Local robust and asymptotically unbiased estimation of conditional Pareto-type tails

ONLINE APPENDIX

Goedele Dierckx[(1)], Yuri Goegebeur[(2)] & Armelle Guillou[(3)]

[(1)] Faculty Economics and Management, Hogeschool-Universiteit Brussel, Warmoesberg 26, 1000 Brussel, Belgium

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

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

Proof of Lemma 1

The case $(s,t) = (0,0)$ is trivial. In case $(s,t) \neq (0,0)$, we obtain, using integration by parts,

$$\begin{aligned} m(u_n, s, t; x) &= \bar{F}(u_n; x) \left\{ \int_1^\infty \left[s z^{s-1} (\ln z)^t + t z^{s-1} (\ln z)^{t-1} \right] \bar{G}(z; \gamma(x), \delta(u_n; x), \rho(x)) dz \right. \\ &\quad \left. + \int_1^\infty \left[s z^{s-1} (\ln z)^t + t z^{s-1} (\ln z)^{t-1} \right] \left[\frac{\bar{F}(u_n z; x)}{\bar{F}(u_n; x)} - \bar{G}(z; \gamma(x), \delta(u_n; x), \rho(x)) \right] dz \right\} \\ &=: \bar{F}(u_n; x)(T_1 + T_2). \end{aligned}$$

By application of Taylor's theorem to $\bar{G}$, we have that

$$\begin{aligned} T_1 &= \int_1^\infty \left[s z^{s-1} (\ln z)^t + t z^{s-1} (\ln z)^{t-1} \right] z^{-1/\gamma(x)} dz \\ &\quad - \frac{\delta(u_n; x)}{\gamma(x)} \int_1^\infty \left[s z^{s-1} (\ln z)^t + t z^{s-1} (\ln z)^{t-1} \right] z^{-1/\gamma(x)} \left[1 - z^{\rho(x)/\gamma(x)} \right] dz + o(\delta(u_n; x)) \\ &=: T_{1,1} - \frac{\delta(u_n; x)}{\gamma(x)} T_{1,2} + o(\delta(u_n; x)). \end{aligned}$$

Straightforward integration then gives

$$
\begin{aligned}
T_{1,1} &= \frac{\gamma^t(x)\Gamma(t+1)}{(1-s\gamma(x))^{t+1}}, \\
T_{1,2} &= \gamma^t(x)\Gamma(t+1)\left[\frac{1}{(1-s\gamma(x))^{t+1}} - \frac{1-\rho(x)}{(1-\rho(x)-s\gamma(x))^{t+1}}\right],
\end{aligned}
$$

and thus

$$
T_1 = \gamma^t(x)\Gamma(t+1)\left\{\frac{1}{(1-s\gamma(x))^{t+1}} - \frac{\delta(u_n;x)}{\gamma(x)}\left[\frac{1}{(1-s\gamma(x))^{t+1}} - \frac{1-\rho(x)}{(1-\rho(x)-s\gamma(x))^{t+1}}\right](1+o(1))\right\}.
$$

A slight modification of Proposition 2.3 in Beirlant *et al.* (2009) gives that

$$
\sup_{z\geq 1} z^{1/\gamma(x)}\left|\frac{\bar{F}(u_n z;x)}{\bar{F}(u_n;x)} - \bar{G}(z;\gamma(x),\delta(u_n;x),\rho(x))\right| = o(\delta(u_n;x)), \qquad u_n\to\infty,
$$

and hence $T_2 = o(\delta(u_n;x))$.

Combining the above results establishes Lemma 1.

Proof of Lemma 2

By application of the rule of repeated expectations we obtain

$$
\begin{aligned}
\widetilde{m}_n(K,s,t;x) &= \mathbb{E}[K_{h_n}(x-X)m(u_n,s,t;X)] \\
&= \int_\Omega K(z)m(u_n,s,t;x-h_n z)b(x-h_n z)dz,
\end{aligned}
$$

so, by straightforward calculations,

$$
\begin{aligned}
&|\widetilde{m}_n(K,s,t;x) - b(x)m(u_n,s,t;x)| \\
&\quad\leq m(u_n,s,t;x)\int_\Omega K(z)|b(x-h_n z)-b(x)|dz \\
&\quad +b(x)\int_\Omega K(z)|m(u_n,s,t;x-h_n z)-m(u_n,s,t;x)|dz \\
&\quad +\int_\Omega K(z)|b(x-h_n z)-b(x)||m(u_n,s,t;x-h_n z)-m(u_n,s,t;x)|dz \\
&\quad =: T_3+T_4+T_5.
\end{aligned}
$$

Concerning T_3, by $(\mathcal{B})$ and $(\mathcal{K})$

$$
\begin{aligned}
T_3 &\leq m(u_n,s,t;x)c_b h_n\int_\Omega K(z)d(0,z)dz \\
&= m(u_n,s,t;x)b(x)O(h_n).
\end{aligned}
$$

The term T_4 can be analyzed by invoking $(\mathcal{M})$ and $(\mathcal{K})$ yielding

$$T_4 = m(u_n, s, t; x)b(x)O(\Phi(u_n, h_n; x)).$$

Finally, applying similar arguments to T_5 gives that $T_5 = m(u_n, s, t; x)b(x)O(h_n\Phi(u_n, h_n; x))$, and the result follows.

Proof of Theorem 1

Note that

$$\begin{aligned}
\mathbb{P}_n^{(j)}(s) &= \sqrt{n}\left[\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\sqrt{h_n^p \bar{F}(u_n;x)b(x)}}K\left(\frac{x-X_i}{h_n}\right)\left(\frac{Y_i}{u_n}\right)^s\left(\ln\frac{Y_i}{u_n}\right)^j \mathbf{1}\{Y_i > u_n\}\right. \\
&\left. -\mathbb{E}\left(\frac{1}{\sqrt{h_n^p \bar{F}(u_n;x)b(x)}}K\left(\frac{x-X}{h_n}\right)\left(\frac{Y}{u_n}\right)^s\left(\ln\frac{Y}{u_n}\right)^j \mathbf{1}\{Y > u_n\}\right)\right], \quad j \in J.
\end{aligned}$$

As such, the empirical processes under consideration fit in the framework of Section 19.5 in van der Vaart (2007) on changing function classes. Indeed, we can consider the classes $\mathcal{W}_n^{(j)} := \{w_{n,s}^{(j)}; s \in [S,0]\}$, where

$$w_{n,s}^{(j)}(v,y) := \frac{1}{\sqrt{h_n^p \bar{F}(u_n;x)b(x)}}K\left(\frac{x-v}{h_n}\right)\left(\frac{y}{u_n}\right)^s\left(\ln\frac{y}{u_n}\right)^j \mathbf{1}\{y > u_n\}, \quad j \in J.$$

So, for the marginal convergence of the processes, it is sufficient to verify the conditions of Theorem 19.28 in van der Vaart (2007).

First, by Lemmas 1 and 2

$$\begin{aligned}
&\mathbb{E}[w_{n,s}^{(j)}(X,Y) - w_{n,t}^{(j)}(X,Y)]^2 \\
&= \frac{\|K\|_2^2}{\bar{F}(u_n;x)b(x)}\mathbb{E}\left(\frac{1}{h_n^p\|K\|_2^2}K^2\left(\frac{x-X}{h_n}\right)\left[\left(\frac{Y}{u_n}\right)^s - \left(\frac{Y}{u_n}\right)^t\right]^2\left(\ln\frac{Y}{u_n}\right)^{2j}\mathbf{1}\{Y > u_n\}\right) \\
&\leq \frac{\|K\|_2^2(s-t)^2}{\bar{F}(u_n;x)b(x)}\mathbb{E}\left(\frac{1}{h_n^p\|K\|_2^2}K^2\left(\frac{x-X}{h_n}\right)\left(\ln\frac{Y}{u_n}\right)^{2(j+1)}\mathbf{1}\{Y > u_n\}\right) \\
&= (2(j+1))!\gamma^{2(j+1)}(x)(s-t)^2\|K\|_2^2(1+o(1)).
\end{aligned}$$

Note that the $o(1)$ term above does not depend on s and t, and therefore

$$\begin{aligned}\sup_{|s-t|<\delta_n} \mathbb{E}[w_{n,s}^{(j)}(X,Y) - w_{n,t}^{(j)}(X,Y)]^2 &\leq (2(j+1))!\gamma^{2(j+1)}(x)\delta_n^2\|K\|_2^2(1+o(1)) \\ &\to 0,\end{aligned}$$

for every sequence $\delta_n \downarrow 0$.

Next we verify the Lindeberg condition. Note that the envelope function $W_n^{(j)}$ for $\mathcal{W}_n^{(j)}$ can be taken as

$$W_n^{(j)}(v,y) = \frac{1}{\sqrt{h_n^p \bar{F}(u_n;x)b(x)}} K\left(\frac{x-v}{h_n}\right)\left(\ln\frac{y}{u_n}\right)^j \mathbf{1}\{y > u_n\}, \;\; j \in J.$$

Using Lemmas 1 and 2, we then have

$$\begin{aligned}\mathbb{E}\left[(W_n^{(j)}(X,Y))^2\right] &= \frac{\|K\|_2^2}{\bar{F}(u_n;x)b(x)}\mathbb{E}\left[\frac{1}{h_n^p\|K\|_2^2}K^2\left(\frac{x-X}{h_n}\right)\left(\ln\frac{Y}{u_n}\right)^{2j}\mathbf{1}\{Y > u_n\}\right] \\ &= \gamma^{2j}(x)(2j)!\|K\|_2^2(1+o(1)) = O(1),\end{aligned}$$

and, for every $\varepsilon, \alpha > 0$,

$$\begin{aligned}&\mathbb{E}\left[(W_n^{(j)}(X,Y))^2\mathbf{1}\{W_n^{(j)}(X,Y) > \varepsilon\sqrt{n}\}\right] \\ &\quad\leq \frac{1}{\varepsilon^\alpha n^{\alpha/2}}\mathbb{E}\left[(W_n^{(j)}(X,Y))^{2+\alpha}\right] \\ &\quad= \frac{\|K^{2+\alpha}\|_1}{\varepsilon^\alpha(nh_n^p\bar{F}(u_n;x)b(x))^{\alpha/2}}\frac{1}{\bar{F}(u_n;x)b(x)}\mathbb{E}\left[\frac{1}{h_n^p\|K^{2+\alpha}\|_1}K^{2+\alpha}\left(\frac{x-X}{h_n}\right)\left(\ln_+\frac{Y}{u_n}\right)^{j(2+\alpha)}\mathbf{1}\{Y > u_n\}\right] \\ &\quad= O\left(\frac{1}{(nh_n^p\bar{F}(u_n;x))^{\alpha/2}}\right) \to 0,\end{aligned}$$

if $nh_n^p\bar{F}(u_n;x) \to \infty$, $j \in J$.

Thirdly, we verify the condition on the bracketing integrals $J_{[\,]}(\delta_n, \mathcal{W}_n^{(j)}, L_2(\mathbb{P}))$, $j \in J$, in Theorem 19.28 of van der Vaart (2007). We have that

$$\begin{aligned}|w_{n,s}^{(j)}(v,y) - w_{n,t}^{(j)}(v,y)| &\leq \frac{|s-t|}{\sqrt{h_n^p\bar{F}(u_n;x)b(x)}}K\left(\frac{x-v}{h_n}\right)\left(\ln\frac{y}{u_n}\right)^{j+1}\mathbf{1}\{y > u_n\}, \\ &=: |s-t|w^{(j)}(v,y).\end{aligned}$$

Note

$$\mathbb{E}\left[\left(w^{(j)}(X,Y)\right)^2\right] = \gamma^{2(j+1)}(x)(2(j+1))!\|K\|_2^2(1+o(1)), \;\; j \in J.$$

So that the condition on $J_{[\,]}(\delta_n, \mathcal{W}_n^{(j)}, L_2(\mathbb{P}))$, $j \in J$, is easy to verify using the result of Example 19.7 in van der Vaart (2007) on parametric function classes.

Finally, we comment on the pointwise convergence of the covariance functions on $[S,0]^2$. For $(s_1,s_2) \in [S,0]^2$ we have that

$$\begin{aligned}
&\mathbb{C}ov(\mathbb{P}_n^{(j)}(s_1), \mathbb{P}_n^{(j)}(s_2)) \\
&= \mathbb{C}ov(w_{n,s_1}^{(j)}(X,Y), w_{n,s_2}^{(j)}(X,Y)) \\
&= \frac{\|K\|_2^2}{\bar{F}(u_n;x)b(x)}\mathbb{E}\left[\frac{1}{h_n^p\|K\|_2^2}K^2\left(\frac{x-X}{h_n}\right)\left(\frac{Y}{u_n}\right)^{s_1+s_2}\left(\ln\frac{Y}{u_n}\right)^{2j}\mathbf{1}\{Y>u_n\}\right] \\
&\quad -\frac{h_n^p}{\bar{F}(u_n;x)b(x)}\mathbb{E}\left[K_{h_n}(x-X)\left(\frac{Y}{u_n}\right)^{s_1}\left(\ln\frac{Y}{u_n}\right)^{j}\mathbf{1}\{Y>u_n\}\right]\times \\
&\quad \mathbb{E}\left[K_{h_n}(x-X)\left(\frac{Y}{u_n}\right)^{s_2}\left(\ln\frac{Y}{u_n}\right)^{j}\mathbf{1}\{Y>u_n\}\right] \\
&\to \frac{\gamma^{2j}(x)(2j)!\|K\|_2^2}{[1-(s_1+s_2)\gamma(x)]^{1+2j}}, \;\; n\to\infty; \;\; j\in J.
\end{aligned}$$

The joint convergence of the empirical processes follows then from the fact that the coordinate classes being Donsker is equivalent to the union of the coordinate classes being Donsker, see van der Vaart p. 270. The pointwise convergence of the covariances between the processes $\mathbb{P}_n^{(j)}$, $j \in J$, can be established along the same line of arguments as above.

Proof of Theorem 2

To prove the existence and consistency of $(\hat{\gamma}_n(x), \hat{\delta}_n(x))$ we adapt the proof of Theorem 5.1 in Chapter 6 of Lehmann and Casella (1998), where existence and consistency of solutions of the likelihood equations is established, to the MDPDE framework. Let Q_r denote the sphere centered at $(\gamma_0(x), 0)$ and radius r, and let $\widehat{\Delta}_\alpha(\gamma,\delta;\rho)$ denote the density power divergence objective function. Note that r should be such that Q_r is a subset of the parameter space. First we rescale $\widehat{\Delta}_\alpha(\gamma,\delta;\rho)$ as $\widetilde{\Delta}_\alpha(\gamma,\delta;\rho) := \widehat{\Delta}_\alpha(\gamma,\delta;\rho)/(\bar{F}(u_n;x)b(x))$, and we show that for any r sufficiently small

$$\mathbb{P}_{(\gamma_0(x),0)}(\widetilde{\Delta}_\alpha(\gamma_0(x),0;\rho_0(x)) < \widetilde{\Delta}_\alpha(\gamma,\delta;\rho_0(x)) \text{ for all } (\gamma,\delta) \text{ on the surface of } Q_r) \to 1.$$

Let $f_s(\gamma,\delta;\rho_0(x))$, $s=1,2$, denote the derivatives of $\widetilde{\Delta}_\alpha(\gamma,\delta;\rho_0(x))$ with respect to γ and δ, respectively, without the common scale factor $1+\alpha$. Similarly, f_{st} and f_{stu}, $s,t,u=1,2$, denote the second and third order derivatives, respectively (again apart from the common scaling by $1+\alpha$).

By Taylor's theorem

$$\begin{aligned}
&\widetilde{\Delta}_\alpha(\gamma,\delta;\rho_0(x)) - \widetilde{\Delta}_\alpha(\gamma_0(x),0;\rho_0(x)) \\
&= (1+\alpha)\{f_1(\gamma_0(x),0;\rho_0(x))(\gamma-\gamma_0(x)) + f_2(\gamma_0(x),0;\rho_0(x))\delta \\
&\quad + \frac{1}{2}\left[f_{11}(\gamma_0(x),0;\rho_0(x))(\gamma-\gamma_0(x))^2 + f_{22}(\gamma_0(x),0;\rho_0(x))\delta^2 + 2f_{12}(\gamma_0(x),0;\rho_0(x))(\gamma-\gamma_0(x))\delta\right] \\
&\quad + \frac{1}{6}\Big[f_{111}(\tilde{\gamma},\tilde{\delta};\rho_0(x))(\gamma-\gamma_0(x))^3 + f_{222}(\tilde{\gamma},\tilde{\delta};\rho_0(x))\delta^3 + 3f_{112}(\tilde{\gamma},\tilde{\delta};\rho_0(x))(\gamma-\gamma_0(x))^2\delta \\
&\quad + 3f_{122}(\tilde{\gamma},\tilde{\delta};\rho_0(x))(\gamma-\gamma_0(x))\delta^2\Big]\Big\} \qquad \text{(A-1)} \\
&=: (1+\alpha)\{S_1+S_2+S_3\},
\end{aligned}$$

where $(\tilde{\gamma},\tilde{\delta})$ is a point on the line segment connecting (γ,δ) and $(\gamma_0(x),0)$. After some tedious, but straightforward derivations one obtains

$$\begin{aligned}
&f_1(\gamma_0(x),0;\rho_0(x)) \\
&= \gamma_0^{-\alpha-2}(x)\Big[-\frac{\alpha\gamma_0(x)(1+\gamma_0(x))}{[1+\alpha(1+\gamma_0(x))]^2}\frac{T_n(K,0,0;x)}{\bar{F}(u_n;x)b(x)} \\
&\quad + \gamma_0(x)\frac{T_n(K,-\alpha(1+\gamma_0(x))/\gamma_0(x),0;x)}{\bar{F}(u_n;x)b(x)} - \frac{T_n(K,-\alpha(1+\gamma_0(x))/\gamma_0(x),1;x)}{\bar{F}(u_n;x)b(x)}\Big], \\
&f_2(\gamma_0(x),0;\rho_0(x)) \\
&= \gamma_0^{-\alpha-1}(x)\Big[-\frac{\alpha\rho_0(x)(1+\gamma_0(x))}{[1+\alpha(1+\gamma_0(x))][1-\rho_0(x)+\alpha(1+\gamma_0(x))]}\frac{T_n(K,0,0;x)}{\bar{F}(u_n;x)b(x)} \\
&\quad + \frac{T_n(K,-\alpha(1+\gamma_0(x))/\gamma_0(x),0;x)}{\bar{F}(u_n;x)b(x)} - (1-\rho_0(x))\frac{T_n(K,-(\alpha(1+\gamma_0(x))-\rho_0(x))/\gamma_0(x),0;x)}{\bar{F}(u_n;x)b(x)}\Big].
\end{aligned}$$

By using the results of Lemmas 1, 2 and Theorem 1, we have that $f_1(\gamma_0(x),0;\rho_0(x)) \xrightarrow{\mathbb{P}} 0$ and $f_2(\gamma_0(x),0;\rho_0(x)) \xrightarrow{\mathbb{P}} 0$, so, for any given $r>0$ we have that $|f_1(\gamma_0(x),0;\rho_0(x))| < r^2$ and

$|f_2(\gamma_0(x),0;\rho_0(x))| < r^2$ with probability tending to 1, and hence, on Q_r, $|S_1| < 2r^3$ with probability tending to 1.

We now focus on the second order derivatives appearing in S_2. Again, by tedious calculus one obtains

$$\begin{aligned}
&f_{11}(\gamma_0(x),0;\rho_0(x))\\
&= \gamma_0^{-\alpha-2}(x)\Bigg[\left(\frac{\alpha+2}{1+\alpha(1+\gamma_0(x))}-\frac{2\alpha+4}{[1+\alpha(1+\gamma_0(x))]^2}+\frac{2\alpha+2}{[1+\alpha(1+\gamma_0(x))]^3}\right)\frac{T_n(K,0,0;x)}{\bar F(u_n;x)b(x)}\\
&\quad -(\alpha+1)\frac{T_n(K,-\alpha(1+\gamma_0(x))/\gamma_0(x),0;x)}{\bar F(u_n;x)b(x)}+\frac{2\alpha+2}{\gamma_0(x)}\frac{T_n(K,-\alpha(1+\gamma_0(x))/\gamma_0(x),1;x)}{\bar F(u_n;x)b(x)}\\
&\quad -\frac{\alpha}{\gamma_0^2(x)}\frac{T_n(K,-\alpha(1+\gamma_0(x))/\gamma_0(x),2;x)}{\bar F(u_n;x)b(x)}\Bigg],\\
&f_{12}(\gamma_0(x),0;\rho_0(x))\\
&= \gamma_0^{-\alpha-2}(x)\Bigg[\Bigg(\frac{1+\alpha(2+\alpha)(1+\gamma_0(x))}{[1+\alpha(1+\gamma_0(x))]^2}\\
&\quad -\frac{(1-\rho_0(x))^2-\alpha[\rho_0(x)(1-\rho_0(x))-2(1+\gamma_0(x))(1-\rho_0(x))]+\alpha^2(1+\gamma_0(x))(1-\rho_0(x))}{[1-\rho_0(x)+\alpha(1+\gamma_0(x))]^2}\Bigg)\\
&\quad \times\frac{T_n(K,0,0;x)}{\bar F(u_n;x)b(x)}-(1+\alpha)\frac{T_n(K,-\alpha(1+\gamma_0(x))/\gamma_0(x),0;x)}{\bar F(u_n;x)b(x)}\\
&\quad +(\alpha+1)(1-\rho_0(x))\frac{T_n(K,-(\alpha(1+\gamma_0(x))-\rho_0(x))/\gamma_0(x),0;x)}{\bar F(u_n;x)b(x)}\\
&\quad +\frac{\alpha}{\gamma_0(x)}\frac{T_n(K,-\alpha(1+\gamma_0(x))/\gamma_0(x),1;x)}{\bar F(u_n;x)b(x)}\\
&\quad -\frac{(\alpha-\rho_0(x))(1-\rho_0(x))}{\gamma_0(x)}\frac{T_n(K,-(\alpha(1+\gamma_0(x))-\rho_0(x))/\gamma_0(x),1;x)}{\bar F(u_n;x)b(x)}\Bigg],\\
&f_{22}(\gamma_0(x),0;\rho_0(x))\\
&= \gamma_0^{-\alpha-2}(x)\Bigg[\Bigg(\frac{1+\alpha+\gamma_0(x)}{1+\alpha(1+\gamma_0(x))}-\frac{2(1-\rho_0(x))(1+\gamma_0(x)+\alpha)}{1-\rho_0(x)+\alpha(1+\gamma_0(x))}\\
&\quad +\frac{(1+\gamma_0(x))(1-2\rho_0(x))+\alpha(1-\rho_0(x))^2}{1-2\rho_0(x)+\alpha(1+\gamma_0(x))}\Bigg)\frac{T_n(K,0,0;x)}{\bar F(u_n;x)b(x)}\\
&\quad -(\alpha+\gamma_0(x))\frac{T_n(K,-\alpha(1+\gamma_0(x))/\gamma_0(x),0;x)}{\bar F(u_n;x)b(x)}\\
&\quad +2(1-\rho_0(x))(\alpha+\gamma_0(x))\frac{T_n(K,-(\alpha(1+\gamma_0(x))-\rho_0(x))/\gamma_0(x),0;x)}{\bar F(u_n;x)b(x)}\\
&\quad -[(1+\gamma_0(x))(1-2\rho_0(x))+(\alpha-1)(1-\rho_0(x))^2]\frac{T_n(K,-(\alpha(1+\gamma_0(x))-2\rho_0(x))/\gamma_0(x),0;x)}{\bar F(u_n;x)b(x)}\Bigg].
\end{aligned}$$

Now, let $f^*_{st}(\gamma_0(x),0;\rho_0(x))$ denote the limits of the random terms $f_{st}(\gamma_0(x),0;\rho_0(x))$, $s,t=1,2$.

These can be obtained from the results of Lemmas 1, 2 and Theorem 1, and are given by

$$f^*_{11}(\gamma_0(x),0;\rho_0(x)) = \gamma_0^{-\alpha-2}(x)\,\frac{1+\alpha^2(1+\gamma_0(x))^2}{[1+\alpha(1+\gamma_0(x))]^3},$$

$$\begin{aligned} &f^*_{12}(\gamma_0(x),0;\rho_0(x)) \\ &\quad = \gamma_0^{-\alpha-2}(x)\,\frac{\rho_0(x)(1-\rho_0(x))[1+\alpha(1+\gamma_0(x))+\alpha^2(1+\gamma_0(x))^2]+\alpha^3\rho_0(x)(1+\gamma_0(x))^3}{[1+\alpha(1+\gamma_0(x))]^2[1-\rho_0(x)+\alpha(1+\gamma_0(x))]^2}, \\ &f^*_{22}(\gamma_0(x),0;\rho_0(x)) \\ &\quad = \gamma_0^{-\alpha-2}(x)\,\frac{(1-\rho_0(x))\rho_0^2(x)+\alpha\rho_0^2(x)(1+\gamma_0(x))[\alpha(1+\gamma_0(x))-\rho_0(x)]}{[1+\alpha(1+\gamma_0(x))][1-\rho_0(x)+\alpha(1+\gamma_0(x))][1-2\rho_0(x)+\alpha(1+\gamma_0(x))]}. \end{aligned}$$

Now, write

$$\begin{aligned} 2S_2 \;=\; & f^*_{11}(\gamma_0(x),0;\rho_0(x))(\gamma-\gamma_0(x))^2+f^*_{22}(\gamma_0(x),0;\rho_0(x))\delta^2+2f^*_{12}(\gamma_0(x),0;\rho_0(x))(\gamma-\gamma_0(x))\delta \\ & +[f_{11}(\gamma_0(x),0;\rho_0(x))-f^*_{11}(\gamma_0(x),0;\rho_0(x))](\gamma-\gamma_0(x))^2 \\ & +[f_{22}(\gamma_0(x),0;\rho_0(x))-f^*_{22}(\gamma_0(x),0;\rho_0(x))]\delta^2 \\ & +2[f_{12}(\gamma_0(x),0;\rho_0(x))-f^*_{12}(\gamma_0(x),0;\rho_0(x))](\gamma-\gamma_0(x))\delta. \end{aligned}$$

Note that the first three terms are in fact a nonrandom positive definite quadratic form in $(\gamma-\gamma_0(x))$ and δ. This can be verified analytically, but the result is not included in the appendix. By the spectral decomposition this quadratic form can be rewritten as $\lambda_1\xi_1^2+\lambda_2\xi_2^2$, where $0<\lambda_1\le\lambda_2$ are the eigenvalues and ξ_1 and ξ_2 are orthogonal transformations of $(\gamma-\gamma_0(x))$ and δ. Note that in this new coordinate system Q_r becomes $\xi_1^2+\xi_2^2=r^2$. Thus, for the quadratic form we have that $\lambda_1\xi_1^2+\lambda_2\xi_2^2\ge\lambda_1(\xi_1^2+\xi_2^2)=\lambda_1 r^2$. For the random part of S_2 we know from Lemmas 1, 2 and Theorem 1 that $f_{st}(\gamma_0(x),0;\rho_0(x))\xrightarrow{\mathbb{P}} f^*_{st}(\gamma_0(x),0;\rho_0(x))$, $s,t=1,2$, and thus in absolute value the random part is less than $4r^3$ with probability tending to 1. Overall, we have that there exists $c>0$ and $r_0>0$ such that for $r<r_0$

$$S_2 > cr^2$$

with probability tending to 1.

For the term S_3, one can show that $|f_{stu}(\gamma,\delta;\rho_0(x))|\le M_{stu}(\boldsymbol{V})$, where $\boldsymbol{V}:=[(X_1,Y_1),\ldots,(X_n,Y_n)]$, for $(\gamma,\delta)\in Q_r$, with $M_{stu}(\boldsymbol{V})\xrightarrow{\mathbb{P}} m_{stu}$, $s,t,u=1,2$, which is bounded. The derivations are straightforward, and are for brevity omitted from the appendix. Thus, with probability tending to 1, $|f_{stu}(\tilde{\gamma}, \tilde{\delta}; \rho_0(x))| < 2m_{stu}$, and hence $|S_3| < er^3$ on Q_r, where

$$e := \frac{1}{3} \sum_{s=1}^{2} \sum_{t=1}^{2} \sum_{u=1}^{2} m_{stu}.$$

Combining the above we find that with probability tending to 1,

$$\min(S_1 + S_2 + S_3) > cr^2 - (2 + e)r^3,$$

where the minimum is over (γ, δ) on the surface of Q_r. Clearly, the right-hand side of the above inequality is positive if $r < c/(2 + e)$.

To complete the proof of the existence and consistency we adjust the line of argumentation of Theorem 3.7 in Chapter 6 of Lehmann and Casella (1998). For $r > 0$, small enough such that Q_r is a subset of the parameter space, consider

$$S_n(r) := \{\boldsymbol{v} : \widetilde{\Delta}_\alpha(\gamma_0(x), 0; \rho_0(x)) < \widetilde{\Delta}_\alpha(\gamma, \delta; \rho_0(x)) \text{ for all } (\gamma, \delta) \text{ on the surface of } Q_r\}.$$

From the above we have that $\mathbb{P}_{(\gamma_0(x),0)}(S_n(r)) \to 1$ for any such r, and hence there exists a sequence $r_n^* \downarrow 0$ such that $\mathbb{P}_{(\gamma_0(x),0)}(S_n(r_n^*)) \to 1$ as $n \to \infty$. By the differentiability of $\widetilde{\Delta}_\alpha(\gamma, \delta; \rho_0(x))$ we have that $\boldsymbol{v} \in S_n(r_n^*)$ implies that there exists a point $(\hat{\gamma}_n(r_n^*), \hat{\delta}_n(r_n^*)) \in Q_{r_n^*}$ for which $\widetilde{\Delta}_\alpha(\gamma, \delta; \rho_0(x))$ attains a local minimum, and thus $f_s(\hat{\gamma}_n(r_n^*), \hat{\delta}_n(r_n^*); \rho_0(x)) = 0$, $s = 1, 2$. Now let $(\hat{\gamma}_n^*(x), \hat{\delta}_n^*(x)) := (\hat{\gamma}_n(r_n^*), \hat{\delta}_n(r_n^*))$ for $\boldsymbol{v} \in S_n(r_n^*)$ and arbitrary otherwise. Clearly

$$\mathbb{P}_{(\gamma_0(x),0)}(f_1(\hat{\gamma}_n^*(x), \hat{\delta}_n^*(x); \rho_0(x)) = 0, f_2(\hat{\gamma}_n^*(x), \hat{\delta}_n^*(x); \rho_0(x)) = 0) \geq \mathbb{P}_{(\gamma_0(x),0)}(S_n(r_n^*)) \to 1,$$

as $n \to \infty$. Thus with probability tending to 1 there exists a sequence of solutions to the estimating equations (7) and (8). Also, for any fixed $r > 0$ and n sufficiently large

$$\begin{aligned} \mathbb{P}_{(\gamma_0(x),0)}(d((\hat{\gamma}_n^*(x), \hat{\delta}_n^*(x)), (\gamma_0(x), 0)) < r) &\geq \mathbb{P}_{(\gamma_0(x),0)}(d((\hat{\gamma}_n^*(x), \hat{\delta}_n^*(x)), (\gamma_0(x), 0)) < r_n^*) \\ &\geq \mathbb{P}_{(\gamma_0(x),0)}(S_n(r_n^*)) \to 1, \end{aligned}$$

which establishes the consistency of the sequence $(\hat{\gamma}_n^*(x), \hat{\delta}_n^*(x))$.

Proof of Corollary 1

We have that

$$\mathbb{S}_n^{(j)}(s) = \mathbb{P}_n^{(j)}(s) + r_n \left[\mathbb{E}\left(\frac{T_n(K,s,j;x)}{\bar{F}(u_n;x)b(x)} \right) - \frac{j!\gamma_0^j(x)}{[1-s\gamma_0(x)]^{1+j}} \right], \quad j \in J.$$

From Lemmas 1 and 2

$$\begin{aligned} r_n & \left[\mathbb{E}\left(\frac{T_n(K,s,j;x)}{\bar{F}(u_n;x)b(x)} \right) - \frac{j!\gamma_0^j(x)}{[1-s\gamma_0(x)]^{1+j}} \right] \\ &= -\lambda\sqrt{b(x)} j! \gamma_0^{j-1}(x) \left[\frac{1}{[1-s\gamma_0(x)]^{j+1}} - \frac{1-\rho_0(x)}{[1-\rho_0(x)-s\gamma_0(x)]^{j+1}} \right] + o(1), \quad j \in J, \end{aligned}$$

where the $o(1)$ terms are uniform in $s \in [S, 0]$.

Proof of Theorem 3

To start we establish the joint limiting distribution of the random terms appearing in $f_s(\gamma_0(x), 0; \rho_0(x))$, $s = 1, 2$, when appropriately normalized. Let

$$\mathbb{T}_n := \frac{1}{\bar{F}(u_n;x)b(x)} \begin{bmatrix} T_n(K,0,0;x) \\ T_n(K,-\alpha(1+\gamma_0(x))/\gamma_0(x),0;x) \\ T_n(K,-(\alpha(1+\gamma_0(x))-\rho_0(x))/\gamma_0(x),0;x) \\ T_n(K,-\alpha(1+\gamma_0(x))/\gamma_0(x),1;x) \end{bmatrix},$$

$$\widetilde{\mathbb{T}} := \begin{bmatrix} 1 \\ \frac{1}{1+\alpha(1+\gamma_0(x))} \\ \frac{1}{1-\rho_0(x)+\alpha(1+\gamma_0(x))} \\ \frac{\gamma_0(x)}{[1+\alpha(1+\gamma_0(x))]^2} \end{bmatrix},$$

and set $\overline{\mathbb{A}}_n(\rho_0(x)) := r_n[\mathbb{T}_n - \widetilde{\mathbb{T}}]$. Thus, from Corollary 1, we get that

$$\overline{\mathbb{A}}_n(\rho_0(x)) \rightsquigarrow N_4(\lambda\sqrt{b(x)}\mathbb{D}, \mathbf{\Sigma}(\rho_0(x))),$$

where $\mathbb{D}$ is a (4×1) vector with elements

$$
\begin{aligned}
D_1 &:= 0, \\
D_2 &:= -\frac{\alpha\rho_0(x)(1+\gamma_0(x))}{\gamma_0(x)[1+\alpha(1+\gamma_0(x))][1-\rho_0(x)+\alpha(1+\gamma_0(x))]}, \\
D_3 &:= -\frac{\rho_0(x)[\alpha(1+\gamma_0(x))-\rho_0(x)]}{\gamma_0(x)[1-\rho_0(x)+\alpha(1+\gamma_0(x))][1-2\rho_0(x)+\alpha(1+\gamma_0(x))]}, \\
D_4 &:= \frac{\rho_0(x)(1-\rho_0(x))-\alpha^2\rho_0(x)(1+\gamma_0(x))^2}{[1+\alpha(1+\gamma_0(x))]^2[1-\rho_0(x)+\alpha(1+\gamma_0(x))]^2},
\end{aligned}
$$

and $\mathbf{\Sigma}(\rho_0(x))$ a symmetric (4×4) matrix with elements

$$
\begin{aligned}
\sigma_{11}(\rho_0(x)) &:= \|K\|_2^2, \\
\sigma_{21}(\rho_0(x)) &:= \frac{\|K\|_2^2}{1+\alpha(1+\gamma_0(x))}, \\
\sigma_{22}(\rho_0(x)) &:= \frac{\|K\|_2^2}{1+2\alpha(1+\gamma_0(x))}, \\
\sigma_{31}(\rho_0(x)) &:= \frac{\|K\|_2^2}{1-\rho_0(x)+\alpha(1+\gamma_0(x))}, \\
\sigma_{32}(\rho_0(x)) &:= \frac{\|K\|_2^2}{1-\rho_0(x)+2\alpha(1+\gamma_0(x))}, \\
\sigma_{33}(\rho_0(x)) &:= \frac{\|K\|_2^2}{1-2\rho_0(x)+2\alpha(1+\gamma_0(x))}, \\
\sigma_{41}(\rho_0(x)) &:= \frac{\gamma_0(x)\|K\|_2^2}{[1+\alpha(1+\gamma_0(x))]^2}, \\
\sigma_{42}(\rho_0(x)) &:= \frac{\gamma_0(x)\|K\|_2^2}{[1+2\alpha(1+\gamma_0(x))]^2}, \\
\sigma_{43}(\rho_0(x)) &:= \frac{\gamma_0(x)\|K\|_2^2}{[1-\rho_0(x)+2\alpha(1+\gamma_0(x))]^2}, \\
\sigma_{44}(\rho_0(x)) &:= \frac{2\gamma_0^2(x)\|K\|_2^2}{[1+2\alpha(1+\gamma_0(x))]^3}.
\end{aligned}
$$

Now, apply a Taylor series expansion of the estimating equations $f_1(\hat{\gamma}_n(x), \hat{\delta}_n(x); \rho_0(x)) = 0$ and

$f_2(\hat{\gamma}_n(x), \hat{\delta}_n(x); \rho_0(x)) = 0$ around $(\gamma_0(x), 0)$. This gives

$$\begin{aligned}
0 &= f_1(\gamma_0(x), 0; \rho_0(x)) + f_{11}(\gamma_0(x), 0; \rho_0(x))(\hat{\gamma}_n(x) - \gamma_0(x)) + f_{12}(\gamma_0(x), 0; \rho_0(x))\hat{\delta}_n(x) \\
&\quad + \frac{1}{2}\Big\{ f_{111}(\breve{\gamma}_n(x), \breve{\delta}_n(x); \rho_0(x))(\hat{\gamma}_n(x) - \gamma_0(x))^2 + f_{122}(\breve{\gamma}_n(x), \breve{\delta}_n(x); \rho_0(x))\hat{\delta}_n^2(x) \\
&\quad + 2 f_{112}(\breve{\gamma}_n(x), \breve{\delta}_n(x); \rho_0(x))(\hat{\gamma}_n(x) - \gamma_0(x))\hat{\delta}_n(x) \Big\}, \\
0 &= f_2(\gamma_0(x), 0; \rho_0(x)) + f_{21}(\gamma_0(x), 0; \rho_0(x))(\hat{\gamma}_n(x) - \gamma_0(x)) + f_{22}(\gamma_0(x), 0; \rho_0(x))\hat{\delta}_n(x) \\
&\quad + \frac{1}{2}\Big\{ f_{211}(\breve{\gamma}_n(x), \breve{\delta}_n(x); \rho_0(x))(\hat{\gamma}_n(x) - \gamma_0(x))^2 + f_{222}(\breve{\gamma}_n(x), \breve{\delta}_n(x); \rho_0(x))\hat{\delta}_n^2(x) \\
&\quad + 2 f_{122}(\breve{\gamma}_n(x), \breve{\delta}_n(x); \rho_0(x))(\hat{\gamma}_n(x) - \gamma_0(x))\hat{\delta}_n(x) \Big\},
\end{aligned}$$

where $(\breve{\gamma}_n(x), \breve{\delta}_n(x))$ is a point on the line segment connecting $(\hat{\gamma}_n(x), \hat{\delta}_n(x))$ and $(\gamma_0(x), 0)$. A straightforward rearrangement gives a set of random equations where interest is in $r_n(\hat{\gamma}_n(x) - \gamma_0(x))$ and $r_n\hat{\delta}_n(x)$:

$$-r_n \begin{bmatrix} f_1(\gamma_0(x), 0; \rho_0(x)) \\ f_2(\gamma_0(x), 0; \rho_0(x)) \end{bmatrix} = \begin{bmatrix} \tilde{f}_{11}(\gamma_0(x), 0; \rho_0(x)) & \tilde{f}_{12}(\gamma_0(x), 0; \rho_0(x)) \\ \tilde{f}_{12}(\gamma_0(x), 0; \rho_0(x)) & \tilde{f}_{22}(\gamma_0(x), 0; \rho_0(x)) \end{bmatrix} \begin{bmatrix} r_n(\hat{\gamma}_n(x) - \gamma_0(x)) \\ r_n\hat{\delta}_n(x) \end{bmatrix} \quad \text{(A-2)}$$

where

$$\begin{aligned}
\tilde{f}_{11}(\gamma_0(x), 0; \rho_0(x)) &:= f_{11}(\gamma_0(x), 0; \rho_0(x)) + \frac{1}{2}\Big[f_{111}(\breve{\gamma}_n(x), \breve{\delta}_n(x); \rho_0(x))(\hat{\gamma}_n(x) - \gamma_0(x)) \\
&\quad + f_{112}(\breve{\gamma}_n(x), \breve{\delta}_n(x); \rho_0(x))\hat{\delta}_n(x) \Big], \\
\tilde{f}_{12}(\gamma_0(x), 0; \rho_0(x)) &:= f_{12}(\gamma_0(x), 0; \rho_0(x)) + \frac{1}{2}\Big[f_{122}(\breve{\gamma}_n(x), \breve{\delta}_n(x); \rho_0(x))\hat{\delta}_n(x) \\
&\quad + f_{112}(\breve{\gamma}_n(x), \breve{\delta}_n(x); \rho_0(x))(\hat{\gamma}_n(x) - \gamma_0(x)) \Big], \\
\tilde{f}_{22}(\gamma_0(x), 0; \rho_0(x)) &:= f_{22}(\gamma_0(x), 0; \rho_0(x)) + \frac{1}{2}\Big[f_{222}(\breve{\gamma}_n(x), \breve{\delta}_n(x); \rho_0(x))\hat{\delta}_n(x) \\
&\quad + f_{122}(\breve{\gamma}_n(x), \breve{\delta}_n(x); \rho_0(x))(\hat{\gamma}_n(x) - \gamma_0(x)) \Big].
\end{aligned}$$

Now, introduce

$$\mathbb{B}(\rho_0(x)) := \gamma_0^{-\alpha-2}(x) \begin{bmatrix} b_{11}(\rho_0(x)) & \gamma_0(x) & 0 & -1 \\ b_{21}(\rho_0(x)) & \gamma_0(x) & -\gamma_0(x)(1 - \rho_0(x)) & 0 \end{bmatrix},$$

with

$$\begin{aligned}
b_{11}(\rho_0(x)) &:= -\frac{\alpha\gamma_0(x)(1 + \gamma_0(x))}{[1 + \alpha(1 + \gamma_0(x))]^2}, \\
b_{21}(\rho_0(x)) &:= -\frac{\alpha\gamma_0(x)\rho_0(x)(1 + \gamma_0(x))}{[1 + \alpha(1 + \gamma_0(x))][1 - \rho_0(x) + \alpha(1 + \gamma_0(x))]},
\end{aligned}$$

so that

$$r_n \begin{bmatrix} f_1(\gamma_0(x), 0; \rho_0(x)) \\ f_2(\gamma_0(x), 0; \rho_0(x)) \end{bmatrix} = \mathbb{B}(\rho_0(x))\overline{\mathbb{A}}_n(\rho_0(x)),$$

is leading to the weak convergence

$$r_n \begin{bmatrix} f_1(\gamma_0(x), 0; \rho_0(x)) \\ f_2(\gamma_0(x), 0; \rho_0(x)) \end{bmatrix} \rightsquigarrow N_2(\lambda\sqrt{b(x)}\mathbb{B}(\rho_0(x))\mathbb{D}, \mathbb{B}(\rho_0(x))\boldsymbol{\Sigma}(\rho_0(x))\mathbb{B}'(\rho_0(x))).$$

Concerning the terms $\tilde{f}_{st}(\gamma_0(x), 0; \rho_0(x))$, $s, t = 1, 2$, we have by Lemmas 1 and 2, Theorem 1, the consistency of $(\hat{\gamma}_n(x), \hat{\delta}_n(x))$ and because $|f_{stu}(\gamma, \delta; \rho_0(x))| \leq M_{stu}(\boldsymbol{V})$, in some open neighborhood of $(\gamma_0(x), 0)$, with $M_{stu}(\boldsymbol{V}) = O_{\mathbb{P}}(1)$, $s, t, u = 1, 2$, that $\tilde{f}_{st}(\gamma_0(x), 0; \rho_0(x)) \xrightarrow{\mathbb{P}} f^*_{st}(\gamma_0(x), 0; \rho_0(x))$, $s, t = 1, 2$. Let

$$\mathbb{C}(\rho_0(x)) := \begin{bmatrix} f^*_{11}(\gamma_0(x), 0; \rho_0(x)) & f^*_{12}(\gamma_0(x), 0; \rho_0(x)) \\ f^*_{12}(\gamma_0(x), 0; \rho_0(x)) & f^*_{22}(\gamma_0(x), 0; \rho_0(x)) \end{bmatrix}.$$

From the proof of the consistency, we know that $\mathbb{C}(\rho_0(x))$ is a positive definite matrix, and thus invertible. Then, according to Lemma 5.2 in Chapter 6 of Lehmann and Casella (1998), for the solution of the system of equations (A-2), we have the following convergence

$$\begin{aligned} r_n \begin{bmatrix} \hat{\gamma}_n(x) - \gamma_0(x) \\ \hat{\delta}_n(x) \end{bmatrix} \rightsquigarrow \; & N_2(-\lambda\sqrt{b(x)}\mathbb{C}^{-1}(\rho_0(x))\mathbb{B}(\rho_0(x))\mathbb{D}, \\ & \mathbb{C}^{-1}(\rho_0(x))\mathbb{B}(\rho_0(x))\boldsymbol{\Sigma}(\rho_0(x))\mathbb{B}'(\rho_0(x))\mathbb{C}^{-1}(\rho_0(x))). \end{aligned}$$

After tedious calculations one can show that $-\mathbb{C}^{-1}(\rho_0(x))\mathbb{B}(\rho_0(x))\mathbb{D} = [0, 1]'$. Taking into account that $r_n\delta(u_n; x) \to \lambda\sqrt{b(x)}$, the theorem follows.

Proof of Proposition 1

The arguments needed to prove the consistency and asymptotic normality are the same as those used in the proofs of Theorem 2 and 3, and therefore we limit ourselves to giving some comments to the main ideas. Concerning the consistency one works with $\widetilde{\Delta}_\alpha(\gamma, \delta; \tilde{\rho}(x))$ and its derivatives. Again by Lemmas 1, 2 and Theorem 1 we have that $f_s(\gamma_0(x), 0; \tilde{\rho}(x)) \xrightarrow{\mathbb{P}} 0$, $s = 1, 2$, and that

$f_{st}(\gamma_0(x), 0; \tilde{\rho}(x)) \xrightarrow{\mathbb{P}} f^*_{st}(\gamma_0(x), 0; \tilde{\rho}(x))$, $s, t = 1, 2$, leading to the results for S_1 and S_2. Also for the third order derivatives we can use the same arguments. This establishes the existence and the consistency. To prove the asymptotic normality one uses the same line of argumentation as in Theorem 3, with $\rho_0(x)$ replaced by $\tilde{\rho}(x)$ in $\overline{\mathbb{A}}_{k,n}(\rho_0(x))$, $\boldsymbol{\Sigma}(\rho_0(x))$, $\mathbb{B}(\rho_0(x))$ and $\mathbb{C}(\rho_0(x))$, and replacing the vector $\mathbb{D}$ by $\widetilde{\mathbb{D}}$, having as elements $\widetilde{D}_1 := D_1$, $\widetilde{D}_2 := D_2$, $\widetilde{D}_4 := D_4$ and

$$\widetilde{D}_3 := -\frac{[\alpha(1+\gamma_0(x)) - \tilde{\rho}(x)]\rho_0(x)}{\gamma_0(x)[1-\tilde{\rho}(x)+\alpha(1+\gamma_0(x))][1-\rho_0(x)-\tilde{\rho}(x)+\alpha(1+\gamma_0(x))]}.$$

Proof of Theorem 4

The proof of Theorem 4 is similar to that of Theorems 2 and 3, and therefore we only give the big lines of argument.

Concerning the existence and consistency of $(\hat{\gamma}(x), \hat{\delta}_n(x))$ as estimators for $(\gamma_0(x), 0)$, we have that by the consistency of $\hat{\rho}_n(x)$ and conditioning on the event $\hat{\rho}_n(x) \in (\rho_0(x)-\varepsilon, \rho_0(x)+\varepsilon)$ for some $\varepsilon > 0$, it is sufficient to show that

$$\mathbb{P}_{(\gamma_0(x),0)}(\widetilde{\Delta}_\alpha(\gamma_0(x), 0; \hat{\rho}_n(x)) < \widetilde{\Delta}_\alpha(\gamma, \delta; \hat{\rho}_n(x)) \\ \text{for all } (\gamma, \delta) \text{ on the surface of } Q_r \mid \hat{\rho}_n(x) \in (\rho_0(x)-\varepsilon, \rho_0(x)+\varepsilon)) \to 1.$$

First make a Taylor series expansion as in (A-1), though now with $\rho_0(x)$ replaced by $\hat{\rho}_n(x)$.

Assume that $(-(\alpha(1+\gamma_0(x)) - (\rho_0(x)-\varepsilon))/\gamma_0(x), -(\alpha(1+\gamma_0(x)) - (\rho_0(x)+\varepsilon))/\gamma_0(x)) \in [S, 0]$. Concerning S_1, we have that $f_1(\gamma_0(x), 0; \hat{\rho}_n(x))$ does not depend on $\hat{\rho}_n(x)$ and therefore obviously

$f_1(\gamma_0, 0; \hat{\rho}_n(x)) \xrightarrow{\mathbb{P}} 0$, whereas for $f_2(\gamma_0, 0; \hat{\rho}_n(x))$ we write

$$
\begin{aligned}
& f_2(\gamma_0(x), 0; \hat{\rho}_n(x)) \\
= \quad & \gamma_0^{-\alpha-1}(x) \Big[- \frac{\alpha \hat{\rho}_n(x)(1+\gamma_0(x))}{[1+\alpha(1+\gamma_0(x))][1-\hat{\rho}_n(x)+\alpha(1+\gamma_0(x))]} \frac{T_n(K,0,0;x)}{\bar{F}(u_n;x)b(x)} \\
& + \frac{T_n(K, -\alpha(1+\gamma_0(x))/\gamma_0(x), 0; x)}{\bar{F}(u_n;x)b(x)} \\
& -(1-\hat{\rho}_n(x)) \left(\frac{T_n(K, -(\alpha(1+\gamma_0(x)) - \rho_0(x))/\gamma_0(x), 0; x)}{\bar{F}(u_n;x)b(x)} - \frac{1}{1-\rho_0(x)+\alpha(1+\gamma_0(x))} \right) \\
& -(1-\hat{\rho}_n(x)) \left(\frac{T_n(K, -(\alpha(1+\gamma_0(x)) - \hat{\rho}_n(x))/\gamma_0(x), 0; x)}{\bar{F}(u_n;x)b(x)} - \frac{T_n(K, -(\alpha(1+\gamma_0(x)) - \rho_0(x))/\gamma_0(x), 0; x)}{\bar{F}(u_n;x)b(x)} \right) \\
& - \frac{1-\hat{\rho}_n(x)}{1-\rho_0(x)+\alpha(1+\gamma_0(x))} \Big] \\
=: \quad & \gamma_0^{\alpha-1}(x) \left[T_1 + T_2 + T_3 + T_4 + T_5 \right].
\end{aligned}
$$

Now use Lemmas 1, 2 and Theorem 1 to obtain

$$
\begin{aligned}
T_1 & \xrightarrow{\mathbb{P}} -\frac{\alpha \rho_0(x)(1+\gamma_0(x))}{[1+\alpha(1+\gamma_0(x))][1-\rho_0(x)+\alpha(1+\gamma_0(x))]}, \\
T_2 & \xrightarrow{\mathbb{P}} \frac{1}{1+\alpha(1+\gamma_0(x))}, \\
T_3 & \xrightarrow{\mathbb{P}} 0, \\
|T_4| & \leq \frac{1-\hat{\rho}_n(x)}{\gamma_0(x)} \, |\hat{\rho}_n(x) - \rho_0(x)| \, \frac{T_n(K,0,1;x)}{\bar{F}(u_n;x)b(x)} = o_{\mathbb{P}}(1), \\
T_5 & \xrightarrow{\mathbb{P}} -\frac{1-\rho_0(x)}{1-\rho_0(x)+\alpha(1+\gamma_0(x))}.
\end{aligned}
$$

Combining these results gives that $f_2(\gamma_0, 0; \hat{\rho}_n(x)) \xrightarrow{\mathbb{P}} 0$.

For S_2, write

$$
\begin{aligned}
2S_2 = \; & f_{11}^*(\gamma_0(x), 0; \rho_0(x))(\gamma - \gamma_0(x))^2 + f_{22}^*(\gamma_0(x), 0; \rho_0(x))\delta^2 \\
& + 2 f_{12}^*(\gamma_0(x), 0; \rho_0(x))(\gamma - \gamma_0(x))\delta \\
& + [f_{11}(\gamma_0(x), 0; \hat{\rho}_n(x)) - f_{11}^*(\gamma_0(x), 0; \rho_0(x))](\gamma - \gamma_0(x))^2 \\
& + [f_{22}(\gamma_0(x), 0; \hat{\rho}_n(x)) - f_{22}^*(\gamma_0(x), 0; \rho_0(x))]\delta^2 \\
& + 2[f_{12}(\gamma_0(x), 0; \hat{\rho}_n(x)) - f_{12}^*(\gamma_0(x), 0; \rho_0(x))](\gamma - \gamma_0(x))\delta.
\end{aligned}
$$

By arguments similar to those used above when treating $f_2(\gamma_0(x), 0; \hat{\rho}_n(x))$, we have that $f_{st}(\gamma_0(x), 0; \hat{\rho}_n(x)) \xrightarrow{\mathbb{P}} f_{st}^*(\gamma_0(x), 0; \rho_0(x))$, $s, t = 1, 2$, and hence we can proceed as in the proof of Theorem 2. Finally, conditionally on $\hat{\rho}_n(x) \in (\rho_0(x) - \varepsilon, \rho_0(x) + \varepsilon)$, also the argument for the third order derivatives holds and the proof for the existence and consistency can be completed in the same way as in the proof of Theorem 2.

The proof of asymptotic normality proceeds along the lines of the proof of Theorem 3. To start we make a Taylor series expansion of the estimating equations, leading to (A-2), though with $\rho_0(x)$ replaced by $\hat{\rho}_n(x)$. Since $\mathbb{P}(\hat{\rho}_n(x) \in (\rho_0(x) - \varepsilon, \rho_0(x) + \varepsilon)) \to 1$, we have that (by an appropriate choice of S in Corollary 1)

$$\overline{\mathbb{A}}_{k,n}(\hat{\rho}_n(x)) \rightsquigarrow N_4(\lambda\sqrt{b(x)}\mathbb{D}, \mathbf{\Sigma}(\rho_0(x))),$$

and hence

$$\begin{aligned} r_n \begin{bmatrix} f_1(\gamma_0(x), 0; \hat{\rho}_n(x)) \\ f_2(\gamma_0(x), 0; \hat{\rho}_n(x)) \end{bmatrix} &= \mathbb{B}(\hat{\rho}_n(x))\overline{\mathbb{A}}_{k,n}(\hat{\rho}_n(x)) \\ &\rightsquigarrow N_2(\lambda\sqrt{b(x)}\mathbb{B}(\rho_0(x))\mathbb{D}, \mathbb{B}(\rho_0(x))\mathbf{\Sigma}(\rho_0(x))\mathbb{B}'(\rho_0(x))). \end{aligned}$$

The rest of the proof is identical to that of Theorem 3.

References

[1] 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.

[2] Lehmann, E.L., Casella, G., 1998. *Theory of Point Estimation.* Springer.

[3] van der Vaart, A.W., 2007. *Asymptotic Statistics.* Cambridge Series in Statistical and Probabilistic Mathematics.