{F}}^u(\boldsymbol{x}) \big) - c_{Y\boldsymbol{X}}^u \big(F_Y^u(t), \boldsymbol{F}^u(\boldsymbol{x}) \big) \right|}{c_{Y\boldsymbol{X}}^u \big(F_Y^u(t), \boldsymbol{F}^u(\boldsymbol{x}) \big)}$$

$$= o_{\mathbb{P}}\left(\frac{n}{a_n^2} \right),$$

provided that assumption (C7)-(i) is satisfied.

Hence, reassembling (A.12), (A.13) and (A.14) yields

$$\frac{a_n^2}{n} \widehat{A}_n(s) = -s \frac{a_n^2}{n} \widehat{A}_{1n} + \frac{s^2}{2} w^{-1}(\boldsymbol{x}) f_{T|\boldsymbol{X}}(m_\tau(\boldsymbol{x})|\boldsymbol{x}) + o_{\mathbb{P}}(1).$$

Therefore, by the quadratic approximation Lemma of a convex function, see e.g. Hjort and Pollard [17], if $\frac{a_n^2}{n} \widehat{A}_{1n} = O_{\mathbb{P}}(1)$, then

$$a_n(\widehat{m}_\tau(\boldsymbol{x}) - m_\tau(\boldsymbol{x})) = \arg\min_s \frac{a_n^2}{n} \widehat{A}_n(s) = \frac{w(\boldsymbol{x})}{f_{T|\boldsymbol{X}}(m_\tau(\boldsymbol{x})|\boldsymbol{x})} \frac{a_n^2}{n} \widehat{A}_{1n} + o_{\mathbb{P}}(1).$$

$$\text{(A.15)}$$

As a consequence, the asymptotic behavior of our estimator will be driven by the asymptotic expression of $\frac{a_n^2}{n} \widehat{A}_{1n}$. Developing the expression of the latter, we will show that

$$\frac{a_n^2}{n} \widehat{A}_{1n} = \frac{a_n}{n} \sum_{i=1}^{n} \psi_\tau(\epsilon_i) W_i(\boldsymbol{x}) \big[\widehat{c}_{Y\boldsymbol{X}}^u \big(F_Y^u(Y_i), \boldsymbol{F}^u(\boldsymbol{x}) \big) - c_{Y\boldsymbol{X}}^u \big(F_Y^u(Y_i), \boldsymbol{F}^u(\boldsymbol{x}) \big) \big]$$

$$+ o_{\mathbb{P}}(1). \qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(A.16)}$$

To that end, note that $\frac{a_n^2}{n} \widehat{A}_{1n}$ may be decomposed as

$$\frac{a_n^2}{n}\widehat{A}_{1n} = \frac{a_n}{n}\sum_{i=1}^{n}\psi_\tau(\epsilon_i)W_i(\boldsymbol{x})\big[\widehat{c}^u_{Y\boldsymbol{X}}\big(F^u_Y(Y_i), \boldsymbol{F}^u(\boldsymbol{x})\big) - c^u_{Y\boldsymbol{X}}\big(F^u_Y(Y_i), \boldsymbol{F}^u(\boldsymbol{x})\big)\big]$$

$$+ \Delta_{3n} + \Delta_{4n} + \Delta_{5n} + \Delta_{6n},$$

where

$$\Delta_{3n} = \frac{a_n}{n}\sum_{i=1}^{n}\psi_\tau(\epsilon_i)(\widehat{W}_i(\boldsymbol{x}) - W_i(\boldsymbol{x}))\widehat{c}^u_{Y\boldsymbol{X}}\big(\widehat{F}^u_Y(Y_i), \widehat{\boldsymbol{F}}^u(\boldsymbol{x})\big) \tag{A.17}$$

$$\Delta_{4n} = \frac{a_n}{n}\sum_{i=1}^{n}\psi_\tau(\epsilon_i)W_i(\boldsymbol{x})\big[\widehat{c}^u_{Y\boldsymbol{X}}\big(\widehat{F}^u_Y(Y_i), \widehat{\boldsymbol{F}}^u(\boldsymbol{x})\big) - \widehat{c}^u_{Y\boldsymbol{X}}\big(F^u_Y(Y_i), \widehat{\boldsymbol{F}}^u(\boldsymbol{x})\big)\big]$$

$$\tag{A.18}$$

$$\Delta_{5n} = \frac{a_n}{n}\sum_{i=1}^{n}\psi_\tau(\epsilon_i)W_i(\boldsymbol{x})\big[\widehat{c}^u_{Y\boldsymbol{X}}\big(F^u_Y(Y_i), \widehat{\boldsymbol{F}}^u(\boldsymbol{x})\big) - \widehat{c}^u_{Y\boldsymbol{X}}\big(F^u_Y(Y_i), \boldsymbol{F}^u(\boldsymbol{x})\big)\big]$$

$$\tag{A.19}$$

$$\Delta_{6n} = \frac{a_n}{n}\sum_{i=1}^{n}\psi_\tau(\epsilon_i)W_i(\boldsymbol{x})c^u_{Y\boldsymbol{X}}\big(F^u_Y(Y_i), \boldsymbol{F}^u(\boldsymbol{x})\big). \tag{A.20}$$

We will show that all these quantities converge to 0 in probability. Starting with Δ_{3n}, we have, for a large n,

$$|\Delta_{3n}| = \left|\frac{a_n}{n}\sum_{i=1}^{n}\psi_\tau(\epsilon_i)W_i(\boldsymbol{x})\widehat{c}^u_{Y\boldsymbol{X}}\big(\widehat{F}^u_Y(Y_i), \widehat{\boldsymbol{F}}^u(\boldsymbol{x})\big)\frac{\widehat{G}_C(Y_i - |\boldsymbol{x}) - G_C(Y_i - |\boldsymbol{x})}{1 - \widehat{G}_C(Y_i - |\boldsymbol{x})}\right|$$

$$\leq a_n \sup_{t\in\mathbb{R}}\widehat{c}^u_{Y\boldsymbol{X}}\big(\widehat{F}^u_Y(t), \widehat{\boldsymbol{F}}^u(\boldsymbol{x})\big) \sup_{\substack{t\leq\max_{1\leq i\leq n} Y_i \\ \Delta_i=1}}\frac{\left|\widehat{G}_C(t|\boldsymbol{x}) - G_C(t|\boldsymbol{x})\right|}{1 - \widehat{G}_C(t|\boldsymbol{x})}$$

$$\times \frac{1}{n}\sum_{i=1}^{n}|\psi_\tau(\epsilon_i)|W_i(\boldsymbol{x})$$

$$= O_{\mathbb{P}}(a_n) \sup_{t\leq\tau_{F_Y}}\frac{\left|\widehat{G}_C(t|\boldsymbol{x}) - G_C(t|\boldsymbol{x})\right|}{1 - \widehat{G}_C(t|\boldsymbol{x})} = o_{\mathbb{P}}(1),$$

under the condition that assumptions (C2), (C4)-(i), (C6) and (C7)-(i) are met.

The proofs of (A.18) and (A.19) are very similar in spirit. Hence, for the sake of brevity, we only consider here (A.18). For any $i = 1, \ldots, n$, there exists a $\theta_i \in (0, 1)$ such that,

$$\widehat{c}^u_{Y\boldsymbol{X}}\big(\widehat{F}^u_Y(Y_i), \widehat{\boldsymbol{F}}^u(\boldsymbol{x})\big) - \widehat{c}^u_{Y\boldsymbol{X}}\big(F^u_Y(Y_i), \widehat{\boldsymbol{F}}^u(\boldsymbol{x})\big) = \big(\widehat{F}^u_Y(Y_i) - F^u_Y(Y_i)\big)\times$$

$$\partial_1\widehat{c}^u_{Y\boldsymbol{X}}\Big(F^u_Y(Y_i) + \theta_i\big(\widehat{F}^u_Y(Y_i) - F^u_Y(Y_i)\big), \widehat{\boldsymbol{F}}^u(\boldsymbol{x})\Big),$$

where ∂_j denotes the partial derivative with respect to the j-th argument. Therefore, for a large n and under the bandwidth requirement,

$$|\Delta_{4n}| \leq a_n \sup_{t\in\mathbb{R}}|\widehat{F}^u_Y(t) - F^u_Y(t)| \sup_{u_0\in(0,1)}\sup_{\boldsymbol{u}\in\mathcal{V}_{\boldsymbol{F}^u(\boldsymbol{x})}}|\partial_1\widehat{c}^u_{Y\boldsymbol{X}}(u_0, \boldsymbol{u})|$$

$$\times \frac{1}{n} \sum_{i=1}^{n} |\psi_\tau(\epsilon_i)| W_i(\boldsymbol{x})$$
$$= o_{\mathbb{P}}(1),$$

provided assumptions (C4)-(i), (C5), and (C7)-(ii) are satisfied. Similarly, $\Delta_{5n} = o_{\mathbb{P}}(1)$.

Lastly, $\Delta_{6n} = o_{\mathbb{P}}(1)$, since

$$\mathbb{E}(\psi_\tau(\epsilon) W(\boldsymbol{x}) c_{Y\boldsymbol{X}}^u (F_Y^u(Y), \boldsymbol{F}^u(\boldsymbol{x}))) = 0,$$

and, by assumption (C4)-(ii),

$$\frac{1}{n} \sum_{i=1}^{n} \psi_\tau(\epsilon_i) W_i(\boldsymbol{x}) c_{Y\boldsymbol{X}}^u (F_Y^u(Y_i), \boldsymbol{F}^u(\boldsymbol{x})) = O_{\mathbb{P}}(n^{-1/2}).$$

Finally, inserting (A.16) in (A.15), we conclude that

$$a_n(\widehat{m}_\tau(\boldsymbol{x}) - m_\tau(\boldsymbol{x})) = \frac{w(\boldsymbol{x})}{f_{T|\boldsymbol{X}}(m_\tau(\boldsymbol{x})|\boldsymbol{x})} \frac{a_n}{n} \sum_{i=1}^{n} \psi_\tau(\epsilon_i) W_i(\boldsymbol{x})$$
$$\times \left[\widehat{c}_{Y\boldsymbol{X}}^u (F_Y^u(Y_i), \boldsymbol{F}^u(\boldsymbol{x})) - c_{Y\boldsymbol{X}}^u (F_Y^u(Y_i), \boldsymbol{F}^u(\boldsymbol{x})) \right]$$
$$+ o_{\mathbb{P}}(1),$$

which completes the proof. $\qquad\square$

Proof of Corollary 4.2

Given that, by the result of Theorem 4.1, the asymptotic behavior of $\widehat{m}_\tau(\boldsymbol{x})$ will be dictated by the expression of the multivariate copula estimator, we will concentrate on the latter. Using the asymptotic expressions of both bivariate copulas along with the multivariate copula decomposition of Section 2.3, under the required bandwidth conditions we have that,

$$a_n \left(\widehat{c}_{Y\boldsymbol{X}}^u (u_0, \boldsymbol{u}) - c_{Y\boldsymbol{X}}^u (u_0, \boldsymbol{u}) - h^2 b_{Y\boldsymbol{X}}(u_0, \boldsymbol{u}) \right) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \widetilde{Z}^{nj}(u_0, \boldsymbol{u}) + o_{\mathbb{P}}(1),$$

$$\forall \boldsymbol{u} \in (0,1)^2, \text{ uniformly in } u_0 \in (0,1),$$

where

$$\widetilde{Z}^{nj}(u_0, \boldsymbol{u}) = \left(Z_1^{nj}(u_0, u_1) c_2^u(u_0, u_2) + Z_2^{nj}(u_0, u_2) c_1^u(u_0, u_1) \right) c_{X_1 X_2 | Y}^u(u_1, u_2 | u_0)$$

$$b_{Y\boldsymbol{X}}(u_0, \boldsymbol{u}) = \left[b_1(u_0, u_1) c_2^u(u_0, u_2) + b_2(u_0, u_2) c_1^u(u_0, u_1) \right] c_{X_1 X_2 | Y}^u(u_1, u_2 | u_0).$$

Therefore, using the result of Theorem 4.1, we may determine that

$$a_n \left(\widehat{m}_\tau(\boldsymbol{x}) - m_\tau(\boldsymbol{x}) - h^2 \frac{w(\boldsymbol{x})}{f_{T|\boldsymbol{X}}(m_\tau(\boldsymbol{x})|\boldsymbol{x})} n^{-1} \right.$$

$$\times \sum_{i=1}^{n} \psi_\tau(\epsilon_i) W_i(\boldsymbol{x}) b_{Y\boldsymbol{X}} \left(F^u(Y_i), \boldsymbol{F}^u(\boldsymbol{x})\right) \Bigg)$$

$$= \frac{w(\boldsymbol{x})}{f_{T|\boldsymbol{X}}(m_\tau(\boldsymbol{x})|\boldsymbol{x})} \sqrt{n} \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \psi_\tau(\epsilon_i) W_i(\boldsymbol{x}) \widetilde{Z}^{nj} \left(F^u(Y_i), \boldsymbol{F}^u(\boldsymbol{x})\right) + o_\mathbb{P}(1).$$

Let us define

$$B(\boldsymbol{x}) = \frac{w(\boldsymbol{x})}{f_{T|\boldsymbol{X}}(m_\tau(\boldsymbol{x})|\boldsymbol{x})} \mathbb{E}\left(\psi_\tau(\epsilon) W(\boldsymbol{x}) b_{Y\boldsymbol{X}} \left(F^u(Y), \boldsymbol{F}^u(\boldsymbol{x})\right)\right).$$

Provided that assumption (C8) holds, we have

$$\frac{w(\boldsymbol{x})}{f_{T|\boldsymbol{X}}(m_\tau(\boldsymbol{x})|\boldsymbol{x})} n^{-1} \sum_{i=1}^{n} \psi_\tau(\epsilon_i) W_i(\boldsymbol{x}) b_{Y\boldsymbol{X}} \left(F^u(Y_i), \boldsymbol{F}^u(\boldsymbol{x})\right) = B(\boldsymbol{x}) + O_\mathbb{P}(n^{-1/2}),$$

from which it results that

$$a_n \left(\widehat{m}_\tau(\boldsymbol{x}) - m_\tau(\boldsymbol{x}) - h^2 B(\boldsymbol{x})\right)$$
$$= \frac{w(\boldsymbol{x})}{f_{T|\boldsymbol{X}}(m_\tau(\boldsymbol{x})|\boldsymbol{x})} \sqrt{n} \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \psi_\tau(\epsilon_i) W_i(\boldsymbol{x}) \widetilde{Z}^{nj} \left(F^u(Y_i), \boldsymbol{F}^u(\boldsymbol{x})\right) + o_\mathbb{P}(1).$$
$$\tag{A.21}$$

The last step of the proof is then to simplify the obtained expression on the right hand side of the last equality. To that end, define $V_n = n^{-2} \sum_{i=1}^{n} \sum_{j=1}^{n} \psi_\tau(\epsilon_i) \times W_i(\boldsymbol{x}) \widetilde{Z}^{nj} \left(F^u(Y_i), \boldsymbol{F}^u(\boldsymbol{x})\right)$, and observe that this is a V-statistic with the symmetric kernel

$$\Psi_n(V_i, V_j) = \frac{1}{2} \Big[\psi_\tau(\epsilon_i) W_i(\boldsymbol{x}) \widetilde{Z}^{nj} \left(F^u(Y_i), \boldsymbol{F}^u(\boldsymbol{x})\right)$$
$$+ \psi_\tau(\epsilon_j) W_j(\boldsymbol{x}) \widetilde{Z}^{ni} \left(F^u(Y_j), \boldsymbol{F}^u(\boldsymbol{x})\right) \Big],$$

where $V_t = (Y_t, \boldsymbol{X}_t, \Delta_t), t = i, j$. As the statistic's kernel depends on n, this suggests to apply Corollary 1 in Martins-Filho and Yao [30] which establishes the $\sqrt{n}$-equivalence between the V-statistic and the Hájek-projection of its corresponding U-statistic (see e.g. Serfling [40], page 189). Therefore, as $\mathbb{E}\big(Z_k^{nj}(u_0, u_k)\big) = 0$ for all u_0, u_k, $k = 1, 2$, and $\mathbb{E}\big(\Psi_n(V_i, V_j)\big) = 0$, we have

$$V_n = n^{-1} \sum_{i=1}^{n} \lambda_n\left(Y_i, \Delta_i, \boldsymbol{X}_i, \boldsymbol{x}\right) + o_\mathbb{P}(n^{-1/2}), \tag{A.22}$$

where $\lambda_n\left(Y_i, \Delta_i, \boldsymbol{X}_i, \boldsymbol{x}\right) = \mathbb{E}\left[\psi_\tau(\epsilon) W(\boldsymbol{x}) \widetilde{Z}^{ni}\left(F^u(Y), \boldsymbol{F}^u(\boldsymbol{x})\right) | Y_i, \Delta_i, \boldsymbol{X}_i\right]$, provided that assumption (C9) holds. Lastly, the result of Corollary 4.2 follows readily from the insertion of (A.22) in (A.21) and the application of Lyapunov's central limit theorem. $\square$

Acknowledgments

References

[1] BRECHMANN, E. C. (2010). Truncated and Simplified Regular Vines and Their Applications Masterarbeit, Technische Universität München.

[2] BÜCHER, A., EL GHOUCH, A. and VAN KEILEGOM, I. (2014). Single-index Quantile Regression models for Censored Data. *Submitted*.

[3] CHARPENTIER, A., FERMANIAN, J. D. and SCAILLET, O. (2006). *The estimation of copulas: Theory and practice.* J. Rank, ed., Risk Books.

[4] CHAUDHURI, P. (1991). Global nonparametric estimation of conditional quantile functions and their derivatives. *Journal of Multivariate Analysis* **39** 246–269. MR1147121

[5] CHEN, S. X. and HUANG, T. M. (2007). Nonparametric estimation of copula functions for dependence modelling. *Canad. J. Statist.* **35** 265–282. MR2393609

[6] CHEN, K. and LO, S. H. (1997). On the rate of uniform convergence of the product-limit estimator: strong and weak laws. *The Annals of Statistics* **25** 1050–1087. MR1447741

[7] CZADO, C. (2010). *Workshop on Copula Theory and Its Applications.* Springer.

[8] DETTE, H., VAN HECKE, R. and VOLGUSHEV, S. (2014). Some comments on copula-based regression. *J. Amer. Statist. Assoc.* **109(507)** 1319–1324. MR3265699

[9] DISSMANN, J., BRECHMANN, E. C., CZADO, C. and KUROWICKA, D. (2013). Selecting and Estimating Regular Vine Copulae and Application to Financial Returns. *Comput. Stat. Data Anal.* **59(0)** 52–69. MR3000041

[10] EL GHOUCH, A. and VAN KEILEGOM, I. (2009). Local Linear quantile Regression with Dependent Censored Data. *Statistica Sinica* **19** 1621–1640. MR2589201

[11] ELSNER, J. B., KOSSIN, J. P. and JAGGER, T. H. (2008). The increasing intensity of the strongest tropical cyclones. *Nature* **455(7209)** 92–95.

[12] EMBRECHTS, P. (2009). Copulas: A Personal View. *Journal of Risk and Insurance* **76(3)** 639–650.

[13] FENG, X., HE, X. and HU, J. (2011). Wild bootstrap for quantile regression. *Biometrika* **98** 995–999. MR2860339

[14] GEENENS, G., CHARPENTIER, A. and PAINDAVEINE, D. (2014). Probit transformation for nonparametric kernel estimation of the copula density. *arXiv:1404.4414 [stat.ME]*. MR3624880

[15] GENEST, C. and NEŠLEHOVÁ, J. (2007). A Primer on Copulas for Count Data. *The ASTIN Bulletin* **37(2)** 475–515. MR2422797

[16] GIJBELS, I. and MIELNICZUK, J. (1990). Estimating the density of a copula function. *Communications in Statistics – Theory and Methods* **19(2)** 445–464. MR1067582

[17] HJORT, N. L. and POLLARD, D. (1993). Asymptotics for minimisers of convex processes Technical Report, Yale University.

[18] HOBÆK HAFF, I., AAS, K. and FRIGESSI, A. (2010). On the simplified pair-copula construction – simply useful or too simplistic? *Journal of Multivariate Analysis* **101(5)** 1296–1310. MR2595309

[19] HOBÆK HAFF, I. and SEGERS, J. (2015). Nonparametric estimation of pair-copula constructions with the empirical pair-copula. *Computational Statistics & Data Analysis* **84** 1–13. MR3292793

[20] HOFERT, M. and PHAM, D. (2013). Densities of nested Archimedian copulas. *Journal of Multivariate Analysis* **118** 37–52. MR3054089

[21] JOE, H. (2014). *Dependence Modeling with Copulas.* Chapman & Hall/CRC Monographs on Statistics & Applied Probability.

[22] KOENKER, R. (2005). *Quantile Regression.* Cambridge Univ. Press. MR2268657

[23] KOENKER, R. and BASSET, G. J. (1978). Regression quantiles. *Econometrica* **46(1)** 33–50. MR0474644

[24] KOENKER, R. and BILIAS, Y. (2001). Quantile regression for duration data: A reappraisal of the Pensylvania employment bonus experiments. *Empirical Economics* **26** 199–220.

[25] KOENKER, R. and GELING, O. (2001). Reappraising medfly longevity: a quantile regression survival analysis. *J. Amer. Statist. Assoc.* **96** 458–468. MR1939348

[26] LANGHOLZ, B. and GOLDSTEIN, L. (1996). Risk set sampling in epidemiologic cohort studies. *Statistical Science* 35–53.

[27] LENG, C. and TONG, X. (2013). A quantile regression estimator for censored data. *Bernoulli* **19(1)** 344–361. MR3019498

[28] LI, Q., LIN, J. and RACINE, J. S. (2013). Optimal bandwidth selection for nonparametric conditional distribution and quantile functions. *Journal of Business & Economic Statistics* **31** 57–65. MR3030798

[29] LUBIN, J. H., BOICE, J. D., EDLING, C., HORNUNG, R. W., HOWE, G. R., KUNZ, E., KUSIAK, R. A., MORRISON, H. I., RADFORD, E. P., SAMET, J. M. et al. (1995). Lung cancer in radon-exposed miners and estimation of risk from indoor exposure. *Journal of the National Cancer Institute* **87** 817–827.

[30] MARTINS-FILHO, C. and YAO, F. (2006). A Note on the use of V and U statistics in nonparametric models of regression. *Annals of the Institute of Statistical Mathematics* **58** 389–406. MR2246162

[31] Nagler, T. (2014). Kernel Methods for Vine Copula Estimation Masterarbeit, Technische Universität München.

[32] Nagler, T. and Czado, C. (2016). Evading the curse of dimensionality in multivariate kernel density estimation with simplified vines. *Journal of Multivariate Analysis* **151** 69–89. MR3545278

[33] Nelsen, R. (2006). *An Introduction to Copulas.* Springer, New York. MR2197664

[34] Noh, H., El Ghouch, A. and Bouezmarni, T. (2013). Copula-Based Regression Estimation and Inference. *J. Amer. Statist. Assoc.* **108** 676–688. MR3174651

[35] Noh, H., El Ghouch, A. and Van Keilegom, I. (2015). Semiparametric Conditional Quantile Estimation through Copula-Based Multivariate Models. *Journal of Business and Economic Statistics* **33(2)** 167–178. MR3337055

[36] Oh, D. H. and Patton, A. (2012). Modelling dependence in high dimensions with factor copulas. *Manuscript, Duke University.*

[37] Portnoy, S. (2003). Censored Regression Quantiles. *J. Amer. Statist. Assoc.* **98(464)** 1001–1012. MR2041488

[38] Portnoy, S. and Koenker, R. (1997). The Gaussian hare and the Laplacian tortoise: computability of squared-error versus absolute-error estimators. *Statistical Science* **12** 279–300. MR1619189

[39] Powell, J. L. (1986). Censored Regression Quantiles. *J. Econometrics.* **32** 143–155. MR0853049

[40] Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics. Wiley Series in Probability and Statistics – Applied Probability and Statistics Section Series.* Wiley. MR0595165

[41] Sklar, M. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Stitst. Univ. Paris* **8** 229–231. MR0125600

[42] Spokoiny, V., Wang, W. and Härdle, W. K. (2013). Local quantile regression. *Journal of Statistical Planning and Inference* **143** 1109–1129. MR3049611

[43] Stöber, J., Joe, H. and Czado, C. (2013). Simplified pair copula constructions – limitations and extensions. *Journal of Multivariate Analysis* **119(0)** 101–118. MR3061418

[44] R Core Team (2014). R: A Language and Environment for Statistical Computing R Foundation for Statistical Computing, Vienna, Austria.

[45] Wang, H. J. and Wang, L. (2009). Locally weighted censored quantile regression. *J. Amer. Statist. Assoc.* **104** 1117–1128. MR2562007

[46] Wu, T., Yu, K. and Yu, Y. (2010). Single-index quantile regression. *Journal of Multivariate Analysis* **101** 1607–1621. MR2610735

[47] Ying, Z., Jung, S. H. and Wei, L. J. (1995). Survival analysis with median regression models. *J. Amer. Statist. Assoc.* **90** 178–184. MR1325125

[48] Zhu, L., Huang, M. and Li, R. (2012). Semiparametric quantile regression with high-dimensional covariates. *Statistica Sinica* **22** 1379–1401. MR3027092