MONASH University

MONASH BUSINESS SCHOOL

ISSN 1440-771X

Department of Econometrics and Business Statistics

http://business.monash.edu/econometrics-and-business-statistics/research/publications

Bootstrapping tail statistics: tail quantile process, Hill estimator, and confidence intervals for high-quantiles of heavy tailed distributions

Svetlana Litvinova and Mervyn J Silvapulle

August 2018

Working Paper 12/18

Bootstrapping tail statistics: tail quantile process, Hill estimator, and confidence intervals for high-quantiles of heavy tailed distributions

Svetlana Litvinova[1] and Mervyn J Silvapulle[2]

Department of Econometrics and Business Statistics,

Monash Business School, Monash University, Australia.

svetlana.litvinova@monash.edu; mervyn.silvapulle@monash.edu

August 2018

Abstract

In risk management areas such as reinsurance, the need often arises to construct a confidence interval for a quantile in the tail of the distribution. While different methods are available for this purpose, doubts have been raised about the validity of full-sample bootstrap. In this paper, we first obtain some general results on the validity of full-sample bootstrap for the tail quantile process. This opens the possibility of developing bootstrap methods based on tail statistics. Second, we develop a bootstrap method for constructing confidence intervals for high-quantiles of heavy-tailed distributions and show that it is consistent. In our simulation study, the bootstrap method for constructing confidence intervals for high quantiles performed overall better than the data tilting method, but none was uniformly the best; the data tilting method appears to be currently the preferred choice. Since the two methods are based on quite different approaches, we recommend that both methods be used side by side in applications.

Keywords: Full-sample bootstrap; Intermediate order statistic; Hill estimator; Extreme value Index; Tail empirical process; Tail quantile process

1 Introduction

Identification of rare and costly extreme events is important in risk management. It helps the decision making process in taking precautionary actions to reduce the potential losses caused by such extreme events. Some examples are: (1) How high is the flood level expected to reach only once in 50 years? (2) What is the level of loss that an investment portfolio is expected to exceed with only 0.005 probability? (3) What is the level of the insurance

[1]Svetlana Litvinova gratefully acknowledges the support from Monash University through a Postgraduate Scholarship.

[2]Mervyn Silvapulle gratefully acknowledges support from the Australian Research Council

claim that is expected to exceed by only 0.1% of the claims? Such inferential topics are at the core of re-insurance and underwriting. A quantity that is regularly used in such areas is a confidence interval for a high-quantile x_p defined by $F(x_p) = 1 - p$, where F is a distribution function and p is a small positive number $(0 < p < 1)$. The best available methods use estimates of the tail of F to construct estimators of x_p; the value of p is too small to rely on the sample quantile as an estimate of x_p. This paper develops a new and improved method for constructing a confidence interval for x_p. Our proposed method is based on full-sample bootstrap. Our main theoretical results clarify doubts that have been raised about the validity of such a full-bootstrap method.

There is a large literature on estimation of confidence intervals for high-quantiles, particularly for heavy-tailed distributions. It is a topic that is carefully studied in textbooks and monographs in this field. For an excellent account of the topic, see Chapter 4 in Beirlant *et al.* (2004), section 4.3 in de Haan and Ferreira (2006), and sections 6.3 and 6.4 in Embrechts *et al.* (1997). Coles (2001) provides an excellent introductory account to the topic. Peng and Qi (2006) studied coverage of confidence intervals for high quantiles of a heavy-tailed distribution, and developed an improved method based on data tilting; in their simulation study, the data-tilting method performed better than the competing ones.

It is well-known that bootstrap has its limitations for inference involving extremes and heavy-tailed distributions. For heavy-tailed distributions, the bootstrap fails to consistently estimate the distribution of a sample mean (Athreya 1987; Knight 1989). Similarly, the bootstrap fails to consistently estimate the distribution of sample minimum or maximum (Angus 1993). A subsample bootstrap, where resamples have a size of order of magnitude smaller than that of the original sample, is asymptotically valid in these cases (for example, Athreya 1987; Deheuvels *et al.* 1993). However, in simulation studies, the subsample bootstrap performed poorly for inference on mean of a heavy-tailed distribution (Hall 1998), and for inference on the Pareto index (Guillou 2000); our simulations results, not reported here, also corroborate the same. Geluk and de Haan (2002) examined validity of the bootstrap method for intermediate order statistics. The question of whether or not full-sample nonparametric bootstrap is consistent for estimating the distribution of extreme statistics, and for inference on high quantiles in heavy-tailed univariate distributions, is nontrivial but deserves a clear answer in view of its potential applications in empirical studies and for the development of bootstrap methods for inference on the tails of distributions.

In this paper, we obtain bootstrap results that correspond to fundamental results on the tail quantile process. First, we show that a bootstrap version of a fundamental asymptotic representation of the tail quantile process holds. This result is useful for developing bootstrap methods for inference on the tail of the distribution. Second, we show that the full sample bootstrap is consistent for estimating the distribution of the Hill estimator of extreme value index, and finally develop a bootstrap method for constructing confidence intervals for high-quantiles of heavy-tailed univariate distributions. These theoretical results provide answers to some of the aforementioned questions that have been raised in the literature about the validity of full-sample nonparametric bootstrap. We carried out a simulation study to compare our bootstrap method with the data-tilting method, which appears to be currently the best available. In our simulation study, the overall performance of our proposed bootstrap method was better than that of the data-tilting method, but the difference was not huge, and none of them performed uniformly the best. Therefore, our recommendation is that both methods be used in empirical studies.

2 Notation and preliminaries

In this section we introduce some notation, state the main regularity conditions, and recall some relevant results from the extreme value theory literature; there are no new results in this section and most of them are from de Haan and Ferreira (2006) which is the main reference for the details in this paper. Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed with common distribution function F, and let $X_{1,,n} \leq X_{2,n} \leq \ldots \leq X_{n,n}$ denote the corresponding order statistics. Let F_n denote the empirical distribution function of $X_1, X_2, \ldots, X_n$. If there exists a sequence of constants $a_n > 0$ and $b_n \in \mathbb{R}$ such that $pr\{(X_{n,n} - b_n)/a_n \leq x\} \to G(x)$ and G is nondegenerate then (a) G is called an *extreme* *value distribution* function, (b) F is said to be in the *maximum domain of attraction* of G, denoted by $F \in D(G)$, and (c) the class of extreme value distributions is $G_\gamma(ax+b)$, where $G_\gamma(x) = \exp\{-(1+\gamma x)^{-1/\gamma}\}$ $(x > -\gamma^{-1}, b \in \mathbb{R}, a > 0)$. The parameter γ, called the *extreme value index* or the *tail index*, characterizes the tail behaviour of F; in what follows, the term 'tail' refers to 'right tail'. If $\gamma < 0$, $\gamma = 0$, or $\gamma > 0$ then F has a finite end point, *light tail*, or *heavy tail*, respectively.

In this paper, we pay particular attention to the case when F belongs to the family of heavy-tailed distributions; thus, $F \in D(G_\gamma)$ where $\gamma > 0$. Therefore, the tail of $\{1 - F(x)\}$ decays as a power function; more specifically, $F(x) = 1 - s(x)x^{-1/\gamma}$ where $s(x)$ is slowly varying at ∞. Examples of distributions in this family include Pareto, Burr, Student's t, α-stable, and loggamma distributions (Matthys and Beirlant 2003). Thus, our assumption that F has heavy tail does not restrict F to a particular parametric family, such as, a t distribution.

Let $U(y) = F^{\leftarrow}(1 - y^{-1})$, where $F^{\leftarrow}$ denotes the inverse function of F. There is a one-to-one correspondence between F and U. Often, regularity conditions on the tails of F are expressed in terms of U. The distribution function $F \in D(G)$ if and only if there exists a positive function $a(t)$ $(t \in \mathbb{R})$ such that (for example, Theorem 1.1.6, de Haan and Ferreira 2006).

$$\lim_{t\to\infty} \{U(tx) - U(t)\}/a(t) = \{x^\gamma - 1\}/\gamma, \quad (x > 0). \tag{1}$$

If (1) holds, then U is said to satisfy the *first-order condition* of regular variation.

To establish convergence in distribution of intermediate order statistics, a second-order refinement is also necessary. The function U is said to satisfy the *second-order condition* of regular variation if there exists a function H that is not a multiple of $(x^\gamma - 1)/\gamma$ and a positive or negative function A with $\lim_{t\to\infty} A(t) = 0$, such that

$$\lim_{t\to\infty} [\{U(tx) - U(t)\}/a(t) - \{x^\gamma - 1\}/\gamma]/A(t) = H(x), \quad (x > 0). \tag{2}$$

For heavy-tailed distributions, namely for the case $\gamma > 0$, this condition reduces to

$$[\{U(tx)/U(t)\} - x^\gamma]/A(t) \to x^\gamma(x^\rho - 1)/\rho \text{ as } t \to \infty \tag{3}$$

for some $\rho \le 0$ $(x > 0, \gamma > 0)$.

Let $k := k(n)$ denote a sequence of positive integers satisfying $k \to \infty$ and $(k/n) \to 0$ as $n \to \infty$; in what follows this is assumed without further comment. Then $X_{n-k,n}$ is called an *intermediate order statistic*, and the sequence $\{k(n)\}$ is called an *intermediate order sequence*. Next, let us define,

Condition A:

(A.1). $k \to \infty$ and $(k/n) \to 0$ as $n \to \infty$. (A.2). The function U satisfies (1) and (2).
(A.3). $k^{1/2}A(n/k) \to \lambda$ as $n \to \infty$, for some finite λ. (A.4). $k^{1/2}A(n/k) \to 0$ as $n \to \infty$.

Now, we recall important results on statistics of extremes. In the next section, we provide the corresponding bootstrap versions of these results. For a proof of the next result, see Drees (1998) and Theorem 2.4.2 and Corollary 2.4.5 in de Haan and Ferreira (2006).

Theorem 1. *Suppose that (A.1), (A.2), and (A.3) are satisfied. Then there exists a sequence of Brownian motions $\{W_n(s) : s > 0\}$ such that for suitably chosen functions a_0 and A_0 and each $\varepsilon > 0$, we have*

$$\sup_{0<s\le 1} s^{\gamma+(1/2)+\varepsilon}\Big|k^{1/2}\left(\frac{X_{n-[ks],n}-B_0(n/k)}{a_0(n/k)} - \frac{s^{-\gamma}-1}{\gamma}\right) - s^{-\gamma-1}W_n(s) \\ -k^{1/2}A_0(n/k)\Psi_{\gamma,\rho}(s^{-1})\Big| = o_p(1).$$

where

$$B_0(n/k) = \begin{cases} U(n/k), & \gamma \ge -1/2, \\ X_{n,n} + \gamma^{-1}a_0(n/k), & \gamma < -1/2 \end{cases}$$

For later reference, we state the following special case; see for example, Theorem 2.4.1 in de Haan and Ferreira (2006). To help interpretation, it is worth noting that $X_{n-k,n} = F_n^{\leftarrow}\{1-(k/n)\}$ and $U(n/k) = F^{\leftarrow}\{1-(k/n)\}$.

Theorem 2. *Suppose that the conditions of Theorem 1 are satisfied. Then $k^{1/2}\{X_{n-k,n} - U(n/k)\}/a(n/k)$ converges, in distribution, to $N(0,1)$.*

To define a high quantile, let p_n be a sequence in the range $0 < p_n < 1$ satisfying $p_n \to 0$ as $n \to \infty$. A $100(1-p_n)\%$ quantile denoted x_{p_n} of F is defined by $F(x_{p_n}) = 1 - p_n$.

Since $p_n \to 0$ as $n \to \infty$ we refer to x_{p_n} as a *high quantile*. The $100(1-p_n)$%-quantile of $\{X_1, \ldots, X_n\}$ is equal to the intermediate order statistic $X_{n-[np_n],n}$ where $[\cdot]$ denotes the integer part. If p_n is close to, or smaller than, n^{-1}, as is the case with high-quantiles, then the foregoing sample quantile is not a good estimator of the population quantile x_{p_n}. In this case, x_{p_n} is typically estimated using information about the behaviour of the tail of F.

Consider the case when F has heavy-tails (i.e $\gamma > 0$), and let

$$\hat{\gamma} = k^{-1} \sum_{i=0}^{k-1} [\log X_{n-i,n} - \log X_{n-k,n}],$$

which is the well-known *Hill estimator* of γ. We estimate x_{p_n} by $\hat{x}_{p_n} = X_{n-k,n}\{k/(np_n)^{\hat{\gamma}}\}$ (for example, see page 1966 in Peng and Qi 2006); the basic idea is to temporarily assume that $F(x) = 1 - cx^{-1/\gamma}$ and construct $\hat{x}_{p_n}$ as $(p_n/\hat{c}_n)^{-\hat{\gamma}}$. Next, we recall two more results in addition to Theorem 2 (see Theorem 3.2.5 in de Haan and Ferreira 2006, and Theorem 1 in Peng and Qi 2006). Let $d_n = k/(np_n)$.

Theorem 3. *Suppose that (A.1), (A.2), and (A.4) are satisfied, and that F has heavy tail. Then $k^{1/2}(\hat{\gamma} - \gamma)$ converges, in distribution, to $N(0, \gamma^2)$ as $n \to \infty$.*

Theorem 4. *Suppose that (A.1), (A.2), and (A.4) are satisfied. Further, suppose that F has heavy tail, $np_n = o(k)$, and $\log d_n = o(k^{1/2})$. Then, $\{k^{1/2}/\log d_n\} \log\{\hat{x}_{p_n}/x_{p_n}\}$ converges, in distribution, to $N(0, \gamma^2)$ as $n \to \infty$.*

Theorem 4 can be used for constructing an asymptotic confidence interval for x_{p_n}. To implement the aforementioned methods, the threshold k must be chosen. There has been extensive studies about various choices of k; for example, see Scarrott and MacDonald (2012), section 2.2.1 in Peng and Qi (2017), Ferreira *et al.* (2003), Gomes and Oliveira (2001), and Danielsson *et al.* (2001). In general, the available automatic methods of choosing k are not that reliable (Danielsson *et al.* 2016).

3 First order consistency of the full-sample bootstrap

Let $X_1^*, \ldots, X_n^*$ denote a simple random sample from the empirical distribution function F_n of $\{X_1, \ldots, X_n\}$, and let $X_{1,n}^*, \ldots, X_{1,n}^*$ denote the corresponding order statistics. Let

$\{\hat{\gamma}^*, \hat{x}^*_{p_n}\}$ denote the bootstrap statistics corresponding to $\{\hat{\gamma}, \hat{x}_{p_n}\}$. Let P denote the probability relating to the original sample, and let P' denote the joint probability measure for the original sample and the bootstrap sample; the precise construction of P' is discussed in the Supplementary Materials. The following results establish the first-order asymptotic validity of the full-sample bootstrap estimators of the distributions of $\{X_{n-k,n}, \hat{\gamma}, \hat{x}_{p_n}\}$ based on $\{X^*_{n-k,n}, \hat{\gamma}^*, \hat{x}^*_{p_n}\}$. We use these to show that a valid asymptotic confidence interval for the high-quantile x_{p_n} can be constructed by full-sample bootstrap. The process $\{X_{n-[ks],n}\}_{s\in[0,1]}$ can also be suitably approximated by its corresponding bootstrap process, as stated in the result below. The consistency of bootstrap is with respect to the joint probability P'; it is not the conditional probability given the sample.

The next result is a bootstrap version of Theorem 1

Theorem 5. *Suppose that the conditions of Theorem 1 are satisfied. Then, for suitably chosen function $a_0(\cdot)$, we have*

$$\sup_{0\le s\le 1} s^{\gamma+(1/2)+\epsilon}\left|k^{1/2}\frac{X^*_{n-[ks],n}-X_{n-[ks],n}-b_0(n)}{a_0(\frac{n}{k})}-\frac{B'_{k,n}(s)}{s^{\gamma+1}}\right| = o_{P'}(1) \tag{4}$$

where

$$b_0(n) = \begin{cases} 0, & \gamma \ge -1/2, \\ X^*_{n,n} - X_{n,n}, & \gamma < -1/2. \end{cases}$$

A corollary of the foregoing result is for a particular intermediate sequence. This is stated below; it can also proved directly without appealing to the above, and proof is somewhat simpler.

Theorem 6. *Suppose that the conditions of Theorem 2 are satisfied. Then,*

$$P'[Y^*_n \le y] - P[Y_n \le y] \to 0 \ \text{as } n\to\infty \quad (y\in\mathbb{R}),$$

where $Y^*_n = k^{1/2}(X^*_{n-k,n} - X_{n-k,n})/a(n/k)$ and $Y_n = k^{1/2}\{X_{n-k,n} - U(n/k)\}/a(n/k)$.

Theorem 7. *Suppose that the conditions of Theorem 3 are satisfied. Then,*

$$P'[k^{1/2}(\hat{\gamma}^* - \hat{\gamma}) \le y] - P[k^{1/2}(\hat{\gamma} - \gamma) \le y] \to 0 \ \text{as } n\to\infty \quad (y\in\mathbb{R}).$$

Theorem 8. *Suppose that the conditions of Theorem 4 are satisfied. Then,*

$$P'[Y_n^* \le y] \; - \; P[Y_n \le y] \to 0 \text{ as } n \to \infty \quad (y \in \mathbb{R}),$$

where $Y_n^* = (k^{1/2}/\log d_n)\log(\hat{x}_{p_n}^*/\hat{x}_{p_n})$ *and* $Y_n = (k^{1/2}/\log d_n)\log(\hat{x}_{p_n}/x_{p_n})$.

Theorem 8 is used for constructing a full-sample bootstrap confidence interval for x_{p_n}. In the simulations study that we report in the next section, we used the percentile method because it performed better than the other methods. Let $\hat{x}_{p_n,\alpha} := \hat{x}_{p_n,\alpha}(X_1, \ldots, X_n)$ denote the α-quantile of $\hat{x}_{p_n}^*$, conditional on $\{X_1, \ldots, X_n\}$, defined by $P^*(\hat{x}_{p_n}^* \le \hat{x}_{p_n,\alpha} \mid X_1, \ldots, X_n) = 1 - \alpha$, where we used the symbol P^* to denote the conditional bootstrap probability. Since $\hat{x}_{p_n,\alpha}$ is a conditional quantile, it can be computed approximately by simulation; for example, generate B bootstrap samples, compute the quantile estimate $\hat{x}_{p_n}^{*(j)}$ for the jth bootstrap sample ($j = 1, \ldots, B$), and approximate $\hat{x}_{p_n,\alpha}$ by the $1 - \alpha$ quantile of $\{\hat{x}_{p_n}^{*(j)}, j = 1, \ldots, B\}$. Then the *percentile* $(1 - \alpha)$-confidence interval for x_{p_n} is $I_{Ep}(\alpha) = (\hat{x}_{p_n,\alpha/2}, \hat{x}_{p_n,1-\alpha/2})$. This is stated in the next result.

Proposition 1. *Suppose that the conditions of Theorem 4 are satisfied. Then,* $P\{x_{p_n} \in I_{Ep}(\alpha)\} \to (1 - \alpha)$ *as* $n \to \infty$ *(*$0 < \alpha < 1$*).*

4 Simulation Study

We carried out a simulation study to compare the performance of our bootstrap method with the data tilting method for constructing a confidence interval for a high-quantile x_{p_n}. We restricted to the data tilting method for comparison mainly because it appears to be the best available based on the simulation study reported in Peng and Qi (2017). We compared the two methods in terms of coverage probability and length of the confidence interval. The results for the following settings are provided: $F = \text{Frechet}(2)$, $t(2)$ and $t(4)$; $np_n = 2, 0.2, 0.02$ when $n = 200$; $np_n = 1, 0.1, 0.01$ when $n = 1000$.

Estimation of tail quantiles requires to choose k. This is a nontrivial problem, as mentioned in section 2. In our study, we evaluated a method of Danielsson *et al.* (2016), which uses a minimum distance criterion between the Pareto and empirical tail quantiles. This

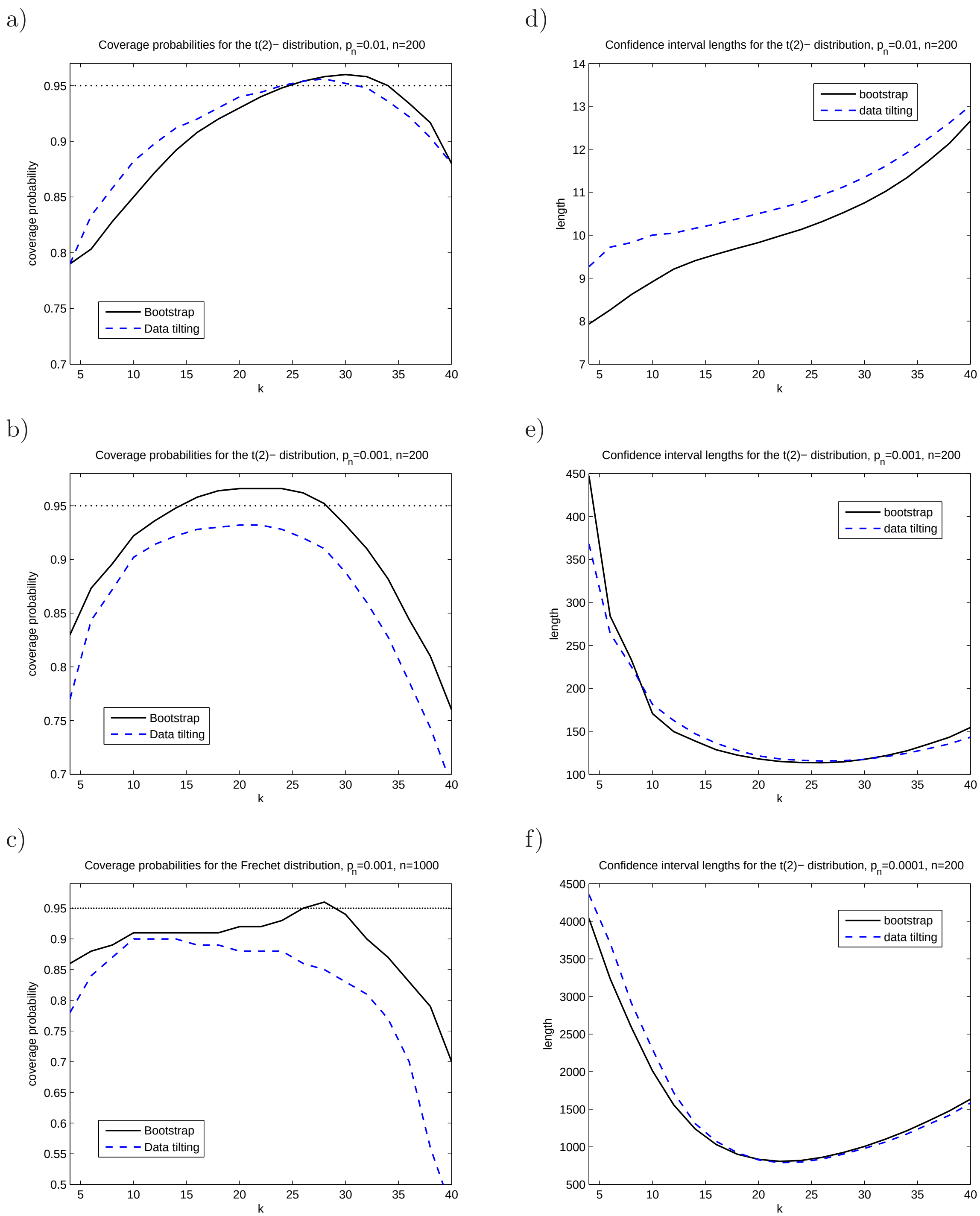

Figure 1: Comparison of bootstrap and data-tilting methods for 95% confidence intervals for high quantiles. The population distribution is t_2 and the sample size is $n = 200$. The continuous and dashed lines correspond to the bootstrap and data-tilting methods respectively. The three panels in the first column provide a plot of coverage rates against k. The three rows of diagrams correspond to $p_n = 0.01, 0.001$, and 0.0001 respectively.

method performed well for the Frechet distribution but was unreliable for other distributions, especially for estimating moderately high quantiles. Therefore, we do not report these results here, instead we provide results for a range of fixed values of k: for $n = 1000$ we considered $k = 5, 10, \dots, 200$, and for $n = 200$ we considered $k = 2, 4, \dots, 70$.

For each distribution F, the bootstrap method was implemented as follows: (1) Draw a simple random sample from F. (2) Draw a simple random sample $X_1^*, \dots, X_n^*$ from the sample in the previous step. (3) Choose a value for k and compute the estimate $\hat{x}_{p_n}^*$ using the bootstrap sample. (4) Repeat the previous two steps $B = 1000$ times. (5) Compute the quantiles $\hat{x}_{p_n,\alpha/2}$ and $\hat{x}_{p_n,1-\alpha/2}$ for the B values of $\hat{x}_{p_n}^*$ in the previous step, construct a $100(1-\alpha)\%$-confidence interval for x_{p_n} as $(\hat{x}_{p_n,\alpha/2}, \hat{x}_{p_n,1-\alpha/2})$, and check whether or not the true quantile is in this confidence interval. (6). Construct a confidence interval by the data-tilting method and check whether or not the true quantile is in this confidence interval. (7). Repeat the previous steps 1000 times and compute the coverage rates and mean lengths of the confidence intervals.

The results for t-distribution with 2 degrees of freedom are presented here in Figure 1; the results for t-distribution with 4 degrees of freedom and the Frechet(2) distribution are presented in the Supplementary Materials. The relative performance of the two methods for the other distributions in the Supplementary Materials are similar.

The main observations may be summarised as follows:

(1). Large values of np_n ($(n, p_n) = (200, 0.01), (1000, 0.001)$): For larger values of k, the bootstrap method performs better in terms of coverage and length. For smaller values of k, the data tilting method has slightly better coverage at the cost of larger confidence intervals.

(2). Small values of np_n ($(n, p_n) = (200, 0.001), (200, 0.0001), (1000, 0.0001), (1000, 0.00001)$: The bootstrap method performs better in terms of coverage and length.

In summary, the bootstrap method performs at least as well as the data tilting method, or better.

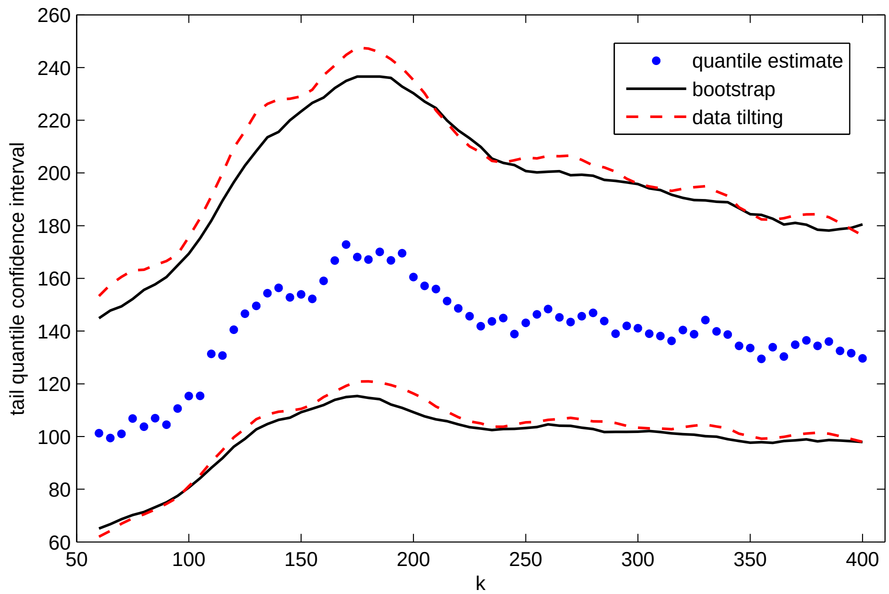

Figure 2: The 90% confidence interval for the 0.01%-upper quantile of fire-loss, based on the Danish fire-loss data which consists of $n = 2156$ observations. The confidence intervals are for each fixed k, $k = 60, 65, \ldots, 400$. The continuous and dashed lines correspond to bootstrap and data-tilting methods.

5 Application to Danish Fire Insurance Data

As an application, we analyze a large set of data on fire insurance claims in Denmark. The data set contains the loss X in Danish Krone, caused by each fire from January 1980 to December 1990. The losses could be to buildings, contents, or profits. In our analysis we consider total losses over one million Danish Krone. There are $n = 2156$ observations in the data set.

These data were analysed in several studies to illustrate different methods of inference based on statistics for extremes. In a detailed study, McNeil (1997) followed by Resnick (1997), concluded that the distribution of losses to be heavy-tailed with the tail index γ being in the range 0.5 to 0.7. We also assume that the loss is heavy-tailed.

Suppose that we wish to construct a 90% confidence interval for the quantile corresponding to $p_n = 0.001$. Peng and Qi (2006) also used the same set of data, and applied data tilting, likelihood ratio, and normal approximation methods for constructing confidence intervals for tail quantile for a range of k; in their computations, the data-tilting confidence intervals turned out to be the shortest for most values of k. Figure 2 provides the tail quantile estimates together with the 90% confidence intervals obtained by the bootstrap and data tilting methods for k, $k = 60, 65, \ldots, 400$. It appears that lengths of confidence intervals obtained by the bootstrap method are slightly smaller than those based on the data tilting method for $k \leq 185$; this is consistent with the simulation results. We observe that the two methods provide similar results for these data. It is of interest to note that the data tilting method and the bootstrap method lead similar results and conclusions even though the two methods adopt quite different techniques for inference. The consistency between the two sets of results provide some assurance of the reliability of the two sets of results. In summary, there is much to be gained in empirical studies by comparing and contrasting the results based on the data tilting and the bootstrap methods.

6 Supplementary Materials

The Supplementary Materials contains the proofs and additional simulation results.

Acknowledgements: The first author is a PhD student and this paper is based on her thesis that is in preparation. We wish to thank professor John Einmahl for some helpful discussions at the early stages of this research. A few days before this paper was submitted to the Working Paper Series, Professor Chen Zhou of Erasmus University presented a seminar at Monash University; there is some overlap between the results of our paper and those presented at this seminar, but the two approaches appear to be different. However, to our knowledge, there is no paper available in the literature, published or unpublished, containing any of the results/proofs in this paper.

References

Angus, J. (1993). Asymptotic theory for bootsrapping the extremes. *Communications in Statistics - Theory and Methods*, **22**(1), 15–30.

Athreya, K. B. (1987). Bootstrap of the mean in the infinite variance case. *The Annals of Statistics*, **15**, 724–731.

Beirlant, J., Goegebeur, Y., Segers, J., and Teugels, J. (2004). *Statistics of Extremes: Theory and Applications*. John Wiley & Sons (Chichester).

Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag Inc.

Csörgő, S. and Mason, D. M. (1989). Bootstrapping empirical functions. *The Annals of statistics*, **17**(4), 1447–1471.

Danielsson, J., de Haan, L., Peng, L., and de Vries, C. G. (2001). Using a bootstrap method to choose the sample fraction in tail index estimation. *Journal of Multivariate Analysis*, **76**(2), 226–248.

Danielsson, J., Ergun, L. M., de Haan, L., and de Vries, C. (2016). Tail index estimation: Quantile driven threshold selection. *SRC Working paper*.

de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction (Springer Series in Operations Research and Financial Engineering)*. Springer.

Deheuvels, P., Mason, D. M., and Shorack, G. R. (1993). Some results on the influence of extremes on the bootstrap. *Annales de l'institut Henri Poincare (B) Probabilites et Statistiques*, **29**(1), 83–103.

Drees, H. (1998). On smooth statistical tail functionals. *Scandinavian Journal of Statistics*, **25**(1), 187–210.

Embrechts, P., Kluppelberg, C., and Mikosch, T. (1997). *Modelling Extremal Events: For Insurance and Finance*. Springer-Verlag Inc.

Ferreira, A., de Haan, L., and Peng, L. (2003). On optimising the estimation of high quantiles of a probability distribution. *Statistics*, **37**(5), 401–434.

Geluk, J. and de Haan, L. (2002). On bootstrap sample size in extreme value theory. *Publ. Inst. Math. (N.S.)*, **71**(85), 21–25.

Gomes, M. I. and Oliveira, O. (2001). The bootstrap methodology in statistics of extremes - choice of the optimal sample fraction. *Extremes*, **4**(4), 331–358.

Guillou, A. (2000). Bootstrap confidence intervals for the pareto index. *Communications in Statistics - Theory and Methods*, **29**(1), 211–226.

Hall, P. (1998). Comparison of bootstrap and asymptotic approximations to the distribution of a heavy-tailed mean. *Statistica Sinica*, **8**, 887–906.

Knight, K. (1989). On the bootstrap of the sample mean in the infinite variance case. *The Annals of Statistics*, **17**, 1168–1175.

Matthys, G. and Beirlant, J. (2003). Estimating the extreme value index and high quantiles with exponential regression models. *Statistica Sinica*, **13**, 853–880.

McNeil, A. (1997). Estimating the tails of loss severity distribution using extreme value theory. *Astin Bulletin: A Journal of the International Actuarial Association*, **27**, 117–137.

Peng, L. and Qi, Y. (2006). Confidence regions for high quantiles of a heavy tailed distribution. *The Annals of Statistics*, **34**(4), 1964–1986.

Peng, L. and Qi, Y. (2017). *Inference for Heavy-Tailed Data: Applications in Insurance and Finance*. Academic Press.

Resnick, S. I. (1997). Discussion of the danish data on large fire insurance losses. *Astin Bulletin: A Journal of the International Actuarial Association*, **27**, 139–151.

Scarrott, C. and MacDonald, A. (2012). A review of extreme value threshold estimation and uncertainty quantification. *Revstat Statistical Journal*, **10**(1), 33–60.

Supplementary Materials for "Bootstrapping tail statistics: tail quantile process, Hill estimator, and confidence intervals for high-quantiles of heavy tailed distributions"

Svetlana Litvinova[1] and Mervyn J Silvapulle[2]

Department of Econometrics and Business Statistics,

Monash Business School, Monash University, Australia.

svetlana.litvinova@monash.edu; mervyn.silvapulle@monash.edu

August 2018

Abstract

This Supplement section provides the proofs of the theorems and the proposition in the main paper. There are five appendices. The first two recall some known results on tail quantile processes and tail-quantile bootstrap processes. These results are stated here for convenience and also to introduce notation. They are crucial for the proofs of our main theorems. The third provides proofs for some preliminary lemmas; these are new results. The fourth appendix provides the proofs of the theorems in the main paper. The fifth provides some simulation results.

Keywords: Full-sample bootstrap; Intermediate order statistic; Hill estimator; Extreme value Index; Tail-empirical process; Tail-quantile process; Tail-quantile bootstrap processes.

[1]Svetlana Litvinova gratefully acknowledges the support from Monash University through a Postgraduate Scholarship.

[2]Silvapulle gratefully acknowledges support from the Australian Research Council

1 Some known results on tail quantile processes.

Let $V_1, \ldots, V_n$ be independent and identically distributed random variables with the Uniform(0,1) distribution. Let $V_{1,n} \leq \ldots, V_{n,n}$ denote order statistics of $V_1, \ldots, V_n$. Let

$$U_n(x) = n^{-1} \sum_{i=1}^{n} I(V_i \leq x)$$

denote the empirical distribution function of $V_1, \ldots, V_n$. Define the *uniform quantile process* of $\{V_1, \ldots, V_n\}$ by

$$\beta_n(s) = n^{1/2}\{s - V_n(s)\} \quad (0 \leq s \leq 1),$$

where

$$V_n(s) = \begin{cases} V_{k,n} & ((k-1)/n < s \leq k/n, k = 1, \ldots, n), \\ V_{1,n} & (s = 0), \end{cases}$$

is the *empirical quantile function.*

Let us recall the following asymptotic approximation of the weighted uniform quantile process:

Theorem A.1 (Theorem 2.1 in Csörgő *et al.* 1986)**.** *There exists a probability space* $(\Omega, \mathcal{A}, P)$ *which carries a sequence of independent uniform (0,1) random variables* $V_1, V_2, \ldots$ *and a sequence of Brownian bridges* $B_1, B_2, \ldots$ *such that*

$$\sup_{\lambda/n \leq s \leq 1-\lambda/n} \frac{|\beta_n(s) - B_n(s)|}{\{s(1-s)\}^{1/2-\nu}} = O_P(n^{-\nu}) \tag{A.1}$$

for all $0 < \lambda < \infty$ *and* $0 \leq \nu < 1/2$.

Csörgő and Mason (1989) established a result similar to (A.1) for the bootstrap version of the uniform empirical quantile process; since this is used in the proof below, we recall it here. First, they showed that there exists a probability space, denoted $(\Omega', \mathcal{A}', P')$ that carries $\{V_i\}$ and $\{B_i\}$, a sequence of independent uniform (0,1) random variables $V'_1, V'_2, \ldots$, and a sequence of Brownian bridges $B'_1, B'_2, \ldots$ such that the two sets of random elements $\{V_i\}_{i=1}^{\infty} \cup \{B_i(s) : \ 0 \leq s \leq 1\}_{i=1}^{\infty}$ and $\{V'_i\}_{i=1}^{\infty} \cup \{B'_i(s) : \ 0 \leq s \leq 1\}_{i=1}^{\infty}$ are independent.

Let $\{V'_{i,n}, \beta'_n(s), V'_n(s), i = 1, \ldots, n, 0 \le s \le 1\}$ be defined in the same way as $\{V_{i,n}, \beta_n(s), V_n(s), i = 1, \ldots, n, 0 \le s \le 1\}$ except that $\{V_i\}_{i=1}^{\infty}$ is replaced by $\{V'_i\}_{i=1}^{\infty}$. Therefore, by (A.1), we have

$$\sup_{\lambda/n \le s \le 1-\lambda/n} \frac{|\beta'_n(s) - B'_n(s)|}{\{s(1-s)\}^{1/2-\nu}} = O_{P'}(n^{-\nu}) \quad (0 < \lambda < \infty, 0 \le \nu < 1/2).$$

Let

$$V_i^* = V_n(V'_i) \quad (i = 1, \ldots, n), \quad V_n^*(s) = V_n\{V'_n(s)\} \quad (0 \le s \le 1).$$

Then $\{V_1^*, \ldots, V_n^*\} \subseteq \{V_1, \ldots, V_n\}$ and $V_n^*(s)$ represents the bootstrap empirical quantile function. Let $\{V_{1,n}^*, \ldots V_{n,n}^*\}$ denote the order statistics of $\{V_{1,n}, \ldots V_{n,n}\}$. Then, we have

$$V_{[ks],n}^* = V_n(V'_{[ks],n}) = V_n(V'_{[nks/n],n}) = V_n\{V'_n(ks/n)\} = V_n^*(ks/n).$$

Define the *bootstrap uniform quantile process* as

$$\beta_n^*(s) = n^{1/2}\{V_n(s) - V_n^*(s)\} \quad (0 \le s \le 1).$$

Next, let us recall the following result.

Theorem A.2 (Theorem 2.1, Csörgő and Mason 1989)**.** *On the probability space* $(\Omega', \mathcal{A}', P')$,

$$\sup_{\lambda/n \le s \le 1-\lambda/n} \frac{|\beta_n^*(s) - B'_n(s)|}{(s(1-s))^{1/2-\nu}} = O_{P'}(n^{-\nu}) \tag{A.2}$$

for all $0 < \lambda < \infty$ *and* $0 \le \nu < 1/4$.

Let k denote an *intermediate order sequence* satisfying

$$k \to \infty, k/n \to 0 \text{ as } n \to \infty. \tag{A.3}$$

Define the *uniform tail quantile process* as

$$\beta_{k,n}(s) = \left(\frac{n}{k}\right)^{1/2} \beta_n\left(\frac{ks}{n}\right) = k^{1/2}\left(s - \frac{n}{k}V_{[ks],n}\right) \quad (1/k \le s \le 1),$$

and also the uniform tail quantile bootstrap processes as

$$\beta^*_{k,n}(s) = \left(\frac{n}{k}\right)^{1/2} \beta^*_n\left(\frac{ks}{n}\right) = k^{1/2}\left(\frac{n}{k}V_{[ks],n} - \frac{n}{k}V^*_{[ks],n}\right) \quad (1/k \le s \le 1) \tag{A.4}$$

where $[x]$ denotes the smallest positive integer larger than or equal to x. The uniform tail quantile process has been studied in the literature (Csörgő and Horváth, 1987). The following result is from Peng and Qi (2017). Let

$$B_{k,n}(s) = (n/k)^{1/2} B_n(ks/n), \tag{A.5}$$

and B_n is as it appears in (A.1).

Theorem A.3 (Theorem 2.3, Peng and Qi 2017)**.** *For any $0 \le \nu < 1/2$, we have*

$$\sup_{1 \le sk \le n-1} \frac{|\beta_{k,n}(s) - B_{k,n}(s)|}{s^\nu} = o_P(1). \tag{A.6}$$

In the following theorem we establish an asymptotic result for the bootstrap uniform quantile process. Let

$$B'_{k,n}(s) = (n/k)^{1/2} B'_n(ks/n),$$

where B'_n is as in (A.2).

Lemma A.1 (Remark 1 in Csörgő and Mason (1989))**.** *For $0 < \delta \le 1/2$*

$$a^{-\delta} \sup_{0 \le s \le a} s^{-1/2+\delta} |B_n(s)| \xrightarrow{d} \sup_{0 \le s \le 1} s^{-1/2+\delta} |W(s)|, \tag{A.7}$$

as $a \downarrow 0$ where W is a standard Wiener process on $[0, 1]$.

Lemma A.2 (Csörgő and Mason (1989), p.1454)**.** *On the probability space $(\Omega', \mathcal{A}', P')$ of Theorem A.2*

$$\sup_{0<s<1} \frac{U_n(s)}{s} + \sup_{0<s\le 1} \frac{1 - U_n(s)}{1 - s} = O_{P'}(1), \tag{A.8}$$

$$\sup_{V_{1,n} \le s < 1} \frac{s}{U_n(s)} + \sup_{0<s<V_{n,n}} \frac{1 - s}{1 - U_n(s)} = O_{P'}(1), \tag{A.9}$$

$$\sup_{0<s<1} \frac{s}{V_n(s)} + \sup_{0<s\le 1} \frac{1 - s}{1 - V_n(s)} = O_{P'}(1), \tag{A.10}$$

$$\sup_{\rho/n \le s < 1} \frac{V_n(s)}{s} + \sup_{0<s\le 1-\rho/n} \frac{1 - V_n(s)}{1 - s} = O_{P'}(1), \qquad \text{for any } 0 < \rho < \infty. \tag{A.11}$$

Lemma A.3 (Theorem 2.3 in Csörgő and Mason (1989)). *On the probability space* $(\Omega', \mathcal{A}', P')$ *of Theorem A.2*

$$\sup_{0<s<1} \frac{U_n^*(s)}{s} + \sup_{0<s\le 1} \frac{1-U_n^*(s)}{1-s} = O_{P'}(1), \tag{A.12}$$

$$\sup_{V_{1,n}^* \le s<1} \frac{s}{U_n^*(s)} + \sup_{0<s<V_{n,n^*}} \frac{1-s}{1-U_n^*(s)} = O_{P'}(1), \tag{A.13}$$

$$\sup_{0<s<1} \frac{s}{V_n^*(s)} + \sup_{0<s\le 1} \frac{1-s}{1-V_n^*(s)} = O_{P'}(1), \tag{A.14}$$

$$\sup_{\rho/n\le s<1} \frac{V_n^*(s)}{s} + \sup_{0<s\le 1-\rho/n} \frac{1-V_n^*(s)}{1-s} = O_{P'}(1), \qquad \textit{for any } 0<\rho<\infty. \tag{A.15}$$

2 Some results on extreme value theory

Recall that

$$U(y) = F^{\leftarrow}\{1-(1/y)\} \quad (y>0). \tag{A.16}$$

where $F^{\leftarrow}$ is the inverse function. If there exists a positive function $a(t)$ $(t \in \mathbb{R})$ such that (for example, Theorem 1.1.6, de Haan and Ferreira 2006)

$$\lim_{t\to\infty} \{U(tx)-U(t)\}/a(t) = \{x^\gamma - 1\}/\gamma, \quad (x>0)$$

then U is said to satisfy the *first-order condition of regular variation*. To establish convergence in distribution of statistics related to the tail of F, a second-order refinement is also necessary. The function U is said to satisfy the *second-order condition of regular variation* if there exists a function H that is not a multiple of $(x^\gamma-1)/\gamma$, and a positive or negative function A with $\lim_{t\to\infty} A(t) = 0$, such that

$$\lim_{t\to\infty} [\{U(tx)-U(t)\}/a(t) - \{x^\gamma-1\}/\gamma]/A(t) = H(x) \quad (x>0). \tag{A.17}$$

For the heavy-tailed distributions, namely for the case $\gamma > 0$, this condition reduces to

$$\{U(tx)/U(t) - x^\gamma\}/A(t) \to x^\gamma(x^\rho - 1)/\rho \text{ as } t\to\infty \tag{A.18}$$

for some $\rho \le 0$.

Proposition 1 (Corollary 2.3.5 in de Haan and Ferreira (2006)). *Suppose that relation (A.17) holds. Then there exist a positive function $a_*(t)$ and a function $A_*(t)$ of constant sign such that*

$$\lim_{t\to\infty} \frac{\frac{U(tx)-U(t)}{a_*(t)} - \frac{x^\gamma-1}{\gamma}}{A_*(t)} = \Psi_{\gamma,\rho}(x), \quad x>0,, \tag{A.19}$$

where

$$\Psi_{\gamma,\rho}(x) = \begin{cases} \frac{x^{\gamma+\rho}-1}{\gamma+\rho} & \gamma+\rho\neq 0,\ \rho<0, \\ \log x, & \gamma+\rho=0,\ \rho<0, \\ \frac{1}{\gamma}x^\gamma \log x, & \rho=0\neq\gamma, \\ \frac{1}{2}(\log x)^2, & \rho=0=\gamma, \end{cases} \tag{A.20}$$

$$a_*(t) = \begin{cases} a(t)(1-\rho^{-1}A(t)) & \rho<0, \\ a(t)(1-\gamma^{-1}A(t)) & \rho=0\neq\gamma, \\ a(t), & \rho=0=\gamma, \end{cases} \tag{A.21}$$

and

$$A_*(t) = \begin{cases} \rho^{-1}A(t) & \rho<0, \\ a(t) & \rho=0. \end{cases} \tag{A.22}$$

The following is a uniform bound for the second order approximation (A.19).

Lemma A.4 (Theorem 2.3.6 in de Haan and Ferreira (2006)). *Suppose that (A.19) holds for some fixed $\gamma\in\mathbb{R}$ and $\rho\leq 0$. Then there are functions a_0 and A_0 satisfying, as $t\to\infty$, $A_0(t)/A_*(t)=o(1)$, $a_0(t)/a_*(t)-1=o(A_*(t))$, with the following property: for any $\epsilon>0$ and $\delta>0$, there exists $t_0=t_0(\epsilon,\delta)$ such that for $t\geq t_0$ and $tx\geq t_0$, we have*

$$\left|\frac{\frac{U(tx)-U(t)}{a_0(t)} - \frac{x^\gamma-1}{\gamma}}{A_0(t)} - \Psi_{\gamma,\rho}(x)\right| \leq \epsilon x^{\gamma+\rho}\max(x^\delta, x^{-\delta}). \tag{A.23}$$

Lemma A.5 (de Haan and Ferreira (2006), p60). *The second-order condition (A.18) is equivalent to*

$$\lim_{t\to\infty} \frac{\log U(tx)-\log U(t)-\gamma\log x}{A(t)} = \frac{x^\rho-1}{\rho}. \tag{A.24}$$

Moreover, the following inequality holds. For a possibly different function $A_0(t)$, with $A_0(t)\sim A(t)$, $t\to\infty$, and for any $\epsilon>0$ and $\delta>0$, there exists a t_0 such that for $t\geq t_0$ and $tx\geq t_0$, we have

$$\left|\frac{\log U(tx)-\log U(t)-\gamma\log x}{A_0(t)} - \frac{x^\rho-1}{\rho}\right| \leq \epsilon x^\rho \max\{x^\delta, x^{-\delta}\}. \tag{A.25}$$

3 Preliminary lemmas to prove the theorems in the paper

Lemma A.6. *For $0 \leq \nu < 1/2$ and $\epsilon > 0$, we have*

$$\sup_{0\leq s\leq 1} s^{-\nu}|B_{k,n}(s)| = O_{P'}(1), \qquad \sup_{0\leq s\leq 1} s^{-\nu}|B'_{k,n}(s)| = O_{P'}(1), \tag{A.26}$$

$$\lim_{\delta\downarrow 0}\liminf_{n\to\infty} P'\left\{\sup_{0\leq s\leq\delta} s^{-\nu}|B_{k,n}(s)| < \epsilon\right\} = 1. \tag{A.27}$$

Proof. Let us note that the value $s^{-\nu}|B_{k,n}(s)|$ at $s = 0$ is defined by continuity.

Proof of (A.26): Substituting for $B_{k,n}$ in (A.5), we have

$$\begin{aligned}
&\sup_{0\leq s\leq 1} s^{-\nu}|B_{k,n}(s)| = \sup_{0\leq s\leq 1} s^{-\nu}(n/k)^{1/2}|B_n(ks/n)| \\
=&\sup_{0\leq t\leq k/n} (nt/k)^{-\nu}(n/k)^{1/2}|B_n(t)| = \sup_{0\leq t\leq k/n} (t)^{-\nu}|B_n(t)|(n/k)^{1/2-\nu} \\
=&(k/n)^{-\mu}\sup_{0\leq t\leq k/n} t^{-1/2+\mu}|B_n(t)|, \qquad [\text{since } 0 < \mu \leq 1/2].
\end{aligned}$$

Then, for any $x \in \mathbb{R}$,

$$\begin{aligned}
P'\left\{\sup_{0\leq s\leq 1} s^{-\nu}|B_{k,n}(s)| < x\right\} &\leq P'\left\{(k/n)^{-\mu}\sup_{0\leq t\leq k/n} t^{-1/2+\mu}|B_n(t)| < x\right\} \\
&\to P'\left\{\sup_{0\leq t\leq 1} t^{-1/2+\mu}|W(t)| < x\right\}
\end{aligned}$$

as $n \to \infty$ by Lemma A.1. The statement (A.26) is immediate since $\sup_{0\leq t\leq 1} t^{-1/2+\mu}|W(t)|$ is bounded in probability.

Proof of (A.27): Let $\mu = 1/2 - \nu$, then

$$\begin{aligned}
&\sup_{0\leq s\leq \delta} s^{-\nu}|B_{k,n}(s)| = \sup_{0\leq s\leq \delta} s^{-\nu}(n/k)^{1/2}|B_n(ks/n)| \\
=&\sup_{0\leq t\leq k\delta/n} (nt/k)^{-\nu}(n/k)^{1/2}|B_n(t)| = \sup_{0\leq t\leq k\delta/n} (t)^{-\nu}|B_n(t)|(n/k)^{1/2-\nu} \\
=&(k/n)^{-\mu}\sup_{0\leq t\leq k\delta/n} t^{-1/2+\mu}|B_n(t)| = \delta^{\mu}(\delta k/n)^{-\mu}\sup_{0\leq t\leq k\delta/n} t^{-1/2+\mu}|B_n(t)|.
\end{aligned}$$

Let $\epsilon_1 > 0$ and $\epsilon > 0$ be given. By Lemma A.1, for any $x \in \mathbb{R}$ and $\delta > 0$,

$$\lim_{n\to\infty} P'\left\{(\delta k/n)^{-\mu}\sup_{0\leq t\leq k\delta/n} t^{-1/2+\mu}|B_n(t)| < x\right\} = P'\left\{\sup_{0\leq t\leq 1} t^{-1/2+\mu}|W(t)| < x\right\}.$$

Therefore, there exists n_0 such that $\forall n \geq n_o$

$$\left| P'\left\{(\delta k/n)^{-\mu} \sup_{0\leq t\leq k\delta/n} t^{-1/2+\mu}|B_n(t)| < x\right\} - P'\left\{\sup_{0\leq t\leq 1} t^{-1/2+\mu}|W(t)| < x\right\}\right| < \frac{\epsilon_1}{2}.$$

Since $\sup_{0\leq t\leq 1} t^{-1/2+\mu}|W(t)|$ is bounded in probability, there exists $M > 0$ such that

$$P'\left\{\sup_{0\leq t\leq 1} t^{-1/2+\mu}|W(t)| < M\right\} > 1 - \frac{\epsilon_1}{2}.$$

Then, for any $n \geq n_0$,

$$\begin{aligned} P'\left\{(\delta k/n)^{-\mu} \sup_{0\leq t\leq k\delta/n} t^{-1/2+\mu}|B_n(t)| < M\right\} &\geq P'\left\{\sup_{0\leq t\leq 1} t^{-1/2+\mu}|W(t)| < M\right\} - \frac{\epsilon_1}{2} \\ &\geq 1 - \frac{\epsilon_1}{2} - \frac{\epsilon_1}{2} = 1 - \epsilon_1. \end{aligned}$$

Therefore,

$$\liminf_{n\to\infty} P'\left\{(\delta k/n)^{-\mu} \sup_{0\leq t\leq k\delta/n} t^{-1/2+\mu}|B_n(t)| < M\right\} \geq 1 - \epsilon_1. \tag{A.28}$$

Since $\delta^\mu \downarrow 0$ as $\delta \downarrow 0$, where $0 < \mu < 1/2$, there exists $\delta_0 > 0$ such that for any $0 < \delta < \delta_0$ $\delta^\mu < \frac{\epsilon}{M}$. Let $0 < \delta < \delta_0$. Then, since $\epsilon\delta^{-\mu} > M$, we have

$$P'\left\{\delta^\mu(\delta k/n)^{-\mu} \sup_{0\leq t\leq k\delta/n} t^{-1/2+\mu}|B_n(t)| < \epsilon\right\} = P'\left\{(\delta k/n)^{-\mu} \sup_{0\leq t\leq k\delta/n} t^{-1/2+\mu}|B_n(t)| < \epsilon\delta^{-\mu}\right\}$$
$$\geq P'\left\{\left((\delta k/n)^{-\mu} \sup_{0\leq t\leq k\delta/n} t^{-1/2+\mu}|B_n(t)| < M\right)\right\}.$$

Take $\liminf_{n\to\infty}$ on both sides of the last inequality with $\delta > 0$ held fixed in the range $0 < \delta < \delta_0$.

$$\begin{aligned} &\liminf_{n\to\infty} P'\left\{\delta^\mu(\delta k/n)^{-\mu} \sup_{0\leq t\leq k\delta/n} t^{-1/2+\mu}|B_n(t)| < \epsilon\right\} \\ \geq &\liminf_{n\to\infty} P'\left\{(\delta k/n)^{-\mu} \sup_{0\leq t\leq k\delta/n} t^{-1/2+\mu}|B_n(t)| < M\right\} \\ \geq &1 - \epsilon_1, \qquad \text{[by (A.28)]} \end{aligned}$$

Take $\lim_{\delta\downarrow 0}$ on both sides of the last inequality.

$$\lim_{\delta\downarrow 0}\liminf_{n\to\infty} P'\left\{\delta^{\mu}(\delta k/n)^{-\mu}\sup_{0\le t\le k\delta/n} t^{-1/2+\mu}|B_n(t)| < \epsilon\right\} \ge 1-\epsilon_1.$$

We obtained

$$1 \ge \lim_{\delta\downarrow 0}\liminf_{n\to\infty} P'\left\{\delta^{\mu}(\delta k/n)^{-\mu}\sup_{0\le t\le k\delta/n} t^{-1/2+\mu}|B_n(t)| < \epsilon\right\} \ge 1-\epsilon_1$$

Since $\epsilon_1 > 0$ is arbitrary,

$$\lim_{\delta\downarrow 0}\liminf_{n\to\infty} P'\left\{\delta^{\mu}(\delta k/n)^{-\mu}\sup_{0\le t\le k\delta/n} t^{-1/2+\mu}|B_n(t)| < \epsilon\right\} = 1.$$

□

Lemma A.7. *On the probability space* $(\Omega', \mathcal{A}', P')$ *of Theorem A.2, for any* $0 \le \nu < 1/2$*, we have*

$$\sup_{k^{-1}<s\le 1} \frac{\left|\beta^*_{k,n}(s) - B'_{k,n}(s)\right|}{s^{\nu}} = o_{P'}(1). \tag{A.29}$$

Proof. Choose any $0 \le \nu < 1/2$. Let $\mu = (1/2) - \nu$ and δ be any number in the range $0 \le \delta < \min(\mu, 1/4)$. Then for any $0 < t \le k/n$, we have

$$(n/k)^{\mu} t^{\mu-1/2} \le (n/k)^{\delta} t^{\delta-1/2} \tag{A.30}$$

since

$$\frac{(n/k)^{\mu} t^{\mu-1/2}}{(n/k)^{\delta} t^{\delta-1/2}} = \left(\frac{n}{k}\right)^{\mu-\delta} t^{\mu-\delta} \le \left(\frac{n}{k}\right)^{\mu-\delta}\left(\frac{k}{n}\right)^{\mu-\delta} = 1.$$

Therefore,

$$\begin{aligned}
\sup_{1/k \le s \le 1} \frac{\left|\beta^*_{k,n}(s) - B'_{k,n}(s)\right|}{s^\nu} &= \sup_{1/k \le s \le 1} (n/k)^{1/2} \frac{|\beta^*_n(ks/n) - B'_n(ks/n)|}{s^\nu} \\
&= \sup_{1/k \le s \le 1} (n/k)^{1/2-\nu} \frac{|\beta^*_n(ks/n) - B'_n(ks/n)|}{(ks/n)^\nu} \\
&= \sup_{1/n \le t \le k/n} (n/k)^{1/2-\nu} \frac{|\beta^*_n(t) - B'_n(t)|}{t^\nu} \\
&= \sup_{1/n \le t \le k/n} (n/k)^{\mu} \frac{|\beta^*_n(t) - B'_n(t)|}{t^{1/2-\mu}} \quad (0 \le \mu < 1/2) \\
&\le \sup_{1/n \le t \le k/n} (n/k)^{\delta} \frac{|\beta^*_n(t) - B'_n(t)|}{t^{1/2-\delta}} \qquad \text{[by (A.30)]} \\
&\le \sup_{1/n \le t \le k/n} n^\delta \frac{|\beta^*_n(t) - B'_n(t)|}{\{t(1-t)\}^{1/2-\delta}} k^{-\delta} \sup_{1/n \le t \le k/n} (1-t)^{1/2-\delta} \\
&= O_{P'}(1) k^{-\delta} \{1 - (1/n)\}^{1/2-\delta} \qquad \text{[by (A.2)]} \\
&= O_{P'}(k^{-\delta}) \\
&= o_{P'}(1).
\end{aligned}$$

□

Lemma A.8. *Let $\epsilon > 0$ be given. Then, on the probability space $(\Omega', \mathcal{A}', P')$ of Theorem A.2, we have*

$$\sup_{1/k \le s \le 1} s^{1/2+\epsilon} \left| k^{1/2} \left(\log\left(\frac{k}{nV^*_{[ks],n}}\right) - \log\left(\frac{k}{nV_{[ks],n}}\right) \right) - \frac{B'_{k,n}(s)}{s} \right| = o_{P'}(1). \quad \text{(A.31)}$$

Proof. Let $s \in [1/k, 1]$,

$$\xi_l(s) = \min\{(n/k)V^*_{[ks],n}, (n/k)V_{[ks],n}\}, \quad \xi_u(s) = \max\{(n/k)V^*_{[ks],n}, (n/k)V_{[ks],n}\}. \tag{A.32}$$

By applying the mean value theorem to $\varphi(t) = \log(1/t)$, we have

$$k^{1/2} \left(\log\left(\frac{k}{nV^*_{[ks],n}}\right) - \log\left(\frac{k}{nV_{[ks],n}}\right) \right) = -\frac{1}{\xi_n} k^{1/2} \left(\frac{n}{k} V^*_{[ks],n} - \frac{n}{k} V_{[ks],n}\right) = \xi_n^{-1}(s) \beta^*_{k,n}(s), \tag{A.33}$$

where $\xi_n(s) \in [\xi_l(s), \xi_u(s)]$. Then,

$$
\begin{aligned}
&\sup_{1/k\le s\le 1} s^{1/2+\epsilon}\left|k^{1/2}\left(\log\left(\frac{k}{nV^*_{[ks],n}}\right)-\log\left(\frac{k}{nV_{[ks],n}}\right)\right)-\frac{B'_{k,n}(s)}{s}\right|\\
=&\sup_{1/k\le s\le 1} s^{1/2+\epsilon}\left|\xi_n^{-1}(s)\beta^*_{k,n}(s)-s^{-1}B'_{k,n}(s)\right| \qquad [\text{by (A.33) }]\\
\le&\sup_{1/k\le s\le 1} s^{1/2+\epsilon}\frac{1}{\xi_n(s)}\left|\beta^*_{k,n}(s)-B'_{k,n}(s)\right|+\sup_{1/k\le s\le 1} s^{1/2+\epsilon}B'_{k,n}(s)\left|\frac{1}{s}-\frac{1}{\xi_n(s)}\right|\\
=&\sup_{1/k\le s\le 1}\frac{s}{\xi_n(s)}\frac{\left|\beta^*_{k,n}(s)-B'_{k,n}(s)\right|}{s^{1/2-\epsilon}}+\sup_{1/k\le<s\le 1}\frac{s^{1/2+\epsilon}}{s}B'_{k,n}(s)\frac{|\xi_n(s)-s|}{\xi_n(s)}\\
=&A_1+A_2,
\end{aligned}
$$

where

$$
A_1=\sup_{1/k\le s\le 1}\frac{s}{\xi_n(s)}\sup_{1/k\le s\le 1}\frac{\left|\beta^*_{k,n}(s)-B'_{k,n}(s)\right|}{s^{1/2-\epsilon}},\quad A_2=\sup_{1/k\le s\le 1}\frac{B'_{k,n}(s)}{s^{1/2-\epsilon}}\frac{|\xi_n(s)-s|}{\xi_n(s)}. \tag{A.34}
$$

First, we obtain bounds for A_1, and then for A_2. By Lemmas A.2 and A.3, we have

$$
\sup_{0<s\le 1}\frac{s}{(nV_{[ks],n})/k}=\sup_{0<s\le 1}\frac{ks}{n}\frac{1}{V_n(ks/n)}=\sup_{0<t\le k/n}\frac{t}{V_n(t)}=O_{P'}(1), \tag{A.35}
$$

$$
\sup_{1/k\le s\le 1}\frac{nV_{[ks],n}}{k}\frac{1}{s}=\sup_{1/k\le s\le 1}V_n(ks/n)\frac{n}{ks}=\sup_{1/n\le t\le k/n}\frac{V_n(t)}{t}=O_{P'}(1), \tag{A.36}
$$

$$
\sup_{0<s\le 1}\frac{s}{(nV^*_{[ks],n})/k}=\sup_{0<s\le 1}\frac{ks}{n}\frac{1}{V^*_n(ks/n)}=\sup_{0<t\le k/n}\frac{t}{V^*_n(t)}=O_{P'}(1), \tag{A.37}
$$

$$
\sup_{1/k\le s\le 1}\frac{nV^*_{[ks],n}}{k}\frac{1}{s}=\sup_{1/k\le s\le 1}V^*_n(ks/n)\frac{n}{ks}=\sup_{1/n\le t\le k/n}\frac{V^*_n(t)}{t}=O_{P'}(1). \tag{A.38}
$$

Therefore, by (A.35) and (A.37),

$$
\sup_{0<s\le 1}\frac{s}{\xi_n(s)}=O_{P'}(1). \tag{A.39}
$$

Similarly, by (A.36) and (A.38),

$$
\sup_{1/k\le s\le 1}\frac{\xi_n(s)}{s}=O_{P'}(1). \tag{A.40}
$$

Hence, $A_1=o_{P'}(1)$ by (A.39) and (A.29).

Next, we obtain a bound for A_2. Let $\delta \in (0,1)$, and choose $k > \delta^{-1}$. Then, $1/k < \delta \le 1$, and we have

$$\begin{aligned}
A_2 &= \sup_{1/k \le s \le 1} \frac{B'_{k,n}(s)}{s^{1/2-\epsilon}} \frac{|\xi_n(s) - s|}{\xi_n(s)} \\
&\le \sup_{1/k \le s \le \delta} \frac{B'_{k,n}(s)}{s^{1/2-\epsilon}} \frac{|\xi_n(s) - s|}{\xi_n(s)} + \sup_{\delta \le s \le 1} \frac{B'_{k,n}(s)}{s^{1/2-\epsilon}} \frac{|\xi_n(s) - s|}{\xi_n(s)} \\
&\le \sup_{1/k \le s \le \delta} \frac{B'_{k,n}(s)}{s^{1/2-\epsilon}} \sup_{1/k \le s \le \delta} \frac{|\xi_n(s) - s|}{\xi_n(s)} + \sup_{\delta < s \le 1} \frac{B'_{k,n}(s)}{s^{1/2-\epsilon}} \sup_{\delta < s \le 1} \frac{|\xi_n(s) - s|}{\xi_n(s)} \\
&\le B_1(n,k,\delta) B_2(n,k) \quad + \quad B_3(n,k) B_4(n,k,\delta),
\end{aligned}$$

where

$$B_1 = \sup_{0 < s \le \delta} s^{\epsilon - 1/2} B'_{k,n}(s), \qquad B_2 = \sup_{1/k \le s \le 1} \frac{|\xi_n(s) - s|}{\xi_n(s)} \tag{A.41}$$

$$B_3 = \sup_{0 < s \le 1} s^{\epsilon - 1/2} B'_{k,n}(s), \qquad B_4 = \sup_{\delta < s \le 1} \frac{|\xi_n(s) - s|}{\xi_n(s)}. \tag{A.42}$$

By (A.39),

$$B_2(n,k) = \sup_{1/k \le s \le 1} \frac{|\xi_n(s) - s|}{\xi_n(s)} = \sup_{1/k \le s \le 1} \left| 1 - \frac{s}{\xi_n(s)} \right| = O_{P'}(1). \tag{A.43}$$

Also, since $\sup_{\delta < s \le 1} s/\xi_n(s) = O_{P'}(1)$ by (A.39), and

$$\sup_{\delta < s \le 1} 1/s = 1/\delta < \infty, \tag{A.44}$$

for any $\delta > 0$, we have

$$\sup_{\delta < s \le 1} 1/\xi_n(s) = O_{P'}(1). \tag{A.45}$$

By Theorem A.3, for any $0 \le \nu < 1/2$

$$\sup_{1 \le sk \le n-1} \frac{|\beta_{k,n}(s) - B_{k,n}(s)|}{s^\nu} = o_{P'}(1).$$

Since $\sup_{\delta < s \le 1} s^{-\nu} B_{k,n}(s) = O_{P'}(1)$ by Lemma A.6 and $(1/k) < \delta$, we have

$$\sup_{\delta < s \le 1} s^{-\nu} |\beta_{k,n}(s)| = O_{P'}(1).$$

Then, using the definition in (A.4), we have

$$\sup_{\delta<s\leq 1} s^{-\nu}\left|s-\frac{n}{k}V_{[ks],n}\right| = O_{P'}(k^{-1/2}) = o_{P'}(1),$$

and hence

$$\sup_{\delta<s\leq 1}\left|s-\frac{n}{k}V_{[ks],n}\right| = o_{P'}(1). \tag{A.46}$$

By Lemma A.7, for any $0\leq\nu<1/2$

$$\sup_{\delta<s\leq 1}\frac{\left|\beta^*_{k,n}(s)-B'_{k,n}(s)\right|}{s^{\nu}} = o_{P'}(1).$$

Since $\sup_{\delta<s\leq 1} s^{-\nu}B'_{k,n}(s) = O_{P'}(1)$ by Lemma A.6,we have

$$\sup_{\delta<s\leq 1} s^{-\nu}|\beta^*_{k,n}(s)| = O_{P'}(1).$$

Therefore, substituting for $\beta^*_{k,n}(s)$ using (A.4), we have

$$\sup_{\delta<s\leq 1} s^{-\nu}\left|\frac{n}{k}V_{[ks],n}-\frac{n}{k}V^*_{[ks],n}\right| = O_{P'}(k^{-1/2}) = o_{P'}(1)$$

and

$$\sup_{\delta<s\leq 1}\left|\frac{n}{k}V_{[ks],n}-\frac{n}{k}V^*_{[ks],n}\right| = o_{P'}(1). \tag{A.47}$$

By (A.46) and (A.47)

$$\sup_{\delta<s\leq 1}\left|s-\frac{n}{k}V^*_{[ks],n}\right| = o_{P'}(1). \tag{A.48}$$

By (A.46) and (A.48),

$$\sup_{\delta<s\leq 1}|s-\xi_n(s)| = o_{P'}(1). \tag{A.49}$$

Thus,

$$B_4(n,k,\delta) = \sup_{\delta<s\leq 1}\frac{|\xi_n(s)-s|}{\xi_n(s)} \leq \left(\sup_{\delta<s\leq 1} 1/\xi_n(s)\right)\left(\sup_{\delta<s\leq 1}|s-\xi_n(s)|\right) = o_{P'}(1) \tag{A.50}$$

by (A.49) and (A.45).

Let $\epsilon_1>0$ and $\epsilon_2>0$ be given. Since $B_2(n,k) = O_{P'}(1)$ as $n\to\infty$ by (A.43), there exists $M_1 = M_1(\epsilon)$ and $n_1 = n_1(\epsilon)$ such that for all $n\geq n_1$

$$P'\left[B_2(n,k)<M_1\right] > 1-\epsilon_1.$$

Therefore,

$$\liminf_{n\to\infty} P'\left[B_2(n,k) < M_1\right] \geq 1-\epsilon_1. \tag{A.51}$$

We have

$$\begin{aligned}
&P'\left\{B_1(n,k,\delta)B_2(n,k) < \epsilon_2\right\}\\
\geq &P'\left\{B_1(n,k,\delta) < \frac{\epsilon_2}{M_1} \cap B_2(n,k) < M_1\right\}\\
\geq &P'\left\{B_1(n,k,\delta) < \frac{\epsilon_2}{M_1}\right\} + P'\left\{B_2(n,k) < M_1\right\} - 1.
\end{aligned}$$

Taking $\liminf_{n\to\infty}$ on both sides of the last inequality, we have

$$\begin{aligned}
&\liminf_{n\to\infty} P'\left\{B_1(n,k,\delta)B_2(n,k) < \epsilon_2\right\}\\
\geq &\liminf_{n\to\infty} P'\left\{B_1(n,k,\delta) < \frac{\epsilon_2}{M_1}\right\} + \liminf_{n\to\infty} P'\left\{B_2(n,k) < M_1\right\} - 1.
\end{aligned}$$

Next, taking $\lim_{\delta\downarrow 0}$ on both sides, we have

$$\begin{aligned}
&\lim_{\delta\downarrow 0}\liminf_{n\to\infty} P'\left\{B_1(n,k,\delta)B_2(n,k) < \epsilon_2\right\}\\
\geq &\lim_{\delta\downarrow 0}\liminf_{n\to\infty} P'\left\{B_1(n,k,\delta) < \frac{\epsilon_2}{M_1}\right\} + \lim_{\delta\downarrow 0}\liminf_{n\to\infty} P'\left\{B_2(n,k) < M_1\right\} - 1\\
= &1 + \lim_{\delta\downarrow 0}\liminf_{n\to\infty} P'\left\{B_2(n,k) < M_1\right\} - 1 \qquad [\text{ by (A.27)}]\\
= &\liminf_{n\to\infty} P'\left\{B_2(n,k) < M_1\right\}\\
\geq &1-\epsilon_1 \qquad\qquad [\text{ by (A.51)}].
\end{aligned}$$

Thus, for any arbitrary $\epsilon_1 > 0$, we have

$$1 \geq \lim_{\delta\downarrow 0}\liminf_{n\to\infty} P'\left\{B_1(n,k,\delta)B_2(n,k) < \epsilon_2\right\} \geq 1-\epsilon_1,$$

and hence

$$\lim_{\delta\downarrow 0}\liminf_{n\to\infty} P'\left\{B_1(n,k,\delta)B_2(n,k) < \epsilon_2\right\} = 1. \tag{A.52}$$

Similarly, let us evaluate $P'\left\{B_3(n,k)B_4(n,k,\delta) < \epsilon_2\right\}$. Since $B_3(n,k) = O_{P'}(1)\ as\ n \to \infty$ by (A.26), there exists $M_2 = M_2(\epsilon_1)$ and $n_2 = n_2(\epsilon_1)$ such that for all $n \geq n_2$

$$P'\left[B_3(n,k) < M_2\right] > 1 - \frac{\epsilon_1}{2}.$$

Therefore,

$$\liminf_{n\to\infty} P'\left[B_3(n,k) < M_2\right] \geq 1 - \frac{\epsilon_1}{2}. \tag{A.53}$$

By (A.50), $B_4(n,k,\delta) = o_{P'}(1)$ as $n \to \infty$ for every $\delta > 0$, and hence there exists $n_3 = n_3(\epsilon_1)$ such that for all $n \geq n_3$

$$P'\left[B_4(n,k,\delta) < \frac{\epsilon_2}{M_1}\right] > 1 - \frac{\epsilon_1}{2}.$$

Therefore,

$$\liminf_{n\to\infty} P'\left[B_4(n,k,\delta) < \frac{\epsilon_2}{M_1}\right] \geq 1 - \frac{\epsilon_1}{2}. \tag{A.54}$$

Next, note that

$$\begin{aligned}
&P'\left\{B_3(n,k)B_4(n,k,\delta) < \epsilon_2\right\} \geq P'\left\{B_3(n,k) < M_2 \cap B_4(n,k,\delta) < \frac{\epsilon_2}{M_2}\right\} \\
\geq &P'\left\{B_3(n,k) < M_2\right\} + P'\left\{B_4(n,k,\delta) < \frac{\epsilon_2}{M_2}\right\} - 1.
\end{aligned}$$

Taking $\liminf_{n\to\infty}$ on both sides of the last inequality with $\delta > 0$ held fixed, we obtain

$$\begin{aligned}
&\liminf_{n\to\infty} P'\left\{B_3(n,k)B_4(n,k,\delta) < \epsilon_2\right\} \\
\geq &\liminf_{n\to\infty} P'\left\{B_3(n,k) < M_2\right\} + \liminf_{n\to\infty} P'\left\{B_4(n,k,\delta) < \frac{\epsilon_2}{M_1}\right\} - 1 \\
\geq &1 - \frac{\epsilon_1}{2} + 1 - \frac{\epsilon_1}{2} - 1 \qquad \text{[by (A.53) and (A.54)]} \\
= &1 - \epsilon_1.
\end{aligned}$$

Since $\epsilon_1 > 0$ is arbitrary, and $\liminf_{n\to\infty} P'\left\{B_3(n,k)B_4(n,k,\delta) < \epsilon_2\right\} \geq 1 - \epsilon_1$, we have

$$\liminf_{n\to\infty} P'\left\{B_3(n,k)B_4(n,k,\delta) < \epsilon_2\right\} = 1$$

for any fixed $\delta \in (0,1)$, and hence

$$\lim_{\delta\downarrow 0}\liminf_{n\to\infty} P'\left\{B_3(n,k)B_4(n,k,\delta) < \epsilon_2\right\} = 1. \tag{A.55}$$

Next, substituting for A_2 in (A.34), we obtain

$$\begin{aligned}
&P'\left\{A_2(n,k) < 2\epsilon_2\right\} \geq P'\left\{B_1(n,k,\delta)B_2(n,k) < \epsilon_2 \cap B_3(n,k)B_4(n,k,\delta) < \epsilon_2\right\} \\
\geq &P'\left\{B_1(n,k,\delta)B_2(n,k) < \epsilon_2\right\} + P'\left\{B_3(n,k)B_4(n,k,\delta) < \epsilon_2\right\} - 1.
\end{aligned}$$

Taking $\liminf_{n\to\infty}$ on both sides of the above inequality, we have

$$\begin{aligned}&\liminf_{n\to\infty} P'\left\{A_2(n,k) < 2\epsilon_2\right\}\\ \geq &\liminf_{n\to\infty} P'\left\{B_1(n,k,\delta)B_2(n,k) < \epsilon_2\right\} + \liminf_{n\to\infty} P'\left\{B_3(n,k)B_4(n,k,\delta) < \epsilon_2\right\} - 1.\end{aligned}$$

Now, take $\lim_{\delta\downarrow 0}$. Since the left hand side of the above inequality does not depend on δ, we have

$$\begin{aligned}&\liminf_{n\to\infty} P'\left\{A_2(n,k) < 2\epsilon_2\right\}\\ \geq &\lim_{\delta\downarrow 0}\liminf_{n\to\infty} P'\left\{B_1(n,k,\delta)B_2(n,k) < \epsilon_2\right\} + \lim_{\delta\downarrow 0}\liminf_{n\to\infty} P'\left\{B_3(n,k)B_4(n,k,\delta) < \epsilon_2\right\} - 1 = 1\end{aligned}$$

by (A.52) and (A.55). Therefore,

$$A_2(n,k) = o_{P'}(1).$$

Since we have already proved that $A_1(n,k) = o_{P'}(1)$, we have $A_1(n,k) + A_2(n,k) = o_{P'}(1)$. This completes the proof of Lemma A.8. ☐

Lemma A.9. *Let $\epsilon > 0$ be given. Then, on the probability space $(\Omega', \mathcal{A}', P')$ of Theorem A.2 we have*

$$\sup_{1/k\leq s\leq 1} s^{\gamma+(1/2)+\epsilon}\left|\frac{k^{1/2}}{\gamma}\left\{\left(\frac{k}{nV^*_{[ks],n}}\right)^{\gamma} - 1\right\} - \frac{k^{1/2}}{\gamma}\left\{\left(\frac{k}{nV_{[ks],n}}\right)^{\gamma} - 1\right\} - \frac{B'_{k,n}(s)}{s^{\gamma+1}}\right| = o_{P'}(1). \tag{A.56}$$

Proof. Let $\xi_l(s)$ and $\xi_u(s)$, $1/k \leq s \leq 1$, be defined in (A.32). Let $\varphi(t) = \gamma^{-1}(t^{-\gamma} - 1)$. Applying the mean value theorem to $\varphi(t)$ and noting that $\dot{\varphi}(t) = -t^{-\gamma-1}$, we obtain

$$\begin{aligned}&\frac{k^{1/2}}{\gamma}\left\{\left(\frac{k}{nV^*_{[ks],n}}\right)^{\gamma} - 1\right\} - \frac{k^{1/2}}{\gamma}\left\{\left(\frac{k}{nV_{[ks],n}}\right)^{\gamma} - 1\right\}\\ =& -\xi_n^{-\gamma-1}(s)k^{1/2}\left(\frac{n}{k}V^*_{[ks],n} - \frac{n}{k}V_{[ks],n}\right) = \xi_n^{-\gamma-1}(s)\beta^*_{k,n}(s),\end{aligned}$$

where $\xi_n(s) \in [\xi_l(s), \xi_u(s)]$. Substituting this in (A.56), we have

$$
\begin{aligned}
&\sup_{1/k\le s\le 1} s^{\gamma+(1/2)+\epsilon}\left|\frac{k^{1/2}}{\gamma}\left\{\left(\frac{k}{nV^*_{[ks],n}}\right)^{\gamma}-1\right\}-\frac{k^{1/2}}{\gamma}\left\{\left(\frac{k}{nV_{[ks],n}}\right)^{\gamma}-1\right\}-\frac{B'_{k,n}(s)}{s^{\gamma+1}}\right|\\
&=\sup_{1/k\le s\le 1} s^{\gamma+(1/2)+\epsilon}\left|\xi_n(s)^{-\gamma-1}\beta^*_{k,n}(s)-\frac{B'_{k,n}(s)}{s^{\gamma+1}}\right|\\
&\le \sup_{1/k\le s\le 1}\frac{s^{\gamma+(1/2)+\epsilon}}{\xi_n^{\gamma+1}(s)}\left|\beta^*_{k,n}(s)-B'_{k,n}(s)\right|+\sup_{1/k\le s\le 1} s^{\gamma+(1/2)+\epsilon}|B'_{k,n}(s)|\left|\frac{1}{s^{\gamma+1}}-\frac{1}{\xi_n^{\gamma+1}(s)}\right|\\
&=\sup_{1/k\le s\le 1}\left(\frac{s}{\xi_n(s)}\right)^{\gamma+1}\frac{\left|\beta^*_{k,n}(s)-B'_{k,n}(s)\right|}{s^{(1/2)-\epsilon}}+\sup_{1/k\le s\le 1}\frac{|B'_{k,n}(s)|}{s^{(1/2)-\epsilon}}s^{\gamma+1}\left|\frac{1}{s^{\gamma+1}}-\frac{1}{\xi_n^{\gamma+1}(s)}\right|\\
&\le A_1+A_2,
\end{aligned}
$$

where

$$
A_1=\sup_{1/k\le s\le 1}\left(\frac{s}{\xi_n(s)}\right)^{\gamma+1}\sup_{1/k\le s\le 1}\frac{\left|\beta^*_{k,n}(s)-B'_{k,n}(s)\right|}{s^{1/2-\epsilon}},
$$

$$
A_2=\sup_{1/k\le s\le 1}\frac{|B'_{k,n}(s)|}{s^{1/2-\epsilon}}s^{\gamma+1}\left|\frac{1}{s^{\gamma+1}}-\frac{1}{\xi_n^{\gamma+1}(s)}\right|.
$$

By (A.39), (A.40) and (A.29), $A_1 = o_{P'}(1)$. Next, we show that A_2 is also $o_{P'}(1)$. Let $\delta \in (0,1)$. Choose $k > \delta^{-1}$, then $1/k < \delta \le 1$. Then,

$$
\begin{aligned}
A_2 &= \sup_{1/k\le s\le 1}\frac{|B'_{k,n}(s)|}{s^{1/2-\epsilon}}s^{\gamma+1}\left|\frac{1}{s^{\gamma+1}}-\frac{1}{\xi_n^{\gamma+1}(s)}\right|\\
&\le \sup_{1/k\le s\le \delta}\frac{|B'_{k,n}(s)|}{s^{1/2-\epsilon}}s^{\gamma+1}\left|\frac{1}{s^{\gamma+1}}-\frac{1}{\xi_n^{\gamma+1}(s)}\right|+\sup_{\delta< s\le 1}\frac{|B'_{k,n}(s)|}{s^{1/2-\epsilon}}s^{\gamma+1}\left|\frac{1}{s^{\gamma+1}}-\frac{1}{\xi_n^{\gamma+1}(s)}\right|\\
&\le \sup_{1/k\le s\le \delta}\frac{|B'_{k,n}(s)|}{s^{1/2-\epsilon}}\sup_{1/k\le s\le \delta}s^{\gamma+1}\left|\frac{1}{s^{\gamma+1}}-\frac{1}{\xi_n^{\gamma+1}(s)}\right|+\sup_{\delta< s\le 1}\frac{|B'_{k,n}(s)|}{s^{1/2-\epsilon}}\sup_{\delta< s\le 1}s^{\gamma+1}\left|\frac{1}{s^{\gamma+1}}-\frac{1}{\xi_n^{\gamma+1}(s)}\right|\\
&\le \sup_{0< s\le \delta}\frac{|B'_{k,n}(s)|}{s^{1/2-\epsilon}}\sup_{1/k\le s\le 1}s^{\gamma+1}\left|\frac{1}{s^{\gamma+1}}-\frac{1}{\xi_n^{\gamma+1}(s)}\right|+\sup_{0< s\le 1}\frac{|B'_{k,n}(s)|}{s^{1/2-\epsilon}}\sup_{\delta< s\le 1}s^{\gamma+1}\left|\frac{1}{s^{\gamma+1}}-\frac{1}{\xi_n^{\gamma+1}(s)}\right|\\
&= B_1(n,k,\delta)B_2(n,k)+B_3(n,k)B_4(n,k,\delta),
\end{aligned}
$$

say. Next we consider B_1, B_2, B_3, and B_4. First, we show that

$$
B_2(n,k) = O_{P'}(1). \tag{A.57}
$$

By the definition of B_2, we have

$$\begin{aligned}B_2(n,k) = &\sup_{1/k\le s\le 1} s^{\gamma+1}\left|\frac{1}{s^{\gamma+1}} - \frac{1}{\xi_n^{\gamma+1}(s)}\right| = \sup_{1/k\le s\le 1}\left|1-\left(\frac{s}{\xi_n(s)}\right)^{\gamma+1}\right| \\ \le& 1 + \sup_{1/k\le s\le 1}\left|\frac{s}{\xi_n(s)}\right|^{\gamma+1} = O_{P'}(1)\end{aligned}$$

by (A.39) and (A.40). Next, we show that

$$B_4(n,k,\delta) = o_{P'}(1). \tag{A.58}$$

Let $\varphi(t) = t^{-\gamma-1}$; then $\dot{\varphi}(t) = (-\gamma-1)t^{-\gamma-2}$. By the mean value theorem applied to $\varphi(t)$, we have

$$\frac{1}{s^{\gamma+1}} - \frac{1}{\xi_n^{\gamma+1}(s)} = (s-\xi_n(s))\frac{(-\gamma-1)}{\tilde{\xi}_n(s)^{\gamma+2}}$$

for some $\tilde{\xi}_n(s) \in [s, \xi_n(s)]$. Then,

$$\begin{aligned}s^{\gamma+1}\left|\frac{1}{s^{\gamma+1}} - \frac{1}{\xi_n^{\gamma+1}(s)}\right| &= |-\gamma-1||s-\xi_n(s)|\frac{s^{\gamma+1}}{\tilde{\xi}_n(s)^{\gamma+2}} \\ &= |-\gamma-1|\frac{|s-\xi_n(s)|}{\tilde{\xi}_n(s)}\left(\frac{s}{\tilde{\xi}_n(s)}\right)^{\gamma+1} \\ &\le |-\gamma-1|\frac{|s-\xi_n(s)|}{\min\{s,\xi_n(s)\}}\left(\frac{s}{\min\{s,\xi_n(s)\}}\right)^{\gamma+1};\end{aligned}$$

the last inequality follows since $0 < \min\{s, \tilde{\xi}_n(s)\} < \tilde{\xi}_n(s)$. Starting with the definition of B_4, we have

$$\begin{aligned}B_4(n,k,\delta) =& (\gamma+1)\sup_{\delta<s\le 1} s^{\gamma+1}\left|\frac{1}{s^{\gamma+1}} - \frac{1}{\xi_n^{\gamma+1}(s)}\right| \\ \le& (\gamma+1)\sup_{\delta<s\le 1}\frac{|s-\xi_n(s)|}{\min\{s,\xi_n(s)\}}\sup_{\delta<s\le 1}\left(\frac{s}{\min\{s,\xi_n(s)\}}\right)^{\gamma+1} \\ \le& (\gamma+1)\sup_{\delta<s\le 1}\frac{1}{\min\{s,\xi_n(s)\}}\sup_{\delta<s\le 1}|s-\xi_n(s)|O_{P'}(1) \qquad \text{[by (A.39), (A.40)]} \\ =& O_{P'}(1)o_{P'}(1)O_{P'}(1) \qquad \text{[by (A.44), (A.45), (A.49)]} \\ =& o_{P'}(1).\end{aligned}$$

Next, we consider B_2. Let $\epsilon_1 > 0$ and $\epsilon_2 > 0$ be given. Since $B_2(n,k) = O_{P'}(1)$ as $n \to \infty$ by (A.57), there exists $M_1 = M_1(\epsilon)$ and $n_1 = n_1(\epsilon)$ such that for all $n \ge n_1$

$$P'\left[B_2(n,k) < M_1\right] > 1-\epsilon_1.$$

Therefore,

$$\liminf_{n\to\infty} P'\left[B_2(n,k) < M_1\right] \geq 1-\epsilon_1. \tag{A.59}$$

Note that

$$\begin{aligned}
&P'\left\{B_1(n,k,\delta)B_2(n,k) < \epsilon_2\right\} \geq P'\left\{B_1(n,k,\delta) < \frac{\epsilon_2}{M_1} \cap B_2(n,k) < M_1\right\} \\
\geq &P'\left\{B_1(n,k,\delta) < \frac{\epsilon_2}{M_1}\right\} + P'\left\{B_2(n,k) < M_1\right\} - 1.
\end{aligned}$$

Taking $\liminf\limits_{n\to\infty}$ on both sides of the last inequality, we obtain

$$\begin{aligned}
&\liminf_{n\to\infty} P'\left\{B_1(n,k,\delta)B_2(n,k) < \epsilon_2\right\} \\
\geq &\liminf_{n\to\infty} P'\left\{B_1(n,k,\delta) < \frac{\epsilon_2}{M_1}\right\} + \liminf_{n\to\infty} P'\left\{B_2(n,k) < M_1\right\} - 1
\end{aligned}$$

Next, take $\lim_{\delta\downarrow 0}$ on both sides.

$$\begin{aligned}
&\lim_{\delta\downarrow 0}\liminf_{n\to\infty} P'\left\{B_1(n,k,\delta)B_2(n,k) < \epsilon_2\right\} \\
\geq &\lim_{\delta\downarrow 0}\liminf_{n\to\infty} P'\left\{B_1(n,k,\delta) < \frac{\epsilon_2}{M_1}\right\} + \lim_{\delta\downarrow 0}\liminf_{n\to\infty} P'\left\{B_2(n,k) < M_1\right\} - 1 \\
= &1 + \lim_{\delta\downarrow 0}\liminf_{n\to\infty} P'\left\{B_2(n,k) < M_1\right\} - 1 \qquad [\text{ by (A.27)}] \\
= &\liminf_{n\to\infty} P'\left\{B_2(n,k) < M_1\right\} \\
\geq &1-\epsilon_1. \qquad [\text{ by (A.59)}]
\end{aligned}$$

Thus, for any arbitrary $\epsilon_1 > 0$ we have

$$1 \geq \lim_{\delta\downarrow 0}\liminf_{n\to\infty} P'\left\{B_1(n,k,\delta)B_2(n,k) < \epsilon_2\right\} \geq 1-\epsilon_1,$$

and hence

$$\lim_{\delta\downarrow 0}\liminf_{n\to\infty} P'\left\{B_1(n,k,\delta)B_2(n,k) < \epsilon_2\right\} = 1. \tag{A.60}$$

Similarly, we evaluate $P'\left\{B_3(n,k)B_4(n,k,\delta) < \epsilon_2\right\}$. Since $B_3(n,k) = O_{P'}(1)\ as\ n \to \infty$ by (A.26), there exists $M_2 = M_2(\epsilon_1)$ and $n_2 = n_2(\epsilon_1)$ such that for all $n \geq n_2$

$$P'\left[B_3(n,k) < M_2\right] > 1 - \frac{\epsilon_1}{2}.$$

Therefore,

$$\liminf_{n\to\infty} P'\left[B_3(n,k) < M_2\right] \geq 1 - \frac{\epsilon_1}{2}. \tag{A.61}$$

Since by (A.58) $B_4(n,k,\delta) = o_{P'}(1)$ as $n \to \infty$ for every $\delta > 0$, there exists $n_3 = n_3(\epsilon_1)$ such that for all $n \geq n_3$

$$P'\left[B_4(n,k,\delta) < \frac{\epsilon_2}{M_1}\right] > 1 - \frac{\epsilon_1}{2}.$$

Therefore,

$$\liminf_{n\to\infty} P'\left[B_4(n,k,\delta) < \frac{\epsilon_2}{M_1}\right] \geq 1 - \frac{\epsilon_1}{2}. \tag{A.62}$$

Note that

$$\begin{aligned}
&P'\left\{B_3(n,k)B_4(n,k,\delta) < \epsilon_2\right\} \geq P'\left\{B_3(n,k) < M_2 \cap B_4(n,k,\delta) < \frac{\epsilon_2}{M_2}\right\} \\
\geq &P'\left\{B_3(n,k) < M_2\right\} + P'\left\{B_4(n,k,\delta) < \frac{\epsilon_2}{M_2}\right\} - 1.
\end{aligned}$$

Then, take $\liminf_{n\to\infty}$ on both sides of the last inequality with $\delta > 0$ held fixed:

$$\begin{aligned}
&\liminf_{n\to\infty} P'\left\{B_3(n,k)B_4(n,k,\delta) < \epsilon_2\right\} \\
\geq &\liminf_{n\to\infty} P'\left\{B_3(n,k) < M_2\right\} + \liminf_{n\to\infty} P'\left\{B_4(n,k,\delta) < \frac{\epsilon_2}{M_1}\right\} - 1 \\
\geq &1 - \frac{\epsilon_1}{2} + 1 - \frac{\epsilon_1}{2} - 1 \qquad \text{[by (A.61) and (A.62)]} \\
= &1 - \epsilon_1.
\end{aligned}$$

Since $\epsilon_1 > 0$ is arbitrary, and $\liminf_{n\to\infty} P'\left\{B_3(n,k)B_4(n,k,\delta) < \epsilon_2\right\} \geq 1 - \epsilon_1$, we have

$$\liminf_{n\to\infty} P'\left\{B_3(n,k)B_4(n,k,\delta) < \epsilon_2\right\} = 1$$

for any fixed $\delta \in (0,1)$, and hence

$$\lim_{\delta\downarrow 0} \liminf_{n\to\infty} P'\left\{B_3(n,k)B_4(n,k,\delta) < \epsilon_2\right\} = 1. \tag{A.63}$$

Now,

$$\begin{aligned}
&P'\left\{A_2(n,k) < 2\epsilon_2\right\} \geq P'\left\{B_1(n,k,\delta)B_2(n,k) < \epsilon_2 \cap B_3(n,k)B_4(n,k,\delta) < \epsilon_2\right\} \\
\geq &P'\left\{B_1(n,k,\delta)B_2(n,k) < \epsilon_2\right\} + P'\left\{B_3(n,k)B_4(n,k,\delta) < \epsilon_2\right\} - 1.
\end{aligned}$$

Take $\liminf_{n\to\infty}$ on both sides of the above inequality:

$$\begin{aligned}
&\liminf_{n\to\infty} P'\left\{A_2(n,k) < 2\epsilon_2\right\} \\
&\geq \liminf_{n\to\infty} P'\left\{B_1(n,k,\delta)B_2(n,k) < \epsilon_2\right\} + \liminf_{n\to\infty} P'\left\{B_3(n,k)B_4(n,k,\delta) < \epsilon_2\right\} - 1.
\end{aligned}$$

Next, take $\lim_{\delta\downarrow 0}$. Since the left hand side of the above inequality does not depend on δ, we have

$$\begin{aligned}
&\liminf_{n\to\infty} P'\left\{A_2(n,k) < 2\epsilon_2\right\} \\
&\geq \lim_{\delta\downarrow 0}\liminf_{n\to\infty} P'\left\{B_1(n,k,\delta)B_2(n,k) < \epsilon_2\right\} + \lim_{\delta\downarrow 0}\liminf_{n\to\infty} P'\left\{B_3(n,k)B_4(n,k,\delta) < \epsilon_2\right\} - 1 = 1
\end{aligned}$$

by (A.60) and (A.63). Therefore,

$$A_2(n,k) = o_{P'}(1).$$

Since we have already shown that $A_1 = o_{P'}(1)$, we have

$$A_1(n,k) + A_2(n,k) = o_{P'}(1).$$

$\square$

4 Proofs of Theorems

Let $X_1, \dots, X_n$ be independent and identically distributed random variables from the distribution function F. Then, $F(X_1), \dots, F(X_n)$ are independent and identically distributed as Uniform(0,1). Therefore, in what follows, we assume without loss of generality that $V_i = F(X_i)$, $i = 1, \dots n$, where $\{V_i\}_{i=1}^n$ are defined at the beginning of the section 1. Let $\{V_i'\}_{i=1}^n$ and $\{V_i^*\}_{i=1}^n$ also be as at the beginning of the section 1. Let

$$X_i^* = F^{\leftarrow}(V_i^*),\ i = 1, \dots n.$$

Since $\{V_1^*, \dots, V_n^*\} \subseteq \{V_1, \dots, V_n\}$, it follows that $\{X_1^*, \dots, X_n^*\} \subseteq \{X_1, \dots, X_n\}$. Let $\{X_{1,n} \leq \dots \leq X_{n,n}\}$ and $\{X_{1,n}^* \leq \dots \leq X_{n,n}^*\}$ denote order statistics of $\{X_1, \dots, X_n\}$ and $\{X_1^*, \dots, X_n^*\}$, respectively. Then, since $U(x) = F^{\leftarrow}(1 - x^{-1})$ by (A.16), $V_{n-[ks],n} = 1 - V_{[ks],n}$ and $V_{n-[ks],n}^* = 1 - V_{[ks],n}^*$ in distribution, we have, in distribution,

$$X_{n-[ks],n} = F^{\leftarrow}(V_{n-[ks],n}) = F^{\leftarrow}(1 - V_{[ks],n}) = U(1/V_{[ks],n}), \tag{A.64}$$

$$X_{n-[ks],n}^* = F^{\leftarrow}(V_{n-[ks],n}^*) = F^{\leftarrow}(1 - V_{[ks],n}^*) = U(1/V_{[ks],n}^*). \tag{A.65}$$

4.1 Proof of Theorem 5

Proof. Let

$$\begin{aligned}\Lambda_n(s;\gamma,\rho,\delta) = \max&\left\{\left(\frac{k}{nV_{[ks],n}^*}\right)^{\gamma+\rho+\delta}, \left(\frac{k}{nV_{[ks],n}^*}\right)^{\gamma+\rho-\delta}\right\} \\ &+ \max\left\{\left(\frac{k}{nV_{[ks],n}}\right)^{\gamma+\rho+\delta}, \left(\frac{k}{nV_{[ks],n}}\right)^{\gamma+\rho-\delta}\right\}\end{aligned} \tag{A.66}$$

Let $\epsilon > 0$ and $\delta > 0$ be as in Lemma A.4. We apply inequality (A.23) with $t = n/k$, $s \in [1/k, 1]$ and $x = k/(nV_{[ks],n})$. Since $\sup_{1/k \leq s \leq 1} V_{[ks],n} \leq V_{k,n} = o_{P'}(1)$, $tx \to \infty$, in probability$[P']$. Hence, for any fixed t_0, $tx > t_0$ with probability approaching 1. For the rest of the proof of this theorem, the inequalities are implicitly assumed to hold with probability approaching 1 as $n \to \infty$, without further comment.

Invoking inequality (A.23) and multiplying by $\{A_0(n/k)\}$, we obtain

$$\begin{aligned}&\left|\frac{U(1/V_{[ks],n}) - U(\frac{n}{k})}{a_0(\frac{n}{k})} - \frac{\big(k/(nV_{[ks],n})\big)^\gamma - 1}{\gamma} - A_0\left(\frac{n}{k}\right)\Psi_{\gamma,\rho}\left(\frac{k}{nV_{[ks],n}}\right)\right| \\ \leq&\varepsilon\left|A_0\left(\frac{n}{k}\right)\right|\left(\frac{k}{nV_{[ks],n}}\right)^{\gamma+\rho}\max\left\{\left(\frac{k}{nV_{[ks],n}}\right)^{\delta}, \left(\frac{k}{nV_{[ks],n}}\right)^{-\delta}\right\}.\end{aligned} \tag{A.67}$$

Similarly, for the bootstrap sample, we apply inequality (A.23) with $t = n/k$ and $x = k/(nV^*_{[ks],n})$. Since $\sup_{1/k\leq s\leq 1} V^*_{[ks],n} \leq V^*_{k,n} = o_{P'}(1)$, we have $tx \to \infty$, in probability $[P']$, as $n \to \infty$. Therefore, for any given t_0, $tx > t_0$ for large n. Hence,

$$\begin{aligned}&\left|\frac{U(1/V^*_{[ks],n}) - U(\frac{n}{k})}{a_0(\frac{n}{k})} - \frac{\left(k/(nV^*_{[ks],n})\right)^\gamma - 1}{\gamma} - A_0\left(\frac{n}{k}\right)\Psi_{\gamma,\rho}\left(\frac{k}{nV^*_{[ks],n}}\right)\right| \\ \leq&\varepsilon\left|A_0\left(\frac{n}{k}\right)\right|\left(\frac{k}{nV^*_{[ks],n}}\right)^{\gamma+\rho}\max\left\{\left(\frac{k}{nV^*_{[ks],n}}\right)^{\delta}, \left(\frac{k}{nV^*_{[ks],n}}\right)^{-\delta}\right\}.\end{aligned} \tag{A.68}$$

Subtracting (A.67) from (A.68), multiplying throughout by $k^{1/2}s^{\gamma+(1/2)+\epsilon}$, and using (A.65), (A.64) and (A.66) we obtain,

$$\begin{aligned}&s^{\gamma+1/2+\epsilon}k^{1/2}\Bigg|\frac{X^*_{n-[ks],n} - X_{n-[ks],n}}{a_0(\frac{n}{k})} - \left(\frac{\left(k/(nV^*_{[ks],n})\right)^\gamma - 1}{\gamma} - \frac{\big(k/(nV_{[ks],n})\big)^\gamma - 1}{\gamma}\right) \\ &- A_0\left(\frac{n}{k}\right)\left\{\Psi_{\gamma,\rho}\left(\frac{k}{nV^*_{[ks],n}}\right) - \Psi_{\gamma,\rho}\left(\frac{k}{nV_{[ks],n}}\right)\right\}\Bigg| \\ \leq&\varepsilon k^{1/2}\left|A_0\left(\frac{n}{k}\right)\right| s^{\gamma+(1/2)+\epsilon}\Lambda_n(s;\gamma,\rho,\delta).\end{aligned} \tag{A.69}$$

We consider $\sup_{1/k\leq s\leq 1}$ and $\sup_{0\leq s\leq k^{-1}}$ of the expression in the Theorem separately. By (A.56),

$$\begin{aligned}&\sup_{1/k\leq s\leq 1} s^{\gamma+(1/2)+\epsilon}\left|k^{1/2}\left(\frac{\left\{k/(nV^*_{[ks],n})\right\}^\gamma - 1}{\gamma} - \frac{\{k/(nV_{[ks],n})\}^\gamma - 1}{\gamma}\right) - s^{-\gamma-1}B'_{k,n}(s)\right| \\ &= o_{P'}(1).\end{aligned} \tag{A.70}$$

By (A.56) with $\gamma = 1$,

$$\sup_{1/k\leq s\leq 1} s^{(3/2)+\epsilon}\left|k^{1/2}\left(k/(nV^*_{[ks],n}) - k/(nV_{[ks],n})\right) - s^{-2}B'_{k,n}(s)\right| = o_{P'}(1).$$

Since $\sup_{1/k \le s \le 1} s^{-(1/2)+\epsilon} B'_{k,n}(s) = O_{P'}(1)$, by (A.26),

$$\sup_{1/k \le s \le 1} s^{(3/2)+\epsilon} k^{1/2} \left| k/(nV^*_{[ks],n}) - k/(nV_{[ks],n}) \right| = O_{P'}(1), \tag{A.71}$$

Next, we show that

$$\sup_{k^{-1} \le s \le 1} s^{\gamma+(1/2)+\varepsilon} [k^{1/2} A_0 \left(\frac{n}{k}\right)] \left\{ \Psi_{\gamma,\rho}\left(\frac{k}{nV^*_{[ks],n}}\right) - \Psi_{\gamma,\rho}\left(\frac{k}{nV_{[ks],n}}\right) \right\} = o_{P'}(k^{-1/2}). \tag{A.72}$$

Let

$$A = \frac{k}{nV^*_{[ks],n}}, \quad B = \frac{k}{nV_{[ks],n}}.$$

By the mean value theorem,

$$\Psi_{\gamma,\rho}(A) - \Psi_{\gamma,\rho}(B) = \dot{\Psi}_{\gamma,\rho}(\zeta_n(s))(A - B)$$

where $\zeta_n(s)$ lies between A and B. By (A.35)-(A.38), we have

$$\sup_{k^{-1} \le s \le 1} s\zeta_n(s) = O_{P'}(1) \tag{A.73}$$

$$\sup_{k^{-1} \le s \le 1} \{s\zeta_n(s)\}^{-1} = O_{P'}(1). \tag{A.74}$$

Next consider, part of the expression in (A.72):

$$\begin{aligned} C_1(s) &:= s^{\gamma+(1/2)+\varepsilon}\{\Psi_{\gamma,\rho}(A) - \Psi_{\gamma,\rho}(B)\} \\ &= s^{\gamma+(1/2)+\varepsilon}\dot{\Psi}_{\gamma,\rho}\{\zeta_n(s)\}(A - B) \\ &= s^{\gamma+(1/2)+\varepsilon}s^{-(3/2)-\varepsilon}\dot{\Psi}_{\gamma,\rho}\{\zeta_n(s)\}\{s^{(3/2)+\varepsilon}(A - B)\} \\ &= s^{\gamma-1}\dot{\Psi}_{\gamma,\rho}\{\zeta_n(s)\}O_{P'}(k^{-1/2}) \end{aligned}$$

Next, we consider the four cases $\gamma + \rho = 0$ and $\rho < 0$, $\gamma + \rho \neq 0$ and $\rho < 0$, $\rho = 0 \neq \gamma$, and $\rho = 0 = \gamma$ separately.

Case 1: $\gamma + \rho \neq 0$ and $\rho < 0$:

$$\Psi_{\gamma,\rho}(x) = (\gamma + \rho)^{-1}(x^{\gamma+\rho} - 1), \quad \dot{\Psi}_{\gamma,\rho}(x) = x^{\gamma+\rho-1}.$$

$$
\begin{aligned}
\sup_{k^{-1}\le s\le 1} C_1(s) &= \sup_{k^{-1}\le s\le 1} s^{\gamma-1}s^{-(\gamma+\rho-1)}\{s\zeta_n(s)\}^{\gamma+\rho-1}O_{P'}(k^{-1/2}) \\
&\le \sup_{k^{-1}\le s\le 1} s^{-\rho} \sup_{k^{-1}\le s\le 1}\{s\zeta_n(s)\}^{\gamma+\rho-1}O_{P'}(k^{-1/2}) = O_{P'}(k^{-1/2}).
\end{aligned}
$$

Case 2: $\gamma+\rho=0$ and $\rho<0$:

$$
\Psi_{\gamma,\rho}(x)=\log(x), \quad \dot{\Psi}_{\gamma,\rho}(x)=x^{-1}.
$$

$$
\begin{aligned}
\sup_{k^{-1}\le s\le 1} C_1(s) &= \sup_{k^{-1}\le s\le 1} s^{\gamma-1}\{\zeta_n(s)\}^{-1}O_{P'}(k^{-1/2}) \\
&\le \sup_{k^{-1}\le s\le 1} s^{\gamma}\{s\zeta_n(s)\}^{-1}O_{P'}(k^{-1/2}) = O_{P'}(k^{-1/2}).
\end{aligned}
$$

Case 3: $\rho=0\neq\gamma$:

$$
\Psi_{\gamma,\rho}(x)=\gamma^{-1}x^{\gamma}\log x, \quad \dot{\Psi}_{\gamma,\rho}(x)=x^{\gamma-1}(\log x+\gamma^{-1}).
$$

In the following arguments, we use (A.73) and (A.74).

$$
\begin{aligned}
\sup_{k^{-1}\le s\le 1} C_1(s) &= \sup_{k^{-1}\le s\le 1} s^{\gamma-1}\zeta_n^{\gamma-1}[\log(\zeta_n)+\gamma^{-1}]O_{P'}(k^{-1/2}) \\
&= \sup_{k^{-1}\le s\le 1} (s\zeta_n)^{\gamma-1}[\log(s\zeta_n)+\gamma^{-1}-\log(s)]O_{P'}(k^{-1/2}) \\
&\le \sup_{k^{-1}\le s\le 1} (s\zeta_n)^{\gamma-1}[\sup_{k^{-1}\le s\le 1}|\log(s\zeta_n)|+|\gamma^{-1}|+|\log(s)|]O_{P'}(k^{-1/2}) \\
&= O_{P'}(1)[O_{P'}(1)+|\gamma^{-1}|+|\log(k)|]O_{P'}(k^{-1/2}) \\
&= O_{P'}(\log(k)k^{-1/2}) \\
&= o_{P'}(1).
\end{aligned}
$$

Case 4: $\rho=0=\gamma$:

$$
\Psi_{\gamma,\rho}(x)=2^{-1}(\log x)^2, \quad \dot{\Psi}_{\gamma,\rho}(x)=x^{-1}\log x.
$$

$$
\begin{aligned}
\sup_{k^{-1}\le s\le 1} C_1(s) &= \sup_{k^{-1}\le s\le 1} s^{\gamma-1}\dot{\Psi}_{\gamma,\rho}\{\zeta_n(s)\}O_{P'}(k^{-1/2}) \\
&= \sup_{k^{-1}\le s\le 1} s^{\gamma-1}\{\Psi_{\gamma,\rho}\{\zeta_n(s)\}\}^{-1}\log\{\Psi_{\gamma,\rho}\{\zeta_n(s)\}\}O_{P'}(k^{-1/2}) \\
&= \sup_{k^{-1}\le s\le 1} \{s\Psi_{\gamma,\rho}\{\zeta_n(s)\}\}^{-1}\log\{s\Psi_{\gamma,\rho}\{\zeta_n(s)\}-\log(s)\}O_{P'}(k^{-1/2}) \\
&\le O_{P'}(1)[O_{P'}(1)+\log(k)]O_{P'}(k^{-1/2}) = O_{P'}(\log(k)k^{-1/2}) \\
&= o_{P'}(1).
\end{aligned}
$$

Now (A.72) follows.

Next, let

$$\sup_{1/k\le s\le 1} s^{\gamma+1/2+\epsilon}\left(\frac{k}{nV_{[ks],n}}\right)^{\gamma+\rho\pm\delta} = \sup_{1/k\le s\le 1}\left(\frac{ks}{nV_{[ks],n}}\right)^{\gamma+\rho\pm\delta} s^{1/2+\epsilon-\rho\mp\delta}.$$

By (A.35) and (A.36), $(ks/(nV_{[ks],n}))^{\gamma+\rho\pm\delta} = O_{P'}(1)$. Choose $\delta > 0$ such as $\delta < 1/2+\epsilon-\rho$, then

$$\sup_{1/k\le s\le 1} s^{1/2+\epsilon-\rho\mp\delta} = 1.$$

Therefore,

$$\sup_{1/k\le s\le 1} s^{\gamma+1/2+\epsilon}\left(\frac{k}{nV_{[ks],n}}\right)^{\gamma+\rho\pm\delta} = O_{P'}(1). \tag{A.75}$$

Similarly, by (A.37) and (A.38) for $0 < \delta < 1/2+\epsilon-\rho$

$$\sup_{1/k\le s\le 1} s^{\gamma+1/2+\epsilon}\left(\frac{k}{nV^*_{[ks],n}}\right)^{\gamma+\rho\pm\delta} = \sup_{1/k\le s\le 1}\left(\frac{k}{nV^*_{[ks],n}}\right)^{\gamma+\rho\pm\delta} s^{1/2+\epsilon-\rho\mp\delta} = O_{P'}(1). \tag{A.76}$$

By (A.75) and (A.76),

$$\sup_{1/k\le s\le 1} s^{\gamma+1/2+\epsilon}\Lambda_n(s;\gamma,\rho,\delta) = O_{P'}(1).$$

Since $k^{1/2}A(n/k) = o(1)$,

$$\sup_{1/k\le s\le 1} \varepsilon k^{1/2}\left|A_0\left(\frac{n}{k}\right)\right| s^{\gamma+1/2+\epsilon}\Lambda_n(s;\gamma,\rho,\delta) = o_{P'}(1). \tag{A.77}$$

By substituting (A.70) and (A.77) in (A.69), we obtain

$$\sup_{1/k\le s\le 1} s^{\gamma+1/2+\epsilon}\left|k^{1/2}\frac{X^*_{n-[ks],n} - X_{n-[ks],n}}{a_0(\frac{n}{k})} - \frac{B'_{k,n}(s)}{s^{\gamma+1}}\right| = o_{P'}(1). \tag{A.78}$$

This completes the first part of the proof for the interval $[k^{-1},1]$. Next consider the interval $[0,k^{-1}]$. We consider the cases $\gamma \ge -1/2$ and $\gamma \le -1/2$ separately. First, consider the case when $\gamma \geq -1/2$. Since $V_{[ks],n} = V_{1,n}$ and $V^*_{[ks],n} = V^*_{1,n}$ for $s \leq 1/k$,

$$
\begin{aligned}
& \sup_{0\leq s\leq 1/k} s^{\gamma+1/2+\epsilon}\left|k^{1/2}\frac{X^*_{n-[ks],n} - X_{n-[ks],n}}{a_0(\frac{n}{k})} - \frac{B'_{k,n}(s)}{s^{\gamma+1}}\right| \\
= & \sup_{0\leq s\leq 1/k} s^{\gamma+1/2+\epsilon}\left|k^{1/2}\frac{U(1/V^*_{[ks],n}) - U(1/V_{[ks],n})}{a_0(\frac{n}{k})} - \frac{B'_{k,n}(s)}{s^{\gamma+1}}\right| \\
= & \sup_{0\leq s\leq 1/k} s^{\gamma+1/2+\epsilon}\left|k^{1/2}\frac{U(1/V^*_{1,n}) - U(1/V_{1,n})}{a_0(\frac{n}{k})} - \frac{B'_{k,n}(s)}{s^{\gamma+1}}\right| \\
\leq & \sup_{0\leq s\leq 1/k} s^{\gamma+1/2+\epsilon}\left|k^{1/2}\frac{X^*_{n-1,n} - X_{n-1,n}}{a_0(\frac{n}{k})} - \frac{B'_{k,n}(k^{-1})}{k^{-(\gamma+1)}}\right| + \sup_{0\leq s\leq 1/k}\frac{\left|B'_{k,n}(s)\right|}{s^{\gamma+1}} \\
\leq & \; k^{-(\gamma+(1/2)+\epsilon)}\left|k^{1/2}\frac{X^*_{n-1,n} - X_{n-1,n}}{a_0(\frac{n}{k})} - \frac{B'_{k,n}(k^{-1})}{k^{-(\gamma+1}}\right| \\
& + \left|\frac{B'_{k,n}(k^{-1})}{k^{-(\gamma+1}}\right| + \sup_{0\leq s\leq 1/k} s^{\gamma+(1/2)+\varepsilon}\frac{\left|B'_{k,n}(s)\right|}{s^{1/2-\epsilon}} \\
= & \; o_{P'}(1)
\end{aligned}
$$

by (A.78) and and (A.27).

Next, consider the case when $\gamma < -1/2$. In the following we show that

$$
\sup_{1/k\leq s\leq 1} s^{\gamma+1/2+\epsilon}\left|k^{1/2}\frac{U(1/V^*_{[ks],n}) - U(1/V_{[ks],n}) - (U(1/V^*_{1,n}) - U(1/V_{1,n}))}{a_0(\frac{n}{k})} - \frac{B'_{k,n}(s)}{s^{\gamma+1}}\right| = o_{P'}(1). \tag{A.79}
$$

$$
\begin{aligned}
& \sup_{1/k\leq s\leq 1} s^{\gamma+1/2+\epsilon}\left|k^{1/2}\frac{U(1/V^*_{[ks],n}) - U(1/V_{[ks],n}) - (U(1/V^*_{1,n}) - U(1/V_{1,n}))}{a_0(\frac{n}{k})} - \frac{B'_{k,n}(s)}{s^{\gamma+1}}\right| \\
\leq & \sup_{1/k\leq s\leq 1} s^{\gamma+1/2+\epsilon}\left|k^{1/2}\frac{U(1/V^*_{[ks],n}) - U(1/V_{[ks],n})}{a_0(\frac{n}{k})} - \frac{B'_{k,n}(s)}{s^{\gamma+1}}\right| \\
& + \sup_{1/k\leq s\leq 1} s^{\gamma+1/2+\epsilon}\left|k^{1/2}\frac{U(1/V^*_{1,n}) - U(1/V_{1,n})}{a_0(\frac{n}{k})}\right|.
\end{aligned}
$$

The first term is $o_{P'}(1)$ by Theorem 5. The second term

$$
\begin{aligned}
& \sup_{1/k\leq s\leq 1} s^{\gamma+1/2+\epsilon}\left|k^{1/2}\frac{U(1/V^*_{1,n}) - U(1/V_{1,n})}{a_0(\frac{n}{k})}\right| \\
\leq & \; k^{-\gamma-1/2-\epsilon}\left|k^{1/2}\frac{U(1/V^*_{1,n}) - U(1/V_{1,n})}{a_0(\frac{n}{k})}\right| \\
\leq & \; k^{-\gamma-1/2-\epsilon}\left|k^{1/2}\frac{U(1/V^*_{1,n}) - U(1/V_{1,n})}{a_0(\frac{n}{k})} - \frac{B'_{k,n}(k^{-1})}{k^{-(\gamma+1)}}\right| + \frac{\left|B'_{k,n}(k^{-1})\right|}{k^{1/2-\epsilon}} = o_{P'}(1)
\end{aligned}
$$

by Theorem 5 with $s = 1/k$ and (A.27). Therefore, (A.79) holds. Now,

$$\begin{aligned}
&\sup_{0\le s\le 1/k} s^{\gamma+1/2+\epsilon}\left|k^{1/2}\frac{U(1/V^*_{[ks],n})-U(1/V_{[ks],n})-(U(1/V^*_{1,n})-U(1/V_{1,n}))}{a_0(\frac{n}{k})}-\frac{B'_{k,n}(s)}{s^{\gamma+1}}\right|\\
=&\sup_{0\le s\le 1/k} s^{\gamma+1/2+\epsilon}\left|k^{1/2}\frac{U(1/V^*_{1,n})-U(1/V_{1,n})-(U(1/V^*_{1,n})-U(1/V_{1,n}))}{a_0(\frac{n}{k})}-\frac{B'_{k,n}(s)}{s^{\gamma+1}}\right|\\
=&\sup_{0\le s\le 1/k}\frac{\left|B'_{k,n}(s)\right|}{s^{1/2-\epsilon}}=o_{P'}(1)
\end{aligned}$$

by (A.27).

□

Proof. (Theorem 6). By Theorem 5 with $s = 1$,

$$\left|k^{1/2}\frac{X^*_{n-k,n}-X_{n-k,n}}{a_0(\frac{n}{k})}-B'_{k,n}(1)\right|=o_{P'}(1).$$

Then,

$$k^{1/2}\frac{X^*_{n-k,n}-X_{n-k,n}}{a_0(\frac{n}{k})}=B'_{k,n}(1)+o_{P'}(1)=(n/k)^{1/2}B'_n(k/n)+o_{P'}(1) \qquad \text{(A.80)}$$

and $(n/k)^{1/2}B'_n(k/n)$ is distributed as $N(0, 1-(k/n))$. Therefore, the expression on the left of (A.80) converges in distribution to $N(0, 1)$, which completes the proof of the theorem. □

4.2 Proof for the Hill estimator, Theorem 7

The proof for Theorem 7 can be deduced from the general Theorem 5. Here we provide a more direct proof. Since the Hill estimator appears frequently in different contexts in statistical inference, it would be helpful and instructive to derive the results without using a powerful general result.

Let

$$\begin{aligned}
\Delta_n(s;\rho,\delta)=&\left(\frac{k}{nV_{[ks],n}}\right)^{\rho}\max\left\{\left(\frac{k}{nV_{[ks],n}}\right)^{\delta},\left(\frac{k}{nV_{[ks],n}}\right)^{-\delta}\right\}\\
&+\left(\frac{k}{nV^*_{[ks],n}}\right)^{\rho}\max\left\{\left(\frac{k}{nV^*_{[ks],n}}\right)^{\delta},\left(\frac{k}{nV^*_{[ks],n}}\right)^{-\delta}\right\}.
\end{aligned}$$

Lemma A.10. *Let $\gamma > 0$ and $\rho \le 0$. Suppose that condition (A.18) is satisfied and $k^{1/2}A(n/k) = o(1)$. Let $\varepsilon > 0$ and $\delta > 0$ be given. Then*

$$\sup_{0\le s\le 1} s^{1/2+\epsilon}\left|k^{1/2}\left(\log X^*_{n-[ks],n} - \log X_{n-[ks],n}\right) - \gamma\frac{B'_{k,n}(s)}{s}\right| = o_{P'}(1). \tag{A.81}$$

Proof. First, we prove that

$$\sup_{1/k\le s\le 1} s^{1/2+\epsilon}\left|k^{1/2}\left(\log X^*_{n-[ks],n} - \log X_{n-[ks],n}\right) - \gamma\frac{B'_{k,n}(s)}{s}\right| = o_{P'}(1). \tag{A.82}$$

We wish to invoke inequality (A.25). Let $t = t(n,k) = n/k$, $s \in [1/k, 1]$ and $x = x(n,k,s) = k/(nV_{[ks],n})$. Then

$$1/V_{[ks],n} = t(n,k)x(n,k,s) \ge t(n,k)x(n,k,1) = 1/V_{k,n} \quad (1/k \le s \le 1).$$

Let $\varepsilon_1 > 0$ and $t_0 > 0$ be a given. Let A denote the event "$1/V_{k,n} > t_0$". Since $1/V_{k,n} \to \infty$ in probability, there exists an n_0 such that $P'[A] > 1 - \varepsilon_1$ for $n \ge n_0$. Suppose that A occurred and that $n \ge n_0$. Then $t(n,k)x(n,k,s) > t_0$ and hence by the inequality (A.25) the following claims, between the lines, hold with probability at least $(1-\varepsilon_1)$ for any $s \in [1/k, 1]$:

$$\left|\frac{\log U\left\{k/(nV_{[ks],n})\right\} - \log U\left(n/k\right) - \gamma\log\left\{k/(nV_{[ks],n})\right\}}{A_0\left(n/k\right)} - \frac{\left\{k/(nV_{[ks],n})\right\}^\rho - 1}{\rho}\right|$$
$$\le \varepsilon\left(\frac{k}{nV_{[ks],n}}\right)^\rho \max\left\{\left(\frac{k}{nV_{[ks],n}}\right)^\delta, \left(\frac{k}{nV_{[ks],n}}\right)^{-\delta}\right\},$$

which is equivalent to

$$\left|\left\{\log U\left(\frac{k}{nV_{[ks],n}}\right) - \log U\left(\frac{n}{k}\right)\right\} - \gamma\log\left(\frac{k}{nV_{[ks],n}}\right) - A_0\left(\frac{n}{k}\right)\frac{\left\{k/(nV_{[ks],n})\right\}^\rho - 1}{\rho}\right|$$
$$\le \varepsilon\left|A_0\left(\frac{n}{k}\right)\right|\left(\frac{k}{nV_{[ks],n}}\right)^\rho \max\left\{\left(\frac{k}{nV_{[ks],n}}\right)^\delta, \left(\frac{k}{nV_{[ks],n}}\right)^{-\delta}\right\}. \tag{A.83}$$

Similarly, for the bootstrap sample, apply inequality (A.25) with $t = n/k$, $s \in [1/k, 1]$ and $x_* = k/(nV^*_{[ks],n})$. Since $V^*_{[ks],n} \le V^*_{k,n} \to 0$ as $n \to \infty$ in probability, $\inf_{0<s\le 1} tx_* = 1/(V^*_{[k],n}) \to \infty$ in probability, therefore, $\inf_{1/k\le s\le 1} tx_* > t_0$ with probability approaching 1. Then we have,

$$\left|\left\{\log U\left(\frac{k}{nV^*_{[ks],n}}\right)-\log U\left(\frac{n}{k}\right)\right\}-\gamma\log\left(\frac{k}{nV^*_{[ks],n}}\right)-A_0\left(\frac{n}{k}\right)\frac{\left\{k/(nV^*_{[ks],n})\right\}^\rho-1}{\rho}\right|$$
$$\le\varepsilon\left|A_0\left(\frac{n}{k}\right)\right|\left(\frac{k}{nV^*_{[ks],n}}\right)^\rho\max\left\{\left(\frac{k}{nV^*_{[ks],n}}\right)^\delta,\left(\frac{k}{nV^*_{[ks],n}}\right)^{-\delta}\right\} \tag{A.84}$$

Subtracting (A.83) from (A.84), we get,

$$\Bigg|\left(\log X^*_{n-[ks],n}-\log X_{n-[ks],n}\right)-\gamma\left\{\log\left(\frac{k}{nV^*_{[ks],n}}\right)-\log\left(\frac{k}{nV_{[ks],n}}\right)\right\}$$
$$-A_0\left(\frac{n}{k}\right)\left(\frac{\left\{k/(nV^*_{[ks],n})\right\}^\rho-1}{\rho}-\frac{\left\{k/(nV_{[ks],n})\right\}^\rho-1}{\rho}\right)\Bigg|\le\varepsilon\left|A_0\left(\frac{n}{k}\right)\right|\Delta_n(s;\rho,\delta)$$

and

$$s^{1/2+\epsilon}k^{1/2}\Bigg|\left(\log X^*_{n-[ks],n}-\log X_{n-[ks],n}\right)-\gamma\left\{\log\left(\frac{k}{nV^*_{[ks],n}}\right)-\log\left(\frac{k}{nV_{[ks],n}}\right)\right\}$$
$$-A_0\left(\frac{n}{k}\right)\left(\frac{\left\{k/(nV^*_{[ks],n})\right\}^\rho-1}{\rho}-\frac{\left\{k/(nV_{[ks],n})\right\}^\rho-1}{\rho}\right)\Bigg|$$
$$\le\varepsilon s^{1/2+\epsilon}k^{1/2}\left|A_0\left(\frac{n}{k}\right)\right|\Delta_n(s;\rho,\delta). \tag{A.85}$$

By (A.31),

$$\sup_{1/k\le s\le 1}s^{1/2+\epsilon}\left|k^{1/2}\left\{\log\left(\frac{k}{nV^*_{[ks],n}}\right)-\log\left(\frac{k}{nV_{[ks],n}}\right)\right\}-\frac{B'_{k,n}(s)}{s}\right|=o_{P'}(1). \tag{A.86}$$

By (A.56),

$$\sup_{1/k\le s\le 1}s^{\rho+1/2+\epsilon}\left|k^{1/2}\left(\frac{\left\{k/(nV^*_{[ks],n})\right\}^\rho-1}{\rho}-\frac{\left\{k/(nV_{[ks],n})\right\}^\rho-1}{\rho}\right)-s^{-\rho-1}B'_{k,n}(s)\right|=o_{P'}(1).$$

Since $\sup_{1/k\le s\le 1}s^{-1/2+\epsilon}B'_{k,n}(s)=O_{P'}(1)$ by (A.26),

$$\sup_{1/k\le s\le 1}s^{\rho+1/2+\epsilon}\left|\frac{\left\{k/(nV^*_{[ks],n})\right\}^\rho-1}{\rho}-\frac{\left\{k/(nV_{[ks],n})\right\}^\rho-1}{\rho}\right|=O_{P'}(k^{-1/2})=o_{P'}(1).$$

Since $\rho \le 0$ and $k^{1/2}A(n/k) = o(1)$,

$$\sup_{1/k\le s\le 1} k^{1/2} A_0\left(\frac{n}{k}\right) s^{1/2+\epsilon} \left| \frac{\left\{k/(nV^*_{[ks],n})\right\}^\rho - 1}{\rho} - \frac{\left\{k/(nV_{[ks],n})\right\}^\rho - 1}{\rho} \right| = o_{P'}(k^{-1/2}). \tag{A.87}$$

By (A.35)–(A.38), for $1/k \le s \le 1$,

$$\left(\frac{k}{nV_{[ks],n}}\right)^{\rho\pm\delta} = O_{P'}(1) s^{-\rho\mp\delta}, \qquad \left(\frac{k}{nV^*_{[ks],n}}\right)^{\rho\pm\delta} = O_{P'}(1) s^{-\rho\mp\delta}.$$

Then

$$\sup_{1/k\le s\le 1} s^{1/2+\epsilon}\Delta_n(s;\rho,\delta) = O_{P'}(1) \sup_{1/k\le s\le 1} s^{1/2+\epsilon-\rho\mp\delta} = O_{P'}(1)$$

by choosing $\delta > 0$ such that $\delta < 1/2 + \epsilon - \rho$. Since $k^{1/2}A(n/k) = o(1)$,

$$\sup_{1/k\le s\le 1} \varepsilon k^{1/2} \left| A_0\left(\frac{n}{k}\right)\right| s^{1/2+\epsilon}\Delta_n(s;\rho,\delta) = o_{P'}(1). \tag{A.88}$$

By substituting (A.86), (A.87) and (A.88) in (A.85), we obtain (A.82).

Next, we show that

$$\sup_{0\le s\le 1/k} s^{1/2+\epsilon} \left| k^{1/2}\left(\log X^*_{n-[ks],n} - \log X_{n-[ks],n}\right) - \gamma \frac{B'_{k,n}(s)}{s} \right| = o_{P'}(1). \tag{A.89}$$

Since $V_{[ks],n} = V_{1,n}$ and $V^*_{[ks],n} = V^*_{1,n}$ for $s \le 1/k$,

$$\begin{aligned}
&\sup_{0\le s\le 1/k} s^{1/2+\epsilon} \left| k^{1/2}\left(\log X^*_{n-[ks],n} - \log X_{n-[ks],n}\right) - \gamma \frac{B'_{k,n}(s)}{s} \right| \\
= &\sup_{0\le s\le 1/k} s^{1/2+\epsilon} \left| k^{1/2}\left(\log U(1/V^*_{[ks],n}) - \log U(1/V_{[ks],n})\right) - \gamma \frac{B'_{k,n}(s)}{s} \right| \\
= &\sup_{0\le s\le 1/k} s^{1/2+\epsilon} \left| k^{1/2}\left(\log U(1/V^*_{1,n}) - \log U(1/V_{1,n})\right) - \gamma \frac{B'_{k,n}(s)}{s} \right| \\
\le &\sup_{0\le s\le 1/k} s^{1/2+\epsilon} \left| k^{1/2}\left(\log U(1/V^*_{1,n}) - \log U(1/V_{1,n})\right) - \gamma \frac{B'_{k,n}(k^{-1})}{k^{-1}} \right| + \gamma \sup_{0\le s\le 1/k} \frac{\left|B'_{k,n}(s)\right|}{s^{1/2-\epsilon}} \\
= &\; k^{-1/2-\epsilon} \left| k^{1/2}\left(\log U(1/V^*_{1,n}) - \log U(1/V_{1,n})\right) - \gamma \frac{B'_{k,n}(k^{-1})}{k^{-1}} \right| + \gamma \sup_{0\le s\le 1/k} \frac{\left|B'_{k,n}(s)\right|}{s^{1/2-\epsilon}} = o_{P'}(1)
\end{aligned}$$

by (A.82) with $s = 1/k$ and (A.27).

□

Proof. (Theorem 7) By definition,

$$\hat{\gamma} = \int_0^1 \left(\log X_{n-[ks],n} - \log X_{n-k,n}\right) ds, \quad \hat{\gamma}^* = \int_0^1 \left(\log X^*_{n-[ks],n} - \log X^*_{n-k,n}\right) ds.$$

By Lemma A.10

$$k^{1/2} \left(\log X^*_{n-[ks],n} - \log X_{n-[ks],n}\right) = \gamma \frac{B'_{k,n}(s)}{s} + o_{P'}(1) s^{-1/2-\epsilon}$$

for $0 \leq s \leq 1$, therefore

$$\begin{aligned} k^{1/2} \left(\hat{\gamma}^* - \hat{\gamma}\right) &= \int_0^1 \left(k^{1/2}(\log X^*_{n-[ks],n} - \log X_{n-[ks],n}) - k^{1/2}(\log X^*_{n-k,n} - \log X_{n-k,n})\right) ds \\ &= \gamma \int_0^1 \left(s^{-1} B'_{k,n}(s) - B'_{k,n}(1)\right) ds + o_{P'}(1) \int_0^1 s^{-1/2-\epsilon} ds \\ &= \gamma \int_0^1 \left(s^{-1} B'_{k,n}(s) - B'_{k,n}(1)\right) ds + o_{P'}(1) \end{aligned}$$

since $\int_0^1 s^{-1/2-\epsilon} ds < \infty$.

Let $W(s)$, $0 \leq s \leq 1$, denote a standard Wiener process. Then,

$$B'_{k,n}(s) = \left(\frac{n}{k}\right)^{1/2} B'_n \left(\frac{ks}{n}\right) = W(s) - \frac{ks}{n} W\left(\frac{n}{k}\right) \quad (0 \leq s \leq 1)$$

in distribution [p.253 in Csörgő and Horváth (1987)], and

$$\frac{B'_{k,n}(s)}{s} - B'_{k,n}(1) = \frac{W(s)}{s} - \frac{k}{n} W\left(\frac{n}{k}\right) - \left(W(1) - \frac{k}{n} W\left(\frac{n}{k}\right)\right) = \frac{W(s)}{s} - W(1) \quad (0 \leq s \leq 1).$$

Therefore, the asymptotic distribution of $k^{1/2} \left(\hat{\gamma}^* - \hat{\gamma}\right)$ is $\gamma \int_0^1 \left(W(s)/s - W(1)\right) ds$, which is a centered normal with variance γ^2 [p.22 in Peng and Qi (2017)].

□

4.3 Proof for bootstrap consistency of high quantile

Proof. (Theorem 8). By definition of $\hat{x}_{p_n}$ and $\hat{x}^*_{p_n}$,

$$\frac{k^{1/2}}{\log d_n}\log\frac{\hat{x}^*_{p_n}}{\hat{x}_{p_n}} = \frac{k^{1/2}}{\log d_n}\log\frac{X^*_{n-k,n}d_n^{\hat{\gamma}^*}}{X_{n-k,n}d_n^{\hat{\gamma}}} = \frac{k^{1/2}}{\log d_n}\log\frac{X^*_{n-k,n}}{X_{n-k,n}} + \frac{k^{1/2}}{\log d_n}\log\frac{d_n^{\hat{\gamma}^*}}{d_n^{\hat{\gamma}}}.$$

The first term is $o_{P'}(1)$ as shown below. By lemma A.10,

$$k^{1/2}\left|\log X^*_{n-k,n} - \log X_{n-k,n} - \gamma W_n(1)\right| = o_{P'}(1),$$

and hence

$$k^{1/2}\left|\log X^*_{n-k,n} - \log X_{n-k,n}\right| = O_{P'}(1).$$

Since $\log d_n \to \infty$ by condition, $k^{1/2}(\log d_n)^{-1}\log(X^*_{n-k,n}/X_{n-k,n}) = o_{P'}(1)$.

The second term

$$\frac{k^{1/2}}{\log d_n}\log\frac{d_n^{\hat{\gamma}^*}}{d_n^{\hat{\gamma}}} = \frac{k^{1/2}}{\log d_n}(\hat{\gamma}^*\log d_n - \hat{\gamma}\log d_n) = k^{1/2}(\hat{\gamma}^* - \hat{\gamma}) \to N(0,\gamma^2)$$

in distribution by theorem 7.

□

Proof. (Proposition 1): Let us recall from section 1 that $\{V_i\}$ and $\{V_i'\}$ are two independent sequences of uniform random variables generating the original random variables $\{X_i\}$ and the bootstrap random variables $\{X_i^*\}$ respectively. To indicate the dependence on $\{V_i\}$ and $\{V_i'\}$ let us use the symbols v and v' respectively. Let $\zeta_{n,\alpha}$ denote the α-quantile of the distribution $P'\left(\log(\hat{x}^*_{p_n}(v,v')/\hat{x}_{p_n}(v)) \le x\right)$. Then

$$P'[\log\hat{x}^*_{p_n}(v,v') - \log\hat{x}_{p_n}(v) \le \zeta_{n,\alpha}] = 1-\alpha.$$

Recall that $\hat{x}_{p_n,\alpha}(v)$ satisfies $P^*\{\hat{x}^*_{p_n}(v,v') \le \hat{x}_{p_n,\alpha}(v) \mid v\} = 1-\alpha$, where P^* denotes the bootstrap probability conditional on $\{X_1,\ldots,X_n\}$. Hence

$$P'\{\hat{x}^*_{p_n}(v,v') \le \hat{x}_{p_n,\alpha}(v)\} = 1-\alpha.$$

Therefore,

$$\log\hat{x}_{p_n,\alpha}(v) = \log\hat{x}_{p_n}(v) + \zeta_{n,\alpha}, \tag{A.90}$$

Let $N_{(0,\gamma)}$ be a random variable having normal distribution with zero mean and variance γ. Let $N_{(0,1)}$ be random variable with the standard normal distribution. By Theorem 4,

$$P\left(\log \frac{\hat{x}_{p_n}(v)}{x_{p_n}} \leq x\right) \to P\left(N_{(0,\gamma)} \leq x\right) = \Psi(x). \tag{A.91}$$

By Theorem 8,

$$P'\left(\log \frac{\hat{x}^*_{p_n}(v,v')}{\hat{x}_{p_n}(v)} \leq x\right) \to P\left(N_{(0,\gamma)} \leq x\right) = \Psi(x). \tag{A.92}$$

Choose a subsequence along which convergence in (A.92) is almost sure. By Lemma 21.2 in van der Vaart (1998), weak convergence of a sequence of distribution functions implies convergence of the corresponding quantile functions at every continuity point. Applying this result to (A.92) at points $\alpha/2$ and $1-\alpha/2$, we obtain, respectively,

$$\zeta_{n,\alpha/2} \to \Psi^{\leftarrow}(\alpha/2) = -z_{\alpha/2}\gamma, \quad \zeta_{n,1-\alpha/2} \to \Psi^{\leftarrow}(1-\alpha/2) = z_{\alpha/2}\gamma. \tag{A.93}$$

Then,

$$\begin{aligned}
&P\left(\hat{x}_{p_n,\alpha/2} \leq x_{p_n} \leq \hat{x}_{p_n,1-\alpha/2}\right) \\
=&P\left(\log \hat{x}_{p_n,\alpha/2} \leq \log x_{p_n} \leq \log \hat{x}_{p_n,1-\alpha/2}\right) \\
=&P\left(\log \hat{x}_{p_n} + \zeta_{n,\alpha/2} \leq \log x_{p_n} \leq \log \hat{x}_{p_n} + \zeta_{n,1-\alpha/2}\right) \quad \text{[by (A.90)]} \\
=&P\left(\zeta_{n,\alpha/2} \leq \log \frac{x_{p_n}}{\hat{x}_{p_n}} \leq \zeta_{n,1-\alpha/2}\right) \\
\to& pr\left(-z_{\alpha/2}\gamma \leq -N_{(0,\gamma^2)} \leq z_{\alpha/2}\gamma\right) \quad \text{[by (A.91), (A.93)]} \\
=&pr\left(-z_{\alpha/2} \leq -N_{(0,1)} \leq z_{\alpha/2}\right) = 1-\alpha.
\end{aligned}$$

□

5 Additional Simulation Results

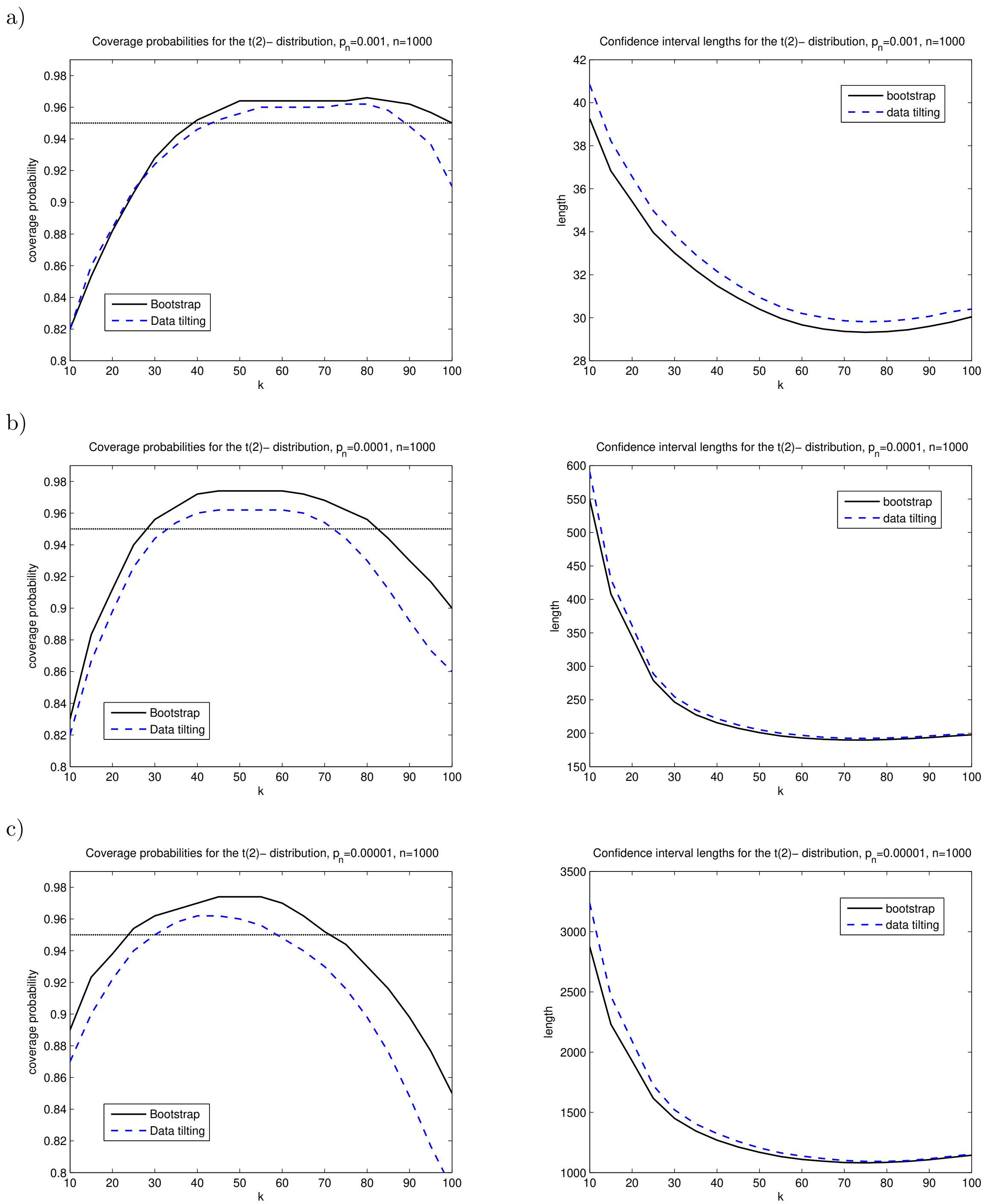

Figure 1: Coverage probabilities and average lengths of the bootstrap and data tilting 95% confidence intervals for the t(2)-distribution; $n = 1000$. The coverage probabilities and average lengths of the confidence intervals for quantiles corresponding to the tail probability $1 - p_n$, where: a) $p_n = 0.001$, b) $p_n = 0.0001$ and c) $p_n = 0.00001$ are plotted as a function of k.

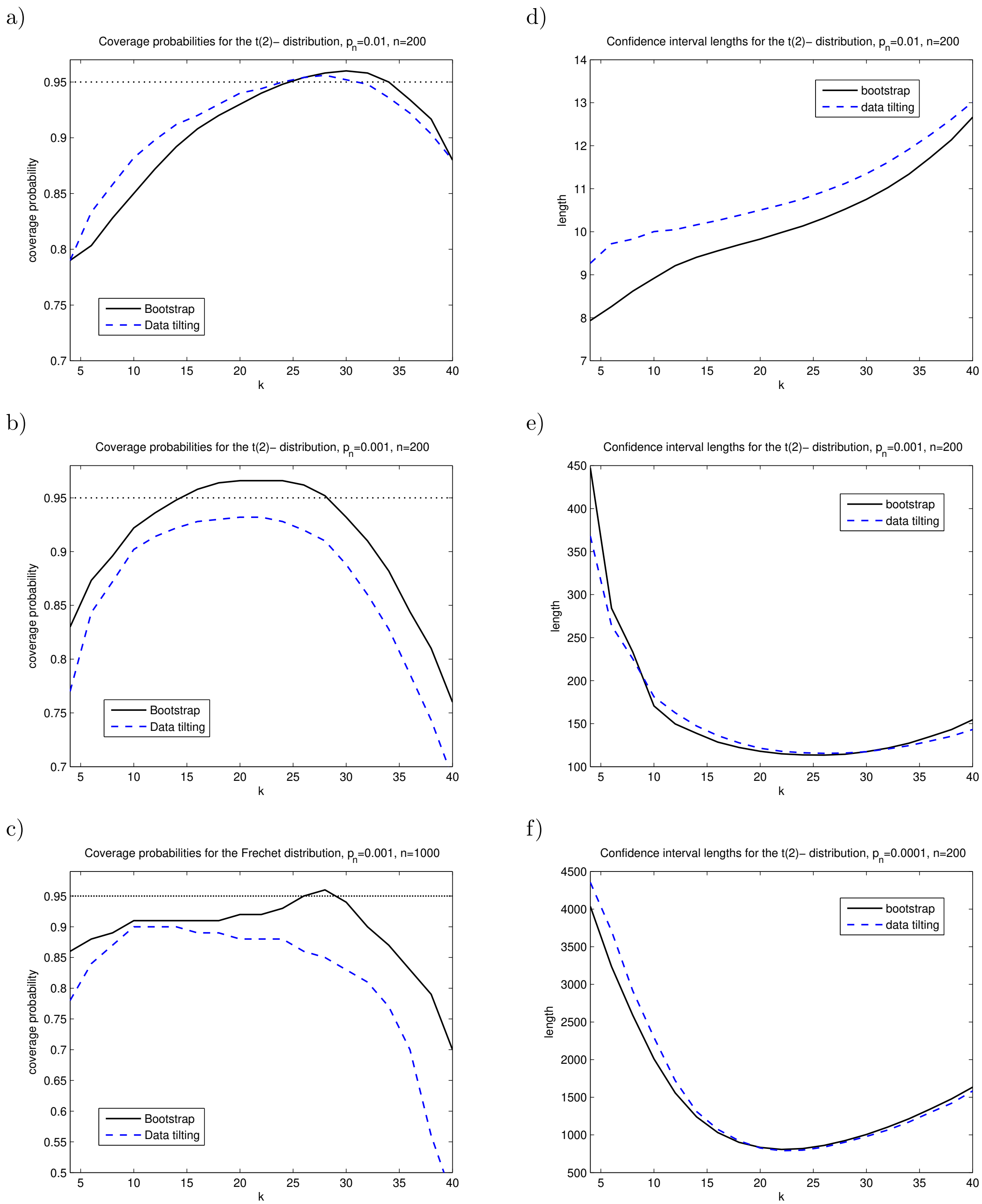

Figure 2: Coverage probabilities and average lengths of the bootstrap and data tilting 95% confidence intervals for the t(2)-distribution; $n = 200$. The coverage probabilities and average lengths of the confidence intervals for quantiles corresponding to the tail probability $1 - p_n$, where: a) $p_n = 0.01$, b) $p_n = 0.001$ and c) $p_n = 0.0001$ are plotted as a function of k.

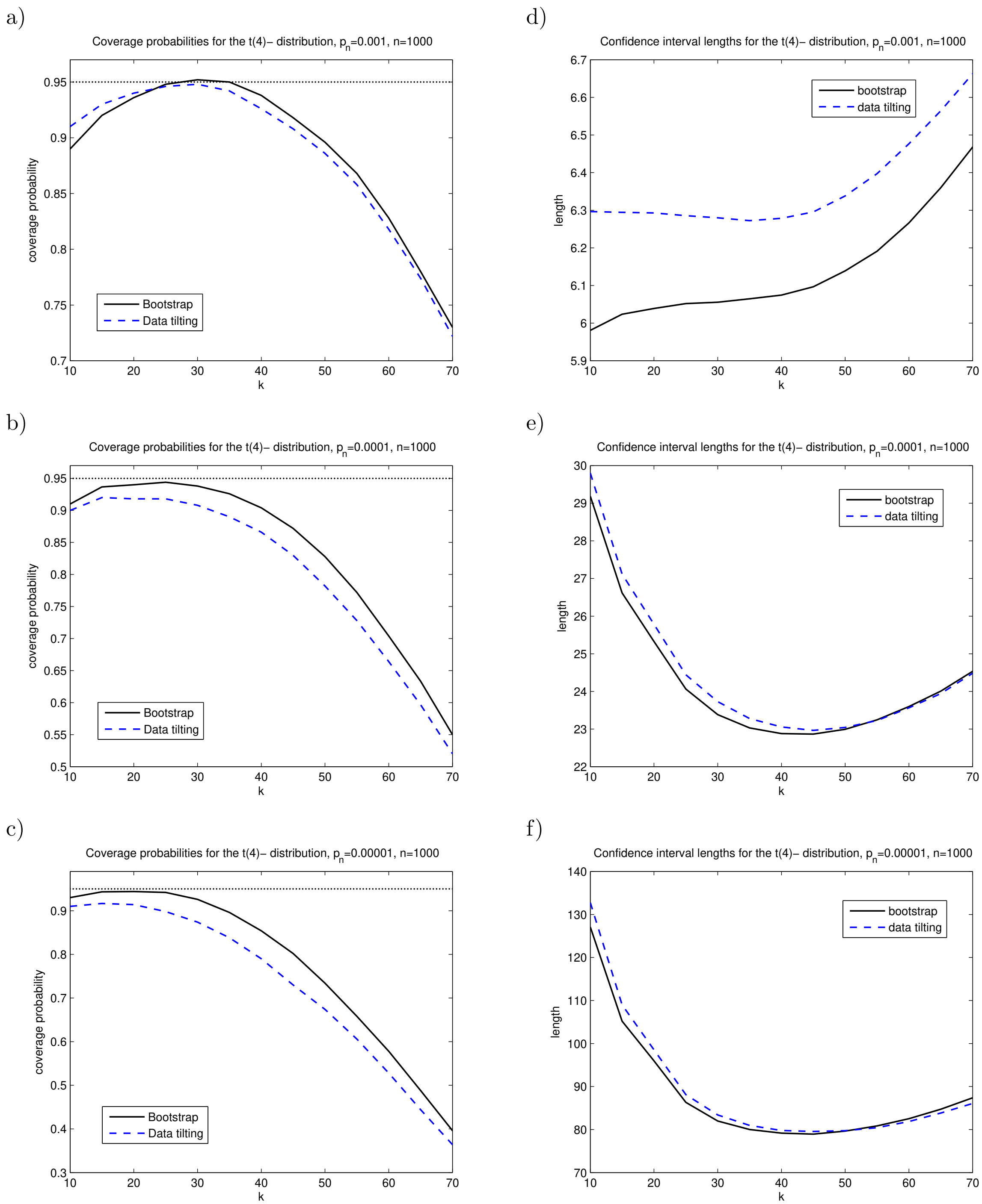

Figure 3: Coverage probabilities and average lengths of the bootstrap and data tilting 95% confidence intervals for the t(4)-distribution; $n = 1000$. The coverage probabilities and average lengths of the confidence intervals for quantiles corresponding to the tail probability $1 - p_n$, where: a) $p_n = 0.001$, b) $p_n = 0.0001$ and c) $p_n = 0.00001$ are plotted as a function of k.

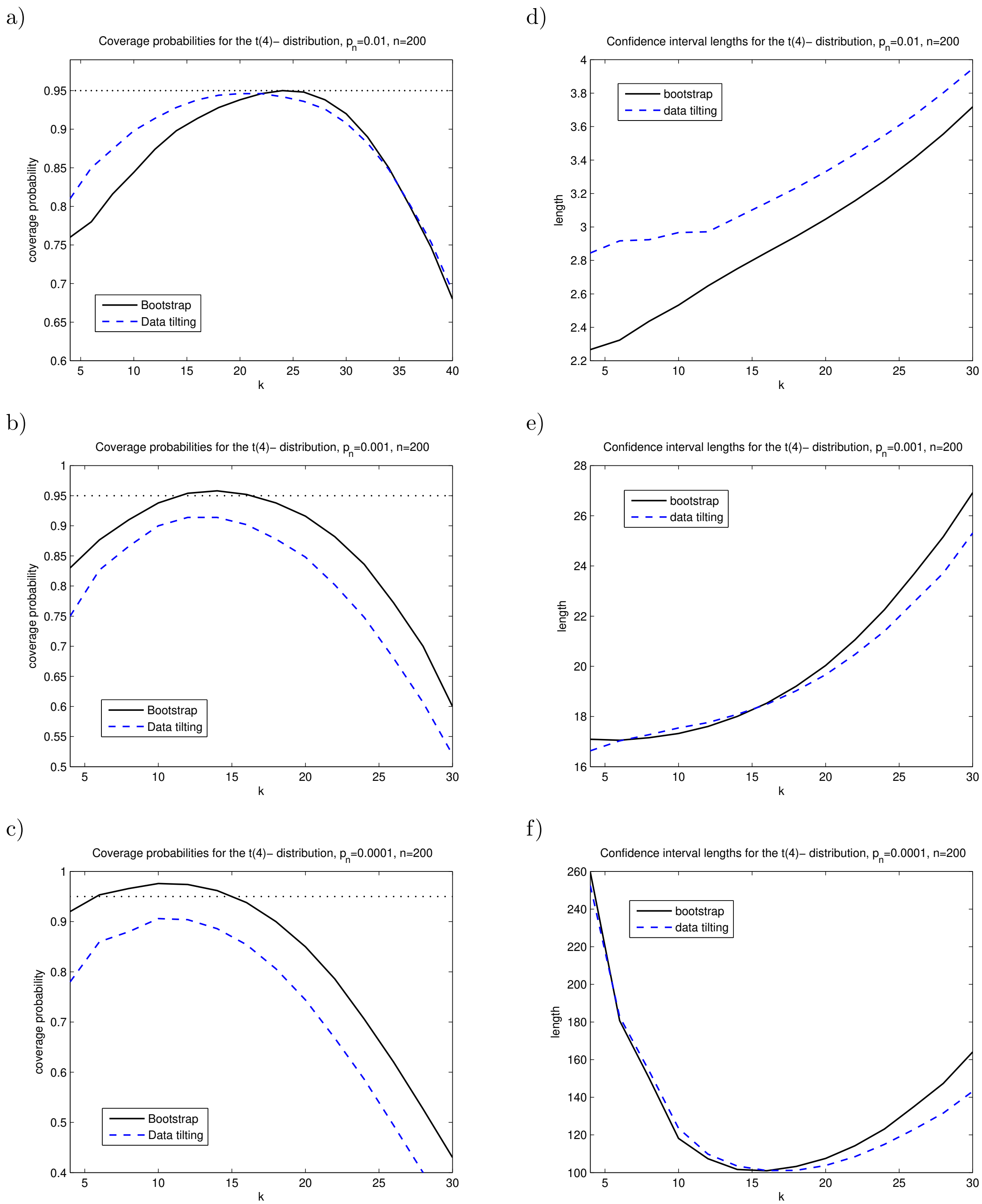

Figure 4: Coverage probabilities and average lengths of the bootstrap and data tilting 95% confidence intervals for the t(4)-distribution; $n = 200$. The coverage probabilities and average lengths of the confidence intervals for quantiles corresponding to the tail probability $1 - p_n$, where: a) $p_n = 0.01$, b) $p_n = 0.001$ and c) $p_n = 0.0001$ are plotted as a function of k.

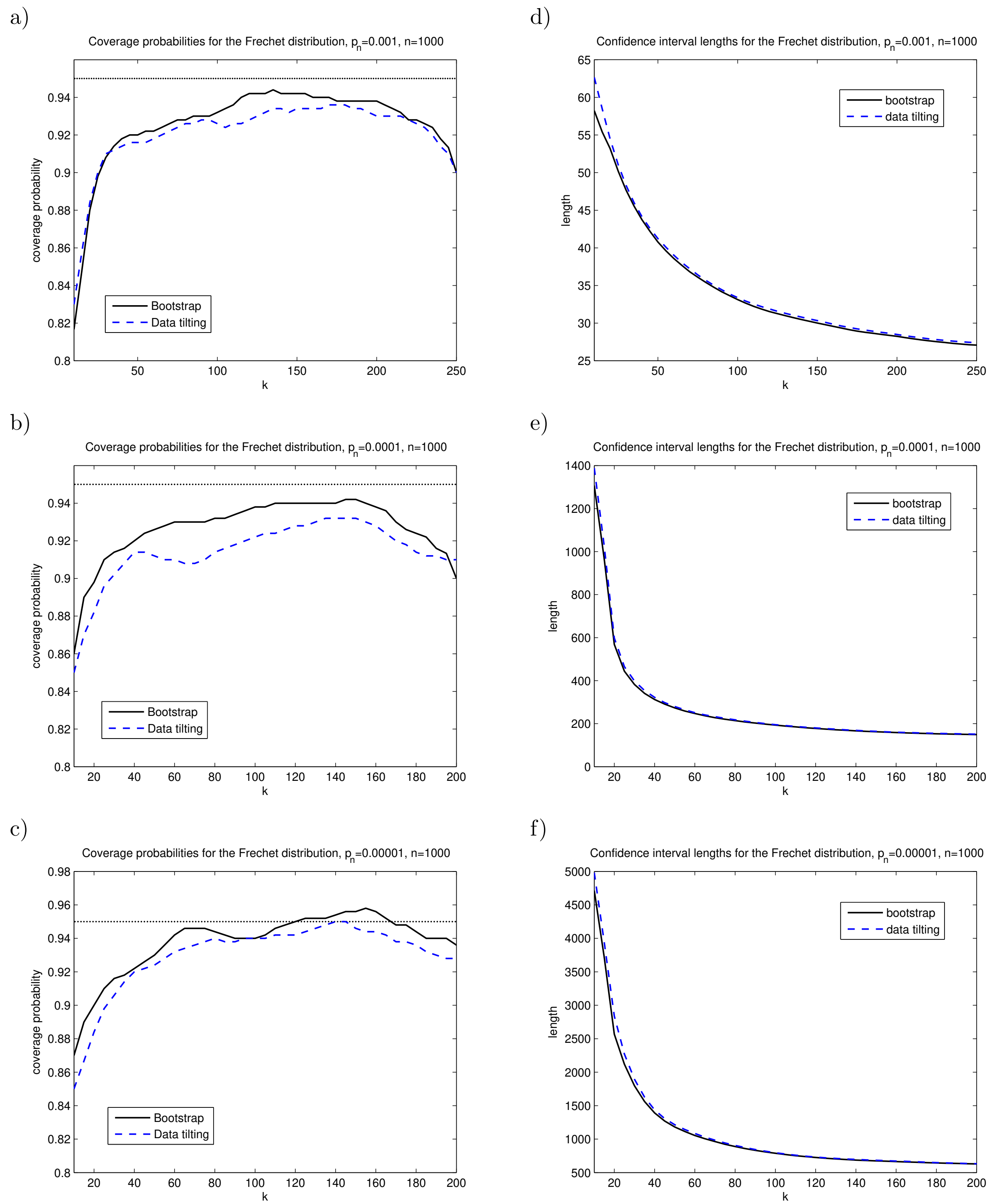

Figure 5: Coverage probabilities and average lengths of the bootstrap and data tilting 95% confidence intervals for the Frechet(2) distribution; $n = 1000$. The coverage probabilities and average lengths of the confidence intervals for quantiles corresponding to the tail probability $1 - p_n$, where: a) $p_n = 0.001$, b) $p_n = 0.0001$ and c) $p_n = 0.00001$ are plotted as a function of k.

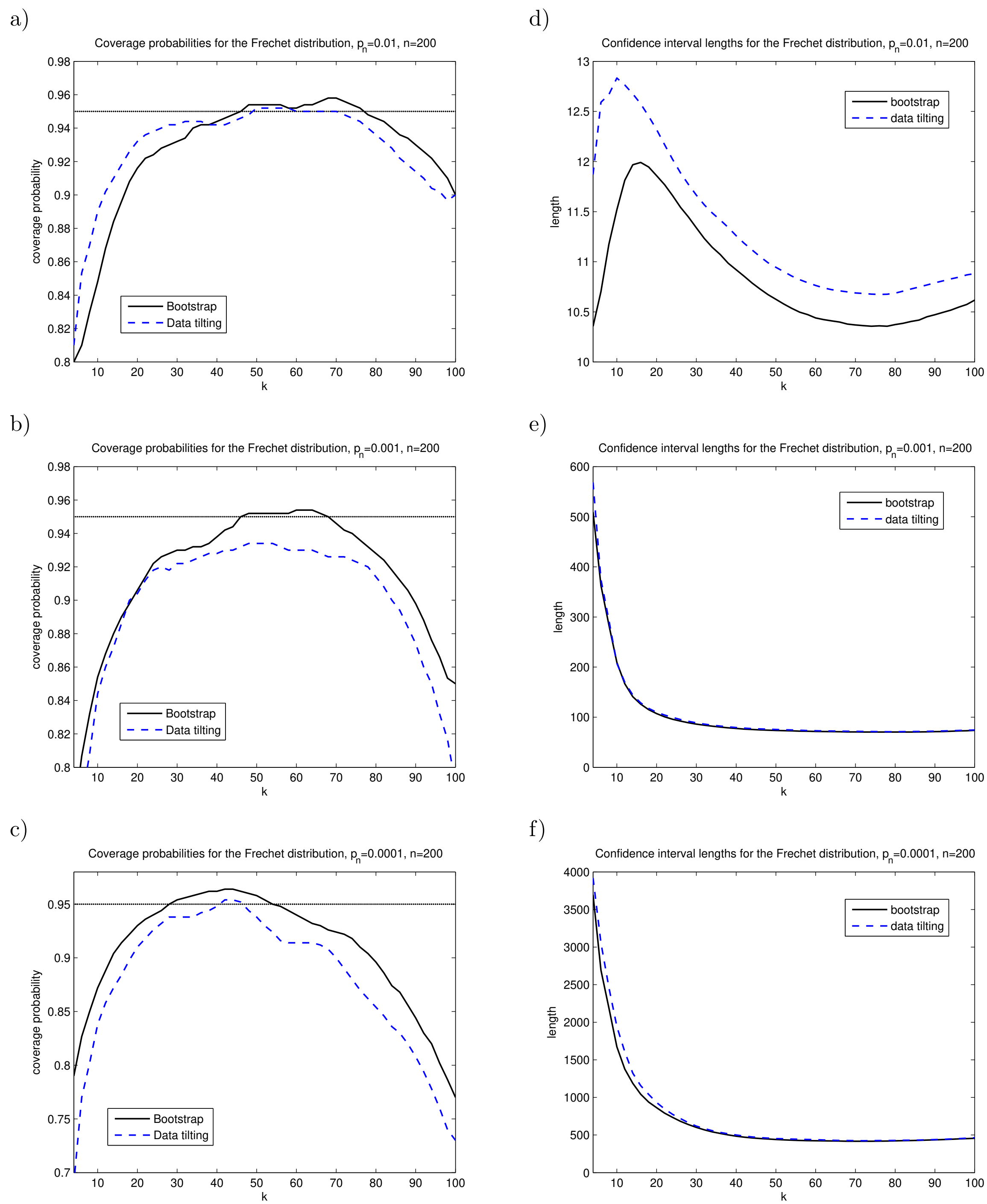

Figure 6: Coverage probabilities and average lengths of the bootstrap and data tilting 95% confidence intervals for the Frechet(2) distribution; $n = 200$. The coverage probabilities and average lengths of the confidence intervals for quantiles corresponding to the tail probability $1 - p_n$, where: a) $p_n = 0.01$, b) $p_n = 0.001$ and c) $p_n = 0.0001$ are plotted as a function of k.

References

Csörgő, M. and Horváth, L. (1987). Approximation of intermediate quantile processes. *Journal of Multivariate Analsis*, **21**, 250–262.

Csörgő, M., Csörgő, S., Horváth, L., and Mason, D. M. (1986). Weighted empirical and quantile processes. *Ann. Probab.*, **14**(1), 31–85.

Csörgő, S. and Mason, D. M. (1989). Bootstrapping empirical functions. *The Annals of statistics*, **17**(4), 1447–1471.

de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction (Springer Series in Operations Research and Financial Engineering)*. Springer.

Peng, L. and Qi, Y. (2017). *Inference for Heavy-Tailed Data: Applications in Insurance and Finance*. Academic Press.

van der Vaart, A. W. (1998). *Asymptotic statistics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.