



HAL
open science

Nonparametric estimation of conditional marginal excess moments

Yuri Goegebeur, Armelle Guillou, Nguyen Khanh Le Ho, Jing Qin

► **To cite this version:**

Yuri Goegebeur, Armelle Guillou, Nguyen Khanh Le Ho, Jing Qin. Nonparametric estimation of conditional marginal excess moments. *Journal of Multivariate Analysis*, 2023, 193, <10.1016/j.jmva.2022.105121>. <hal-03271590v2>

HAL Id: hal-03271590

<https://hal.science/hal-03271590v2>

Submitted on 4 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Nonparametric estimation of conditional marginal excess moments

Yuri Goegebeur⁽¹⁾, Armelle Guillou⁽²⁾, Nguyen Khanh Le Ho⁽¹⁾, Jing Qin⁽¹⁾

⁽¹⁾ Department of Mathematics and Computer Science, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark

⁽²⁾ Institut Recherche Mathématique Avancée, UMR 7501, Université de Strasbourg et CNRS, 7 rue René Descartes, 67084 Strasbourg cedex, France

Abstract

Several risk measures have been proposed in the literature, among them the marginal mean excess, defined as $MME_p = \mathbb{E}[(Y^{(1)} - Q_1(1-p))_+ | Y^{(2)} > Q_2(1-p)]$, provided $\mathbb{E}|Y^{(1)}| < \infty$, where $(Y^{(1)}, Y^{(2)})$ denotes a pair of risk factors, $y_+ := \max(0, y)$, Q_j the quantile function of $Y^{(j)}$, $j = 1, 2$, and $p \in (0, 1)$. In this paper we consider a generalization of this measure, where the random variables of main interest $(Y^{(1)}, Y^{(2)})$ are observed together with a random covariate $X \in \mathbb{R}^d$, and where the $Y^{(1)}$ excess is also power transformed. This leads to the concept of conditional marginal excess moment for which an estimator is proposed allowing extrapolation outside the data range. The main asymptotic properties of this estimator have been established, using empirical processes arguments combined with the multivariate extreme value theory. The finite sample behavior of the estimator is evaluated by a simulation experiment. We apply also our method on a vehicle insurance customer dataset.

1 Introduction

In many scientific disciplines, quantifying risks related to extreme events is of crucial importance. For instance, insurance companies are once in a while faced with extreme claims which can jeopardise the solvency of a portfolio, and hence accurate modelling of the upper tail of the claim size distribution assumes a central place in their risk management. Other examples of disciplines include environmental science (storms, rainfall), geology (severe earthquakes), hydrology (flooding) and telecommunication (network load). The quantification of the risk of a risk factor Y with distribution function F_Y is done by so-called risk measures. In the univariate context, the most commonly used risk measures are the Value-at-Risk (VaR), defined as $VaR_p = Q(1-p)$, where Q denotes the quantile function of Y , i.e.,

$$Q(p) := \inf\{y : F_Y(y) \geq p\}, \quad p \in (0, 1),$$

see, e.g., Jorion (2007) for a review, and the conditional tail expectation (CTE), given by

$$CTE_p = \mathbb{E}(Y | Y > Q(1-p)), \quad p \in (0, 1),$$

provided $\mathbb{E}|Y| < \infty$. During recent years, the CTE became a popular alternative to VaR since, compared to VaR it is more conservative and it is also a coherent risk measure. We refer to, e.g., Artzner et al. (1999), Cai and Tan (2007) and Brazauskas et al. (2008). In practice, risk is often related to several risk factors, and hence the above discussed risk measures need to be adjusted for

this multivariate context. For a pair of risk factors $(Y^{(1)}, Y^{(2)})$, with $\mathbb{E}|Y^{(1)}| < \infty$, the CTE can be generalized to the marginal expected shortfall (MES), defined as

$$MES_p = \mathbb{E}(Y^{(1)} | Y^{(2)} > Q_2(1-p)), \quad p \in (0, 1), \quad (1)$$

where Q_2 denotes the quantile function of the risk factor $Y^{(2)}$. This measure was introduced by Acharya et al. (2010), to measure the contribution of a financial firm to an overall systemic risk. For a financial firm, the MES is defined as its short-run expected equity loss conditional on the market taking a loss greater than its VaR. Cai et al. (2015) studied the MES in a bivariate extreme value framework, and proposed an estimator for it when the $Y^{(2)}$ quantile is extreme, i.e., when $p < 1/n$. See also Di Bernardino and Prieur (2018), and Das and Fasen-Hartmann (2018, 2019) for related analyses of the MES in a multivariate extreme value setting. By replacing $Y^{(1)}$ in (1) by an excess over a high quantile of its distribution one obtains the marginal mean excess (MME):

$$MME_p := \mathbb{E}[(Y^{(1)} - Q_1(1-p))_+ | Y^{(2)} > Q_2(1-p)], \quad p \in (0, 1),$$

provided $\mathbb{E}|Y^{(1)}| < \infty$, and where $y_+ := \max(0, y)$ and Q_1 is the quantile function of $Y^{(1)}$. Das and Fasen-Hartmann (2018, 2019) study a slightly different version of MME under multivariate regular variation, introduced an estimator for it based on extreme value arguments, and established its consistency.

Recently, several of the above mentioned risk measures have been generalized to the situation where the variable(s) of main interest is (are) observed together with a random covariate. Daouia et al. (2011, 2013) studied the estimation of extreme conditional quantiles, while the CTE was extended to the regression case in El Methni et al. (2014, 2018). Goegebeur et al. (2021a, b) considered the estimation of the MES in presence of random covariates.

In this paper we consider a generalization of the marginal mean excess, where the random variables of main interest $(Y^{(1)}, Y^{(2)})$ are observed together with a random covariate $X \in \mathbb{R}^d$, and where the $Y^{(1)}$ excess is also power transformed. We assume that the covariate X has a density function f_X , with support $S_X \subset \mathbb{R}^d$, and we denote by $F_j(\cdot|x)$ the conditional distribution function of $Y^{(j)}$ given $X = x$ and by $U_j(\cdot|x)$ the associated conditional tail quantile function, i.e., $U_j(\cdot|x) := \inf\{y : F_j(y|x) \geq 1 - 1/\cdot\}$, $j = 1, 2$. In particular, we introduce the conditional marginal excess moment (CMEM), defined as

$$\theta_{\beta,p}(x_0) = \mathbb{E} \left[(Y^{(1)} - U_1(1/p|x_0))_+^\beta | Y^{(2)} > U_2(1/p|x_0), X = x_0 \right], \quad (2)$$

where $\beta > 0$, provided $\mathbb{E}(|Y^{(1)}|^\beta | X = x_0) < \infty$, and where x_0 is a reference position such that $x_0 \in \text{Int}(S_X)$, the interior of the support of f_X , assumed to be non-empty. The motivation for introducing this power β is that, in an insurance context for instance, with $Y^{(1)}$ denoting the claim size, $(Y^{(1)} - U_1(1/p|x_0))_+$ can be viewed as the payment by the reinsurer. Thus, different values of β allow us to compute, among others, the expectation or the variance of this payment. We are interested in the situation where p is small, i.e., $p < 1/n$, with n denoting the sample size. To obtain this extrapolation, we will work in two steps, where in a first step we study an intermediate case, which allows for constructing an empirical estimator for (2). In a second step this intermediate estimator will be extrapolated outside the data range by a Weissman-type construction.

The remainder of the paper is organized as follows. In Section 2 we consider the CMEM in the intermediate case. We introduce a locally weighted average as estimator for the CMEM and derive its limiting distribution under suitable conditions. This intermediate estimator is then extrapolated outside the data range in Section 3, where we study again the asymptotic properties. In Section 4 we evaluate the finite sample performance with a simulation experiment, while in Section 5 we illustrate the method on a vehicle insurance customer dataset. Some auxiliary results and their proofs are given in Section 6. Section 7 contains the proofs of the main results.

2 Estimator for the intermediate case

Let $(Y_i^{(1)}, Y_i^{(2)}, X_i)$, $i = 1, \dots, n$, be independent copies of $(Y^{(1)}, Y^{(2)}, X)$. We start with considering an estimator for $\theta_{\beta,p}(x_0)$ when $p \downarrow 0$ at an intermediate rate, i.e., when $p = k/n$, where $k \rightarrow \infty$ as $n \rightarrow \infty$ but in such a way that $k/n \rightarrow 0$. In this case it is natural to consider

$$\hat{\theta}_n := \frac{1}{k} \sum_{i=1}^n K_{h_n}(x_0 - X_i) \left[Y_i^{(1)} - \hat{U}_1 \left(\frac{n}{k} | x_0 \right) \right]_+^\beta \mathbb{1}_{\{Y_i^{(2)} > \hat{U}_2(\frac{n}{k} | x_0)\}},$$

where $K_{h_n}(\cdot) := K(\cdot/h_n)/h_n^d$, with K a joint density function on \mathbb{R}^d , $(h_n)_{n \geq 1}$ is a positive non-random sequence of bandwidths with $h_n \rightarrow 0$ if $n \rightarrow \infty$, $\mathbb{1}_A$ the indicator function on the event A , and $\hat{U}_j(\cdot | x_0)$ is an estimator for $U_j(\cdot | x_0)$, defined as $\hat{U}_j(\cdot | x_0) = \inf\{y : \hat{F}_{n,j}(y | x_0) \geq 1 - 1/\cdot\}$, where $\hat{F}_{n,j}(y | x_0)$ is a classical kernel estimator (Rosenblatt, 1956, Parzen, 1962), given by

$$\hat{F}_{n,j}(y | x_0) = \frac{\frac{1}{n} \sum_{i=1}^n K_{h_n}(x_0 - X_i) \mathbb{1}_{\{Y_i^{(j)} \leq y\}}}{\hat{f}_n(x_0)}, \quad (3)$$

for $j = 1, 2$, with

$$\hat{f}_n(x_0) := \frac{1}{n} \sum_{i=1}^n K_{h_n}(x_0 - X_i),$$

being a density estimator.

To study the asymptotic behaviour of this estimator we need to introduce some conditions.

We assume that $Y^{(1)}$ and $Y^{(2)}$ are positive random variables, and that they follow a conditional Pareto-type model. Let RV_ψ denote the class of regularly varying functions at infinity with index ψ , i.e., positive measurable functions f satisfying $f(tx)/f(x) \rightarrow x^\psi$, as $t \rightarrow \infty$, for all $x > 0$. If $\psi = 0$, then we call f a slowly varying function at infinity.

Assumption (D) For all $x \in S_X$, the conditional survival function of $Y^{(j)}$, $j = 1, 2$, given $X = x$, satisfies

$$\bar{F}_j(y | x) = A_j(x) y^{-1/\gamma_j(x)} \left(1 + \frac{1}{\gamma_j(x)} \delta_j(y | x) \right),$$

where $A_j(x) > 0$, $\gamma_j(x) > 0$, and $|\delta_j(\cdot | x)|$ is normalized regularly varying at infinity with index $-\beta_j(x)$, $\beta_j(x) > 0$, i.e.,

$$\delta_j(y | x) = B_j(x) \exp \left(\int_1^y \frac{\varepsilon_j(u | x)}{u} du \right),$$

with $B_j(x) \in \mathbb{R}$ and $\varepsilon_j(y | x) \rightarrow -\beta_j(x)$ as $y \rightarrow \infty$. Moreover, we assume $y \rightarrow \varepsilon_j(y | x)$ to be a continuous function.

Under Assumption (D) we have that $U_j(\cdot | x)$, $j = 1, 2$, satisfy

$$U_j(y | x) = [A_j(x)]^{\gamma_j(x)} y^{\gamma_j(x)} (1 + a_j(y | x)), \quad (4)$$

where $a_j(y | x) := \delta_j(U_j(y | x) | x) (1 + o(1))$, and thus $|a_j(\cdot | x)| \in RV_{-\beta_j(x)\gamma_j(x)}$.

We also need a condition that describes the right-hand upper tail dependence of the joint distribution of $(Y^{(1)}, Y^{(2)})$ given $X = x$. Let $R_t(y_1, y_2 | x) := t\mathbb{P}(\bar{F}_1(Y^{(1)} | x) \leq y_1/t, \bar{F}_2(Y^{(2)} | x) \leq y_2/t | X = x)$.

Assumption (\mathcal{R}) For all $x \in S_X$ we have as $t \rightarrow \infty$

$$R_t(y_1, y_2|x) \rightarrow R(y_1, y_2|x),$$

uniformly in $y_1, y_2 \in (0, T]$, for any $T > 0$, and in $x \in B(x_0, r)$, for some $r > 0$.

Note that this condition reflects in fact the asymptotic behavior of the conditional copula function. It is an adjustment of the usual first order condition in multivariate extreme value theory (see, e.g., Chapter 6 in de Haan and Ferreira, 2006, and Cai et al., 2015) to the regression context.

Based on conditions (\mathcal{D}) and (\mathcal{R}) we can already obtain the following theoretical approximation to $\theta_{\beta,p}(x_0)$, which is a result that is in fact valid for any $p \downarrow 0$.

Lemma 1 Assume (\mathcal{D}), (\mathcal{R}) and $\gamma_1(x_0) < 1/\beta$. Then, as $p \downarrow 0$ we have

$$\frac{\theta_{\beta,p}(x_0)}{[U_1(1/p|x_0)]^\beta} \rightarrow - \int_0^\infty R \left(\left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}}, 1 \middle| x_0 \right) du^{-\gamma_1(x_0)}.$$

In order to deal with the regression context, $f_X(\cdot)$, $R(y_1, y_2|\cdot)$ and the functions appearing in $\bar{F}_j(y|\cdot)$, $j = 1, 2$, are assumed to satisfy the following Hölder conditions. Let $\|\cdot\|$ denote some norm on \mathbb{R}^d .

Assumption (\mathcal{H}) There exist positive constants M_{f_X} , M_R , M_{A_j} , M_{γ_j} , M_{B_j} , M_{ε_j} , η_{f_X} , η_R , η_{A_j} , η_{γ_j} , η_{B_j} , η_{ε_j} , where $j = 1, 2$, and $\kappa > \gamma_1(x_0)\beta$, such that for all $x, z \in S_X$:

$$\begin{aligned} |f_X(x) - f_X(z)| &\leq M_{f_X} \|x - z\|^{\eta_{f_X}}, \\ \sup_{y_1 > 0, \frac{1}{2} \leq y_2 \leq 2} \frac{|R(y_1, y_2|x) - R(y_1, y_2|z)|}{y_1^\kappa \wedge 1} &\leq M_R \|x - z\|^{\eta_R}, \\ |A_j(x) - A_j(z)| &\leq M_{A_j} \|x - z\|^{\eta_{A_j}}, \\ |\gamma_j(x) - \gamma_j(z)| &\leq M_{\gamma_j} \|x - z\|^{\eta_{\gamma_j}}, \\ |B_j(x) - B_j(z)| &\leq M_{B_j} \|x - z\|^{\eta_{B_j}}, \\ \sup_{y \geq 1} |\varepsilon_j(y|x) - \varepsilon_j(y|z)| &\leq M_{\varepsilon_j} \|x - z\|^{\eta_{\varepsilon_j}}. \end{aligned}$$

We also impose a condition on the kernel function K , which is a standard condition in local estimation, see, e.g., Daouia et al. (2011, 2013) and Escobar-Bach et al. (2018a).

Assumption (\mathcal{K}) K is a bounded density function on \mathbb{R}^d , with support S_K included in the unit ball in \mathbb{R}^d , with respect to the norm $\|\cdot\|$.

Before stating the weak convergence of $\hat{\theta}_n$, we need to introduce a second order condition.

Assumption (\mathcal{S}). There exist $\kappa > \gamma_1(x_0)\beta$ and $\tau < 0$ such that, as $t \rightarrow \infty$

$$\sup_{x \in B(x_0, r)} \sup_{y_1 > 0, \frac{1}{2} \leq y_2 \leq 2} \frac{|R_t(y_1, y_2|x) - R(y_1, y_2|x)|}{y_1^\kappa \wedge 1} = O(t^\tau),$$

for some $r > 0$.

This second order condition specifies the rate of convergence of $R_t(y_1, y_2|x)$ to its limit $R(y_1, y_2|x)$ as $t \rightarrow \infty$. It is an adjustment of the second order condition in Cai et al. (2015) to the regression

context, by assuming that the order of approximation is also uniform in $x \in B(x_0, r)$. Note that the uniform requirement in the second order condition excludes the case where $(Y^{(1)}, Y^{(2)})$ are asymptotically upper tail independent given $X = x_0$, which corresponds to $R(y_1, y_2|x_0) = 0$.

We can now state our first main result, the weak convergence of $\hat{\theta}_n/\hat{f}_n(x_0)$, properly normalized. Weak convergence is denoted by the arrow \rightsquigarrow .

Theorem 1 *Assume (\mathcal{D}) , (\mathcal{H}) , (\mathcal{K}) , (\mathcal{S}) with $x \rightarrow R(y_1, y_2|x)$ being a continuous function, $R(y, 1|x_0)$ is continuously differentiable in y , and $y \rightarrow F_j(y|x_0)$, $j = 1, 2$, are strictly increasing. Let $x_0 \in \text{Int}(S_X)$ such that $f_X(x_0) > 0$. Consider sequences $k = \lfloor n^\alpha \ell_1(n) \rfloor$ and $h_n = n^{-\Delta} \ell_2(n)$, where ℓ_1 and ℓ_2 are slowly varying functions at infinity, with $\alpha \in (0, 1)$ and*

$$\max \left(\frac{\alpha}{d + 2(\eta_R \wedge \eta_{f_X} \wedge \eta_{A_1} \wedge \eta_{\gamma_1})}, \frac{\alpha - 2\beta_1(x_0)\gamma_1(x_0)(1 - \alpha)}{d}, \frac{\alpha}{d + 2(\eta_{A_2} \wedge \eta_{\gamma_2})[1 - \beta\gamma_1(x_0)]}, \right. \\ \left. \frac{\alpha - 2\beta_2(x_0)\gamma_2(x_0)(1 - \alpha)(1 - \beta\gamma_1(x_0))}{d + 2(\eta_{\varepsilon_2} \wedge \eta_{B_2})[1 - \beta\gamma_1(x_0)]}, \frac{\alpha}{d} - \frac{2\beta(1 - \alpha)\gamma_1^2(x_0)\beta_1(x_0)}{d[1 + \gamma_1(x_0)(\beta_1(x_0) - \varepsilon)]}, \right. \\ \left. \frac{\alpha + 2(1 - \alpha)\tau}{d} \right) < \Delta < \frac{\alpha}{d},$$

where $0 < \varepsilon < \beta_1(x_0)$.

Then, for $\gamma_1(x_0) < 1/(2\beta)$, we have

$$\sqrt{kh_n^d} \left(\frac{\hat{\theta}_n}{\hat{f}_n(x_0)\theta_{\beta, k/n}(x_0)} - 1 \right) \rightsquigarrow \frac{1}{f_X(x_0) \int_0^\infty R([1 + u^{-\frac{\gamma_1(x_0)}{\beta}}]^{-\frac{1}{\gamma_1(x_0)}}, 1|x_0) du^{-\gamma_1(x_0)}} \\ \times \left\{ \int_0^\infty W([1 + u^{-\frac{\gamma_1(x_0)}{\beta}}]^{-\frac{1}{\gamma_1(x_0)}}, 1) du^{-\gamma_1(x_0)} \right. \\ \left. - (W(1, \infty) - W(\infty, 1)) \int_0^\infty R_1([1 + u^{-\frac{\gamma_1(x_0)}{\beta}}]^{-\frac{1}{\gamma_1(x_0)}}, 1|x_0) (1 + u^{-\frac{\gamma_1(x_0)}{\beta}})^{-\frac{1}{\gamma_1(x_0)}} du^{-\gamma_1(x_0)} \right\} \\ - \frac{W(\infty, 1)}{f_X(x_0)} + \frac{\beta\gamma_1(x_0)}{f_X(x_0)} W(1, \infty),$$

where $W(y_1, y_2)$ is a zero centered Gaussian process with covariance function

$$\mathbb{E}(W(y_1, y_2)W(\bar{y}_1, \bar{y}_2)) = \|K\|_2^2 f_X(x_0) R(y_1 \wedge \bar{y}_1, y_2 \wedge \bar{y}_2|x_0),$$

with $\|K\|_2 := \sqrt{\int_{\mathbb{R}^d} K^2(u) du}$, and $W(\infty, y)$ and $W(y, \infty)$ are zero centered Gaussian processes with the same covariance function

$$\mathbb{E}(W(\infty, y)W(\infty, \bar{y})) = \mathbb{E}(W(y, \infty)W(\bar{y}, \infty)) = \|K\|_2^2 f_X(x_0) (y \wedge \bar{y}),$$

and $R_1(y, 1|x_0)$ denotes the derivative of $R(y, 1|x_0)$ with respect to y .

The variance of the limiting random variable in Theorem 1, denoted \mathbb{W}_1 , is given by

$$\text{Var}(\mathbb{W}_1) = \frac{\|K\|_2^2}{f_X(x_0)} \left\{ \beta^2 \gamma_1^2(x_0) - 1 - \frac{I_1}{I_2^2} + 2[1 - R(1, 1|x_0)] \right. \\ \left. \times \left(\left(\frac{I_3}{I_2} \right)^2 - [1 + \beta\gamma_1(x_0)] \frac{I_3}{I_2} + \beta\gamma_1(x_0) \right) \right\},$$

where

$$\begin{aligned} I_1 &:= \int_0^\infty R \left(\left[1 + u^{-\gamma_1(x_0)/\beta} \right]^{-1/\gamma_1(x_0)}, 1 \mid x_0 \right) du^{-2\gamma_1(x_0)}, \\ I_2 &:= \int_0^\infty R \left(\left[1 + u^{-\gamma_1(x_0)/\beta} \right]^{-1/\gamma_1(x_0)}, 1 \mid x_0 \right) du^{-\gamma_1(x_0)}, \\ I_3 &:= \int_0^\infty R_1 \left(\left[1 + u^{-\gamma_1(x_0)/\beta} \right]^{-1/\gamma_1(x_0)}, 1 \mid x_0 \right) \left[1 + u^{-\gamma_1(x_0)/\beta} \right]^{-1/\gamma_1(x_0)} du^{-\gamma_1(x_0)}. \end{aligned}$$

3 Estimator for the extreme case

We now turn to the estimation of $\theta_{\beta,p}(x_0)$ under extrapolation. Assuming (\mathcal{D}) , (\mathcal{R}) and $\gamma_1(x_0) < 1/\beta$, according to Lemma 1, we have the following approximation

$$\theta_{\beta,p}(x_0) \sim \frac{[U_1(1/p|x_0)]^\beta}{[U_1(n/k|x_0)]^\beta} \theta_{\beta, \frac{k}{n}}(x_0) \sim \left(\frac{k}{np} \right)^{\beta\gamma_1(x_0)} \theta_{\beta, \frac{k}{n}}(x_0).$$

To estimate $\theta_{\beta,p}(x_0)$, we can use the estimator $\bar{\theta}_{\beta, \frac{k}{n}}(x_0) := \hat{\theta}_n / \hat{f}_n(x_0)$ of $\theta_{\beta, \frac{k}{n}}(x_0)$ combined with an estimator of the extreme value index $\gamma_1(x_0)$. For the latter, we can use the local Hill (1975) estimator

$$\hat{\gamma}_{1,k_1}(x_0) := \frac{\frac{1}{k_1} \sum_{i=1}^n K_{h_n}(x_0 - X_i) \left(\ln Y_i^{(1)} - \ln \hat{U}_1(n/k_1|x_0) \right) \mathbb{1}_{\{Y_i^{(1)} \geq \hat{U}_1(n/k_1|x_0)\}}}{\hat{f}_n(x_0)},$$

already studied in Goegebeur et al. (2021b), and based on an intermediate sequence k_1 , possibly different to k , such that $k_1 \rightarrow \infty$ with $k_1/n \rightarrow 0$. This yields the Weissman-type estimator for $\theta_{\beta,p}(x_0)$

$$\hat{\theta}_{\beta,p}(x_0) = \left(\frac{k}{np} \right)^{\beta\hat{\gamma}_{1,k_1}(x_0)} \bar{\theta}_{\beta, \frac{k}{n}}(x_0).$$

The estimator is said to be of Weissman-type, as it is in nature similar to an estimator for an extreme quantile proposed by Weissman (1978).

The next theorem states our main result, namely the weak convergence of our Weissman-type estimator $\hat{\theta}_{\beta,p}(x_0)$, properly normalized.

Theorem 2 *Assume (\mathcal{D}) , (\mathcal{H}) , (\mathcal{K}) , (\mathcal{S}) with $x \rightarrow R(y_1, y_2|x)$ being a continuous function, $R(y, 1|x_0)$ is continuously differentiable in y , and $y \rightarrow F_j(y|x_0)$, $j = 1, 2$, are strictly increasing. Let $x_0 \in \text{Int}(S_X)$ such that $f_X(x_0) > 0$. Consider sequences $k = \lfloor n^\alpha \ell_1(n) \rfloor$, $k_1 = \lfloor n^{\alpha_1} \ell_2(n) \rfloor$ and $h_n = n^{-\Delta} \ell_3(n)$, where ℓ_1 , ℓ_2 and ℓ_3 are slowly varying functions at infinity, with $\alpha \in (0, 1)$ and*

$$\alpha \leq \alpha_1 < \min \left(\frac{\alpha}{d} [d + 2(\eta_{f_X} \wedge \eta_{A_1} \wedge \eta_{\gamma_1})], \frac{\alpha + 2\gamma_1(x_0)\beta_1(x_0)}{1 + 2\gamma_1(x_0)\beta_1(x_0)} \right),$$

and

$$\begin{aligned} \max \left(\frac{\alpha}{d + 2(\eta_R \wedge \eta_{f_X} \wedge \eta_{A_1} \wedge \eta_{\gamma_1})}, \frac{\alpha - 2\beta_1(x_0)\gamma_1(x_0)(1 - \alpha)}{d}, \frac{\alpha}{d + 2(\eta_{A_2} \wedge \eta_{\gamma_2})[1 - \beta\gamma_1(x_0)]}, \right. \\ \left. \frac{\alpha - 2\beta_2(x_0)\gamma_2(x_0)(1 - \alpha)(1 - \beta\gamma_1(x_0))}{d + 2(\eta_{\varepsilon_2} \wedge \eta_{B_2})[1 - \beta\gamma_1(x_0)]}, \frac{\alpha}{d} - \frac{2\beta(1 - \alpha)\gamma_1^2(x_0)\beta_1(x_0)}{d[1 + \gamma_1(x_0)(\beta_1(x_0) - \varepsilon)]}, \right. \\ \left. \frac{\alpha + 2(1 - \alpha)\tau}{d}, \frac{\alpha_1}{d + 2(\eta_{f_X} \wedge \eta_{\gamma_1} \wedge \eta_{A_1})}, \frac{\alpha_1 - 2\gamma_1(x_0)\beta_1(x_0)(1 - \alpha_1)}{d} \right) < \Delta < \frac{\alpha}{d}, \end{aligned}$$

where $0 < \varepsilon < \beta_1(x_0)$.

Then, for $\gamma_1(x_0) < 1/(2\beta)$ and p satisfying $p \leq \frac{k}{n}$ such that $\frac{\ln k/(np)}{\sqrt{k_1 h_n^d}} \rightarrow 0$ and $\sqrt{\frac{k}{k_1}} \ln \frac{k}{np} \rightarrow r \in [0, \infty]$, we have

$$\begin{aligned} & \min \left(\sqrt{k h_n^d}, \frac{\sqrt{k_1 h_n^d}}{\ln k/(np)} \right) \left(\frac{\hat{\theta}_{\beta,p}(x_0)}{\theta_{\beta,p}(x_0)} - 1 \right) \rightsquigarrow \min(r, 1) \frac{\beta \gamma_1(x_0)}{f_X(x_0)} \left(\int_0^1 W(u, \infty) \frac{1}{u} du - W(1, \infty) \right) \\ & + \min \left(1, \frac{1}{r} \right) \left\{ \frac{1}{f_X(x_0) \int_0^\infty R([1 + u^{-\frac{\gamma_1(x_0)}{\beta}}]^{-\frac{1}{\gamma_1(x_0)}}, 1|x_0) du^{-\gamma_1(x_0)}} \right. \\ & \times \left[\int_0^\infty W([1 + u^{-\frac{\gamma_1(x_0)}{\beta}}]^{-\frac{1}{\gamma_1(x_0)}}, 1) du^{-\gamma_1(x_0)} \right. \\ & \left. \left. - (W(1, \infty) - W(\infty, 1)) \int_0^\infty R_1([1 + u^{-\frac{\gamma_1(x_0)}{\beta}}]^{-\frac{1}{\gamma_1(x_0)}}, 1|x_0) (1 + u^{-\frac{\gamma_1(x_0)}{\beta}})^{-\frac{1}{\gamma_1(x_0)}} du^{-\gamma_1(x_0)} \right] \right. \\ & \left. - \frac{W(\infty, 1)}{f_X(x_0)} + \frac{\beta \gamma_1(x_0)}{f_X(x_0)} W(1, \infty) \right\}. \end{aligned}$$

The variance of the limiting random variable in Theorem 2, denoted \mathbb{W}_2 , is given by

$$\begin{aligned} Var(\mathbb{W}_2) &= \frac{\|K\|_2^2}{f_X(x_0)} \{ [\min(r, 1)]^2 \beta^2 \gamma_1^2(x_0) \\ & + \left[\min \left(1, \frac{1}{r} \right) \right]^2 \left[\beta^2 \gamma_1^2(x_0) - 1 - \frac{I_1}{I_2^2} + 2[1 - R(1, 1|x_0)] \left(\left(\frac{I_3}{I_2} \right)^2 - [1 + \beta \gamma_1(x_0)] \frac{I_3}{I_2} + \beta \gamma_1(x_0) \right) \right] \\ & + 2 \min \left(r, \frac{1}{r} \right) \beta \gamma_1(x_0) \left[\frac{\frac{1}{\beta} I_4 + I_5}{\gamma_1(x_0) I_2} + \left(R(1, 1|x_0) - \int_0^1 \frac{R(u, 1|x_0)}{u} du \right) \left(1 - \frac{I_3}{I_2} \right) - 1 \right] \}, \end{aligned}$$

where I_1, I_2 and I_3 are defined as in the definition of the variance of \mathbb{W}_1 and

$$\begin{aligned} I_4 &:= \int_0^\infty R \left([1 + u^{-\gamma_1(x_0)/\beta}]^{-1/\gamma_1(x_0)}, 1|x_0 \right) \frac{u^{-\gamma_1(x_0)/\beta}}{1 + u^{-\gamma_1(x_0)/\beta}} du^{-\gamma_1(x_0)}, \\ I_5 &:= \int_0^\infty R \left([1 + u^{-\gamma_1(x_0)/\beta}]^{-1/\gamma_1(x_0)}, 1|x_0 \right) \ln(1 + u^{-\gamma_1(x_0)/\beta}) du^{-\gamma_1(x_0)}. \end{aligned}$$

4 Simulation

In this section we illustrate the finite sample performance of $\hat{\theta}_{\beta,p}(x_0)$ with a simulation experiment. We simulate from the following models:

Model 1. We consider the logistic copula model

$$C(u_1, u_2|x) = e^{-[(-\ln u_1)^x + (-\ln u_2)^x]^{1/x}}, \quad u_1, u_2 \in [0, 1], x \geq 2. \quad (5)$$

We take $X \sim U[2, 10]$, and combine this copula model with a Burr(ζ, λ, τ) distribution for $Y^{(1)}$:

$$F_1(y) = 1 - \left(\frac{\zeta}{\zeta + y^\tau} \right)^\lambda, \quad y > 0; \zeta, \lambda, \tau > 0,$$

with $\zeta = 1, \lambda = 0.5, \tau = 8$, giving $\gamma_1 = 0.25$, and a Fréchet distribution for $Y^{(2)}$:

$$F_2(y) = e^{-y^{-1/\gamma_2}}, \quad y > 0,$$

with $\gamma_2 = 0.5$. This model satisfies (\mathcal{S}) with $R(y_1, y_2|x) = y_1 + y_2 - (y_1^x + y_2^x)^{1/x}$, $\tau = -1$ and $\kappa = 1 - \varepsilon$ for some small $\varepsilon > 0$.

Model 2. The conditional distribution of $(Y^{(1)}, Y^{(2)})$ given $X = x$ is that of

$$(|Z_1|^{\gamma_1(x)}, |Z_2|^{\gamma_2(x)}),$$

where (Z_1, Z_2) follow a bivariate standard Cauchy distribution with density function

$$f(z_1, z_2) = \frac{1}{2\pi}(1 + z_1^2 + z_2^2)^{-3/2}, \quad (z_1, z_2) \in \mathbb{R}^2.$$

We take $X \sim U[0, 1]$, and set

$$\gamma_1(x) = 0.25 + 0.125 \sin(2\pi x), \quad \gamma_2(x) = 0.1 + 0.1x.$$

This model satisfies (\mathcal{S}) with $R(y_1, y_2|x) = y_1 + y_2 - \sqrt{y_1^2 + y_2^2}$, $\tau = -1$ and $\beta = 2$ (see, e.g., Cai et al. 2015, in the context without covariates).

Model 3: We consider again the conditional logistic copula model defined in (5) combined with conditional Burr distributions for $Y^{(1)}$ and $Y^{(2)}$ given $X = x$, where we take $\zeta_1 = \zeta_2 = 1$, $\lambda_1 = 1$, $\lambda_2 = 0.5$, and

$$\tau_1(x) = 2e^{0.2x}, \quad \tau_2(x) = 8/\sin(0.3x).$$

For this model $\gamma_1(x) = 0.5e^{-0.2x}$. Similarly to Model 1, this model satisfies (\mathcal{S}) .

For all the models, the conditional marginal distributions $F_j(\cdot|x_0)$, $j = 1, 2$, satisfy Assumption (\mathcal{D}) (see, e.g., Beirlant et al., 2009, Table 1), as well as Assumption (\mathcal{H}) .

We implement our estimators $\hat{\theta}_{\beta,p}(x_0)$ and $\hat{\gamma}_{1,k_1}(x_0)$ with a bi-quadratic kernel function, given by

$$K(x) = \frac{15}{16}(1 - x^2)^2 \mathbb{1}_{\{x \in [-1, 1]\}},$$

which clearly satisfies Assumption (\mathcal{K}) . Related to this, we need to select also a bandwidth h_n . To this aim, we use the cross-validation procedure introduced by Yao (1999), and already used in the extreme value framework by Daouia et al. (2011, 2013) and Escobar-Bach et al. (2018a), and defined as:

$$h_{cv} := \operatorname{argmin}_{h_n \in \mathcal{H}} \sum_{i=1}^n \sum_{j=1}^n \left(\mathbb{1}_{\{Y_i^{(2)} \leq Y_j^{(2)}\}} - \hat{F}_{n,h_n,2,-i}(Y_j^{(2)}|X_i) \right)^2,$$

where \mathcal{H} is the grid of values defined as $R_X \times \{0.05, 0.10, \dots, 0.30\}$, with R_X the range of the covariate X , and

$$\hat{F}_{n,h_n,2,-i}(y|x) := \frac{\sum_{k=1, k \neq i}^n K_{h_n}(x - X_k) \mathbb{1}_{\{Y_k^{(2)} \leq y\}}}{\sum_{k=1, k \neq i}^n K_{h_n}(x - X_k)}.$$

We simulate 500 datasets of sizes $n = 500$ and 1000 from each model. For each sample, we compute $\hat{\theta}_{\beta,p}(x_0)$ for two different values of p : $1/n$ and $1/(2n)$, and for two values of β , where the latter is chosen such that the condition $\gamma_1(x_0) < 1/(2\beta)$ is satisfied: $\beta = 1$ for all models and $\beta = 1.5, 1.25, 1.2$, for Models 1-3, respectively.

The intermediate sequence k_1 on which the estimator for $\gamma_1(x_0)$ is based is selected by a graphical assessment, where k_1 is chosen as the smallest value after which the median of $\hat{\gamma}_{1,k_1}(x_0)$, computed over the 500 replications, shows a stable part.

In Figure 1 we show for Model 1 the boxplots of $\hat{\theta}_{1,p}(x_0)$, where $k = 0.15n$, at various positions of the covariate x_0 , for $n = 500$ (top row) and $n = 1000$ (bottom row), and for $p = 1/n$ (left) and $p = 1/(2n)$ (right). The red curve shows the true value of $\theta_{1,p}(x_0)$. The layout of Figure 2 is similar but with $\beta = 1.5$. Figures 3 and 4 show the corresponding results for Model 2, and Figures 5 and 6 for Model 3. From these simulations we can draw the following conclusions:

- Overall, the estimator performs quite well, but of course the results depend on the model and the covariate position. In Model 1, $R(y_1, y_2|x_0)$ depends on the covariate, but the marginal distributions do not. For Model 2, $R(y_1, y_2|x_0)$ does not depend on the value of x_0 but the marginal distributions do, and for Model 3 both $R(y_1, y_2|x_0)$ and the two marginals depend on x_0 . As is clear from the plots of the true function $\theta_{\beta,p}(x_0)$, this quantity does not change a lot in the covariate for Model 1, while on the contrary it changes a lot in the covariate for Model 2, where there is a maximum and a minimum. In Model 3, $\theta_{\beta,p}(x_0)$ decreases in x_0 . Hence, the estimation under Model 1 is easier than for the other two models.
- The function $\theta_{\beta,p}(x_0)$ follows the pattern of $\gamma_1(x_0)$ rather closely, see Figure 7 where we graph the functions $\gamma_1(x_0)$ for Models 2 and 3.
- Of the three models considered, Model 2 is most challenging for estimation as $\theta_{\beta,p}(x_0)$ changes a lot with the covariate. Near the x_0 value where $\theta_{\beta,p}(x_0)$ attains its maximum, the method tends to underestimate the true value. This can be explained by the local nature of the estimation: indeed, for such positions, local estimation will be based on $Y^{(1)}$ data coming from distributions with a lighter tail, leading to an underestimation.
- Note that the estimation with $p = 1/n$ corresponds already with extrapolation, as the estimation is done locally, and hence based on fewer observations than n .
- For a given n , smaller values of p lead, as expected, to more variable estimates, as do larger values of β .
- The variability of $\hat{\theta}_{\beta,p}(x_0)$ is larger at x_0 where $\gamma_1(x_0)$ is large. This can be expected since at such positions the $Y^{(1)}$ -data are more heavy tailed, and hence show large variability.
- The cross-validation procedure that was used here leads to a global bandwidth, which gives a reasonable performance for the whole covariate range, and which also works quite well for a wide variety of models. For situations where $\theta_{\beta,p}(x_0)$ varies a lot with the covariate it can be advantageous to use a local bandwidth, since at such positions a smaller bandwidth will lead to less bias in the estimation. Evaluating the performance of bandwidth selection criteria is a topic for future research.

5 Real data analysis

In this section we illustrate our method on a real dataset. We analyze the dataset Vehicle Insurance Customer Data, which is publicly available at <https://www.kaggle.com/ranja7/vehicle-insurance-customer-data>. The dataset contains socio-economic data of insurance customers as well as details about the insured vehicle. We focus on estimating the first and second marginal excess moments of the cumulative claim amount throughout the contract, $Y^{(1)}$, given a customer's lifetime value, $Y^{(2)}$, exceeding a high quantile of its distribution and given an observation for the covariate income, X . In our analysis, we only consider the data for which the customer's income is not zero, which leads to a total of $n = 6817$ observations.



Figure 1: Model 1. Boxplots of $\hat{\theta}_{1,p}(x_0)$ for $n = 500$ (top row) and $n = 1000$ (bottom row), and $p = 1/n$ (left) and $p = 1/(2n)$ (right), along with $\theta_{1,p}(x_0)$ (red curve).

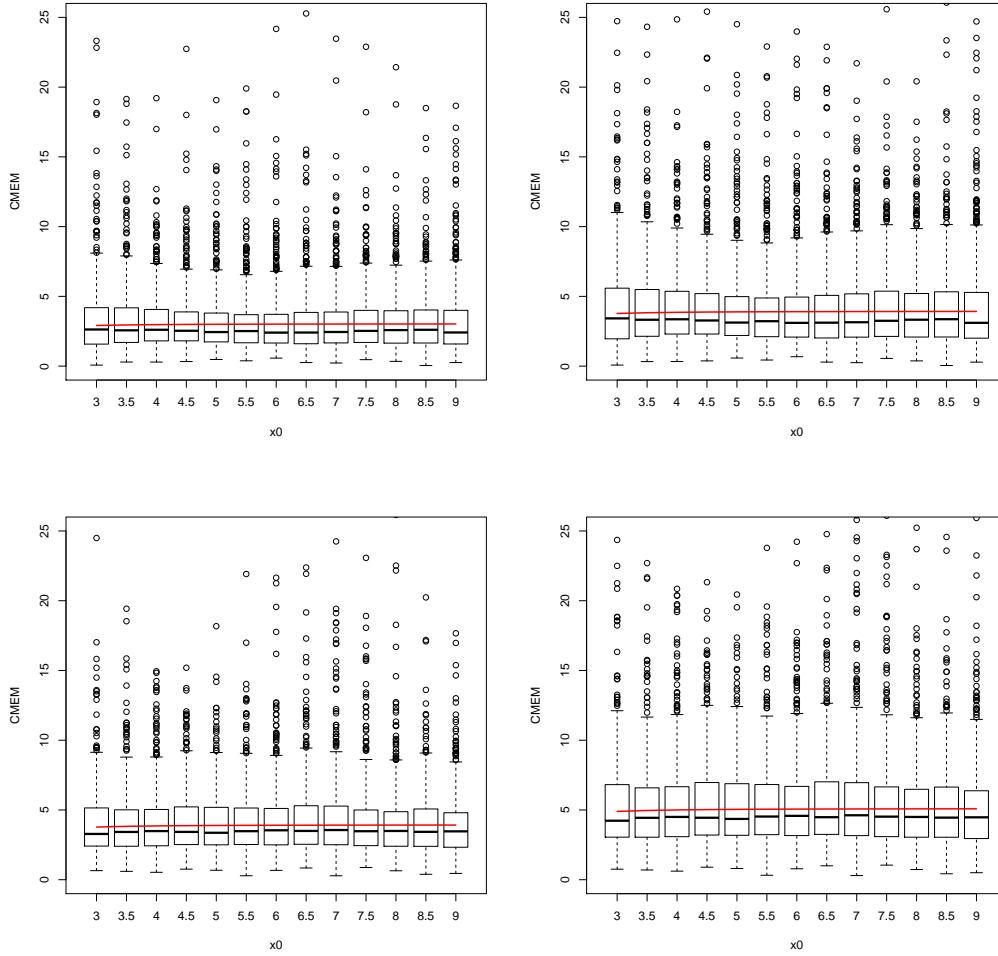


Figure 2: Model 1. Boxplots of $\hat{\theta}_{1.5,p}(x_0)$ for $n = 500$ (top row) and $n = 1000$ (bottom row), and $p = 1/n$ (left) and $p = 1/(2n)$ (right), along with $\theta_{1.5,p}(x_0)$ (red curve).

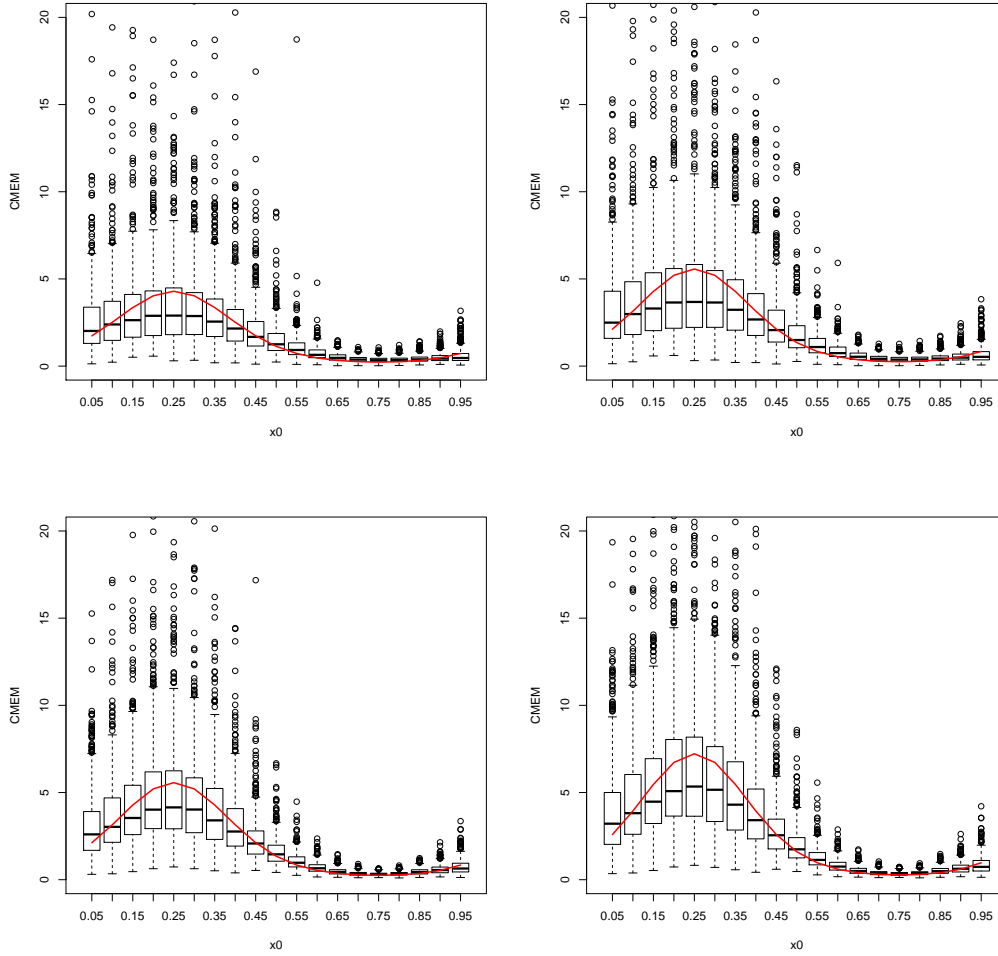


Figure 3: Model 2. Boxplots of $\hat{\theta}_{1,p}(x_0)$ for $n = 500$ (top row) and $n = 1000$ (bottom row), and $p = 1/n$ (left) and $p = 1/(2n)$ (right), along with $\theta_{1,p}(x_0)$ (red curve).

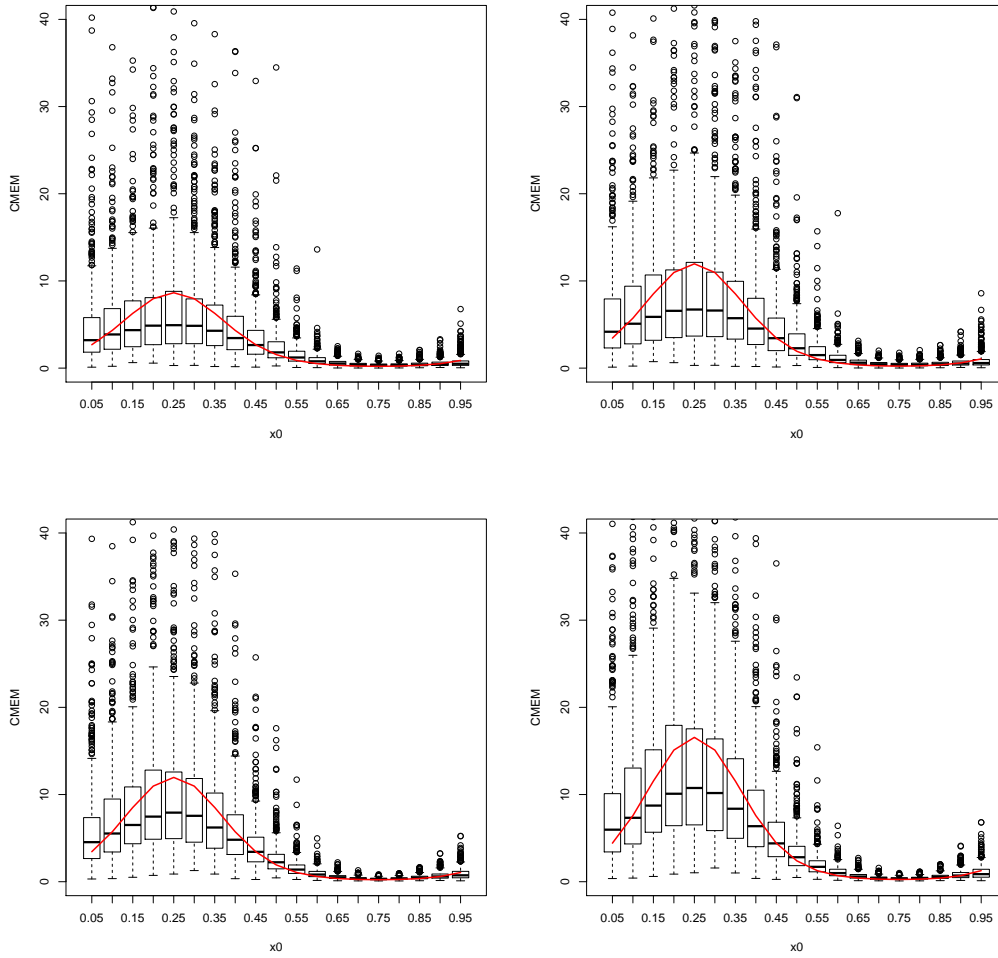


Figure 4: Model 2. Boxplots of $\hat{\theta}_{1.25,p}(x_0)$ for $n = 500$ (top row) and $n = 1000$ (bottom row), and $p = 1/n$ (left) and $p = 1/(2n)$ (right), along with $\theta_{1.25,p}(x_0)$ (red curve).

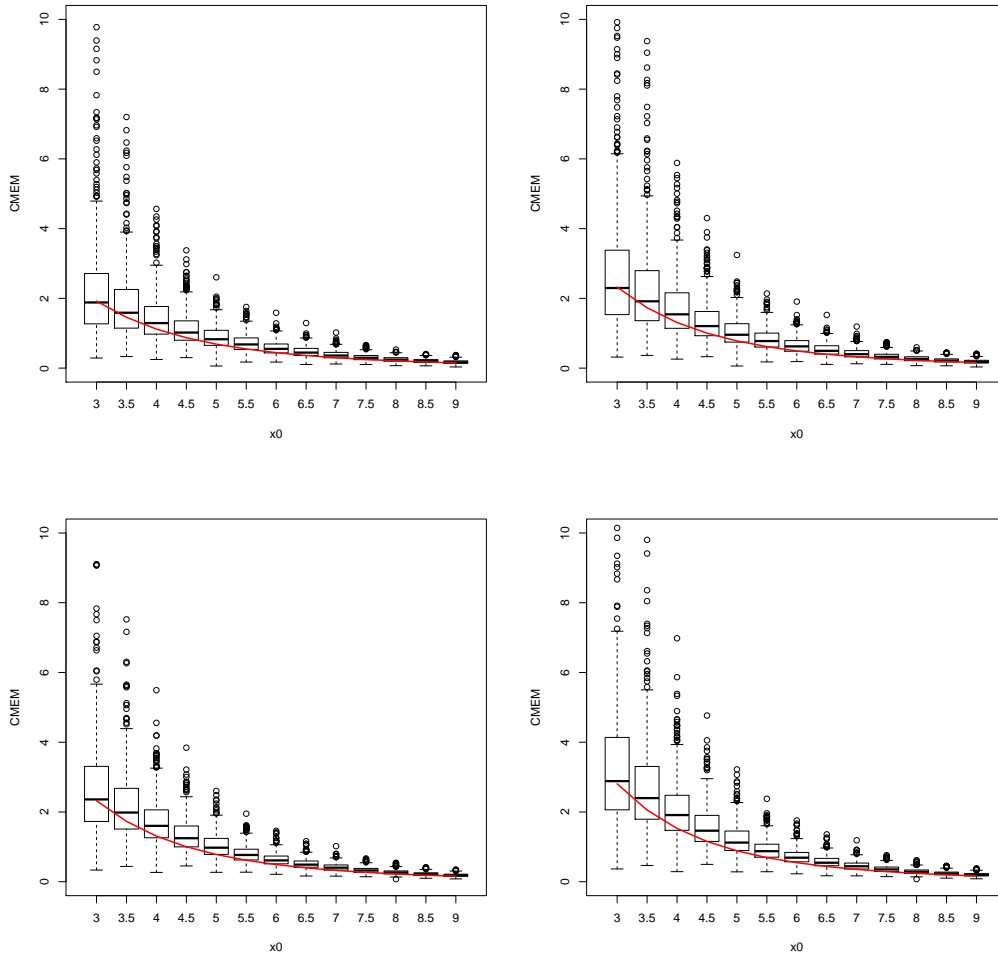


Figure 5: Model 3. Boxplots of $\hat{\theta}_{1,p}(x_0)$ for $n = 500$ (top row) and $n = 1000$ (bottom row), and $p = 1/n$ (left) and $p = 1/(2n)$ (right), along with $\theta_{1,p}(x_0)$ (red curve).

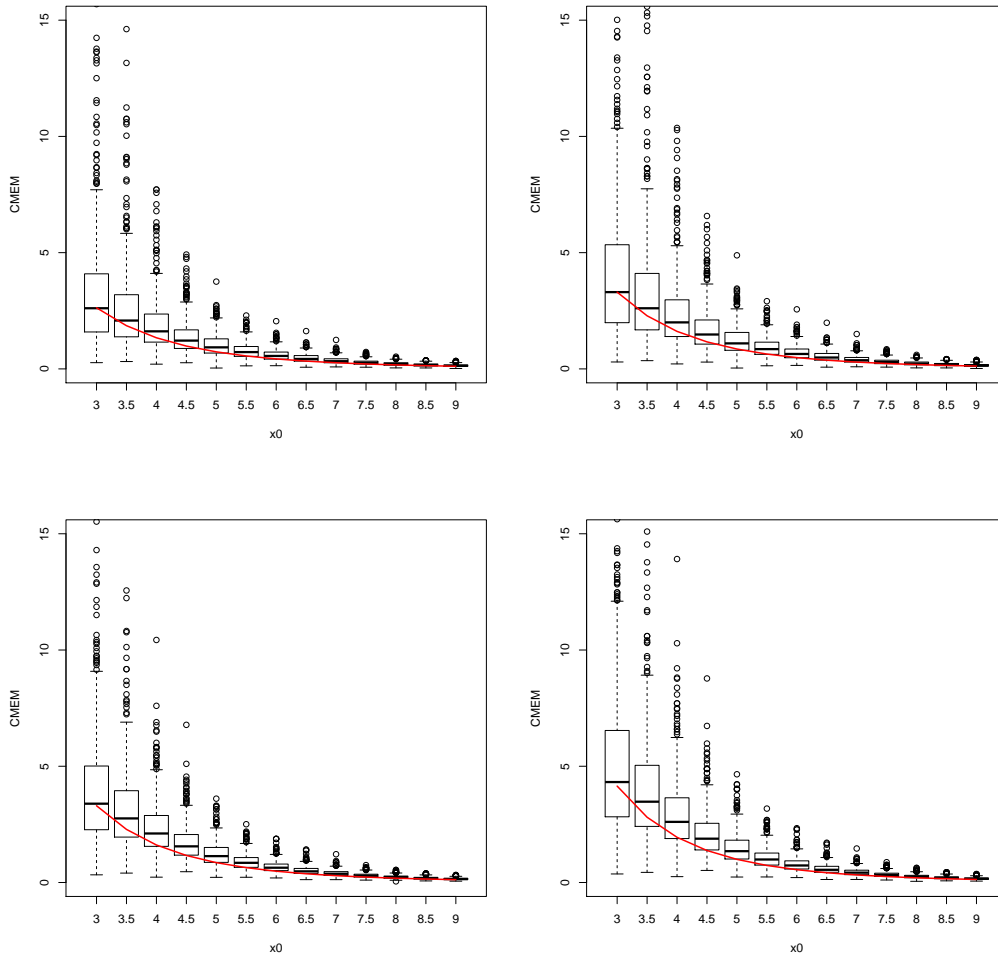


Figure 6: Model 3. Boxplots of $\hat{\theta}_{1,2,p}(x_0)$ for $n = 500$ (top row) and $n = 1000$ (bottom row), and $p = 1/n$ (left) and $p = 1/(2n)$ (right), along with $\theta_{1,2,p}(x_0)$ (red curve).

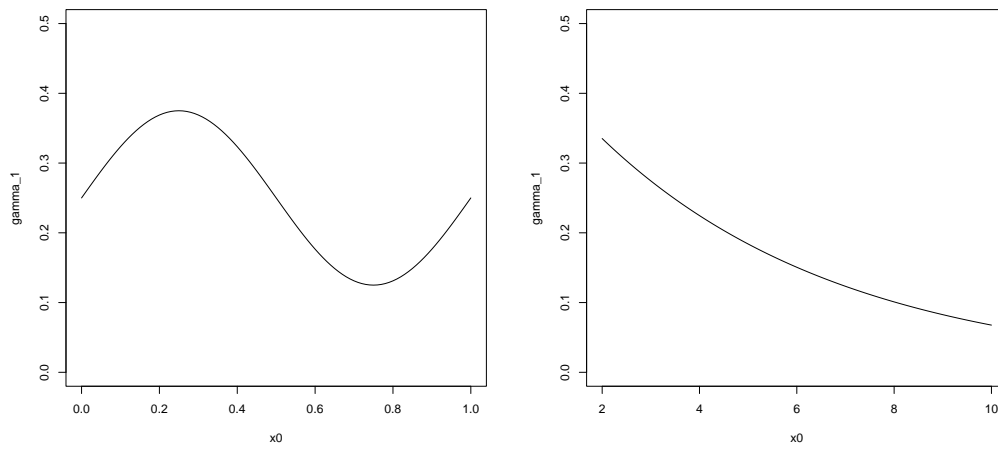


Figure 7: $\gamma_1(x_0)$ as a function of x_0 for Model 2 (left) and Model 3 (right).

In Figure 8, we show a scatterplot of the total claim amount versus the customer's lifetime value, which suggests a positive dependence between the two variables. The model assumptions of asymptotic dependence and marginal conditional distributions of Pareto-type with moreover $\gamma_1(x_0) < 1/(2\beta)$ can be verified by estimating $R(1, 1|x_0)$, the construction of local Pareto quantile plots for the $Y^{(1)}$ and $Y^{(2)}$ data, and the estimation of $\gamma_1(x_0)$. We refer for this analysis to Goegebeur et al. (2021b), where in particular $\gamma_1(x_0)$ estimates were obtained in the range 0.30-0.35, and this for several values of x_0 .

From Theorem 2, by applying the log-transformation, we obtain the following pointwise $(1 - \alpha)$ -confidence interval for $\theta_{1,p}(x_0)$:

$$\left[\hat{\theta}_{1,p}(x_0) \exp \left\{ z_{\alpha/2} \frac{\sqrt{\text{Var}(\mathbb{W}_2)}}{a_n} \right\}, \hat{\theta}_{1,p}(x_0) \exp \left\{ -z_{\alpha/2} \frac{\sqrt{\text{Var}(\mathbb{W}_2)}}{a_n} \right\} \right],$$

where $z_{\alpha/2}$ is the $\alpha/2$ -quantile of the standard normal distribution and $a_n := \min \left(\sqrt{kh_n^d}, \frac{\sqrt{k_1 h_n^d}}{\ln k/(np)} \right)$.

Note that a confidence interval for $\theta_{2,p}(x_0)$ cannot be obtained due to the condition $\gamma_1(x_0) < 1/4$ in Theorem 2, which is not satisfied for the present data. However, it can be shown that for consistency of $\hat{\theta}_{\beta,p}(x_0)$ one has the weaker constraint $\gamma_1(x_0) < 1/\beta$, meaning that for the present data we can still estimate $\theta_{2,p}(x_0)$ in a consistent way.

As for the confidence intervals, we estimate the integrals involved in the variance of the limiting random variable \mathbb{W}_2 by replacing the function R with an adjusted version of the estimator proposed by Escobar-Bach et al. (2018b) for the stable tail dependence function:

$$\hat{R}(y_1, y_2|x_0) = \frac{\frac{1}{k} \sum_{i=1}^n K_{h_n}(x_0 - X_i) \mathbb{1}_{\left\{ \hat{F}_{n,1}(Y_i^{(1)}|X_i) \leq \frac{k}{n} y_1, \hat{F}_{n,2}(Y_i^{(2)}|X_i) \leq \frac{k}{n} y_2 \right\}}}{\hat{f}_n(x_0)},$$

where $\hat{F}_{n,1}$ and $\hat{F}_{n,2}$ are given in (3). The values of k and k_1 are determined by using a data driven procedure proposed in Goegebeur et al. (2019), after which we estimate r by $\sqrt{\frac{k_1}{k} \ln \frac{k}{np}}$. As for the choice of the bandwidth h_n , we again determine it by applying the cross-validation criterion described in Section 4.

In Figure 9, we show the estimates for $\theta_{1,p}(x_0)$ together with the pointwise 95%-confidence intervals for different values of income x_0 , with $p = 0.001$ and $p = 0.0005$. As expected, the estimates of $\theta_{1,p}(x_0)$ with $p = 0.0005$ are higher than those with $p = 0.001$, while at the same time $p = 0.0005$ leads also to wider confidence intervals than $p = 0.001$, reflecting the higher uncertainty of the estimate, which is due to the fact that the estimation is based on fewer observations. Combining these with the estimates for second marginal excess moments we obtain estimates for the standard deviation of the excess $(Y^{(1)} - U_1(1/p|x_0))_+$, given that $Y^{(2)} > U_2(1/p|x_0)$ and $X = x_0$, which is shown in Figure 10.

6 Preliminary results

In order to prove our main theorems, we need some auxiliary results. For any $u > 0$ and $v \in S_K$, define

$$\tilde{s}_n(y_1, u|x_0 - h_n v) = \frac{n}{k} \bar{F}_1 \left(U_1 \left(\frac{n}{k} |x_0 \right) y_1 \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right] |x_0 - h_n v \right),$$

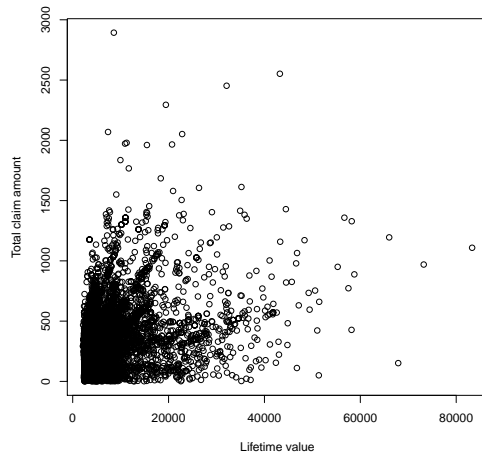


Figure 8: Vehicle insurance dataset. Scatterplot for the total claim amount versus the customer's lifetime value.

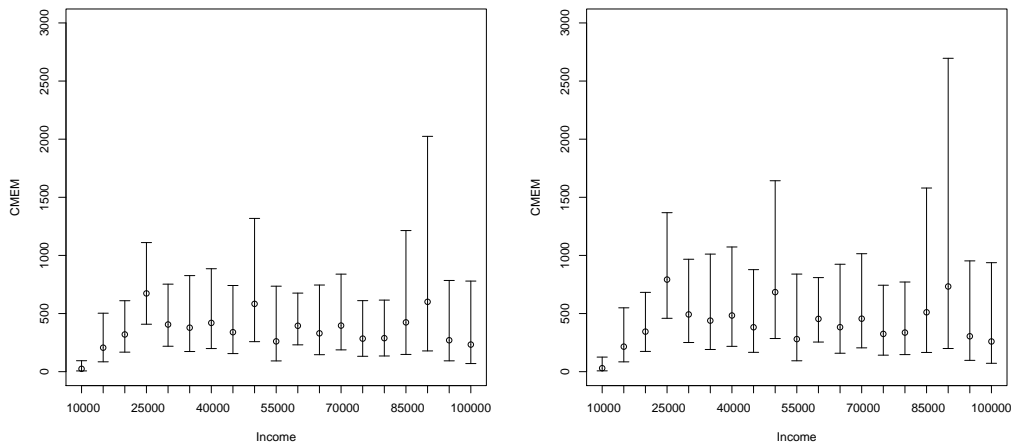


Figure 9: Vehicle insurance dataset. Estimates for $\theta_{1,p}(x_0)$ together with the pointwise 95%-confidence interval for $p = 0.001$ (left) and $p = 0.0005$ (right).

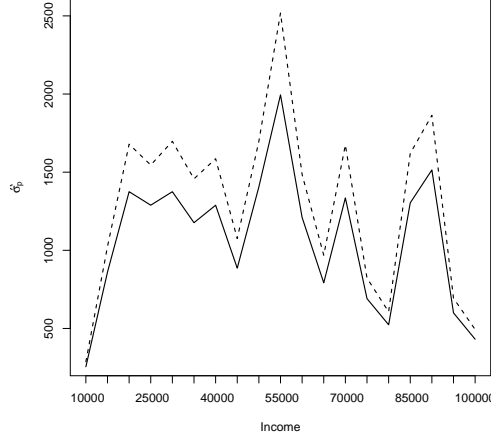


Figure 10: Vehicle insurance dataset. Estimates for the standard deviation of $(Y^{(1)} - U_1(1/p|x_0))_+$ given that $Y^{(2)} > U_2(1/p|x_0)$ and $X = x_0$ as function of income for $p = 0.001$ (solid line) and $p = 0.0005$ (dashed line).

and

$$\tilde{t}_n(y_2|x_0 - h_nv) = \frac{n}{k} \bar{F}_2 \left(U_2 \left(\frac{n}{ky_2} | x_0 \right) | x_0 - h_nv \right).$$

Lemma 2 Assume (\mathcal{D}) and (\mathcal{H}) and $x_0 \in \text{Int}(S_X)$. Consider sequences $k \rightarrow \infty$ and $h_n \rightarrow 0$ as $n \rightarrow \infty$, in such a way that $k/n \rightarrow 0$ and $h_n^{\eta_{\varepsilon_j} \wedge \eta_{\gamma_j}} \ln \frac{n}{k} \rightarrow 0$, $j = 1, 2$. Then, we have, for n large, that

$$\begin{aligned} & \left| \tilde{s}_n(y_1, u|x_0 - h_nv) - \left(y_1 \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right] \right)^{-\frac{1}{\gamma_1(x_0)}} \right| \\ & \leq C \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-1/\gamma_1(x_0)} \left\{ h_n^{\eta_{A_1}} + h_n^{\eta_{\gamma_1}} \ln \frac{n}{k} + h_n^{\eta_{\gamma_1}} \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{Ch_n^{\eta_{\gamma_1}}} \ln \left(1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right) \right. \\ & \quad + \left| \delta_1 \left(U_1 \left(\frac{n}{k} | x_0 \right) | x_0 \right) \right| \times \left[1 + h_n^{\eta_{\varepsilon_1}} \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{Ch_n^{\eta_{\varepsilon_1}}} \left(\ln \frac{n}{k} + \ln \left(1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right) \right) \right. \\ & \quad \left. \left. + h_n^{\eta_{\varepsilon_1} + \eta_{\gamma_1}} \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{Ch_n^{\eta_{\varepsilon_1}} + Ch_n^{\eta_{\gamma_1}}} \ln \left(1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right) \left(\ln \frac{n}{k} + \ln \left(1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right) \right) \right] \right\} \end{aligned}$$

and

$$|\tilde{t}_n(y_2|x_0 - h_nv) - y_2| \leq C \left\{ h_n^{\eta_{A_2}} + h_n^{\eta_{\gamma_2}} \ln \frac{n}{k} + \left| \delta_2 \left(U_2 \left(\frac{n}{k} | x_0 \right) | x_0 \right) \right| \left[h_n^{\eta_{B_2}} + h_n^{\eta_{\varepsilon_2}} \ln \frac{n}{k} \right] \right\},$$

uniformly in $u > 0$, $y_1, y_2 \in [1/2, 2]$ and in $v \in S_K$.

Lemma 3 Assume (\mathcal{D}) , (\mathcal{H}) , $\gamma_1(x_0) < 1/\beta$ and $x_0 \in \text{Int}(S_X)$. For sequences $k = \lfloor n^\alpha \ell_1(n) \rfloor$ and $h_n = n^{-\Delta} \ell_2(n)$, where ℓ_1 and ℓ_2 are slowly varying functions at infinity, with $\alpha \in (0, 1)$ and

$$\begin{aligned} & \max \left(\frac{\alpha}{d + 2(\eta_{A_1} \wedge \eta_{\gamma_1})}, \frac{\alpha - 2\beta_1(x_0)\gamma_1(x_0)(1 - \alpha)}{d}, \frac{\alpha}{d + 2(\eta_{A_2} \wedge \eta_{\gamma_2})[1 - \beta\gamma_1(x_0)]}, \right. \\ & \quad \left. \frac{\alpha - 2\beta_2(x_0)\gamma_2(x_0)(1 - \alpha)(1 - \beta\gamma_1(x_0))}{d + 2(\eta_{\varepsilon_2} \wedge \eta_{B_2})[1 - \beta\gamma_1(x_0)]} \right) < \Delta < \frac{\alpha}{d} \end{aligned}$$

one has that

$$\sup_{v \in S_K} \sup_{y_1, y_2 \in [1/2, 2]} \sqrt{kh_n^d} \left| \int_0^\infty \left[R\left(\tilde{s}_n(y_1, u|x_0 - h_nv), \tilde{t}_n(y_2|x_0 - h_nv)|x_0\right) - R\left(\left(y_1 \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}}\right]\right)^{-\frac{1}{\gamma_1(x_0)}}, y_2|x_0\right) \right] du^{-\gamma_1(x_0)} \right| \rightarrow 0.$$

Introduce now the notation

$$\hat{\theta}_n(y_1, y_2) := \frac{1}{k} \sum_{i=1}^n K_{h_n}(x_0 - X_i) \left[Y_i^{(1)} - U_1\left(\frac{n}{k}|x_0\right) y_1 \right]_+^\beta \mathbb{1}_{\{Y_i^{(2)} > U_2(\frac{n}{ky_2}|x_0)\}}.$$

Assuming $F_1(y|x_0)$ strictly increasing in y , we have

$$\hat{\theta}_n(y_1, y_2) = - \left[y_1 U_1\left(\frac{n}{k}|x_0\right) \right]^\beta \int_0^\infty T_n(\hat{s}_n(y_1, u|x_0), y_2|x_0) du^{-\gamma_1(x_0)}, \quad (6)$$

where

$$T_n(y_1, y_2|x_0) := \frac{1}{k} \sum_{i=1}^n K_{h_n}(x_0 - X_i) \mathbb{1}_{\{\bar{F}_1(Y_i^{(1)}|x_0) \leq \frac{k}{n} y_1, \bar{F}_2(Y_i^{(2)}|x_0) \leq \frac{k}{n} y_2\}}$$

and

$$\hat{s}_n(y_1, u|x_0) := \frac{n}{k} \bar{F}_1\left(U_1\left(\frac{n}{k}|x_0\right) y_1 \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}}\right]|x_0\right).$$

Remark that our final estimator $\hat{\theta}_n = \hat{\theta}_n(\hat{u}_{n,1}, \hat{e}_{n,2})$, where $\hat{u}_{n,i} := \frac{\hat{U}_i(n/k|x_0)}{\bar{U}_i(n/k|x_0)}$, $i = 1, 2$, and $\hat{e}_{n,2} := \frac{n}{k} \bar{F}_2(\hat{u}_{n,2} U_2(n/k|x_0)|x_0)$.

Define now

$$E_n(y_1, y_2) = \mathbb{E} \left(\frac{\hat{\theta}_n(y_1, y_2)}{\left[y_1 U_1\left(\frac{n}{k}|x_0\right) \right]^\beta} \right).$$

Our next proposition establishes the uniform asymptotic behaviour of $\hat{\theta}_n(y_1, y_2)$ in $y_1, y_2 \in [1/2, 2]$, after proper normalization. The proof of the proposition uses a process convergence result for $T_n(y_1, y_2|x_0)$ from Goegebeur et al. (2021a), which is repeated below for completeness.

Theorem 3 (Goegebeur et al., 2021a) *Assume (\mathcal{D}) , (\mathcal{H}) , (\mathcal{K}) , (\mathcal{R}) with $x \rightarrow R(y_1, y_2|x)$ being a continuous function, $x_0 \in \text{Int}(S_X)$ with $f_X(x_0) > 0$, and $y \rightarrow F_j(y|x_0)$, $j = 1, 2$, are strictly increasing. Consider sequences $k \rightarrow \infty$ and $h_n \rightarrow 0$ as $n \rightarrow \infty$, in such a way that $k/n \rightarrow 0$, $kh_n^d \rightarrow \infty$ and $h_n^{\eta_{\gamma_1} \wedge \eta_{\gamma_2} \wedge \eta_{\varepsilon_1} \wedge \eta_{\varepsilon_2}} \ln n/k \rightarrow 0$. Then for $\eta \in [0, 1/2)$, we have,*

$$\sqrt{kh_n^d} \frac{T_n(y_1, y_2|x_0) - \mathbb{E}(T_n(y_1, y_2|x_0))}{y_1^\eta} \rightsquigarrow \frac{W(y_1, y_2)}{y_1^\eta},$$

in $D((0, T]^2)$, for any $T > 0$.

For convenient representation, the limiting process in Theorem 3 will be defined on the same probability space as $T_n(y_1, y_2|x_0)$, via the Skorohod construction, but it should be kept in mind that it is only in distribution equal to the original process. The Skorohod representation theorem gives then (with keeping the same notation)

$$\sup_{y_1, y_2 \in (0, T]} \left| \frac{\sqrt{kh_n^d} [T_n(y_1, y_2|x_0) - \mathbb{E}(T_n(y_1, y_2|x_0))] - W(y_1, y_2)}{y_1^\eta} \right| \rightarrow 0, \text{ a.s. .}$$

Proposition 1 Assume (\mathcal{D}) , (\mathcal{H}) , (\mathcal{K}) , (\mathcal{R}) with $x \rightarrow R(y_1, y_2|x)$ being a continuous function, $x_0 \in \text{Int}(S_X)$ with $f_X(x_0) > 0$, and $y \rightarrow F_j(y|x_0)$, $j = 1, 2$, are strictly increasing. Consider sequences $k \rightarrow \infty$ and $h_n \rightarrow 0$ as $n \rightarrow \infty$, in such a way that $k/n \rightarrow 0$, $kh_n^d \rightarrow \infty$ and $h_n^{\eta_{\gamma_1} \wedge \eta_{\gamma_2} \wedge \eta_{\varepsilon_1} \wedge \eta_{\varepsilon_2}} \ln n/k \rightarrow 0$. Then, for $\gamma_1(x_0) < 1/(2\beta)$, we have

$$\sup_{y_1, y_2 \in [\frac{1}{2}, 2]} \left| \frac{\sqrt{kh_n^d}}{[U_1(n/k|x_0)y_1]^\beta} \left[\hat{\theta}_n(y_1, y_2) - \mathbb{E} \left(\hat{\theta}_n(y_1, y_2) \right) \right] \right. \\ \left. + \int_0^\infty W \left(\left[y_1 \left(1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right) \right]^{-\frac{1}{\gamma_1(x_0)}}, y_2 \right) du^{-\gamma_1(x_0)} \right| \xrightarrow{\mathbb{P}} 0.$$

6.1 Proof of Lemma 2

According to Assumption (\mathcal{D}) , using the inequality $|e^x - 1| \leq |x|e^{|x|}$, we have for some positive constant C

$$\left| \tilde{s}_n(y_1, u|x_0 - h_nv) - \left(y_1 \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right] \right)^{-\frac{1}{\gamma_1(x_0)}} \right| \\ \leq \left| \tilde{s}_n(y_1, u|x_0 - h_nv) - \left(y_1 \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right] \right)^{-\frac{1}{\gamma_1(x_0 - h_nv)}} \right| \\ + C \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}} h_n^{\eta_{\gamma_1}} \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{Ch_n^{\eta_{\gamma_1}}} \ln \left(1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right) \\ \leq C \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}} \times \left\{ h_n^{\eta_{\gamma_1}} \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{Ch_n^{\eta_{\gamma_1}}} \ln \left(1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right) \right. \\ \left. + \left(1 + h_n^{\eta_{\gamma_1}} \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{Ch_n^{\eta_{\gamma_1}}} \ln \left(1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right) \right) \right. \\ \left. \times \left[\frac{A_1(x_0 - h_nv)}{A_1(x_0)} - 1 \right] \left[U_1 \left(\frac{n}{k} | x_0 \right) \right]^{\frac{1}{\gamma_1(x_0)} - \frac{1}{\gamma_1(x_0 - h_nv)}} \right. \\ \left. \times \left| \frac{1 + \frac{1}{\gamma_1(x_0 - h_nv)} \delta_1 \left(U_1 \left(\frac{n}{k} | x_0 \right) y_1 \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right] | x_0 - h_nv \right)}{1 + \frac{1}{\gamma_1(x_0)} \delta_1 \left(U_1 \left(\frac{n}{k} | x_0 \right) | x_0 \right)} \right| \right. \\ \left. + \left[U_1 \left(\frac{n}{k} | x_0 \right) \right]^{\frac{1}{\gamma_1(x_0)} - \frac{1}{\gamma_1(x_0 - h_nv)}} - 1 \right| \times \left| \frac{1 + \frac{1}{\gamma_1(x_0 - h_nv)} \delta_1 \left(U_1 \left(\frac{n}{k} | x_0 \right) y_1 \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right] | x_0 - h_nv \right)}{1 + \frac{1}{\gamma_1(x_0)} \delta_1 \left(U_1 \left(\frac{n}{k} | x_0 \right) | x_0 \right)} \right| \right. \\ \left. + \left| \frac{1 + \frac{1}{\gamma_1(x_0 - h_nv)} \delta_1 \left(U_1 \left(\frac{n}{k} | x_0 \right) y_1 \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right] | x_0 - h_nv \right)}{1 + \frac{1}{\gamma_1(x_0)} \delta_1 \left(U_1 \left(\frac{n}{k} | x_0 \right) | x_0 \right)} - 1 \right| \right\}.$$

Using Assumption (\mathcal{H}) and again the inequality $|e^x - 1| \leq |x|e^{|x|}$, we deduce that, for n large,

$$\left| \frac{A_1(x_0 - h_nv)}{A_1(x_0)} - 1 \right| \leq C h_n^{\eta_{A_1}} \quad (7)$$

$$\left| \left(U_1 \left(\frac{n}{k} | x_0 \right) \right)^{\frac{1}{\gamma_1(x_0)} - \frac{1}{\gamma_1(x_0 - h_nv)}} - 1 \right| \leq C h_n^{\eta_{\gamma_1}} \ln \frac{n}{k}. \quad (8)$$

Now, using the explicit expression for $\delta_1(y|x)$, Assumption (\mathcal{H}) , and Proposition B.1.10 in de Haan and Ferreira (2006) with $0 < \varepsilon < \beta_1(x_0)$, we obtain, for n large, that

$$\begin{aligned}
& \left| \frac{1 + \frac{1}{\gamma_1(x_0 - h_n v)} \delta_1 \left(U_1 \left(\frac{n}{k} \middle| x_0 \right) y_1 \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right] \middle| x_0 - h_n v \right)}{1 + \frac{1}{\gamma_1(x_0)} \delta_1 \left(U_1 \left(\frac{n}{k} \middle| x_0 \right) \middle| x_0 \right)} - 1 \right| \\
& \leq C \left| \delta_1 \left(U_1 \left(\frac{n}{k} \middle| x_0 \right) \middle| x_0 \right) \right| \left\{ \left| \frac{\gamma_1(x_0)}{\gamma_1(x_0 - h_n v)} - 1 \right| \left| \frac{\delta_1 \left(U_1 \left(\frac{n}{k} \middle| x_0 \right) y_1 \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right] \middle| x_0 - h_n v \right)}{\delta_1 \left(U_1 \left(\frac{n}{k} \middle| x_0 \right) \middle| x_0 \right)} \right| \right. \\
& \quad + \left| \frac{\delta_1 \left(U_1 \left(\frac{n}{k} \middle| x_0 \right) y_1 \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right] \middle| x_0 - h_n v \right)}{\delta_1 \left(U_1 \left(\frac{n}{k} \middle| x_0 \right) y_1 \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right] \middle| x_0 \right)} - 1 \right| \left| \frac{\delta_1 \left(U_1 \left(\frac{n}{k} \middle| x_0 \right) y_1 \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right] \middle| x_0 \right)}{\delta_1 \left(U_1 \left(\frac{n}{k} \middle| x_0 \right) \middle| x_0 \right)} \right| \\
& \quad + \left| \frac{\delta_1 \left(U_1 \left(\frac{n}{k} \middle| x_0 \right) y_1 \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right] \middle| x_0 \right)}{\delta_1 \left(U_1 \left(\frac{n}{k} \middle| x_0 \right) \middle| x_0 \right)} - \left(y_1 \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right] \right)^{-\beta_1(x_0)} \right| \\
& \quad \left. + \left| \left(y_1 \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right] \right)^{-\beta_1(x_0)} - 1 \right| \right\} \\
& \leq C \left| \delta_1 \left(U_1 \left(\frac{n}{k} \middle| x_0 \right) \middle| x_0 \right) \right| \times \left[1 + h_n^{\eta_{\varepsilon_1}} \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{Ch_n^{\eta_{\varepsilon_1}}} \left(\ln \frac{n}{k} + \ln \left(1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right) \right) \right], \tag{9}
\end{aligned}$$

uniformly in $y_1 \in [1/2, 2]$ and $v \in S_K$. Combining (7), (8) with (9), the first statement of Lemma 2 follows.

The proof of the second statement of Lemma 2 is similar, although simpler.

6.2 Proof of Lemma 3

We have the decomposition

$$\begin{aligned}
& \sqrt{kh_n^d} \left| \int_0^\infty [R(\tilde{s}_n(y_1, u|x_0 - h_nv), \tilde{t}_n(y_2|x_0 - h_nv)|x_0) \right. \\
& \quad \left. - R\left(\left(y_1 \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}}\right]\right)^{-\frac{1}{\gamma_1(x_0)}}, y_2|x_0\right) du^{-\gamma_1(x_0)} \right| \\
\leq & \sqrt{kh_n^d} \left| \int_0^{\delta_n} [R(\tilde{s}_n(y_1, u|x_0 - h_nv), \tilde{t}_n(y_2|x_0 - h_nv)|x_0) \right. \\
& \quad \left. - R\left(\left(y_1 \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}}\right]\right)^{-\frac{1}{\gamma_1(x_0)}}, y_2|x_0\right) du^{-\gamma_1(x_0)} \right| \\
& + \sqrt{kh_n^d} \left| \int_{\delta_n}^\infty [R(\tilde{s}_n(y_1, u|x_0 - h_nv), \tilde{t}_n(y_2|x_0 - h_nv)|x_0) \right. \\
& \quad \left. - R\left(\left(y_1 \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}}\right]\right)^{-\frac{1}{\gamma_1(x_0)}}, y_2|x_0\right) du^{-\gamma_1(x_0)} \right| \\
\leq & \sqrt{kh_n^d} \left| \int_0^{\delta_n} R(\tilde{s}_n(y_1, u|x_0 - h_nv), \tilde{t}_n(y_2|x_0 - h_nv)|x_0) du^{-\gamma_1(x_0)} \right| \\
& + \sqrt{kh_n^d} \left| \int_0^{\delta_n} R\left(\left(y_1 \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}}\right]\right)^{-\frac{1}{\gamma_1(x_0)}}, y_2|x_0\right) du^{-\gamma_1(x_0)} \right| \\
& - \sqrt{kh_n^d} \int_0^\infty \left| \tilde{s}_n(y_1, u|x_0 - h_nv) - \left(y_1 \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}}\right]\right)^{-\frac{1}{\gamma_1(x_0)}} \right| du^{-\gamma_1(x_0)} \\
& - \sqrt{kh_n^d} \int_{\delta_n}^\infty |\tilde{t}_n(y_2|x_0 - h_nv) - y_2| du^{-\gamma_1(x_0)} \\
=: & T_1 + T_2 + T_3 + T_4,
\end{aligned}$$

where δ_n is a sequence such that $\delta_n \rightarrow 0$.

As for T_1 , since $R(y_1, y_2|x_0) \leq y_1 \wedge y_2$, we have

$$\begin{aligned}
T_1 & \leq -\sqrt{kh_n^d} \int_0^{\delta_n} \tilde{s}_n(y_1, u|x_0 - h_nv) du^{-\gamma_1(x_0)} \\
& \leq -2^{\frac{1}{\gamma_1(x_0)}} \sqrt{kh_n^d} \int_0^{\delta_n} \left(1 + u^{-\frac{\gamma_1(x_0)}{\beta}}\right)^{-\frac{1}{\gamma_1(x_0)}} du^{-\gamma_1(x_0)} \\
& \quad - \sqrt{kh_n^d} \int_0^{\delta_n} \left| \tilde{s}_n(y_1, u|x_0 - h_nv) - \left(y_1 \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}}\right]\right)^{-\frac{1}{\gamma_1(x_0)}} \right| du^{-\gamma_1(x_0)} \\
=: & T_{1,1} + T_{1,2}.
\end{aligned}$$

Clearly,

$$T_{1,1} \leq C \sqrt{kh_n^d} \delta_n^{\frac{1}{\beta} - \gamma_1(x_0)},$$

and, by using the first uniform bound from Lemma 2, and assuming that $h_n^{\eta_{\varepsilon_1} \wedge \eta_{\gamma_1}} |\ln \delta_n| \rightarrow 0$, we have, after tedious computations,

$$T_{1,2} = o\left(\sqrt{kh_n^d} \delta_n^{\frac{1}{\beta} - \gamma_1(x_0)}\right),$$

leading to, for n large,

$$T_1 \leq C \sqrt{kh_n^d} \delta_n^{\frac{1}{\beta} - \gamma_1(x_0)}.$$

The term T_2 is similar, but simpler, and hence

$$T_2 \leq C \sqrt{kh_n^d} \delta_n^{\frac{1}{\beta} - \gamma_1(x_0)}.$$

Concerning T_3 , we can use again the first bound from Lemma 2, and obtain, for n large

$$T_3 \leq C \sqrt{kh_n^d} \left(h_n^{\eta_{A_1}} + h_n^{\eta_{\gamma_1}} \ln \frac{n}{k} + \left| \delta_1 \left(U_1 \left(\frac{n}{k} \middle| x_0 \right) \middle| x_0 \right) \right| \right),$$

while for T_4 we use the second bound from Lemma 2, and obtain, for n large

$$T_4 \leq C \sqrt{kh_n^d} \delta_n^{-\gamma_1(x_0)} \left[h_n^{\eta_{A_2}} + h_n^{\eta_{\gamma_2}} \ln \frac{n}{k} + \left| \delta_2 \left(U_2 \left(\frac{n}{k} \middle| x_0 \right) \middle| x_0 \right) \right| \left(h_n^{\eta_{B_2}} + h_n^{\eta_{\varepsilon_2}} \ln \frac{n}{k} \right) \right].$$

Take $\delta_n = h_n^\xi$. Remember $k = \lfloor n^\alpha \ell_1(n) \rfloor$ and $h_n = n^{-\Delta} \ell_2(n)$, where ℓ_1 and ℓ_2 are slowly varying functions at infinity. Note that $\alpha \in (0, 1)$ by the conditions $k \rightarrow \infty$ and $k/n \rightarrow 0$. The upper bounds for T_1 , T_2 and T_4 , which must tend to zero, lead then to the following bounds for ξ

$$\xi < \min \left(\frac{2\Delta(\eta_{A_2} \wedge \eta_{\gamma_2}) - (\alpha - \Delta d)}{2\Delta\gamma_1(x_0)}, \frac{2\beta_2(x_0)\gamma_2(x_0)(1 - \alpha) + 2\Delta(\eta_{\varepsilon_2} \wedge \eta_{B_2}) - (\alpha - \Delta d)}{2\Delta\gamma_1(x_0)} \right),$$

and

$$\xi > \frac{\alpha - \Delta d}{2\Delta[1/\beta - \gamma_1(x_0)]}.$$

For the parameter Δ , we obtain bounds from the above bounds for ξ , by requiring that the lower bound for ξ is always below the upper bound. This, combined with the bounds coming from the upper bound for T_3 and the condition $kh_n^d \rightarrow \infty$, leads to

$$\max \left(\frac{\alpha}{d + 2(\eta_{A_1} \wedge \eta_{\gamma_1})}, \frac{\alpha - 2\beta_1(x_0)\gamma_1(x_0)(1 - \alpha)}{d}, \frac{\alpha}{d + 2(\eta_{A_2} \wedge \eta_{\gamma_2})[1 - \beta\gamma_1(x_0)]}, \frac{\alpha - 2\beta_2(x_0)\gamma_2(x_0)(1 - \alpha)(1 - \beta\gamma_1(x_0))}{d + 2(\eta_{\varepsilon_2} \wedge \eta_{B_2})[1 - \beta\gamma_1(x_0)]} \right) < \Delta < \frac{\alpha}{d}.$$

So one can first choose $\alpha \in (0, 1)$, followed by selecting Δ satisfying its bounds, and finally the ξ satisfying its bounds.

6.3 Proof of Proposition 1

Note that, using (6), we have

$$\begin{aligned} & \left| \frac{\sqrt{kh_n^d}}{[U_1(n/k|x_0)y_1]^\beta} \left[\hat{\theta}_n(y_1, y_2) - \mathbb{E} \left(\hat{\theta}_n(y_1, y_2) \right) \right] \right. \\ & \quad \left. + \int_0^\infty W \left(\left[y_1 \left(1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right) \right]^{-\frac{1}{\gamma_1(x_0)}}, y_2 \right) du^{-\gamma_1(x_0)} \right| \\ & \leq \left| \int_0^\infty \left[\sqrt{kh_n^d} (T_n(\hat{s}_n(y_1, u|x_0), y_2|x_0) - \mathbb{E}(T_n(\hat{s}_n(y_1, u|x_0), y_2|x_0))) \right. \right. \\ & \quad \left. \left. - W(\hat{s}_n(y_1, u|x_0), y_2) \right] du^{-\gamma_1(x_0)} \right| \\ & \quad + \left| \int_0^\infty \left[W(\hat{s}_n(y_1, u|x_0), y_2) - W \left(\left[y_1 \left(1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right) \right]^{-\frac{1}{\gamma_1(x_0)}}, y_2 \right) \right] du^{-\gamma_1(x_0)} \right| \\ & =: T_5 + T_6. \end{aligned}$$

From Assumption (\mathcal{D}) , one obtains for $n > n_0$ that $\hat{s}_n(y_1, u|x_0) \leq 2^{1+1/\gamma_1(x_0)}$. Hence, for $n > n_0$ and $\eta \in (\gamma_1(x_0)\beta, 1/2)$

$$T_5 \leq \sup_{y_1 \in (0, 2^{1+1/\gamma_1(x_0)}], y_2 \in [1/2, 2]} \frac{\left| \sqrt{kh_n^d} (T_n(y_1, y_2|x_0) - \mathbb{E}(T_n(y_1, y_2|x_0))) - W(y_1, y_2) \right|}{y_1^\eta} \\ \times \sup_{y_1 \in [1/2, 2]} \left(- \int_0^\infty [\hat{s}_n(y_1, u|x_0)]^\eta du^{-\gamma_1(x_0)} \right).$$

Using Proposition B.1.10 in de Haan and Ferreira (2006) one obtains that for $n > n_1$

$$\sup_{y_1 \in [1/2, 2]} \left(- \int_0^\infty [\hat{s}_n(y_1, u|x_0)]^\eta du^{-\gamma_1(x_0)} \right) \leq C \int_0^\infty \left(1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right)^{-\frac{\eta}{\gamma_1(x_0)}} u^{-\gamma_1(x_0)-1} du.$$

The latter integral is finite since $\gamma_1(x_0) < 1/(2\beta)$. Using Theorem 1 in Goegebeur et al. (2021a), see also Theorem 3, we deduce that $T_5 = o_{\mathbb{P}}(1)$.

Now consider T_6 . We have for any $\delta > 0$

$$T_6 \leq \sup_{y_1, y_2 \in [1/2, 2], u \leq \delta} \frac{|W(\hat{s}_n(y_1, u|x_0), y_2)|}{[\hat{s}_n(y_1, u|x_0)]^\eta} \sup_{y_1 \in [1/2, 2]} \left(- \int_0^\delta [\hat{s}_n(y_1, u|x_0)]^\eta du^{-\gamma_1(x_0)} \right) \\ + \sup_{y_1, y_2 \in [1/2, 2], u \leq \delta} \frac{\left| W \left(\left[y_1 \left(1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right) \right]^{-\frac{1}{\gamma_1(x_0)}}, y_2 \right) \right|}{\left[y_1 \left(1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right) \right]^{-\frac{\eta}{\gamma_1(x_0)}}} \\ \times \sup_{y_1 \in [1/2, 2]} \left(- \int_0^\delta \left[y_1 \left(1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right) \right]^{-\frac{\eta}{\gamma_1(x_0)}} du^{-\gamma_1(x_0)} \right) \\ + \sup_{y_1, y_2 \in [1/2, 2]} \left(- \int_\delta^\infty \left| W(\hat{s}_n(y_1, u|x_0), y_2) - W \left(\left[y_1 \left(1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right) \right]^{-\frac{1}{\gamma_1(x_0)}}, y_2 \right) \right| du^{-\gamma_1(x_0)} \right).$$

Note that

$$\sup_{y_1, y_2 \in [1/2, 2], u \leq \delta} \frac{|W(\hat{s}_n(y_1, u|x_0), y_2)|}{[\hat{s}_n(y_1, u|x_0)]^\eta} \leq \sup_{y_1 \in (0, 2^{1+1/\gamma_1(x_0)}], y_2 \in [1/2, 2]} \frac{|W(y_1, y_2)|}{y_1^\eta}, \text{ for } n > n_0,$$

and

$$\sup_{y_1, y_2 \in [1/2, 2], u \leq \delta} \frac{\left| W \left(\left[y_1 \left(1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right) \right]^{-\frac{1}{\gamma_1(x_0)}}, y_2 \right) \right|}{\left[y_1 \left(1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right) \right]^{-\frac{\eta}{\gamma_1(x_0)}}} \leq \sup_{y_1 \in (0, 2^{1+1/\gamma_1(x_0)}], y_2 \in [1/2, 2]} \frac{|W(y_1, y_2)|}{y_1^\eta},$$

these two bounds being finite with probability 1 according to Lemma 2 in Cai et al. (2015). Also, remark that

$$\sup_{y_1 \in [1/2, 2]} \left(- \int_0^\delta [\hat{s}_n(y_1, u|x_0)]^\eta du^{-\gamma_1(x_0)} \right) \leq C \delta^{\eta/\beta - \gamma_1(x_0)} \quad \text{for } n > n_1 \\ = C\varepsilon, \quad \text{by taking } \delta = \varepsilon^{1/(\eta/\beta - \gamma_1(x_0))},$$

and

$$\sup_{y_1 \in [1/2, 2]} \left(- \int_0^\delta \left[y_1 \left(1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right) \right]^{-\frac{\eta}{\gamma_1(x_0)}} du^{-\gamma_1(x_0)} \right) \leq C\varepsilon.$$

Finally,

$$\begin{aligned} & \sup_{y_1, y_2 \in [1/2, 2]} \left(- \int_{\delta}^{\infty} \left| W(\widehat{s}_n(y_1, u|x_0), y_2) - W \left(\left[y_1 \left(1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right) \right]^{-\frac{1}{\gamma_1(x_0)}}, y_2 \right) \right| du^{-\gamma_1(x_0)} \right) \\ & \leq \sup_{y_1, y_2 \in [1/2, 2], u \geq \delta} \left| W(\widehat{s}_n(y_1, u|x_0), y_2) - W \left(\left[y_1 \left(1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right) \right]^{-\frac{1}{\gamma_1(x_0)}}, y_2 \right) \right| \delta^{-\gamma_1(x_0)}. \end{aligned}$$

Since

$$\sup_{y_1 \in [1/2, 2], u > 0} \left| \widehat{s}_n(y_1, u|x_0) - \left[y_1 \left(1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right) \right]^{-\frac{1}{\gamma_1(x_0)}} \right| = o(1),$$

and by using the uniform continuity of W on compact sets, one has for $n > n_2$ that

$$\sup_{y_1, y_2 \in [1/2, 2], u \geq \delta} \left| W(\widehat{s}_n(y_1, u|x_0), y_2) - W \left(\left[y_1 \left(1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right) \right]^{-\frac{1}{\gamma_1(x_0)}}, y_2 \right) \right| \leq \varepsilon.$$

Combining the above one obtains that $T_6 = o_{\mathbb{P}}(1)$. This achieves the proof of Proposition 1.

7 Proof of the main results

7.1 Proof of Lemma 1

We have

$$\begin{aligned} \theta_{\beta, p}(x_0) &= \beta \int_{U_1(1/p|x_0)}^{\infty} \left[y_1 - U_1 \left(\frac{1}{p} | x_0 \right) \right]^{\beta-1} \frac{1}{p} \mathbb{P} \left(Y^{(1)} > y_1, Y^{(2)} > U_2 \left(\frac{1}{p} | x_0 \right) \middle| X = x_0 \right) dy_1 \\ &= \beta \left[U_1 \left(\frac{1}{p} | x_0 \right) \right]^{\beta} \int_1^{\infty} [z - 1]^{\beta-1} \frac{1}{p} \mathbb{P} \left(Y^{(1)} > z U_1 \left(\frac{1}{p} | x_0 \right), Y^{(2)} > U_2 \left(\frac{1}{p} | x_0 \right) \middle| X = x_0 \right) dz. \end{aligned}$$

Note that

$$\lim_{p \downarrow 0} \frac{1}{p} \mathbb{P} \left(Y^{(1)} > z U_1 \left(\frac{1}{p} | x_0 \right), Y^{(2)} > U_2 \left(\frac{1}{p} | x_0 \right) \middle| X = x_0 \right) = R \left(z^{-1/\gamma_1(x_0)}, 1 | x_0 \right),$$

thus we only need to interchange the limit with the integral. To this aim, we can apply the dominated convergence theorem since for $z \geq 1$

$$\begin{aligned} \frac{1}{p} \mathbb{P} \left(Y^{(1)} > z U_1 \left(\frac{1}{p} | x_0 \right), Y^{(2)} > U_2 \left(\frac{1}{p} | x_0 \right) \middle| X = x_0 \right) &\leq \min \left(1, \frac{1}{p} \overline{F}_1 \left(z U_1 \left(\frac{1}{p} | x_0 \right) \middle| x_0 \right) \right) \\ &= \frac{1}{p} \overline{F}_1 \left(z U_1 \left(\frac{1}{p} | x_0 \right) \middle| x_0 \right), \end{aligned}$$

and according to Proposition B.1.9.5 in de Haan and Ferreira (2006), for $z \geq 1, \varepsilon, \delta > 0$, there exists a p_0 such that for $p < p_0$ we have

$$\frac{1}{p} \overline{F}_1 \left(z U_1 \left(\frac{1}{p} | x_0 \right) \middle| x_0 \right) \leq (1 + \delta) z^{-\frac{1}{\gamma_1(x_0)} + \varepsilon}.$$

This implies that, for $p < p_0$,

$$\frac{\theta_{\beta, p}(x_0)}{[U_1(1/p|x_0)]^{\beta}} \leq \beta(1 + \delta) \int_1^{\infty} [z - 1]^{\beta-1} z^{-\frac{1}{\gamma_1(x_0)} + \varepsilon} dz = \beta(1 + \delta) \int_0^1 z^{\beta-1} (1 - z)^{\frac{1}{\gamma_1(x_0)} - \beta - \varepsilon - 1} dz < \infty,$$

as soon as $\varepsilon \in (0, 1/\gamma_1(x_0) - \beta)$.

This yields

$$\frac{\theta_{\beta,p}(x_0)}{[U_1(1/p|x_0)]^\beta} \rightarrow \beta \int_1^\infty [z-1]^{\beta-1} R\left(z^{-1/\gamma_1(x_0)}, 1|x_0\right) dz,$$

from which Lemma 1 follows, after a change of variable.

7.2 Proof of Theorem 1

Our quantity of interest is

$$\begin{aligned} \sqrt{kh_n^d} \left(\frac{\hat{\theta}_n}{\widehat{f}_n(x_0)\theta_{\beta,k/n}(x_0)} - 1 \right) &= \sqrt{kh_n^d} \left(\frac{\hat{\theta}_n}{f_X(x_0)\theta_{\beta,k/n}(x_0)} - 1 \right) \frac{f_X(x_0)}{\widehat{f}_n(x_0)} \\ &\quad - \frac{1}{\widehat{f}_n(x_0)} \sqrt{\frac{k}{n}} \sqrt{nh_n^d} (\widehat{f}_n(x_0) - f_X(x_0)) \\ &= \sqrt{kh_n^d} \left(\frac{\hat{\theta}_n}{f_X(x_0)\theta_{\beta,k/n}(x_0)} - 1 \right) (1 + o_{\mathbb{P}}(1)) + o_{\mathbb{P}}(1). \end{aligned}$$

We have the decomposition

$$\begin{aligned} &\sqrt{kh_n^d} \left(\frac{\hat{\theta}_n}{f_X(x_0)\theta_{\beta,k/n}(x_0)} - 1 \right) \\ &= \sqrt{kh_n^d} (\hat{u}_{n,1}^\beta - 1) \\ &\quad + \frac{[\hat{u}_{n,1} U_1(\frac{n}{k}|x_0)]^\beta}{\theta_{\beta,k/n}(x_0)} \left\{ \frac{\sqrt{kh_n^d}}{f_X(x_0)} \left(\frac{\hat{\theta}_n(\hat{u}_{n,1}, \hat{e}_{n,2})}{[\hat{u}_{n,1} U_1(\frac{n}{k}|x_0)]^\beta} - E_n(\hat{u}_{n,1}, \hat{e}_{n,2}) \right) \right. \\ &\quad \left. + \frac{\sqrt{kh_n^d}}{f_X(x_0)} (E_n(\hat{u}_{n,1}, \hat{e}_{n,2}) - E_n(1, 1)) \right. \\ &\quad \left. + \frac{\sqrt{kh_n^d}}{f_X(x_0)} \left(E_n(1, 1) - \frac{f_X(x_0)\theta_{\beta,k/n}(x_0)}{[U_1(\frac{n}{k}|x_0)]^\beta} \right) \right\} \\ &=: \sqrt{kh_n^d} (\hat{u}_{n,1}^\beta - 1) + \frac{[\hat{u}_{n,1} U_1(\frac{n}{k}|x_0)]^\beta}{\theta_{\beta,k/n}(x_0)} \{T_{1,n} + T_{2,n} + T_{3,n}\}. \end{aligned} \tag{10}$$

This decomposition can be motivated by Proposition 1 and the convergences

$$\begin{aligned} \sqrt{kh_n^d} (\hat{u}_{n,1} - 1) &\rightsquigarrow \frac{\gamma_1(x_0)}{f_X(x_0)} W(1, \infty), \\ \sqrt{kh_n^d} (\hat{e}_{n,2} - 1) &\rightsquigarrow -\frac{W(\infty, 1)}{f_X(x_0)}, \end{aligned}$$

see Lemma 6 and the proof of Theorem 3, respectively, in Goegebeur et al. (2021a).

Then

$$\sqrt{kh_n^d} (\hat{u}_{n,1}^\beta - 1) \rightsquigarrow \frac{\beta\gamma_1(x_0)}{f_X(x_0)} W(1, \infty),$$

and from our Lemma 1, we have

$$\frac{[\hat{u}_{n,1} U_1(\frac{n}{k}|x_0)]^\beta}{\theta_{\beta,k/n}(x_0)} \xrightarrow{\mathbb{P}} -\frac{1}{\int_0^\infty R\left([1+u^{-\frac{\gamma_1(x_0)}{\beta}}]^{-\frac{1}{\gamma_1(x_0)}}, 1|x_0\right) du - \gamma_1(x_0)}.$$

We need to study the three terms $T_{1,n}$, $T_{2,n}$ and $T_{3,n}$.

Concerning $T_{1,n}$, by Proposition 1 and following the lines of proof of Theorem 3 in Goegebeur et al. (2021a) we obtain

$$\frac{[\hat{u}_{n,1}U_1(\frac{n}{k}|x_0)]^\beta}{\theta_{\beta,k/n}(x_0)} T_{1,n} \rightsquigarrow \frac{1}{f_X(x_0)} \frac{\int_0^\infty W([1 + u^{-\frac{\gamma_1(x_0)}{\beta}}]^{-\frac{1}{\gamma_1(x_0)}}, 1) du^{-\gamma_1(x_0)}}{\int_0^\infty R([1 + u^{-\frac{\gamma_1(x_0)}{\beta}}]^{-\frac{1}{\gamma_1(x_0)}}, 1|x_0) du^{-\gamma_1(x_0)}}. \quad (11)$$

Concerning $T_{2,n}$, according to (6)

$$T_{2,n} = \frac{\sqrt{kh_n^d}}{f_X(x_0)} \left\{ \int_0^\infty \mathbb{E}[T_n(\hat{s}_n(1, u|x_0), 1|x_0)] du^{-\gamma_1(x_0)} - \int_0^\infty \mathbb{E}[T_n(\hat{s}_n(\hat{u}_{n,1}, u|x_0), \hat{e}_{n,2}|x_0)] du^{-\gamma_1(x_0)} \right\}.$$

For y_1 either 1 or $\hat{u}_{n,1}$ and y_2 either 1 or $\hat{e}_{n,2}$, we have

$$\begin{aligned} & \int_0^\infty \mathbb{E}[T_n(\hat{s}_n(y_1, u|x_0), y_2|x_0)] du^{-\gamma_1(x_0)} \\ = & \int_0^\infty R\left(y_1^{-\frac{1}{\gamma_1(x_0)}} \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}}\right]^{-\frac{1}{\gamma_1(x_0)}}, y_2|x_0\right) du^{-\gamma_1(x_0)} \int_{S_K} K(v) f_X(x_0 - h_n v) dv \\ & + \int_{S_K} K(v) \int_0^\infty \left[R(\tilde{s}_n(y_1, u|x_0 - h_n v), \tilde{t}_n(y_2|x_0 - h_n v)|x_0) \right. \\ & \quad \left. - R\left(y_1^{-\frac{1}{\gamma_1(x_0)}} \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}}\right]^{-\frac{1}{\gamma_1(x_0)}}, y_2|x_0\right) \right] du^{-\gamma_1(x_0)} f_X(x_0 - h_n v) dv \\ & + \int_0^\infty \int_{S_K} K(v) \left[R_{\frac{n}{k}}(\tilde{s}_n(y_1, u|x_0 - h_n v), \tilde{t}_n(y_2|x_0 - h_n v)|x_0 - h_n v) \right. \\ & \quad \left. - R(\tilde{s}_n(y_1, u|x_0 - h_n v), \tilde{t}_n(y_2|x_0 - h_n v)|x_0 - h_n v) \right] f_X(x_0 - h_n v) dv du^{-\gamma_1(x_0)} \\ & + \int_0^\infty \int_{S_K} K(v) \left[R(\tilde{s}_n(y_1, u|x_0 - h_n v), \tilde{t}_n(y_2|x_0 - h_n v)|x_0 - h_n v) \right. \\ & \quad \left. - R(\tilde{s}_n(y_1, u|x_0 - h_n v), \tilde{t}_n(y_2|x_0 - h_n v)|x_0) \right] f_X(x_0 - h_n v) dv du^{-\gamma_1(x_0)} \\ = & f_X(x_0) \int_0^\infty R\left(y_1^{-\frac{1}{\gamma_1(x_0)}} \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}}\right]^{-\frac{1}{\gamma_1(x_0)}}, y_2|x_0\right) du^{-\gamma_1(x_0)} \\ & + O_{\mathbb{P}}\left(h_n^{\eta_{f_X}}\right) + O_{\mathbb{P}}\left(\left(\frac{n}{k}\right)^\tau\right) + O_{\mathbb{P}}\left(h_n^{\eta_R}\right) + o_{\mathbb{P}}\left(\frac{1}{\sqrt{kh_n^d}}\right), \quad (12) \end{aligned}$$

by Assumptions (S) and (H) combined with Lemma 3.

This implies that

$$\begin{aligned}
T_{2,n} &= \sqrt{kh_n^d} \left\{ \int_0^\infty R \left(\left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}}, 1 \middle| x_0 \right) du^{-\gamma_1(x_0)} \right. \\
&\quad \left. - \int_0^\infty R \left(\frac{\hat{u}_{n,1}^{-\frac{1}{\gamma_1(x_0)}}}{\hat{e}_{n,2}} \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}}, \hat{e}_{n,2} \middle| x_0 \right) du^{-\gamma_1(x_0)} \right\} + o_{\mathbb{P}}(1) \\
&= \sqrt{kh_n^d} \left\{ \hat{e}_{n,2} \int_0^\infty \left[R \left(\left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}}, 1 \middle| x_0 \right) \right. \right. \\
&\quad \left. \left. - R \left(\frac{\hat{u}_{n,1}^{-\frac{1}{\gamma_1(x_0)}}}{\hat{e}_{n,2}} \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}}, 1 \middle| x_0 \right) \right] du^{-\gamma_1(x_0)} \right. \\
&\quad \left. + (1 - \hat{e}_{n,2}) \int_0^\infty R \left(\left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}}, 1 \middle| x_0 \right) du^{-\gamma_1(x_0)} \right\} + o_{\mathbb{P}}(1) \\
&= \hat{e}_{n,2} \sqrt{kh_n^d} \left[H(1) - H \left(\frac{\hat{u}_{n,1}^{-\frac{1}{\gamma_1(x_0)}}}{\hat{e}_{n,2}} \right) \right] + \sqrt{kh_n^d} (1 - \hat{e}_{n,2}) H(1) + o_{\mathbb{P}}(1),
\end{aligned}$$

where

$$H(z) := \int_0^\infty R \left(z \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}}, 1 \middle| x_0 \right) du^{-\gamma_1(x_0)}.$$

Assuming $R(y, 1|x_0)$ differentiable in y with a derivative $R_1(y, 1|x_0)$, we have

$$H'(z) = \int_0^\infty R_1 \left(z \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}}, 1 \middle| x_0 \right) \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}} du^{-\gamma_1(x_0)},$$

by using the Lipschitz property of the R -function combined with the dominated convergence theorem. Then by the mean value theorem

$$T_{2,n} = \left[\sqrt{kh_n^d} (\hat{e}_{n,2} - 1) - \sqrt{kh_n^d} \left(\hat{u}_{n,1}^{-\frac{1}{\gamma_1(x_0)}} - 1 \right) \right] H'(\hat{\xi}_n) - \sqrt{kh_n^d} (\hat{e}_{n,2} - 1) H(1) + o_{\mathbb{P}}(1),$$

where $\hat{\xi}_n$ is an intermediate random value between 1 and $\frac{\hat{u}_{n,1}^{-\frac{1}{\gamma_1(x_0)}}}{\hat{e}_{n,2}}$.

Using again the dominated convergence theorem, we have that $H'(z)$ is continuous at $z = 1$, by the assumed continuity of $R_1(y, 1|x_0)$ and the fact that $R_1(y, 1|x_0)$ is decreasing in y , $R(y, 1|x_0)$ being concave. Then, by the continuous mapping theorem $H'(\hat{\xi}_n) \xrightarrow{\mathbb{P}} H'(1)$.

All in all, we have

$$\begin{aligned}
T_{2,n} &\rightsquigarrow \frac{W(\infty, 1)}{f_X(x_0)} \int_0^\infty R \left(\left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}}, 1 \middle| x_0 \right) du^{-\gamma_1(x_0)} \\
&\quad + \frac{W(1, \infty) - W(\infty, 1)}{f_X(x_0)} \int_0^\infty R_1 \left(\left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}}, 1 \middle| x_0 \right) \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}} du^{-\gamma_1(x_0)},
\end{aligned}$$

from which we deduce that

$$\begin{aligned}
\frac{[\hat{u}_{n,1} U_1(\frac{n}{k} | x_0)]^\beta}{\theta_{\beta, k/n}(x_0)} T_{2,n} &\rightsquigarrow -\frac{W(\infty, 1)}{f_X(x_0)} - \frac{W(1, \infty) - W(\infty, 1)}{f_X(x_0)} \\
&\quad \times \frac{\int_0^\infty R_1 \left(\left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}}, 1 \middle| x_0 \right) \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}} du^{-\gamma_1(x_0)}}{\int_0^\infty R \left(\left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}}, 1 \middle| x_0 \right) du^{-\gamma_1(x_0)}}. \tag{13}
\end{aligned}$$

Now, concerning $T_{3,n}$, we have

$$\begin{aligned}
T_{3,n} &= \frac{\sqrt{kh_n^d}}{f_X(x_0)} \left\{ - \int_0^\infty \mathbb{E} [T_n(\hat{s}_n(1, u|x_0), 1|x_0)] du^{-\gamma_1(x_0)} - \frac{f_X(x_0)\theta_{\beta, k/n}(x_0)}{[U_1(n/k|x_0)]^\beta} \right\} \\
&= \sqrt{kh_n^d} \left\{ - \int_0^\infty R \left(\left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}}, 1|x_0 \right) du^{-\gamma_1(x_0)} \right. \\
&\quad \left. + \int_0^\infty R_{n/k}(\hat{s}_n(1, u|x_0), 1|x_0) du^{-\gamma_1(x_0)} \right\} + o(1) \\
&= \sqrt{kh_n^d} \int_0^\infty [R_{n/k}(\hat{s}_n(1, u|x_0), 1|x_0) - R(\hat{s}_n(1, u|x_0), 1|x_0)] du^{-\gamma_1(x_0)} \\
&\quad + \sqrt{kh_n^d} \int_0^\infty \left[R(\hat{s}_n(1, u|x_0), 1|x_0) - R \left(\left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}}, 1|x_0 \right) \right] du^{-\gamma_1(x_0)} + o(1) \\
&=: T_{3,1,n} + T_{3,2,n} + o(1).
\end{aligned}$$

As for $T_{3,1,n}$ we use Assumption (\mathcal{S}) to obtain

$$\begin{aligned}
|T_{3,1,n}| &\leq \sqrt{kh_n^d} \sup_{x \in B(x_0, h_n)} \sup_{0 < y_1 < \infty, \frac{1}{2} \leq y_2 \leq 2} \frac{|R_{n/k}(y_1, y_2|x) - R(y_1, y_2|x)|}{y_1^\kappa \wedge 1} \\
&\quad \times \left| \int_0^\infty [(\hat{s}_n(1, u|x_0))^\kappa \wedge 1] du^{-\gamma_1(x_0)} \right| \\
&= O \left(\sqrt{kh_n^d} \left(\frac{n}{k} \right)^\tau \right) = o(1),
\end{aligned}$$

by our assumptions on the sequence k and since $\kappa > \beta\gamma_1(x_0)$. Then, for $T_{3,2,n}$ we use the decomposition

$$\begin{aligned}
|T_{3,2,n}| &\leq \sqrt{kh_n^d} \left| \int_0^{T_n} \left[R(\hat{s}_n(1, u|x_0), 1|x_0) - R \left(\left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}}, 1|x_0 \right) \right] du^{-\gamma_1(x_0)} \right| \\
&\quad + \sqrt{kh_n^d} \left| \int_{T_n}^\infty \left[R(\hat{s}_n(1, u|x_0), 1|x_0) - R \left(\left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}}, 1|x_0 \right) \right] du^{-\gamma_1(x_0)} \right|,
\end{aligned}$$

where T_n is a sequence that tends to infinity. Consequently, by using the Lipschitz continuity of $R(\cdot, \cdot|x_0)$ in the first integral and a straightforward bound for the second integral we have

$$\begin{aligned}
|T_{3,2,n}| &\leq C \sqrt{kh_n^d} \int_0^{T_n} u^{-\gamma_1(x_0)-1} \left| \hat{s}_n(1, u|x_0) - \left[1 + u^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}} \right| du \\
&\quad + 2 \sqrt{kh_n^d} \sup_{y_1 > 0} R(y_1, 1|x_0) T_n^{-\gamma_1(x_0)} \\
&= o(1),
\end{aligned}$$

for a sequence T_n such that

$$\sqrt{kh_n^d} T_n^{-\gamma_1(x_0)} \longrightarrow 0, \tag{14}$$

$$\sqrt{kh_n^d} \left| \delta_1 \left(U_1 \left(\frac{n}{k} |x_0 \right) |x_0 \right) \right| T_n^{\frac{1-\gamma_1(x_0)[\beta-\beta_1(x_0)+\varepsilon]}{\beta}} \longrightarrow 0. \tag{15}$$

We have then

$$\frac{[\hat{u}_{n,1} U_1 \left(\frac{n}{k} |x_0 \right)]^\beta}{\theta_{\beta, k/n}(x_0)} T_{3,n} = o_{\mathbb{P}}(1). \tag{16}$$

To conclude the proof we comment on the sequence T_n . Take $T_n = n^\psi$. Then, (14) is satisfied if

$$\psi > \frac{\alpha - \Delta d}{2\gamma_1(x_0)},$$

and (15) is satisfied if

$$\psi < \frac{\beta[2(1 - \alpha)\gamma_1(x_0)\beta_1(x_0) - (\alpha - \Delta d)]}{2[1 - \gamma_1(x_0)(\beta - \beta_1(x_0) + \varepsilon)]}.$$

Requiring that the lower bound for ψ is below its upper bound gives the following constraint on Δ

$$\Delta > \frac{\alpha}{d} - \frac{2\beta(1 - \alpha)\gamma_1^2(x_0)\beta_1(x_0)}{d[1 + \gamma_1(x_0)(\beta_1(x_0) - \varepsilon)]}.$$

From (12), we have the following further conditions

$$\sqrt{kh_n^d} h_n^{\eta_R \wedge \eta_{f_X}} \longrightarrow 0, \quad \sqrt{kh_n^d} \left(\frac{n}{k}\right)^\tau \longrightarrow 0,$$

leading to the additional bounds on Δ

$$\Delta > \max\left(\frac{\alpha}{d + 2(\eta_R \wedge \eta_{f_X})}, \frac{\alpha + 2(1 - \alpha)\tau}{d}\right).$$

Combining (10), (11), (13), (16), Theorem 1 follows.

7.3 Proof of Theorem 2

We use the decomposition

$$\frac{\hat{\theta}_{\beta,p}(x_0)}{\theta_{\beta,p}(x_0)} = \underbrace{\left\{ \left(\frac{k}{np}\right)^{\beta[\hat{\gamma}_{1,k_1}(x_0) - \gamma_1(x_0)]} \right\}}_{T_{\beta,n}^{(1)}} \underbrace{\left\{ \frac{\bar{\theta}_{\beta,\frac{k}{n}}(x_0)}{\theta_{\beta,\frac{k}{n}}(x_0)} \right\}}_{T_{\beta,n}^{(2)}} \underbrace{\left\{ \frac{\left(\frac{k}{np}\right)^{\beta\gamma_1(x_0)} \theta_{\beta,\frac{k}{n}}(x_0)}{\theta_{\beta,p}(x_0)} \right\}}_{T_{\beta,n}^{(3)}},$$

from which we deduce that

$$\begin{aligned} \frac{\hat{\theta}_{\beta,p}(x_0)}{\theta_{\beta,p}(x_0)} - 1 &= \left\{ \left(\frac{k}{np}\right)^{\beta[\hat{\gamma}_{1,k_1}(x_0) - \gamma_1(x_0)]} - 1 \right\} T_{\beta,n}^{(2)} T_{\beta,n}^{(3)} + \left\{ \frac{\bar{\theta}_{\beta,\frac{k}{n}}(x_0)}{\theta_{\beta,\frac{k}{n}}(x_0)} - 1 \right\} T_{\beta,n}^{(3)} \\ &\quad + \left\{ \frac{\left(\frac{k}{np}\right)^{\beta\gamma_1(x_0)} \theta_{\beta,\frac{k}{n}}(x_0)}{\theta_{\beta,p}(x_0)} - 1 \right\}. \end{aligned} \quad (17)$$

We will study the three terms $(T_{\beta,n}^{(i)} - 1)$, $i = 1, 2, 3$, separately.

Using the assumption $\frac{\ln k/(np)}{\sqrt{k_1 h_n^d}} \longrightarrow 0$ combined with Theorem 2.1 in Goegebeur et al. (2021b), we have

$$\frac{\sqrt{k_1 h_n^d}}{\ln k/(np)} \left\{ T_{\beta,n}^{(1)} - 1 \right\} \rightsquigarrow \frac{\beta \gamma_1(x_0)}{f_X(x_0)} \left[\int_0^1 W(z, \infty) \frac{1}{z} dz - W(1, \infty) \right]. \quad (18)$$

The asymptotic behavior of the term $(T_{\beta,n}^{(2)} - 1)$ has been already established in Theorem 1. Now, according to (4), we have

$$\begin{aligned}
T_{\beta,n}^{(3)} - 1 &= \left(\frac{\theta_{\beta,k/n}(x_0)/[U_1(n/k|x_0)]^\beta}{\theta_{\beta,p}(x_0)/[U_1(1/p|x_0)]^\beta} - 1 \right) \frac{[U_1(n/k|x_0)]^\beta (\frac{k}{np})^{\beta\gamma_1(x_0)}}{[U_1(1/p|x_0)]^\beta} + \frac{[U_1(n/k|x_0)]^\beta (\frac{k}{np})^{\beta\gamma_1(x_0)}}{[U_1(1/p|x_0)]^\beta} - 1 \\
&= \left(\frac{\theta_{\beta,k/n}(x_0)/[U_1(n/k|x_0)]^\beta}{\theta_{\beta,p}(x_0)/[U_1(1/p|x_0)]^\beta} - 1 \right) \frac{[U_1(n/k|x_0)]^\beta (\frac{k}{np})^{\beta\gamma_1(x_0)}}{[U_1(1/p|x_0)]^\beta} + \left[\frac{1 + a_1(n/k|x_0)}{1 + a_1(1/p|x_0)} \right]^\beta - 1 \\
&= \left(\frac{\theta_{\beta,k/n}(x_0)/[U_1(n/k|x_0)]^\beta}{\theta_{\beta,p}(x_0)/[U_1(1/p|x_0)]^\beta} - 1 \right) \left(1 + o\left(\frac{1}{\sqrt{kh_n^d}}\right) \right) + o\left(\frac{1}{\sqrt{kh_n^d}}\right),
\end{aligned}$$

since $\sqrt{kh_n^d}|\delta_1(U_1(n/k|x_0)|x_0)| \rightarrow 0$ and $\sqrt{kh_n^d}|\delta_1(U_1(1/p|x_0)|x_0)| \rightarrow 0$ under our assumptions. Remark now that

$$\begin{aligned}
\frac{\theta_{\beta,k/n}(x_0)}{[U_1(n/k|x_0)]^\beta} &= - \int_0^\infty R_{n/k}(\hat{s}_n(1, z|x_0), 1|x_0) dz^{-\gamma_1(x_0)} \\
&= - \int_0^\infty R \left(\left[1 + z^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}}, 1|x_0 \right) dz^{-\gamma_1(x_0)} \\
&\quad - \int_0^\infty [R_{n/k}(\hat{s}_n(1, z|x_0), 1|x_0) - R(\hat{s}_n(1, z|x_0), 1|x_0)] dz^{-\gamma_1(x_0)} \\
&\quad - \int_0^\infty \left[R(\hat{s}_n(1, z|x_0), 1|x_0) - R \left(\left[1 + z^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}}, 1|x_0 \right) \right] dz^{-\gamma_1(x_0)} \\
&=: - \int_0^\infty R \left(\left[1 + z^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}}, 1|x_0 \right) dz^{-\gamma_1(x_0)} + T_{\beta,n}^{(4)} + T_{\beta,n}^{(5)}.
\end{aligned}$$

Note that $T_{\beta,n}^{(4)}$ is similar to $T_{3,1,n}$ and hence $T_{\beta,n}^{(4)} = o(1/\sqrt{kh_n^d})$. Further, $T_{\beta,n}^{(5)}$ is similar to $T_{3,2,n}$ and hence, for a sequence T_n as used for $T_{3,2,n}$, we have $T_{\beta,n}^{(5)} = o(1/\sqrt{kh_n^d})$. Thus

$$\frac{\theta_{\beta,k/n}(x_0)}{[U_1(n/k|x_0)]^\beta} = - \int_0^\infty R \left(\left[1 + z^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}}, 1|x_0 \right) dz^{-\gamma_1(x_0)} + o\left(\frac{1}{\sqrt{kh_n^d}}\right).$$

Similarly, we have

$$\begin{aligned}
\frac{\theta_{\beta,p}(x_0)}{[U_1(1/p|x_0)]^\beta} &= - \int_0^\infty R \left(\left[1 + z^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}}, 1|x_0 \right) dz^{-\gamma_1(x_0)} \\
&\quad - \int_0^\infty \left[R_{1/p} \left(\frac{1}{p} \left[\bar{F}_1 \left(\left[1 + z^{-\frac{\gamma_1(x_0)}{\beta}} \right] U_1 \left(\frac{1}{p} | x_0 \right) | x_0 \right) \right], 1|x_0 \right) \right. \\
&\quad \quad \left. - R \left(\frac{1}{p} \left[\bar{F}_1 \left(\left[1 + z^{-\frac{\gamma_1(x_0)}{\beta}} \right] U_1 \left(\frac{1}{p} | x_0 \right) | x_0 \right) \right], 1|x_0 \right) \right] dz^{-\gamma_1(x_0)} \\
&\quad - \int_0^\infty \left[R \left(\frac{1}{p} \left[\bar{F}_1 \left(\left[1 + z^{-\frac{\gamma_1(x_0)}{\beta}} \right] U_1 \left(\frac{1}{p} | x_0 \right) | x_0 \right) \right], 1|x_0 \right) \right. \\
&\quad \quad \left. - R \left(\left[1 + z^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}}, 1|x_0 \right) \right] dz^{-\gamma_1(x_0)} \\
&=: - \int_0^\infty R \left(\left[1 + z^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}}, 1|x_0 \right) dz^{-\gamma_1(x_0)} + T_{\beta,n}^{(6)} + T_{\beta,n}^{(7)}.
\end{aligned}$$

Clearly $T_{\beta,n}^{(6)} = O(p^{-\tau}) = o\left(\frac{1}{\sqrt{kh_n^d}}\right)$, by our assumptions on k and p .

Finally, similarly as for $T_{\beta,n}^{(5)}$, for T_n a sequence that tends to infinity, we have, for n large,

$$\begin{aligned}
\sqrt{kh_n^d} |T_{\beta,n}^{(7)}| &\leq -\sqrt{kh_n^d} \int_0^{T_n} \left| \frac{1}{p} \left[\bar{F}_1 \left(\left[1 + z^{-\frac{\gamma_1(x_0)}{\beta}} \right] U_1 \left(\frac{1}{p} \middle| x_0 \right) \middle| x_0 \right) \right] - \left[1 + z^{-\frac{\gamma_1(x_0)}{\beta}} \right]^{-\frac{1}{\gamma_1(x_0)}} \right| dz^{-\gamma_1(x_0)} \\
&\quad + 2 \sup_{z_1 \geq 0} R(z_1, 1 | x_0) \sqrt{kh_n^d} T_n^{-\gamma_1(x_0)} \\
&\leq C \sqrt{kh_n^d} \left| \delta_1 \left(U_1 \left(\frac{1}{p} \middle| x_0 \right) \middle| x_0 \right) \right| T_n^{\frac{1-\gamma_1(x_0)[\beta-\beta_1(x_0)+\varepsilon]}{\beta}} + C \sqrt{kh_n^d} T_n^{-\gamma_1(x_0)} \\
&\leq C \sqrt{kh_n^d} \left| \delta_1 \left(U_1 \left(\frac{n}{k} \middle| x_0 \right) \middle| x_0 \right) \right| T_n^{\frac{1-\gamma_1(x_0)[\beta-\beta_1(x_0)+\varepsilon]}{\beta}} + C \sqrt{kh_n^d} T_n^{-\gamma_1(x_0)} \\
&= o(1)
\end{aligned}$$

under the same conditions (14) and (15) on the sequence T_n .

Combining all these results together yield

$$T_{\beta,n}^{(3)} - 1 = o\left(\frac{1}{\sqrt{kh_n^d}}\right). \quad (19)$$

Finally, decomposition (17) combined with Theorem 1, (18) and (19) achieves the proof of Theorem 2.

Acknowledgement

The research of Armelle Guillou was supported by the French National Research Agency under the grant ANR-19-CE40-0013-01/ExtremReg project and an International Emerging Action (IEA-00179). Computation/simulation for the work described in this paper was supported by the DeIC National HPC Centre, SDU.

References

- [1] ACHARYA, V., PEDERSEN, L., PHILIPPON, T., RICHARDSON, M. (2010). Measuring systemic risk. FRB of Cleveland Working Paper No. 10-02. <https://ssrn.com/abstract=1595075>.
- [2] ARTZNER, P., DELBAEN, F., EBER, J-M., HEATH, D. (1999). Coherent measures of risk. *Mathematical Finance* **9** 203–228.
- [3] BEIRLANT, J., JOOSSENS, E., SEGERS, J. (2009). Second-order refined peaks-over-threshold modelling for heavy-tailed distributions. *Journal of Statistical Planning and Inference* **139** 2800–2815.
- [4] BRAZAUKAS, V., JONES, B., PURI, L., ZITIKIS, R. (2008). Estimating conditional tail expectation with actuarial applications in view. *Journal of Statistical Planning and Inference* **128** 3590–3604.
- [5] CAI, J.J., EINMAHL, J.H.J., DE HAAN, L., ZHOU, C. (2015). Estimation of the marginal expected shortfall: the mean when a related variable is extreme. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **77** 417–442.

- [6] CAI, J., TAN, K.S. (2007). Optimal retention for a stop-loss reinsurance under the VaR and CTE risk measures. *Astin Bulletin* **37** 93–112.
- [7] DAOUIA, A., GARDES, L., GIRARD, S. (2013). On kernel smoothing for extremal quantile regression. *Bernoulli* **19** 2557–2589.
- [8] DAOUIA, A., GARDES, L., GIRARD, S., LEKINA, A. (2011). Kernel estimators of extreme level curves. *TEST* **20** 311–333.
- [9] DAS, B., FASEN-HARTMANN, V. (2018). Risk contagion under regular variation and asymptotic tail independence. *Journal of Multivariate Analysis* **165** 194–215.
- [10] DAS, B., FASEN-HARTMANN, V. (2019). Conditional excess risk measures and multivariate regular variation. *Statistics & Risk Modeling* **36** 1–23.
- [11] DI BERNARDINO, E., PRIEUR, C. (2018). Estimation of the multivariate conditional tail expectation for extreme risk levels: Illustration on environmental data sets. *Environmetrics* **29** 1–22.
- [12] EL METHNI, J., GARDES, L., GIRARD, S. (2014). Non-parametric estimation of extreme risk measures from conditional heavy-tailed distributions. *Scandinavian Journal of Statistics* **41** 988–1012.
- [13] EL METHNI, J., GARDES, L., GIRARD, S. (2018). Kernel estimation of extreme regression risk measures. *Electronic Journal of Statistics* **12** 359–398.
- [14] ESCOBAR-BACH, M., GOEGBEUR, Y., GUILLOU, A. (2018a). Local robust estimation of the Pickands dependence function. *Annals of Statistics* **46** 2806–2843.
- [15] ESCOBAR-BACH, M., GOEGBEUR, Y., GUILLOU, A. (2018b). Local estimation of the conditional stable tail dependence function. *Scandinavian Journal of Statistics* **45** 590–617.
- [16] GOEGBEUR, Y., GUILLOU, A., HO, N.K.L., QIN, J. (2021a). Conditional marginal expected shortfall. To appear in *Extremes* <https://doi.org/10.1007/s10687-020-00403-1>.
- [17] GOEGBEUR, Y., GUILLOU, A., HO, N.K.L., QIN, J. (2021b). A Weissman-type estimator of the conditional marginal expected shortfall. To appear in *Econometrics and Statistics* <https://hal.archives-ouvertes.fr/hal-02613135v3>.
- [18] GOEGBEUR, Y., GUILLOU, A., QIN, J. (2019). Bias-corrected estimation for conditional Pareto-type distributions with random right censoring. *Extremes* **22** 459–498.
- [19] DE HAAN, L., FERREIRA, A. (2006). *Extreme value theory, an introduction*. Springer.
- [20] HILL, B.M. (1975). A simple general approach to inference about the tail of a distribution. *Annals of Statistics* **5** 1163–1174.
- [21] JORION, P. (2007). *Value at risk: the new benchmark for managing financial risk*. McGraw-Hill New York.
- [22] PARZEN, E. (1962). On estimation of a probability density function and mode. *Annals of Mathematical Statistics* **33** 1065–1076.
- [23] ROSENBLATT, M. (1956). Remarks on some nonparametric estimates of a density function. *Annals of Mathematical Statistics* **27** 832–837.
- [24] WEISSMAN, I. (1978). Estimation of parameters and large quantiles based on the k largest observations. *Journal of the American Statistical Association* **73** 812–815.

- [25] YAO, Q. (1999). Conditional predictive regions for stochastic processes. Technical report of University of Kent at Canterbury. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.45.2449&rep=rep1&type=pdf>