Journal of the Korean Statistical Society (2004), **33**: *3, pp 339–351*

GENERALIZING THE REFINED PICKANDS ESTIMATOR OF THE EXTREME VALUE INDEX†

SEOKHOON YUN[1]

ABSTRACT

In this paper we generalize and improve the refined Pickands estimator of Drees (1995) for the extreme value index. The finite-sample performance of the refined Pickands estimator is not good particularly when the sample size n is small. For each fixed $k = 1, 2, \ldots$, a new estimator is defined by a convex combination of k different generalized Pickands estimators and its asymptotic normality is established. Optimal weights defining the estimator are also determined to minimize the asymptotic variance of the estimator. Finally, letting k depend upon n, we see that the resulting estimator has a better finite-sample behavior as well as a better asymptotic efficiency than the refined Pickands estimator.

AMS 2000 subject classifications. Primary 62G32; Secondary 62F12.
Keywords. Extreme value index, refined Pickands estimator, asymptotic normality, adaptive estimator.

1. INTRODUCTION

Let $X_1, \ldots, X_n$ be an *iid* sample from an unknown *df* F. Suppose that F belongs to the domain of attraction of an extreme value distribution G_β for some $\beta \in \mathbb{R}$ $[F \in \mathcal{D}(G_\beta)]$, where $G_\beta(x) := \exp\{-(1+\beta x)^{-1/\beta}\}$, $1 + \beta x > 0$. The case $\beta = 0$ is always interpreted as the limit when $\beta \to 0$, and so $G_0(x) = \exp(-e^{-x})$, $x \in \mathbb{R}$. The parameter β called the extreme value index (or tail index) of F is a kind of measure indicating how heavy the right tail of the underlying *df* F is. A lot of efforts have been made in the literature to estimate β using the sample $X_1, \ldots, X_n$ (*e.g.* see Hill, 1975; Pickands, 1975; Smith, 1987; Dekkers and de Haan, 1989; Dekkers *et al.*, 1989; Drees, 1995). In statistical extreme value

Received May 2004; accepted July 2004.
†This work was supported by grant No. 2000-1-10400-009-2 from the Basic Research Program of the Korea Science & Engineering Foundation.
[1]Department of Applied Statistics, University of Suwon, Suwon 445-743, Korea

analysis, the estimation of β is deeply related to the estimation of high quantiles of F.

For heavy tails, *i.e.* for $\beta > 0$, the most prominent estimator of β is the Hill (1975) estimator, which is however unfortunately not consistent for $\beta < 0$. As a consistent estimator for any $\beta \in \mathbb{R}$, an alternative is the Pickands (1975) estimator, which is, if it is assumed that only the m upper order statistics from the sample are to be used, defined by

$$\hat{\beta}^{(P)}_{n,m} := \frac{1}{\log 2} \log \frac{X^{(n)}_{\lfloor m/4 \rfloor} - X^{(n)}_{\lfloor m/2 \rfloor}}{X^{(n)}_{\lfloor m/2 \rfloor} - X^{(n)}_{m}}, \; 4 \le m \le n,$$

where $X^{(n)}_1 \ge X^{(n)}_2 \ge \cdots \ge X^{(n)}_n$ and $\lfloor x \rfloor$ denote the descending order statistics of $X_1, \ldots, X_n$ and the integer part of $x \in \mathbb{R}$, respectively. The Pickands estimator is easy to compute and invariant under location and scale transformations of the sample, but its main disadvantage is its poor asymptotic efficiency.

To overcome this drawback, Drees (1995) considered a refined Pickands estimator of form

$$\hat{\beta}^{(D)}_{n,m}(q_0, \ldots, q_{i(m)}) := \sum_{j=0}^{i(m)} q_j \hat{\beta}^{(P)}_{n,\lfloor m/2^j \rfloor}, \; q_j \ge 0, \; \sum_{j=0}^{i(m)} q_j = 1, \tag{1.1}$$

where $i(m) := \lfloor \log m / \log 2 - 2 \rfloor$, which is a convex combination of several different Pickands estimators (see Drees, 1995 for a more general form of a refined Pickands estimator). Under a certain second order condition on the quantile function of F, he established asymptotic normality of $\hat{\beta}^{(D)}_{n,m}(q_0, \ldots, q_{i(m)})$ for some intermediate sequence $m = m(n)$ and then, for $\beta \ne -1/2$, he also computed optimal weights $q^*_j(\beta)$, $j = 0, 1, \ldots, i(m)$, minimizing its asymptotic variance as follows:

$$\begin{cases} q^*_j(\beta) = \dfrac{1 - 2^{-(j+1)g(\beta)}}{2^{g(\beta)} - 1} \left(2^{g(\beta)+1} - 1\right) 2^{-j-2}, \; j = 0, 1, \ldots, i(m) - 1, \\ q^*_{i(m)}(\beta) = \dfrac{2^{g(\beta)+1} - 2^{-i(m)g(\beta)} - 1}{2^{g(\beta)} - 1} 2^{-i(m)-1}, \end{cases} \tag{1.2}$$

where $g(\beta) := |\beta + 1/2| - 1/2$, the case $\beta = -1$ being interpreted as the limit when $\beta \to -1$. By intermediate sequence we mean a sequence of positive integers $m = m(n)$ such that $m \to \infty$ and $m/n \to 0$ as $n \to \infty$. The resulting minimal asymptotic variance of $\sqrt{m}(\hat{\beta}^{(D)}_{n,m}(q_0, \ldots, q_{i(m)}) - \beta)$ is then given by

$$\left(\sigma^{(D)}_{\beta}\right)^2 := \frac{\beta^2 \left(2^{g(\beta)+1} - 1\right)^2}{\left(2^{\beta} - 1\right)^2 2^{g(\beta)-\beta+2} (\log 2)^2}, \tag{1.3}$$

which is valid even for $\beta = -1/2$. Since Drees did not find optimal weights $q_j^*(\beta)$ for $\beta = -1/2$ which lead to (1.3), he modified the asymptotically optimal weights for β on a small neighborhood of $-1/2$: for $j = 0, 1, \ldots, i(m)$,

$$q_j(\beta) := \begin{cases} q_j^*(\beta), & \text{if } \left|\beta + \frac{1}{2}\right| > \delta\,, \\ q_j^*(-\frac{1}{2} + \delta), & \text{if } \left|\beta + \frac{1}{2}\right| \leq \delta\,, \end{cases} \tag{1.4}$$

for some small $\delta > 0$. Then the asymptotic variance of

$$\sqrt{m}\left(\hat{\beta}_{n,m}^{(D)}(q_0(\beta), \ldots, q_{i(m)}(\beta)) - \beta\right)$$

can be made arbitrarily close to $(\sigma_\beta^{(D)})^2$ for all $\beta \in \mathbb{R}$ if a suitable $\delta > 0$ is chosen. Since $q_j(\beta)$ depends upon β, a data-driven version of $q_j(\beta)$ is finally plugged into $\hat{\beta}_{n,m}^{(D)}(q_0(\beta), \ldots, q_{i(m)}(\beta))$, with the resulting estimator having the same asymptotic performance as $\hat{\beta}_{n,m}^{(D)}(q_0(\beta), \ldots, q_{i(m)}(\beta))$.

The refined Pickands estimator in (1.1) with almost optimal weights given by (1.4) clearly improves the asymptotic efficiency of the original Pickands estimator. However, since the refined Pickands estimator is eventually based on infinite number of weights as $n \to \infty$ and since the optimal weights in (1.2) are determined so that the infinitely many $q_j^*(\beta)$, $j = 0, 1, \ldots$, minimize the asymptotic variance of the estimator, its finite-sample performance is not good enough particularly when the sample size n is small. Notice that the value of $i(m)$ is only 7 even for a fairly large $n = m = 1000$.

The purpose of this paper is to generalize and improve the refined Pickands estimator in such a way that new estimator should have a better finite-sample behavior as well as a better asymptotic efficiency than the refined Pickands estimator. This can be achieved by first replacing $i(m)$ in (1.1) by a fixed i to derive asymptotic normality of $\hat{\beta}_{n,m}^{(D)}(q_0, \ldots, q_i)$ and then letting $i = i(m) \to \infty$ as $n \to \infty$, where $m = m(n)$ is an intermediate sequence.

Specifically, we first consider a class of estimators defined by, for $k = 1, 2, \ldots$,

$$\hat{\beta}_{n,m,\theta}\,(p_{k0}, \ldots, p_{kk}) := \sum_{j=0}^{k} p_{kj}\hat{\beta}_{n,\lfloor m\theta^j \rfloor}(\theta), \tag{1.5}$$

where $\theta \in (0,1)$, $p_{kj} \geq 0$, $\sum_{j=0}^{k} p_{kj} = 1$ and

$$\hat{\beta}_{n,m}(\theta) := \frac{1}{\log\theta} \log \frac{X_m^{(n)} - X_{\lfloor m\theta \rfloor}^{(n)}}{X_{\lfloor m\theta \rfloor}^{(n)} - X_{\lfloor m\theta^2 \rfloor}^{(n)}}, \quad 1 \leq \lfloor m\theta^2 \rfloor < m \leq n.$$

Note that $\hat{\beta}_{n,m}(\theta)$ is a generalization of the Pickands estimator $\hat{\beta}^{(P)}_{n,m}$ ($\hat{\beta}_{n,m}(1/2) = \hat{\beta}^{(P)}_{n,m}$) and is known to be consistent for any $\beta \in \mathbb{R}$ and any intermediate sequence $m = m(n)$ (*cf.* Pereira, 1994; Yun, 2002a).

In Section 2, asymptotic normality of $\hat{\beta}_{n,m,\theta}(p_{k0}, \ldots, p_{kk})$ is established. In Section 3, we derive explicit formulas for the optimal weights $p^*_{kj}(\theta, \beta)$, $j = 0, 1, \ldots, k$, that minimize the asymptotic variance of $\sqrt{m}(\hat{\beta}_{n,m,\theta}(p_{k0}, \ldots, p_{kk}) - \beta)$. Then it turns out that $\lim_{k\to\infty} p^*_{kj}(1/2, \beta) = q^*_j(\beta)$, $j = 0, 1, \ldots$, for $\beta \neq -1/2$ and that the asymptotic variance of

$$\sqrt{m}\left\{\hat{\beta}_{n,m,1/2}\Big(p^*_{k0}(\frac{1}{2}, \beta), \ldots, p^*_{kk}(\frac{1}{2}, \beta)\Big) - \beta\right\}$$

converges to $(\sigma^{(D)}_\beta)^2$ in (1.3) as $k \to \infty$ for all $\beta \in \mathbb{R}$. As a result, if we let $k = k(\theta, m) := \lfloor -\log m / \log\theta - 2 \rfloor$ where $m = m(n)$ is an intermediate sequence, we have that the asymptotic distribution of

$$\sqrt{m}\left\{\hat{\beta}_{n,m,1/2}\Big(p^*_{k(1/2,m),0}(\frac{1}{2}, \beta), \ldots, p^*_{k(1/2,m),k(1/2,m)}(\frac{1}{2}, \beta)\Big) - \beta\right\}$$

coincides with that of $\sqrt{m}(\hat{\beta}^{(D)}_{n,m}(q^*_0(\beta), \ldots, q^*_{i(m)}(\beta)) - \beta)$ for $\beta \neq -1/2$, which is normal. Finally, a data-driven version of $p^*_{k(\theta,m),j}(\theta, \beta)$ is plugged into

$$\hat{\beta}_{n,m,\theta}(p^*_{k(\theta,m),0}(\theta, \beta), \ldots, p^*_{k(\theta,m),k(\theta,m)}(\theta, \beta)),$$

with its asymptotic performance unchanged. Unlike the refined Pickands estimator, the optimal weights $p^*_{kj}(\theta, \beta)$ are well defined for all $\beta \in \mathbb{R}$. Moreover, the new estimator has a better finite-sample behavior than the refined Pickands estimator. If $\theta > 1/2$, then the new estimator is also asymptotically better than the refined Pickands estimator.

2. ASYMPTOTIC NORMALITY

Let F^{-1} denote the quantile function of the underlying *df* F and define the function U by $U(x) := F^{-1}(1 - 1/x)$, $x > 1$. Then the condition $F \in \mathcal{D}(G_\beta)$ holds for some $\beta \in \mathbb{R}$ if and only if there exists a positive function $a(t)$ such that, for $x > 0$,

$$\frac{U(tx) - U(t)}{a(t)} = \frac{x^\beta - 1}{\beta} + R(t, x), \;\; R(t, x) = o(1) \text{ as } t \to \infty \tag{2.1}$$

(*cf.* de Haan, 1984). In this case, the function $a(t)$ must be regularly varying at infinity with index β [$a(t) \in RV_\beta$], that is, $a(t)$ must satisfy

$$\lim_{t\to\infty} \frac{a(tx)}{a(t)} = x^\beta, \ x > 0.$$

For asymptotic normality of $\hat{\beta}_{n,m,\theta}(p_{k0}, \ldots, p_{kk})$, one needs to consider the second order behavior of the function U. Among several second order conditions introduced in the literature (*cf.* Smith, 1987; Dekkers and de Haan, 1989; Pereira, 1994; Drees, 1995), the most general form was given by de Haan and Stadtmüller (1996), who assumed the existence of a function $A(t)$ of constant sign for large values of t such that, for $x > 0$,

$$R(t,x) = A(t)H(x) + o(A(t)), \ A(t) = o(1) \text{ as } t \to \infty, \tag{2.2}$$

where

$$H(x) := \frac{1}{\rho}\left(\frac{x^{\beta+\rho}-1}{\beta+\rho} - \frac{x^\beta - 1}{\beta}\right)$$

for some $\rho \le 0$. In this case, we must have $|A(t)| \in RV_\rho$. As for the case $\beta = 0$, the case $\rho = 0$ is interpreted as the limit when $\rho \to 0$.

We work with a slightly stronger assumption than $F \in \mathcal{D}(G_\beta)$. We assume that F belongs to the differentiable domain of attraction of G_β for some $\beta \in \mathbb{R}$ [$F \in \mathcal{D}_{dif}(G_\beta)$], *i.e.* we assume that F is differentiable in a left neighborhood of the right endpoint of F and that there exist constants $a_n > 0$ and $b_n \in \mathbb{R}$ such that

$$\lim_{n\to\infty} \frac{d}{dx} F^n(a_n x + b_n) = G'_\beta(x)$$

locally uniformly in x with $1 + \beta x > 0$. If $F \in \mathcal{D}_{dif}(G_\beta)$, then $F \in \mathcal{D}(G_\beta)$. A useful necessary and sufficient condition for $F \in \mathcal{D}_{dif}(G_\beta)$ is that $U(t)$ is differentiable for all sufficiently large t and $tU'(t) \in RV_\beta$ (*cf.* Corollary 2.1 of Pereira, 1994).

The following theorem establishes asymptotic normality of $\hat{\beta}_{n,m,\theta}(p_{k0}, \ldots, p_{kk})$, which can be proved using a method similar to one used in Yun (2002b) and so its proof is omitted. By $\xrightarrow{d}$ and $\xrightarrow{p}$ we denote convergence in distribution and convergence in probability, respectively.

THEOREM 2.1. *Suppose $F \in \mathcal{D}_{dif}(G_\beta)$ for some $\beta \in \mathbb{R}$, so that (2.1) holds. Let $\theta \in (0,1)$. Then, for $k = 1, 2, \ldots$ and any $p_{k0}, \ldots, p_{kk} \ge 0$ with $\sum_{j=0}^k p_{kj} = 1$, the following hold.*

(a) *If* $R(t,x) \equiv 0$, *then*

$$\sqrt{m}\left(\hat{\beta}_{n,m,\theta}(p_{k0},\dots,p_{kk}) - \beta\right) \xrightarrow{d} N\big(0, \sigma^2_{\theta,\beta}(p_{k0},\dots,p_{kk})\big) \quad \text{as } n \to \infty$$

for any intermediate sequence $m = m(n)$.

(b) *For* $R(t,x)$ *not being identically zero, assume further that (2.2) holds for some* $\rho \le 0$. *Let* $m = m(n)$ *be an intermediate sequence such that*

$$\sqrt{m} A\left(\frac{n}{m}\right) \to \lambda$$

as $n \to \infty$ *for some* $\lambda \in [-\infty, \infty]$. *Then,*

(i) *for* $\lambda \in (-\infty, \infty)$,

$$\sqrt{m}\left(\hat{\beta}_{n,m,\theta}(p_{k0},\dots,p_{kk}) - \beta\right)$$
$$\xrightarrow{d} N\big(\lambda b_{\theta,\beta,\rho}(p_{k0},\dots,p_{kk}), \sigma^2_{\theta,\beta}(p_{k0},\dots,p_{kk})\big)$$

as $n \to \infty$,

(ii) *for* $\lambda = \pm\infty$,

$$\left(A\left(\frac{n}{m}\right)\right)^{-1}\left(\hat{\beta}_{n,m,\theta}(p_{k0},\dots,p_{kk}) - \beta\right) \xrightarrow{p} b_{\theta,\beta,\rho}(p_{k0},\dots,p_{kk})$$

as $n \to \infty$, *where*

$$b_{\theta,\beta,\rho}(p_{k0},\dots,p_{kk}) := \frac{\beta(1-\theta^{-\rho})(1-\theta^{-\beta-\rho})}{(1-\theta^{-\beta})\rho(\beta+\rho)\log\theta}\sum_{j=0}^{k}\theta^{-j\rho}p_{kj}$$

and

$$\sigma^2_{\theta,\beta}(p_{k0},\dots,p_{kk}) := \frac{\beta^2\theta^{-\beta-1}(1-\theta)}{(1-\theta^{-\beta})^2(\log\theta)^2} \times \left\{\theta^{\beta}\left(1+\theta^{-2\beta-1}\right)\sum_{j=0}^{k}\theta^{-j}p_{kj}^2 - 2\sum_{j=1}^{k}\theta^{-j}p_{k,j-1}p_{kj}\right\}.$$

It is worthy of noting that (2.1) holds with $R(t,x) \equiv 0$ if and only if F is a generalized Pareto *df* up to a scale and location parameter (*cf.* Theorem 3.1 of Pereira, 1994).

3. OPTIMAL WEIGHTS

The optimal values of $p_{k0}, \ldots, p_{kk}$ minimizing the asymptotic mean squared error of the estimator $\hat{\beta}_{n,m,\theta}(p_{k0}, \ldots, p_{kk})$ generally depend upon ρ as well as θ and β, and it is even untractable to derive their explicit forms. For simplicity we therefore consider only the estimators of β which are asymptotically unbiased. Then the optimal weights are the values of $p_{k0}, \ldots, p_{kk}$ minimizing the asymptotic variance of the estimator $\hat{\beta}_{n,m,\theta}(p_{k0}, \ldots, p_{kk})$, which depend only upon θ and β as follows. Recall that $g(\beta) = |\beta + 1/2| - 1/2$, $\beta \in \mathbb{R}$ and that the cases $\beta = 0$ and $\beta = -1$ are interpreted as the limits when $\beta \to 0$ and $\beta \to -1$, respectively. Further, the case $\beta = -1/2$ is also interpreted as the limit when $\beta \to -1/2$.

LEMMA 3.1. *Let $\theta \in (0,1)$. Then, for $k = 1, 2, \ldots$ and $\beta \in \mathbb{R}$, $\sigma^2_{\theta,\beta}(p_{k0}, \ldots, p_{kk})$ in Theorem 2.1 is minimized at $p_{kj} = p^*_{kj}(\theta, \beta)$, $j = 0, 1, \ldots, k$, where*

$$\begin{aligned}
&p^*_{kj}(\theta,\beta) \\
&:= \frac{(1-\theta)}{d_k(\theta,\beta)}\left(1-\theta^{g(\beta)}\right)\left(1-\theta^{g(\beta)+1}\right) \\
&\quad\times\left\{\left(1-\theta^{(k+2)(2g(\beta)+1)}\right)\theta^{j} - \left(1-\theta^{(k+2)(g(\beta)+1)}\right)\theta^{j(g(\beta)+1)+g(\beta)}\right. \\
&\quad\left. - \left(1-\theta^{(k+2)g(\beta)}\right)\theta^{-jg(\beta)+(k+1)(g(\beta)+1)}\right\}
\end{aligned}$$

and

$$\begin{aligned}
&d_k(\theta,\beta) \\
&:= \left(1-\theta^{g(\beta)}\right)\left(1-\theta^{(k+2)(g(\beta)+1)}\right) \\
&\quad\times\left\{\left(1-\theta^{g(\beta)+1}\right)\left(1-\theta^{k+1}\right) - (1-\theta)\left(1-\theta^{(k+1)(g(\beta)+1)}\right)\theta^{g(\beta)}\right\} \\
&\quad+\left(1-\theta^{g(\beta)+1}\right)\left(1-\theta^{(k+2)g(\beta)}\right)\theta^{k+g(\beta)+1} \\
&\quad\times\left\{\left(1-\theta^{g(\beta)}\right)\left(1-\theta^{k+1}\right)\theta^{(k+1)g(\beta)+1} - (1-\theta)\left(1-\theta^{(k+1)g(\beta)}\right)\right\}.
\end{aligned}$$

PROOF. Write

$$V_{\theta,\beta}(p_{k0}, \ldots, p_{kk}, \eta) := a\sum_{j=0}^{k}\theta^{-j}p_{kj}^2 - 2\sum_{j=1}^{k}\theta^{-j}p_{k,j-1}p_{kj} + \eta\Big(\sum_{j=0}^{k}p_{kj} - 1\Big),$$

where $a = a(\theta,\beta) = \theta^{\beta}(1+\theta^{-2\beta-1})$ and η is a Lagrange multiplier. Then, minimizing $\sigma^2_{\theta,\beta}(p_{k0}, \ldots, p_{kk})$ with respect to $p_{k0}, \ldots, p_{kk}$ is equivalent to minimizing

$V_{\theta,\beta}(p_{k0},\dots,p_{kk},\eta)$ with respect to $p_{k0},\dots,p_{kk}$ and η. For this, we need the condition that

$$\frac{\partial V_{\theta,\beta}}{\partial p_{k0}} = \cdots = \frac{\partial V_{\theta,\beta}}{\partial p_{kk}} = \frac{\partial V_{\theta,\beta}}{\partial \eta} = 0,$$

which leads to

$$\begin{gathered} 2ap_{k0} - 2\theta^{-1}p_{k1} + \eta = 0, \\ 2a\theta^{-j}p_{kj} - 2\left(\theta^{-j}p_{k,j-1} + \theta^{-j-1}p_{k,j+1}\right) + \eta = 0, \quad 1 \le j \le k-1, \\ 2a\theta^{-k}p_{kk} - 2\theta^{-k}p_{k,k-1} + \eta = 0, \\ \sum_{j=0}^{k} p_{kj} - 1 = 0, \end{gathered}$$

or equivalently,

$$2ap_{k0} - 2\theta^{-1}p_{k1} + \eta = 0,$$

$$(1+a\theta)\theta p_{k0} - (1+a)\theta p_{k1} + p_{k2} = 0, \tag{3.1}$$

$$\theta^2 p_{k,j-2} - (1+a\theta)\theta p_{k,j-1} + (1+a)\theta p_{kj} - p_{k,j+1} = 0, \quad 2 \le j \le k-1, \tag{3.2}$$

$$\theta p_{k,k-2} - (1+a\theta)p_{k,k-1} + (1+a)p_{kk} = 0, \tag{3.3}$$

$$\sum_{j=0}^{k} p_{kj} = 1. \tag{3.4}$$

To solve homogeneous difference equation (3.2), trying $p_{kj} = \xi^j$ ($\neq 0$) leads to $(\xi-\theta)(\xi^2 - a\theta\xi + \theta) = 0$, which has solutions $\xi_1 = \theta$, $\xi_2 = (a\theta - \sqrt{a^2\theta^2 - 4\theta})/2 = \theta^{g(\beta)+1}$ and $\xi_3 = (a\theta + \sqrt{a^2\theta^2 - 4\theta})/2 = \theta^{-g(\beta)}$.

(i) Suppose $\beta \neq -1, -1/2, 0$. Then the solutions ξ_1, ξ_2, ξ_3 are all different, and so we have

$$\begin{aligned} p_{kj} &= a_k\xi_1^j + b_k\xi_2^j + c_k\xi_3^j \\ &= a_k\theta^j + b_k\theta^{j(g(\beta)+1)} + c_k\theta^{-jg(\beta)}, \ j = 0,1,\dots,k, \end{aligned}$$

for some constants a_k, b_k and c_k. From boundary conditions (3.1), (3.3) and (3.4), we finally have

$$\begin{aligned} a_k &= \frac{(1-\theta)}{d_k(\theta,\beta)}\left(1-\theta^{g(\beta)}\right)\left(1-\theta^{g(\beta)+1}\right)\left(1-\theta^{(k+2)(2g(\beta)+1)}\right), \\ b_k &= -\frac{(1-\theta)}{d_k(\theta,\beta)}\left(1-\theta^{g(\beta)}\right)\left(1-\theta^{g(\beta)+1}\right)\left(1-\theta^{(k+2)(g(\beta)+1)}\right)\theta^{g(\beta)}, \\ c_k &= -\frac{(1-\theta)}{d_k(\theta,\beta)}\left(1-\theta^{g(\beta)}\right)\left(1-\theta^{g(\beta)+1}\right)\left(1-\theta^{(k+2)g(\beta)}\right) \\ &\quad \times\theta^{(k+1)(g(\beta)+1)}. \end{aligned} \tag{3.5}$$

Thus, for $j = 0, 1, \ldots, k$,

$$p_{kj} = a_k \theta^j + b_k \theta^{j(g(\beta)+1)} + c_k \theta^{-jg(\beta)} = p^*_{kj}(\theta, \beta).$$

(ii) Suppose $\beta = -1$ or 0. Then $g(\beta) = 0$, and so $\xi_1 = \xi_2 = \theta$ and $\xi_3 = 1$, which enables us to write

$$p_{kj} = (ja_k + b_k)\,\theta^j + c_k,\ j = 0, 1, \ldots, k,$$

for some constants a_k, b_k and c_k. Boundary conditions (3.1), (3.3) and (3.4) then yield

$$\begin{aligned}
a_k &= \frac{(1-\theta)^2}{e_k(\theta)}\left(1 - \theta^{k+2}\right),\\
b_k &= \frac{(1-\theta)^2}{e_k(\theta)}\left\{1 + (k+1)\,\theta^{k+2}\right\},\\
c_k &= -\frac{(k+2)}{e_k(\theta)}(1-\theta)^2\,\theta^{k+1},
\end{aligned}$$

where

$$e_k(\theta) = 1 - (k+2)^2\left(1+\theta^2\right)\theta^{k+1} + 2\,(k+1)\,(k+3)\,\theta^{k+2} + \theta^{2k+4}.$$

Thus, for $j = 0, 1, \ldots, k$,

$$\begin{aligned}
p_{kj} &= (ja_k + b_k)\theta^j + c_k\\
&= \frac{(1-\theta)^2}{e_k(\theta)}\left[\left\{1 + (k+1)\theta^{k+2} + j(1-\theta^{k+2})\right\}\theta^j - (k+2)\theta^{k+1}\right],
\end{aligned}$$

which is equal to $\lim_{\beta\to -1} p^*_{kj}(\theta,\beta) = \lim_{\beta\to 0} p^*_{kj}(\theta,\beta)$.

(iii) Suppose $\beta = -1/2$. Then $g(\beta) = -1/2$, and so $\xi_1 = \theta$ and $\xi_2 = \xi_3 = \theta^{1/2}$ from which we have

$$p_{kj} = a_k\theta^j + (jb_k + c_k)\theta^{j/2},\ j = 0, 1, \ldots, k,$$

for some constants a_k, b_k and c_k. Again, boundary conditions (3.1), (3.3) and (3.4) yield

$$\begin{aligned}
a_k &= -\frac{(k+2)}{f_k(\theta)}(1-\theta)\left(1-\theta^{1/2}\right)^2\theta^{1/2},\\
b_k &= -\frac{(1-\theta)}{f_k(\theta)}\left(1-\theta^{1/2}\right)^2\left(1-\theta^{k/2+1}\right),\\
c_k &= \frac{(1-\theta)}{f_k(\theta)}\left(1-\theta^{1/2}\right)^2\left(k+1+\theta^{k/2+1}\right),
\end{aligned}$$

where

$$f_k(\theta) = (k+1)\left(1-\theta^{k+3}\right) - 2(k+2)\left(1-\theta^{k+2}\right)\theta^{1/2} + (k+3)\left(1-\theta^{k+1}\right)\theta + 2(1-\theta)\theta^{k/2+1}.$$

Thus, for $j = 0, 1, \ldots, k$,

$$\begin{aligned} p_{kj} &= a_k\theta^j + (jb_k + c_k)\theta^{j/2} \\ &= \frac{(1-\theta)}{f_k(\theta)}\left(1-\theta^{1/2}\right)^2\theta^{j/2} \\ &\quad \times \left\{k+1+\theta^{k/2+1} - j\left(1-\theta^{k/2+1}\right) - (k+2)\theta^{(j+1)/2}\right\}, \end{aligned}$$

which is equal to $\lim_{\beta\to-1/2} p^*_{kj}(\theta,\beta)$. This completes the proof. □

For $k = 1$, the expression of $p^*_{1j}(1/2,\beta)$ was also derived by Falk (1994). Now it can be seen that for $\theta \in (0,1)$, $k = 1, 2, \ldots$ and $\beta \in \mathbb{R}$,

$$\begin{aligned} &\sigma^2_{\theta,\beta}\left(p^*_{k0}(\theta,\beta), \ldots, p^*_{kk}(\theta,\beta)\right) \\ &= \frac{\beta^2\theta^{-\beta-1}(1-\theta)}{(1-\theta^{-\beta})^2(\log\theta)^2}\left[\frac{a_k^2}{1-\theta}\left\{\theta^{g(\beta)}\left(1+\theta^{-2g(\beta)-1}\right)\left(1-\theta^{k+1}\right) - 2\left(1-\theta^k\right)\right\}\right. \\ &\quad + \frac{b_k^2}{1-\theta^{2g(\beta)+1}}\left\{\theta^{g(\beta)}\left(1+\theta^{-2g(\beta)-1}\right)\left(1-\theta^{(k+1)(2g(\beta)+1)}\right)\right. \\ &\qquad \left. - 2\theta^{g(\beta)}\left(1-\theta^{k(2g(\beta)+1)}\right)\right\} \\ &\quad + \frac{c_k^2}{1-\theta^{-2g(\beta)-1}}\left\{\theta^{g(\beta)}\left(1+\theta^{-2g(\beta)-1}\right)\left(1-\theta^{-(k+1)(2g(\beta)+1)}\right)\right. \\ &\qquad \left. - 2\theta^{-g(\beta)-1}\left(1-\theta^{-k(2g(\beta)+1)}\right)\right\} \\ &\quad + \frac{2a_kb_k}{1-\theta^{g(\beta)+1}}\left\{\theta^{g(\beta)}\left(1+\theta^{-2g(\beta)-1}\right)\left(1-\theta^{(k+1)(g(\beta)+1)}\right)\right. \\ &\qquad \left. - \left(1+\theta^{g(\beta)}\right)\left(1-\theta^{k(g(\beta)+1)}\right)\right\} + 2b_kc_k\theta^{g(\beta)}\left(1+\theta^{-2g(\beta)-1}\right) \\ &\quad + \frac{2a_kc_k}{1-\theta^{-g(\beta)}}\left\{\theta^{g(\beta)}\left(1+\theta^{-2g(\beta)-1}\right)\left(1-\theta^{-(k+1)g(\beta)}\right)\right. \\ &\qquad \left.\left. - \left(1+\theta^{-g(\beta)-1}\right)\left(1-\theta^{-kg(\beta)}\right)\right\}\right], \end{aligned}$$

where a_k, b_k and c_k are defined as in (3.5). Here, notice that

$$\lim_{k\to\infty} a_k = \frac{(1-\theta)\left(1-\theta^{g(\beta)+1}\right)}{\left(1-\theta^{g(\beta)}\right)},$$
$$\lim_{k\to\infty} b_k = -\frac{(1-\theta)\left(1-\theta^{g(\beta)+1}\right)\theta^{g(\beta)}}{\left(1-\theta^{g(\beta)}\right)},$$
$$\lim_{k\to\infty} c_k = 0,$$

and thus that

$$\lim_{k\to\infty} \sigma^2_{\theta,\beta}\left(p^*_{k0}(\theta,\beta),\ldots,p^*_{kk}(\theta,\beta)\right) = \frac{\beta^2(1-\theta)^2\left(1-\theta^{g(\beta)+1}\right)^2}{(1-\theta^{-\beta})^2\theta^{g(\beta)+\beta+2}(\log\theta)^2} =: \sigma^2_{\theta,\beta},\ \theta\in(0,1),\ \beta\in\mathbb{R}.$$

Moreover, for $j=0,1,\ldots,$ we have

$$\lim_{k\to\infty} p^*_{kj}(\theta,\beta) = \frac{(1-\theta)\left(1-\theta^{g(\beta)+1}\right)\left(1-\theta^{(j+1)g(\beta)}\right)\theta^j}{1-\theta^{g(\beta)}} =: p^*_j(\theta,\beta),\ \theta\in(0,1),\ \beta\in\mathbb{R}.$$

These expressions are generalizations of (1.2) and (1.3) derived by Drees (1995), since $p^*_j(1/2,\beta)=q^*_j(\beta)$, $j=0,1,\ldots,$ for $\beta\neq -1/2$ and $\sigma^2_{1/2,\beta}=(\sigma^{(D)}_\beta)^2$ for all $\beta\in\mathbb{R}$.

In other words, for $\beta\neq-1/2$ the estimator in (1.5) with $\theta=1/2$ and weights $p_{kj}=p^*_{kj}(1/2,\beta)$ has the same asymptotic performance as the refined Pickands estimator in (1.1) with weights $q_j=q^*_j(\beta)$ as $k\to\infty$, but the former has a much better performance for each finite k which is typical for finite samples. Also, note that unlike the Drees' weights $q^*_j(\beta)$, the optimal weights $p^*_{kj}(\theta,\beta)$ and their limits $p^*_j(\theta,\beta)$ are well defined for all $\beta\in\mathbb{R}$. In fact, if we let $k=k(\theta,m)=\lfloor -\log m/\log\theta-2\rfloor$ where $m=m(n)$ is an intermediate sequence, then by using a method similar to one used in Drees (1995) it can be shown that for $\theta\in(0,1)$ and $\beta\in\mathbb{R}$, the asymptotic distribution (as $n\to\infty$) of $\sqrt{m}\{\hat\beta_{n,m,\theta}(p^*_0(\theta,\beta),\ldots,p^*_{k-1}(\theta,\beta),1-\sum_{j=0}^{k-1}p^*_j(\theta,\beta))-\beta\}$ is normal with mean zero and variance $\sigma^2_{\theta,\beta}$ under the condition of Theorem 2.1 (a) or the conditions of Theorem 2.1 (b) with $m=m(n)$ satisfying $\sqrt{m}A(n/m)\to 0$ as $n\to\infty$. Since $\lim_{n\to\infty}p^*_{k(\theta,m),j}(\theta,\beta)=p^*_j(\theta,\beta)$, a slight modification of the proof thus leads to the following result.

THEOREM 3.1. *Suppose $F \in \mathcal{D}_{dif}(G_\beta)$ for some $\beta \in \mathbb{R}$, so that (2.1) holds. Let $\theta \in (0,1)$.*

(a) *If $R(t,x) \equiv 0$, then*

$$\sqrt{m}\left(\hat{\beta}_{n,m,\theta}(p^*_{k(\theta,m),0}(\theta,\beta),\dots,p^*_{k(\theta,m),k(\theta,m)}(\theta,\beta)) - \beta\right) \xrightarrow{d} N(0,\sigma^2_{\theta,\beta}) \tag{3.6}$$

as $n \to \infty$ for any intermediate sequence $m = m(n)$.

(b) *For $R(t,x)$ not being identically zero, assume further that (2.2) holds for some $\rho \le 0$. Then (3.6) holds for any intermediate sequence $m = m(n)$ such that $\sqrt{m}A(n/m) \to 0$ as $n \to \infty$.*

Since the optimal weights $p^*_{k(\theta,m),j}(\theta,\beta)$ depend upon the unknown parameter β, it is reasonable to utilize the data-driven estimators $p^*_{k(\theta,m),j}(\theta,\tilde{\beta}_n)$, where $\tilde{\beta}_n$ is any initial estimator of β which is weakly consistent. Then, since

$$\lim_{n\to\infty} p^*_{k(\theta,m),j}(\theta,\beta) = p^*_j(\theta,\beta)$$

and each $p^*_j(\theta,\cdot)$ is a continuous function, it can be seen that

$$\hat{\beta}_{n,m,\theta}(p^*_{k(\theta,m),0}(\theta,\tilde{\beta}_n),\dots,p^*_{k(\theta,m),k(\theta,m)}(\theta,\tilde{\beta}_n)) \tag{3.7}$$

is asymptotically equivalent to $\hat{\beta}_{n,m,\theta}(p^*_{k(\theta,m),0}(\theta,\beta),\dots,p^*_{k(\theta,m),k(\theta,m)}(\theta,\beta))$ and thus that Theorem 3.1 still holds with $p^*_{k(\theta,m),j}(\theta,\beta)$ replaced by $p^*_{k(\theta,m),j}(\theta,\tilde{\beta}_n)$ for any weakly consistent estimator $\tilde{\beta}_n$ of β.

Finally, we consider a practical problem of choosing θ in an optimal way. One way is to choose θ which minimizes the asymptotic variance $\sigma^2_{\theta,\beta}$ in (3.6). However, this is not feasible since $\sigma^2_{\theta,\beta}$ is strictly decreasing, as a function of $\theta \in (0,1)$, for each $\beta \in \mathbb{R}$. This fact is easily understood by noticing that the larger the value of θ is, the more number of convexing the estimator in (3.7) has. In practice, one may choose the largest value of θ which well defines the estimator in (3.7). For a comparison with the refined Pickands estimator in (1.1) with weights $q_j = q^*_j(\beta)$, let us consider the ratio $\sigma^2_{\theta,\beta}/(\sigma^{(D)}_\beta)^2$, which is, for $\beta \ne -1/2$, the asymptotic relative efficiency of the refined Pickands estimator with respect to the estimator in (3.7). If $\theta > 1/2$, then $\sigma^2_{\theta,\beta}/(\sigma^{(D)}_\beta)^2 < 1$ for all $\beta \in \mathbb{R}$. Also, note that

$$\lim_{\theta\uparrow 1} \frac{\sigma^2_{\theta,\beta}}{\left(\sigma^{(D)}_\beta\right)^2} = \frac{\left(2^\beta - 1\right)^2 (g(\beta)+1)^2\, 2^{g(\beta)-\beta+2} (\log 2)^2}{\beta^2 \left(2^{g(\beta)+1} - 1\right)^2}$$

$$= \begin{cases} 2(\log 2)^2, & \text{if } \beta \leq -\frac{1}{2}, \\ \dfrac{4\left(2^{\beta}-1\right)^2(\beta+1)^2(\log 2)^2}{\beta^2\left(2^{\beta+1}-1\right)^2}, & \text{if } \beta > -\frac{1}{2}. \end{cases}$$

REFERENCES

BLOOMFIELD, P. J., ROYLE, J. A., STEINBER, L. J. AND YANG, Q. (1996). "Accounting for meteorological effects in measuring urban ozone levels and trends", *Atmospheric Environment*, **30**, 3067–3077.

DE HAAN, L. (1984). "Slow variation and characterization of domains of attraction", In *Statistical Extremes and Applications* (J. Tiago de Oliveira, ed.), 31–48, Reidel, Dordrecht.

DE HAAN, L. AND STADTMÜLLER, U. (1996). "Generalized regular variation of second order", *Journal of the Australian Mathematical Society*, **A61**, 381–395.

DEKKERS, A. L. M. AND DE HAAN, L. (1989). "On the estimation of the extreme-value index and large quantile estimation", *The Annals of Statistics*, **17**, 1795–1832.

DEKKERS, A. L. M., EINMAHL, J. H. J. AND DE HAAN, L. (1989). "A moment estimator for the index of an extreme-value distribution", *The Annals of Statistics*, **17**, 1833–1855.

DREES, H. (1995). "Refined Pickands estimators of the extreme value index", *The Annals of Statistics*, **23**, 2059–2080.

FALK, M. (1994). "Efficiency of convex combinations of Pickands estimator of the extreme value index", *Journal of Nonparametric Statistics*, **4**, 133–147.

HILL, B. M. (1975). "A simple general approach to inference about the tail of a distribution", *The Annals of Statistics*, **3**, 1163–1174.

PEREIRA, T. T. (1994). "Second order behavior of domains of attraction and the bias of generalized Pickands' estimator", In *Extreme Value Theory and Applications III* (J. Galambos, J. Lechner and E. Simiu, eds.), 165–177, NIST.

PICKANDS, J. (1975). "Statistical inference using extreme order statistics", *The Annals of Statistics*, **3**, 119–131.

SMITH, R. L. (1987). "Estimating tails of probability distributions", *The Annals of Statistics*, **15**, 1174–1207.

YUN, S. (2002a). "On a generalized Pickands estimator of the extreme value index", *Journal of Statistical Planning and Inference*, **102**, 389–409.

YUN, S. (2002b). "Minimax choice and convex combinations of generalized Pickands estimator of the extreme value index", *Journal of the Korean Statistical Society*, **31**, 315–328.