SUPPLEMENTARY MATERIAL FOR "EXTREME VALUE INFERENCE FOR HETEROGENEOUS POWER LAW DATA"

BY JOHN H.J. EINMAHL[1,a], YI HE[2,b]

[1]Department of Econometrics & OR and CentER, Tilburg University, Tilburg, The Netherlands, [a]j.h.j.einmahl@uvt.nl

[2] Amsterdam School of Economics, University of Amsterdam, Amsterdam, The Netherlands, [b]yi.he@uva.nl

The supplementary material consist of two sections. Section 1 provides more simulation results for the extreme quantile estimator. Section 2 provides proofs of Theorems 2.3 and 4.1 and Lemmas 7.9–7.11.

1. More Simulation Results. Boxplots for the extreme quantile estimator $\widehat{x}_\tau$ for $\tau = 1/200$ for the distributions in Subsection 5.1.

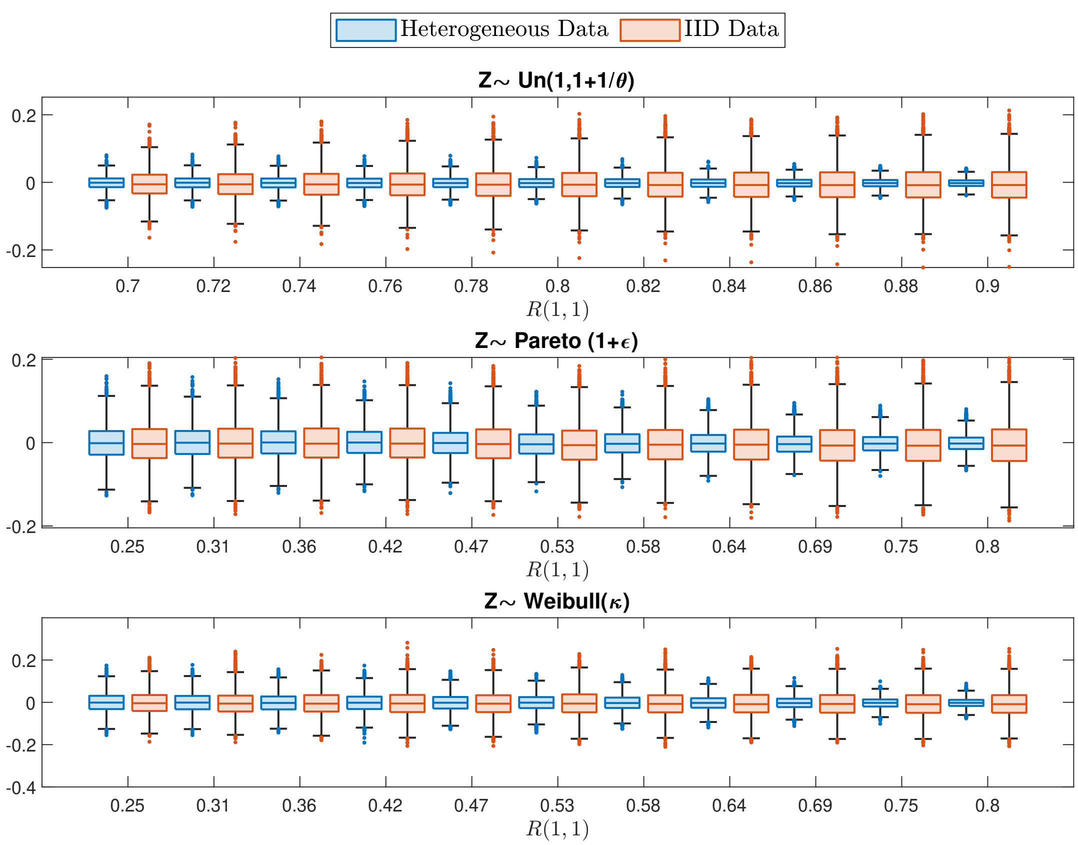

FIG 1. *Boxplots of Relative Estimation Error* $\log \widehat{x}_\tau / \log x_\tau - 1$ *for* $\tau = 1/200$

Boxplots for the extreme quantile estimator $\widehat{x}_\tau$ for $\tau = 1/200$ for the distributions in Subsection 5.2.

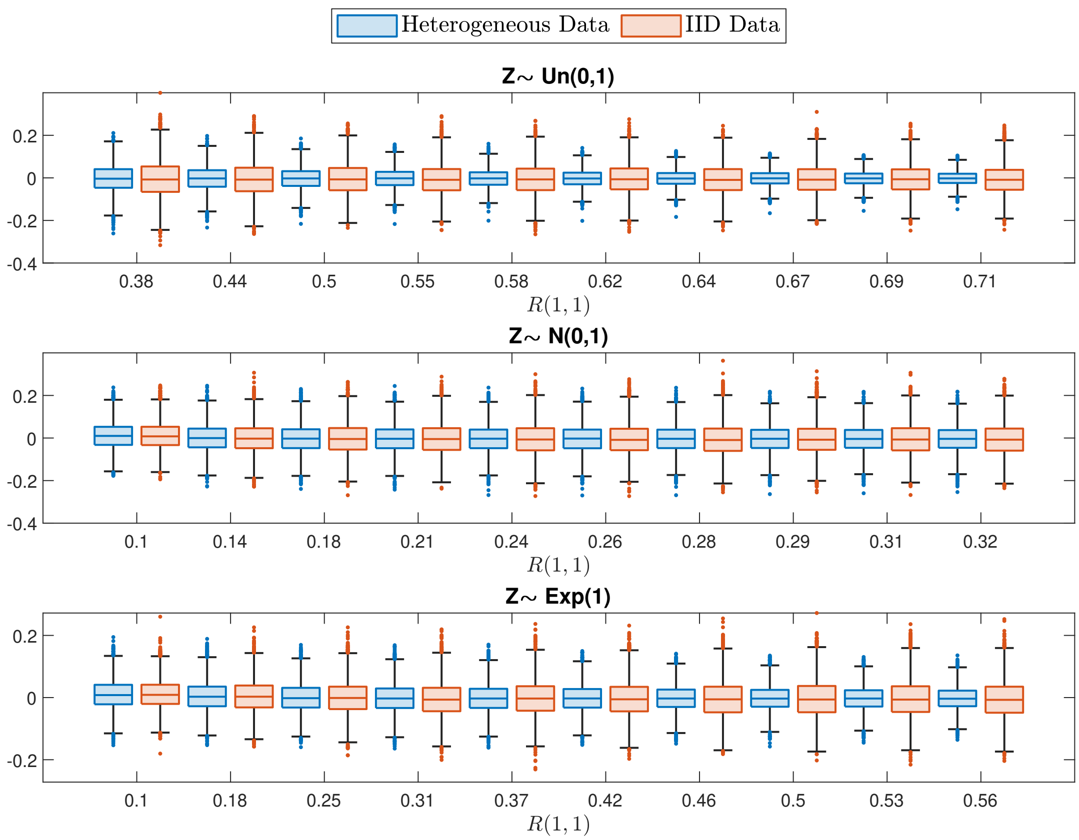

FIG 2. *Boxplots of Relative Estimation Error* $\log \widehat{x}_\tau / \log x_\tau - 1$ *for* $\tau = 1/200$

2. Additional Proofs.

PROOF OF THEOREM 2.3. We only need to show the pointwise consistency of $\widehat{R}(x,y)$ for $x, y > 0$; the uniform consistency follows by the monotonicity of $\widehat{R}$ and R and the homogeneity of R as in the proof of Theorem 7.2.1 in de Haan and Ferreira (2006).

Define $U_i^{(p)} = T_p(X_i^{(p)})$ and $V_i^{(p)} = T_p(\widetilde{X}_i^{(p)})$. Let $U_{1,p} \le U_{2,p} \ldots \le U_{p,p}$ be the order statistics of $\{U_i^{(p)}\}$, and define $V_{i,p}$ similarly. Observe that, with probability 1, we have

$$\widehat{R}(x,y) = \frac{1}{k}\sum_{i=1}^{p} \mathbb{1}\left[U_i^{(p)} < U_{\lceil kx+1\rceil,p}, V_i^{(p)} < V_{\lceil ky+1\rceil,p}\right] = R_{p,k}\left(\frac{p}{k}U_{\lceil kx+1\rceil,p}, \frac{p}{k}V_{\lceil ky+1\rceil,p}\right),$$

where

$$R_{p,k}(x,y) = \frac{1}{k}\sum_{i=1}^{p} \mathbb{1}\left[U_i^{(p)} < kx/p, V_i^{(p)} < ky/p\right].$$

Using (1) from Lemma 2.1, for every $x, y > 0$,

$$\mathbb{E}R_{p,k}(x,y) = \frac{1}{k}\sum_{i=1}^{p} \mathbb{P}\left(U_i^{(p)} < kx/p\right)\mathbb{P}\left(U_i^{(p)} < ky/p\right) \to R(x,y).$$

Furthermore, for every $x, y > 0$,

$$\operatorname{var}\left(R_{p,k}(x,y)\right) \le \frac{1}{k^2}\sum_{i=1}^{p} \mathbb{P}\left(U_i^{(p)} < kx/p, V_i^{(p)} < ky/p\right) = \frac{1}{k}\mathbb{E}R_{p,k}(x,y) \to 0.$$

This implies that $R_{p,k}(x,y) \xrightarrow{\mathbb{P}} R(x,y)$ pointwise. The convergence is then uniform by continuity of R and the monotonicity of $R_{p,k}$ and R.

It remains to show that

$$\frac{p}{k}U_{\lceil kx+1\rceil,p} \xrightarrow{\mathbb{P}} x, \quad \frac{p}{k}V_{\lceil ky+1\rceil,p} \xrightarrow{\mathbb{P}} y. \tag{2.1}$$

We only need to prove the first one; the second one then follows. Take the intermediate sequence $\widetilde{k} = \lceil kx+1\rceil$ and accordingly the intermediate quantile $\widetilde{u}_p = Q_p(1-\widetilde{k}/p)$. Observe that Theorem 2.1 implies that $X_{p-\widetilde{k}:p}/\widetilde{u}_p \xrightarrow{\mathbb{P}} 1$. This implies that, for any $c > 1$, with probability tending to 1

$$\frac{p}{\widetilde{k}}T_p\left(X_{p-\widetilde{k}:p}\right) \ge \frac{p}{\widetilde{k}}T_p(c\widetilde{u}_p) = \frac{p}{\widetilde{k}}T_p(\widetilde{u}_p)\cdot\frac{T_p(c\widetilde{u}_p)}{T_p(\widetilde{u}_p)} \to 1\cdot c^{-1/\gamma},$$

and similarly

$$\frac{p}{\widetilde{k}}T_p\left(X_{p-\widetilde{k}:p}\right) \le \frac{p}{\widetilde{k}}T_p(c^{-1}\widetilde{u}_p) = \frac{p}{\widetilde{k}}T_p(\widetilde{u}_p)\cdot\frac{T_p(c^{-1}\widetilde{u}_p)}{T_p(\widetilde{u}_p)} \to 1\cdot c^{1/\gamma}.$$

As c can be arbitrarily close to 1, $\frac{p}{k}U_{\widetilde{k},p} = \frac{\widetilde{k}}{k}\cdot\frac{p}{\widetilde{k}}T_p\left(X_{p-\widetilde{k}:p}\right) \xrightarrow{\mathbb{P}} x\cdot 1 = x$. □

PROOF OF THEOREM 4.1. Write $d_p = k/(p\tau)$. Then

$$\frac{\sqrt{k}}{\log(k/(p\tau))}\left(\frac{\widehat{x}_\tau}{x_\tau}-1\right) = \frac{\sqrt{k}}{x_\tau \log d_p}\left(X_{p-k:p}d_p^{\widehat{\gamma}} - x_\tau\right)$$

$$= \frac{d_p^{\gamma}u_p}{x_\tau}\cdot\frac{\sqrt{k}}{\log d_p}\left(\left(\frac{X_{p-k:p}}{u_p}-1\right)d_p^{\widehat{\gamma}-\gamma} + \left(d_p^{\widehat{\gamma}-\gamma}-1\right) - \left(\frac{x_\tau d_p^{-\gamma}}{u_p}-1\right)\right).$$

Using condition (8), we obtain that

$$\frac{d_p^{\gamma}u_p}{x_\tau} \to 1 \quad \text{and} \quad \frac{\sqrt{k}}{\log d_p}\left(\frac{x_\tau d_p^{-\gamma}}{u_p}-1\right) \to 0.$$

Hence it remains to show that

$$\frac{\sqrt{k}}{\log d_p}\left(\frac{X_{p-k:p}}{u_p}-1\right)d_p^{\widehat{\gamma}-\gamma} + \frac{\sqrt{k}}{\log d_p}\left(d_p^{\widehat{\gamma}-\gamma}-1\right)$$

converges in distribution to the normal specified in the theorem. Now, Theorem 2.2 readily yields

$$\frac{\sqrt{k}}{\log d_p}\left(d_p^{\widehat{\gamma}-\gamma}-1\right) \xrightarrow{a.s} \gamma\left(-V(1) + \int_0^1 V(x)\frac{dx}{x}\right)$$

and from Corollary 2.2 we obtain

$$\frac{\sqrt{k}}{\log d_p}\left(\frac{X_{p-k:p}}{u_p}-1\right)d_p^{\widehat{\gamma}-\gamma}\xrightarrow{a.s.}\frac{\gamma V(1)}{\log 1/\nu}.$$

The limiting variance follows by computing the variance of the sum of the two limits. □

PROOF OF LEMMA 7.9. There exists $\varepsilon > 0$ such that $S(\varepsilon) > 0$. Furthermore, for some constant M

$$M\varepsilon^{-1/\gamma}(U_\sigma(t))^{-1/\gamma}\geq T(\varepsilon U_\sigma(t))\geq T_\sigma(U_\sigma(t))S(\varepsilon)=t^{-1}S(\varepsilon),$$

where the first inequality is due to (12). Rewriting the inequality completes the proof of the first part. The result for S can be seen to hold by interchanging T_σ and S in the definition of T. □

PROOF OF LEMMA 7.10. Let $t = t(p) \to \infty$ be any intermediate threshold sequence such that $pT(t) \to \infty$. It is possible to take a small (probability) sequence $\beta = \beta(p) = o(T(t))$ but $p\beta \to \infty$. Consider the truncated function $T_p^+(t) = \frac{1}{p}\sum_{i=\lfloor p\beta\rfloor+1}^{p} S\left(\frac{t}{Q_\sigma(1-i/p)}\right)$ and $T^+(t) = \int_{\lfloor p\beta\rfloor/p}^{1} S\left(\frac{t}{Q_\sigma(1-u)}\right)du$. Since survival probabilities are bounded by 1,

$$|T_p(t)-T_p^+(t)|+|T-T^+(t)|\leq\beta+\beta=o(T(t)).$$

For Assumption 2.1, it remains to show that

$$|T_p^+(t)-T^+(t)|=o(T(t)).$$

For $i = \lfloor p\beta\rfloor + 1, \ldots, p$, Lemma 7.3 gives that

$$\int_{(i-1)/p}^{i/p} S\left(\frac{t}{Q_\sigma(1-u)}\right)-S\left(\frac{t}{Q_\sigma(1-i/p)}\right)du\leq M\int_{(i-1/p}^{i/p} T\left(\frac{t}{Q_\sigma(1-u)}\right)\log\frac{Q_\sigma(1-u)}{Q_\sigma(1-i/p)}\,du.$$

By the Lipschitz-continuity of $\log U_\sigma(e^t)$, for some constants K, M and for $\lfloor p\beta\rfloor \leq (i-1) \leq pu \leq i$,

$$\log\frac{Q_\sigma(1-u)}{Q_\sigma(1-i/p)}\leq K\log\frac{i}{i-1}\leq\frac{M}{i}\leq\frac{M}{pu}.$$

Together with Lemma 7.9 and (12), for some constant C

$$\int_{(i-1)/p}^{i/p}\left(S\left(\frac{t}{Q_\sigma(1-u)}\right)-S\left(\frac{t}{Q_\sigma(1-i/p)}\right)\right)du$$

$$\leq\frac{M}{p}\int_{(i-1)/p}^{i/p}\frac{1}{u}T\left(\frac{t}{Q_\sigma(1-u)}\right)du\leq\frac{C}{p}\int_{(i-1)/p}^{i/p}\frac{1}{u}\left(tu^\gamma\right)^{-1/\gamma}du=\frac{Ct^{-1/\gamma}}{p}\int_{(i-1)/p}^{i/p}\frac{1}{u^2}du.$$

Summing up the bounds over i,

$$\left|T_p^+(t)-T^+(t)\right|\leq\frac{Ct^{-1/\gamma}}{p}\int_{\beta/2}^{1}\frac{1}{u^2}du\leq\frac{2Ct^{-1/\gamma}}{p\beta}=o(T(t)).$$

□

PROOF OF LEMMA 7.11. By monotonicity of h and Q_σ, for all $x > 0$

$$\begin{aligned} x f_p(x) &= \frac{1}{p}\sum_{i=1}^{p} g\left(\frac{x}{Q_\sigma(1-i/p)}\right)\frac{x}{Q_\sigma(1-i/p)} \leq \frac{1}{p}\sum_{i=1}^{p} h\left(\frac{x}{Q_\sigma(1-i/p)}\right) \\ &\leq \int_0^1 h\left(\frac{x}{Q_\sigma(1-u)}\right) du = \int_0^\infty h\left(\frac{x}{u}\right) dF_\sigma(u) \leq M T(x). \end{aligned}$$

Similarly, using monotonicity of S, for $x > 0$

$$T_p(x) := \frac{1}{p}\sum_{i=1}^{p} S\left(\frac{x}{Q_\sigma(1-i/p)}\right) \leq \int_0^1 S\left(\frac{x}{Q_\sigma(1-u)}\right) du = T(x).$$

□

REFERENCES

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