Supplementary Material to “Extreme value statistics in semi-supervised models”

Hanan Ahmed* John H.J. Einmahl†

Tilburg University Tilburg University

Chen Zhou‡

Erasmus University Rotterdam

March 15, 2024

Abstract

Appendix A provides details in the proof of Proposition 6.1. Appendix B provides covariance calculation needed in the proof of Proposition 2.1. Appendix C provides proofs of all results in results in Section 2.2. Appendix D provides additional simulation results.

A Details in the proof of Proposition 6.1

Based on the first order conditions derived from the log-likelihood function, the MLE $(\hat{\gamma}, \hat{\sigma})$ can be obtained as follows. Firstly, solve the following equation

$$h_n(t) := f_n(t)g_n(t) - 1 = 0,$$

*Email: h.e.g.ahmed@tilburguniversity.edu

†John Einmahl holds the Arie Kapteyn Chair 2019–2022 and gratefully acknowledges the corresponding research support. Email: j.h.j.einmahl@uvt.nl

‡Postal address: Econometric Institute, Erasmus University Rotterdam, 3000DR Rotterdam, The Netherlands. Phone: +31(10)4082269. Email: zhou@ese.eur.nl

where

$$f_n(t) := 1 + \frac{1}{k}\sum_{i=1}^{k}\log(1 + t(X_{n-i+1:n} - X_{n-k:n})),$$

$$g_n(t) := \frac{1}{k}\sum_{i=1}^{k}\frac{1}{1 + t(X_{n-i+1:n} - X_{n-k:n})},$$

to obtain a solution $t^* \neq 0$. Then $\hat{\gamma} = f_n(t^*) - 1$ and $\hat{\sigma} = \hat{\gamma}/t^*$; see Zhou (2009) and Grimshaw (1993). Note that $t^* = \hat{\gamma}/\hat{\sigma}$.

In the proof of Theorem 4.1 in Zhou (2009) provided two bounds for t^*, denoted as $\underline{t}$ and $\bar{t}$.[1] Using the tail quantile process result in Lemma 6.3, following exactly the same steps in the proof of Theorem 4.1 in Zhou (2009), we can show that as $n \to \infty$, with probability tending to 1, $h_n(\underline{t}) < 0 < h_n(\bar{t})$. Together with the continuity of the function h_n, we obtain the existence of a solution t^* in between $\underline{t}$ and $\bar{t}$.

In addition, based on the specific construction of the two bounds, using the tail quantile process result in Lemma 6.3, we can verify that both $|\underline{t}a(n/k) - \gamma| = O_p(1/\sqrt{k})$ and $\left|\bar{t}a(n/k) - \gamma\right| = O_p(1/\sqrt{k})$ as $n \to \infty$. This again follows exactly the same steps in the proof of Theorem 4.1 in Zhou (2009). As a consequence, since $t^* = \hat{\gamma}/\hat{\sigma}$, the solution $(\hat{\gamma}, \hat{\sigma})$ satisfies that $\left|\frac{\hat{\gamma}}{\hat{\sigma}/a(n/k)} - \gamma\right| = O_p(1/\sqrt{k})$ and $\log\frac{\hat{\sigma}}{a(n/k)} = O_P(1)$. Notice that these two conditions are required in the proof in Drees et al. (2004).

Next, one can linearize $f_n(\cdot)$ and $g_n(\cdot)$ in the neighborhood of $\gamma/a(n/k)$; see Proposition

[1] The two bounds differ for the three cases $\gamma < 0$, $\gamma = 0$ and $\gamma > 0$.

3.1 in Drees et al. (2004). For $\gamma \neq 0$, the solution $(\hat{\gamma}, \hat{\sigma})$ satisfies that

$$
\begin{aligned}
\hat{\gamma} &= f_n(\hat{\gamma}/\hat{\sigma}) - 1 \\
&= \int_0^1 \log\left(1 + \frac{\hat{\gamma}}{\hat{\sigma}/a(n/k)} \frac{X_{n-[ks]:n} - X_{n-k:n}}{a(n/k)}\right) ds \\
&= \int_0^{1/(2k)} \log\left(1 + \frac{\hat{\gamma}}{\hat{\sigma}/a(n/k)} \frac{X_{n-[ks]:n} - X_{n-k:n}}{a(n/k)}\right) ds + \int_{1/(2k)}^1 \log(s^{-\gamma}) ds \\
&\quad + \int_{1/(2k)}^1 \log\left(s^{\gamma}\left(1 + \frac{\hat{\gamma}}{\hat{\sigma}/a(n/k)} \frac{X_{n-[ks]:n} - X_{n-k:n}}{a(n/k)}\right)\right) ds \\
&= I_1 + \gamma(1 - O(k^{-1}\log k)) + I_2.
\end{aligned}
$$

The first term I_1 is dealt with by an inequality in Lemma 3.2 in Drees et al. (2004). We get that $I_1 = o_p(1/\sqrt{k})$ as $n \to \infty$. In addition, I_2 is handled by combining the fact that $\left|\frac{\gamma}{\hat{\sigma}/a(n/k)} - \gamma\right| = O_p(1/\sqrt{k})$ and

$$
\begin{aligned}
&\log\left(s^{\gamma}\left(1 + \frac{\hat{\gamma}}{\hat{\sigma}/a(n/k)} \frac{X_{n-[ks]:n} - X_{n-k:n}}{a(n/k)}\right)\right) \\
=&\log\left(s^{\gamma}\left(1 + \gamma\frac{X_{n-[ks]:n} - X_{n-k:n}}{a(n/k)}\right)\right) + \left(\frac{\hat{\gamma}}{\hat{\sigma}/a(n/k)} - \gamma\right) \frac{\frac{X_{n-[ks]:n} - X_{n-k:n}}{a(n/k)}}{1 + \gamma\frac{X_{n-[ks]:n} - X_{n-k:n}}{a(n/k)}} + o_p(1/\sqrt{k}) \\
=&\log\left(s^{\gamma}\left(s^{-\gamma} + \frac{Z_n(s)}{\sqrt{k}}\right)\right) + \left(\frac{\hat{\gamma}}{\hat{\sigma}/a(n/k)} - \gamma\right) \frac{\frac{s^{-\gamma}-1}{\gamma} + \frac{Z_n(s)}{\sqrt{k}}}{s^{-\gamma} + \frac{Z_n(s)}{\sqrt{k}}} + o_p(1/\sqrt{k}) \\
=&s^{\gamma}\frac{Z_n(s)}{\sqrt{k}} + \left(\frac{\hat{\gamma}}{\hat{\sigma}/a(n/k)} - \gamma\right) \frac{1 - s^{\gamma}}{\gamma} + o_p(1/\sqrt{k}),
\end{aligned}
$$

uniformly for all $s \in [1/(2k), 1]$. Eventually, we obtain that as $n \to \infty$,

$$
\hat{\gamma} - \gamma = \frac{1}{\sqrt{k}} \int_0^1 s^{\gamma} Z_n(s) ds + \frac{1}{\gamma+1}\left(\frac{\hat{\gamma}}{\hat{\sigma}/a(n/k)} - \gamma\right) + o_p(1/\sqrt{k}). \tag{S.1}
$$

A similar linearization for the function g_n leads to that as $n \to \infty$,

$$
\frac{1}{\hat{\gamma}+1} - \frac{1}{\gamma+1} = -\frac{1}{(\gamma+1)(2\gamma+1)}\left(\frac{\hat{\gamma}}{\hat{\sigma}/a(n/k)} - \gamma\right) - \frac{1}{\sqrt{k}} \int_0^1 s^{2\gamma} Z_n(s) ds + o_p(1/\sqrt{k}). \tag{S.2}
$$

Multiplying both sides of the equation (S.2) by $(2\gamma+1)$ and adding to the equation (S.1), we get that, as $n \to \infty$,

$$\hat{\gamma} + \frac{2\gamma+1}{\hat{\gamma}+1} - \gamma - \frac{2\gamma+1}{\gamma+1} = \frac{1}{\sqrt{k}} \int_0^1 (s^\gamma - (2\gamma+1)s^{2\gamma}) Z_n(s) ds + o_p(1/\sqrt{k}).$$

The proposition is then proved by applying the Delta method. The case $\gamma = 0$ is handled in a similar way.

B Calculation of covariance in Proposition 2.1

We calculate the covariance of Ω and $\tilde{\Omega}$ as follows. Note that $Cov(W_1(s), -\tilde{W}(t)) = (1-\nu^2)R(s,t)$. Then, we have that

$$\begin{aligned}
&Cov(\Omega, \tilde{\Omega}) = (1-\nu^2)\frac{(\gamma+1)^2(g+1)^2}{\gamma g} \\
&\cdot \int_0^1 \int_0^1 (s^\gamma - (2\gamma+1)s^{2\gamma})(t^g - (2g+1)t^{2g})\Big(\frac{R(s,t)}{s^{\gamma+1}t^{g+1}} - \frac{R(s,1)}{s^{\gamma+1}} - \frac{R(1,t)}{t^{g+1}} + R(1,1)\Big) dsdt \\
&= (1-\nu^2)\frac{(\gamma+1)^2(g+1)^2}{\gamma g} \int_0^1 \int_0^1 \left(\frac{1}{st} - \frac{(2g+1)}{st^{1-g}} - \frac{(2\gamma+1)}{s^{1-\gamma}t} + \frac{(2\gamma+1)(2g+1)}{s^{1-\gamma}t^{1-g}} \right) R(s,t) \\
&\qquad - \left(\frac{t^g}{s} - \frac{(2g+1)t^{2g}}{s} - \frac{(2\gamma+1)t^g}{s^{1-\gamma}} + \frac{(2\gamma+1)(2g+1)t^{2g}}{s^{1-\gamma}} \right) R(s,1) \\
&\qquad - \left(\frac{s^\gamma}{t} - \frac{(2g+1)s^\gamma}{t^{1-g}} - \frac{(2\gamma+1)s^{2\gamma}}{t} + \frac{(2\gamma+1)(2g+1)s^{2\gamma}}{t^{1-g}} \right) R(1,t) \\
&\qquad + \left(s^\gamma t^g - (2g+1)s^\gamma t^{2g} - (2\gamma+1)s^{2\gamma}t^g + (2\gamma+1)(2g+1)s^{2\gamma}t^{2g} \right) R(1,1) dsdt.
\end{aligned}$$

Using a change of variables and the first order homogeneity of R, we obtain:

$$\begin{aligned}
&\int_0^1 \int_0^1 \frac{R(s,t)}{st} dsdt = \int_0^1 \int_0^t \frac{R(s,t)}{st} dsdt + \int_0^1 \int_0^s \frac{R(s,t)}{st} dtds \\
&\quad = \int_0^1 \int_0^1 \frac{R(ts,t)}{st} dsdt + \int_0^1 \int_0^1 \frac{R(s,ts)}{st} dtds = \int_0^1 \frac{R(s,1)}{s} ds + \int_0^1 \frac{R(1,t)}{t} dt,
\end{aligned}$$

$$
\begin{aligned}
&\int_0^1\int_0^1 \frac{R(s,t)}{s^{1-\gamma}t} dsdt = \int_0^1\int_0^t \frac{R(s,t)}{s^{1-\gamma}t} dsdt + \int_0^1\int_0^s \frac{R(s,t)}{s^{1-\gamma}t} dtds \\
&\quad = \int_0^1\int_0^1 \frac{R(ts,t)}{(st)^{1-\gamma}} dsdt + \int_0^1\int_0^1 \frac{R(s,ts)}{s^{1-\gamma}t} dtds = \frac{1}{1+\gamma}\left[\int_0^1 \frac{R(s,1)}{s^{1-\gamma}} ds + \int_0^1 \frac{R(1,t)}{t} dt\right],
\end{aligned}
$$

and similarly

$$
\int_0^1\int_0^1 \frac{R(s,t)}{st^{1-g}} dsdt = \frac{1}{1+g}\left[\int_0^1 \frac{R(s,1)}{s} ds + \int_0^1 \frac{R(1,t)}{t^{1-g}} dt\right].
$$

Also,

$$
\begin{aligned}
&\int_0^1\int_0^1 \frac{R(s,t)}{s^{1-\gamma}t^{1-g}} dsdt = \int_0^1\int_0^t \frac{R(s,t)}{s^{1-\gamma}t^{1-g}} dsdt + \int_0^1\int_0^s \frac{R(s,t)}{s^{1-\gamma}t^{1-g}} dtds \\
&= \int_0^1\int_0^1 \frac{R(ts,t)}{s^{1-\gamma}t^{1-\gamma-g}} dsdt + \int_0^1\int_0^1 \frac{R(s,ts)}{s^{1-\gamma-g}t^{1-g}} dtds \\
&= \frac{1}{\gamma+g+1}\left[\int_0^1 \frac{R(s,1)}{s^{1-\gamma}} ds + \int_0^1 \frac{R(1,t)}{t^{1-g}} dt\right].
\end{aligned}
$$

Substituting the expressions for these four integrals involving $R(s,t)$ in the formula for $Cov(\Omega,\tilde{\Omega})$ above, we obtain that this covariance is equal to $(1-\nu^2)(\gamma+1)(g+1)R_g$.

C Proofs of results in Subsection 2.2

For the proof of Theorem 2.2 we first need an approximations of $\hat{\sigma}$ and $\tilde{\sigma}_g$, similar to those obtained for $\hat{\gamma}$ and $\hat{g}$.

Proposition C.1. *Assume that (3) and (4) hold and $\sqrt{k}A\left(\frac{n}{k}\right) = O(1)$, as $n \to \infty$. Then for $\gamma > -1/2$, and $\gamma \neq 0$, with probability tending to 1, there exist a unique maximizer of the likelihood functions based on $\{X_i\}_{i=1}^n$ denoted as $\hat{\sigma}$, such that as $n \to \infty$,*

$$
\sqrt{k}\left(\frac{\hat{\sigma}}{a\left(\frac{n}{k}\right)} - 1\right) - \frac{\gamma+1}{\gamma}\int_0^1 ((\gamma+1)(2\gamma+1)s^{2\gamma} - s^{\gamma})Z_n(s)ds = o_{\mathbb{P}}(1),
$$

and, for $\gamma = 0$,

$$\sqrt{k}\left(\frac{\hat{\sigma}}{a\left(\frac{n}{k}\right)} - 1\right) - \int_0^1 (3 + \log s) Z_n(s) ds = o_{\mathbb{P}}(1).$$

Proof. The existence of $\hat{\sigma}$ follows from Theorem 4.1 in Zhou (2009); see the explanation in the proof of Proposition 6.1 above. Next, from the expansion of $\hat{\gamma} - \gamma$ and the relation (S.1), we can derive the expansion of $\frac{\hat{\gamma}}{\hat{\sigma}/a(n/k)} - \gamma$. The proposition is proved by further calculation using the Delta method. □

Proposition C.2. *Assume that* F_2 *is continuous and* k *satisfies (3). For* $g > -1/2$ *and* $g \neq 0$, *with probability tending to 1, there exists a unique maximizer of the likelihood function based on* $\{\tilde{Y}_i\}_{i=1}^n$, *denoted as* $\tilde{\sigma}_g$, *such that, as* $n \to \infty$,

$$\sqrt{k}\left(\frac{\tilde{\sigma}_g}{\left(\frac{n}{k}\right)^g} - 1\right) - \frac{g+1}{g}\int_0^1 ((g+1)(2g+1)s^{2g} - s^g) H_n(s) ds = o_{\mathbb{P}}(1),$$

and, for $g = 0$,

$$\sqrt{k}\left(\tilde{\sigma}_g - 1\right) - \int_0^1 (3 + \log s) H_n(s) ds = o_{\mathbb{P}}(1).$$

Proof. The proof follows exactly the same step as that of Proposition C.1. The main difference is again that $\{\tilde{Y}_i\}_{i=1}^n$ are not i.i.d. observations. Nevertheless, all asymptotic expansions such as that of $\hat{g} - g$ are still valid with the only difference that the random limit is driven by a proper functional of $\tilde{W}$ instead. Such asymptotic expansions are sufficient to ensure that the proof can still be realized. □

Proof of Theorem 2.2. From Proposition C.1, it follows that for $\gamma > -\frac{1}{2}$ and $\gamma \neq 0$,

$$\sqrt{k}\left(\frac{\hat{\sigma}}{a\left(\frac{n}{k}\right)} - 1\right) - \frac{\gamma+1}{\gamma}\sqrt{k}A\left(\frac{n}{k}\right)\int_0^1 ((\gamma+1)(2\gamma+1)s^{2\gamma} - s^{\gamma})\Psi(s^{-1}) ds$$
$$\xrightarrow{\mathbb{P}} \frac{\gamma+1}{\gamma}\int_0^1 ((\gamma+1)(2\gamma+1)s^{2\gamma} - s^{\gamma})(s^{-\gamma-1}W_1(s) - W_1(1)) ds =: \Sigma_\gamma\,,$$

as $n \to \infty$. Similarly it follows from Proposition C.2 that for $g > -\frac{1}{2}$ and $g \neq 0$,

$$\sqrt{k}\left(\frac{\tilde{\sigma}_g}{\left(\frac{n}{k}\right)^g} - 1\right) \xrightarrow{\mathbb{P}} \frac{g+1}{g}\int_0^1 ((g+1)(2g+1)s^{2g} - s^g)\left(\tilde{W}(1) - s^{-g-1}\tilde{W}(s)\right) ds =: \tilde{\Sigma}_g\,.$$

Hence, as $n \to \infty$,

$$\left(\sqrt{k}\left(\frac{\hat{\sigma}}{a\left(\frac{n}{k}\right)} - 1\right), \sqrt{k}\left(\frac{\tilde{\sigma}_g}{\left(\frac{n}{k}\right)^g} - 1\right)\right) \xrightarrow{\mathbb{P}} \left(\Sigma_\gamma - \frac{\lambda\rho}{(1-\rho)(1+\gamma-\rho)}, \tilde{\Sigma}_g\right). \tag{S.3}$$

It follows from the uniform consistency of $\hat{R}$ on $[0,1]^2$, that $\hat{S}_g \xrightarrow{\mathbb{P}} S_g$. Using this, $\hat{\sigma}/a\left(\frac{n}{k}\right) \xrightarrow{\mathbb{P}} 1$, and (S.3), we obtain, as $n \to \infty$,

$$\begin{aligned}\sqrt{k}\left(\frac{\hat{\sigma}_g}{a\left(\frac{n}{k}\right)} - 1\right) &= \sqrt{k}\left(\frac{\hat{\sigma}}{a\left(\frac{n}{k}\right)} - 1\right) - \frac{S_g}{1+(1+g)^2}\sqrt{k}\left(\frac{\tilde{\sigma}_g}{(\frac{n}{k})^g} - 1\right)(1 + o_{\mathbb{P}}(1)) \\ &\xrightarrow{\mathbb{P}} \Sigma_\gamma - \frac{\lambda\rho}{(1-\rho)(1+\gamma-\rho)} - \frac{S_g}{1+(1+g)^2}\tilde{\Sigma}_g\,. \end{aligned} \tag{S.4}$$

From (36), as $n \to \infty$,

$$\sqrt{k}\left(\hat{\gamma}_g - \gamma\right) \xrightarrow{\mathbb{P}} \Omega - \frac{1+\gamma}{1+g}R_g\tilde{\Omega}\,; \tag{S.5}$$

see the proof of Proposition 2.1 for the definitions of Ω and $\tilde{\Omega}$, R_g is defined as in (5). By Lemma 6.3 with $s = 1$, as $n \to \infty$,

$$\sqrt{k}\left(\frac{X_{n-k:n} - U_1\left(\frac{n}{k}\right)}{a(\frac{n}{k})}\right) \xrightarrow{\mathbb{P}} W_1(1). \tag{S.6}$$

Combining (S.4)–(S.6), we have that, as $n \to \infty$

$$\begin{aligned}&\sqrt{k}\left(\hat{\gamma}_g - \gamma, \frac{\hat{\sigma}_g}{a\left(\frac{n}{k}\right)} - 1, \frac{X_{n-k:n} - U_1\left(\frac{n}{k}\right)}{a\left(\frac{n}{k}\right)}\right) \\ &\quad\xrightarrow{d} \left(\Omega - \frac{1+\gamma}{1+g}R_g\tilde{\Omega}, \Sigma_\gamma - \frac{\lambda\rho}{(1-\rho)(1+\gamma-\rho)} - \frac{S_g}{1+(1+g)^2}\tilde{\Sigma}_g, W_1(1)\right). \end{aligned} \tag{S.7}$$

It remains to be shown that the random vector on the right in (S.7) has the trivariate normal distribution specified in (11). The trivariate normality and the mean vector are straightforward from the definitions. The variance of the first component is given in (8). The variance of the third component is obviously equal to 1. For the variance of the second component we readily obtain (and it is well-known) that $Var(\Sigma_\gamma) = 1 + (1+\gamma)^2$ and $Var(\tilde{\Sigma}_g) = (1-\nu^2)(1+(1+g)^2)$. Moreover,

$$
\begin{aligned}
&Cov(\Sigma_\gamma, \tilde{\Sigma}_g) = (1-\nu^2)\frac{(\gamma+1)(g+1)}{\gamma g}\int_0^1\int_0^1 ((\gamma+1)(2\gamma+1)s^{2\gamma} - s^\gamma)((g+1)(2g+1)t^{2g} - t^g) \\
&\left(\frac{R(s,t)}{s^{\gamma+1}t^{g+1}} - \frac{R(s,1)}{s^{\gamma+1}} - \frac{R(1,t)}{t^{g+1}} + R(1,1)\right) dsdt \\
&= (1-\nu^2)\frac{(\gamma+1)(g+1)}{\gamma g} \\
&\cdot \int_0^1\int_0^1 \Big(\frac{(\gamma+1)(2\gamma+1)(g+1)(2g+1)}{s^{1-\gamma}t^{1-g}} - \frac{(\gamma+1)(2\gamma+1)}{s^{1-\gamma}t} - \frac{(g+1)(2g+1)}{t^{1-g}s} + \frac{1}{st}\Big) R(s,t) \\
&\quad -\Big(\frac{(\gamma+1)(g+1)(2g+1)}{t^{1-g}} - \frac{(\gamma+1)}{t} - \frac{(g+1)(2g+1)}{(\gamma+1)t^{1-g}} + \frac{1}{(\gamma+1)t}\Big) R(1,t) \\
&\quad -\Big(\frac{(g+1)(\gamma+1)(2\gamma+1)}{s^{1-\gamma}} - \frac{(g+1)}{s} - \frac{(\gamma+1)(2\gamma+1)}{(g+1)s^{1-\gamma}} + \frac{1}{(g+1)s}\Big) R(s,1) dsdt \\
&\quad +(1-\nu^2)\frac{(\gamma+1)(g+1)}{\gamma g}\left((\gamma+1)(g+1) - \frac{(\gamma+1)}{(g+1)} - \frac{(g+1)}{(\gamma+1)} + \frac{1}{(\gamma+1)(g+1)}\right) R(1,1) \\
&= (1-\nu^2)S_g.
\end{aligned}
$$

This yields the variance of the second component.

We finally consider the three covariance terms in the matrix K in (11). We begin with the covariance of the first and second component on the right in (S.7). It easily follows (and is again well-known) that $Cov(\Omega, \Sigma_\gamma) = -(1+\gamma)$ and $Cov(\tilde{\Omega}, \tilde{\Sigma}_g) = -(1-\nu^2)(1+g)$.

Further,

$$Cov(\Omega, \tilde{\Sigma}_g) = (1-\nu^2)\frac{(1+\gamma)^2(1+g)}{\gamma g}\Big[\Big(\frac{(g+1)^2+\gamma-(g+1)^2(\gamma+1)}{(g+1)(\gamma+1)}\Big)R(1,1)$$
$$+\frac{g^2}{g+1}\int_0^1 \frac{R(s,1)}{s}ds + \Big(\frac{2\gamma+1}{\gamma+1} - \frac{(2\gamma+1)(g+1)(2g+1)}{\gamma+g+1} + (2\gamma+1)(g+1) - \frac{2\gamma+1}{g+1}\Big)$$
$$\int_0^1 \frac{R(s,1)}{s^{1-\gamma}}ds + \Big((2g+1) - \frac{(2\gamma+1)(g+1)(2g+1)}{\gamma+g+1} + \frac{\gamma(g+1)(2g+1)}{\gamma+1}\Big)\int_0^1 \frac{R(1,t)}{t^{1-g}}dt\Big]$$
$$= (1-\nu^2)(1+\gamma)Q_2,$$

$$Cov(\tilde{\Omega}, \Sigma_\gamma) = (1-\nu^2)\frac{(1+\gamma)(1+g)^2}{\gamma g}\Big[\Big(\frac{(\gamma+1)^2+g-(\gamma+1)^2(g+1)}{(g+1)(\gamma+1)}\Big)R(1,1)$$
$$+\frac{\gamma^2}{\gamma+1}\int_0^1 \frac{R(1,t)}{t}dt + \Big((2\gamma+1) - \frac{(2g+1)(\gamma+1)(2\gamma+1)}{\gamma+g+1} + \frac{g(\gamma+1)(2\gamma+1)}{g+1}\Big)$$
$$\int_0^1 \frac{R(s,1)}{s^{1-\gamma}}ds + \Big(\frac{2g+1}{g+1} - \frac{(2g+1)(\gamma+1)(2\gamma+1)}{\gamma+g+1} + (2g+1)(\gamma+1) - \frac{2g+1}{\gamma+1}\Big)\int_0^1 \frac{R(1,t)}{t^{1-g}}dt\Big]$$
$$= (1-\nu^2)(1+\gamma)Q_1.$$

Hence,

$$Cov\left(\Omega - \frac{1+\gamma}{1+g}R_g\tilde{\Omega}, \Sigma_\gamma - \frac{S_g}{1+(1+g)^2}\tilde{\Sigma}_g\right)$$
$$= -(1+\gamma)\left[1+(1-\nu^2)\left(\frac{R_gS_g}{1+(1+g)^2} + \frac{1+\gamma}{1+g}R_gQ_1 + \frac{S_gQ_2}{1+(1+g)^2}\right)\right]$$
$$= -(1+\gamma)[1+(1-\nu^2)Q].$$

Since Ω and $W_1(1)$ are uncorrelated, we have

$$Cov\left(\Omega - \frac{1+\gamma}{1+g}R_g\tilde{\Omega}, W_1(1)\right) = Cov\left(-\frac{1+\gamma}{1+g}R_g\tilde{\Omega}, W_1(1)\right)$$
$$= (1-\nu^2)\frac{(1+\gamma)(1+g)}{g}R_g\Big[(2g+1)\int_0^1 \frac{R(1,s)}{s^{1-g}}ds - \int_0^1 \frac{R(1,s)}{s}ds - \frac{g}{g+1}R(1,1)\Big] = (1-\nu^2)M_1.$$

Note that $Cov(\Sigma_\gamma, W_1(1)) = \gamma$ (and not 0 as incorrectly assumed in de Haan and Ferreira

(2006), page 139). Hence

$$
\begin{aligned}
&Cov\left(\Sigma_\gamma - \frac{S_g}{1+(1+g)^2}\tilde{\Sigma}_g, W_1(1)\right)\\
&= \gamma - (1-\nu^2)\frac{(1+g)}{g(1+(1+g)^2)}S_g\Big[(2g+1)(g+1)\int_0^1 \frac{R(1,s)}{s^{1-g}}ds - \int_0^1 \frac{R(1,s)}{s}ds - \frac{g(g+2)}{g+1}R(1,1)\Big]\\
&= \gamma + (1-\nu^2)M_2.
\end{aligned}
$$

□

Proof of Theorem 2.3. Using Theorem 2.2 in combination with Theorem 4.3.1 in de Haan and Ferreira (2006) yields, as $n \to \infty$,

$$
\sqrt{k}\frac{\hat{x}_{p_g} - x_p}{a\left(\frac{n}{k}\right)q_\gamma\left(\frac{k}{np}\right)} \xrightarrow{d} \Omega - \frac{1+\gamma}{1+g}R_g\tilde{\Omega} - \gamma_-\left(\Sigma_\gamma - \frac{\lambda\rho}{(1-\rho)(1+\gamma-\rho)} - \frac{S_g}{1+(1+g)^2}\tilde{\Sigma}_g\right)
$$
$$
+(\gamma_-)^2 W_1(1) - \lambda\frac{\gamma_-}{\gamma_-+\rho},
$$

where $\gamma_- := \min(0,\gamma)$. The distribution of the limiting random variable is easily seen to be that in (12). □

D Additional simulation results

We provide additional simulation results when the target variable and the covariate are tail independent. More specifically, we begin with simulating data from the bivariate normal distribution with standard normal marginals and correlation 0.5, restricted to the first quadrant. Again, these data are denoted by $(\check{X}_i, Y_i)$. To obtain the data for estimation, (X_i, Y_i) with X_i having extreme value index $\gamma = -0.3, 0, 0.3$, we transform $\check{X}_i$ in the same way as in (16).

Table 1 shows the empirical percentages of variance reduction for different values of g. The results are based on 10,000 replications with $n = 500$, $m = 1000$ and $k = 125$. There is a variance inflation ranging from less than 5% to slightly more than 10%.

Table 1: Variance reduction for different extreme value indices under tail independence

g	γ		
	-0.3	0	0.3
-0.25	-12.1%	-9.0%	-5.6%
-0.125	-12.9%	-9.4%	-5.6%
0	-12.2%	-8.9%	-5.0%
0.125	-11.3%	-8.1%	-4.3%
0.25	-10.4%	-7.4%	-3.7%

References

de Haan, L. and A. Ferreira (2006). *Extreme Value Theory: an Introduction*. Springer.

Drees, H., A. Ferreira, and L. de Haan (2004). On maximum likelihood estimation of the extreme value index. *The Annals of Applied Probability 14*, 1179–1201.

Grimshaw, S. D. (1993). Computing maximum likelihood estimates for the generalized Pareto distribution. *Technometrics 35*(2), 185–191.

Zhou, C. (2009). Existence and consistency of the maximum likelihood estimator for the extreme value index. *Journal of Multivariate Analysis 100*(4), 794–815.