Summer 5-15-2018

Statistical Inference for the Haezendonck-Goovaerts Risk Measure

Xing Wang

STATISTICAL INFERENCE FOR THE HAEZENDONCK-GOOVAERTS RISK MEASURE

BY

Xing Wang

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree

Of

Doctor of Philosophy

In the Robinson College of Business

Of

Georgia State University

GEORGIA STATE UNIVERSITY
ROBINSON COLLEGE OF BUSINESS
2018

ACCEPTANCE

This dissertation was prepared under the direction of the *Xing Wang* Dissertation Committee. It has been approved and accepted by all members of that committee, and it has been accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Business Administration in the J. Mack Robinson College of Business of Georgia State University.

Richard Phillips, Dean

DISSERTATION COMMITTEE

Liang Peng

Haci Akcin

Samuel Harry Cox

Gengsheng Qin

ABSTRACT

STATISTICAL INFERENCE FOR THE HAEZENDONCK-GOOVAERTS RISK MEASURE

BY

Xing Wang

May 15th, 2018

Committee Chair: *Liang Peng*

Major Academic Unit: *Department of Risk Management and Insurance*

Recently the Haezendonck-Goovaerts (H-G) risk measure is receiving much attention in actuarial science with applications to optimal portfolios and optimal reinsurance because of its advantage in well quantifying the tail behavior of losses. This thesis systematically studies statistical inferences of the H-G risk measure under various settings including heavy-tailed losses, fixed and intermediate risk levels.

The thesis starts by proposing an empirical likelihood inference for the H-G risk measure for two different risk levels—fixed risk level and intermediate risk level. More specifically, Chapter 2 considers the case of fixed risk level, and the derived asymptotic limit of a nonparametric inference is employed to construct an interval for the H-G risk measure. Chapter 3 considers the case of intermediate risk level, i.e., the level is treated as a function of the sample size and goes to one as the sample size tends to infinity. The proposed maximum empirical likelihood estimator for the H-G risk measure has a different limit from that for the case of a fixed level. But the proposed empirical likelihood method indeed gives a unified interval estimation for both cases.

Chapter 4 proposes a two-part estimation for the H-G risk measure and the proposed estimators always have an asymptotic normal distribution regardless of the moment conditions. To achieve this, we separately estimate the tail part by extreme value theory and the middle part non-parametrically.

The above chapters focus on independent data. In Chapter 5, we extend our methodology from

independent data to dependent data and conduct the sensitivity analysis of a portfolio under the H-G risk measure. We first derive an expression for computing the sensitivity of the H-G risk measure, which enables us to estimate the sensitivity non-parametrically via the H-G risk measure. Second, we derive the asymptotic distributions of the nonparametric estimators for the H-G risk measure and its sensitivity by assuming that loss variables in the portfolio follow from a strictly stationary α-mixing sequence. Finally, this estimation combining with a bootstrap method is applied to a real dataset.

Besides the study of the H-G risk measure, we investigate the estimation of the finite endpoint of a distribution function when normally distributed measurement errors contaminate the observations. Under the framework of extreme value theory, we propose a class of estimators for the standard deviation of the measurement errors as well as for the endpoint. Asymptotic properties of the proposed estimators are established and simulations demonstrate their good finite sample performance.

ACKNOWLEDGEMENTS

First and foremost, I am deeply grateful to my advisor, Professor Liang Peng, for his valuable guidance and countless encouragement during my master and Ph.D. period. His continuous aid, motivation, and immense knowledge guide me throughout all the stages of writing this thesis. His words and advice profoundly influenced my research. His kindness, patience, and enthusiasm have made a significant impact on my personality. I could not imagine having a better advisor and mentor for my Ph.D. study.

Second, my sincere thanks go to my dissertation committee members: Professor Haci Akcin, Professor Samuel Harry Cox and Professor Gengsheng Qin for their time and insightful comments that improved my work in various perspectives. I would like to express my gratitude for their continuous support and for evaluating my dissertation.

I would also like to express my heartfelt thanks to the professors and staff of the Risk Management and Insurance Department. I owe the most profound gratitude for their patience and constant help on many academic affairs such as courses, research, and teaching. It is fantastic to have the opportunity to study, work and interact with the faculty, staff and students at Georgia State University. I would like to express my deep appreciation to Professor Stephen H. Shore, Professor Ajay Subramanian, Professor Ruodu Wang and Professor Xingxing Yu for their kindness and assistance throughout my doctoral studies and job search. Moreover, I am thankful to all my friends who have shared moments of happiness and sadness with me and have made my life colorful.

I am deeply grateful for my beloved and supportive husband, Xiaowei Yue, without whom this journey would not have been possible. Last but not the least, I would like to acknowledge my parents to whom no amount of gratitude is adequate for their love, emotional support and confidence in me throughout my studies and my life in general.

Thanks for all your encouragement!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

This dissertation systematically addresses some statistical inference problems for a class of risk measures called Haezendonck-Goovaerts (H-G) risk measures. These risk measures are used to quantify the tail property of loss distributions and are receiving much attention in actuarial science with applications to reinsurance policy and optimal portfolios. However, there are few efficient inference methods for estimating the H-G risk measure when the loss is heavy-tailed distribution. This thesis is dedicated to developing efficient methods to estimate the H-G risk measure under various scenarios including heavy-tailed distributions.

This thesis starts by proposing an empirical likelihood inference for the H-G risk measure at two different risk levels separately. Chapter 2 considers the case that the H-G risk measure is defined at a fixed level, where a nonparametric estimation method is proposed and the derived asymptotic limit is employed to construct an interval for the H-G risk measure. Compared to the nonparametric estimation proposed by Ahn and Shyamalkumar (2014), this method shows a better performance empirically.

Chapter 3 is dedicated to extending the statistical inference from a fixed level to an intermediate level where the level is treated as a function of the sample size. Since the intermediate level tends to one as the sample size goes to infinity, the proposed maximum empirical likelihood estimator for the H-G risk measure has a different limit from that for a fixed level. Interestingly, the proposed empirical likelihood method indeed gives a unified interval estimation for both cases. A simulation study is conducted to examine the finite sample performance of the proposed method.

In Chapter 4, motivated by the fact that many loss variables in insurance and finance could have a heavy tail including infinite variance, we propose a two-part estimation for the H-G risk measure. The proposed estimators always have an asymptotic normal distribution regardless of the moment conditions. To achieve this, we separately estimate the tail part by extreme value theory and the

middle part non-parametrically. A simulation study and real data analysis confirm the effectiveness of the proposed new inference procedure for estimating the H-G risk measure.

Chapters $2 - 4$ focus on statistical inferences for the H-G risk measure when the losses are independent. When this risk measure is applied to insurance or a financial portfolio with several loss variables, sensitivity analysis becomes useful in managing the portfolio, and the assumption of independent observations may not be reasonable. Thus, in Chapter 5, we extend our estimation methodology from independent data to dependent data. First, we derive the theoretical expression for computing the sensitivity of the H-G risk measure, which enables us to estimate the sensitivity non-parametrically via the H-G risk measure. Further, we derive the asymptotic distributions of the nonparametric estimators for the H-G risk measure and its sensitivity by assuming that loss variables in the portfolio follow a strictly stationary α-mixing sequence. A simulation study is provided to examine the finite sample performance of the proposed nonparametric estimators. Also, the method combining with a bootstrap method is applied to a real dataset.

Chapter 6 investigates the estimation of the finite endpoint of a distribution function when normally distributed measurement errors contaminate the observations. Under the framework of extreme value theory, we propose a class of estimators for the standard deviation of the measurement errors as well as for the endpoint. Asymptotic properties of the proposed estimators are established and simulations demonstrate their good finite sample performance. Also, we apply the proposed methods to the outdoor long jump data to estimate the ultimate limit for human beings in the long jump.

To give an overview of the proposed study and employed techniques, Section 1.1 defines the risk measure, including the H-G risk measure and some related problems; the maximum empirical likelihood method which is employed in the statistical inference is presented in Section 1.2; extreme value theory is reviewed in Section 1.3.

1.1 Risk Measure

1.1.1 Definition of the Risk Measure

Risk management generally involves risk identification, risk quantification, and risk prediction. As one of the important parts of risk management, risk quantification is a process of using the observed data to evaluate the risks quantitatively and arrange the risks in the order of importance. Quantifying the risk is associated with capital allocation, decision-making and actuarial premium calculation which are important in risk management and actuarial science. Quantitative techniques help enhance the credibility and the quality of decision-making significantly. During the process of qualitative analysis, risk measures are necessary tools that help with the risk quantification and forecast.

Generally speaking, a risk measure is a function that maps a random variable to a real number. It allows us to link the uncertainty of the loss to some real numbers so as to express the riskiness. The formal definition of risk measure is as follows:

Definition 1.1.1. *A risk measure ρ is a mapping from a set of random variables to the real numbers:*

$$\rho : \mathcal{L} \longrightarrow R,$$
$$X \longmapsto \rho(X),$$

where $\mathcal{L}$ is the L^p space, i.e. $\mathcal{L} = \{X : E||X||^p < \infty\}$.

Since a risk measure can be an arbitrary function that maps a space of probability distributions to real numbers, infinite choices of functions could serve as risk measures. Examples are the mean and the variance, which are commonly used to measure the centrality and dispersion of the risk separately. In order to find a proper function to evaluate the risk and fulfill the needs of risk management, different criteria are used for choosing the optimal risk measure in practice. When risk measures are used as capital requirements to regulate the risk, some desired properties are raised by market participants or the insurance underwriters.

If we treat the random variable X as the position of the risk, it is apparent that the riskier a position is, the higher its risk measure should be. When X is positive, $\rho(X)$ is interpreted as the amount of cash needed to add to the risky position X to make it an acceptable position. On the contrary, if $\rho(X) < 0$, the capital amount $-\rho(X)$ could be taken out from the already being acceptable position to be invested in a more profitable way. Thus, the concept of risk measure is strictly related to that of acceptability. The coherent risk measure introduced by Artzner et al. (1999) is a fundamental concept related to the acceptability of a risk measure. It represents a subset of risk measures achieving the highest status in theoretical studies as well as industry regulation.

Definition 1.1.2. *A risk measure $\rho : \mathcal{L} \longrightarrow R$ is called a coherent risk measure if the following four properties hold:*

- *translation invariance: $X \in \mathcal{L}, \alpha \in R \Rightarrow \rho(X + \alpha) = \rho(X) + \alpha$.*

- *positive homogeneity: $X \in \mathcal{L}, \alpha > 0 \Rightarrow \rho(\alpha X) = \alpha \rho(X)$.*

- *monotonicity: $X_1, X_2 \in \mathcal{L}, X_1 < X_2 \Rightarrow \rho(X_1) \leq \rho(X_2)$.*

- *sub-additivity: $X_1, X_2 \in \mathcal{L} \Rightarrow \rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$.*

In the literature, properties of different kinds of risk measures are studied, and the behavior of risk measures for the heavy-tailed distributions receives a lot of attention.

There are two risk measures emphasizing the tail behavior of the loss: Value at Risk (VaR) and Condition Value at Risk (CVaR). VaR is known as the quantile risk measure or quantile premium principle, which quantifies the value of an asset's tail risk and is always specified with a given confidence level α.

Definition 1.1.3. *The Value-at-Risk (VaR) of random variable X at level α is defined as the lower α-quantile of X*

$$VaR_\alpha(X) := \inf\{x \geq 0 : 1 - F_X(x) \leq \alpha\},$$

where F_X is the cumulative distribution function of X.

4

VaR (in general) is not a coherent risk measure because the sub-additivity does not always hold. Using this risk measure, portfolio diversifications may not lead to risk reduction. From this point of view, VaR is not acceptable for determining regulator capital for financial institutions.

As one of the typically used coherent risk measures, Conditional Value at Risk (CVaR) measures the expectation of the loss under the condition that the loss is above the VaR at a given confidence level α.

Definition 1.1.4. *Let X be a continuous loss random variable. Given a parameter α, $0 < \alpha < 1$, the α-CVaR of X (Conditional Value-at-Risk (CVaR) in the continuous case) is*

$$CVaR_\alpha(X) = E[X|X \geq VaR_\alpha(X)].$$

1.1.2   The Haezondonck-Goovaerts Risk Measure

The so-called Haezendonck-Goovaerts (H-G) risk measure originates from Haezendonck and Goovaerts (1982) by considering the premium calculation principle—Orlicz premium. It is related to the normalized Young function $\psi(.)$ (see Krasnoselskiĭ and Rutitskiĭ (1961) and Skii and Ruttcki (1961)) which generates the H-G risk measure and is defined as follows:

Definition 1.1.5. *Let $\psi : [0, \infty] \to [0, \infty]$ be a convex function. If*

$$\psi(0) = 0, \quad \psi(1) = 1 \quad and \quad \psi(\infty) = \infty,$$

then ψ is called a normalized / generalized Young function.

Definition 1.1.6. *Let $\psi : [0, \infty] \to [0, \infty]$ be a normalized Young function. Suppose X is a real number loss variable. For a fixed number $q \in (0, 1)$ and each $\beta > 0$, let $\alpha = \alpha(\beta)$ be a solution to*

$$E\{\psi(\frac{(X - \beta)_+}{\alpha})\} = 1 - q, \tag{1.1.1}$$

where $x_+ = \max(x, 0)$. We call the unique solution $\alpha(\beta)$ the Orlicz premium corresponding to

X, β and q. Then, the so-called Haezendonck-Goovaerts risk measure with level q is defined as

$$\theta_q = \inf_{\beta > 0} \{\beta + \alpha(\beta)\}. \qquad (1.1.2)$$

Bellini and Gianin (2012) showed that the minimizer exists for all $q \in (0, 1)$, and it is unique when $\psi(.)$ is strictly convex. This minimizer β^* is called the Orlicz quantile of the loss.

From Zhu, Zhang, and Zhang (2013), two groups of generalized Young functions appear frequently in the actuarial literature:

(i) The power Young function: $\psi(x) = x^p$, $p \geq 1$;

(ii) The exponential Young function: $\psi(x) = (e^{\beta x}/(e^\beta - 1))$, $\beta > 0$.

If $\psi(x) = x$, then $\alpha(\beta) = \frac{1}{1-q}E\{(X - \beta)_+\}$ and $\theta = \frac{1}{1-q}E\{(X - F^-(q))_+\}$, where $F(x) = P(X \leq x)$ and $F^-(x)$ denotes the inverse function of $F(x)$. In this case, the H-G risk measure is reduced to the Conditional Value at Risk (CVaR).

The actuarial intuition of the H-G risk measure is provided by Bellini and Gianin (2008b). If we treat $(X - \beta)_+$ as the payment for the loss X when applying franchise deductible β, $\alpha(x)$ represents the corresponding Orlicz premium of this insurance contract. The minimization construction in the definition minimizes the consumption of the insurer. The minimizer β^* represents the optimal choice of the franchise deductible from the point of view of the insurer.

In the study of the inference issue, we formulate the H-G risk measure as a solution to the following estimating equations. Suppose $X_1, ..., X_n$ are independent and identically distributed (i.i.d.) random variables with distribution function $F(x)$. The H-G risk measure is equivalent to solving the following estimating equations

$$\begin{cases} E\{\psi(\frac{X_i - \beta}{\theta - \beta})I(X_i > \beta)\} = 1 - q, \\ E\{\psi'(\frac{X_i - \beta}{\theta - \beta})(X_i - \theta)I(X_i > \beta)\} = 0 \end{cases} \qquad (1.1.3)$$

for some β and $\theta > \beta$ under some conditions (see Tang and Yang (2014)). This transformation enables us to employ the empirical likelihood method in Qin and Lawless (1994) to estimate the

H-G risk measure. The statistical inference of the H-G risk measure can be done by combining techniques of the empirical process and the empirical likelihood method.

1.2 Empirical Likelihood (EL) Methods

When we employ risk measures to the sample with unknown underlying distributions, the empirical likelihood method could provide an efficient way to estimate the H-G risk measure. This section gives an introduction to the empirical likelihood method that will be employed in the statistical inference about risk measures.

1.2.1 Classical Likelihood Ratio Methods

In parametric likelihood methods, we suppose that the joint distribution of all available data has a known parametric form. Let $X_1, X_2, X_3, \ldots, X_n$ be $i.i.d.$ observations with underlying density distribution function $f_X(x; \theta)$, where $\theta \in \Theta$ is q-dimensional parameter and the parameter θ takes its values in the set $\Theta \subseteq R^q$. The likelihood function for these n observations is the joint distribution of $(X_1, X_2, X_3, \ldots, X_n)$ given by

$$L(X_1, X_2, \ldots, X_n; \theta) := \Pi_{j=1}^n f_X(X_j; \theta). \tag{1.2.1}$$

The corresponding log-likelihood function is

$$l(X_1, X_2, \ldots, X_n; \theta) := \sum_{j=1}^n \log(f_X(X_j; \theta)). \tag{1.2.2}$$

Since the principle of maximum likelihood is choosing the estimator $\hat{\theta}$ as the value for the parameter that makes the observed data most probable, the maximum likelihood estimator is the value $\hat{\theta}_n$ such that

$$l(X_1, X_2, \ldots, X_n; \hat{\theta}_n) = \sup_{\theta \in \Theta} l(X_1, X_2, \ldots, X_n; \theta). \tag{1.2.3}$$

7

In order to test the hypothesis $H_0 : \theta = \theta_0$, the likelihood ratio statistic is defined as

$$R(\theta_0) = \frac{L(\theta_0; X_1, X_2, \ldots, X_n)}{L(\hat{\theta}; X_1, X_2, \ldots, X_n)}.$$

Wilks' theorem is often proved and used to construct a confidence interval.

Theorem 1.2.1. *(Wilks' theorem) Under mild regularity conditions, if $\theta = \theta_0$, then*

$$-2 \log(R(\theta_0)) \xrightarrow{d} \chi^2_q,$$

and the likelihood ratio confidence region for θ is

$$\{\theta : -2 \log(R(\theta)) \leq \chi^2_{q;\alpha}\},$$

where $\chi^2_{q;\alpha}$ is the upper α-quantile.

1.2.2 Empirical Maximum Likelihood Methods

The empirical likelihood method was introduced by Owen (1990). It is a nonparametric maximum likelihood estimation. Without assuming the form of the underlying distribution, this method provides data-determined shapes for the confidence region, and it can easily incorporate known constraints on parameters, and adjust for biased sampling schemes.

Let $X_1, X_2, X_3, \ldots$ be $i.i.d.$ observations with unknown underlying distribution function $F_0(x)$. As the point estimator of $F_0(x)$, the empirical cumulative distribution function (ECDF) $F_n(x)$ is defined as

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I(X_i \leq x), \tag{1.2.4}$$

where $I(.)$ is the indicator function. Then the nonparametric likelihood function for these n observations is

$$L(F) := \Pi_{j=1}^{n}(F(X_i) - F(X_i^-)). \tag{1.2.5}$$

8

Here the value $L(F)$ is the probability of getting exactly the observed sample values $X_1, \ldots, X_n$ from the CDF $F(.)$. If F is a continuous distribution function, $L(F) = 0$. If F is a discrete distribution function on $\{X_1, \ldots, X_n\}$ with $p_i = F(X_i)$ for $i = 1, 2, \ldots, n$, where $p_i \geq 0$ and $\sum_{i=1}^{n} p_i = 1$, then the empirical likelihood is $L(F) = \Pi_{i=1}^{n} p_i$.

We also use ratios of the nonparametric likelihood as a basis for hypothesis testing and constructing confidence intervals. We define the ratio as follows:

$$\tilde{R}(F) = \frac{L(F)}{\sup_F \{L(F)\}}.$$

Lemma 1.2.2. *Let $X_1, X_2, \ldots, X_n \in R$ be independent random variables with a common CDF F_0. Let F_n be their ECDF and F be any CDF. If $F \neq F_n$, then $L(F) < L(F_n)$.*

From Lemma 1.2.2, we know $F_n = \arg\max_F \{L(F)\}$, and the EL ratio is

$$\tilde{R}(F) = \frac{L(F)}{L(F_n)} = \Pi_{i=1}^{n} n p_i,$$

where p_i is the probability mass function of X_i, and satisfies $p_i \geq 0$, $\sum_{i=1}^{n} p_i \leq 1$. To maximize $\tilde{R}(F)$, we need consider the support of F on the data, i.e., $\sum_{i=1}^{n} p_i = 1$.

Consider no tied data $X_1, X_2, \ldots, X_n$, i.e., $X_i \neq X_j$ when $i \neq j$. Suppose the distribution function F puts probability $p_i \geq 0$ on the value $X_i \in R^d$, then

$$L(F) = \Pi_{i=1}^{n} p_i \quad \text{and} \quad \tilde{R}(F) = \Pi_{i=1}^{n} n p_i = n^n \Pi_{i=1}^{n} p_i.$$

If we are interested in a parameter $\theta = T(F), F \in \mathcal{F}$ for the functional T, we choose $\mathcal{F}$ containing all distributions that have support on $X_i, i = 1, 2, \ldots, n$. The likelihood ratio function is defined as

$$R(\theta) = \sup_{\{(p_1, p_2, \ldots, p_n)\}} \{n^n \Pi_{i=1}^{n} p_i \mid p_i = f(X_i); T(F) = \theta\}. \tag{1.2.6}$$

We accept $T(F) = \theta_0$ when $R(\theta_0) \geq r_0$ for some threshold r_0, and the corresponding EL confidence

region is $\{\theta : R(\theta) \leq r_0\}$ with r_0 chosen via an EL analogue of Wilks' theorem.

Owen (1990) studied the EL for a univariate mean as an example of function at T, i.e., $T(F) = E[X]$. After adding the restriction $E[X] = \mu_0$, the equation (1.2.6) can be rewritten to the following

$$R(\mu_0) = \sup_{\{(p_1, p_2, \ldots, p_n)\}} \{\Pi_{i=1}^n np_i \mid p_i \geq 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i X_i = \mu_0\}. \qquad (1.2.7)$$

Theorem 1.2.3. *Let $X_1, \ldots, X_n$ be independent random variables with common distribution F_0. Let $\mu_0 = E(X_i)$, and suppose that $0 < Var(X_i) < \infty$. Then*

$$-2\log(R(\mu_0)) \xrightarrow{d} \chi^2_{(1)} \quad as \quad n \to \infty.$$

If $X_1, \ldots, X_n$ are independent random vectors with finite covariance matrix of rank $q > 0$, and $\mu_0 \in R^d$, then the Theorem 1.2.3 could be extended to multivariate data and

$$-2\log(R(\mu_0)) \xrightarrow{d} \chi^2_{(q)} \quad \text{as} \quad n \to \infty.$$

Besides the univariate mean example, Qin and Lawless (1994) extended empirical likelihood methods to estimating equations. If we have a smooth function of means, $\theta = h(\mu)$ which is implicitly defined by $E(g(X, \theta)) = 0$, where $g(X, \theta) = (g_1(X, \theta), g_2(X, \theta), \ldots, g_r(X, \theta))^T$ is the estimating function, the empirical likelihood function ratio with estimating equations is defined as

$$R(\theta) = \sup_{\{(p_1, p_2, \ldots, p_n)\}} \{\Pi_{i=1}^n np_i \mid p_i \geq 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i g(X_i, \theta) = 0\}. \qquad (1.2.8)$$

The estimator of θ is obtained by optimizing $L(\theta)$, i.e., $\hat{\theta}^{EML} := \arg\max_{\theta \in \Theta} R(\theta)$.

Theorem 1.2.4. *Let $X_1, \ldots, X_n$ be i.i.d. random variables, and suppose $g(X, \theta)$ has finite covariance matrix of rank $q > 0$. If $E[g(X, \theta)] = 0$, then*

$$-2\log(R(\theta_0)) \xrightarrow{d} \chi^2_{(1)} \quad as \quad n \to \infty,$$

10

where θ_0 is the true value of θ.

This extension enables us to infer the H-G risk measure after it is transformed to the equation set (1.1.3). Also, the theorem could help with constructing confidence intervals.

1.3 Extreme Value Theory

In this section, a brief introduction to extreme value theory is given. Let $X_1, X_2, X_3, \ldots$ be $i.i.d.$ random variables. In contrast to the central limit theorem where the limit behavior of the partial sum $X_1 + X_2 + \cdots + X_n$ as $n \to \infty$ is studied, the theory of extremes is concerned with the limit behavior of the sample extremes $\max(X_1, X_2, \ldots, X_n)$ or $\min(X_1, X_2, \ldots, X_n)$ as $n \to \infty$. Extreme value theory (EVT) is focused on the possible limit distributions for sample maxima of independent and identically distributed random variables.

In this section, we are interested in (right-) tail properties of distributions. Let F be the underlying cumulative distribution function. Suppose x^* is the right endpoint of X, i.e.,

$$x^* = \sup\{x : F(x) < 1\}, \tag{1.3.1}$$

where x^* could be finite or infinite. From De Haan and Ferreira (2007), there exists a sequence of constants $a_n > 0$, and $b_n \in R$ $(n = 1, 2, \ldots)$ such that

$$\frac{\max_{1 \le i \le n} X_i - b_n}{a_n} \xrightarrow{d} Y \text{ as } n \to \infty, \tag{1.3.2}$$

where Y is a non-degenerate random variable. That is, there exists a nondegenerate distribution function G such that

$$\lim_{n \to \infty} F^n(a_n x + b_n) = G(x) \tag{1.3.3}$$

for every continuity point x of G. Distribution functions G that can occur as a limit in (1.3.3) are called extreme value distributions. The related distribution F satisfying (1.3.3) is called in the maximum domain of attraction or simply domain of attraction of G.

11

From Fisher and Tippett (1928) and Gnedenko (1943), it turns out that G is determined by a single parameter γ which we call the extreme value index.

Theorem 1.3.1. *The class of extreme value distributions is $G_\gamma(ax + b)$ with $a > 0$, $b \in R$, where*

$$G_\gamma(x) = \exp(-(1 + \gamma x)^{-1/\gamma}), \quad 1 + \gamma x > 0. \tag{1.3.4}$$

Von Mises (1936) and Jenkinson (1955) explored the parametrization in Theorem 1.3.1. They classified the distribution F into three different groups according to the index of the extreme value distribution.

Definition 1.3.2. *(1) For $\gamma > 0$, use $G_\gamma((x - 1)/\gamma)$, let $\alpha = 1/\gamma > 0$,*

$$\Phi_\alpha(x) := \begin{cases} 0, & x \leq 0, \\ \exp(-x^{-\alpha}), & x > 0. \end{cases}$$

This class is often called the Fréchet class of distributions

(2) For $\gamma < 0$, use $G_\gamma(-(x + 1)/\gamma)$, let $\alpha = -1/\gamma > 0$,

$$\Psi_\alpha(x) := \begin{cases} \exp(-(-x)^\alpha), & x < 0, \\ 1, & x \geq 0. \end{cases}$$

This class is often called the Reverse-Weibull class of distributions.

(3) For $\gamma = 0$,

$$G_0(x) = \exp(-e^{-x}), x \in R. \tag{1.3.5}$$

This class is often called the Gumbel class of distributions.

When we are using a heavy-tailed distribution to model the risk, we always care about the right part of the loss. When $\gamma > 0$, it is trivial that $G_\gamma(x) > 0$ for all $x > 0$, i.e., the right endpoint of the distribution is infinity and the distribution has a rather heavy right tail. This indicates that the moments of order greater than or equal to $1/\gamma$ do not exist. For $\gamma = 0$, the right endpoint of the

distribution equals infinity. However, the distributions are rather light-tailed and all moments exist. When $\gamma < 0$, the right endpoint of the distribution is $-1/\gamma$, so it has a short right tail. Thus, the bigger the γ, the heavier the right tail.

A simple estimator for estimating the tail index $\gamma \in R$ is the Pickands estimator from Pickands III (1975), defined as

$$\hat{\gamma}_P := (\log 2)^{-1} \log\left(\frac{X_{n-k;n} - X_{n-2k;n}}{X_{n-2k;n} - X_{n-4k;n}}\right), \tag{1.3.6}$$

where $X_{n;n} \geq X_{n-1;n} \geq \cdots \geq X_{1;n}$ is the ordered sample of $X_1, X_2, \ldots, X_n$.

A more widely used way to estimate the tail index for a distribution having a right tail is the Hill estimator in Hill (1975). From the first order variation, we know $F \in D(G_\gamma)$ for $\gamma > 0$ if and only if

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

The Hill estimator $\hat{\gamma}_H$ is defined as

$$\hat{\gamma}_H := \frac{1}{k} \sum_{i=0}^{k-1} \log X_{n-i;n} - \log X_{n-k;n}. \tag{1.3.7}$$

EMPIRICAL LIKELIHOOD INFERENCE FOR THE HAEZENDONCK-GOOVAERTS RISK MEASURE

Recently Haezendonck-Goovaerts risk measure is receiving much attention in actuarial science with applications in the study of optimal portfolio and optimal reinsurance policy. Nonparametric estimation is proposed by Ahn and Shyamalkumar (2014), where the derived asymptotic limit can be employed to construct an interval for the Haezendonck-Goovaerts risk measure. In this chapter, we propose an alternative empirical likelihood inference for this risk measure. A simulation study shows the good performance of the proposed method. The content of this chapter is based on the following joint work:

L. Peng, X. Wang and Y. Zheng (2015). Empirical likelihood inference for Haezendonck-Goovaerts risk measure. European Actuarial Journal 5, 427-445.

2.1 Motivation and Introduction

Let $\psi : [0, \infty] \to [0, \infty]$ be a convex function satisfying $\psi(0) = 0, \psi(1) = 1$ and $\psi(\infty) = \infty$, i.e., a normalized Young function. Suppose X is a loss variable. For a fixed number $q \in (0, 1)$ and each $\beta > 0$, let $\alpha = \alpha(\beta)$ be a solution to

$$E\{\psi(\frac{(X - \beta)_+}{\alpha})\} = 1 - q, \tag{2.1.1}$$

where $x_+ = \max(x, 0)$. Then, the so-called Haezendonck-Goovaerts risk measure with level q is defined as

$$\theta_q = \inf_{\beta > 0}\{\beta + \alpha(\beta)\}. \tag{2.1.2}$$

This risk measure originates from Haezendonck and Goovaerts (1982) by considering the premium calculation principle induced by an Orlicz norm.

Recently there has been an increasing interest in studying Haezendonck-Goovaerts risk measure with applications in actuarial science. For example, Goovaerts et al. (2004) showed that this risk measure preserves the convex order property; Bellini and Gianin (2008a) and Bellini and Gianin (2008b) provided a dual representation for this risk measure; Goovaerts et al. (2012) investigated a relationship between this risk measure and others; Cheung and Lo (2013) obtained a lower bound for this risk measure when a sum of random variables is concerned; studies of optimal portfolio and optimal reinsurance under this risk measure are given by Bellini and Gianin (2008b) and Zhu, Zhang, and Zhang (2013), respectively; Tang and Yang (2012) and Tang and Yang (2014) derived a first order approximation for this risk measure when the underlying distribution is in the domain of attraction of an extreme value distribution, which is of importance in predicting extreme risks; a second order approximation for this risk measure is obtained by Mao and Hu (2012), which is necessary for the study of estimating this risk measure nonparametrically when the level q depends on the sample size and goes to one as the sample size tends to infinity; nonparametric estimation for this risk measure is proposed by Ahn and Shyamalkumar (2014) and its asymptotic limit is derived too.

Although some nice theoretical properties and applications of this Haezendonck-Goovaerts risk measure have been found in the literature, statistical inference is quite underdeveloped. For example, how does one effectively construct a confidence interval for the Haezendonck-Goovaerts risk measure θ_q at a given level $q \in (0, 1)$? Quantifying variability of a risk measure is of importance in risk management such as backtesting. A simple way to obtain an interval for θ_q is to either estimate the asymptotic variance of the nonparametric estimator of θ_q in Ahn and Shyamalkumar (2014) or use a bootstrap method. In general this simple method does not lead to an accurate interval. Alternatively one can investigate the possibility of developing an empirical likelihood method for this risk measure since empirical likelihood methods are powerful in interval estimation and hypothesis tests. We refer to Owen (2001) for an overview on empirical likelihood methods and their advantages. Recently empirical likelihood methods have been proposed for constructing intervals for some risk measures in the literature; see Peng and Qi (2006) for high quantiles; Chan

et al. (2007) for conditional Value-at-Risk; Baysal and Staum (2008) for Value-at-Risk and expected shortfall. A standard way to formulate an empirical likelihood function is via estimating equations; see Qin and Lawless (1994). By noting that the Haezendonck-Goovaerts risk measure can be written as a solution to two estimating equations, we are able to employ the empirical likelihood method in Qin and Lawless (1994) to estimate this risk measure and to construct a confidence interval for it. However the results in Qin and Lawless (1994) can not be applied due to the involved non-smoothing functionals when the Haezendonck-Goovaerts risk measure is written as a solution to estimation equations. Instead, we develop our theoretical results by combining techniques in the empirical process and the empirical likelihood method.

This chapter is organized as follows. Section 2.2 presents the methodology and main results, where the imposed regularity conditions are different from those in Ahn and Shyamalkumar (2014) since we focus on the case of having a normal limit. These conditions can be verified straightforwardly. A simulation study is given in Section 2.3, which shows that the new method has good finite sample performance and provides a more accurate interval than the normal approximation method based on the nonparametric estimator in Ahn and Shyamalkumar (2014). All proofs are put in Section 2.4. Some conclusions are made in Section 2.5.

2.2 Nonparametric Maximum Empirical Likelihood Estimation for Fixed Quantile Level

Throughout suppose $X_1, \cdots, X_n$ are independent random variables with common distribution function $F(x)$, and we use notations $\xrightarrow{p}$, $\xrightarrow{d}$, $\xrightarrow{a.s.}$, $o_p(1)$, $O_p(1)$ and $I(\cdot)$ to denote convergence in probability, convergence in distribution, convergence almost surely, small order in probability, bounded in probability and indicate function, respectively. The nonparametric estimator for the Haezendonck-Goovaerts risk measure proposed by Ahn and Shyamalkumar (2014) first solve the following equation with respect to α for each fixed β:

$$\frac{1}{n} \sum_{i=1}^{n} \psi\left(\frac{(X_i - \beta)_+}{\alpha}\right) = 1 - q. \tag{2.2.1}$$

This equation is the sample version of equation (2.1.1). Denote this solution by $\hat{\alpha}(\beta)$. Next, using (2.1.2), Ahn and Shyamalkumar (2014) defined their nonparametric estimator for θ_q as

$$\hat{\theta}_q^{AS} = \inf_{\beta > 0}\{\beta + \hat{\alpha}(\beta)\}, \tag{2.2.2}$$

and derived its asymptotic limit. As shown by Ahn and Shyamalkumar (2014), the limit could be non-normal. Under some conditions, the limit is normal, and Ahn and Shyamalkumar (2014) proposed an estimator for the asymptotic variance and stated that it is important to study methods for interval estimation such as bootstrap method, but they did not conduct any empirical/theoretical investigation.

Although equations (2.1.1) and (2.1.2) have a unique solution for a given $q \in (0, 1)$ when ψ is strictly convex (see Bellini and Gianin (2012)), $\hat{\theta}_q^{AS}$ may not exist for a large q and finite n due to the first step estimation $\hat{\alpha}(\beta)$; see the simulation results in Table 2.1 below.

By taking derivative with respect to β in (2.1.1), we obtain

$$E\{\psi'(\frac{X - \beta}{\alpha(\beta)})\frac{-\alpha(\beta) - (X - \beta)\alpha'(\beta)}{\alpha^2(\beta)}I(X > \beta)\} = 0. \tag{2.2.3}$$

Equation (2.2.2) implies that we have to solve the equation $1 + \alpha'(\beta) = 0$, which, combining with (2.2.3), results in the following estimating equation

$$E\{\psi'(\frac{X - \beta}{\alpha(\beta)})(X - \beta - \alpha(\beta))I(X > \beta)\} = 0. \tag{2.2.4}$$

Hence, it follows from (2.2.3) and (2.2.4) that $\theta_q(> \beta)$ and β satisfy the following estimating equations:

$$\begin{cases} E\{\psi(\frac{X_i - \beta}{\theta_q - \beta})I(X_i > \beta)\} = 1 - q, \\ E\{\psi'(\frac{X_i - \beta}{\theta_q - \beta})(X_i - \theta_q)I(X_i > \beta)\} = 0. \end{cases} \tag{2.2.5}$$

A rigorous derivation can be found in Tang and Yang (2014) under some conditions. The above view of Haezendonck-Goovaerts risk measure motivates us to consider the following maximum

empirical likelihood estimator for θ_q and empirical likelihood based confidence intervals. Note that moment estimator based on (2.2.5) can be employed too, but its asymptotic behavior will be the same as that of the proposed maximum empirical likelihood estimator.

For $i = 1, \cdots, n$, put

$$Y_i(\theta_q, \beta) = \left(\psi(\frac{X_i - \beta}{\theta_q - \beta})I(X_i > \beta) - 1 + q, \quad \psi'(\frac{X_i - \beta}{\theta_q - \beta})(X_i - \theta_q)I(X_i > \beta) \right)^T.$$

Then it follows from Qin and Lawless (1994) that the empirical likelihood function for (θ_q, β) is defined as

$$L(\theta_q, \beta) = \sup\{\prod_{i=1}^{n}(np_i) : p_1 \geq 0, \cdots, p_n \geq 0, \sum_{i=1}^{n} p_i = 1, \sum_{i=1}^{n} p_i Y_i(\theta_q, \beta) = 0\}.$$

By the Lagrange multiplier technique, we have

$$l(\theta_q, \beta) := -2\log L(\theta_q, \beta) = 2\sum_{i=1}^{n}\log(1 + \lambda^T Y_i(\theta_q, \beta)), \tag{2.2.6}$$

where $\lambda = \lambda(\theta_q, \beta)$ satisfies

$$\sum_{i=1}^{n} \frac{Y_i(\theta_q, \beta)}{1 + \lambda^T Y_i(\theta_q, \beta)} = 0. \tag{2.2.7}$$

As in Qin and Lawless (1994), the maximum empirical likelihood estimator for (θ_q, β) is defined as

$$(\hat{\theta}_q^{MEL}, \hat{\beta}^{MEL}) = \arg\min_{\theta_q > \beta > 0} l(\theta_q, \beta).$$

When an interval for θ_q is concerned, one needs to consider the profile empirical likelihood ratio function $l^P(\theta_q) = \min_{\beta < \theta_q} l(\theta_q, \beta)$.

In order to derive the asymptotic limit of $(\hat{\theta}_q^{MEL}, \hat{\beta}^{MEL})$ and to show that Wilks theorem holds for the above empirical likelihood method, conditions and theorems in Qin and Lawless (1994) can not be applied since our functionals are non-smooth due to the factor $I(X_i > \beta)$. One way to overcome this issue is to smooth the indicator function as Chen and Hall (1993) for quantile

estimation and Chen, Peng, and Zhao (2009) for copulas. Unfortunately this smoothing technique can not be employed here due to the fact that $\psi(t)$ is defined only for $t \geq 0$. Recently Molanes Lopez, Keilegom, and Veraverbeke (2009) gave some general regularity conditions to show that Wilks theorem holds for non-smooth functionals, but did not provide the asymptotic limit of the maximum empirical likelihood estimator. Here we prove our results by combining expansions in empirical processes and empirical likelihood method, which results in the following regularity conditions:

- C2.1) ψ is a strictly convex function on $[0, \infty]$ with $\psi(0) = 0, \psi(1) = 1, \psi(\infty) = \infty$, and $\psi(t)$ has a continuous second derivative on $(0, \infty)$ with $|\psi'(0+)| < \infty$ and $0 \leq \psi''(0+) < \infty$;

- C2.2) F is continuous;

- C2.3) $E\{\sup_{(\theta_q, \beta)^T \in \Omega} |\psi(\frac{X-\beta}{\theta_q-\beta})|^{2\delta_1} I(X > \beta)\} < \infty$ and $E\{\sup_{(\theta_q, \beta)^T \in \Omega} |\psi'(\frac{X-\beta}{\theta_q-\beta})|^{2\delta_1} |X - \theta_q|^{2\delta_1} I(X > \beta)\} < \infty$ for some $\delta_1 > 1$,

$$\sup_{(\theta_q, \beta)^T \in \Omega} \int_\beta^\infty F^{\delta_2}(x)\{1 - F(x)\}^{\delta_2} \Big\{ |\psi'(\tfrac{x-\beta}{\theta_q-\beta})| + \psi''(\tfrac{x-\beta}{\theta_q-\beta})x + \psi(\tfrac{x-\beta}{\theta_q-\beta})|\psi'(\tfrac{x-\beta}{\theta_q-\beta})|$$
$$+ |\psi'(\tfrac{x-\beta}{\theta_q-\beta})|\psi''(\tfrac{x-\beta}{\theta_q-\beta})(x - \theta_q)^2 + (\psi'(\tfrac{x-\beta}{\theta_q-\beta}))^2 |x - \theta_q|$$
$$+ \psi(\tfrac{x-\beta}{\theta_q-\beta})\psi''(\tfrac{x-\beta}{\theta_q-\beta})(x - \beta)|x - \theta| \Big\} dx < \infty$$

for some $\delta_2 \in (0, 1/2)$,

$$\sup_{(\theta_q, \beta)^T \in \Omega} \Big\{ |\int_\beta^\infty \psi''(\tfrac{x-\beta}{\theta_q-\beta})(x - \theta_q)^2 \, dF(x)|$$
$$+ |\int_\beta^\infty \psi''(\tfrac{x-\beta}{\theta_q-\beta})(x - \theta_q)(x - \beta) \, dF(x)|$$
$$+ |\int_\beta^\infty \psi'(\tfrac{x-\beta}{\theta_q-\beta})\psi''(\tfrac{x-\beta}{\theta_q-\beta})(x - \theta_q)^3 \, dF(x)|$$
$$+ |\int_\beta^\infty \psi'(\tfrac{x-\beta}{\theta_q-\beta})\psi''(\tfrac{x-\beta}{\theta_q-\beta})(x - \theta_q)^2(x - \beta) \, dF(x)|$$
$$+ |\int_\beta^\infty \psi(\tfrac{x-\beta}{\theta_q-\beta})\psi''(\tfrac{x-\beta}{\theta_q-\beta})(x - \theta_q)^2 \, dF(x)|$$
$$+ |\int_\beta^\infty \psi(\tfrac{x-\beta}{\theta_q-\beta})\psi''(\tfrac{x-\beta}{\theta_q-\beta})(x - \theta_q)(x - \beta) \, dF(x)| \Big\}$$
$$< \infty,$$

where Ω is an open set including $(\theta_{0,q}, \beta_0)^T$. Here $(\theta_{0,q}, \beta_0)^T$ is the solution to equations

19

Theorem 2.2.1. *Under conditions C2.1)–C2.3), we have*

i)

$$\Sigma_1^T \Sigma_0^{-1} \Sigma_1 \sqrt{n} \begin{pmatrix} \hat{\beta}^{MEL} - \beta_0 \\ \hat{\theta}_q^{MEL} - \theta_{0,q} \end{pmatrix} \xrightarrow{d} N(0, \Sigma_1^T \Sigma_0^{-1} \Sigma_1)$$

as $n \to \infty$, *where* $\Sigma_1 = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$ *and* $\Sigma_0 = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}$ *with*

$$a_1 = \int_{\beta_0}^{\infty} \psi'(\frac{x - \beta_0}{\theta_{0,q} - \beta_0}) \frac{x - \theta_{0,q}}{(\theta_{0,q} - \beta_0)^2} \, dF(x), \quad b_1 = \int_{\beta_0}^{\infty} \psi'(\frac{x - \beta_0}{\theta_{0,q} - \beta_0}) \frac{\beta_0 - x}{(\theta_{0,q} - \beta_0)^2} \, dF(x),$$

$$a_2 = -\psi'(0+)(\beta_0 - \theta_{0,q}) + \int_{\beta_0}^{\infty} \psi''(\frac{x - \beta_0}{\theta_{0,q} - \beta_0}) \frac{(x - \theta_{0,q})^2}{(\theta_{0,q} - \beta_0)^2} \, dF(x),$$

$$b_2 = \int_{\beta_0}^{\infty} \{\psi''(\frac{x - \beta_0}{\theta_{0,q} - \beta_0}) \frac{(\beta_0 - x)(x - \theta_{0,q})}{(\theta_{0,q} - \beta_0)^2} - \psi'(\frac{x - \beta_0}{\theta_{0,q} - \beta_0})\} \, dF(x),$$

and

$$\sigma_1^2 = E\{\psi^2(\frac{X - \beta_0}{\theta_{0,q} - \beta_0})I(X > \beta_0)\} - (1 - q)^2,$$

$$\sigma_{12} = E\{\psi(\frac{X - \beta_0}{\theta_{0,q} - \beta_0})\psi'(\frac{X - \beta_0}{\theta_{0,q} - \beta_0})(X - \theta_{0,q})I(X > \beta_0)\},$$

$$\sigma_2^2 = E\{(\psi'(\frac{X - \beta_0}{\theta_{0,q} - \beta_0}))^2(X - \theta_{0,q})^2 I(X > \beta_0)\}.$$

ii) $l^P(\theta_{0,q})$ *converges in distribution to a chi-squared limit with one degree of freedom as* $n \to \infty$, *which ensures that the proposed empirical likelihood confidence interval below has an asymptotically correct level.*

Remark 2.2.2. *When* Σ_1 *has rank 2, then i) in Theorem 2.2.1 becomes*

$$\sqrt{n} \begin{pmatrix} \hat{\beta} - \beta_0 \\ \hat{\theta}_q^{MEL} - \theta_{0,q} \end{pmatrix} \xrightarrow{d} N(0, \Sigma_1^{-1} \Sigma_0 (\Sigma_1^{-1})^T).$$

In Figure 2.1 below, we plot the determinant of Σ_1 for the uniform distribution and Pareto distributions used in the simulation study, which are positive, i.e., Σ_1 has rank 2.
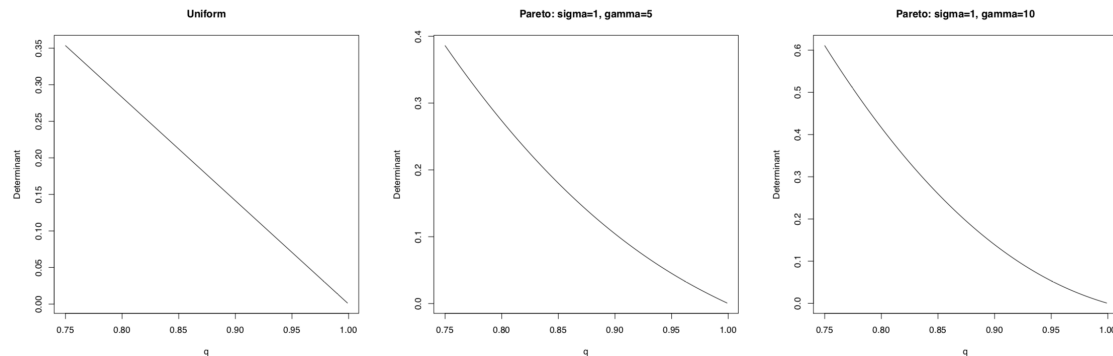
Figure 2.1: Determinant of Σ_1 in Theorem 2.2.1

Remark 2.2.3. *Note that we do not assume $\psi'(0+) = 0$. Instead we assume F is continuous to ensure (2.2.5) holds. So conditions C2.1) and C2.2) appear in Tang and Yang (2014). The first two inequalities with respect to δ_1 in C2.3) ensure Lemmas 2.4.2 and 2.4.3, which are standard for an empirical likelihood method. The other two inequalities in C2.3) are similar to the bounded conditions for partial derivatives with respect to parameters in Qin and Lawless (1994), which are employed in the proof of Lemma 2.4.1. We employ these different conditions due to non-differentiability. All conditions C2.1)–C2.3) can be checked straightforward.*

Based on the above theorem, a confidence interval for $\theta_{0,q}$ with level ξ is obtained as

$$I_{\xi}^{EL} = \{\theta_q : l^P(\theta_q) \leq \chi_{1,\xi}^2\},$$

where $\chi_{1,\xi}^2$ denotes the ξ−th quantile of a chi-squared distribution with one degree of freedom.

We remark that the above regularity conditions are different from those in Ahn and Shyamalkumar (2014). A theoretical comparison for these two estimators is hard due to their complicated asymptotic variances. Instead a simulation comparison is given in Section 2.3, which shows that the new method has some advantages. Moreover, if one is interested in a confidence region for

risk measures $\theta_{q_1}, \cdots, \theta_{q_m}$ at several different levels $q_1, q_2, \cdots, q_m$, the above empirical likelihood method can easily be extended by considering corresponding $2m$ equations. We skip details.

2.3 Simulation Study

In this section, we examine the finite sample behavior of the proposed maximum empirical likelihood estimator and the empirical likelihood based confidence interval, and compare them with the nonparametric estimator in Ahn and Shyamalkumar (2014) in terms of mean squared errors and coverage accuracy. First we compare the finite sample behavior of these two estimators $\hat{\theta}_q^{MEL}$ and $\hat{\theta}_q^{AS}$ in terms of mean squared errors and biases. For computing these quantities, we employ $\psi(x) = \frac{x^2+x}{2}I(x > 0)$ and draw $10,000$ random samples with sample size $n = 500$ and $2,000$ from one of the following two distributions

$$F_1(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } 0 < x < 1 \\ 1 & \text{if } x \geq 1 \end{cases} \quad \text{and } F_2(x; \gamma, \sigma) = \begin{cases} 0 & \text{if } x \leq 0 \\ 1 - (1 + \sigma x)^{-\gamma} & \text{if } x > 0, \end{cases}$$

where $\sigma > 0$ and $\gamma > 4$. For these two distributions, an explicit formula for θ_q is available in Ahn and Shyamalkumar (2014). It is easy to check that conditions C2.1–C2.3) in Theorem 2.2.1 are satisfied. For example, one can choose any $\delta_1 > 1$ and $\delta_2 \in (0, \frac{1}{2})$ in C2.3) for distribution $F_1(x)$, and choose any $1 < \delta_1 < \frac{\gamma+1}{4}$ and $\frac{2}{\gamma} < \delta_2 < \frac{1}{2}$ in C2.3) for distribution $F_2(x; \gamma, \sigma)$ when Ω is chosen small enough. In Table 2.1 we report the bias, standard deviation and square root of mean squared error for these two estimators at different levels $q = 0.9, 0.95, 0.99$. We also report the number of times when the minimization fails to give a solution. From Table 2.1, we observe that i) $\hat{\theta}_q^{MEL}$ has a smaller mean squared error than $\hat{\theta}_q^{AS}$ for distribution $F_1(x)$ except the case $n = 500$ and $q = 0.99$, where $\hat{\theta}_q^{AS}$ can not be calculated for 517 out of $10,000$ times; ii) $\hat{\theta}_q^{AS}$ has a smaller mean squared error than $\hat{\theta}_q^{MEL}$ for distribution $F_2(x; 1, 15)$, but sometimes has a larger mean squared error for distribution $F_2(x; 1, 5)$; iii) $\hat{\theta}_q^{AS}$ has a computational issue especially when $q = 0.99$, i.e., minimization fails sometimes.

Next we compare the proposed empirical likelihood based confidence interval with the normal approximation method based on $\hat{\theta}_q^{AS}$ in terms of coverage probability by drawing $1,000$ random samples with sample size $n = 500$ and $2,000$. We employ the same Young function $\psi(x)$ and distribution functions $F_1(x)$ and $F_2(x; \gamma, \sigma)$ as above. For computing the empirical coverage probability of the proposed empirical likelihood method, we first use the R package 'emplik' to compute $l(\theta_{0,q}, \beta)$ for each β, and then use the R package 'nlm' to minimize $l(\theta_{0,q}, \beta)$ over $\beta < \theta_{0,q}$ so as to get $l^P(\theta_{0,q})$. For comparison with the interval, denoted by I_ξ^{AS}, obtained from the nonparametric estimator $\hat{\theta}_q^{AS}$, we employ the naive bootstrap method by drawing $1,000$ resamples from the original sample to construct the bootstrap confidence interval. We also compute the bootstrap calibrated empirical likelihood based confidence interval, denoted by I_ξ^{BEL}, by drawing $1,000$ resamples from the original sample and using these $1,000$ bootstrapped versions of $l^P(\hat{\theta}_q^{MEL})$ to obtain the critical value; see Owen (2001) for details on calibration for empirical likelihood methods. We report the empirical coverage probabilities for these three intervals with levels $\xi = 0.9$ and 0.95 for different $q = 0.9, 0.95, 0.99$ in Table 2.2, which show that i) the proposed empirical likelihood method performs better than the normal approximation method based on $\hat{\theta}_q^{AS}$ in most cases; ii) the proposed bootstrap calibrated empirical likelihood method gives most accurate coverage probability; iii) coverage accuracy for these three intervals improves when either the sample size increases or γ in the distribution $F_2(x; \sigma, \gamma)$ increases, i.e., tail becomes lighter.

In summary, the proposed maximum empirical likelihood estimator $\hat{\theta}_q^{MEL}$ and empirical likelihood based confidence interval I_ξ^{EL} perform well in comparison with the corresponding methods based on the nonparametric estimator $\hat{\theta}_q^{AS}$ in Ahn and Shyamalkumar (2014) in terms of mean squared error, coverage probability and computational difficulty.

2.4 Proofs

Throughout we define the empirical distribution as $F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I(X_i \leq x)$ and empirical process as $\alpha_n(x) = \sqrt{n}\{F_n(x) - F(x)\}$. Then by the classical theory in empirical processes and

23

Skorohod construction, we have

$$\sup_{-\infty < x < \infty} |\alpha_n(x) - B(x)| \overset{a.s.}{\to} 0 \quad \text{and} \quad \sup_{-\infty < x < \infty} \frac{|\alpha_n(x)|}{F^\nu(x)(1 - F(x))^\nu} = O_p(1) \qquad (2.4.1)$$

for any $\nu \in (0, \frac{1}{2})$, where $B(x)$ is a Gaussian process with zero mean and covariance

$$E\{B(x_1)B(x_2)\} = F(x_1 \wedge x_2) - F(x_1)F(x_2);$$

see Shorack and Wellner (1986).

Lemma 2.4.1. *Under conditions of Theorem 2.2.1, when $|\beta - \beta_0| + |\theta_q - \theta_{0,q}| = \Delta_n \overset{p}{\to} 0$ as $n \to \infty$, we have*

$$\begin{aligned}
& \tfrac{1}{n} \sum_{i=1}^n \psi(\tfrac{X_i - \beta}{\theta_q - \beta}) I(X_i > \beta) - 1 + q \\
= \ & \int_{\beta_0}^\infty \{F(x) - F_n(x)\} \psi'(\tfrac{x - \beta_0}{\theta_{0,q} - \beta_0}) \tfrac{1}{\theta_{0,q} - \beta_0} \, dx \\
& + (\beta - \beta_0) \int_{\beta_0}^\infty \psi'(\tfrac{x - \beta_0}{\theta_{0,q} - \beta_0}) \tfrac{x - \theta_{0,q}}{(\theta_{0,q} - \beta_0)^2} \, dF(x) \\
& + (\theta_q - \theta_{0,q}) \int_{\beta_0}^\infty \psi'(\tfrac{x - \beta_0}{\theta_{0,q} - \beta_0}) \tfrac{\beta_0 - x}{(\theta_{0,q} - \beta_0)^2} \, dF(x) + o_p(\tfrac{1}{\sqrt{n}} + \Delta_n),
\end{aligned}$$

$$\begin{aligned}
& \tfrac{1}{n} \sum_{i=1}^n \psi'(\tfrac{X_i - \beta}{\theta_q - \beta})(X_i - \theta_q) I(X_i > \beta) \\
= \ & \psi'(0+)(\beta_0 - \theta_{0,q})\{F(\beta_0) - F_n(\beta_0)\} \\
& + \int_{\beta_0}^\infty \{F(x) - F_n(x)\}\{\psi''(\tfrac{x - \beta_0}{\theta_{0,q} - \beta_0}) \tfrac{x - \theta_{0,q}}{\theta_{0,q} - \beta_0} + \psi'(\tfrac{x - \beta_0}{\theta_{0,q} - \beta_0})\} \, dx \\
& + (\beta - \beta_0)\{-\psi'(0+)(\beta_0 - \theta_{0,q}) + \int_{\beta_0}^\infty \psi''(\tfrac{x - \beta_0}{\theta_{0,q} - \beta_0}) \tfrac{(x - \theta_{0,q})^2}{(\theta_{0,q} - \beta_0)^2} \, dF(x)\} \\
& + (\theta_q - \theta_{0,q}) \int_{\beta_0}^\infty \{\psi''(\tfrac{x - \beta_0}{\theta_{0,q} - \beta_0}) \tfrac{(x - \theta_{0,q})(\beta_0 - x)}{(\theta_{0,q} - \beta_0)^2} - \psi'(\tfrac{x - \beta_0}{\theta_{0,q} - \beta_0})\} \, dF(x) + o_p(\tfrac{1}{\sqrt{n}} + \Delta_n),
\end{aligned}$$

$$\begin{aligned}
& \tfrac{1}{n} \sum_{i=1}^n \psi^2(\tfrac{X_i - \beta}{\theta_q - \beta}) I(X_i > \beta) - \int_{\beta_0}^\infty \psi^2(\tfrac{x - \beta_0}{\theta_{0,q} - \beta_0}) \, dF(x) \\
= \ & 2 \int_{\beta_0}^\infty \{F(x) - F_n(x)\} \psi(\tfrac{x - \beta_0}{\theta_{0,q} - \beta_0}) \psi'(\tfrac{x - \beta_0}{\theta_{0,q} - \beta_0}) \tfrac{1}{\theta_{0,q} - \beta_0} \, dx \\
& + (\beta - \beta_0) \int_{\beta_0}^\infty 2\psi(\tfrac{x - \beta_0}{\theta_{0,q} - \beta_0}) \psi'(\tfrac{x - \beta_0}{\theta_{0,q} - \beta_0}) \tfrac{x - \theta_{0,q}}{(\theta_{0,q} - \beta_0)^2} \, dF(x) \\
& + (\theta_q - \theta_{0,q}) \int_{\beta_0}^\infty 2\psi(\tfrac{x - \beta_0}{\theta_{0,q} - \beta_0}) \psi'(\tfrac{x - \beta_0}{\theta_{0,q} - \beta_0}) \tfrac{\beta_0 - x}{(\theta_{0,q} - \beta_0)^2} \, dF(x) + o_p(\tfrac{1}{\sqrt{n}} + \Delta_n),
\end{aligned}$$

24

$$\frac{1}{n}\sum_{i=1}^{n}\{\psi'(\frac{X_i-\beta}{\theta_q-\beta})\}^2(X_i-\theta_q)^2 I(X_i > \beta) - \int_{\beta_0}^{\infty}\{\psi'(\frac{x-\beta_0}{\theta_{0,q}-\beta_0})\}^2(x-\theta_{0,q})^2\,dF(x)$$

$$= \{F(\beta_0) - F_n(\beta_0)\}\{\psi'(0+)\}^2(\beta_0 - \theta_{0,q})^2$$

$$+ 2\int_{\beta_0}^{\infty}\{F(x) - F_n(x)\}\{\psi'(\frac{x-\beta_0}{\theta_{0,q}-\beta_0})\psi''(\frac{x-\beta_0}{\theta_{0,q}-\beta_0})\frac{(x-\theta_{0,q})^2}{\theta_{0,q}-\beta_0} + (\psi'(\frac{x-\beta_0}{\theta_{0,q}-\beta_0}))^2(x-\theta_{0,q})\}\,dx$$

$$+ (\beta - \beta_0)\{-(\psi'(0+))^2(\beta_0-\theta_{0,q})^2 + 2\int_{\beta_0}^{\infty}\psi'(\frac{x-\beta_0}{\theta_{0,q}-\beta_0})\psi''(\frac{x-\beta_0}{\theta_{0,q}-\beta_0})\frac{(x-\theta_{0,q})^3}{(\theta_{0,q}-\beta_0)^2}\}\,dF(x)$$

$$+ (\theta_q - \theta_{0,q})\int_{\beta_0}^{\infty}\{2\psi'(\frac{x-\beta_0}{\theta_{0,q}-\beta_0})\psi''(\frac{x-\beta_0}{\theta_{0,q}-\beta_0})\frac{(\beta_0-x)(x-\theta_{0,q})^2}{(\theta_{0,q}-\beta_0)^2} + 2(\psi'(\frac{x-\beta_0}{\theta_{0,q}-\beta_0}))^2(\theta_{0,q}-x)\}\,dF(x)$$

$$+ o_p(\frac{1}{\sqrt{n}} + \Delta_n)$$

and

$$\frac{1}{n}\sum_{i=1}^{n}\psi(\frac{X_i-\beta}{\theta_q-\beta})\psi'(\frac{X_i-\beta}{\theta_q-\beta})(X_i-\theta_q)I(X_i > \beta) - \int_{\beta_0}^{\infty}\psi(\frac{x-\beta_0}{\theta_{0,q}-\beta_0})\psi'(\frac{x-\beta_0}{\theta_{0,q}-\beta_0})(x-\theta_{0,q})\,dF(x)$$

$$= \int_{\beta_0}^{\infty}\{F(x)-F_n(x)\}\{(\psi'(\frac{x-\beta_0}{\theta_{0,q}-\beta_0}))^2\frac{x-\theta_{0,q}}{\theta_{0,q}-\beta_0} + \psi(\frac{x-\beta_0}{\theta_{0,q}-\beta_0})\psi''(\frac{x-\beta_0}{\theta_{0,q}-\beta_0})\frac{x-\theta_{0,q}}{\theta_{0,q}-\beta_0}$$

$$+ \psi(\frac{x-\beta_0}{\theta_{0,q}-\beta_0})\psi'(\frac{x-\beta_0}{\theta_{0,q}-\beta_0})\}\,dx$$

$$+ (\beta - \beta_0)\int_{\beta_0}^{\infty}\{(\psi'(\frac{x-\beta_0}{\theta_{0,q}-\beta_0}))^2\frac{(x-\theta_{0,q})^2}{(\theta_{0,q}-\beta_0)^2} + \psi(\frac{x-\beta_0}{\theta_{0,q}-\beta_0})\psi''(\frac{x-\beta_0}{\theta_{0,q}-\beta_0})\frac{(x-\theta_{0,q})^2}{(\theta_{0,q}-\beta_0)^2}\}\,dF(x)$$

$$+ (\theta_q - \theta_{0,q})\int_{\beta_0}^{\infty}\{(\psi'(\frac{x-\beta_0}{\theta_{0,q}-\beta_0}))^2\frac{(\beta_0-x)(x-\theta_{0,q})}{(\theta_{0,q}-\beta_0)^2} + \psi(\frac{x-\beta_0}{\theta_{0,q}-\beta_0})\psi''(\frac{x-\beta_0}{\theta_{0,q}-\beta_0})\frac{(\beta_0-x)(x-\theta_{0,q})}{(\theta_{0,q}-\beta_0)^2}$$

$$- \psi(\frac{x-\beta_0}{\theta_{0,q}-\beta_0})\psi'(\frac{x-\beta_0}{\theta_{0,q}-\beta_0})\}\,dF(x) + o_p(\frac{1}{\sqrt{n}} + \Delta_n).$$

Proof. It follows from the Taylor expansion that

$$\frac{1}{n}\sum_{i=1}^{n}\psi(\frac{X_i-\beta}{\theta_q-\beta})I(X_i > \beta) - 1 + q$$

$$= \int_{\beta}^{\infty}\psi(\frac{x-\beta}{\theta_q-\beta})\,dF_n(x) - \int_{\beta_0}^{\infty}\psi(\frac{x-\beta_0}{\theta_{0,q}-\beta_0})\,dF(x)$$

$$= -\int_{\beta}^{\infty}\psi(\frac{x-\beta}{\theta_q-\beta})\,d\{F(x)-F_n(x)\} + \int_{\beta}^{\infty}\psi(\frac{x-\beta}{\theta_q-\beta})\,dF(x) - \int_{\beta_0}^{\infty}\psi(\frac{x-\beta_0}{\theta_{0,q}-\beta_0})\,dF(x)$$

$$= \int_{\beta}^{\infty}\{F(x)-F_n(x)\}\psi'(\frac{x-\beta}{\theta_q-\beta})\frac{1}{\theta_q-\beta}\,dx + (\beta-\beta_0)\int_{\beta_1}^{\infty}\psi'(\frac{x-\beta_1}{\theta_1-\beta_1})\frac{x-\theta_1}{(\theta_1-\beta_1)^2}\,dF(x)$$

$$+ (\theta-\theta_{0,q})\int_{\beta_1}^{\infty}\psi'(\frac{x-\beta_1}{\theta_1-\beta_1})\frac{\beta_1-x}{(\theta_1-\beta_1)^2}\,dF(x)$$

$$= \int_{\beta_0}^{\infty}\{F(x)-F_n(x)\}\psi'(\frac{x-\beta_0}{\theta_{0,q}-\beta_0})\frac{1}{\theta_{0,q}-\beta_0}\,dx + (\beta-\beta_0)\{F_n(\beta_2)-F(\beta_2)\}\psi'(0+)\frac{1}{\theta_2-\beta_2}$$

$$+ (\beta-\beta_0)\int_{\beta_2}^{\infty}\{F(x)-F_n(x)\}\{\psi''(\frac{x-\beta_2}{\theta_2-\beta_2})\frac{x-\theta_2}{(\theta_2-\beta_2)^3} + \psi'(\frac{x-\beta_2}{\theta_2-\beta_2})\frac{1}{(\theta_2-\beta_2)^2}\}\,dx$$

$$+ (\theta_q-\theta_{0,q})\int_{\beta_2}^{\infty}\{F(x)-F_n(x)\}\{\psi''(\frac{x-\beta_2}{\theta_2-\beta_2})\frac{\beta_2-x}{(\theta_2-\beta_2)^2} + \psi'(\frac{x-\beta_2}{\theta_2-\beta_2})\frac{1}{(\theta_2-\beta_2)^2}\}\,dx$$

$$+ (\beta-\beta_0)\int_{\beta_1}^{\infty}\psi'(\frac{x-\beta_1}{\theta_1-\beta_1})\frac{x-\theta_1}{(\theta_1-\beta_1)^2}\,dF(x) + (\theta-\theta_{0,q})\int_{\beta_1}^{\infty}\psi'(\frac{x-\beta_1}{\theta_1-\beta_1})\frac{\beta_1-x}{(\theta_1-\beta_1)^2}\,dF(x)$$

$$:= I_1 + \cdots + I_6,$$

$$(2.4.2)$$

where $(\theta_1, \beta_1)^T = \lambda_1(\theta_q, \beta)^T + (1 - \lambda_1)(\theta_{0,q}, \beta_0)^T$ and $(\theta_2, \beta_2)^T = \lambda_2(\theta_q, \beta)^T + (1 - \lambda_2)(\theta_{0,q}, \beta_0)^T$ for some $\lambda_1, \lambda_2 \in [0, 1]$. It follows from (2.4.1) that

$$I_2 = O_p(\frac{1}{\sqrt{n}}\Delta_n) = o_p(\frac{1}{\sqrt{n}} + \Delta_n). \qquad (2.4.3)$$

By (2.4.1) and condition C2.3), we have

$$I_3 = O_p(\frac{1}{\sqrt{n}}\Delta_n) = o_p(\frac{1}{\sqrt{n}} + \Delta_n) \quad \text{and} \quad I_4 = O_p(\frac{1}{\sqrt{n}}\Delta_n) = o_p(\frac{1}{\sqrt{n}} + \Delta_n). \qquad (2.4.4)$$

Note that the condition $E\{\sup_{(\theta_q,\beta)^T \in \Omega} |\psi'(\frac{X-\beta}{\theta_q-\beta})|^{2\delta_1}|X - \theta_q|^{2\delta_1}I(X > \beta)\} < \infty$ for some $\delta_1 > 1$ in condition C2.3) implies that

$$\begin{cases} \sup_{(\theta_q,\beta)^T \in \Omega} \int_\beta^\infty |\psi'(\frac{x-\beta}{\theta_q-\beta})|\, dF(x) < \infty \\ \sup_{(\theta_q,\beta)^T \in \Omega} \int_\beta^\infty |\psi'(\frac{x-\beta}{\theta_q-\beta})|x\, dF(x)\} < \infty \end{cases} \qquad (2.4.5)$$

by noting that $|\psi'(0+)| < \infty$. Similarly, the condition $\sup_{(\theta_q,\beta)^T \in \Omega} \int_\beta^\infty \psi''(\frac{x-\beta}{\theta_q-\beta})(x - \theta_q)^2\, dF(x) < \infty$ in C2.3) implies that

$$\begin{cases} \sup_{(\theta_q,\beta)^T \in \Omega} \int_\beta^\infty \psi''(\frac{x-\beta}{\theta_q-\beta})x^2\, dF(x) < \infty \\ \sup_{(\theta_q,\beta)^T \in \Omega} \int_\beta^\infty \psi''(\frac{x-\beta}{\theta_q-\beta})x\, dF(x) < \infty \\ \sup_{(\theta_q,\beta)^T \in \Omega} \int_\beta^\infty \psi''(\frac{x-\beta}{\theta_q-\beta})\, dF(x) < \infty. \end{cases} \qquad (2.4.6)$$

Hence it follows from (2.4.5), (2.4.6) and the Taylor expansion that

$$\begin{cases} I_5 = (\beta - \beta_0) \int_{\beta_0}^\infty \psi'(\frac{x-\beta_0}{\theta_{0,q}-\beta_0})\frac{x-\theta_{0,q}}{(\theta_{0,q}-\beta_0)^2}\, dF(x) + O_p(\Delta_n^2) \\ I_6 = (\theta_q - \theta_{0,q}) \int_{\beta_0}^\infty \psi'(\frac{x-\beta_0}{\theta_{0,q}-\beta_0})\frac{\beta_0-x}{(\theta_{0,q}-\beta_0)^2}\, dF(x) + O_p(\Delta_n^2). \end{cases} \qquad (2.4.7)$$

Therefore, the first equation in Lemma 2.4.1 follows from (2.4.2), (2.4.3), (2.4.4) and (2.4.7). The rest can be shown similarly. ∎

Lemma 2.4.2. *Under conditions of Theorem 2.2.1, we have*

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i(\theta_{0,q}, \beta_0) \xrightarrow{d} N(0, \Sigma_0)$$

and

$$\frac{1}{n} \sum_{i=1}^{n} Y_i(\theta_{0,q}, \beta_0) Y_i^T(\theta_{0,q}, \beta_0) \xrightarrow{p} \Sigma_0,$$

where Σ_0, given in Theorem 2.2.1, is positive definite.

Proof. We only need to show that Σ_0 is positive definite since the rest directly follows from the central limit theorem and the weak law of large numbers, or by using Lemma 2.4.1 and (2.4.1). Hence, we need to show that $Var\Big((a,b)Y_i(\theta_{0,q}, \beta_0)\Big) > 0$ for any $a^2 + b^2 \neq 0$.

If $(a,b)Y_i(\theta_{0,q}, \beta_0)$ is degenerate, then

$$a\psi(\frac{x - \beta_0}{\theta_{0,q} - \beta_0}) + b\psi'(\frac{x - \beta_0}{\theta_{0,q} - \beta_0})(x - \theta_{0,q}) = c \tag{2.4.8}$$

for some constant c and all $x > \beta_0$. Obviously, when $b = 0$, (2.4.8) can not be true since ψ is a strictly convex function. By assuming $b \neq 0$, it follows from (2.4.8) that

$$a\psi'(\frac{x - \beta_0}{\theta_{0,q} - \beta_0})\frac{1}{\theta_{0,q} - \beta_0} + b\psi''(\frac{x - \beta_0}{\theta_{0,q} - \beta_0})\frac{x - \theta_{0,q}}{\theta_{0,q} - \beta_0} + b\psi'(\frac{x - \beta_0}{\theta_{0,q} - \beta_0}) = 0$$

for all $x > \beta_0$, i.e.,

$$(\log \psi'(\frac{x - \beta_0}{\theta_{0,q} - \beta_0}))' = -(\frac{a}{b(\theta_{0,q} - \beta_0)} + 1)\frac{1}{x - \theta_{0,q}} \quad \text{for} \quad x > \beta_0,$$

i.e.,

$$\log \psi'(\frac{x - \beta_0}{\theta_{0,q} - \beta_0}) = -(\frac{a}{b(\theta_{0,q} - \beta_0)} + 1) \log |x - \theta_{0,q}| + c_1$$

for some constant c_1 and all $x > \beta_0$, which is impossible since the left hand side is an increasing function of x, but the right hand side is not. Hence (2.4.8) can not be true, i.e., Σ_0 is positive definite.

Lemma 2.4.3. *Under conditions of Theorem 2.2.1, we have*

$$\sup_{1 \le i \le n} \sup_{(\theta,\beta)^T \in \Omega} ||Y_i(\theta,\beta)|| = o_p(n^{\frac{1}{2\gamma}})$$

for some $\gamma \in (1,\delta_1)$, where $||\cdot||$ denotes L_2 norm.

Proof. Note that

$$
\begin{aligned}
&P(\sup_{1\le i\le n}\sup_{(\theta,\beta)^T\in\Omega}\psi(\tfrac{X_i-\beta}{\theta-\beta})I(X_i>\beta)\ge n^{\frac{1}{2\gamma}})\\
\le\ & \textstyle\sum_{i=1}^n P(\sup_{(\theta,\beta)^T\in\Omega}\psi(\tfrac{X_i-\beta}{\theta-\beta})I(X_i>\beta)\ge n^{\frac{1}{2\gamma}})\\
\le\ & \tfrac{n}{n^{\delta_1/\gamma}}E\sup_{(\theta,\beta)^T\in\Omega}\psi^{2\delta_1}(\tfrac{X_1-\beta}{\theta-\beta})I(X_1>\beta)\\
\to\ & 0.
\end{aligned}
$$

Similarly

$$P(\sup_{1\le i\le n}\sup_{(\theta,\beta)^T\in\Omega}|\psi'(\frac{X_i-\beta}{\theta-\beta})(X_i-\theta)I(X_i>\beta)|\ge n^{\frac{1}{2\gamma}})\to 0.$$

Hence, the lemma follows. ∎

Proof of Theorem 2.2.1. i) Like the proof of Owen (1990), it follows from Lemmas 2.4.1–2.4.3 and C2.3) that

$$\lambda = \{\frac{1}{n}\sum_{i=1}^n Y_i(\theta_q,\beta)Y_i^T(\theta_q,\beta)\}^{-1}\frac{1}{n}\sum_{i=1}^n Y_i(\theta_q,\beta)(1+o_p(1)),$$

and further

$$
\begin{aligned}
l(\theta_q,\beta) &= 2\textstyle\sum_{i=1}^n \lambda^T Y_i(\theta_q,\beta) - \sum_{i=1}^n \lambda^T Y_i(\theta_q,\beta)Y_i^T(\theta_q,\beta)\lambda + o_p(1)\\
&= \{\tfrac{1}{\sqrt{n}}\textstyle\sum_{i=1}^n Y_i(\theta_q,\beta)\}^T\{\tfrac{1}{n}\sum_{i=1}^n Y_i(\theta_q,\beta)Y_i^T(\theta_q,\beta)\}^{-1}\{\tfrac{1}{\sqrt{n}}\sum_{i=1}^n Y_i(\theta_q,\beta)\} + o_p(1)\\
&= \{\tfrac{1}{\sqrt{n}}\textstyle\sum_{i=1}^n Y_i(\theta_q,\beta)\}^T\Sigma_0^{-1}\{\tfrac{1}{\sqrt{n}}\sum_{i=1}^n Y_i(\theta_q,\beta)\} + o_p(1).
\end{aligned}
$$

$$\tag{2.4.9}$$

28

Put $\nu/\sqrt{n} = (\beta - \beta_0, \theta_q - \theta_{0,q})^T$. Then it follows from (2.4.9) and Lemmas 2.4.1–2.4.2 that

$$l(\theta_q, \beta) = \{\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i(\theta_{0,q}, \beta_0) + \Sigma_1 \nu\}^T \Sigma_0^{-1} \{\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i(\theta_{0,q}, \beta_0) + \Sigma_1 \nu\} + o_p(1),$$

which is minimized at

$$\Sigma_1^T \Sigma_0^{-1} \Sigma_1 \nu = -\Sigma_1^T \Sigma_0^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i(\theta_{0,q}, \beta_0) + o_p(1),$$

i.e., i) holds.

ii) Put $\nu_1/\sqrt{n} = \beta - \beta_0$ and $a = (a_1, a_2)^T$. As above, we can show that

$$l(\theta_{0,q}, \beta_0) = \{\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i(\theta_{0,q}, \beta_0)\}^T \Sigma_0^{-1} \{\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i(\theta_{0,q}, \beta_0)\} + o_p(1)$$

and

$$l(\theta_{0,q}, \beta) = \{\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i(\theta_{0,q}, \beta_0) + \nu a\}^T \Sigma_0^{-1} \{\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i(\theta_{0,q}, \beta_0) + \nu a\} + o_p(1).$$

Hence

$$l(\theta_{0,q}, \beta) - l(\theta_{0,q}, \beta_0)$$
$$= \nu a^T \Sigma_0^{-1} \{\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i(\theta_{0,q}, \beta_0)\} + \{\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i(\theta_{0,q}, \beta_0)\}^T \Sigma_0^{-1} \{\nu a\}$$
$$+ \nu a^T \Sigma_0^{-1} \{\nu a\} + o_p(1),$$

which is minimized at
$$\nu = \frac{-a^T \Sigma_0^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i(\theta_{0,q}, \beta_0)}{a^T \Sigma_0^{-1} a} + o_p(1),$$

i.e.,

$$
\begin{aligned}
l^P(\theta_{0,q}) \quad &= l(\theta_{0,q}, \beta_0) - \frac{a^T \Sigma_0^{-1} \{\frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i(\theta_{0,q}, \beta_0)\} a^T \Sigma_0^{-1} \{\frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i(\theta_{0,q}, \beta_0)\}}{a^T \Sigma_0^{-1} a} + o_p(1) \\
&= \{\frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i(\theta_{0,q}, \beta_0)\}^T \Sigma_0^{-1/2} \Sigma_0^{-1/2} \{\frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i(\theta_{0,q}, \beta_0)\} \\
&\quad - \{\frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i(\theta_{0,q}, \beta_0)\}^T \Sigma_0^{-1/2} \frac{\Sigma_0^{-1/2} a a^T \Sigma_0^{-1/2}}{a^T \Sigma_0^{-1} a} \Sigma_0^{-1/2} \{\frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i(\theta_{0,q}, \beta_0)\} + o_p(1) \\
&= \{\frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i(\theta_{0,q}, \beta_0)\}^T \Sigma_0^{-1/2} \{I_{2\times2} - \frac{\Sigma_0^{-1/2} a a^T \Sigma_0^{-1/2}}{a^T \Sigma_0 a}\} \Sigma_0^{-1/2} \{\frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i(\theta_{0,q}, \beta_0)\} \\
&\quad + o_p(1),
\end{aligned}
$$

where $I_{2\times2}$ denotes the 2 by 2 identity matrix. Since $I_{2\times2} - \frac{\Sigma_0^{-1/2} a a^T \Sigma_0^{-1/2}}{a^T \Sigma_0 a}$ is symmetric, idempotent and its trace equals to one, ii) follows from Lemma 2.4.2. ∎

2.5 Conclusions

By writing the Haezendonck-Goovaerts risk measure as a solution to two estimating equations, we study the maximum empirical likelihood estimator and the empirical likelihood based confidence interval for this risk measure. Due to non-differentiability, conditions and theorems in Qin and Lawless (1994) can not be applied. Instead results are derived by combining techniques in empirical processes and empirical likelihood method, which results in some different regularity conditions from those in Ahn and Shyamalkumar (2014). The imposed regularity conditions are straightforward to check such as uniform distribution, Pareto distribution and exponential distribution. Comparison with the nonparametric estimator in Ahn and Shyamalkumar (2014) shows that the proposed empirical likelihood inference has good finite sample performance. Moreover, the new method is easy to implement by using existing R packages 'emplik' and 'nlm', and to extend to a joint inference for several levels $(q_1, \cdots, q_m)$ by using $2m$ estimating equations.

Table 2.1: *Estimation.* We report the bias (Bias), standard deviation (SD), square root of mean squared error (SRMSE) for both estimators $\hat{\theta}_q^{MEL}$ and $\hat{\theta}_q^{AS}$ at different levels $q = 0.9, 0.95, 0.99$ and with sample size $n = 500$ and $2,000$. We also report the number of times when the minimization fails (NoNS).

CDF	(n,q)	$\hat{\theta}_q^{MEL}$ Bias	$\hat{\theta}_q^{MEL}$ SD	$\hat{\theta}_q^{MEL}$ SRMSE	$\hat{\theta}_q^{MEL}$ NoNS	$\hat{\theta}_q^{AS}$ Bias	$\hat{\theta}_q^{AS}$ SD	$\hat{\theta}_q^{AS}$ SRMSE	$\hat{\theta}_q^{AS}$ NoNS
$F_1(\cdot)$	(500, 0.9)	-1.107e-4	6.745e-3	6.746e-3	0	-8.223e-4	7.242e-3	7.288e-3	0
$F_1(\cdot)$	(500, 0.95)	-1.630e-4	4.895e-3	4.898e-3	0	-8.597e-4	5.311e-3	5.380e-3	0
$F_1(\cdot)$	(500, 0.99)	-6.728e-4	1.296e-2	1.298e-2	0	-1.105e-3	2.644e-3	2.865e-3	517
$F_1(\cdot)$	(2000, 0.9)	-2.584e-5	3.439e-3	3.439e-3	0	-2.506e-4	3.611e-3	3.619e-3	0
$F_1(\cdot)$	(2000, 0.95)	6.258e-5	2.424e-3	2.425e-3	0	-2.082e-4	2.559e-3	2.568e-3	0
$F_1(\cdot)$	(2000, 0.99)	-1.082e-6	1.174e-3	1.174e-3	0	-2.118e-4	1.224e-3	1.243e-3	10
$F_2(\cdot;1,5)$	(500, 0.9)	-1.523e-2	1.451e-1	1.459e-1	0	-1.405e-2	1.669e-1	1.675e-1	0
$F_2(\cdot;1,5)$	(500, 0.95)	-3.104e-2	2.187e-1	2.209e-1	0	-3.249e-2	2.195e-1	2.219e-1	0
$F_2(\cdot;1,5)$	(500, 0.99)	-1.243e-1	5.574e-1	5.711e-1	0	-1.068e-1	5.311e-1	5.418e-1	76
$F_2(\cdot;1,5)$	(2000, 0.9)	-5.211e-3	8.629e-2	8.645e-2	0	-5.118e-3	8.043e-2	8.059e-2	0
$F_2(\cdot;1,5)$	(2000, 0.95)	-1.253e-2	1.200e-1	1.206e-1	0	-1.218e-2	1.216e-1	1.222e-1	0
$F_2(\cdot;1,5)$	(2000, 0.99)	-4.879e-2	3.429e-1	3.464e-1	0	-4.770e-2	3.290e-1	3.324e-1	45
$F_2(\cdot;1,15)$	(500, 0.9)	-1.101e-3	2.294e-2	2.297e-2	0	-1.629e-3	2.147e-2	2.154e-2	3
$F_2(\cdot;1,15)$	(500, 0.95)	-2.839e-3	3.159e-2	3.172e-2	0	-3.583e-3	3.091e-2	3.112e-2	2
$F_2(\cdot;1,15)$	(500, 0.99)	-1.198e-2	7.219e-2	7.318e-2	0	-1.621e-2	7.116e-2	7.298e-2	14
$F_2(\cdot;1,15)$	(2000, 0.9)	-1.110e-4	1.455e-2	1.456e-2	0	-4.944e-4	1.085e-2	1.086e-2	0
$F_2(\cdot;1,15)$	(2000, 0.95)	7.363e-4	1.708e-2	1.710e-2	0	-1.023e-3	1.586e-2	1.589e-2	0
$F_2(\cdot;1,15)$	(2000, 0.99)	-3.981e-3	3.929e-2	3.949e-2	0	-4.747e-3	3.902e-2	3.931e-2	15

Table 2.2: *Coverage accuracy*. We report coverage probabilities for intervals I_ξ^{EL}, I_ξ^{BEL} and I_ξ^{AS} with levels $\xi = 0.9$ and 0.95 for different $q = 0.9, 0.95, 0.99$ and sample size $n = 500$ and $2,000$.

CDF	(n, q)	$I_{0.9}^{EL}$	$I_{0.9}^{BEL}$	$I_{0.9}^{AS}$	$I_{0.95}^{EL}$	$I_{0.95}^{BEL}$	$I_{0.95}^{AS}$
$F_1(\cdot)$	$(500, 0.9)$	0.901	0.890	0.883	0.951	0.949	0.930
$F_1(\cdot)$	$(500, 0.95)$	0.904	0.897	0.852	0.950	0.948	0.902
$F_1(\cdot)$	$(500, 0.99)$	0.846	0.933	0.809	0.883	0.959	0.825
$F_1(\cdot)$	$(2000, 0.9)$	0.892	0.890	0.887	0.951	0.944	0.929
$F_1(\cdot)$	$(2000, 0.95)$	0.907	0.901	0.888	0.946	0.943	0.933
$F_1(\cdot)$	$(2000, 0.99)$	0.916	0.890	0.861	0.954	0.940	0.904
$F_2(\cdot; 1, 5)$	$(500, 0.9)$	0.780	0.842	0.793	0.861	0.903	0.840
$F_2(\cdot; 1, 5)$	$(500, 0.95)$	0.753	0.816	0.751	0.838	0.890	0.812
$F_2(\cdot; 1, 5)$	$(500, 0.99)$	0.557	0.781	0.634	0.606	0.825	0.704
$F_2(\cdot; 1, 5)$	$(2000, 0.9)$	0.831	0.871	0.837	0.905	0.923	0.896
$F_2(\cdot; 1, 5)$	$(2000, 0.95)$	0.825	0.868	0.818	0.895	0.914	0.875
$F_2(\cdot; 1, 5)$	$(2000, 0.99)$	0.765	0.833	0.771	0.838	0.908	0.861
$F_2(\cdot; 1, 15)$	$(500, 0.9)$	0.868	0.879	0.845	0.929	0.940	0.912
$F_2(\cdot; 1, 15)$	$(500, 0.95)$	0.864	0.883	0.813	0.912	0.929	0.883
$F_2(\cdot; 1, 15)$	$(500, 0.99)$	0.642	0.818	0.703	0.691	0.898	0.757
$F_2(\cdot; 1, 15)$	$(2000, 0.9)$	0.873	0.887	0.865	0.929	0.937	0.929
$F_2(\cdot; 1, 15)$	$(2000, 0.95)$	0.872	0.886	0.864	0.936	0.942	0.917
$F_2(\cdot; 1, 15)$	$(2000, 0.99)$	0.866	0.894	0.824	0.917	0.940	0.878

CHAPTER 3

INFERENCE FOR THE INTERMEDIATE HAEZENDONCK-GOOVAERTS RISK MEASURE

Nonparametric inference of the Haezendonck-Goovaerts (H-G) risk measure has been studied by Ahn and Shyamalkumar (2014) and Peng, Wang, and Zheng (2015) when the risk measure is defined at a fixed level. In risk management, the level is usually set to be quite near one by regulators. Therefore, especially when the sample size is not large enough, it is useful to treat the level as a function of the sample size, which diverges to one as the sample size goes to infinity. In this chapter, we extend the results in Peng, Wang, and Zheng (2015) from a fixed level to an intermediate level. Although the proposed maximum empirical likelihood estimator for the H-G risk measure has a different limit for a fixed level and an intermediate level, the proposed empirical likelihood method indeed gives a unified interval estimation for both cases. A simulation study is conducted to examine the finite sample performance of the proposed method. The content of this chapter is based on the joint work:

X. Wang and L. Peng (2016). Inference for intermediate Haezendonck-Goovaerts risk measure. Insurance: Mathematics and Economics 68, 231-240.

3.1 Motivation and Introduction

Let X denote a loss variable with distribution F. Then the q-th quantile of F is defined as $F^{\leftarrow}(q) = \inf\{x|F(x) \geq q\}$, which is also called Value-at-Risk (VaR) in risk management. A simple nonparametric estimator for a quantile is the so-called empirical quantile. However, in risk management, the level q is usually set to be quite near one by regulators. Therefore, when q is close to one and n is not large enough, it may be useful to model $1 - q$ as a function of n, which goes to zero as n turns to infinity, so as to improve the quantile estimation. Here we have to consider two situations separately: intermediate quantile (i.e., $q = q_n \to 1$ and $n(1 - q_n) \to \infty$ as $n \to \infty$)

and extreme quantile (i.e., $q = q_n \to 1$ and $n(1 - q_n) \to c \geq 0$ as $n \to \infty$). For an extreme quantile, it usually involves extrapolation of data, say assuming the underlying distribution is in the domain of attraction of an extreme value distribution; see De Haan and Ferreira (2007) . We refer to Matthys et al. (2004) and Vandewalle and Beirlant (2006) for applications to actuarial science. For an intermediate quantile, the empirical quantile is still consistent, but has a different asymptotic limit from the case of a fixed quantile. This complicates interval estimation since distinguishing a fixed level and an intermediate level is practically impossible. Therefore finding a unified inference is of importance. It is also known that a bootstrap method and the delete-1 jackknife method do not lead to a consistent interval estimation for empirical quantiles; see Shao and Tu (2012).

Recently Li, Gong, and Peng (2010) showed that the empirical likelihood method gives a unified interval estimation for quantiles at a fixed level and an intermediate level. As quantile is one of many commonly employed risk measures in insurance and finance, one may wonder whether empirical likelihood inference can unify other risk measures at a fixed level and an intermediate level. In this chapter, we investigate this possibility for the so-called Haezendonck-Goovaerts (H-G) risk measure, which has been studied a lot in the literature of actuarial science recently.

Let $\psi : [0, \infty] \to [0, \infty]$ be a convex function satisfying $\psi(0) = 0, \psi(1) = 1$ and $\psi(\infty) = \infty$, i.e., ψ is a normalized Young function. For a number $q \in (0, 1)$ and each $\beta > 0$, let $\alpha = \alpha(\beta)$ be a solution to

$$E\{\psi(\frac{(X - \beta)_+}{\alpha})\} = 1 - q, \tag{3.1.1}$$

where $x_+ = \max(x, 0)$. Then, the H-G risk measure with level q is defined as

$$\theta = \inf_{\beta > 0}\{\beta + \alpha(\beta)\}, \tag{3.1.2}$$

see Haezendonck and Goovaerts (1982) for details.

Recently properties and applications to reinsurance and risk management of the H-G risk measure have been studied in the literature of actuarial science; see Bellini and Gianin (2008a); Bellini and Gianin (2008b); Bellini and Gianin (2012) , Cheung and Lo (2013) , Goovaerts et

al. (2004), Goovaerts et al. (2012) and Zhu, Zhang, and Zhang (2013). For nonparametric inference, we refer to Ahn and Shyamalkumar (2014) and Peng, Wang, and Zheng (2015). All these papers consider a fixed level q. When $q = q_n \to 1$ as $n \to \infty$, Tang and Yang (2012); Tang and Yang (2014) derived asymptotic approximations for the H-G risk measure by assuming that the underlying distribution belongs to the domain of attraction of an extreme value distribution, and a second order approximation is derived in Mao and Hu (2012). However, nonparametric estimation for the H-G risk measure with either an intermediate level or an extreme level remains unknown in the literature. For the extreme quantile, it seems that approximations in Tang and Yang (2012); Tang and Yang (2014) can be employed to derive an estimator for the H-G risk measure by combining them with extreme value statistics, and results in Mao and Hu (2012) are useful in deriving the asymptotic limit of the proposed estimator. In this chapter we mainly concern with an intermediate level, that is, one expects nonparametric estimators for a fixed level are still valid, but the asymptotic limit for a fixed level and an intermediate level will be quite different. For the purpose of giving a unified inference, we will investigate the possibility of extending the empirical likelihood inference for the H-G risk measure at a fixed level in Peng, Wang, and Zheng (2015) to an intermediate level.

We organize this chapter as follows. Section 3.2 presents the empirical likelihood method and asymptotic results. A simulation study and data analysis are given in Section 3.3. All proofs are put in Section 3.4. Some conclusions are summarized in Section 3.5.

3.2 Nonparametric Maximum Empirical Likelihood Estimation for Intermediate Quantile Level

Throughout suppose $X, X_1, ..., X_n$ are independent and identically distributed random variables with distribution function $F(x)$. Notations $\xrightarrow{p}$, $\xrightarrow{d}$, $o_p(1)$, $O_p(1)$ and $I(\cdot)$ denote convergence in probability, convergence in distribution, small order in probability, bounded in probability and indicator function, respectively.

Since the H-G risk measure is equivalent to solving the following estimating equations

$$
\begin{cases}
E\{\psi(\frac{X_i - \beta}{\theta - \beta})I(X_i > \beta)\} = 1 - q, \\
E\{\psi'(\frac{X_i - \beta}{\theta - \beta})(X_i - \theta)I(X_i > \beta)\} = 0
\end{cases}
\tag{3.2.1}
$$

for some β and $\theta > \beta$ under some conditions (see Tang and Yang (2014)), Peng, Wang, and Zheng (2015) proposed to employ the empirical likelihood inference in Qin and Lawless (1994) to the above estimating equations, derived the asymptotic limit of the maximum empirical likelihood estimator and proved that Wilk's theorem holds for the empirical likelihood method. Note that results in Qin and Lawless (1994) are not applicable due to the non-differentiable issue caused by the involved indicator function. We refer to Owen (2001) for an overview on empirical likelihood methods.

In this chapter, we extend the results in Peng, Wang, and Zheng (2015) to the case of intermediate quantile, i.e., the case of $q_n \to 1$ and $n(1 - q_n) \to \infty$ as $n \to \infty$. As in Peng, Wang, and Zheng (2015), we put

$$
Y_{ni}(\theta, \beta) = \left(\psi(\frac{X_i - \beta}{\theta - \beta})I(X_i > \beta) - 1 + q_n, \quad \psi'(\frac{X_i - \beta}{\theta - \beta})(X_i - \theta)I(X_i > \beta) \right)^T
$$

for $i = 1, \cdots, n$, and define the empirical likelihood function for (θ, β) as

$$
L_n(\theta, \beta) = \sup\{\prod_{i=1}^{n}(np_i) : p_1 \geq 0, \cdots, p_n \geq 0, \sum_{i=1}^{n} p_i = 1, \sum_{i=1}^{n} p_i Y_{ni}(\theta, \beta) = 0\}.
$$

By the Lagrange multiplier technique, we have

$$
l_n(\theta, \beta) := -2\log L_n(\theta, \beta) = 2\sum_{i=1}^{n} \log(1 + \lambda_n^T Y_{ni}(\theta, \beta)),
\tag{3.2.2}
$$

where $\lambda_n = \lambda_n(\theta, \beta)$ satisfies

$$
\sum_{i=1}^{n} \frac{Y_{ni}(\theta, \beta)}{1 + \lambda_n^T Y_{ni}(\theta, \beta)} = 0.
\tag{3.2.3}
$$

Therefore the maximum empirical likelihood estimator for (θ, β) is defined as

$$(\hat{\theta}_n^{MEL}, \hat{\beta}_n^{MEL}) = \arg\min_{\theta > \beta} l_n(\theta, \beta).$$

It is known that the asymptotic limit of empirical quantile for an intermediate quantile requires some restrictions on the tail behavior of the underlying distribution (see Drees and Haan (1999)). Hence the study of intermediate H-G risk measure needs some conditions on the tail behavior of F too. Here we focus on the case of heavy-tailed distribution and normalized Young function for ψ. Throughout let $(\beta_{n0}, \theta_{n0})$ denote the true values of (β, θ) determined by (3.1.1) and (3.1.2) with $q = q_n$, and assume the following conditions hold.

- C3.1) $\psi(t) = t^k$ for some $k > 1$. Hence ψ is a normalized Young function.

- C3.2) $\lim_{t \to \infty} \frac{tF'(t)}{1 - F(t)} = \gamma$ for some $\gamma > 2k$, where k is given in C3.1), which implies that $\lim_{t \to \infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-\gamma}$ for $x > 0$.

- C3.3) As $n \to \infty$, $q_n \to 1$ and $\sqrt{n(1 - q_n)}/F^{\leftarrow}(q_n) \to \infty$, which imply $n(1 - q_n) \to \infty$.

Theorem 3.2.1. *Under conditions C3.1)–C3.3), we have*

$$\frac{\sqrt{n(1 - q_n)}}{F^{\leftarrow}(q_n)} \begin{pmatrix} \hat{\beta}_n^{MEL} - \beta_{n0} \\ \hat{\theta}_n^{MEL} - \theta_{n0} \end{pmatrix} \xrightarrow{d} N(0, \Sigma_1^{-1}\Sigma_0(\Sigma_1^{-1})^T),$$

as $n \to \infty$, where $\Sigma_1 = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$ and $\Sigma_0 = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}$ with

$$a_1 = \int_0^{c_2^{-\gamma}} k(\frac{x^{-1/\gamma} - c_2}{c_1 - c_2})^{k-1} \frac{x^{-1/\gamma} - c_1}{(c_1 - c_2)^2} \, dx,$$

$$b_1 = \int_0^{c_2^{-\gamma}} k(\frac{x^{-1/\gamma} - c_2}{c_1 - c_2})^{k-1} \frac{c_2 - x^{-1/\gamma}}{(c_1 - c_2)^2} \, dx,$$

$$a_2 = \int_0^{c_2^{-\gamma}} k(k - 1)(\frac{x^{-1/\gamma} - c_2}{c_1 - c_2})^{k-2} \frac{(x^{-1/\gamma} - c_1)^2}{(c_1 - c_2)^2} \, dx,$$

37

$$b_2 = \int_0^{c_2^{-\gamma}} \{k(k-1)(\frac{x^{-1/\gamma} - c_2}{c_1 - c_2})^{k-2} \frac{(c_2 - x^{-1/\gamma})(x^{-1/\gamma} - c_1)}{(c_1 - c_2)^2} - k(\frac{x^{-1/\gamma} - c_2}{c_1 - c_2})^{k-1}\} \, dx,$$

$$\sigma_1^2 = \int_0^{c_2^{-\gamma}} (\frac{x^{-1/\gamma} - c_2}{c_1 - c_2})^{2k} \, dx,$$

$$\sigma_2^2 = \int_0^{c_2^{-\gamma}} k^2 (\frac{x^{-1/\gamma} - c_2}{c_1 - c_2})^{2k-2} (x^{-1/\gamma} - c_1)^2 \, dx,$$

$$\sigma_{12} = \int_0^{c_2^{-\gamma}} (\frac{x^{-1/\gamma} - c_2}{c_1 - c_2})^k k(\frac{x^{-1/\gamma} - c_2}{c_1 - c_2})^{k-1} (x^{-1/\gamma} - c_1) \, dx,$$

$$c_1 = c_1(\gamma, k) = \frac{\gamma(\gamma - k)^{k/\gamma - 1}}{k^{(k-1)/\gamma}} (B(\gamma - k, k))^{1/\gamma},$$

$$c_2 = c_2(\gamma, k) = \frac{(\gamma - k)^{k/\gamma}}{k^{(k-1)/\gamma}} (B(\gamma - k, k))^{1/\gamma},$$

and $B(a, b) = \int_0^1 x^{a-1}(1-x)^{b-1} \, dx$.

By comparing the above limit with the asymptotic limit in Peng, Wang, and Zheng (2015) for a fixed level q, the maximum empirical likelihood estimator for the H-G risk measure is always consistent, but has a different limit for a fixed level and an intermediate level. However, for interval estimation, the following theorem shows that the proposed empirical likelihood method provides a unified interval estimation for a fixed level and an intermediate level by combining it with the result in Peng, Wang, and Zheng (2015) for a fixed level.

Theorem 3.2.2. *Under conditions C3.1)–C3.3), $l_n^P(\theta_{n0}) = \min_{\beta < \theta_{n0}} l_n(\beta, \theta_{n0})$ converges in distribution to a chi-squared limit with one degree of freedom as $n \to \infty$, which ensures that the empirical likelihood confidence interval $I^{EL}(\alpha) = \{\theta : l_n^P(\theta) \leq \chi_{1,\alpha}^2\}$ has an asymptotically correct level α, where $\chi_{1,\alpha}^2$ denotes the αth quantile of a chi-squared limit with one degree of freedom.*

3.3   Simulation Study and Data Analysis

3.3.1   Simulation Study

In this subsection we examine the finite sample performance of the proposed empirical likelihood inference in terms of mean and standard deviation for the point estimation and coverage probability

for the interval estimation.

We draw $10,000$ random samples with sample size $n = 200$ and $n = 2000$ from the Pareto distribution function $F(x) = 1 - x^{-\gamma}$ for $x \geq 1$ with $\gamma = 3$ and $\gamma = 5$, the t-distribution with degrees of freedom $\gamma = 3$ and $\gamma = 5$, and the standard exponential distribution. We take $\psi(x) = x^{1.1}$ and choose q_n to satisfy

$$\frac{\sqrt{n(1-q_n)}}{F^{\leftarrow}(q_n)} = n^d \quad \text{with} \quad d = \frac{2}{5}, \frac{1}{4}, \frac{1}{8}.$$

In Tables 3.1–3.3 below we report the true values $(\theta_{n0}, \beta_{n0})$, the means and standard deviations of $(\frac{\hat{\theta}_n^{MEL}}{\theta_{n0}}, \frac{\hat{\beta}_n^{MEL}}{\beta_{n0}})$, and the coverage probabilities for $I^{EL}(0.9)$ and $I^{EL}(0.95)$. The case of $d = \frac{2}{5}$ gives a level q a bit away from one, which may be treated as a fixed level given the considered sample size n. As shown in Tables 3.1 and 3.2, the maximum empirical likelihood estimator is consistent, and the empirical likelihood inference gives a unified interval estimation. For the small $d = \frac{1}{8}$, coverage accuracy improves as the sample size becomes larger. Although we focus on developing our method for the case of heavy-tailed distributions, Table 3.3 does indicate the method may work for light tailed distributions too. A future project is to prove the conjecture that the proposed empirical likelihood method is indeed valid for any distribution in the domain of attraction of an extreme value distribution.

In conclusion, the simulation results do show that the proposed empirical likelihood method provides a unified inference for the H-G risk measure at both a fixed level and an intermediate level.

3.3.2   Data Analysis

Typically utilization of nursing home care is measured in patient days, called TPY. The nursing home data analyzed in Frees (2009) reports TPY for 362 facilities in year 2000 and 355 facilities in year 2001, which gives a total $n = 717$ observations, say $X_1, \cdots, X_n$; see Figure 3.1 below. This data set is available at:

http://instruction.bus.wisc.edu/jfrees/jfreesbooks/Regression%20Modeling/BookWebDec2010/

To estimate γ in Condition C3.2), we compute the Hill's estimator

$$\hat{\gamma}(m) = \{\frac{1}{m} \sum_{i=1}^{m} \log X_{n,n-i+1} - \log X_{n,n-m}\}^{-1}$$

for $m = 10, 11, \cdots, 200$ in Figure 3.1, where $X_{n,1} \leq \cdots \leq X_{n,n}$ denote the order statistics of $X_1, \cdots, X_n$. From Figure 3.1 we conclude that γ in Condition C3.2) is larger than 3 and the method is applicable with $k = 1.1$ in Condition C3.1).

Since solving the equation $\sqrt{n(1-q_n)}/F_n^{\leftarrow}(q_n) = n^d$ for $q_n = 0.7, 0.9, 0.95, 0.99$ gives negative values of d, where F_n is the empirical distribution of F defined in the beginning of Section 3.4 below, Condition C3.3) can not be satisfied for these levels q_n. Here we study the H-G risk at the above levels for the transformed data $Y_i = X_i/100$. It is easy to see that the H-G risk for X_i's is 100 times the H-G risk for Y_i's. Table 3.4 below reports the values of d, $\hat{\beta}_n^{MEL}$, $\hat{\theta}_n^{MEL}$, and the empirical VaR at level q_n, which shows that H-G risk $\hat{\theta}_n^{MEL}$ at level q_n is significantly larger than its corresponding VaR $F_n^{\leftarrow}(q_n)$.

For constructing the empirical likelihood confidence regions for the H-G risk measure at q_n, we plot the profile empirical likelihood ratio function $l_n^P(\delta \hat{\theta}_n^{MEL})$ for different δ; See Figure 3.2. Confidence regions at level 90% and 95% can be obtained from those values below the two straight lines in Figure 3.2, respectively. Figure 3.2 also shows that the plot becomes bumpier for some values of $\delta > 1$ as q_n is larger, which may be explained by the fact that the estimator for the H-G risk at a large q_n has a big variance.

3.4 Proofs

Throughout we define $F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I(X_i \leq x)$, and put

$$\alpha_n(t) = \sqrt{n}\{\frac{1}{n} \sum_{i=1}^{n} I(F(X_i) \leq t) - t\} = \sqrt{n}\{F_n(F^{\leftarrow}(t)) - t\}$$

40

and $w_n(t) = (1 - q_n)^{-1/2}\alpha_n(1 - t(1 - q_n))$. Although Theorem 2.1 of Einmahl (1992) studied the intermediate left tail process $(1 - q_n)^{-1/2}\alpha_n(t(1 - q_n))$, it holds for the intermediate right tail process $w_n(t)$ too. That is, there exists a sequence of standard Wiener processes $\{W_n(t)\}$ such that for any $\eta \in [0, 1/2)$ and $M > 0$

$$\sup_{0 < t \leq M} \frac{|w_n(t) - W_n(t)|}{t^\eta} \xrightarrow{p} 0 \quad \text{as} \quad n \to \infty. \tag{3.4.1}$$

We also have

$$\sup_{a_n \leq t \leq 1 - a_n} \frac{|F_n(F^\leftarrow(t)) - t|}{t(1 - t)} \xrightarrow{p} 0 \quad \text{when} \quad a_n \to 0 \quad \text{and} \quad na_n \to \infty. \tag{3.4.2}$$

Moreover, under conditions C3.1)–C3.3), it follows from Theorem 4.1 of Tang and Yang (2012) that

$$\lim_{n \to \infty} \frac{\theta_{n0}}{c_1 F^\leftarrow(q_n)} = 1 \quad \text{and} \quad \lim_{n \to \infty} \frac{\beta_{n0}}{c_2 F^\leftarrow(q_n)} = 1, \tag{3.4.3}$$

where c_1 and c_2 are given in Theorem 3.2.1.

Before proving our theorems, we need some lemmas.

Lemma 3.4.1. *Under conditions of Theorem 3.2.1, when* $|\beta/\beta_{n0} - 1| + |\theta/\theta_{n0} - 1| = \Delta_n \xrightarrow{p} 0$ *as* $n \to \infty$, *we have*

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^n \psi(\frac{X_i - \beta}{\theta - \beta})I(X_i > \beta) - 1 + q_n \\
&= -(1 + o_p(1))n^{-1/2}(1 - q_n)^{1/2}\int_0^{c_2^{-\gamma}} W_n(x)k(\frac{x^{-1/\gamma} - c_2}{c_1 - c_2})^{k-1}\frac{1}{c_1 - c_2}\frac{1}{\gamma x^{1+1/\gamma}}\,dx \\
&\quad + (1 + O_p(\Delta_n))\frac{1 - q_n}{F^\leftarrow(q_n)}(\beta - \beta_{n0})\int_0^{c_2^{-\gamma}} k(\frac{x^{-1/\gamma} - c_2}{c_1 - c_2})^{k-1}\frac{x^{-1/\gamma} - c_1}{(c_1 - c_2)^2}\,dx \\
&\quad + (1 + O_p(\Delta_n))\frac{1 - q_n}{F^\leftarrow(q_n)}(\theta - \theta_{n0})\int_0^{c_2^{-\gamma}} k(\frac{x^{-1/\gamma} - c_2}{c_1 - c_2})^{k-1}\frac{c_2 - x^{-1/\gamma}}{(c_1 - c_2)^2}\,dx,
\end{aligned}
\tag{3.4.4}
$$

$$\frac{1}{n} \sum_{i=1}^{n} \psi'(\tfrac{X_i-\beta}{\theta-\beta})(X_i - \theta)I(X_i > \beta)$$

$$
\begin{aligned}
= \quad & -(1+o_p(1))n^{-1/2}(1-q_n)^{1/2}F^{\leftarrow}(q_n)\int_0^{c_2^{-\gamma}} W_n(x) \\
& \times \{k^2(\tfrac{x^{-1/\gamma}-c_2}{c_1-c_2})^{k-1} - k(k-1)(\tfrac{x^{-1/\gamma}-c_2}{c_1-c_2})^{k-2}\}\tfrac{1}{\gamma x^{1+1/\gamma}}\, dx \\
& +(1+O_p(\Delta_n))(1-q_n)(\beta - \beta_{n0})\int_0^{c_2^{-\gamma}} k(k-1)(\tfrac{x^{-1/\gamma}-c_2}{c_1-c_2})^{k-2}\tfrac{(x^{-1/\gamma}-c_1)^2}{(c_1-c_2)^2}\, dx \\
& +(1+O_p(\Delta_n))(1-q_n)(\theta - \theta_{n0})\int_0^{c_2^{-\gamma}} \{k(k-1)(\tfrac{x^{-1/\gamma}-c_2}{c_1-c_2})^{k-2}\tfrac{(c_2-x^{-1/\gamma})(x^{-1/\gamma}-c_1)}{(c_1-c_2)^2} \\
& -k(\tfrac{x^{-1/\gamma}-c_2}{c_1-c_2})^{k-1}\}\, dx,
\end{aligned}
\tag{3.4.5}
$$

$$\frac{1}{n(1-q_n)}\sum_{i=1}^{n}\psi^2(\frac{X_i-\beta}{\theta-\beta})I(X_i > \beta) \xrightarrow{p} \int_0^{c_2^{-\gamma}}\psi^2(\frac{t^{-1/\gamma}-c_2}{c_1-c_2})\, dt \tag{3.4.6}$$

$$
\begin{aligned}
& \frac{1}{n(1-q_n)(F^{\leftarrow}(q_n))^2}\sum_{i=1}^{n}\{\psi'(\tfrac{X_i-\beta}{\theta-\beta})\}^2(X_i-\theta)^2 I(X_i > \beta) \\
& \xrightarrow{p} \int_0^{c_2^{-\gamma}}\{\psi'(\tfrac{t^{-1/\gamma}-c_2}{c_1-c_2})\}^2(t^{-1/\gamma}-c_1)^2\, dt,
\end{aligned}
\tag{3.4.7}
$$

$$
\begin{aligned}
& \frac{1}{n(1-q_n)F^{\leftarrow}(q_n)}\sum_{i=1}^{n}\psi(\tfrac{X_i-\beta}{\theta-\beta})\psi'(\tfrac{X_i-\beta}{\theta-\beta})(X_i-\theta)I(X_i > \beta) \\
& \xrightarrow{p} \int_0^{c_2^{-\gamma}}\psi(\tfrac{t^{-1/\gamma}-c_2}{c_1-c_2})\psi'(\tfrac{t^{-1/\gamma}-c_2}{c_1-c_2})(t^{-1/\gamma}-c_1)\, dt
\end{aligned}
\tag{3.4.8}
$$

and

$$
\begin{cases}
E\{\psi^d(\tfrac{X_i-\beta}{\theta-\beta})I(X_i > \beta)\} = O(1-q_n) \\
E\{|\psi'(\tfrac{X_i-\beta}{\theta-\beta})|^d|X_i-\theta|^d I(X_i > \beta)\} = O((1-q_n)(F^{\leftarrow}(q_n))^d)
\end{cases}
\tag{3.4.9}
$$

for any $d \in [1, \gamma/k)$.

42

Proof. By the mean value theorem, we can write

$$
\frac{1}{n}\sum_{i=1}^{n}\psi(\tfrac{X_i-\beta}{\theta-\beta})I(X_i>\beta)-1+q_n
$$

$$
=\ \int_{\beta}^{\infty}\psi(\tfrac{x-\beta}{\theta-\beta})\,dF_n(x)-\int_{\beta_{n0}}^{\infty}\psi(\tfrac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}})\,dF(x)
$$

$$
=\ \int_{\beta}^{\infty}\psi(\tfrac{x-\beta}{\theta-\beta})\,d\{F_n(x)-F(x)\}+\int_{\beta}^{\infty}\psi(\tfrac{x-\beta}{\theta-\beta})\,dF(x)-\int_{\beta_{n0}}^{\infty}\psi(\tfrac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}})\,dF(x)
$$

$$
=\ -\int_{\beta}^{\infty}\{F_n(x)-F(x)\}\psi'(\tfrac{x-\beta}{\theta-\beta})\tfrac{1}{\theta-\beta}\,dx+(\beta-\beta_{n0})\int_{\beta_1}^{\infty}\psi'(\tfrac{x-\beta_1}{\theta_1-\beta_1})\tfrac{x-\theta_1}{(\theta_1-\beta_1)^2}\,dF(x)
$$

$$
+(\theta-\theta_{n0})\int_{\beta_1}^{\infty}\psi'(\tfrac{x-\beta_1}{\theta_1-\beta_1})\tfrac{\beta_1-x}{(\theta_1-\beta_1)^2}\,dF(x) \tag{3.4.10}
$$

$$
=\ -\int_{\beta}^{X_{n,n}}\{F_n(x)-F(x)\}\psi'(\tfrac{x-\beta}{\theta-\beta})\tfrac{1}{\theta-\beta}\,dx-\int_{X_{n,n}}^{\infty}\{1-F(x)\}\psi'(\tfrac{x-\beta}{\theta-\beta})\tfrac{1}{\theta-\beta}\,dx
$$

$$
+(\beta-\beta_{n0})\int_{\beta_1}^{\infty}\psi'(\tfrac{x-\beta_1}{\theta_1-\beta_1})\tfrac{x-\theta_1}{(\theta_1-\beta_1)^2}\,dF(x)
$$

$$
+(\theta-\theta_{n0})\int_{\beta_1}^{\infty}\psi'(\tfrac{x-\beta_1}{\theta_1-\beta_1})\tfrac{\beta_1-x}{(\theta_1-\beta_1)^2}\,dF(x)
$$

$$
:=\ I_1+I_2+I_3+I_4,
$$

where $X_{n,n}=\max(X_1,\cdots,X_n)$, β_1 lies between β and β_{n0}, and θ_1 lies between θ and θ_{n0}. Define $f(x)=F'(x)$. It follows from (3.4.1), (3.4.3), Condition C3.2) and Potter's bound for regular variations (see Bingham, Goldie, and Teugels (1987)) that

$$
I_1\ =\ -\int_{1-F(\beta)}^{1-F(X_{n,n})}\{F_n(F^{\leftarrow}(1-y))-1+y\}\psi'(\tfrac{F^{\leftarrow}(1-y)-\beta}{\theta-\beta})\tfrac{1}{\theta-\beta}\left(-\tfrac{1}{f(F^{\leftarrow}(1-y))}\right)dy
$$

$$
=\int_{\frac{1-F(\beta)}{1-q_n}}^{\frac{1-F(X_{n,n})}{1-q_n}}\tfrac{1}{\sqrt{n}}\alpha_n(1-t(1-q_n))\psi'(\tfrac{F^{\leftarrow}(1-t(1-q_n))-\beta}{\theta-\beta})\tfrac{1}{\theta-\beta}\tfrac{1}{f(F^{\leftarrow}(1-t(1-q_n)))}(1-q_n)\,dt
$$

$$
=\int_{\frac{1-F(\beta)}{1-q_n}}^{\frac{1-F(X_{n,n})}{1-q_n}}n^{-1/2}(1-q_n)^{1/2}(W_n(t)+o_p(t^{\eta}))\psi'(\tfrac{F^{\leftarrow}(1-t(1-q_n))-\beta}{\theta-\beta})\tfrac{1}{\theta-\beta}\times
$$

$$
\tfrac{F^{\leftarrow}(1-t(1-q_n))}{F^{\leftarrow}(1-t(1-q_n))f(F^{\leftarrow}(1-t(1-q_n)))}(1-q_n)\,dt
$$

$$
=\int_{\frac{1-F(\beta)}{1-q_n}}^{\frac{1-F(X_{n,n})}{1-q_n}}n^{-1/2}(1-q_n)^{1/2}(W_n(t)+o_p(t^{\eta}))\psi'(\tfrac{F^{\leftarrow}(1-t(1-q_n))-\beta}{\theta-\beta})\tfrac{1}{\theta-\beta}\times
$$

$$
\tfrac{F^{\leftarrow}(1-t(1-q_n))}{\gamma\{1-F(F^{\leftarrow}(1-t(1-q_n)))\}}(1+o(1))(1-q_n)\,dt
$$

$$
=\int_{\frac{1-F(\beta)}{1-q_n}}^{\frac{1-F(X_{n,n})}{1-q_n}}n^{-1/2}(1-q_n)^{1/2}(W_n(t)+o_p(t^{\eta}))\psi'(\tfrac{F^{\leftarrow}(1-t(1-q_n))-\beta}{\theta-\beta})\tfrac{1}{\theta-\beta}\times
$$

$$
\tfrac{F^{\leftarrow}(q_n)t^{-1/\gamma}}{\gamma t(1-q_n)}(1+o(1))(1-q_n)\,dt
$$

$$
=(1+o_p(1))n^{-1/2}(1-q_n)^{1/2}\int_{c_2^{-\gamma}}^{0}W_n(t)k(\tfrac{t^{-1/\gamma}-c_2}{c_1-c_2})^{k-1}\tfrac{1}{(c_1-c_2)}\tfrac{1}{\gamma t^{1+1/\gamma}}\,dt
$$

$$
=-(1+o_p(1))n^{-1/2}(1-q_n)^{1/2}\int_{0}^{c_2^{-\gamma}}W_n(t)k(\tfrac{t^{-1/\gamma}-c_2}{c_1-c_2})^{k-1}\tfrac{1}{c_1-c_2}\tfrac{1}{\gamma t^{1+1/\gamma}}\,dt
$$

$$
\tag{3.4.11}
$$

by noting that $1 - F(X_{n,n}) = O_p(1/n)$ which implies that

$$\frac{1 - F(X_{n,n})}{1 - q_n} = O_p(\frac{1}{n(1 - q_n)}) = o_p(1). \tag{3.4.12}$$

By Potter's bound, (3.4.12) and Condition C3.3), we have for any $0 < \epsilon < \frac{1}{2} - \frac{k}{\gamma}$

$$
\begin{aligned}
I_2 &= -\int_{\frac{X_{n,n}}{F^{\leftarrow}(q_n)}}^{\infty} \frac{1 - F(F^{\leftarrow}(q_n)t)}{1 - F(F^{\leftarrow}(q_n))} (1 - q_n)\psi'(\frac{F^{\leftarrow}(q_n)t - \beta}{\theta - \beta})\frac{F^{\leftarrow}(q_n)}{\theta - \beta} \, dt \\
&= -\int_{\frac{X_{n,n}}{F^{\leftarrow}(q_n)}}^{\infty} t^{-\gamma}(1 - q_n)k(\frac{t - c_2}{c_1 - c_2})^{k-1}\frac{1}{c_1 - c_2}(1 + o(1)) \, dt \\
&= (1 - q_n)(1 + o(1))O_p((\frac{X_{n,n}}{F^{\leftarrow}(q_n)})^{-\gamma + k}) \\
&= (1 - q_n)(1 + o(1))O_p((\frac{F^{\leftarrow}(1 - (1 - q_n)\frac{1 - F(X_{n,n})}{1 - q_n})}{F^{\leftarrow}(1 - (1 - q_n))})^{-\gamma + k}) \\
&= (1 - q_n)(1 + o(1))O_p((\frac{1 - F(X_{n,n})}{1 - q_n})^{1 - k/\gamma - \epsilon}) \\
&= (1 - q_n)(1 + o(1))O_p((n(1 - q_n))^{-1 + k/\gamma + \epsilon}) \\
&= n^{-1/2}(1 - q_n)^{1/2}(1 + o(1))O_p((n(1 - q_n))^{-1/2 + k/\gamma + \epsilon}) \\
&= o_p(n^{-1/2}(1 - q_n)^{1/2}),
\end{aligned}
\tag{3.4.13}
$$

$$
\begin{aligned}
I_3 &= (\beta - \beta_{n0})\int_0^{(1 - F(\beta_1))/(1 - q_n)} \psi'(\frac{F^{\leftarrow}(1 - y(1 - q_n)) - \beta_1}{\theta_1 - \beta_1})\frac{F^{\leftarrow}(1 - y(1 - q_n)) - \theta_1}{(\theta_1 - \beta_1)^2}(1 - q_n) \, dy \\
&= (1 + O_p(\Delta_n))(\beta - \beta_{n0})\frac{1 - q_n}{F^{\leftarrow}(q_n)}\int_0^{c_2^{-\gamma}} k(\frac{y^{-1/\gamma} - c_2}{c_1 - c_2})^{k-1}\frac{y^{-1/\gamma} - c_1}{(c_1 - c_2)^2} \, dy
\end{aligned}
\tag{3.4.14}
$$

and

$$
\begin{aligned}
I_4 &= (\theta - \theta_{n0})\int_0^{(1 - F(\beta_1))/(1 - q_n)} \psi'(\frac{F^{\leftarrow}(1 - y(1 - q_n)) - \beta_1}{\theta_1 - \beta_1})\frac{\beta_1 - F^{\leftarrow}(1 - y(1 - q_n))}{(\theta_1 - \beta_1)^2}(1 - q_n) \, dy \\
&= (1 + O_p(\Delta_n))(\theta - \theta_{n0})\frac{1 - q_n}{F^{\leftarrow}(q_n)}\int_0^{c_2^{-\gamma}} k(\frac{y^{-1/\gamma} - c_2}{c_1 - c_2})^{k-1}\frac{c_2 - y^{-1/\gamma}}{(c_1 - c_2)^2} \, dy.
\end{aligned}
\tag{3.4.15}
$$

Hence (3.4.4) follows from (3.4.11) and (3.4.13)–(3.4.15).

Using the same arguments as above, we can show that (3.4.5) holds by writing

$$
\frac{1}{n} \sum_{i=1}^{n} \psi'(\frac{X_i-\beta}{\theta-\beta})(X_i-\theta)I(X_i>\beta)
$$
$$
= \int_{\beta}^{\infty} \psi'(\frac{x-\beta}{\theta-\beta})(x-\theta)\, d\{F_n(x)-F(x)\} + \int_{\beta}^{\infty} \psi'(\frac{x-\beta}{\theta-\beta})(x-\theta)\, dF(x)
$$
$$
= -\int_{\beta}^{\infty}\{F_n(x)-F(x)\}\{\psi''(\frac{x-\beta}{\theta-\beta})\frac{x-\theta}{\theta-\beta}+\psi'(\frac{x-\beta}{\theta-\beta})\}\, dx
$$
$$
+(\beta-\beta_{n0})\int_{\beta_1}^{\infty}\psi''(\frac{x-\beta_1}{\theta_1-\beta_1})\frac{(x-\theta_1)^2}{(\theta_1-\beta_1)^2}\, dF(x)
$$
$$
+(\theta-\theta_{n0})\int_{\beta_1}^{\infty}\{\psi''(\frac{x-\beta_1}{\theta_1-\beta_1})\frac{(\beta_1-x)(x-\theta_1)}{(\theta_1-\beta_1)^2}-\psi'(\frac{x-\beta_1}{\theta_1-\beta_1})\}\, dF(x)
$$

since

$$
0 = \int_{\beta_{n0}}^{\infty} \psi'(\frac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}})(x-\theta_{n0})\, dF(x).
$$

By the mean value theorem again, we can write

$$
\frac{1}{n}\sum_{i=1}^{n}\psi^2(\frac{X_i-\beta}{\theta-\beta})I(X_i>\beta)-\int_{\beta_{n0}}^{\infty}\psi^2(\frac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}})\, dF(x)
$$
$$
= -\int_{\beta}^{\infty}\{F_n(x)-F(x)\}2\psi(\frac{x-\beta}{\theta-\beta})\psi'(\frac{x-\beta}{\theta-\beta})\frac{1}{\theta-\beta}\, dx
$$
$$
+(\beta-\beta_{n0})\int_{\beta_1}^{\infty}2\psi(\frac{x-\beta_1}{\theta_1-\beta_1})\psi'(\frac{x-\beta_1}{\theta_1-\beta_1})\frac{x-\theta_1}{(\theta_1-\beta_1)^2}\, dF(x)
$$
$$
+(\theta-\theta_{n0})\int_{\beta_1}^{\infty}2\psi(\frac{x-\beta_1}{\theta_1-\beta_1})\psi'(\frac{x-\beta_1}{\theta_1-\beta_1})\frac{\beta_1-x}{(\theta_1-\beta_1)^2}\, dF(x)
$$
$$
= -\int_{\beta}^{\delta_n}\{F_n(x)-F(x)\}2\psi(\frac{x-\beta}{\theta-\beta})\psi'(\frac{x-\beta}{\theta-\beta})\frac{1}{\theta-\beta}\, dx \tag{3.4.16}
$$
$$
-\int_{\delta_n}^{X_{n,n}}\{F_n(x)-F(x)\}2\psi(\frac{x-\beta}{\theta-\beta})\psi'(\frac{x-\beta}{\theta-\beta})\frac{1}{\theta-\beta}\, dx
$$
$$
-\int_{X_{n,n}}^{\infty}\{1-F(x)\}2\psi(\frac{x-\beta}{\theta-\beta})\psi'(\frac{x-\beta}{\theta-\beta})\frac{1}{\theta-\beta}\, dx
$$
$$
+(\beta-\beta_{n0})\int_{\beta_1}^{\infty}2\psi(\frac{x-\beta_1}{\theta_1-\beta_1})\psi'(\frac{x-\beta_1}{\theta_1-\beta_1})\frac{x-\theta_1}{(\theta_1-\beta_1)^2}\, dF(x)
$$
$$
+(\theta-\theta_{n0})\int_{\beta_1}^{\infty}2\psi(\frac{x-\beta_1}{\theta_1-\beta_1})\psi'(\frac{x-\beta_1}{\theta_1-\beta_1})\frac{\beta_1-x}{(\theta_1-\beta_1)^2}\, dF(x)
$$
$$
:= II_1 + II_2 + II_3 + II_4 + II_5,
$$

where δ_n is chosen to satisfy $n\{1-F(\delta_n)\}=(n(1-q_n))^{1/2-k/\gamma}$, which implies that $\delta_n \to \infty$ and

$$
\frac{1-F(X_{n,n})}{1-F(\delta_n)}=O_p(\frac{1}{n(1-F(\delta_n))})=o_p(1), \quad \text{i.e.,} \quad \frac{X_{n,n}}{\delta_n}\xrightarrow{p}\infty. \tag{3.4.17}
$$

It follows from (3.4.2), (3.4.3), (3.4.17), Condition C3.2) with $a_n=1-F(\delta_n)$ and Potter's bound

for regular variations that

$$
\begin{aligned}
II_1 &= \int_{\frac{1-F(\beta)}{1-q_n}}^{\frac{1-F(\delta_n)}{1-q_n}} \{F_n(F^{\leftarrow}(1-(1-q_n)t)) - (1-(1-q_n)t)\} \times \\
&\quad 2\psi\left(\frac{F^{\leftarrow}(1-(1-q_n)t)-\beta}{\theta-\beta}\right)\psi'\left(\frac{F^{\leftarrow}(1-(1-q_n)t)-\beta}{\theta-\beta}\right)\frac{1}{\theta-\beta}\frac{1-q_n}{f(F^{\leftarrow}(1-(1-q_n)t))}\, dt \\
&= \int_{\frac{1-F(\beta)}{1-q_n}}^{\frac{1-F(\delta_n)}{1-q_n}} o_p(1)(1-q_n)t 2\psi\left(\frac{t^{-1/\gamma}-c_2}{c_1-c_2}\right)\psi'\left(\frac{t^{-1/\gamma}-c_2}{c_1-c_2}\right)\frac{1}{c_1-c_2}\frac{1}{\gamma t^{1+1/\gamma}}\, dt \qquad (3.4.18) \\
&= o_p(1)(1-q_n)O_p\left(\int_0^{c_2^{-\gamma}}(t^{-1/\gamma})^{2k}\, dt\right) \\
&= o_p(1-q_n),
\end{aligned}
$$

$$
\begin{aligned}
|II_2| &\leq \int_{\delta_n}^{X_{n,n}} \{1-F_n(x)+1-F(x)\}2\psi\left(\frac{x-\beta}{\theta-\beta}\right)\psi'\left(\frac{x-\beta}{\theta-\beta}\right)\frac{1}{\theta-\beta}\, dx \\
&\leq \int_{\delta_n}^{X_{n,n}} \{1-F_n(\delta_n)+1-F(\delta_n)\}2\psi\left(\frac{x-\beta}{\theta-\beta}\right)\psi'\left(\frac{x-\beta}{\theta-\beta}\right)\frac{1}{\theta-\beta}\, dx \\
&= \int_{\delta_n}^{X_{n,n}} O_p(1-F(\delta_n))2\psi\left(\frac{x-\beta}{\theta-\beta}\right)\psi'\left(\frac{x-\beta}{\theta-\beta}\right)\frac{1}{\theta-\beta}\, dx \\
&= O_p(1-F(\delta_n))\int_{\frac{1-F(X_{n,n})}{1-q_n}}^{\frac{1-F(\delta_n)}{1-q_n}} 2\psi\left(\frac{t^{-1/\gamma}-c_2}{c_1-c_2}\right)\psi'\left(\frac{t^{-1/\gamma}-c_2}{c_1-c_2}\right)\frac{1}{c_1-c_2}\frac{1}{\gamma t^{1+1/\gamma}}\, dt \\
&= O_p(1-F(\delta_n))\int_{\frac{1-F(X_{n,n})}{1-q_n}}^{\frac{1-F(\delta_n)}{1-q_n}} O((t^{-1/\gamma})^{2k}t^{-1})\, dt \\
&= O_p(1-F(\delta_n))O_p\left(\left(\frac{1-F(X_{n,n})}{1-q_n}\right)^{-2k/\gamma-\epsilon}\right) \qquad (3.4.19) \\
&= O_p((1-F(\delta_n))(n(1-q_n))^{2k/\gamma+\epsilon}) \\
&= O_p(n^{-1}(n(1-q_n))^{1/2-k/\gamma+2k/\gamma+\epsilon}) \\
&= o_p(n^{-1}n(1-q_n)) \\
&= o_p(1-q_n)
\end{aligned}
$$

for any $0 < \epsilon < \frac{1}{2} - \frac{k}{\gamma}$, and

$$
\begin{aligned}
II_3 &= \int_{\frac{1-F(X_{n,n})}{1-q_n}}^{0} (1-q_n)t 2\psi\left(\frac{F^{\leftarrow}(1-(1-q_n)t)-\beta}{\theta-\beta}\right)\psi'\left(\frac{F^{\leftarrow}(1-(1-q_n)t)-\beta}{\theta-\beta}\right)\frac{1}{\theta-\beta}\frac{1-q_n}{f(F^{\leftarrow}(1-(1-q_n)t))}\, dt \\
&= (1-q_n)O_p\left(\int_0^{\frac{1}{n(1-q_n)}} t(t^{-1/\gamma})^{2k-1}t^{-1-1/\gamma}\, dt\right) \\
&= o_p(1-q_n).
\end{aligned}
$$

$$(3.4.20)$$

Like the proofs of (3.4.14) and (3.4.15), we have

$$\begin{cases} II_4 = o_p(1 - q_n), \quad II_5 = o_p(1 - q_n) \\ \frac{1}{1-q_n} \int_{\beta_{n0}}^{\infty} \psi^2\left(\frac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}}\right) dF(x) \to \int_0^{c_2^{-\gamma}} \psi^2\left(\frac{t^{-1/\gamma}-c_2}{c_1-c_2}\right) dt. \end{cases} \tag{3.4.21}$$

Hence (3.4.6) follows from (3.4.16)–(3.4.21). Similarly we can show (3.4.7) and (3.4.8). The proof of (3.4.9) follows from the same arguments in proving (3.4.6) and (3.4.7). ∎

Lemma 3.4.2. *Put*

$$\begin{cases} g_1(x) = k\left(\frac{x^{-1/\gamma}-c_2}{c_1-c_2}\right)^{k-1} \frac{1}{c_1-c_2} \frac{1}{\gamma x^{1+1/\gamma}}, \\ g_2(x) = \left\{k^2\left(\frac{x^{-1/\gamma}-c_2}{c_1-c_2}\right)^{k-1} - k(k-1)\left(\frac{x^{-1/\gamma}-c_2}{c_1-c_2}\right)^{k-2}\right\} \frac{1}{\gamma x^{1+1/\gamma}}. \end{cases}$$

Then we have

$$E\left\{\int_0^{c_2^{-\gamma}} W_n(t)g_1(t)\,dt\right\}^2 = \int_0^{c_2^{-\gamma}} \left(\frac{t^{-1/\gamma}-c_2}{c_1-c_2}\right)^{2k} dt,$$

$$E\left\{\int_0^{c_2^{-\gamma}} W_n(t)g_2(t)\,dt\right\}^2 = \int_0^{c_2^{-\gamma}} k^2\left(\frac{t^{-1/\gamma}-c_2}{c_1-c_2}\right)^{2k-2}(t^{-1/\gamma}-c_1)^2\,dt,$$

$$E\left\{\int_0^{c_2^{-\gamma}} W_n(t)g_1(t)\,dt\right\}\left\{\int_0^{c_2^{-\gamma}} W_n(t)g_2(t)\,dt\right\}$$

$$= \int_0^{c_2^{-\gamma}} k(c_1-c_2)\left\{\left(\frac{t^{-1/\gamma}-c_2}{c_1-c_2}\right)^{2k} - \left(\frac{t^{-1/\gamma}-c_2}{c_1-c_2}\right)^{2k-1}\right\} dt.$$

Proof. Note that

$$G_1(x) = \int_x^{c_2^{-\gamma}} g_1(t)\,dt = \left(\frac{x^{-1/\gamma}-c_2}{c_1-c_2}\right)^k$$

and

$$G_2(x) = \int_x^{c_2^{-\gamma}} g_2(t)\,dt = k(c_1-c_2)\left\{\left(\frac{x^{-1/\gamma}-c_2}{c_1-c_2}\right)^k - \left(\frac{x^{-1/\gamma}-c_2}{c_1-c_2}\right)^{k-1}\right\}.$$

Hence,

$$E\{\int_0^{c_2^{-\gamma}} W_n(t)g_1(t)\, dt\}^2$$

$$= \int_0^{c_2^{-\gamma}} \int_0^{c_2^{-\gamma}} E\{W_n(t)W_n(s)\}g_1(t)g_1(s)\, dt\, ds$$

$$= \int_0^{c_2^{-\gamma}} \int_0^{c_2^{-\gamma}} (t \wedge s)g_1(t)g_1(s)\, dt\, ds$$

$$= 2\int_0^{c_2^{-\gamma}} (\int_t^{c_2^{-\gamma}} g_1(s)\, ds)tg_1(t)\, dt$$

$$= -\int_0^{c_2^{-\gamma}} t\, dG_1^2(t)$$

$$= \int_0^{c_2^{-\gamma}} G_1^2(t)\, dt,$$

$$E\{\int_0^{c_2^{-\gamma}} W_n(t)g_2(t)\, dt\}^2 = \int_0^{c_2^{-\gamma}} G_2^2(t)\, dt,$$

$$E\{\int_0^{c_2^{-\gamma}} W_n(t)g_1(t)\, dt\}\{\int_0^{c_2^{-\gamma}} W_n(t)g_2(t)\, dt\}$$

$$= -\int_0^{c_2^{-\gamma}} G_1(t)t\, dG_2(t) - \int_0^{-c_2^{-\gamma}} G_2(t)t\, dG_1(t)$$

$$= -\frac{1}{2}\int_0^{c_2^{-\gamma}} tk(c_1 - c_2)\, d(\tfrac{t^{-1/\gamma}-c_2}{c_1-c_2})^{2k}$$

$$+ \int_0^{c_2^{-\gamma}} tk(c_1 - c_2)\tfrac{k-1}{2k-1}\, d(\tfrac{x^{-1/\gamma}-c_2}{c_1-c_2})^{2k-1}$$

$$- \frac{1}{2}\int_0^{c_2^{-\gamma}} tk(c_1 - c_2)\, d(\tfrac{t^{-1/\gamma}-c_2}{c_1-c_2})^{2k}$$

$$+ \int_0^{c_2^{-\gamma}} tk(c_1 - c_2)\tfrac{k}{2k-1}\, d(\tfrac{t^{-1/\gamma}-c_2}{c_1-c_2})^{2k-1}$$

$$= \int_0^{c_2^{-\gamma}} k(c_1 - c_2)\{(\tfrac{t^{-1/\gamma}-c_2}{c_1-c_2})^{2k} - (\tfrac{t^{-1/\gamma}-c_2}{c_1-c_2})^{2k-1}\}\, dt,$$

i.e., the lemma holds. ∎

Lemma 3.4.3. *Under conditions of Theorem 3.2.1, we have*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n \widetilde{Y}_{ni}(\theta_{n0}, \beta_{n0}) \xrightarrow{d} N(0, \Sigma_0), \tag{3.4.22}$$

$$\frac{1}{n}\{\sum_{i=1}^n \widetilde{Y}_{ni}(\theta_{n0}, \beta_{n0})\widetilde{Y}_{ni}^T(\theta_{n0}, \beta_{n0})\} \xrightarrow{p} \Sigma_0 \tag{3.4.23}$$

and

$$\max_{1 \le i \le n} ||\tilde{Y}_{ni}(\theta_{n0}, \beta_{n0})|| = o_p(\sqrt{n}), \tag{3.4.24}$$

48

where $\widetilde{Y}_{ni}(\theta_{n0}, \beta_{n0}) = A_n Y_{ni}(\theta_{n0}, \beta_{n0})$,

$$A_n = \begin{pmatrix} (1-q_n)^{-1/2} & 0 \\ 0 & (1-q_n)^{-1/2}/F^{\leftarrow}(q_n) \end{pmatrix} \text{ and } \Sigma_0 \text{ is given in Theorem 3.2.1.}$$

Proof. Equations (3.4.22) and (3.4.23) follow from Lemmas 3.4.1 and 3.4.2 directly. For proving (3.4.24), write $\tilde{Y}_{ni}(\theta_{n0}, \beta_{n0}) = (\tilde{Y}_{ni,1}, \tilde{Y}_{ni,2})^T$. Choose $d = 2 + \frac{\gamma/k - 2}{2}$ and $\delta = \frac{d/2-1}{2d}$. Then it follows from (3.4.9) that

$$P(\max_{1\leq i \leq n} |\tilde{Y}_{ni,1}| > n^{1/2}(n(1-q_n))^{-\delta})$$

$$\leq \sum_{i=1}^n P(|\tilde{Y}_{ni,1}| > n^{1/2}(n(1-q_n))^{-\delta})$$

$$\leq \sum_{i=1}^n n^{-d/2}(n(1-q_n))^{\delta d} E|\tilde{Y}_{ni,1}|^d$$

$$= n^{1-d/2}(n(1-q_n))^{\delta d}(1-q_n)^{-d/2}O(1-q_n)$$

$$= O((n(1-q_n))^{1-d/2+\delta d})$$

$$= o(1),$$

i.e., $\max_{1\leq i \leq n} |\tilde{Y}_{ni,1}| = o_p(\sqrt{n})$. Similarly we have $\max_{1\leq i \leq n} |\tilde{Y}_{ni,2}| = o_p(\sqrt{n})$. Hence (3.4.24) holds. ∎

Proof of Theorem 3.2.1. Note that (3.2.2) and (3.2.3) still hold, but with a different λ_n when $Y'_{ni}s$ are replaced by $\tilde{Y}'_{ni}s$. For the simplicity of notation, we still use λ_n. Hence, like the proof of Owen (1990), it follows from Lemmas 3.4.1 and 3.4.3 that

$$\lambda_n = \{\frac{1}{n}\sum_{i=1}^n \tilde{Y}_{ni}(\theta, \beta)\tilde{Y}_{ni}^T(\theta, \beta)\}^{-1}\frac{1}{n}\sum_{i=1}^n \tilde{Y}_{ni}(\theta, \beta)(1 + o_p(1)),$$

and further

$$\begin{aligned} l_n(\theta, \beta) &= 2\sum_{i=1}^n \lambda_n^T \tilde{Y}_{ni}(\theta, \beta) - \sum_{i=1}^n \lambda_n^T \tilde{Y}_{ni}(\theta, \beta)\tilde{Y}_{ni}^T(\theta, \beta)\lambda_n + o_p(1) \\ &= \{\frac{1}{\sqrt{n}}\sum_{i=1}^n \tilde{Y}_{ni}(\theta, \beta)\}^T\{\frac{1}{n}\sum_{i=1}^n \tilde{Y}_{ni}(\theta, \beta)\tilde{Y}_{ni}^T(\theta, \beta)\}^{-1}\{\frac{1}{\sqrt{n}}\sum_{i=1}^n \tilde{Y}_{ni}(\theta, \beta)\} \\ &\quad + o_p(1) \\ &= \{\frac{1}{\sqrt{n}}\sum_{i=1}^n \widetilde{Y}_{ni}(\theta, \beta)\}^T\Sigma_0^{-1}\{\frac{1}{\sqrt{n}}\sum_{i=1}^n \widetilde{Y}_{ni}(\theta, \beta)\} + o_p(1). \end{aligned}$$

$$(3.4.25)$$

Put $\nu/\{\sqrt{n(1-q_n)}/F^{\leftarrow}(q_n)\} = (\beta - \beta_{n0}, \theta - \theta_{n0})^T$. Then it follows from (3.4.26) and

Lemmas [3.4.1](#) and [3.4.3](#) that

$$l_n(\theta, \beta) = \{\frac{1}{\sqrt{n}} \sum_{i=1}^n \widetilde{Y}_{ni}(\theta_{n0}, \beta_{n0}) + \Sigma_1 \nu\}^T \Sigma_0^{-1} \{\frac{1}{\sqrt{n}} \sum_{i=1}^n \widetilde{Y}_{ni}(\theta_{n0}, \beta_{n0}) + \Sigma_1 \nu\} + o_p(1),$$

which is minimized at

$$\Sigma_1^T \Sigma_0^{-1} \Sigma_1 \nu = -\Sigma_1^T \Sigma_0^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^n \widetilde{Y}_{ni}(\theta_{n0}, \beta_{n0}) + o_p(1). \tag{3.4.26}$$

Hence Theorem [3.2.1](#) follows from [(3.4.26)](#) and Lemma [3.4.3](#). ∎

Proof of Theorem [3.2.2](#). Put $\nu_1 / \{\sqrt{n(1 - q_n)} / F^{\leftarrow}(q_n)\} = \beta - \beta_{n0}$ and $a = (a_1, a_2)^T$. As above, we can show that

$$l_n(\theta_{n0}, \beta_{n0}) = \{\frac{1}{\sqrt{n}} \sum_{i=1}^n \widetilde{Y}_{ni}(\theta_{n0}, \beta_{n0})\}^T \Sigma_0^{-1} \{\frac{1}{\sqrt{n}} \sum_{i=1}^n \widetilde{Y}_{ni}(\theta_{n0}, \beta_{n0})\} + o_p(1)$$

and

$$l_n(\theta_{n0}, \beta) = \{\frac{1}{\sqrt{n}} \sum_{i=1}^n \widetilde{Y}_{ni}(\theta_{n0}, \beta_{n0}) + \nu_1 a\}^T \Sigma_0^{-1} \{\frac{1}{\sqrt{n}} \sum_{i=1}^n \widetilde{Y}_{ni}(\theta_{n0}, \beta_{n0}) + \nu_1 a\} + o_p(1).$$

Hence

$$
\begin{aligned}
&l_n(\theta_{n0}, \beta) - l_n(\theta_{n0}, \beta_{n0}) \\
&= \nu_1 a^T \Sigma_0^{-1} \{\frac{1}{\sqrt{n}} \sum_{i=1}^n \widetilde{Y}_{ni}(\theta_{n0}, \beta_{n0})\} + \{\frac{1}{\sqrt{n}} \sum_{i=1}^n \widetilde{Y}_{ni}(\theta_{n0}, \beta_{n0})\}^T \Sigma_0^{-1} \{\nu_1 a\} \\
&\quad + \nu_1 a^T \Sigma_0^{-1} \{\nu_1 a\} + o_p(1),
\end{aligned}
$$

which is minimized at

$$\nu_1 = \frac{-a^T \Sigma_0^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^n \widetilde{Y}_{ni}(\theta_{n0}, \beta_{n0})}{a^T \Sigma_0^{-1} a} + o_p(1),$$

50

i.e.,

$$
\begin{aligned}
& l_n^P(\theta_{n0}) \\
= & \ l_n(\theta_{n0}, \beta_{n0}) - \frac{a^T \Sigma_0^{-1} \{\frac{1}{\sqrt{n}} \sum_{i=1}^n \widetilde{Y}_{ni}(\theta_{n0},\beta_{n0})\} a^T \Sigma_0^{-1} \{\frac{1}{\sqrt{n}} \sum_{i=1}^n \widetilde{Y}_{ni}(\theta_{n0},\beta_{n0})\}}{a^T \Sigma_0^{-1} a} + o_p(1) \\
= & \ \{\tfrac{1}{\sqrt{n}} \sum_{i=1}^n \widetilde{Y}_{ni}(\theta_{n0}, \beta_{n0})\}^T \Sigma_0^{-1/2} \Sigma_0^{-1/2} \{\tfrac{1}{\sqrt{n}} \sum_{i=1}^n \widetilde{Y}_{ni}(\theta_{n0}, \beta_{n0})\} \\
& -\{\tfrac{1}{\sqrt{n}} \sum_{i=1}^n \widetilde{Y}_{ni}(\theta_{n0}, \beta_{n0})\}^T \Sigma_0^{-1/2} \frac{\Sigma_0^{-1/2} a a^T \Sigma_0^{-1/2}}{a^T \Sigma_0^{-1} a} \Sigma_0^{-1/2} \{\tfrac{1}{\sqrt{n}} \sum_{i=1}^n \widetilde{Y}_{ni}(\theta_{n0}, \beta_{n0})\} \\
= & \ \{\tfrac{1}{\sqrt{n}} \sum_{i=1}^n \widetilde{Y}_{ni}(\theta_{n0}, \beta_{n0})\}^T \Sigma_0^{-1/2} \{I_{2\times 2} - \frac{\Sigma_0^{-1/2} a a^T \Sigma_0^{-1/2}}{a^T \Sigma_0 a}\} \Sigma_0^{-1/2} \{\tfrac{1}{\sqrt{n}} \sum_{i=1}^n \widetilde{Y}_{ni}(\theta_{n0}, \beta_{n0})\} \\
& + o_p(1),
\end{aligned}
$$

where $I_{2\times 2}$ denotes the 2 by 2 identity matrix. Since $I_{2\times 2} - \frac{\Sigma_0^{-1/2} a a^T \Sigma_0^{-1/2}}{a^T \Sigma_0 a}$ is symmetric, idempotent and its trace equals to one, the theorem follows from Lemma 3.4.3. ∎

3.5 Conclusions

Although the H-G risk measure has been studied extensively in actuarial science, statistical inference remains unknown when the level $q = q_n$ depends on the sample size n and is an intermediate one, i.e., $q_n \to 1$ and $n(1 - q_n) \to \infty$ as $n \to \infty$. This chapter extends the empirical likelihood inference for a fixed level in Peng, Wang, and Zheng (2015) to an intermediate level. The proposed maximum empirical likelihood estimator is always consistent, but has a different asymptotic distribution for a fixed level and an intermediate level. The proposed empirical likelihood method provides a unified interval without knowing whether the level for computing the H-G risk measure is a fixed one or an intermediate one.

Table 3.1: *Estimation and coverage probability for Pareto distribution.* We report the mean and standard deviation in brackets for both $\hat{\theta}_n^{MEL}/\theta_{n0}$ and $\hat{\beta}_n^{MEL}/\beta_{n0}$ at different level q_n satisfying $\sqrt{n(1-q_n)}/F^{\leftarrow}(q_n) = n^d$ with $d = 2/5, 1/4, 1/8$ and $F(x) = 1 - x^{-\gamma}$ for $x \geq 1$. We also report the coverage probabilities for $I^{EL}(0.9)$ and $I^{EL}(0.95)$.

(n,d,γ)	q_n	θ_{n0}	β_{n0}	$\dfrac{\hat{\theta}_n^{MEL}}{\theta_{n0}}$	$\dfrac{\hat{\beta}_n^{MEL}}{\beta_{n0}}$	$I^{EL}(0.9)$	$I^{EL}(0.95)$
$(200,\frac{2}{5},3)$	0.4705	1.8968	1.2013	0.9995(0.0645)	1.0020(0.0206)	0.8511	0.9110
$(200,\frac{1}{4},3)$	0.7960	2.6067	1.6509	1.0008(0.1094)	1.0111(0.0770)	0.8604	0.9139
$(200,\frac{1}{8},3)$	0.9078	3.3974	2.1517	1.0018(0.1596)	1.0233(0.1212)	0.8469	0.8948
$(2000,\frac{2}{5},3)$	0.5983	2.0798	1.3172	1.0006(0.0304)	1.0018(0.0186)	0.8762	0.9356
$(2000,\frac{1}{4},3)$	0.8977	3.2817	2.0784	1.0027(0.0680)	1.0089(0.0680)	0.8723	0.9274
$(2000,\frac{1}{8},3)$	0.9673	4.7990	3.0393	1.0045(0.1101)	1.0203(0.1176)	0.8754	0.9288
$(200,\frac{2}{5},5)$	0.5309	1.4700	1.1466	1.0000(0.0326)	1.0012(0.0151)	0.8734	0.9287
$(200,\frac{1}{4},5)$	0.8493	1.8448	1.4389	1.0017(0.0640)	1.0081(0.0530)	0.8752	0.9274
$(200,\frac{1}{8},5)$	0.9415	2.2290	1.7387	1.0028(0.1000)	1.0179(0.0768)	0.8527	0.8976
$(2000,\frac{2}{5},5)$	0.6624	1.5700	1.2246	1.0002(0.0143)	1.0006(0.0093)	0.8896	0.9437
$(2000,\frac{1}{4},5)$	0.9338	2.1746	1.6962	1.0014(0.0359)	1.0056(0.0405)	0.8974	0.9461
$(2000,\frac{1}{8},5)$	0.9830	2.8528	2.2251	1.0030(0.0660)	1.0100(0.0691)	0.8893	0.9373

Table 3.2: *Estimation and coverage probability for t distribution.* We report the mean and standard deviation in brackets for both $\hat{\theta}_n^{MEL}/\theta_{n0}$ and $\hat{\beta}_n^{MEL}/\beta_{n0}$ at different level q_n satisfying $\sqrt{n(1-q_n)}/F^{\leftarrow}(q_n) = n^d$ with $d = 2/5, 1/4, 1/8$ and distribution $t(\gamma)$. We also report the coverage probabilities for $I^{EL}(0.9)$ and $I^{EL}(0.95)$.

(n,d,γ)	q_n	θ_{n0}	β_{n0}	$\dfrac{\hat{\theta}_n^{MEL}}{\theta_{n0}}$	$\dfrac{\hat{\beta}_n^{MEL}}{\beta_{n0}}$	$I^{EL}(0.9)$	$I^{EL}(0.95)$
$(200,\frac{2}{5},3)$	0.7648	1.9724	0.7557	0.9969(0.1446)	1.0057(0.1731)	0.8768	0.9301
$(200,\frac{1}{4},3)$	0.8678	2.6375	1.2963	0.9981(0.1707)	1.0134(0.1652)	0.8685	0.9229
$(200,\frac{1}{8},3)$	0.9277	3.4275	1.8805	1.0012(0.2106)	1.0388(0.2040)	0.8455	0.8972
$(2000,\frac{2}{5},3)$	0.7969	2.1361	0.8953	1.0008(0.0524)	1.0031(0.0613)	0.8840	0.9416
$(2000,\frac{1}{4},3)$	0.9213	3.3105	1.7964	1.0030(0.0795)	1.0110(0.0903)	0.8879	0.9382
$(2000,\frac{1}{8},3)$	0.9705	4.8560	2.8694	1.0077(0.1208)	1.0264(0.1572)	0.8787	0.9310
$(200,\frac{2}{5},5)$	0.7726	1.6655	0.7585	0.9989(0.1051)	1.0056(0.1669)	0.8969	0.9472
$(200,\frac{1}{4},5)$	0.8774	2.1724	1.2667	0.9997(0.1159)	1.0151(0.1466)	0.8885	0.9368
$(200,\frac{1}{8},5)$	0.9368	2.7292	1.7815	0.9996(0.1348)	1.0246(0.1366)	0.8592	0.9110
$(2000,\frac{2}{5},5)$	0.8055	1.7940	0.8924	1.0009(0.0382)	1.0027(0.0562)	0.8999	0.9504
$(2000,\frac{1}{4},5)$	0.9307	2.6495	1.7097	1.0016(0.0480)	1.0051(0.0577)	0.9007	0.9491
$(2000,\frac{1}{8},5)$	0.9768	3.6421	2.5763	1.0027(0.0689)	1.0113(0.0776)	0.9014	0.9489

Table 3.3: *Estimation and coverage probability for exponential distribution.* We report the mean and standard deviation in brackets for both $\hat{\theta}_n^{MEL}/\theta_{n0}$ and $\hat{\beta}_n^{MEL}/\beta_{n0}$ at different level q_n satisfying $\sqrt{n(1-q_n)}/F^{\leftarrow}(q_n) = n^d$ with $d = 2/5, 1/4, 1/8$ and the standard exponential distribution. We also report the coverage probabilities for $I^{EL}(0.9)$ and $I^{EL}(0.95)$.

(n,d)	q_n	θ_{n0}	β_{n0}	$\frac{\hat{\theta}_n^{MEL}}{\theta_{n0}}$	$\frac{\hat{\beta}_n^{MEL}}{\beta_{n0}}$	$I^{EL}(0.9)$	$I^{EL}(0.95)$
$(200,\frac{2}{5})$	0.6394	2.0606	0.9606	0.9977(0.0748)	1.0024(0.0969)	0.8898	0.9408
$(200,\frac{1}{4})$	0.8077	2.6895	1.5895	0.9978(0.0822)	1.0027(0.0993)	0.8964	0.9457
$(200,\frac{1}{8})$	0.9002	3.3448	2.2448	0.9987(0.0932)	1.0098(0.1126)	0.8910	0.9361
$(2000,\frac{2}{5})$	0.6937	2.2240	1.1240	0.9998(0.0256)	1.0005(0.0339)	0.9008	0.9507
$(2000,\frac{1}{4})$	0.8906	3.2529	2.1529	0.9999(0.0305)	1.0008(0.0334)	0.8963	0.9505
$(2000,\frac{1}{8})$	0.9634	4.3486	3.2486	1.0007(0.0413)	1.0033(0.0486)	0.9052	0.9551
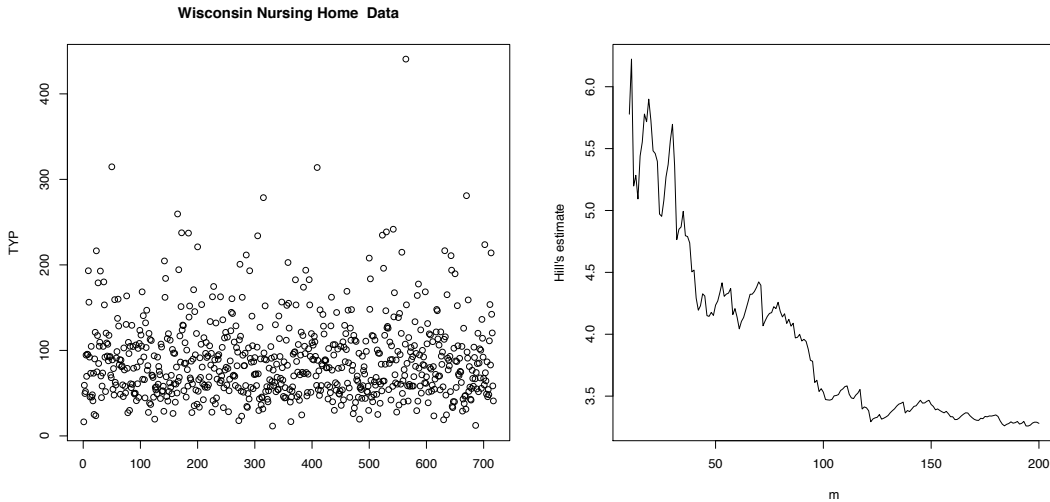
Figure 3.1: Left panel: Measures of the utilization of nursing home care in patients days for 362 facilities in year 2000 and 355 facilities in year 2001. Right panel: Hill's estimate $\hat{\gamma}(m)$ for $m = 10, 11, \cdots, 200$

Table 3.4: *Wisconsin nursing home data.* We report $\hat{\theta}_n^{MEL}$, $\hat{\beta}_n^{MEL}$, VaR $F_n^{\leftarrow}(q_n)$ and d satisfying $\sqrt{n(1-q_n)}/F_n^{\leftarrow}(q_n) = n^d$ for $q_n = 0.7, 0.9, 0.95, 0.99$, where F_n is the empirical distribution function of F.

q_n	0.7	0.9	0.95	0.99
d	0.4034	0.2639	0.1823	0.0177
$\hat{\theta}_n^{MEL}$	1.4668	1.9603	2.2821	3.1679
$\hat{\beta}_n^{MEL}$	1.0029	1.4704	1.7829	2.3859
$F_n^{\leftarrow}(q_n)$	1.0338	1.4932	1.8059	2.3842
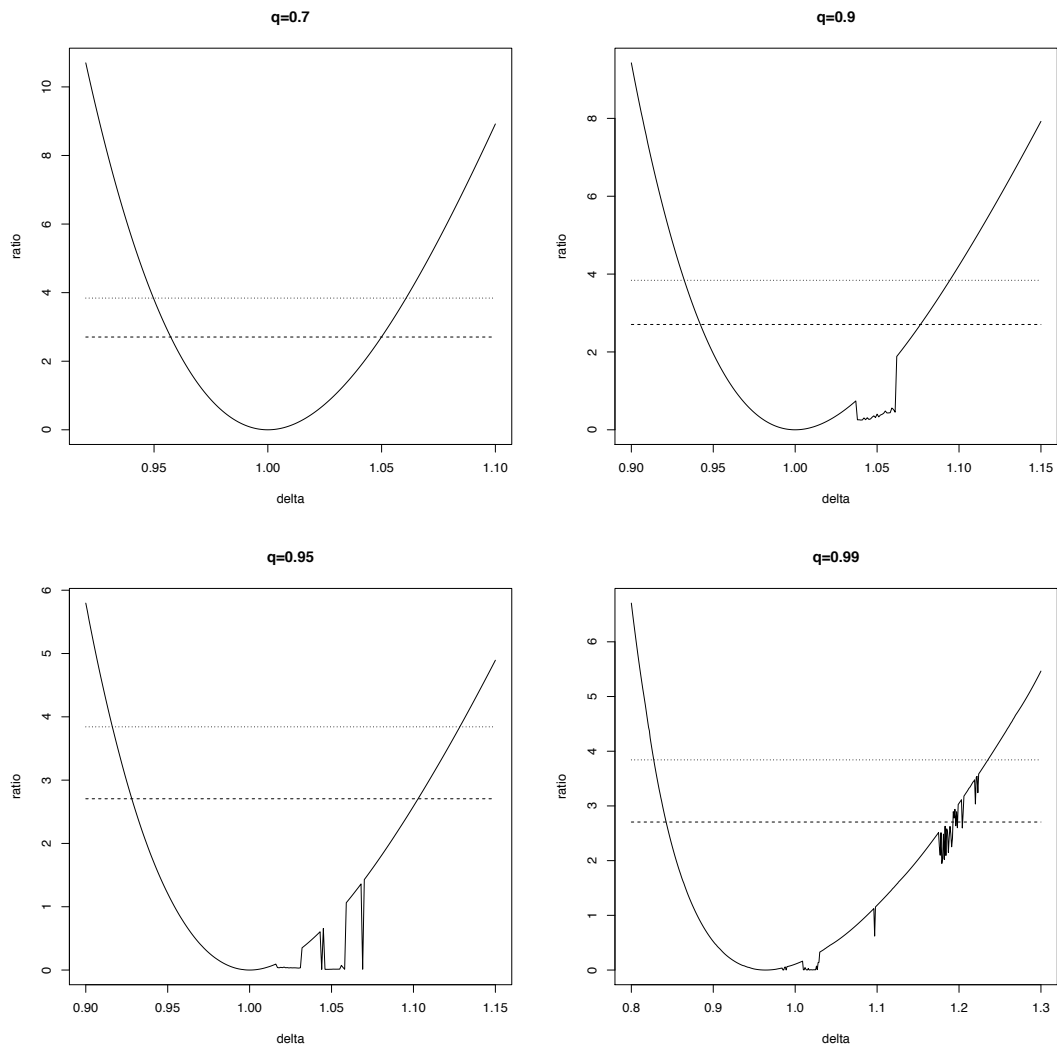
Figure 3.2: Profile empirical likelihood ratios $l_n^P(\delta\hat{\theta}_n^{MEL})$ are plotted against different values of δ for $q_n = 0.7, 0.9, 0.95, 0.99$. Two straight lines are the 90% and 95% quantiles of $\chi^2(1)$ distribution, respectively.

HAEZENDONCK-GOOVAERTS RISK MEASURE WITH A HEAVY-TAILED LOSS

Haezendonck-Goovaerts (H-G) Risk Measure depends on the involved Young function.When employing Haezendonck-Goovaerts Risk Measure to the loss variable does who not have enough moments, the nonparametric estimator in Ahn and Shyamalkumar (2014) has a non-normal limit, which challenges interval estimation. Motivated by the fact that many loss variables in insurance and finance could have a heavier tail such as an infinite variance, this chapter proposes a new estimator which estimates the tail by extreme value theory and the middle part non-parametrically. It turns out that the proposed new estimator always has a normal limit regardless of the tail heaviness of the loss variable. Hence an interval with asymptotically correct confidence level can be obtained easily either by the normal approximation method via estimating the asymptotic variance or by a bootstrap method. A simulation study and real data analysis confirm the effectiveness of the proposed new inference procedure for estimating the H-G risk measure. The content of this chapter is based on joint work:

Q. Liu, L. Peng and X. Wang (2017). Haezendonck-Goovaerts risk measure with a heavy tailed loss. Insurance: Mathematics and Economics 76, 28-47.

4.1 Motivation and Introduction

Risk management generally involves risk identification, risk quantification, and risk prediction. Measuring a risk and quantifying its uncertainty is an important task. Recently Haezendonck-Goovaerts (H-G) risk measure has received much attention in actuarial science with applications to optimal portfolio management and optimal reinsurance policy; see Bellini and Gianin (2008a); Bellini and Gianin (2008b), Cheung and Lo (2013). Zhu, Zhang, and Zhang (2013), and references therein.

Let $\psi : [0, \infty] \to [0, \infty]$ be a convex function satisfying $\psi(0) = 0, \psi(1) = 1$ and $\psi(\infty) = \infty$,

i.e., ψ is a so-called normalized Young function. For a number $q \in (0,1)$ and each $\beta > 0$, let $\alpha = \alpha(\beta)$ be a solution to

$$E\{\psi(\frac{(X-\beta)_+}{\alpha})\} = 1-q, \tag{4.1.1}$$

where $x_+ = \max(x,0)$. Then, Haezendonck and Goovaerts (1982) proposed the so-called H-G risk measure at level q as

$$\theta = \inf_{\beta>0}\{\beta + \alpha(\beta)\}. \tag{4.1.2}$$

Some important properties and connections with other risk measures are given in Goovaerts et al. (2012). For example, if $\psi(x) = x$, then $\alpha(\beta) = \frac{1}{1-q}E\{(X-\beta)_+\}$ and $\theta = \frac{1}{1-q}E\{(X - F^-(q))_+\}$, where $F(x) = P(X \le x)$ and $F^-(x)$ denotes the inverse function of $F(x)$. Hence, in this case, the H-G risk measure equals the expected shortfall.

In order to employ this risk measure in practice, an efficient statistical inference is needed. Ahn and Shyamalkumar (2014) first proposed a nonparametric estimation and derived its asymptotic limit, which may be nonnormal when the loss variable has no enough moments, which depends on the involved Young function ψ. When the limit is normal, Peng, Wang, and Zheng (2015) developed an empirical likelihood method to effectively construct an interval when the H-G risk measure is defined at a fixed level. Further,Wang and Peng (2016) showed that this empirical likelihood method is still valid for an intermediate level, which leads to a unified interval estimator of the H-G risk measure at either a fixed level or an intermediate level. We refer to Owen (2001) for an overview of empirical likelihood methods, which has been shown to be quite effective in interval estimation and hypothesis test. Properties of the H-G risk measure at an extreme level are available in Tang and Yang (2012); Tang and Yang (2014) and Mao and Hu (2012).

One can estimate β and θ by solving

$$\begin{cases} \frac{1}{n}\sum_{i=1}^n \psi(\frac{X_i-\beta}{\theta-\beta})I(X_i > \beta) = 1-q, \\ \frac{1}{n}\sum_{i=1}^n \psi'(\frac{X_i-\beta}{\theta-\beta})(X_i-\theta)I(X_i > \beta) = 0, \end{cases} \tag{4.1.3}$$

which will result in a nonnormal limit when either

$$E\{\psi(\frac{X_i - \beta_0}{\theta_0 - \beta_0})I(X_i > \beta_0)\}^2 = \infty \quad \text{or} \quad E\{\psi'(\frac{X_i - \beta_0}{\theta_0 - \beta_0})(X_i - \theta_0)I(X_i > \beta_0)\}^2 = \infty, \quad (4.1.4)$$

where θ_0 and β_0 denote the true values of θ and β, respectively. This makes interval estimation nontrivial since one has to employ different methods to separately deal with the cases of having a normal limit and a nonnormal limit.

Practically it is often observed that loss data in insurance have a heavy-tailed distribution and even have an infinite variance, which implies that (4.1.4) holds quite frequently. Particularly this chapter is motivated by analyzing the Danish fire loss data (see left panel in Figure 4.1), which consists of losses to building and losses to contents. The data were collected at the Copenhagen Reinsurance Company and comprise 2167 fire losses over the period 1980 to 1990. By assuming that

$$\lim_{t \to \infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-1/\gamma} \quad \text{for} \quad x > 0, \quad (4.1.5)$$

i.e., $1 - F$ has a heavy tail with tail index $1/\gamma$, γ can be estimated by the well-known Hill's estimator

$$\hat{\gamma}(k) = \frac{1}{k} \sum_{i=1}^{k} \log \frac{X_{n,n-i+1}}{X_{n,n-k}}, \quad (4.1.6)$$

where $X_{n,1} \leq \cdots \leq X_{n,n}$ denote the order statistics of $X_1, \cdots, X_n$, $k = k(n) \to \infty$ and $k/n \to 0$ as $n \to \infty$; see Hill (1975) for details. Note that (4.1.5) implies that $EX_+^d < \infty$ for $d < 1/\gamma$ and $EX_+^d = \infty$ for $d > 1/\gamma$. Moreover (4.1.5) holds for many commonly used loss distributions in insurance such as Pareto distribution, inverse gamma distribution, student t distribution, Cauchy distribution, Burr distribution, Log-gamma distribution, etc.. The middle and right panels in Figure 4.1 show that γ is between 0.5 and 1, which implies that $EX_+ < \infty$ but $EX_+^2 = \infty$. Therefore, when $\psi(x) = \psi_r(x) = x^r$ with some $r > 1$, the nonparametric estimator of the H-G risk measure based on (4.1.3) has a nonnormal limit, which makes interval estimation nontrivial and it generally requires a subsample bootstrap method.
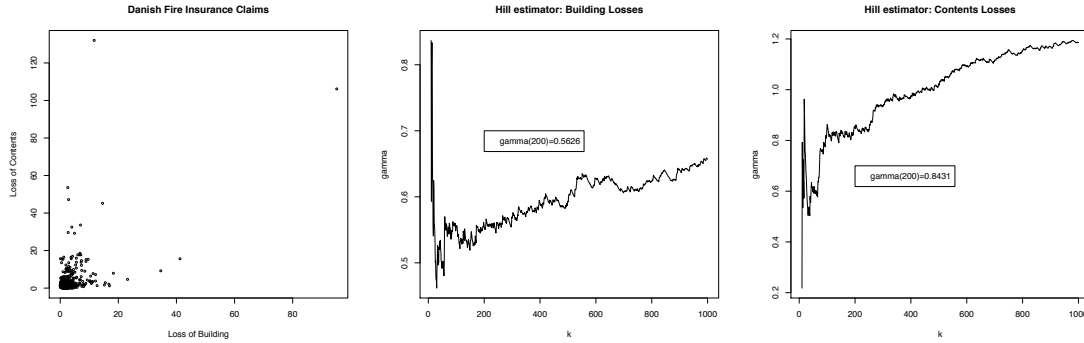
Figure 4.1: Left panel: Danish fire losses to building and contents; Middle panel: Hill's estimator for losses to building; Right panel: Hill's estimator for losses to contents.

Motivated by the idea of estimating the mean of a heavy-tailed distribution in Peng (2001); Peng (2004) and the expected shortfall of a heavy-tailed loss variable in Necir and Meraghni (2009), this chapter proposes to separately estimate the expectations in (4.1.1) by two parts: semi-parametric estimation for the tail and nonparametric estimation for the middle part. It turns out the proposed new estimator will always have a normal limit regardless of the tail heaviness of X. Hence interval estimation can be done by using either the normal approximation method via estimating the asymptotic variance or a bootstrap method. In the simulation study and data analysis below, we simply employ the naive bootstrap method, i.e., resample directly from original data, and a comparison study shows that a blind application of methods without considering a nonnormal limit would forecast risk inaccurately.

We organize this chapter as follows. Section 4.2 presents the new methodologies and main results for estimating the H-G risk measure at both a fixed level and an intermediate level. A simulation study is given in Section 4.3. Analysis of the Danish fire loss data is presented in Section 4.4. All proofs are put in Section 4.5.

4.2 Nonparametric Estimation for Heazondonck-Gooverts Risk Measure with A Heavy-Tailed Loss

Throughout we assume $X, X_1, \cdots, X_n$ are independent and identically distributed random variables with distribution function F satisfying (4.1.5), and

$$\left\{ \begin{array}{l} \psi(x) \text{ is a normalized Young function with } \psi'(0) < \infty \text{ and continuous second} \\ \text{derivatives on } (0, \infty), \text{ and satisfies } \lim_{x \to \infty} \frac{\psi''(x)}{r(r-1)x^{r-2}} = d_0 > 0 \text{ for some } r > 1, \end{array} \right. \quad (4.2.1)$$

Since we want to estimate the tail semiparametrically, it is necessary to specify an approximation rate in (4.1.5) as usual in the context of extreme value theory for controlling the bias of an employed tail probability estimator. Put $\bar{F}(x) = 1 - F(x)$ and let $\bar{F}^-(t)$ denote the inverse function of $\bar{F}(t)$. Then it is known that (4.1.5) is equivalent to

$$\lim_{t \to 0} \frac{\bar{F}^-(tx)}{\bar{F}^-(t)} = x^{-\gamma} \quad \text{for} \quad x > 0.$$

Hence we assume there exists a function $A(t) \to 0$ with a constant sign near zero such that

$$\lim_{t \to 0} \frac{\frac{\bar{F}^-(tx)}{\bar{F}^-(t)} - x^{-\gamma}}{A(t)} = x^{-\gamma} \frac{x^\rho - 1}{\rho} \quad (4.2.2)$$

for some $\rho \geq 0$. We refer to Haan and Stadtmüller (1996) for details on the second order regular variation condition (4.2.2). A well-known subclass of (4.2.2) is the so-called Hall's model (Hall (1982)):

$$1 - F(x) = cx^{-1/\gamma}\{1 + dx^{-\rho/\gamma} + o(x^{-\rho/\gamma})\}$$

for some $c > 0, d \neq 0$ and $\rho > 0$ as $x \to \infty$, which implies that

$$\bar{F}^-(t) = c^\gamma t^{-\gamma}\{1 + \gamma dc^{-\rho}t^\rho + o(t^\rho)\} \text{ as } t \to 0,$$

and (4.2.2) holds with $A(t) = \rho\gamma d c^{\gamma-\rho} t^{\rho}$. This subclass includes many commonly employed heavy-tailed loss distributions in insurance. Under condition (4.2.2), it is known that we could estimate $\bar{F}(x)$ for a larger x by $\hat{\bar{F}}(x) = \frac{k}{n} X_{n,n-k}^{1/\hat{\gamma}(k)} x^{-1/\hat{\gamma}(k)}$, where $\hat{\gamma}(k)$ is defined in (4.1.6) and $X_{n,1} \leq \cdots \leq X_{n,n}$ denote the order statistics of $X_1, \cdots, X_n$; see De Haan and Ferreira (2007) for details on consistency and asymptotic normality when $x = x(n)$ diverges. Throughout for ease of notation, we do not emphasize the dependence on n and k for estimators and some other quantities if there is no confusion.

Due to the different asymptotic behavior of estimating the H-G risk measure at a fixed level and an intermediate level, we study these two cases separately.

4.2.1 Fixed Confidence Level

In this subsection we assume the level $q \in (0, 1)$ is a fixed constant. Write

$$E\{\psi(\tfrac{X-\beta}{\theta-\beta})I(X > \beta)\}$$
$$= \int_{\beta}^{\bar{F}^-(k/n)} \psi(\tfrac{x-\beta}{\theta-\beta})\, dF(x) + \int_{\bar{F}^-(k/n)}^{\infty} \psi(\tfrac{x-\beta}{\theta-\beta})\, dF(x)$$

and

$$E\{\psi'(\tfrac{X-\beta}{\theta-\beta})(X-\theta)I(X > \beta)\}$$
$$= \int_{\beta}^{\bar{F}^-(k/n)} \psi'(\tfrac{x-\beta}{\theta-\beta})(x-\theta)\, dF(x) + \int_{\bar{F}^-(k/n)}^{\infty} \psi'(\tfrac{x-\beta}{\theta-\beta})(x-\theta)\, dF(x),$$

which motivate us to estimate θ and β by solving

$$
\begin{cases}
1-q &= \int_{\beta}^{X_{n,n-k}} \psi(\tfrac{x-\beta}{\theta-\beta})\, dF_n(x) + \int_{X_{n,n-k}}^{\infty} \psi(\tfrac{x-\beta}{\theta-\beta})\, d\{1 - \hat{\bar{F}}(x)\} \\
&:= \Delta_{11}(\theta, \beta) + \Delta_{12}(\theta, \beta) \\
0 &= \int_{\beta}^{X_{n,n-k}} \psi'(\tfrac{x-\beta}{\theta-\beta})(x-\theta)\, dF_n(x) + \int_{X_{n,n-k}}^{\infty} \psi'(\tfrac{x-\beta}{\theta-\beta})(x-\theta)\, d\{1 - \hat{\bar{F}}(x)\} \\
&:= \Delta_{21}(\theta, \beta) + \Delta_{22}(\theta, \beta),
\end{cases}
\tag{4.2.3}
$$

where $F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I(X_i \leq x)$. That is, we estimate the integral with respect to the tail semiparametrically and the integral with respect to the middle nonparametrically. Denote the resulted estimators for θ and β by $\hat{\theta}$ and $\hat{\beta}$, respectively.

Theorem 4.2.1. *Assume (4.2.2) holds with $\alpha := \frac{1}{\gamma} > r$ and $\rho > 0$, and (4.2.1) holds. Let $k = k(n)$ satisfy*

$$k = k(n) \to \infty, \quad k/n \to 0, \quad \sqrt{k}A(k/n) \to \lambda \in (-\infty, \infty) \quad as \quad n \to \infty. \tag{4.2.4}$$

Then, for a fixed $q \in (0,1)$, we have

$$\frac{\sqrt{n}}{\sigma(k/n)} \begin{pmatrix} \hat{\theta} - \theta_0 \\ \hat{\beta} - \beta_0 \end{pmatrix} \xrightarrow{d} N(C^{-1}\Delta, C^{-1}\Sigma(C^{-1})^T),$$

where $\sigma^2(k/n) = \int_{F(\beta_0)}^{1-k/n} \int_{F(\beta_0)}^{1-k/n} \{\min(s,t) - st\} \, d\psi(\frac{F^-(s)-\beta_0}{\theta_0-\beta_0}) d\psi(\frac{F^-(t)-\beta_0}{\theta_0-\beta_0})$,

$$C = \begin{pmatrix} a_1 & 0 \\ a_2 & b_2 \end{pmatrix}, \quad \Delta = \begin{pmatrix} -\sqrt{\frac{2r-\alpha}{2r}}I(\alpha < 2r)\frac{r\alpha^2\lambda}{(r-\alpha)^2(1+\rho)} \\ -\sqrt{\frac{2r-\alpha}{2r}}I(\alpha < 2r)\frac{\alpha^2 r^2\lambda}{(r-\alpha)^2(1+\rho)} \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix},$$

with

$$a_1 = E\{\psi'(\frac{X-\beta_0}{\theta_0-\beta_0})\frac{X-\beta_0}{(\theta_0-\beta_0)^2}I(X > \beta_0)\},$$

$$a_2 = E\{\psi''(\frac{X-\beta_0}{\theta_0-\beta_0})\frac{(X-\beta_0)(X-\theta_0)}{(\theta_0-\beta_0)^3}I(X > \beta_0) + \psi'(\frac{X-\beta_0}{\theta_0-\beta_0})\frac{I(X > \beta_0)}{\theta_0-\beta_0}\},$$

$$b_2 = -E\{\psi''(\frac{X-\beta_0}{\theta_0-\beta_0})\frac{(X-\theta_0)^2}{(\theta_0-\beta_0)^3}I(X > \beta_0)\},$$

$$\sigma_1^2 = 1 + (2r - \alpha)\{\frac{r^3}{2(r-\alpha)^4} + \frac{\alpha r}{(\alpha-r)^3} + \frac{1}{\alpha-r}\}I(\alpha < 2r),$$

$$\sigma_{12} = \sigma_{21} = r + r(2r - \alpha)\{\frac{r^3}{2(r-\alpha)^4} + \frac{\alpha r}{(\alpha-r)^3} + \frac{1}{\alpha-r}\}I(\alpha < 2r),$$

$$\sigma_2^2 = r^2 + r^2(2r - \alpha)\{\frac{r^3}{2(r-\alpha)^4} + \frac{\alpha r}{(\alpha-r)^3} + \frac{1}{\alpha-r}\}I(\alpha < 2r).$$

Remark 4.2.2. *A theoretical optimal choice of k is to minimize the asymptotic mean squared error in the above limit. Since the assumption (4.2.4) is the standard one for tail index estimation, one could simply employ an existing data-driven method in choosing k for Hill's estimator such as*

the bootstrap method in Danielsson et al. (2001) or the method in Drees and Kaufmann (1998) .
Another commonly used technique is to plot the estimator against different $k's$ and pick up a k in a
relatively stable region.

Remark 4.2.3. *For constructing a confidence interval or region, one needs to choose a smaller*
k such that λ in (4.2.4) is zero, and then estimates a_1, a_2, b_1, b_2 by replacing the expectations by
their corresponding averages and replacing θ_0, β_0 by $\hat{\theta}$ and $\hat{\beta}$, and estimates Σ by replacing α by
$\hat{\alpha}$. Alternatively one may simply use a bootstrap method to estimate the asymptotic variance as
we do in the simulation study and data analysis. Like any inference for extreme value statistics,
it is extremely challenging to develop a data-driven method for choosing k in terms of coverage
probability. Although bootstrap method is not applicable to maximum/minimum, it is generally valid
for an extreme value statistic which involves an upper k order statistics with $k = k(n) \to \infty$ as
$n \to \infty$ when the asymptotic bias in the statistic is negligible; see Li and Peng (2012) and Qi (2008)
for the validation of applying a bootstrap method to extreme value statistics.

4.2.2   Intermediate Quantile

In this subsection we consider the case of intermediate quantile, i.e.,

$$q = q_n \to 1 \quad \text{and} \quad n(1 - q_n) \to \infty \quad \text{as} \quad n \to \infty. \tag{4.2.5}$$

Note that the above conditions imply that the number of observations above β_0 becomes smaller by
the first condition, but it tends to infinity by the second equation. Therefore, one could expect the
estimation procedure for a fixed level should still be valid for such an intermediate level.

To emphasize the dependence on the sample size n, write $(\theta_{n0}, \beta_{n0})^T$ as the true value of $(\theta, \beta)^T$,
which is determined by the following equations with $q = q_n$ as derived before:

$$\begin{cases} E\{\psi(\frac{X_i - \beta_{n0}}{\theta_{n0} - \beta_{n0}})I(X_i > \beta_{n0})\} = 1 - q_n, \\ E\{\psi'(\frac{X_i - \beta_{n0}}{\theta_{n0} - \beta_{n0}})(X_i - \theta_{n0})I(X_i > \beta_{n0})\} = 0. \end{cases} \tag{4.2.6}$$

Using (4.2.1), like the proof of Corollary 6.1 of Tang and Yang (2014), we have

$$\lim_{n\to\infty} \frac{\theta_{n0}}{c_1 F^-(q_n)} = 1 \text{ and } \lim_{n\to\infty} \frac{\beta_{n0}}{c_2 F^-(q_n)} = 1, \tag{4.2.7}$$

where

$$c_2 = \left(\int_0^\infty (1 + \frac{\gamma}{\lambda} x)^{-1/\gamma} \, d\psi(x) \right)^\gamma, \quad c_1 = (1 + \frac{\gamma}{\lambda})c_2, \text{ and } \lambda(> 0) \text{ satisfies}$$

$$\int_0^\infty \psi'(\lambda y)(1 + \gamma y)^{-1/\gamma - 1} \, dy = \int_0^\infty \psi'(\lambda y)\lambda y(1 + \gamma y)^{-1/\gamma - 1} \, dy.$$

Following the same arguments in Section 4.2.1, our new estimator $(\hat{\theta}, \hat{\beta})^T$ is defined to solve

$$
\begin{cases}
1 - q_n &= \int_\beta^{X_{n,n-k}} \psi(\frac{x-\beta}{\theta-\beta}) \, dF_n(x) + \int_{X_{n,n-k}}^\infty \psi(\frac{x-\beta}{\theta-\beta}) \, d\{1 - \hat{\bar{F}}(x)\} \\
&:= \Delta_{I11}(\theta, \beta) + \Delta_{I12}(\theta, \beta), \\
0 &= \int_\beta^{X_{n,n-k}} \psi'(\frac{x-\beta}{\theta-\beta})(x - \theta) \, dF_n(x) + \int_{X_{n,n-k}}^\infty \psi'(\frac{x-\beta}{\theta-\beta})(x - \theta) \, d\{1 - \hat{\bar{F}}(x)\} \\
&:= \Delta_{I21}(\theta, \beta) + \Delta_{I22}(\theta, \beta).
\end{cases}
\tag{4.2.8}
$$

Theorem 4.2.4. *Assume (4.2.2) holds with $\alpha := \frac{1}{\gamma} > r$ and $\rho > 0$, and (4.2.1) holds. Consider the intermediate quantile*

$$q_n \to 1 \quad \text{and} \quad n(1 - q_n) \to \infty. \tag{4.2.9}$$

Further assume for n large enough

$$n^{-\gamma}\{n(1 - q_n)\}^{1+\gamma - r\gamma I(2r\gamma>1) - \frac{1}{2}I(2r\gamma\leq 1)} \geq n^\delta \quad \text{for some} \quad \delta > 0, \tag{4.2.10}$$

and let $k = k(n)$ satisfy

$$k = k(n) \to \infty, \quad k/n \to 0, \quad \sqrt{k}A(k/n) \to \lambda \in (-\infty, \infty), \quad \frac{k}{n(1 - q_n)} \to 0. \tag{4.2.11}$$

Then

$$\frac{\sqrt{n}}{\tilde{\sigma}(k/n)}\frac{1-q_n}{\overline{F}^-(1-q_n)}\begin{pmatrix}\hat{\theta}-\theta_0\\\hat{\beta}-\beta_0\end{pmatrix}\xrightarrow{d}N(C_I^{-1}\Delta_I,C_I^{-1}\Sigma_I(C_I^{-1})^T),$$

where $\tilde{\sigma}^2(k/n)=\int_{F(\beta_{n0})}^{1-k/n}\int_{F(\beta_{n0})}^{1-k/n}\{\min(s,t)-st\}\,d\psi(\frac{F^-(s)-\beta_{n0}}{\theta_{n0}-\beta_{n0}})d\psi(\frac{F^-(t)-\beta_{n0}}{\theta_{n0}-\beta_{n0}}),$

$$C_I=\begin{pmatrix}a_{I1}&0\\a_{I2}&b_{I2}\end{pmatrix},\quad\Sigma_I=\begin{pmatrix}\sigma_{I1}^2&\sigma_{I12}\\\sigma_{I21}&\sigma_{I2}^2\end{pmatrix},$$

$$\Delta_I=\begin{pmatrix}-\sqrt{\frac{2r-\alpha}{2r}}I(\alpha<2r)\frac{r\alpha^2\lambda}{(\alpha-r)^2(1+\rho)}\\-\sqrt{\frac{2r-\alpha}{2r}}I(\alpha<2r)\frac{r^2\alpha^2\lambda}{(r-\alpha)^2(1+\rho)}\end{pmatrix},$$

with

$$a_{I1}=\alpha c_2^{-\alpha-1}\int_0^\infty(1+\frac{c_1-c_2}{c_2}x)^{-\alpha-1}\psi'(x)\,dx,$$

$$a_{I2}=\alpha c_2^{-\alpha-1}\int_0^\infty(1+\frac{c_1-c_2}{c_2}x)^{-\alpha-1}\{x(x-1)\psi''(x)+\psi'(x)\}\,dx,$$

$$b_{I2}=-\alpha c_2^{-\alpha-1}\int_0^\infty(x-1)^2\psi''(x)(1+\frac{c_1-c_2}{c_2}x)^{-\alpha-1}\,dx;$$

$$\sigma_{I1}^2=1+(2r-\alpha)\{\frac{r^3}{2(r-\alpha)^4}-\frac{\alpha r}{(r-\alpha)^3}-\frac{1}{r-\alpha}\}I(\alpha<2r),$$

$$\sigma_{I2}^2=r^2I(\alpha\leq2r)+r^2(2r-\alpha)\{\frac{r^3}{2(r-\alpha)^4}-\frac{\alpha r}{(r-\alpha)^3}-\frac{1}{r-\alpha}\}I(\alpha<2r)+\frac{d_2}{d_1}I(\alpha>2r),$$

$$\sigma_{I12}=\sigma_{I21}=rI(\alpha\leq2r)+r(2r-\alpha)\{\frac{r^3}{2(r-\alpha)^4}-\frac{\alpha r}{(r-\alpha)^3}-\frac{1}{r-\alpha}\}I(\alpha<2r)+\frac{d_3}{d_1}I(\alpha>2r),$$

where d_1, d_2 *and* d_3 *are defined in the proof.*

Remark 4.2.5. *It follows from Lemma 4.5.4 in Section 4.5 that (4.2.10) ensures that the normalizing constant* $\frac{\sqrt{n}}{\tilde{\sigma}(k/n)}\frac{1-q_n}{\overline{F}^-(1-q_n)}\to\infty$ *for k small enough, i.e., the proposed estimators are consistent.*

Remark 4.2.6. *If* $n(1-q_n)=d_1^*n^{\delta_1}$ *for some* $\delta_1\in(0,1)$ *and* $d_1^*>0$, *then we could choose* $k=d_2^*n^{\delta_2}$ *for some* $d_2^*>0$ *and* $\delta_2\in(0,\delta_1)$. *To ensure* $\sqrt{k}A(k/n)\to\lambda$, *we need* $\delta_2\leq\frac{2\rho}{1+2\rho}$. *To ensure* $\frac{\sqrt{n}}{\tilde{\sigma}(k/n)}\frac{1-q_n}{\overline{F}^-(1-q_n)}\to\infty$, *it follows from Lemma 4.5.4 in Section 4.5 that we need* $\delta_2<\frac{-\gamma+(1+\gamma-r\gamma)\delta_1}{1/2-r\gamma}$

for $\alpha < 2r$. In summary, we require

$$\delta_2 < I(\alpha < 2r) \min(\delta_1, \frac{-\gamma + (1 + \gamma - r\gamma)\delta_1}{1/2 - r\gamma}) + I(\alpha \geq 2r)\delta_1 \quad and \quad \delta_2 \leq \frac{2\rho}{1 + 2\rho}.$$

As in extreme value statistics, it is always challenging to choose an optimal k in terms of either mean squared error or coverage probability, which we will not address here. Instead we plot our estimator or coverage probability against different k in the simulation study and data analysis below.

4.3 Simulation Study

We conduct a small scale simulation study to evaluate the finite-sample performance of the proposed method with a comparison with the nonparametric estimator proposed by Ahn and Shyamalkumar (2014), which indeed solves equations (4.2.3) with $X_{n,n-k}$ replaced by ∞. We denote this estimator by $\hat{\theta}^{AH}$.

We draw $1,000$ random samples from t-distribution with degrees of freedom 1.5, 2.3 and 5 with sample size $n = 2000$ and 4000. Take $\psi(x) = \psi_r(x) = x^r$ with $r = 1.1$ and consider $q = 0.9$. We compute the coverage probability of the constructed confidence interval with level 0.9 by using the bootstrap method with 1000 repetitions. We plot the means and standard deviations of ratios of estimators to true values and coverage probabilities of intervals for $\hat{\theta}$ and $\hat{\theta}^{AH}$ in Figures 4.2–4.4 respectively, where $k = 0$ represents results for $\hat{\theta}^{AH}$. From these figures, we observe that the proposed estimator performs better than the estimator in Ahn and Shyamalkumar (2014) when $X_i \sim t(1.5)$ and $X_i \sim t(2.3)$. Note that $\hat{\theta}^{AH}$ has a nonnormal limit in case of $X_i \sim t(1.5)$ and has a normal limit in case of $X_i \sim t(2.3)$. When $X_i \sim t(5)$, the new estimator has a similar standard deviation as $\hat{\theta}^{AH}$, but a slightly larger bias and the new estimator prefers a smaller k since the tail part does not play a role theoretically in this case and larger k introduces a big bias.

In summary, the proposed new estimator performs well regardless of the tail heaviness of the loss variable.

4.4 Data Analysis

We apply the proposed estimation procedure to the Danish fire losses discussed in the introduction, which has an infinite variance for both losses to contents and losses to building. For comparison, we also compute the estimator in Ahn and Shyamalkumar (2014) denoted by $\hat{\theta}^{AH}$ as above. Like the above simulation study, we simply employ the bootstrap method with $1,000$ repetitions to construct confidence intervals for θ and β respectively. Note that these intervals based on $\hat{\theta}^{Ah}$ are theoretically incorrect due to nonnormal limits.

For computing the proposed estimators at levels $q = 0.9$ and $q = 0.95$ and constructing confidence intervals with level 90%, we follow the standard practice in extreme value statistics by using different k from 100 to 200 with step 5. How to choose an optimal k in terms of either mean squared error or coverage probability is quite challenging and beyond the scope of this chapter. By taking $\psi(x) = \psi_r(x) = x^r$ with $r = 1.1$, the estimators and corresponding intervals are plotted in Figure 4.5, where we use $k = 0$ to denote results with respect to $\hat{\theta}^{AH}$. Panels with $q = 0.95$ show that the intervals based on the new estimators are quite different from those based on $\hat{\theta}^{AH}$ especially for the losses to contents, which has a heavier tail. Therefore a blind application of a method without considering the fact of having a nonnormal limit may lead to an under-predicted risk, i.e., smaller θ.

4.5 Proofs

Since the main idea in this chapter is to estimate the tail semi-parametrically and the middle nonparametrically, a key technique employed to link both parts is the following approximations for the empirical distribution process and empirical quantile process:

$$\sup_{U_{n,1} \leq u \leq U_{n,n}} \frac{n^{\delta}|\alpha_n(u) - B_n(u)|}{u^{1/2-\delta}(1-u)^{1/2-\delta}} = O_p(1) \tag{4.5.1}$$

and

$$\sup_{\lambda/n \le s \le 1-\lambda/n} \frac{n^\delta |\beta_n(s) + B_n(s)|}{s^{1/2-\delta}(1-s)^{1/2-\delta}} = O_p(1) \tag{4.5.2}$$

for any $\delta \in [0, 1/4)$, where $\{B_n(s)\}$ is a sequence of Brownian bridges, $G_n(u) = \frac{1}{n}\sum_{i=1}^n I(U_i \le u)$, $\alpha_n(u) = \sqrt{n}\{G_n(u) - u\}$, $Q_n(s) = U_{n,k}$ if $\frac{k-1}{n} < s \le \frac{k}{n}$, $Q_n(0) = U_{n,1}$, $\beta_n(s) = \sqrt{n}\{Q_n(s) - s\}$, $U_i = F(X_i)$ for $i = 1, \cdots, n$ and $U_{n,1} \le \cdots \le U_{n,n}$ denote the order statistics of $U_1, \cdots, U_n$; see Csorgo et al. (1986) for details.

Before we prove Theorem 4.2.1, we need some lemmas.

Lemma 4.5.1. *Under conditions of Theorem 4.2.1, we have*

$$\lim_{n\to\infty} \frac{\sqrt{k/n}\{\bar{F}^-(k/n)\}^r}{\sigma(k/n)} = \frac{(\theta_0 - \beta_0)^r}{d_0}\sqrt{\frac{2r-\alpha}{2r}}I(\alpha < 2r).$$

Proof. Note that (4.2.1) implies that

$$\lim_{x\to\infty} \frac{\psi(x)}{x^r} = \lim_{x\to\infty} \frac{\psi'(x)}{rx^{r-1}} = d_0.$$

Write

$$\begin{aligned}
\sigma^2(k/n) &= 2\int_{\beta_0}^{\bar{F}^-(k/n)} \int_{\beta_0}^s F(t)\bar{F}(s)\, d\psi(\tfrac{t-\beta_0}{\theta_0-\beta_0})d\psi(\tfrac{s-\beta_0}{\theta_0-\beta_0}) \\
&= 2\int_{\beta_0}^{\bar{F}^-(k/n)} \bar{F}(s)\psi(\tfrac{s-\beta_0}{\theta_0-\beta_0})\, d\psi(\tfrac{s-\beta_0}{\theta_0-\beta_0}) \\
&\quad -2\int_{\beta_0}^{\bar{F}^-(k/n)} \int_{\beta_0}^s \bar{F}(t)\bar{F}(s)\, d\psi(\tfrac{t-\beta_0}{\theta_0-\beta_0})d\psi(\tfrac{s-\beta_0}{\theta_0-\beta_0}) \\
&=: I_1 - I_2.
\end{aligned}$$

When $\alpha < 2r$, write

$$\begin{aligned}
I_1 &= 2\frac{k}{n}\int_{\beta_0}^{\bar{F}^-(k/n)} \frac{\bar{F}\{s\}}{\bar{F}\{\bar{F}^-(k/n)\}}\psi(\tfrac{s-\beta_0}{\theta_0-\beta_0})\, d\psi(\tfrac{s-\beta_0}{\theta_0-\beta_0}) \\
&= 2\frac{k}{n}\int_{\beta_0}^{\bar{F}^-(k/n)} \{\tfrac{s}{\bar{F}^-(k/n)}\}^{-\alpha}\psi(\tfrac{s-\beta_0}{\theta_0-\beta_0})\, d\psi(\tfrac{s-\beta_0}{\theta_0-\beta_0}) \\
&\quad +2\frac{k}{n}\int_{\beta_0}^{\bar{F}^-(k/n)} \left(\frac{\bar{F}\{s\}}{\bar{F}\{\bar{F}^-(k/n)\}} - \{\tfrac{s}{\bar{F}^-(k/n)}\}^{-\alpha}\right)\psi(\tfrac{s-\beta_0}{\theta_0-\beta_0})\, d\psi(\tfrac{s-\beta_0}{\theta_0-\beta_0}) \\
&=: I_{1,1} + I_{1,2}.
\end{aligned}$$

It is easy to see that

$$\lim_{n\to\infty} \frac{I_{1,1}}{(k/n)\{\bar{F}^-(k/n)\}^{2r}} = \lim_{n\to\infty} \frac{2}{\theta_0-\beta_0} \frac{\int_{\beta_0}^{\bar{F}^-(k/n)} s^{-\alpha}\psi(\frac{s-\beta_0}{\theta_0-\beta_0})\psi'(\frac{s-\beta_0}{\theta_0-\beta_0})\,ds}{\{\bar{F}^-(k/n)\}^{2r-\alpha}}$$

$$= (\theta_0-\beta_0)^{-2r}\frac{2rd_0^2}{2r-\alpha}.$$

For $I_{1,2}$, it follows from Theorem B.1.10 of De Haan and Ferreira (2007) and a similar argument as the last equation above that

$$\lim_{n\to\infty} \frac{I_{1,2}}{(k/n)\{\bar{F}^-(k/n)\}^{2r}} = 0.$$

Since $\alpha > r$, we have

$$|I_2| \le 2\{\int_{\beta_0}^{\infty} \bar{F}(s)\,d\psi(\frac{s-\beta_0}{\theta_0-\beta_0})\}^2 < \infty,$$

which implies that

$$\lim_{n\to\infty} \frac{|I_2|}{(k/n)\{\bar{F}^-(k/n)\}^{2r}} = 0,$$

i.e., the lemma holds for $\alpha < 2r$.

When $\alpha = 2r$, the lemma follows from the facts that $\lim_{n\to\infty} \sqrt{k/n}\{\bar{F}^-(k/n)\}^r$ is a constant and $\lim_{n\to\infty} \sigma^2(k/n) = \infty$. When $\alpha > 2r$, the lemma follows from $\lim_{n\to\infty} \sqrt{k/n}\{\bar{F}^-(k/n)\}^r = 0$ and $\lim_{n\to\infty} \sigma^2(k/n)$ is a positive constant. Hence the lemma follows. ∎

Lemma 4.5.2. *Under conditions of Theorem 4.2.1, we have*

$$\frac{\sqrt{n}}{\sigma(k/n)}\{\Delta_{11}(\theta_0,\beta_0) - \int_{\beta_0}^{\bar{F}^-(k/n)} \psi(\frac{x-\beta_0}{\theta_0-\beta_0})\,dF(x)\}$$
$$= -\frac{\int_{F(\beta_0)}^{1-k/n} B_n(s)\,d\psi(\frac{F^-(s)-\beta_0}{\theta_0-\beta_0})}{\sigma(k/n)} + o_p(1), \tag{4.5.3}$$

$$\frac{\sqrt{n}}{\sigma(k/n)}\{\Delta_{12}(\theta_0,\beta_0) - \int_{\bar{F}^-(k/n)}^{\infty} \psi(\frac{x-\beta_0}{\theta_0-\beta_0})\,dF(x)\}$$
$$= -\sqrt{\frac{2r-\alpha}{2r}}I(\alpha<2r)\frac{r\alpha}{(r-\alpha)^2}\sqrt{\frac{n}{k}}\int_0^1\{\frac{B_n(1-\frac{k}{n}s)}{s} - B_n(1-\frac{k}{n})\}ds$$
$$+\sqrt{\frac{2r-\alpha}{2r}}I(\alpha<2r)\frac{r}{r-\alpha}\sqrt{\frac{n}{k}}B_n(1-\frac{k}{n}) \tag{4.5.4}$$
$$-\sqrt{\frac{2r-\alpha}{2r}}I(\alpha<2r)\frac{r\alpha^2\lambda}{(\alpha-r)^2(1+\rho)} + o_p(1),$$

$$\frac{\partial\Delta_{11}(\theta_0,\beta_0)}{\partial\theta} = E\{-\psi'(\frac{X-\beta_0}{\theta_0-\beta_0})\frac{X-\beta_0}{(\theta_0-\beta_0)^2}I(X>\beta_0)\} + o_p(1), \tag{4.5.5}$$

$$\frac{\partial \Delta_{11}(\theta_0, \beta_0)}{\partial \beta} = E\{\psi'(\frac{X-\beta_0}{\theta_0-\beta_0})\frac{X-\theta_0}{(\theta_0-\beta_0)^2}I(X>\beta_0)\} + o_p(1) = o_p(1), \tag{4.5.6}$$

$$\frac{\partial \Delta_{12}(\theta_0, \beta_0)}{\partial \theta} = o_p(1), \quad \frac{\partial \Delta_{12}(\theta_0, \beta_0)}{\partial \beta} = o_p(1). \tag{4.5.7}$$

Proof. Write

$$
\begin{aligned}
& \int_{\beta_0}^{X_{n,n-k}} \psi(\tfrac{x-\beta_0}{\theta_0-\beta_0})\, dF_n(x) \\
={}& -\psi(\tfrac{X_{n,n-k}-\beta_0}{\theta_0-\beta_0})\bar{F}_n(X_{n,n-k}) + \int_{\beta_0}^{X_{n,n-k}} \bar{F}_n(x)\, d\psi(\tfrac{x-\beta_0}{\theta_0-\beta_0}) \\
={}& -\psi(\tfrac{X_{n,n-k}-\beta_0}{\theta_0-\beta_0})\tfrac{k}{n} + \int_{\beta_0}^{X_{n,n-k}} \{\bar{F}_n(x)-\bar{F}(x)\}\, d\psi(\tfrac{x-\beta_0}{\theta_0-\beta_0}) + \int_{\beta_0}^{X_{n,n-k}} \bar{F}(x)\, d\psi(\tfrac{x-\beta_0}{\theta_0-\beta_0}) \\
={}& -\tfrac{k}{n}\{\psi(\tfrac{X_{n,n-k}-\beta_0}{\theta_0-\beta_0}) - \psi(\tfrac{\bar{F}^-(k/n)-\beta_0}{\theta_0-\beta_0})\} - \tfrac{k}{n}\psi(\tfrac{\bar{F}^-(k/n)-\beta_0}{\theta_0-\beta_0}) \\
& + \int_{\beta_0}^{X_{n,n-k}} \{\bar{F}_n(x)-\bar{F}(x)\}\, d\psi(\tfrac{x-\beta_0}{\theta_0-\beta_0}) + \int_{\beta_0}^{\bar{F}^-(k/n)} \bar{F}(x)\, d\psi(\tfrac{x-\beta_0}{\theta_0-\beta_0}) \\
& + \int_{\bar{F}^-(k/n)}^{X_{n,n-k}} \bar{F}(x)\, d\psi(\tfrac{x-\beta_0}{\theta_0-\beta_0}) \\
={}& \int_{\beta_0}^{\bar{F}^-(k/n)} \psi(\tfrac{x-\beta_0}{\theta_0-\beta_0})\, dF(x) - \tfrac{k}{n}\{\psi(\tfrac{X_{n,n-k}-\beta_0}{\theta_0-\beta_0}) - \psi(\tfrac{\bar{F}^-(k/n)-\beta_0}{\theta_0-\beta_0})\} \\
& - \int_{F(\beta_0)}^{1-k/n} \{G_n(x)-x\}\, d\psi(\tfrac{F^-(x)-\beta_0}{\theta_0-\beta_0}) - \int_{1-k/n}^{U_{n,n-k}} \{G_n(x)-x\}\, d\psi(\tfrac{F^-(x)-\beta_0}{\theta_0-\beta_0}) \\
& + \int_{\bar{F}^-(k/n)}^{X_{n,n-k}} \bar{F}(x)\, d\psi(\tfrac{x-\beta_0}{\theta_0-\beta_0}) \\
={}& I_1 + \cdots + I_5.
\end{aligned}
\tag{4.5.8}
$$

It follows from (4.5.2) that

$$\sqrt{k}\{\frac{n}{k}(1 - U_{n,n-k}) - 1\} = \sqrt{\frac{n}{k}}B_n(1-\frac{k}{n}) + o_p(1). \tag{4.5.9}$$

Hence, by (4.5.9) and Lemma 4.5.1,

$$
\begin{aligned}
\frac{\sqrt{n}I_2}{\sigma(k/n)} &= -\frac{\sqrt{n}}{\sigma(k/n)}\frac{k}{n}\psi'\left(\frac{\bar{F}^-(k/n)-\beta_0}{\theta_0-\beta_0}\right)\frac{X_{n,n-k}-\bar{F}^-(k/n)}{\theta_0-\beta_0}\{1+o_p(1)\} \\
&= -rd_0(\theta_0-\beta_0)^{-r}\frac{\sqrt{n}}{\sigma(k/n)}\frac{k}{n}\{\bar{F}^-(k/n)\}^r\left(\frac{X_{n,n-k}}{\bar{F}^-(k/n)}-1\right)\{1+o_p(1)\} \\
&= -rd_0(\theta_0-\beta_0)^{-r}\frac{\sqrt{n}}{\sigma(k/n)}\frac{k}{n}\{\bar{F}^-(k/n)\}^r\left(\frac{\bar{F}^-(1-U_{n,n-k})}{\bar{F}^-(k/n)}-1\right)\{1+o_p(1)\} \\
&= -rd_0(\theta_0-\beta_0)^{-r}\frac{\sqrt{n}}{\sigma(k/n)}\frac{k}{n}\{\bar{F}^-(k/n)\}^r\left\{\left\{\frac{n}{k}(1-U_{n,n-k})\right\}^{-1/\alpha}-1\right\}\{1+o_p(1)\} \\
&= \frac{rd_0}{\alpha}(\theta_0-\beta_0)^{-r}\frac{\sqrt{n}}{\sigma(k/n)}\frac{k}{n}\{\bar{F}^-(k/n)\}^r\left\{\frac{n}{k}(1-U_{n,n-k})-1\right\}\{1+o_p(1)\} \\
&= \frac{rd_0}{\alpha}(\theta_0-\beta_0)^{-r}\frac{\sqrt{k/n}\{\bar{F}^-(k/n)\}^r}{\sigma(k/n)}\sqrt{\frac{n}{k}}B_n\left(1-\frac{k}{n}\right)\{1+o_p(1)\} \\
&= \frac{r}{\alpha}\sqrt{\frac{2r-\alpha}{2r}}I(\alpha<2r)\sqrt{\frac{n}{k}}B_n\left(1-\frac{k}{n}\right)+o_p(1).
\end{aligned}
$$
(4.5.10)

By Lemma 4.5.1 and (4.5.1), we have

$$
\frac{\sqrt{n}I_3}{\sigma(k/n)} = -\frac{\int_{F(\beta_0)}^{1-k/n}B_n(s)\,d\psi\left(\frac{F^-(s)-\beta_0}{\theta_0-\beta_0}\right)}{\sigma(k/n)}+o_p(1).
$$
(4.5.11)

Note that for any ξ between $U_{n,n-k}$ and $1-k/n$, we have $\frac{\sqrt{n}}{k}B_n(\xi)\xrightarrow{p}0$ and $U_{n,n-k}/(1-k/n)\xrightarrow{p}1$.
Hence, by Lemma 4.5.1 and (4.5.10),

$$
\begin{aligned}
\frac{\sqrt{n}I_4}{\sigma(k/n)} &= -\frac{\sqrt{n}}{\sigma(k/n)}\int_{1-k/n}^{U_{n,n-k}}\{G_n(x)-x\}d\psi\left(\frac{F^-(x)-\beta_0}{\theta_0-\beta_0}\right) \\
&= -\frac{1}{\sigma(k/n)}\int_{1-k/n}^{U_{n,n-k}}B_n(x)d\psi\left(\frac{F^-(x)-\beta_0}{\theta_0-\beta_0}\right)\{1+o_p(1)\} \\
&= -\frac{B_n(\xi)}{\sigma(k/n)}\left\{\psi\left(\frac{X_{n,n-k}-\beta_0}{\theta_0-\beta_0}\right)-\psi\left(\frac{\bar{F}^-(k/n)-\beta_0}{\theta_0-\beta_0}\right)\right\}\{1+o_p(1)\} \\
&= \frac{\sqrt{n}}{\sigma(k/n)}I_2\frac{\sqrt{n}}{k}B_n(\xi)\{1+o_p(1)\} \\
&= o_p(1).
\end{aligned}
$$
(4.5.12)

Using (4.5.10), we have

$$
\begin{aligned}
\frac{\sqrt{n}I_5}{\sigma(k/n)} &= \frac{\sqrt{n}}{\sigma(k/n)}\int_{\bar{F}^-(k/n)}^{X_{n,n-k}}\bar{F}(x)d\psi\left(\frac{x-\beta_0}{\theta_0-\beta_0}\right) \\
&= \frac{\sqrt{n}}{\sigma(k/n)}\bar{F}(\bar{F}^-(k/n))\int_{\bar{F}^-(k/n)}^{X_{n,n-k}}d\psi\left(\frac{x-\beta_0}{\theta_0-\beta_0}\right)\{1+o_p(1)\} \\
&= \frac{\sqrt{n}}{\sigma(k/n)}\frac{k}{n}\left\{\psi\left(\frac{X_{n,n-k}-\beta_0}{\theta_0-\beta_0}\right)-\psi\left(\frac{\bar{F}^-(k/n)-\beta_0}{\theta_0-\beta_0}\right)\right\}\{1+o_p(1)\} \\
&= -\frac{\sqrt{n}I_2}{\sigma(k/n)}\{1+o_p(1)\}.
\end{aligned}
$$
(4.5.13)

70

Therefore, (4.5.3) follows from (4.5.8), (4.5.10)–(4.5.13).

For proving (4.5.4), we write

$$
\begin{aligned}
& \int_{X_{n,n-k}}^{\infty} \psi\left(\tfrac{x-\beta_0}{\theta_0-\beta_0}\right) d\{1 - \hat{\bar{F}}(x)\} \\
=\ & \int_{X_{n,n-k}}^{\infty} \psi\left(\tfrac{x-\beta_0}{\theta_0-\beta_0}\right) d\{1 - \tfrac{k}{n}\left(\tfrac{x}{X_{n,n-k}}\right)^{-1/\hat{\gamma}(k)} - 1 + \tfrac{k}{n}\left(\tfrac{x}{X_{n,n-k}}\right)^{-1/\gamma}\} \\
& + \int_{X_{n,n-k}}^{\infty} \psi\left(\tfrac{x-\beta_0}{\theta_0-\beta_0}\right) d\{1 - \tfrac{k}{n}\left(\tfrac{x}{X_{n,n-k}}\right)^{-1/\gamma} - F(x)\} \\
& + \int_{X_{n,n-k}}^{\bar{F}^{-}(k/n)} \psi\left(\tfrac{x-\beta_0}{\theta_0-\beta_0}\right) dF(x) \\
& + \int_{\bar{F}^{-}(k/n)}^{\infty} \psi\left(\tfrac{x-\beta_0}{\theta_0-\beta_0}\right) dF(x) \\
=\ & II_1 + \cdots + II_4.
\end{aligned}
\tag{4.5.14}
$$

Note that (4.2.2) implies that

$$
\lim_{t\to 0} \frac{\log \frac{\bar{F}^{-}(tx)}{\bar{F}^{-}(t)} + \gamma \log x}{A(t)} = \frac{x^{\rho} - 1}{\rho}.
$$

Hence, using (4.5.2) and standard arguments in proving the asymptotic limit of Hill's estimator (see De Haan and Ferreira (2007)), we can show that

$$
\begin{aligned}
& \sqrt{k}\{\hat{\gamma}(k) - \gamma\} \\
=\ & \sqrt{k}\{-\tfrac{1}{k}\sum_{i=1}^{k} \gamma \log \tfrac{1-U_{n,n-i+1}}{1-U_{n,n-k}} - \gamma\} \\
& + \sqrt{k} A(1-U_{n,n-k})\{1 + o_p(1)\} \tfrac{1}{k} \sum_{i=1}^{k} \frac{\left(\tfrac{1-U_{n,n-i+1}}{1-U_{n,n-k}}\right)^{\rho} - 1}{\rho} \\
=\ & -\sqrt{k}\tfrac{1}{k}\sum_{i=1}^{k} \gamma \log \tfrac{\frac{n}{i}(1-U_{n,n-i+1})}{\frac{n}{k}(1-U_{n,n-k})} + \sqrt{k}\{\tfrac{1}{k}\sum_{i=1}^{k} \gamma \log \tfrac{k}{i} - \gamma\} \\
& + \sqrt{k} A(1-U_{n,n-k})\{1 + o_p(1)\} \tfrac{1}{k} \sum_{i=1}^{k} \frac{\left(\tfrac{1-U_{n,n-i+1}}{1-U_{n,n-k}}\right)^{\rho} - 1}{\rho} \\
=\ & -\sqrt{k}\tfrac{1}{k}\sum_{i=1}^{k} \gamma\{\tfrac{n}{i}(1-U_{n,n-i+1}) - 1\} + \gamma\sqrt{k}\{\tfrac{n}{k}(1-U_{n,n-k}) - 1\} + \lambda \int_0^1 \tfrac{s^{\rho}-1}{\rho}\, ds \\
& + o_p(1) \\
=\ & -\gamma\sqrt{\tfrac{n}{k}} \int_0^1 \{\tfrac{B_n(1-\frac{k}{n}s)}{s} - B_n(1-\tfrac{k}{n})\}\, ds - \tfrac{\lambda}{1+\rho} + o_p(1).
\end{aligned}
\tag{4.5.15}
$$

It follows from Lemma 4.5.1, (4.5.9) and (4.5.15) that

$$
\begin{aligned}
\frac{\sqrt{n}II_1}{\sigma(k/n)} &= -\frac{\sqrt{n}}{\sigma(k/n)}\frac{k}{n}\int_1^\infty \psi\Big(\frac{X_{n,n-k}x-\beta_0}{\theta_0-\beta_0}\Big)d\{x^{-1/\hat\gamma(k)}-x^{-1/\gamma}\} \\
&= \frac{\sqrt{n}}{\sigma(k/n)}\frac{k}{n}\Big(\frac{1}{\hat\gamma(k)}-\frac{1}{\gamma}\Big)\int_1^\infty \psi\Big(\frac{X_{n,n-k}x-\beta_0}{\theta_0-\beta_0}\Big)d\{x^{-1/\gamma}\log x\}\{1+o_p(1)\} \\
&= \frac{\sqrt{n}}{\sigma(k/n)}\frac{k}{n}\Big(\frac{1}{\hat\gamma(k)}-\frac{1}{\gamma}\Big)d_0\int_1^\infty \Big(\frac{X_{n,n-k}x-\beta_0}{\theta_0-\beta_0}\Big)^r d\{x^{-1/\gamma}\log x\}\{1+o_p(1)\} \\
&= d_0(\theta_0-\beta_0)^{-r}\frac{\sqrt{n}}{\sigma(k/n)}\frac{k}{n}\{\bar F^-(k/n)\}^r\Big(\frac{1}{\hat\gamma(k)}-\frac{1}{\gamma}\Big)\int_1^\infty x^r d\{x^{-\alpha}\log x\}\{1+o_p(1)\} \\
&= -\sqrt{\frac{2r-\alpha}{2r}}\frac{r}{(\alpha-r)^2}I(\alpha<2r)\Big(\alpha\sqrt{\tfrac{n}{k}}\int_0^1\{\tfrac{B_n(1-\frac{k}{n}s)}{s}-B_n(1-\tfrac{k}{n})\}\,ds+\tfrac{\alpha^2\lambda}{1+\rho}\Big) \\
&\quad +o_p(1),
\end{aligned}
$$

$$(4.5.16)$$

$$
\begin{aligned}
\frac{\sqrt{n}II_2}{\sigma(k/n)} &= \frac{\sqrt{n}}{\sigma(k/n)}\int_{X_{n,n-k}}^\infty \psi\Big(\frac{x-\beta_0}{\theta_0-\beta_0}\Big)d\{\bar F(x)-\tfrac{k}{n}\big(\tfrac{x}{X_{n,n-k}}\big)^{-1/\gamma}\} \\
&= \frac{\sqrt{n}}{\sigma(k/n)}\frac{k}{n}\int_1^\infty \psi\Big(\frac{X_{n,n-k}x-\beta_0}{\theta_0-\beta_0}\Big)d\{\tfrac{\bar F(X_{n,n-k}x)}{k/n}-x^{-\alpha}\} \\
&= \frac{\sqrt{n}}{\sigma(k/n)}\frac{k}{n}\{\big(\tfrac{X_{n,n-k}}{\bar F^-(k/n)}\big)^{-\alpha}-1\}\int_1^\infty \psi\Big(\frac{X_{n,n-k}x-\beta_0}{\theta_0-\beta_0}\Big)d\{x^{-\alpha}\}\{1+o_p(1)\} \\
&= d_0\frac{\sqrt{n}}{\sigma(k/n)}\frac{k}{n}\{\big(\tfrac{X_{n,n-k}}{\bar F^-(k/n)}\big)^{-\alpha}-1\}\int_1^\infty \Big(\frac{X_{n,n-k}x-\beta_0}{\theta_0-\beta_0}\Big)^r d\{x^{-\alpha}\}\{1+o_p(1)\} \\
&= d_0(\theta_0-\beta_0)^{-r}\frac{\sqrt{n}}{\sigma(k/n)}\frac{k}{n}\{\bar F^-(k/n)\}^r\{\big(\tfrac{X_{n,n-k}}{\bar F^-(k/n)}\big)^{-\alpha}-1\}\int_1^\infty x^r\,dx^{-\alpha}\{1+o_p(1)\} \\
&= \frac{\alpha d_0}{r-\alpha}(\theta_0-\beta_0)^{-r}\frac{\sqrt{n}}{\sigma(k/n)}\frac{k}{n}\{\bar F^-(k/n)\}^r\{\tfrac{n}{k}(1-U_{n,n-k})-1\}\{1+o_p(1)\} \\
&= \sqrt{\frac{2r-\alpha}{2r}}\frac{\alpha}{r-\alpha}I(\alpha<2r)\sqrt{\tfrac{n}{k}}B_n(1-\tfrac{k}{n})+o_p(1),
\end{aligned}
$$

$$(4.5.17)$$

and

$$
\begin{aligned}
\frac{\sqrt{n}II_3}{\sigma(k/n)} &= \frac{\sqrt{n}}{\sigma(k/n)}\psi\Big(\frac{\bar F^-(k/n)-\beta_0}{\theta_0-\beta_0}\Big)\int_{\bar F^-(k/n)}^{X_{n,n-k}}dF(x)\{1+o_p(1)\} \\
&= d_0(\theta_0-\beta_0)^{-r}\frac{\sqrt{n}}{\sigma(k/n)}\{\bar F^-(k/n)\}^r(1-\tfrac{k}{n}-U_{n,n-k})\{1+o_p(1)\} \\
&= d_0(\theta_0-\beta_0)^{-r}\frac{\sqrt{k/n}\{\bar F^-(k/n)\}^r}{\sigma(k/n)}\sqrt{k}\{\tfrac{n}{k}(1-U_{n,n-k})-1\}\{1+o_p(1)\} \\
&= \sqrt{\frac{2r-\alpha}{2r}}I(\alpha<2r)\sqrt{\tfrac{n}{k}}B_n(1-\tfrac{k}{n})+o_p(1).
\end{aligned}
$$

$$(4.5.18)$$

Hence (4.5.4) follows from (4.5.14), (4.5.16)–(4.5.18).

Note that

$$
\frac{\partial\Delta_{11}(\theta_0,\beta_0)}{\partial\theta}=-\int_{\beta_0}^{X_{n,n-k}}\psi'\Big(\frac{x-\beta_0}{\theta_0-\beta_0}\Big)\frac{x-\beta_0}{(\theta_0-\beta_0)^2}\,dF_n(x)
$$

and

$$\frac{\partial \Delta_{12}(\theta_0, \beta_0)}{\partial \theta} = -\int_{X_{n,n-k}}^{\infty} \psi'(\frac{x - \beta_0}{\theta_0 - \beta_0}) \frac{x - \beta_0}{(\theta_0 - \beta_0)^2} \, d\hat{\bar{F}}(x).$$

Similar to the proofs above, one can show that

$$\int_{\beta_0}^{X_{n,n-k}} \psi'(\frac{x - \beta_0}{\theta_0 - \beta_0}) \frac{x - \beta_0}{(\theta_0 - \beta_0)^2} \, dF_n(x) - \int_{\beta_0}^{\bar{F}^-(\frac{k}{n})} \psi'(\frac{x - \beta_0}{\theta_0 - \beta_0}) \frac{x - \beta_0}{(\theta_0 - \beta_0)^2} \, dF(x) \xrightarrow{p} 0.$$

Then (4.5.5) and the first part of (4.5.7) follow from

$$\int_{\beta_0}^{\bar{F}^-(\frac{k}{n})} \psi'(\frac{x - \beta_0}{\theta_0 - \beta_0}) \frac{x - \beta_0}{(\theta_0 - \beta_0)^2} \, dF(x) \to E\{\psi'(\frac{X - \beta_0}{\theta_0 - \beta_0}) \frac{X - \beta_0}{(\theta_0 - \beta_0)^2} I(X > \beta_0)\}$$

and

$$\begin{aligned}
& \int_{X_{n,n-k}}^{\infty} \psi'(\frac{x-\beta_0}{\theta_0-\beta_0}) \frac{x-\beta_0}{(\theta_0-\beta_0)^2} \, d\{1 - \hat{\bar{F}}(x)\} \\
= \ & \int_{X_{n,n-k}}^{\infty} \psi'(\frac{x-\beta_0}{\theta_0-\beta_0}) \frac{x-\beta_0}{(\theta_0-\beta_0)^2} \frac{k}{n} X_{n,n-k}^{1/\hat{\gamma}} \frac{1}{\hat{\gamma}} x^{-1/\hat{\gamma}-1} \, dx \\
= \ & O_p(\frac{k}{n} X_{n,n-k}^{1/\hat{\gamma}} X_{n,n-k}^{r-1/\hat{\gamma}}) \\
= \ & O_p((\frac{k}{n})^{1-r\gamma-\epsilon}) = o_p(1)
\end{aligned}$$

by using the Potter's bound in Bingham, Goldie, and Teugels (1987) and choosing $\epsilon > 0$ small enough. The rest assertions follow from the same arguments as above, and thus we skip the details.

∎

Lemma 4.5.3. *Under conditions of Theorem 4.2.1, we have*

$$\begin{aligned}
& \frac{\sqrt{n}}{\sigma(k/n)}\{\Delta_{21}(\theta_0, \beta_0) - \int_{\beta_0}^{\bar{F}^-(k/n)} \psi'(\frac{x-\beta_0}{\theta_0-\beta_0})(x - \theta_0) \, dF(x)\} \\
= \ & -\frac{\int_{F(\beta_0)}^{1-k/n} B_n(s) \, d\{\psi'(\frac{F^-(s)-\beta_0}{\theta_0-\beta_0})(F^-(s)-\theta_0)\}}{\sigma(k/n)} + o_p(1),
\end{aligned} \tag{4.5.19}$$

$$\begin{aligned}
& \frac{\sqrt{n}}{\sigma(k/n)}\{\Delta_{22}(\theta_0, \beta_0) - \int_{\bar{F}^-(k/n)}^{\infty} \psi'(\frac{x-\beta_0}{\theta_0-\beta_0})(x - \theta_0) \, dF(x)\} \\
= \ & -(\theta_0 - \beta_0)\sqrt{\frac{2r-\alpha}{2r}} I(\alpha < 2r) \frac{\alpha r^2}{(r-\alpha)^2} \sqrt{\frac{n}{k}} \int_0^1 \{\frac{B_n(1-\frac{k}{n}s)}{s} - B_n(1-\frac{k}{n})\} ds \\
& +(\theta_0 - \beta_0)\sqrt{\frac{2r-\alpha}{2r}} I(\alpha < 2r) \frac{r^2}{r-\alpha} \sqrt{\frac{n}{k}} B_n(1-\frac{k}{n}) \\
& -(\theta_0 - \beta_0)\sqrt{\frac{2r-\alpha}{2r}} I(\alpha < 2r) \frac{\alpha^2 r^2 \lambda}{(r-\alpha)^2(1+\rho)} + o_p(1),
\end{aligned} \tag{4.5.20}$$

$$\frac{\partial \Delta_{21}(\theta_0, \beta_0)}{\partial \theta} = -E\{\psi''(\tfrac{X-\beta_0}{\theta_0-\beta_0})\tfrac{(X-\beta_0)(X-\theta_0)}{(\theta_0-\beta_0)^2}I(X > \beta_0)$$
$$+\psi'(\tfrac{X-\beta_0}{\theta_0-\beta_0})I(X > \beta_0)\} + o_p(1), \qquad (4.5.21)$$

$$\frac{\partial \Delta_{21}(\theta_0, \beta_0)}{\partial \beta} = E\{\psi''(\frac{X-\beta_0}{\theta_0-\beta_0})\frac{(X-\theta_0)^2}{(\theta_0-\beta_0)^2}I(X > \beta_0)\} + o_p(1), \qquad (4.5.22)$$

$$\frac{\partial \Delta_{22}(\theta_0, \beta_0)}{\partial \theta} = o_p(1), \quad \frac{\partial \Delta_{22}(\theta_0, \beta_0)}{\partial \beta} = o_p(1). \qquad (4.5.23)$$

Proof. The proof is similar to that of Lemma 4.5.2, and so we skip the details. ∎

Proof of Theorem 4.2.1. By Taylor expansions and Lemmas 4.5.1 – 4.5.3, we have

$$\frac{\sqrt{n}}{\sigma(k/n)}\{\Delta_{11}(\theta, \beta) - \Delta_{11}(\theta_0, \beta_0)\} + \frac{\sqrt{n}}{\sigma(k/n)}\{\Delta_{11}(\theta_0, \beta_0) - \int_{\beta_0}^{\bar{F}^-(k/n)} \psi(\tfrac{x-\beta_0}{\theta_0-\beta_0}) \, dF(x)\}$$
$$+ \frac{\sqrt{n}}{\sigma(k/n)}\{\Delta_{12}(\theta, \beta) - \Delta_{12}(\theta_0, \beta_0)\} + \frac{\sqrt{n}}{\sigma(k/n)}\{\Delta_{12}(\theta_0, \beta_0) - \int_{\bar{F}^-(k/n)}^{\infty} \psi(\tfrac{x-\beta_0}{\theta_0-\beta_0}) \, dF(x)\}$$
$$= -\frac{\sqrt{n}}{\sigma(k/n)}(\theta - \theta_0)E\{\psi'(\tfrac{X-\beta_0}{\theta_0-\beta_0})\tfrac{X-\beta_0}{(\theta_0-\beta_0)^2}I(X > \beta_0)\}$$
$$+ \frac{\sqrt{n}}{\sigma(k/n)}(\beta - \beta_0)E\{\psi'(\tfrac{X-\beta_0}{\theta_0-\beta_0})\tfrac{X-\theta_0}{(\theta_0-\beta_0)^2}I(X > \beta_0)\}$$
$$- \frac{\int_{F(\beta_0)}^{1-k/n} B_n(s) \, d\psi(\tfrac{F^-(s)-\beta_0}{\theta_0-\beta_0})}{\sigma(k/n)} + \sqrt{\tfrac{2r-\alpha}{2r}}I(\alpha < 2r)\tfrac{r^2}{(r-\alpha)^2}\sqrt{\tfrac{n}{k}}B_n(1 - \tfrac{k}{n})$$
$$- \sqrt{\tfrac{2r-\alpha}{2r}}I(\alpha < 2r)\tfrac{r\alpha}{(r-\alpha)^2}\sqrt{\tfrac{n}{k}}\int_0^1 \tfrac{B_n(1-\tfrac{k}{n}s)}{s} ds$$
$$- \sqrt{\tfrac{2r-\alpha}{2r}}I(\alpha < 2r)\tfrac{r\alpha^2\lambda}{(\alpha-r)^2(1+\rho)} + o_p(1).$$

Similarly

$$
\begin{aligned}
&\frac{\sqrt{n}}{\sigma(k/n)}\{\Delta_{21}(\theta,\beta)-\Delta_{21}(\theta_0,\beta_0)\}+\frac{\sqrt{n}}{\sigma(k/n)}\{\Delta_{21}(\theta_0,\beta_0)\\
&\quad-\int_{\beta_0}^{\bar{F}^-(k/n)}\psi'(\tfrac{x-\beta_0}{\theta_0-\beta_0})(x-\theta_0)\,dF(x)\}\quad+\frac{\sqrt{n}}{\sigma(k/n)}\{\Delta_{22}(\theta,\beta)-\Delta_{22}(\theta_0,\beta_0)\}\\
&\quad+\frac{\sqrt{n}}{\sigma(k/n)}\{\Delta_{22}(\theta_0,\beta_0)-\int_{\bar{F}^-(k/n)}^{\infty}\psi'(\tfrac{x-\beta_0}{\theta_0-\beta_0})(x-\theta_0)\,dF(x)\}\\
&=-\frac{\sqrt{n}}{\sigma(k/n)}(\theta-\theta_0)E\{\psi''(\tfrac{X-\beta_0}{\theta_0-\beta_0})\tfrac{(X-\beta_0)(X-\theta_0)}{(\theta_0-\beta_0)^2}I(X>\beta_0)+\psi'(\tfrac{X-\beta_0}{\theta_0-\beta_0})I(X>\beta_0)\}\\
&\quad+\frac{\sqrt{n}}{\sigma(k/n)}(\beta-\beta_0)E\{\psi''(\tfrac{X-\beta_0}{\theta_0-\beta_0})\tfrac{(X-\theta_0)^2}{(\theta_0-\beta_0)^2}I(X>\beta_0)\}\\
&\quad-\frac{\int_{F(\beta_0)}^{1-k/n}B_n(s)\,d\{\psi'(\tfrac{F^-(s)-\beta_0}{\theta_0-\beta_0})(F^-(s)-\theta_0)\}}{\sigma(k/n)}\\
&\quad+(\theta_0-\beta_0)\sqrt{\tfrac{2r-\alpha}{2r}}I(\alpha<2r)\tfrac{r^3}{(r-\alpha)^2}\sqrt{\tfrac{n}{k}}B_n(1-\tfrac{k}{n})\\
&\quad-(\theta_0-\beta_0)\sqrt{\tfrac{2r-\alpha}{2r}}I(\alpha<2r)\tfrac{\alpha r^2}{(r-\alpha)^2}\sqrt{\tfrac{n}{k}}\int_0^1\tfrac{B_n(1-\tfrac{k}{n}s)}{s}ds\\
&\quad-(\theta_0-\beta_0)\sqrt{\tfrac{2r-\alpha}{2r}}I(\alpha<2r)\tfrac{\alpha^2 r^2\lambda}{(r-\alpha)^2(1+\rho)}+o_p(1).
\end{aligned}
$$

Hence

$$
\begin{cases}
\dfrac{\sqrt{n}}{\sigma(k/n)}(\hat{\theta}-\theta_0)a_1=\xi_n+o_p(1)\\[2mm]
\dfrac{\sqrt{n}}{\sigma(k/n)}(\hat{\theta}-\theta_0)a_2+\dfrac{\sqrt{n}}{\sigma(k/n)}(\hat{\beta}-\beta_0)b_2=\eta_n+o_p(1),
\end{cases}
\tag{4.5.24}
$$

where

$$
\begin{aligned}
\xi_n&=-\frac{\int_{F(\beta_0)}^{1-k/n}B_n(s)\,d\psi(\tfrac{F^-(s)-\beta_0}{\theta_0-\beta_0})}{\sigma(k/n)}+\sqrt{\tfrac{2r-\alpha}{2r}}I(\alpha<2r)\tfrac{r^2}{(r-\alpha)^2}\sqrt{\tfrac{n}{k}}B_n(1-\tfrac{k}{n})\\
&\quad-\sqrt{\tfrac{2r-\alpha}{2r}}I(\alpha<2r)\tfrac{r\alpha}{(r-\alpha)^2}\sqrt{\tfrac{n}{k}}\int_0^1\tfrac{B_n(1-\tfrac{k}{n}s)}{s}ds-\sqrt{\tfrac{2r-\alpha}{2r}}I(\alpha<2r)\tfrac{r\alpha^2\lambda}{(r-\alpha)^2(1+\rho)},
\end{aligned}
\tag{4.5.25}
$$

$$
\begin{aligned}
\eta_n&=-\frac{\int_{F(\beta_0)}^{1-k/n}B_n(s)\,d\{\psi'(\tfrac{F^-(s)-\beta_0}{\theta_0-\beta_0})(F^-(s)-\theta_0)\}}{\sigma(k/n)(\theta_0-\beta_0)}+\sqrt{\tfrac{2r-\alpha}{2r}}I(\alpha<2r)\tfrac{r^3}{(r-\alpha)^2}\sqrt{\tfrac{n}{k}}B_n(1-\tfrac{k}{n})\\
&\quad-\sqrt{\tfrac{2r-\alpha}{2r}}I(\alpha<2r)\tfrac{\alpha r^2}{(r-\alpha)^2}\sqrt{\tfrac{n}{k}}\int_0^1\tfrac{B_n(1-\tfrac{k}{n}s)}{s}ds-\sqrt{\tfrac{2r-\alpha}{2r}}I(\alpha<2r)\tfrac{\alpha^2 r^2\lambda}{(r-\alpha)^2(1+\rho)}.
\end{aligned}
\tag{4.5.26}
$$

Note that

$$
(\xi_n,\eta_n)^T\xrightarrow{d}N(\Delta,\Sigma),\quad n\to\infty,
\tag{4.5.27}
$$

where $\Delta=\begin{pmatrix}-\sqrt{\tfrac{2r-\alpha}{2r}}I(\alpha<2r)\tfrac{r\alpha^2\lambda}{(r-\alpha)^2(1+\rho)}\\ -\sqrt{\tfrac{2r-\alpha}{2r}}I(\alpha<2r)\tfrac{\alpha^2 r^2\lambda}{(r-\alpha)^2(1+\rho)}\end{pmatrix}$ and $\Sigma=\begin{pmatrix}\sigma_1^2 & \sigma_{12}\\ \sigma_{21} & \sigma_2^2\end{pmatrix}$ with σ_1^2,σ_2^2 and σ_{12} calculated as follows.

75

First we have

$$\sigma_1^2 = \lim_{n\to\infty} Var(\xi_n)$$

$$= 1 + \frac{2r-\alpha}{2r}\frac{r^4}{(r-\alpha)^4}I(\alpha < 2r)$$

$$+ \frac{2r-\alpha}{2r}\frac{r^2\alpha^2}{(r-\alpha)^4}I(\alpha < 2r)\lim_{n\to\infty}\frac{n}{k}\int_0^1\int_0^1\frac{E\{B_n(1-\frac{k}{n}s)B_n(1-\frac{k}{n}t)\}}{st}\,dsdt$$

$$- \sqrt{\frac{2r-\alpha}{2r}}\frac{2r^2}{(r-\alpha)^2}I(\alpha < 2r)\lim_{n\to\infty}\frac{\sqrt{\frac{k}{n}}\int_{F(\beta_0)}^{1-k/n}s\,d\psi(\frac{F^-(s)-\beta_0}{\theta_0-\beta_0})}{\sigma(k/n)}$$

$$+ \sqrt{\frac{2r-\alpha}{2r}}\frac{2r\alpha}{(r-\alpha)^2}I(\alpha < 2r)\lim_{n\to\infty}\frac{\sqrt{\frac{k}{n}}\int_{F(\beta_0)}^{1-k/n}s\,d\psi(\frac{F^-(s)-\beta_0}{\theta_0-\beta_0})}{\sigma(k/n)}$$

$$- \frac{2r-\alpha}{2r}\frac{2r^3\alpha}{(r-\alpha)^4}I(\alpha < 2r)$$

$$= 1 + \frac{2r-\alpha}{2r}\frac{r^4}{(r-\alpha)^4}I(\alpha < 2r) + \frac{2r-\alpha}{2r}\frac{2r^2\alpha^2}{(r-\alpha)^4}I(\alpha < 2r)$$

$$- \frac{2r-\alpha}{2r}\frac{2r^2}{(r-\alpha)^2}I(\alpha < 2r) + \frac{2r-\alpha}{2r}\frac{2r\alpha}{(r-\alpha)^2}I(\alpha < 2r) - \frac{2r-\alpha}{2r}\frac{2\alpha r^3}{(r-\alpha)^4}I(\alpha < 2r)$$

$$= 1 + (2r-\alpha)\{\frac{r^3}{2(r-\alpha)^4} - \frac{\alpha r}{(r-\alpha)^3} - \frac{1}{r-\alpha}\}I(\alpha < 2r),$$

where the third equality holds since

$$\lim_{n\to\infty}\frac{\sqrt{k/n}\int_{F(\beta_0)}^{1-k/n}s\,d\psi(\frac{F^-(s)-\beta_0}{\theta_0-\beta_0})}{\sigma(k/n)}$$

$$= \lim_{n\to\infty}\frac{\sqrt{\frac{k}{n}}\psi(\frac{\bar{F}^-(k/n)-\beta_0}{\theta_0-\beta_0})}{\sigma(k/n)} - \lim_{n\to\infty}\frac{\sqrt{k/n}}{\sigma(k/n)}\int_{\beta_0}^{\bar{F}^-(\frac{k}{n})}\bar{F}(s)d\psi(\frac{s-\beta_0}{\theta_0-\beta_0})$$

$$= \sqrt{\frac{2r-\alpha}{2r}}I(\alpha < 2r) - \lim_{n\to\infty}\frac{\sqrt{k/n}}{\sigma(k/n)}\int_{\beta_0}^{\bar{F}^-(\frac{k}{n})}\bar{F}(s)d\psi(\frac{s-\beta_0}{\theta_0-\beta_0})$$

$$= \sqrt{\frac{2r-\alpha}{2r}}I(\alpha < 2r).$$

Similarly we have

$$\sigma_2^2 = \lim_{n\to\infty} Var(\eta_n)$$

$$= r^2 + \frac{2r-\alpha}{2r}\frac{r^6}{(r-\alpha)^4}I(\alpha < 2r) + \frac{2r-\alpha}{2r}\frac{2r^4\alpha^2}{(r-\alpha)^4}I(\alpha < 2r)$$

$$- \sqrt{\frac{2r-\alpha}{2r}}\frac{2r^3}{(r-\alpha)^2}I(\alpha < 2r)\lim_{n\to\infty}\frac{\sqrt{\frac{k}{n}}\int_{F(\beta_0)}^{1-k/n}s\,d\{\psi'(\frac{F^-(s)-\beta_0}{\theta_0-\beta_0})(F^-(s)-\theta_0)\}}{\sigma(k/n)(\theta_0-\beta_0)}$$

$$+ \sqrt{\frac{2r-\alpha}{2r}}\frac{2r^2\alpha}{(r-\alpha)^2}I(\alpha < 2r)\lim_{n\to\infty}\frac{\sqrt{\frac{k}{n}}\int_{F(\beta_0)}^{1-k/n}s\,d\{\psi'(\frac{F^-(s)-\beta_0}{\theta_0-\beta_0})(F^-(s)-\theta_0)\}}{\sigma(k/n)(\theta_0-\beta_0)}$$

$$- \frac{2r-\alpha}{2r}\frac{2r^5\alpha}{(r-\alpha)^4}I(\alpha < 2r)$$

$$= r^2 + r^2(2r-\alpha)\{\frac{r^3}{2(r-\alpha)^4} - \frac{\alpha r}{(r-\alpha)^3} - \frac{1}{r-\alpha}\}I(\alpha < 2r)$$

76

and

$$\begin{aligned}
\sigma_{12} \;&= \sigma_{21} = \lim_{n\to\infty} Cov(\xi_n, \eta_n)\\
&= r - \tfrac{2r-\alpha}{2r}\tfrac{r^3}{(r-\alpha)^2}I(\alpha < 2r) + \tfrac{2r-\alpha}{2r}\tfrac{r^2\alpha}{(r-\alpha)^2}I(\alpha < 2r)\\
&\quad - \tfrac{2r-\alpha}{2r}\tfrac{r^3}{(r-\alpha)^2}I(\alpha < 2r) + \tfrac{2r-\alpha}{2r}\tfrac{r^5}{(r-\alpha)^4}I(\alpha < 2r) - \tfrac{2r-\alpha}{2r}\tfrac{r^4\alpha}{(r-\alpha)^4}I(\alpha < 2r)\\
&\quad + \tfrac{2r-\alpha}{2r}\tfrac{r^2\alpha}{(r-\alpha)^2}I(\alpha < 2r) - \tfrac{2r-\alpha}{2r}\tfrac{r^4\alpha}{(r-\alpha)^4}I(\alpha < 2r) + \tfrac{2r-\alpha}{2r}\tfrac{2r^3\alpha^2}{(r-\alpha)^4}I(\alpha < 2r)\\
&= r + r(2r-\alpha)\{\tfrac{r^3}{2(r-\alpha)^4} - \tfrac{\alpha r}{(r-\alpha)^3} - \tfrac{1}{r-\alpha}\}I(\alpha < 2r).
\end{aligned}$$

Hence, the theorem follows from the above equations.

∎

Next for proving Theorem 4.2.4, we first show some lemmas.

Lemma 4.5.4. *Under conditions of Theorem 4.2.4, we have, as $n \to \infty$*

$$\frac{\sqrt{k/n}\{\bar{F}^-(k/n)/\bar{F}^-(1-q_n)\}^r}{\tilde{\sigma}(k/n)} \to \frac{(c_1-c_2)^r}{d_0}\sqrt{\frac{2r-\alpha}{2r}}I(\alpha < 2r),$$

$$\frac{\{(1-q_n)\log(\bar{F}^-(k/n)/\bar{F}^-(1-q_n))\}^{1/2}}{\tilde{\sigma}(k/n)} \to \frac{(c_1-c_2)^r}{d_0\sqrt{2r}} \quad when \quad \alpha = 2r,$$

$$\frac{\sqrt{1-q_n}}{\tilde{\sigma}(k/n)} \to \frac{c_2^{\frac{\alpha}{2}}}{\sqrt{2\int_0^\infty(1+\frac{c_1-c_2}{c_2}s)^{-\alpha}\psi(s)\psi'(s)\,ds}} \quad when \quad \alpha > 2r.$$

Proof. Write

$$\begin{aligned}
\tilde{\sigma}^2(k/n) \;&= 2\int_{\beta_{n0}}^{\bar{F}^-(k/n)}\int_{\beta_{n0}}^s F(t)\bar{F}(s)\,d\psi(\tfrac{t-\beta_{n0}}{\theta_{n0}-\beta_{n0}})d\psi(\tfrac{s-\beta_{n0}}{\theta_{n0}-\beta_{n0}})\\
&= 2\int_{\beta_{n0}}^{\bar{F}^-(k/n)}\bar{F}(s)\psi(\tfrac{s-\beta_{n0}}{\theta_{n0}-\beta_{n0}})\,d\psi(\tfrac{s-\beta_{n0}}{\theta_{n0}-\beta_{n0}})\\
&\quad - 2\int_{\beta_{n0}}^{\bar{F}^-(k/n)}\int_{\beta_{n0}}^s \bar{F}(t)\bar{F}(s)\,d\psi(\tfrac{t-\beta_{n0}}{\theta_{n0}-\beta_{n0}})d\psi(\tfrac{s-\beta_{n0}}{\theta_{n0}-\beta_{n0}})\\
&=: I_1 - I_2.
\end{aligned}$$

77

By noting that

$$I_2 \leq 2 \int_{\beta_{n0}}^{\bar{F}^-(k/n)} \int_{\beta_{n0}}^{s} \bar{F}(\beta_{n0})\bar{F}(s) \, d\psi(\tfrac{t-\beta_{n0}}{\theta_{n0}-\beta_{n0}}) d\psi(\tfrac{s-\beta_{n0}}{\theta_{n0}-\beta_{n0}})$$

$$= \bar{F}(\beta_{n0})I_1 = O(1-q_n)I_1 = o(I_1),$$

we have

$$\tilde{\sigma}^2(k/n) = I_1\{1 + o(1)\}. \tag{4.5.28}$$

To derive the asymptotic behavior of I_1, we write

$$
\begin{aligned}
I_1 &= \tfrac{2\beta_{n0}}{\theta_{n0}-\beta_{n0}} \int_1^{\frac{\bar{F}^-(k/n)}{\beta_{n0}}} \bar{F}(\beta_{n0}s)\psi(\tfrac{s-1}{\theta_{n0}/\beta_{n0}-1})\psi'(\tfrac{s-1}{\theta_{n0}/\beta_{n0}-1}) \, ds \\
&= \tfrac{2\beta_{n0}}{\theta_{n0}-\beta_{n0}} \bar{F}(\beta_{n0}) \int_1^{\frac{\bar{F}^-(k/n)}{\beta_{n0}}} s^{-\alpha}\psi(\tfrac{s-1}{\theta_{n0}/\beta_{n0}-1})\psi'(\tfrac{s-1}{\theta_{n0}/\beta_{n0}-1}) \, ds \\
&\quad + \tfrac{2\beta_{n0}}{\theta_{n0}-\beta_{n0}} \bar{F}(\beta_{n0}) \int_1^{\frac{\bar{F}^-(k/n)}{\beta_{n0}}} \left(\tfrac{\bar{F}(\beta_{n0}s)}{\bar{F}(\beta_{n0})s} - s^{-\alpha} \right) \psi(\tfrac{s-1}{\theta_{n0}/\beta_{n0}-1})\psi'(\tfrac{s-1}{\theta_{n0}/\beta_{n0}-1}) \, ds \\
&=: I_{1,1} + I_{1,2}.
\end{aligned}
$$

When $\alpha < 2r$, it is easy to see that

$$
\begin{aligned}
&\lim_{n\to\infty} \tfrac{I_{1,1}}{(k/n)\{\bar{F}^-(k/n)/\bar{F}^-(1-q_n)\}^{2r}} \\
&= \lim_{n\to\infty} \tfrac{2\beta_{n0}}{\theta_{n0}-\beta_{n0}} \tfrac{\bar{F}(\beta_{n0})}{k/n} \left(\tfrac{\bar{F}^-(1-q_n)}{\bar{F}^-(k/n)}\right)^{2r} \int_1^{\frac{\bar{F}^-(k/n)}{\beta_{n0}}} s^{-\alpha}\psi(\tfrac{s-1}{\theta_{n0}/\beta_{n0}-1})\psi'(\tfrac{s-1}{\theta_{n0}/\beta_{n0}-1}) \, ds \\
&= \lim_{n\to\infty} \tfrac{2\beta_{n0}c_2^{-2r}}{\theta_{n0}-\beta_{n0}} \tfrac{\int_1^{\frac{\bar{F}^-(k/n)}{\beta_{n0}}} s^{-\alpha}\psi(\tfrac{s-1}{\theta_{n0}/\beta_{n0}-1})\psi'(\tfrac{s-1}{\theta_{n0}/\beta_{n0}-1}) \, ds}{\{\bar{F}^-(k/n)/\beta_{n0})\}^{2r-\alpha}} \\
&= \tfrac{2rd_0^2}{2r-\alpha}(c_1 - c_2)^{-2r}.
\end{aligned}
$$

For $I_{1,2}$, it follows from Theorem B.1.10 of De Haan and Ferreira (2007) and a similar argument as the last equation above that

$$\lim_{n\to\infty} \frac{I_{1,2}}{(k/n)\{\bar{F}^-(k/n)/\bar{F}^-(1-q_n)\}^{2r}} = 0,$$

i.e., the lemma holds for $\alpha < 2r$. When $\alpha = 2r$, by noting that

$$\lim_{x\to\infty} \frac{\int_1^x s^{-\alpha}(s-1)^{\alpha-1} \, ds}{\log x} = \lim_{x\to\infty} \frac{x^{-\alpha}(x-1)^{\alpha-1}}{1/x} = 1,$$

78

we have

$$\lim_{n\to\infty} \frac{I_{1,1}}{(1-q_n)\log\{\bar{F}^-(k/n)/\bar{F}^-(1-q_n)\}}$$

$$= \lim_{n\to\infty} \frac{I_{1,1}}{(1-q_n)\log\{\bar{F}^-(k/n)/\beta_{n0}\}} \qquad (4.5.29)$$

$$= \frac{2rd_0^2}{(c_1-c_2)^{2r}}.$$

Applying the inequality for a second order regular variation in Draisma et al. (1999) to (4.2.2), we can show that

$$I_{1,2} = o\left((1-q_n)\log\frac{\bar{F}^-(k/n)}{\bar{F}^-(1-q_n)}\right). \qquad (4.5.30)$$

Hence, it follows from (4.5.29) and (4.5.30) that

$$\lim_{n\to\infty} \frac{I_1}{(1-q_n)\log(\bar{F}^-(k/n)/\bar{F}^-(1-q_n))} = \frac{2rd_0^2}{(c_1-c_2)^{2r}} \quad \text{when} \quad \alpha = 2r. \qquad (4.5.31)$$

When $\alpha > 2r$, it follows from Potter's bound that

$$
\begin{aligned}
\frac{I_1}{1-q_n} &= 2\frac{\beta_{n0}}{\theta_{n0}-\beta_{n0}}\frac{\bar{F}(\beta_{n0})}{1-q_n}\int_1^{\frac{\bar{F}^-(k/n)}{\beta_{n0}}} \frac{\bar{F}(\beta_{n0}s)}{\bar{F}(\beta_{n0})}\psi(\frac{s-1}{\theta_{n0}/\beta_{n0}-1})\psi'(\frac{s-1}{\theta_{n0}/\beta_{n0}-1})\,ds \\
&= 2\frac{\beta_{n0}}{\theta_{n0}-\beta_{n0}}\frac{\bar{F}(\beta_{n0})}{1-q_n}\int_1^{\frac{\bar{F}^-(k/n)}{\beta_{n0}}} s^{-\alpha}\psi(\frac{s-1}{\theta_{n0}/\beta_{n0}-1})\psi'(\frac{s-1}{\theta_{n0}/\beta_{n0}-1})\,ds\,\{1+o(1)\} \\
&= \frac{2c_2^{-\alpha+1}}{(c_1-c_2)}\int_1^\infty s^{-\alpha}\psi(\frac{s-1}{c_1/c_2-1})\psi'(\frac{s-1}{c_1/c_2-1})\,ds\,\{1+o(1)\} \\
&= 2c_2^{-\alpha}\int_0^\infty (1+\frac{c_1-c_2}{c_2}s)^{-\alpha}\psi(s)\psi'(s)\,ds\,\{1+o(1)\},
\end{aligned}
\qquad (4.5.32)
$$

then (4.5.28) and (4.5.32) imply that

$$\lim_{n\to\infty} \frac{\tilde{\sigma}^2(k/n)}{1-q_n} = 2c_2^{-\alpha}\int_0^\infty (1+\frac{c_1-c_2}{c_2}s)^{-\alpha}\psi(s)\psi'(s)\,ds.$$

Hence the lemma follows from (4.5.28)–(4.5.32) and the fact that

$$
\begin{aligned}
\sqrt{\frac{k}{n}}\left(\frac{\bar{F}^-(k/n)}{\bar{F}^-(1-q_n)}\right)^r &= O\left(\sqrt{\frac{k}{n}}\left(\frac{k}{n(1-q_n)}\right)^{-r/\alpha}\right) \\
&= \begin{cases} o(\sqrt{1-q_n}) & \text{if} \quad \alpha > 2r, \\ O(\sqrt{1-q_n}) & \text{if} \quad \alpha = 2r. \end{cases}
\end{aligned}
$$

∎

Lemma 4.5.5. *Under conditions of Theorem 4.2.4, we have*

$$\frac{\sqrt{n}}{\tilde{\sigma}(k/n)}\left\{\Delta_{I11}(\theta_{n0},\beta_{n0}) - \int_{\beta_{n0}}^{\bar{F}^-(k/n)} \psi\left(\frac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}}\right) dF(x)\right\}$$
$$= -\frac{\int_{F(\beta_{n0})}^{1-\frac{k}{n}} B_n(x) \, d\psi\left(\frac{F^-(x)-\beta_{n0}}{\theta_{n0}-\beta_{n0}}\right)}{\tilde{\sigma}(k/n)} + o_p(1), \tag{4.5.33}$$

$$\frac{\sqrt{n}}{\tilde{\sigma}(k/n)}\left\{\Delta_{I12}(\theta_{n0},\beta_{n0}) - \int_{\bar{F}^-(k/n)}^{\infty} \psi\left(\frac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}}\right) dF(x)\right\}$$
$$= -\sqrt{\frac{2r-\alpha}{2r}}I(\alpha < 2r)\frac{r\alpha}{(r-\alpha)^2}\sqrt{\frac{n}{k}}\int_0^1 \frac{B_n(1-\frac{k}{n}s)}{s} \, ds$$
$$+\sqrt{\frac{2r-\alpha}{2r}}I(\alpha < 2r)\frac{r^2}{r-\alpha}\sqrt{\frac{n}{k}}B_n(1-\frac{k}{n})$$
$$-\sqrt{\frac{2r-\alpha}{2r}}I(\alpha < 2r)\frac{r\alpha^2\lambda}{(\alpha-r)^2(1+\rho)} + o_p(1), \tag{4.5.34}$$

$$\frac{\sqrt{n}}{\tilde{\sigma}(k/n)}\frac{\partial\{\Delta_{I11}(\theta_{n0},\beta_{n0})+\Delta_{I12}(\theta_{n0},\beta_{n0})\}}{\partial\theta}$$
$$= -\frac{1-q_n}{\bar{F}^-(1-q_n)}\alpha c_2^{-\alpha-1}\int_0^{\infty}(1+\frac{c_1-c_2}{c_2}x)^{-\alpha-1}\psi'(x) \, dx \, \{1+o(1)\}, \tag{4.5.35}$$

$$\frac{\sqrt{n}}{\tilde{\sigma}(k/n)}\frac{\partial\{\Delta_{I11}(\theta_{n0},\beta_{n0})+\Delta_{I12}(\theta_{n0},\beta_{n0})\}}{\partial\beta} = o_p(1). \tag{4.5.36}$$

Proof. Write

$$\int_{\beta_{n0}}^{X_{n,n-k}} \psi(\tfrac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}}) \, dF_n(x) - \int_{\beta_{n0}}^{\bar{F}^-(\frac{k}{n})} \psi(\tfrac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}}) \, dF(x)$$

$$= \int_{\beta_{n0}}^{X_{n,n-k}} \psi(\tfrac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}}) \, d\{F_n(x) - F(x)\} - \int_{X_{n,n-k}}^{\bar{F}^-(\frac{k}{n})} \psi(\tfrac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}}) \, dF(x)$$

$$= \psi(\tfrac{X_{n,n-k}-\beta_{n0}}{\theta_{n0}-\beta_{n0}})\{F_n(X_{n,n-k}) - F(X_{n,n-k})\}$$

$$\quad - \int_{\beta_{n0}}^{X_{n,n-k}} \{F_n(x) - F(x)\} \, d\psi(\tfrac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}})$$

$$\quad - \int_{X_{n,n-k}}^{\bar{F}^-(\frac{k}{n})} \psi(\tfrac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}}) \, dF(x)$$

$$= \psi(\tfrac{X_{n,n-k}-\beta_{n0}}{\theta_{n0}-\beta_{n0}})\{\bar{F}(X_{n,n-k}) - \tfrac{k}{n}\}$$

$$\quad - \int_{\beta_{n0}}^{X_{n,n-k}} \{F_n(x) - F(x)\} \, d\psi(\tfrac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}}) \qquad\qquad (4.5.37)$$

$$\quad - \int_{X_{n,n-k}}^{\bar{F}^-(\frac{k}{n})} \bar{F}(x) \, d\psi(\tfrac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}}) + \tfrac{k}{n}\psi(\tfrac{\bar{F}^-(\frac{k}{n})-\beta_{n0}}{\theta_{n0}-\beta_{n0}}) - \bar{F}(X_{n,n-k})\psi(\tfrac{X_{n,n-k}-\beta_{n0}}{\theta_{n0}-\beta_{n0}})$$

$$= -\tfrac{k}{n}\{\psi(\tfrac{X_{n,n-k}-\beta_{n0}}{\theta_{n0}-\beta_{n0}}) - \psi(\tfrac{\bar{F}^-(\frac{k}{n})-\beta_{n0}}{\theta_{n0}-\beta_{n0}})\}$$

$$\quad - \int_{\beta_{n0}}^{\bar{F}^-(\frac{k}{n})} \{F_n(x) - F(x)\} \, d\psi(\tfrac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}})$$

$$\quad - \int_{\bar{F}^-(\frac{k}{n})}^{X_{n,n-k}} \{F_n(x) - F(x)\} \, d\psi(\tfrac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}})$$

$$\quad - \int_{X_{n,n-k}}^{\bar{F}^-(\frac{k}{n})} \bar{F}(x) \, d\psi(\tfrac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}})$$

$$= I_1 + I_2 + I_3 + I_4.$$

It follows from Lemma 4.5.4 that

$$\tfrac{\sqrt{n}}{\tilde{\sigma}(k/n)} I_1 = -\tfrac{\sqrt{n}}{\tilde{\sigma}(k/n)} \tfrac{k}{n} \psi'(\tfrac{\bar{F}^-(\frac{k}{n})-\beta_{n0}}{\theta_{n0}-\beta_{n0}})\{\tfrac{X_{n,n-k}-\bar{F}^-(\frac{k}{n})}{\theta_{n0}-\beta_{n0}}\}\{1 + o_p(1)\}$$

$$= -rd_0 \tfrac{\sqrt{n}}{\tilde{\sigma}(k/n)} \tfrac{k}{n} (\tfrac{\bar{F}^-(\frac{k}{n})}{\theta_{n0}-\beta_{n0}})^r (\tfrac{X_{n,n-k}}{\bar{F}^-(\frac{k}{n})} - 1)\{1 + o_p(1)\}$$

$$= -rd_0 \tfrac{\sqrt{n}}{\tilde{\sigma}(k/n)} \tfrac{k}{n} (\tfrac{\bar{F}^-(\frac{k}{n})}{\bar{F}^-(1-q_n)})^r (c_1 - c_2)^{-r} (\tfrac{\bar{F}^-(1-U_{n,n-k})}{\bar{F}^-(\frac{k}{n})} - 1)\{1 + o_p(1)\}$$

$$= -rd_0 \tfrac{\sqrt{n}}{\tilde{\sigma}(k/n)} \tfrac{k}{n} (\tfrac{\bar{F}^-(\frac{k}{n})}{\bar{F}^-(1-q_n)})^r (c_1 - c_2)^{-r}\{[\tfrac{n}{k}(1 - U_{n,n-k})]^{-1/\alpha} - 1\}\{1 + o_p(1)\}$$

$$= -rd_0 \tfrac{\sqrt{n}}{\tilde{\sigma}(k/n)} \tfrac{k}{n} (\tfrac{\bar{F}^-(\frac{k}{n})}{\bar{F}^-(1-q_n)})^r (c_1 - c_2)^{-r}(-\tfrac{1}{\alpha})\{\tfrac{n}{k}(1 - U_{n,n-k}) - 1\}\{1 + o_p(1)\}$$

$$= \tfrac{rd_0}{\alpha} \tfrac{\sqrt{n}}{\tilde{\sigma}(k/n)} \tfrac{k}{n} (\tfrac{\bar{F}^-(\frac{k}{n})}{\bar{F}^-(1-q_n)})^r (c_1 - c_2)^{-r} \tfrac{1}{\sqrt{k}}\{\sqrt{\tfrac{n}{k}} B_n(1 - \tfrac{k}{n}) + o_p(1)\}\{1 + o_p(1)\}$$

$$= \tfrac{rd_0}{\alpha} \tfrac{\sqrt{k/n}\{\bar{F}^-(k/n)/\bar{F}^-(1-q_n)\}^r}{\tilde{\sigma}(k/n)} (c_1 - c_2)^{-r} \sqrt{\tfrac{n}{k}} B_n(1 - \tfrac{k}{n})\{1 + o_p(1)\}$$

$$= \tfrac{r}{\alpha}\sqrt{\tfrac{2r-\alpha}{2r}} I(\alpha < 2r)\sqrt{\tfrac{n}{k}} B_n(1 - \tfrac{k}{n}) + o_p(1).$$

$$(4.5.38)$$

81

By Lemma 4.5.4 and (4.5.1), we have

$$\frac{\sqrt{n}}{\tilde{\sigma}(k/n)}I_2 = -\frac{\sqrt{n}}{\tilde{\sigma}(k/n)}\int_{\beta_{n0}}^{\bar{F}^{-}(\frac{k}{n})}\{F_n(x)-F(x)\}\,d\psi(\tfrac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}})$$
$$= -\frac{\int_{F(\beta_{n0})}^{1-\frac{k}{n}}B_n(x)\,d\psi(\frac{F^{-}(x)-\beta_{n0}}{\theta_{n0}-\beta_{n0}})}{\tilde{\sigma}(k/n)} + o_p(1). \tag{4.5.39}$$

Note that for any ϵ between $U_{n,n-k}$ and $1-\frac{k}{n}$, we have $\frac{\sqrt{n}}{k}B_n(\epsilon)\xrightarrow{p}0$, which implies that

$$\frac{\sqrt{n}}{\tilde{\sigma}(k/n)}I_3 = -\frac{\sqrt{n}}{\tilde{\sigma}(k/n)}\int_{\bar{F}^{-}(\frac{k}{n})}^{X_{n,n-k}}\{F_n(x)-F(x)\}\,d\psi(\tfrac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}})$$
$$= -\frac{1}{\tilde{\sigma}(k/n)}\int_{1-\frac{k}{n}}^{U_{n,n-k}}\sqrt{n}\{G_n(x)-x\}\,d\psi(\tfrac{F^{-}(x)-\beta_{n0}}{\theta_{n0}-\beta_{n0}})$$
$$= -\frac{1}{\tilde{\sigma}(k/n)}B_n(\epsilon)\{\psi(\tfrac{X_{n,n-k}-\beta_{n0}}{\theta_{n0}-\beta_{n0}})-\psi(\tfrac{\bar{F}^{-}(\frac{k}{n})-\beta_{n0}}{\theta_{n0}-\beta_{n0}})\}(1+o_p(1))$$
$$= \frac{1}{\tilde{\sigma}(k/n)}B_n(\epsilon)\tfrac{n}{k}I_2(1+o_p(1))$$
$$= [\tfrac{\sqrt{n}}{\tilde{\sigma}(k/n)}I_2]\tfrac{\sqrt{n}}{k}B_n(\epsilon)(1+o_p(1))$$
$$= o_p(1), \tag{4.5.40}$$

and

$$\frac{\sqrt{n}}{\tilde{\sigma}(k/n)}I_4 = -\frac{\sqrt{n}}{\tilde{\sigma}(k/n)}\int_{X_{n,n-k}}^{\bar{F}^{-}(\frac{k}{n})}\bar{F}(x)\,d\psi(\tfrac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}})$$
$$= -\frac{\sqrt{n}}{\tilde{\sigma}(k/n)}\bar{F}(\bar{F}^{-}(\tfrac{k}{n}))\int_{X_{n,n-k}}^{\bar{F}^{-}(\frac{k}{n})}\,d\psi(\tfrac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}})(1+o_p(1))$$
$$= \frac{\sqrt{n}}{\tilde{\sigma}(k/n)}\tfrac{k}{n}\{\psi(\tfrac{X_{n,n-k}-\beta_{n0}}{\theta_{n0}-\beta_{n0}})-\psi(\tfrac{\bar{F}^{-}(\frac{k}{n})-\beta_{n0}}{\theta_{n0}-\beta_{n0}})\}(1+o_p(1))$$
$$= -\frac{\sqrt{n}}{\tilde{\sigma}(k/n)}I_1 + o_p(1). \tag{4.5.41}$$

Hence (4.5.33) follows from (4.5.38)–(4.5.41).

For proving (4.5.34), write

$$\int_{X_{n,n-k}}^{\infty}\psi(\tfrac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}})\,d\{1-\hat{\bar{F}}(x)\} - \int_{\bar{F}^{-}(k/n)}^{\infty}\psi(\tfrac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}})\,dF(x)$$
$$= \int_{X_{n,n-k}}^{\infty}\psi(\tfrac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}})\,d\{1-\tfrac{k}{n}(\tfrac{x}{X_{n,n-k}})^{-1/\hat{\gamma}(k)}-1+\tfrac{k}{n}(\tfrac{x}{X_{n,n-k}})^{-1/\gamma}\}$$
$$+\int_{X_{n,n-k}}^{\infty}\psi(\tfrac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}})\,d\{1-\tfrac{k}{n}(\tfrac{x}{X_{n,n-k}})^{-1/\gamma}-F(x)\}$$
$$+\int_{X_{n,n-k}}^{\bar{F}^{-}(k/n)}\psi(\tfrac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}})\,dF(x)$$
$$= II_1 + II_2 + II_3. \tag{4.5.42}$$

82

Since

$$
\begin{aligned}
&\sqrt{k}\{\hat{\gamma}(k) - \gamma\} \\
&= -\gamma\sqrt{\tfrac{n}{k}} \int_0^1 \{\tfrac{B_n(1-\tfrac{k}{n}s)}{s} - B_n(1-\tfrac{k}{n})\}\, ds - \tfrac{\lambda}{1+\rho} + o_p(1),
\end{aligned} \tag{4.5.43}
$$

we have

$$
\begin{aligned}
\tfrac{\sqrt{n}}{\tilde{\sigma}(k/n)} II_1
&= -\tfrac{\sqrt{n}}{\tilde{\sigma}(k/n)} \tfrac{k}{n} \int_1^\infty \psi(\tfrac{X_{n,n-k}x - \beta_{n0}}{\theta_{n0}-\beta_{n0}})\, d\{x^{-1/\hat{\gamma}(k)} - x^{-1/\gamma}\} \\
&= -\tfrac{\sqrt{k/n}}{\tilde{\sigma}(k/n)} \sqrt{k}(\tfrac{1}{\hat{\gamma}(k)} - \tfrac{1}{\gamma}) \int_1^\infty \psi(\tfrac{X_{n,n-k}x - \beta_{n0}}{\theta_{n0}-\beta_{n0}})\, d\{x^{-\alpha}\log x\}\{1 + o_p(1)\} \\
&= -\tfrac{\sqrt{k/n}}{\tilde{\sigma}(k/n)} \sqrt{k}(\tfrac{1}{\hat{\gamma}(k)} - \tfrac{1}{\gamma})d_0 \int_1^\infty (\tfrac{X_{n,n-k}x - \beta_{n0}}{\theta_{n0}-\beta_{n0}})^r\, d\{x^{-\alpha}\log x\}\{1 + o_p(1)\} \\
&= \tfrac{\sqrt{k/n}}{\tilde{\sigma}(k/n)} \{\tfrac{\bar{F}^-(k/n)}{\bar{F}^-(1-q_n)}\}^r \tfrac{d_0}{(c_1-c_2)^r} \sqrt{k}(\tfrac{1}{\hat{\gamma}(k)} - \tfrac{1}{\gamma}) \int_1^\infty x^r d\{x^{-\alpha}\log x\}\{1 + o_p(1)\} \\
&= -\sqrt{\tfrac{2r-\alpha}{2r}} \tfrac{r}{(\alpha-r)^2} I(\alpha < 2r) \\
&\quad \times \{\alpha\sqrt{\tfrac{n}{k}} \int_0^1 \{\tfrac{B_n(1-\tfrac{k}{n}s)}{s} - B_n(1 - \tfrac{k}{n})\}\, ds + \tfrac{\alpha^2\lambda}{1+\rho}\}\{1 + o_p(1)\} \\
&= -\sqrt{\tfrac{2r-\alpha}{2r}} \tfrac{\alpha r}{(\alpha-r)^2} I(\alpha < 2r)\sqrt{\tfrac{n}{k}} \int_0^1 \{\tfrac{B_n(1-\tfrac{k}{n}s)}{s} - B_n(1 - \tfrac{k}{n})\}\, ds \\
&\quad - \tfrac{r\alpha^2\lambda}{(\alpha-r)^2(1+\rho)} \sqrt{\tfrac{2r-\alpha}{2r}} I(\alpha < 2r) + o_p(1),
\end{aligned} \tag{4.5.44}
$$

$$
\begin{aligned}
\tfrac{\sqrt{n}}{\tilde{\sigma}(k/n)} II_2
&= \tfrac{\sqrt{n}}{\tilde{\sigma}(k/n)} \int_1^\infty \psi(\tfrac{X_{n,n-k}x - \beta_{n0}}{\theta_{n0}-\beta_{n0}})\, d\{\bar{F}(X_{n,n-k}x) - \tfrac{k}{n}x^{-1/\gamma}\} \\
&= \tfrac{\sqrt{n}}{\tilde{\sigma}(k/n)} \tfrac{k}{n} \int_1^\infty \psi(\tfrac{X_{n,n-k}x - \beta_{n0}}{\theta_{n0}-\beta_{n0}})\, d\{\tfrac{\bar{F}(X_{n,n-k}x)}{k/n} - x^{-\alpha}\} \\
&= \tfrac{\sqrt{n}}{\tilde{\sigma}(k/n)} \tfrac{k}{n} \{(\tfrac{X_{n,n-k}}{\bar{F}^-(\tfrac{k}{n})})^{-\alpha} - 1\} \int_1^\infty \psi(\tfrac{X_{n,n-k}x - \beta_{n0}}{\theta_{n0}-\beta_{n0}})\, d\{x^{-\alpha}\}\{1 + o_p(1)\} \\
&= \tfrac{\sqrt{n}}{\tilde{\sigma}(k/n)} \tfrac{k}{n} \{(\tfrac{X_{n,n-k}}{\bar{F}^-(\tfrac{k}{n})})^{-\alpha} - 1\}d_0 \int_1^\infty (\tfrac{X_{n,n-k}x - \beta_{n0}}{\theta_{n0}-\beta_{n0}})^r\, d\{x^{-\alpha}\}\{1 + o_p(1)\} \\
&= \tfrac{\sqrt{k/n}\{\bar{F}^-(k/n)/\bar{F}^-(1-q_n)\}^r}{\tilde{\sigma}(k/n)} \tfrac{d_0}{(c_1-c_2)^r} \\
&\quad \times \sqrt{k}\{\tfrac{n}{k}(1 - U_{n.n-k}) - 1\} \int_1^\infty x^r\, d\{x^{-\alpha}\}\{1 + o_p(1)\} \\
&= \sqrt{\tfrac{2r-\alpha}{2r}} \tfrac{\alpha}{r-\alpha} I(\alpha < 2r)\sqrt{\tfrac{n}{k}} B_n(1 - \tfrac{k}{n}) + o_p(1),
\end{aligned} \tag{4.5.45}
$$

$$
\begin{aligned}
\frac{\sqrt{n}}{\tilde{\sigma}(k/n)} II_3 \;&=\; \frac{\sqrt{n}}{\tilde{\sigma}(k/n)} \int_{X_{n,n-k}}^{\bar{F}^-(k/n)} \psi\Big(\frac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}}\Big)\, dF(x) \\
&=\; \frac{\sqrt{n}}{\tilde{\sigma}(k/n)} \psi\Big(\frac{\bar{F}^-(k/n)-\beta_{n0}}{\theta_{n0}-\beta_{n0}}\Big)\{F(\bar{F}^-(k/n)) - F(X_{n,n-k})\}\{1+o_p(1)\} \\
&=\; \frac{\sqrt{n}}{\tilde{\sigma}(k/n)} \Big(\frac{\bar{F}^-(k/n)-\beta_{n0}}{\bar{F}^-(\frac{k}{n})}\Big)^r \Big\{\frac{\bar{F}^-(\frac{k}{n})}{\bar{F}^-(1-q_n)}\Big\}^r \frac{d_0}{(c_1-c_2)^r}\big(1-\tfrac{k}{n}-U_{n,n-k}\big)\{1+o_p(1)\} \quad (4.5.46) \\
&=\; \frac{\sqrt{k/n}\{\bar{F}^-(k/n)/\bar{F}^-(1-q_n)\}^r}{\tilde{\sigma}(k/n)} \frac{d_0}{(c_1-c_2)^r} \sqrt{k}\big\{\tfrac{n}{k}(1-U_{n,n-k})-1\big\}\{1+o_p(1)\} \\
&=\; \sqrt{\frac{2r-\alpha}{2r}}\, I(\alpha<2r)\sqrt{\tfrac{n}{k}}\, B_n\big(1-\tfrac{k}{n}\big) + o_p(1).
\end{aligned}
$$

Hence (4.5.34) follows from (4.5.44)–(4.5.46).

Note that

$$
\frac{\partial \Delta_{I11}(\theta_{n0},\beta_{n0})}{\partial \theta} = -\int_{\beta_{n0}}^{X_{n,n-k}} \psi'\Big(\frac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}}\Big)\frac{x-\beta_{n0}}{(\theta_{n0}-\beta_{n0})^2}\, dF_n(x)
$$

and

$$
\frac{\partial \Delta_{I12}(\theta_{n0},\beta_{n0})}{\partial \theta} = -\int_{X_{n,n-k}}^{\infty} \psi'\Big(\frac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}}\Big)\frac{x-\beta_{n0}}{(\theta_{n0}-\beta_{n0})^2}\, d\hat{\bar{F}}(x).
$$

Similar to the proof of Lemma 4.5.3, we can show that

$$
\int_{\beta_{n0}}^{X_{n,n-k}} \psi'\Big(\frac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}}\Big)\frac{x-\beta_{n0}}{(\theta_{n0}-\beta_{n0})^2}\, dF_n(x) - \int_{\beta_{n0}}^{\bar{F}^-(\frac{k}{n})} \psi'\Big(\frac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}}\Big)\frac{x-\beta_{n0}}{(\theta_{n0}-\beta_{n0})^2}\, dF(x) \xrightarrow{p} 0
$$

and

$$
\int_{X_{n,n-k}}^{\infty} \psi'\Big(\frac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}}\Big)\frac{x-\beta_{n0}}{(\theta_{n0}-\beta_{n0})^2}\, d\{1-\hat{\bar{F}}(x)\} \xrightarrow{p} 0.
$$

In the same way as above, we can prove the results for the derivatives with respect to β. Furthermore, we have

$$
\begin{aligned}
&\int_{\beta_{n0}}^{\infty} \psi'\Big(\frac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}}\Big)\frac{x-\beta_{n0}}{(\theta_{n0}-\beta_{n0})^2}\, dF(x) \\
&=\; -\frac{\bar{F}(\beta_{n0})}{\theta_{n0}-\beta_{n0}} \int_1^{\infty} \psi'\Big(\frac{x-1}{\theta_{n0}/\beta_{n0}-1}\Big)\, d\Big\{\frac{\bar{F}(\beta_0 x)}{\bar{F}(\beta_0)}\Big\} \\
&=\; \alpha\frac{\bar{F}(\beta_{n0})}{\theta_{n0}-\beta_{n0}} \int_1^{\infty} x^{-\alpha-1}\psi'\Big(\frac{x-1}{c_1/c_2-1}\Big)\, dx\, \{1+o(1)\} \\
&=\; \frac{1-q_n}{\bar{F}^-(1-q_n)} \alpha c_2^{-\alpha-1} \int_0^{\infty} \big(1+\tfrac{c_1-c_2}{c_2}x\big)^{-\alpha-1}\psi'(x)\, dx\, \{1+o(1)\}
\end{aligned}
\tag{4.5.47}
$$

and

$$\int_{\beta_{n0}}^{\infty} \psi'\left(\frac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}}\right)\frac{x-\theta_{n0}}{(\theta_{n0}-\beta_{n0})^2}\,dF(x)$$
$$= (\theta_{n0}-\beta_{n0})^{-2}\int_{\beta_{n0}}^{\infty}\psi'\left(\frac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}}\right)(x-\theta_{n0})\,dF(x)$$
$$= (\theta_{n0}-\beta_{n0})^{-2}E[\psi'\left(\frac{X-\beta_{n0}}{\theta_{n0}-\beta_{n0}}\right)(X-\theta_{n0})I(X>\beta_{n0})]$$
$$= 0, \tag{4.5.48}$$

which imply (4.5.35) and (4.5.36). ∎

Lemma 4.5.6. *Under conditions of Theorem 4.2.4, we have*

$$\frac{\sqrt{n}}{\tilde\sigma(k/n)\bar F^-(1-q_n)}\{\Delta_{I21}(\theta_{n0},\beta_{n0})-\int_{\beta_{n0}}^{\bar F^-(k/n)}\psi'\left(\frac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}}\right)(x-\theta_{n0})\,dF(x)\}$$
$$= -\frac{(c_1-c_2)\int_{F(\beta_{n0})}^{1-\frac{k}{n}}B_n(x)\,d\{\psi'\left(\frac{F^-(x)-\beta_{n0}}{\theta_{n0}-\beta_{n0}}\right)\frac{F^-(x)-\theta_{n0}}{\theta_{n0}-\beta_{n0}}\}}{\tilde\sigma(k/n)}+o_p(1), \tag{4.5.49}$$

$$\frac{\sqrt{n}}{\tilde\sigma(k/n)\bar F^-(1-q_n)}\{\Delta_{I22}(\theta_{n0},\beta_{n0})-\int_{\bar F^-(k/n)}^{\infty}\psi'\left(\frac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}}\right)(x-\theta_{n0})\,dF(x)\}$$
$$= -(c_1-c_2)\sqrt{\frac{2r-\alpha}{2r}}I(\alpha<2r)\frac{\alpha r^2}{(r-\alpha)^2}\sqrt{\frac{n}{k}}\int_0^1\frac{B_n(1-\frac{k}{n}s)}{s}\,ds$$
$$+(c_1-c_2)\sqrt{\frac{2r-\alpha}{2r}}I(\alpha<2r)\frac{r^3}{r-\alpha}\sqrt{\frac{n}{k}}B_n(1-\frac{k}{n})$$
$$-(c_1-c_2)\sqrt{\frac{2r-\alpha}{2r}}I(\alpha<2r)\frac{\alpha^2 r^2\lambda}{(r-\alpha)^2(1+\rho)}+o_p(1), \tag{4.5.50}$$

$$\frac{\partial}{\partial\theta}\{\Delta_{I21}(\theta_{n0},\beta_{n0})+\Delta_{I22}(\theta_{n0},\beta_{n0})\}$$
$$= -(1-q_n)\alpha c_2^{-\alpha-1}(c_1-c_2)\{\int_0^{\infty}x(x-1)\psi''(x)(1+\frac{c_1-c_2}{c_2}x)^{-\alpha-1}\,dx$$
$$+\int_0^{\infty}\psi'(x)(1+\frac{c_1-c_2}{c_2}x)^{-\alpha-1}\,dx\}\{1+o_p(1)\}, \tag{4.5.51}$$

$$\frac{\partial}{\partial\beta}\{\Delta_{I21}(\theta_{n0},\beta_{n0})+\Delta_{I22}(\theta_{n0},\beta_{n0})\}$$
$$= (1-q_n)\alpha c_2^{-\alpha-1}(c_1-c_2)\int_0^{\infty}(x-1)^2\psi''(x)(1+\frac{c_1-c_2}{c_2}x)^{-\alpha-1}\,dx\{1+o_p(1)\}. \tag{4.5.52}$$

Proof. The proofs of the first two equations are similar to the proofs of Lemma 4.5.5, thus are

omitted. Note that

$$\frac{\partial}{\partial\theta}\left\{\Delta_{I21}(\theta_{n0},\beta_{n0})+\Delta_{I22}(\theta_{n0},\beta_{n0})\right\}$$

$$= E\left\{-\psi''(\tfrac{X-\beta_{n0}}{\theta_{n0}-\beta_{n0}})\tfrac{(X-\beta_{n0})(x-\theta_{n0})}{(\theta_{n0}-\beta_{n0})^2}I(X>\beta_{n0})-\psi'(\tfrac{X-\beta_{n0}}{\theta_{n0}-\beta_{n0}})I(X>\beta_{n0})\right\}\{1+o_p(1)\}$$

$$= \bar{F}(\beta_{n0})\left(\int_1^\infty\psi''(\tfrac{x-1}{c_1/c_2-1})\tfrac{(x-1)(x-\frac{c_1}{c_2})}{(c_1/c_2-1)^2}\,d\{x^{-\alpha}\}+\int_1^\infty\psi'(\tfrac{x-1}{c_1/c_2-1})\,d\{x^{-\alpha}\}\right)\{1+o_p(1)\}$$

$$= -(1-q_n)\alpha c_2^{-\alpha-1}(c_1-c_2)$$

$$\times\int_0^\infty(1+\tfrac{c_1-c_2}{c_2}x)^{-\alpha-1}\{x(x-1)\psi''(x)+\psi'(x)\}\,dx\,\{1+o_p(1)\}$$

(4.5.53)

and

$$\frac{\partial}{\partial\beta}\left\{\Delta_{I21}(\theta_{n0},\beta_{n0})+\Delta_{I22}(\theta_{n0},\beta_{n0})\right\}$$

$$= E\left\{\psi''(\tfrac{X-\beta_{n0}}{\theta_{n0}-\beta_{n0}})\tfrac{(X-\theta_{n0})^2}{(\theta_{n0}-\beta_{n0})^2}I(X>\beta_{n0})\right\}\{1+o_p(1)\}$$

$$= -\bar{F}(\beta_{n0})\int_1^\infty\psi''(\tfrac{x-1}{c_1/c_2-1})(\tfrac{x-\frac{c_1}{c_2}}{c_1/c_2-1})^2\,d\{x^{-\alpha}\}\{1+o_p(1)\}$$

$$= (1-q_n)\alpha c_2^{-\alpha-1}(c_1-c_2)\int_0^\infty(x-1)^2\psi''(x)(1+\tfrac{c_1-c_2}{c_2}x)^{-\alpha-1}\,dx\,\{1+o_p(1)\},$$

(4.5.54)

which lead to the last two equations in the lemma. ∎

Proof of Theorem 4.2.4. When $|\beta/\beta_{n0}-1|+|\theta/\theta_{n0}-1|=\epsilon_n\to 0$, by Taylor expansions and Lemmas 4.5.5 and 4.5.6, we have

$$\frac{\sqrt{n}}{\tilde{\sigma}(k/n)}\{\Delta_{I11}(\theta,\beta)-\Delta_{I11}(\theta_{n0},\beta_{n0})\}+\frac{\sqrt{n}}{\tilde{\sigma}(k/n)}\{\Delta_{I11}(\theta_{n0},\beta_{n0})-\int_{\beta_{n0}}^{\bar{F}^-(k/n)}\psi(\tfrac{x-\beta_{n0}}{\theta_{n0}-\beta_{nn0}})\,dF(x)\}$$

$$+\frac{\sqrt{n}}{\tilde{\sigma}(k/n)}\{\Delta_{I12}(\theta,\beta)-\Delta_{I12}(\theta_{n0},\beta_0)\}+\frac{\sqrt{n}}{\tilde{\sigma}(k/n)}\{\Delta_{I12}(\theta_{n0},\beta_{n0})-\int_{\bar{F}^-(k/n)}^\infty\psi(\tfrac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}})\,dF(x)\}$$

$$= -\frac{\sqrt{n}}{\tilde{\sigma}(k/n)}\frac{1-q_n}{\bar{F}^-(1-q_n)}(\theta-\theta_{n0})a_{I1}-\frac{\int_{F(\beta_{n0})}^{1-\frac{k}{n}}B_n(x)\,d\psi(\frac{F^-(x)-\beta_{n0}}{\theta_{n0}-\beta_{n0}})}{\tilde{\sigma}(k/n)}$$

$$-\sqrt{\tfrac{2r-\alpha}{2r}}I(\alpha<2r)\tfrac{r\alpha}{(r-\alpha)^2}\sqrt{\tfrac{n}{k}}\int_0^1\tfrac{B_n(1-\frac{k}{n}s)}{s}\,ds$$

$$+\sqrt{\tfrac{2r-\alpha}{2r}}I(\alpha<2r)\tfrac{r^2}{r-\alpha}\sqrt{\tfrac{n}{k}}B_n(1-\tfrac{k}{n})$$

$$-\sqrt{\tfrac{2r-\alpha}{2r}}I(\alpha<2r)\tfrac{r\alpha^2\lambda}{(\alpha-r)^2(1+\rho)}+o_p(1).$$

86

Similarly

$$\frac{\sqrt{n}}{\tilde{\sigma}(k/n)\bar{F}^-(1-q_n)}\{\Delta_{I21}(\theta,\beta)-\Delta_{I21}(\theta_{n0},\beta_{n0})\}$$

$$+\frac{\sqrt{n}}{\tilde{\sigma}(k/n)\bar{F}^-(1-q_n)}\{\Delta_{I21}(\theta_{n0},\beta_{n0})-\int_{\beta_{n0}}^{\bar{F}^-(k/n)}\psi'(\frac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}})(x-\theta_{n0})\,dF(x)\}$$

$$+\frac{\sqrt{n}}{\tilde{\sigma}(k/n)\bar{F}^-(1-q_n)}\{\Delta_{I22}(\theta,\beta)-\Delta_{I22}(\theta_{n0},\beta_{n0})\}$$

$$+\frac{\sqrt{n}}{\tilde{\sigma}(k/n)\bar{F}^-(1-q_n)}\{\Delta_{I22}(\theta_{n0},\beta_{n0})-\int_{\bar{F}^-(k/n)}^{\infty}\psi'(\frac{x-\beta_{n0}}{\theta_{n0}-\beta_{n0}})(x-\theta_{n0})\,dF(x)\}$$

$$=\ -\frac{\sqrt{n}}{\tilde{\sigma}(k/n)}\frac{1-q_n}{\bar{F}^-(1-q_n)}(\theta-\theta_{n0})(c_1-c_2)a_{I2}-\frac{\sqrt{n}}{\tilde{\sigma}(k/n)}\frac{1-q_n}{\bar{F}^-(1-q_n)}(\beta-\beta_{n0})(c_1-c_2)b_{I2}$$

$$-\frac{c_1-c_2}{\tilde{\sigma}(k/n)}\int_{F(\beta_{n0})}^{1-\frac{k}{n}}B_n(x)\,d\{\psi'(\frac{F^-(x)-\beta_{n0}}{\theta_{n0}-\beta_{n0}})\frac{F^-(x)-\theta_{n0}}{\theta_{n0}-\beta_{n0}}\}$$

$$-(c_1-c_2)\sqrt{\frac{2r-\alpha}{2r}}I(\alpha<2r)\frac{\alpha r^2}{(r-\alpha)^2}\sqrt{\frac{n}{k}}\int_0^1\frac{B_n(1-\frac{k}{n}s)}{s}\,ds$$

$$+(c_1-c_2)\sqrt{\frac{2r-\alpha}{2r}}I(\alpha<2r)\frac{r^3}{r-\alpha}\sqrt{\frac{n}{k}}B_n(1-\frac{k}{n})$$

$$-(c_1-c_2)\sqrt{\frac{2r-\alpha}{2r}}I(\alpha<2r)\frac{\alpha^2 r^2\lambda}{(r-\alpha)^2(1+\rho)}+o_p(1).$$

Hence

$$\begin{cases}\frac{\sqrt{n}}{\tilde{\sigma}(k/n)}\frac{1-q_n}{\bar{F}^-(1-q_n)}(\hat{\theta}-\theta_{n0})a_{I1}=\xi_{In}+o_p(1)\\[2mm]\frac{\sqrt{n}}{\tilde{\sigma}(k/n)}\frac{1-q_n}{\bar{F}^-(1-q_n)}(\hat{\theta}-\theta_{n0})a_{I2}+\frac{\sqrt{n}}{\tilde{\sigma}(k/n)}\frac{1-q_n}{\bar{F}^-(1-q_n)}(\hat{\beta}-\beta_{n0})b_{I2}=\eta_{In}+o_p(1),\end{cases}\tag{4.5.55}$$

where

$$\begin{aligned}\xi_{In}\ =\ &-\frac{\int_{F(\beta_{n0})}^{1-\frac{k}{n}}B_n(x)\,d\psi(\frac{F^-(x)-\beta_{n0}}{\theta_{n0}-\beta_{n0}})}{\tilde{\sigma}(k/n)}\\&-\sqrt{\frac{2r-\alpha}{2r}}I(\alpha<2r)\frac{r\alpha}{(r-\alpha)^2}\sqrt{\frac{n}{k}}\int_0^1\frac{B_n(1-\frac{k}{n}s)}{s}\,ds\\&+\sqrt{\frac{2r-\alpha}{2r}}I(\alpha<2r)\frac{r^2}{r-\alpha}\sqrt{\frac{n}{k}}B_n(1-\frac{k}{n})\\&-\sqrt{\frac{2r-\alpha}{2r}}I(\alpha<2r)\frac{r\alpha^2\lambda}{(\alpha-r)^2(1+\rho)}\end{aligned}\tag{4.5.56}$$

and

$$\begin{aligned}\eta_{In}\ =\ &-\frac{\int_{F(\beta_{n0})}^{1-\frac{k}{n}}B_n(x)\,d\{\psi'(\frac{F^-(x)-\beta_{n0}}{\theta_{n0}-\beta_{n0}})\frac{F^-(x)-\theta_{n0}}{\theta_{n0}-\beta_{n0}}\}}{\tilde{\sigma}(k/n)}\\&-\sqrt{\frac{2r-\alpha}{2r}}I(\alpha<2r)\frac{\alpha r^2}{(r-\alpha)^2}\sqrt{\frac{n}{k}}\int_0^1\frac{B_n(1-\frac{k}{n}s)}{s}\,ds\\&+\sqrt{\frac{2r-\alpha}{2r}}I(\alpha<2r)\frac{r^3}{r-\alpha}\sqrt{\frac{n}{k}}B_n(1-\frac{k}{n})\\&-\sqrt{\frac{2r-\alpha}{2r}}I(\alpha<2r)\frac{\alpha^2 r^2\lambda}{(r-\alpha)^2(1+\rho)}.\end{aligned}\tag{4.5.57}$$

Note that

$$(\xi_{In},\eta_{In})^T\xrightarrow{d}N(\Delta_I,\Sigma_I),\quad n\to\infty,\tag{4.5.58}$$

where

$$\Delta_I = \begin{pmatrix} -\sqrt{\frac{2r-\alpha}{2r}}I(\alpha < 2r)\frac{r\alpha^2\lambda}{(\alpha-r)^2(1+\rho)} \\ -\sqrt{\frac{2r-\alpha}{2r}}I(\alpha < 2r)\frac{r^2\alpha^2\lambda}{(r-\alpha)^2(1+\rho)} \end{pmatrix}$$

and

$$\Sigma_I = \begin{pmatrix} \sigma_{I1}^2 & \sigma_{I12} \\ \sigma_{I21} & \sigma_{I2}^2 \end{pmatrix}$$

with $\sigma_{I1}^2, \sigma_{I2}^2$ and σ_{I12} calculated as follows. Note that

$$\lim_{n\to\infty} \frac{\sqrt{k/n}\int_{F(\beta_{n0})}^{1-\frac{k}{n}} x\,d\psi(\frac{F^-(x)-\beta_{n0}}{\theta_{n0}-\beta_{n0}})}{\tilde{\sigma}(k/n)}$$

$$= \lim_{n\to\infty} \frac{\sqrt{k/n}}{\tilde{\sigma}(k/n)}\psi(\frac{\bar{F}^-(k/n)-\beta_{n0}}{\theta_{n0}-\beta_{n0}})$$

$$= d_0\lim_{n\to\infty} \frac{\sqrt{k/n}}{\tilde{\sigma}(k/n)}(\frac{\bar{F}^-(k/n)-\beta_{n0}}{\theta_{n0}-\beta_{n0}})^r$$

$$= \sqrt{\frac{2r-\alpha}{2r}}I(\alpha < 2r).$$

Similarly, we can obtain

$$\lim_{n\to\infty} \frac{\sqrt{\frac{k}{n}}\int_{F(\beta_{n0})}^{1-k/n} s\,d\{\psi'(\frac{F^-(s)-\beta_{n0}}{\theta_{n0}-\beta_{n0}})\frac{F^-(s)-\beta_{n0}}{\theta_{n0}-\theta_{n0}}\}}{\tilde{\sigma}(k/n)} = r\sqrt{\frac{2r-\alpha}{2r}}I(\alpha < 2r).$$

Then we have

$$\sigma_{I1}^2 = \lim_{n\to\infty} Var(\xi_{In})$$

$$= 1 + \frac{2r-\alpha}{2r}\frac{r^4}{(r-\alpha)^4}I(\alpha < 2r) + \frac{2r-\alpha}{2r}\frac{2r^2\alpha^2}{(r-\alpha)^4}I(\alpha < 2r)$$

$$-\sqrt{\frac{2r-\alpha}{2r}}\frac{2r^2}{(r-\alpha)^2}I(\alpha < 2r)\lim_{n\to\infty}\frac{\sqrt{\frac{k}{n}}\int_{F(\beta_{n0})}^{1-k/n} s\,d\psi(\frac{F^-(s)-\beta_{n0}}{\theta_{n0}-\beta_{n0}})}{\tilde{\sigma}(k/n)}$$

$$+\sqrt{\frac{2r-\alpha}{2r}}\frac{2r\alpha}{(r-\alpha)^2}I(\alpha < 2r)\lim_{n\to\infty}\frac{\sqrt{\frac{k}{n}}\int_{F(\beta_{n0})}^{1-k/n} s\,d\psi(\frac{F^-(s)-\beta_{n0}}{\theta_{n0}-\beta_{n0}})}{\tilde{\sigma}(k/n)}$$

$$-\frac{2r-\alpha}{2r}\frac{2r^3\alpha}{(r-\alpha)^4}I(\alpha < 2r)$$

$$= 1 + (2r - \alpha)\{\frac{r^3}{2(r-\alpha)^4} - \frac{\alpha r}{(r-\alpha)^3} - \frac{1}{r-\alpha}\}I(\alpha < 2r).$$

Define

$$\tilde{\sigma}_1^2(\tfrac{k}{n})$$
$$= \int_{F(\beta_{n0})}^{1-\frac{k}{n}} \int_{F(\beta_{n0})}^{1-\frac{k}{n}} \{s \wedge t - st\} \, d\{\psi'(\tfrac{F^-(s)-\beta_{n0}}{\theta_{n0}-\beta_{n0}}) \tfrac{F^-(s)-\theta_{n0}}{\theta_{n0}-\beta_{n0}}\} d\{\psi'(\tfrac{F^-(t)-\beta_{n0}}{\theta_{n0}-\beta_{n0}}) \tfrac{F^-(t)-\theta_{n0}}{\theta_{n0}-\beta_{n0}}\},$$

$$\tilde{\sigma}_2^2(\tfrac{k}{n})$$
$$= \int_{F(\beta_{n0})}^{1-\frac{k}{n}} \int_{F(\beta_{n0})}^{1-\frac{k}{n}} \{s \wedge t - st\} \, d\psi(\tfrac{F^-(s)-\beta_{n0}}{\theta_{n0}-\beta_{n0}}) d\{\psi'(\tfrac{F^-(t)-\beta_{n0}}{\theta_{n0}-\beta_{n0}}) \tfrac{F^-(t)-\theta_{n0}}{\theta_{n0}-\beta_{n0}}\}$$

Similar to the proof of Lemma 4.5.4, we can show that

$$\lim_{n \to \infty} \frac{\tilde{\sigma}_1^2(k/n)}{(k/n)\{\bar{F}^-(k/n)/\bar{F}^-(1-q_n)\}^{2r}} = \frac{2r^3 d_0^2}{2r-\alpha}(c_1 - c_2)^{-2r}$$

and

$$\lim_{n \to \infty} \frac{\tilde{\sigma}_2^2(k/n)}{(k/n)\{\bar{F}^-(k/n)/\bar{F}^-(1-q_n)\}^{2r}} = \frac{2r^2 d_0^2}{2r-\alpha}(c_1 - c_2)^{-2r}$$

for $\alpha < 2r$,

$$\lim_{n \to \infty} \frac{\tilde{\sigma}_1^2(k/n)}{(1-q_n)\log\{\bar{F}^-(k/n)/\bar{F}^-(1-q_n)\}} = \frac{2r^3 d_0^2}{(c_1 - c_2)^{2r}}$$

and

$$\lim_{n \to \infty} \frac{\tilde{\sigma}_2^2(k/n)}{(1-q_n)\log\{\bar{F}^-(k/n)/\bar{F}^-(1-q_n)\}} = \frac{2r^2 d_0^2}{(c_1 - c_2)^{2r}}$$

for $\alpha = 2r$. For $\alpha > 2r$, recall that in the proof of Lemma 4.5.4

$$d_1 =: \lim_{n \to \infty} \frac{\tilde{\sigma}^2(k/n)}{1-q_n}$$
$$= 2c_2^{-\alpha} \int_0^\infty (1 + \tfrac{c_1-c_2}{c_2}x)^{-\alpha} \psi(x)\psi'(x) \, dx,$$

and similar arguments lead to

$$d_2 =: \lim_{n \to \infty} \frac{\tilde{\sigma}_1^2(k/n)}{1-q_n}$$
$$= 2c_2^{-\alpha} \int_0^\infty (x-1)\psi'(x)(1 + \tfrac{c_1-c_2}{c_2}x)^{-\alpha}\{(x-1)\psi''(x) + \psi'(x)\} \, dx$$

and

$$d_3 \ =: \lim_{n\to\infty} \frac{\tilde{\sigma}_2^2(k/n)}{1-q_n}$$

$$= c_2^{-\alpha} \int_0^\infty (1 + \tfrac{c_1-c_2}{c_2}x)^{-\alpha}\{\psi(x)\psi'(x) + (x-1)\psi(x)\psi''(x) + (x-1)(\psi'(x))^2\}\,dx.$$

Combining the equations above with Lemma 4.5.4, we have

$$\lim_{n\to\infty} \frac{\tilde{\sigma}_1^2(k/n)}{\tilde{\sigma}^2(k/n)} = r^2 I(\alpha \le 2r) + \frac{d_2}{d_1}I(\alpha > 2r)$$

and

$$\lim_{n\to\infty} \frac{\tilde{\sigma}_2^2(k/n)}{\tilde{\sigma}^2(k/n)} = r I(\alpha \le 2r) + \frac{d_3}{d_1}I(\alpha > 2r).$$

Therefore

$$\sigma_{I2}^2 \ = \lim_{n\to\infty} Var(\eta_{In})$$

$$= r^2 I(\alpha \le 2r) + \tfrac{d_2}{d_1}I(\alpha > 2r) \tag{4.5.59}$$

$$+r^2(2r-\alpha)\{\tfrac{r^3}{2(r-\alpha)^4} - \tfrac{\alpha r}{(r-\alpha)^3} - \tfrac{1}{r-\alpha}\}I(\alpha < 2r).$$

A similar calculation leads to

$$\sigma_{I12} \ = \sigma_{I21} = \lim_{n\to\infty} Cov(\xi_{In}, \eta_{In})$$

$$= r I(\alpha \le 2r) + \tfrac{d_3}{d_1}I(\alpha > 2r)$$

$$+r(2r-\alpha)\{\tfrac{r^3}{2(r-\alpha)^4} - \tfrac{\alpha r}{(r-\alpha)^3} - \tfrac{1}{r-\alpha}\}I(\alpha < 2r).$$

Hence, the theorem follows. ∎
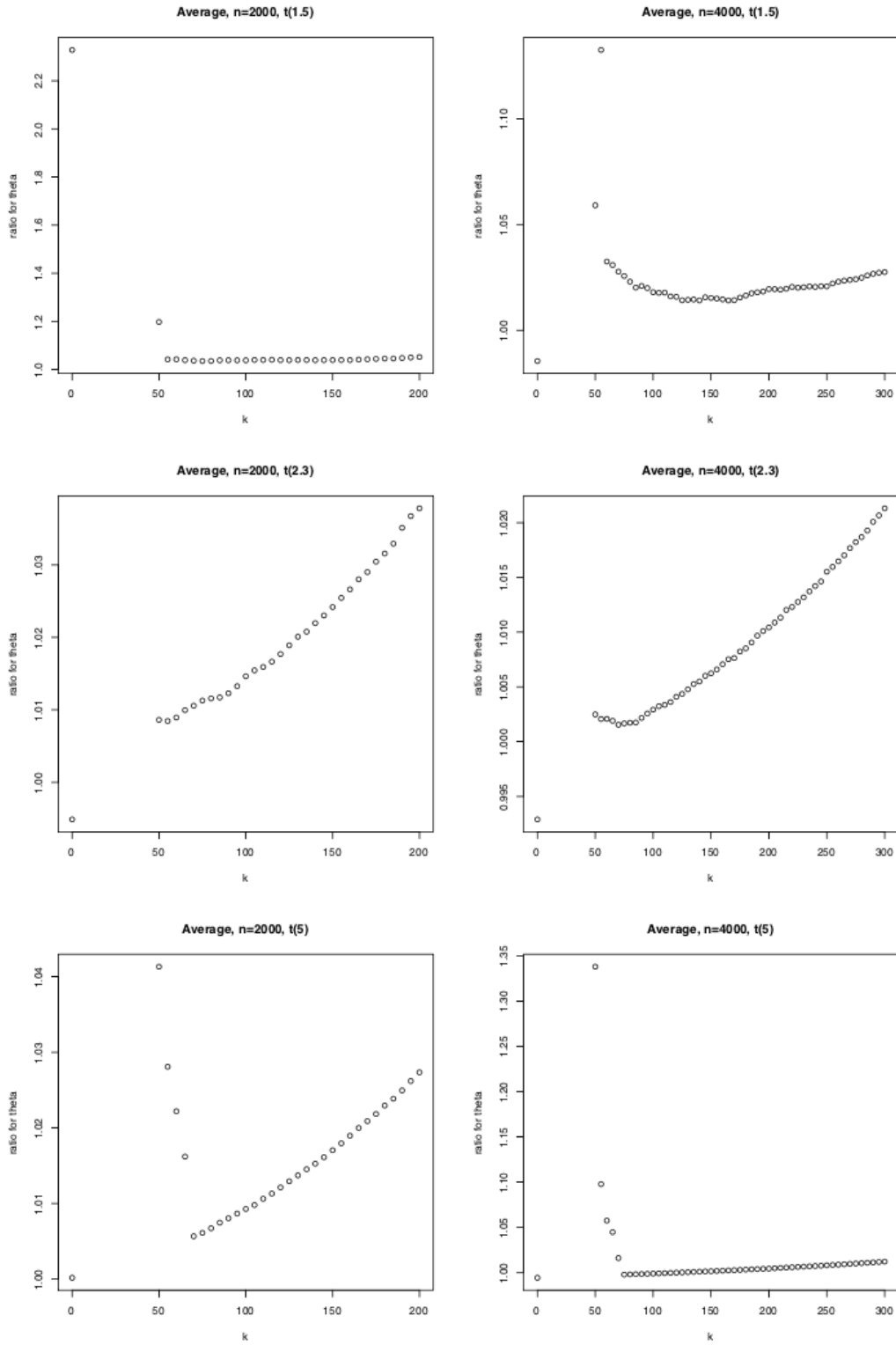
Figure 4.2: Averages of $\hat{\theta}/\theta_0$ and $\hat{\theta}^{AH}/\theta_0$ are plotted against $k = 50, 55, \cdots, 200$ for $n = 2000$ and against $k = 50, 55, \cdots, 300$ for $n = 4000$, where $k = 0$ represents the average of $\hat{\theta}^{AH}/\theta_0$.
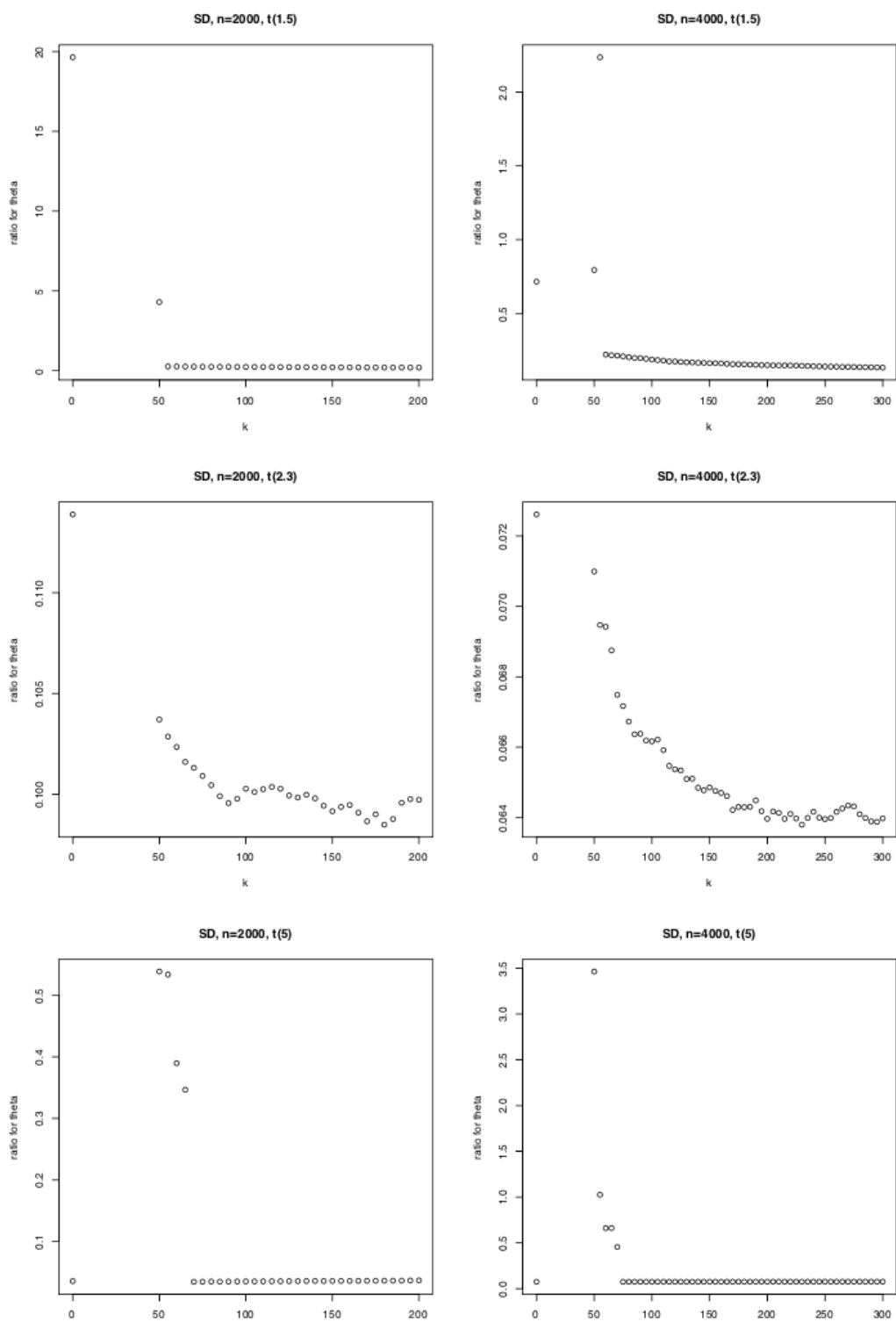
Figure 4.3: Standard deviations of $\hat{\theta}/\theta_0$ and $\hat{\theta}^{AH}/\theta_0$ are plotted against $k = 50, 55, \cdots, 200$ for $n = 2000$ and against $k = 50, 55, \cdots, 300$ for $n = 4000$, where $k = 0$ represents the standard deviation of $\hat{\theta}^{AH}/\theta_0$.

92
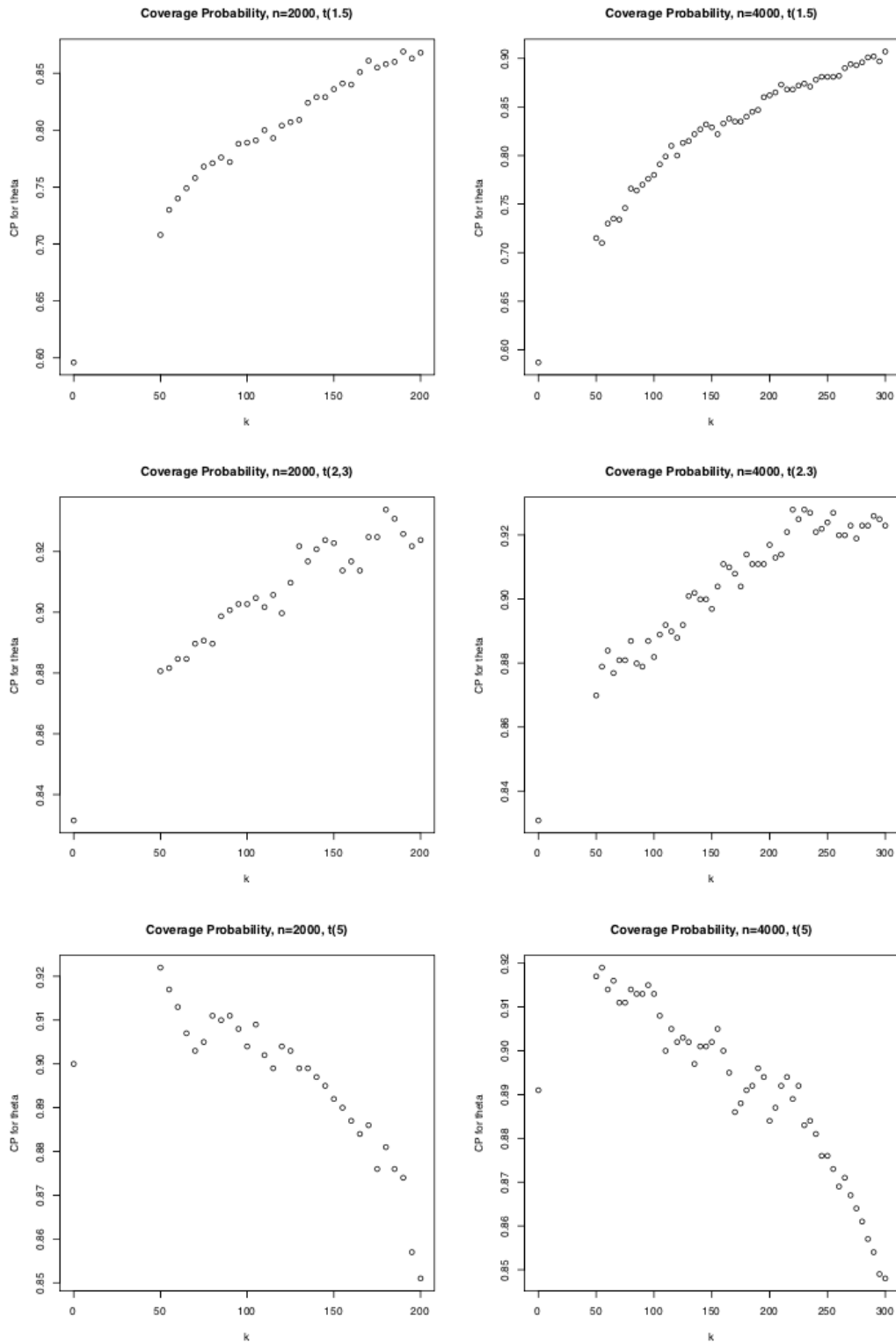
Figure 4.4: Coverage probabilities for intervals with level 90% based on $\hat{\theta}$, $\hat{\theta}^{AH}$ and bootstrap method with 1000 repetitions are plotted against $k = 50, 55, \cdots, 200$ for $n = 2000$ and against $k = 50, 55, \cdots, 300$ for $n = 4000$, where $k = 0$ represents the coverage probability based on $\hat{\theta}^{AH}$.
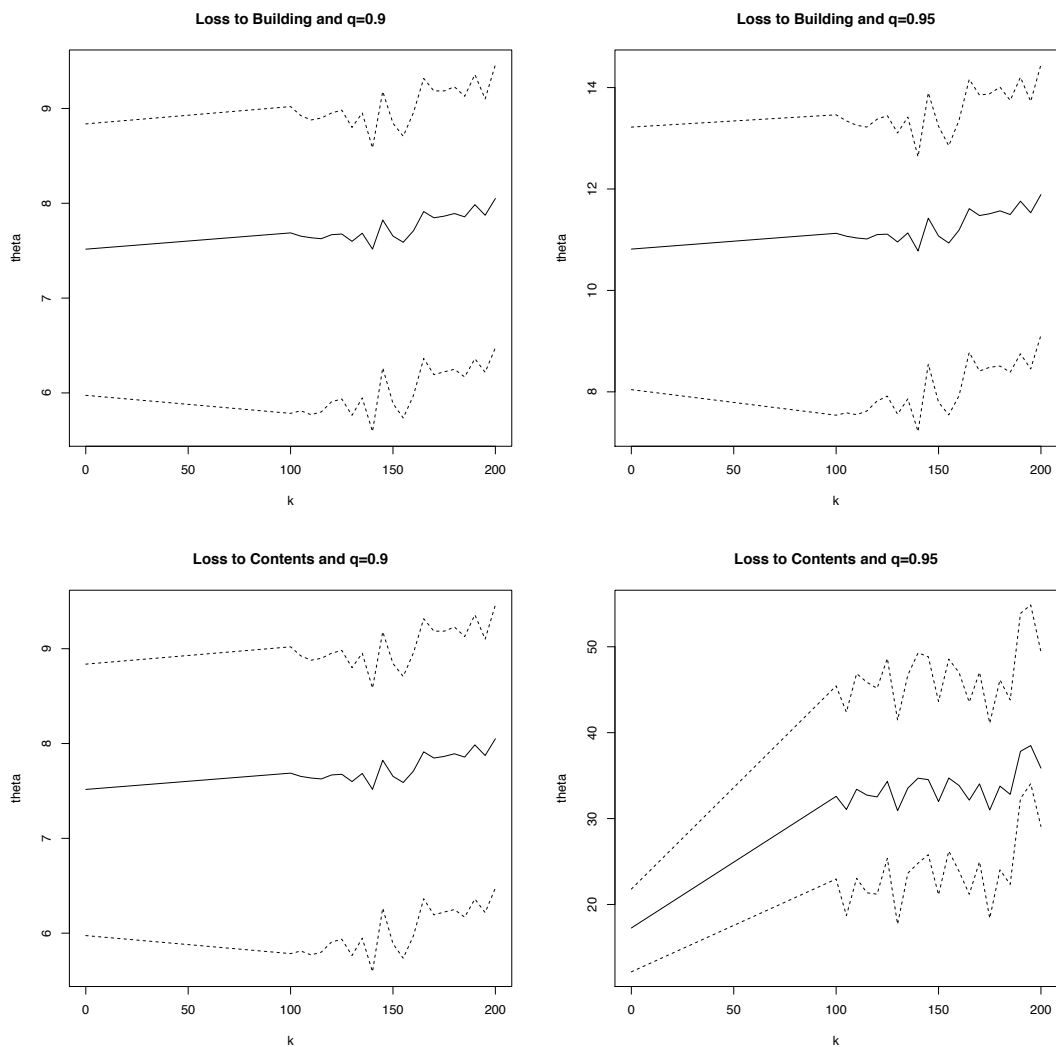
Figure 4.5: Estimators (solid line) and confidence intervals (dotted line) with level 0.9 for $q = 0.9, 0.95$ based on bootstrap method with 1000 repetitions are plotted against $k = 100, 105, \cdots, 200$ for losses to building, where $k = 0$ represents the estimator $\hat{\theta}^{AH}$.

CHAPTER 5

NONPARAMETRIC INFERENCE FOR SENSITIVITY OF THE

HAEZENDONCK-GOOVAERTS RISK MEASURE

When H-G risk measure is applied to an insurance or a financial portfolio with several loss variables, sensitivity analysis becomes useful in managing the portfolio, and the assumption of independent observations may not be reasonable. This chapter (Chapter 5) first derives an expression for computing the sensitivity of the H-G risk measure, which enables us to estimate the sensitivity non-parametrically via the H-G risk measure. Further, we derive the asymptotic distributions of the nonparametric estimators for the H-G risk measure and the sensitivity by assuming that loss variables in the portfolio follow from a strictly stationary α-mixing sequence. A simulation study is provided to examine the finite sample performance of the proposed nonparametric estimators. Finally, the method is applied to a real data set. The content of this chapter is based on the joint work:

X. Wang, Q. Liu, Y. Hou and L. Peng (2018). Nonparametric inference for sensitivity of Haezendonck-Goovaerts risk measure. Scandinavian Actuarial Journal. To appear.

5.1 Introduction to Sensitivity Analysis of H-G Risk Measure

Haezendonck-Goovaerts (H-G) risk measure originates from the premium calculation principle induced by an Orlicz norm in Haezendonck and Goovaerts (1982). More specifically, let ψ : $[0, \infty] \to [0, \infty]$ be a convex function satisfying $\psi(0) = 0, \psi(1) = 1$ and $\psi(\infty) = \infty$, i.e., ψ is a so-called normalized Young function. For a number $q \in (0, 1)$ and each $\beta \in \mathbb{R}$, let $\alpha = \alpha(\beta)$ be a solution to

$$E\left\{\psi(\frac{(Y-\beta)_+}{\alpha})\right\} = 1 - q, \tag{5.1.1}$$

where $x_+ = \max(x, 0)$ and Y is a loss variable. Then, the H-G risk measure at level q with respect to the loss variable Y is defined as

$$\theta = \inf_{\beta \in \mathbb{R}} \{\beta + \alpha(\beta)\}. \tag{5.1.2}$$

When $\psi(x) = x$, we have $\alpha(\beta) = \frac{1}{1-q} E\{(Y - \beta)_+\}$ and $\theta = \frac{1}{1-q} E\{(Y - F_Y^-(q))_+\}$, where $F_Y(y) = P(Y \leq y)$ and $F_Y^-(y)$ denotes the inverse function of $F_Y(y)$. Hence, in this case, the H-G risk measure equals the expected shortfall.

Recently this risk measure has received much attention in the literature of actuarial science with applications in (re)insurance and portfolio management. For example, Goovaerts et al. (2004) showed that the H-G risk measure preserves the convex order property; a dual representation of this risk measure is given in Bellini and Gianin (2008a); Bellini and Gianin (2012); Goovaerts et al. (2012) investigated the relationship between this risk measure and others; Cheung and Lo (2013) obtained a lower bound for this risk measure when the loss variable Y is a sum of random variables; Optimal portfolio and optimal reinsurance under this risk measure are investigated by Bellini and Gianin (2008b) and Zhu, Zhang, and Zhang (2013) respectively. For statistical inference of this risk measure, Ahn and Shyamalkumar (2014) proposed a nonparametric estimation and derived its asymptotic distribution, which may be a non-normal distribution when the loss variable has no enough finite moments. When the limit is a normal distribution, Peng, Wang, and Zheng (2015) andWang and Peng (2016) developed an empirical likelihood method to construct an interval when the H-G risk measure is defined at a fixed level and an intermediate level, respectively. All these statistical inference procedures are built upon the assumption of independent observations.

Consider the total loss of an insurance or a financial portfolio $Y = \sum_{i=1}^{d} a_i X_i$ with X_i being the loss of the ith asset or ith business line in an insurance company. In this case, it is more realistic to assume that $(X_1, \cdots, X_d)^T$ comes from a strictly stationary sequence or a time series model. Throughout we use A^T to denote the transpose of the vector or matrix A. On the other hand, since Y is the loss of a portfolio, an interesting question is about the sensitivity analysis of the

portfolio under the H-G risk measure, which is defined as the partial derivatives of the portfolio with respect to a_i's; see Section 5.2 for explicit formula. The sensitivity analysis for the two commonly employed risk measures, Value-at-Risk and expected shortfall, has been studied in the literature (see Gourieroux, Laurent, and Scaillet (2000); Scaillet (2004); Hong (2009); Jiang and Fu (2015); Hong and Liu (2009); Liu and Hong (2009); Fu, Hong, and Hu (2009). Tsanakas and Millossovich (2016) provided a general study of sensitivity analysis using risk measure. References on sensitivity analysis for utility optimization can be found in the recent paper Cao and Wan (2017). As argued in Scaillet (2004), knowledge of the sensitivity of a risk measure is helpful in reducing the amount of computational time needed to process a large portfolio and in characterizing and evaluating capital allocations under this risk measure. Moreover, the expression of sensitivity analysis for Value-at-Risk and expected shortfall is quite related to the Euler capital allocation rule and risk capital allocation; see Kalkbrener (2005) and Fischer (2003). For a credit portfolio, some simulation methods have been proposed to compute the capital allocation based on either Value-at-Risk or expected shortfall; see Glasserman (2005).

This chapter studies the nonparametric inference for the sensitivity of the H-G risk measure in portfolios. More specifically, first we derive an expression for the sensitivity of a portfolio return under the H-G risk measure, which is a function of the H-G risk measure itself. In order to estimate the sensitivity nonparametrically, second we derive the asymptotic distribution of the nonparametric estimator for the H-G risk measure under the assumption of $\alpha-$mixing sequence, which generalizes the result in Ahn and Shyamalkumar (2014) for independent observations. Third, using the derived results in the above two steps, we propose a nonparametric estimator for the sensitivity and derive its asymptotic distribution. Since the obtained asymptotic variance is quite complicated, it remains challenging to construct a confidence interval/region for the proposed nonparametric estimators for the sensitivity and the H-G risk measure itself. Fourth, we propose to fit an AR-GARCH model to each asset or loss variable and then employ a bootstrap method based on residuals to construct a confidence interval/region for the nonparametric estimator of the sensitivity and/or the H-G risk measure. Note that a blockwise bootstrap method for dependent data is not feasible here since a risk

97

measure is usually set at a high level such as 99% and the effective sample size in each block may not be large enough to nonparametrically estimate the risk measure reasonably well.

We organize this chapter as follows. Methodologies and main results are presented in Section 5.2. A simulation study and a data analysis are given in Sections 5.3 and 5.4, respectively. All proofs are put in Section 5.5. Some conclusions are summarized in Section 5.6.

5.2 Methodologies and Main Results

Suppose the returns/losses of an insurance or a financial portfolio are $Y_t = \sum_{i=1}^d a_i X_{i,t}$ for $t = 1, \cdots, n$, where $X_{i,t}$ is the return/loss of the ith asset or the loss of the ith business line in an insurance company at time t and $\boldsymbol{a} = (a_1, a_2, \ldots, a_d)^T$ is the allocation of the portfolio. Then the H-G risk measure θ at level q with respect to Y_t is equivalent to solving the following equations under some regularity conditions given in Tang and Yang (2014):

$$
\begin{cases}
E\left\{\psi(\frac{Y_t - \beta}{\theta - \beta})I(Y_t > \beta)\right\} = 1 - q, \\
E\left\{\psi'(\frac{Y_t - \beta}{\theta - \beta})(Y_t - \theta)I(Y_t > \beta)\right\} = 0
\end{cases}
\tag{5.2.1}
$$

for some β and $\theta > \beta$, where $I(\cdot)$ is the indicator function. Obviously θ and β depend on $\boldsymbol{a} = (a_1, \cdots, a_d)^T$, and so we will write $\theta = \theta(\boldsymbol{a})$ and $\beta = \beta(\boldsymbol{a})$.

When a risk measure is applied to a portfolio, a quantity called sensitivity of the risk measure, which is defined as the partial derivative of the risk measure with respect to the allocation $\boldsymbol{a}$, becomes useful in managing the portfolio. Therefore the first question is how to compute the sensitivity of a portfolio under the H-G risk measure. Note that such a sensitivity essentially is equivalent to the so-called risk capital allocation.

Let $\theta^{(j)}(\boldsymbol{a})$ denote the sensitivity of the H-G risk measure for the portfolio return Y_t with respect to the allocation a_j. That is $\theta^{(j)}(\boldsymbol{a}) = \frac{\partial}{\partial a_j}\theta(\boldsymbol{a})$. Then we have the following expression for the sensitivity.

Theorem 5.2.1. *Assume the normalized Young function $\psi(x)$ satisfies $\psi'(0) = 0$. Then for $j =$*

$1, \cdots, d$

$$\theta^{(j)}(\boldsymbol{a}) = \frac{E\left\{X_{j,t}\psi'(\frac{Y_t - \beta(\boldsymbol{a})}{\theta(\boldsymbol{a}) - \beta(\boldsymbol{a})})I\big(Y_t > \beta(\boldsymbol{a})\big)\right\}}{E\left\{\psi'(\frac{Y_t - \beta(\boldsymbol{a})}{\theta(\boldsymbol{a}) - \beta(\boldsymbol{a})})I\big(Y_t > \beta(\boldsymbol{a})\big)\right\}}, \tag{5.2.2}$$

where $\theta(\boldsymbol{a})$ and $\beta(\boldsymbol{a})$ are the unique solution to (5.2.1).

Based on the above theorem, a simple nonparametric estimator for the sensitivity $\theta^{(j)}(\boldsymbol{a})$ is

$$\hat{\theta}^{(j)}(\boldsymbol{a}) = \frac{\sum_{t=1}^{n} X_{j,t}\psi'(\frac{Y_t - \hat{\beta}(\boldsymbol{a})}{\hat{\theta}(\boldsymbol{a}) - \hat{\beta}(\boldsymbol{a})})I\big(Y_t > \hat{\beta}(\boldsymbol{a})\big)}{\sum_{t=1}^{n} \psi'(\frac{Y_t - \hat{\beta}(\boldsymbol{a})}{\hat{\theta}(\boldsymbol{a}) - \hat{\beta}(\boldsymbol{a})})I\big(Y_t > \hat{\beta}(\boldsymbol{a})\big)},$$

where $\hat{\theta}(\boldsymbol{a})$ and $\hat{\beta}(\boldsymbol{a})$ solve the following estimating equations for $\theta > \beta$:

$$\begin{cases} \frac{1}{n}\sum_{t=1}^{n}\psi(\frac{Y_t - \beta}{\theta - \beta})I(Y_t > \beta) = 1 - q, \\ \frac{1}{n}\sum_{t=1}^{n}\psi'(\frac{Y_t - \beta}{\theta - \beta})(Y_t - \theta)I(Y_t > \beta) = 0. \end{cases} \tag{5.2.3}$$

Although the joint asymptotic distribution of $\hat{\theta}(\boldsymbol{a})$ and $\hat{\beta}(\boldsymbol{a})$ has been derived in Ahn and Shyamalkumar (2014) for independent observations, we need the asymptotic distribution based on dependent data before deriving the asymptotic distribution of the nonparametric estimator for the sensitivity. Here we focus on $\alpha-$mixing data defined as follows.

For $-\infty \leq a < b \leq \infty$, let $\mathcal{F}_a^b$ denote the σ-field generated by

$$\left\{\boldsymbol{X}_t = (X_{1,t}, \cdots, X_{d,t})^T : a \leq t \leq b\right\}$$

and define for $k \geq 1$

$$\alpha_{\boldsymbol{X}}(k) = \sup\left\{|P(A \cap B) - P(A)P(B)| : A \in \mathcal{F}_{-\infty}^{i}, B \in \mathcal{F}_{i+k}^{\infty}, -\infty < i < \infty\right\}.$$

Then the sequence $\{\boldsymbol{X}_t\}$ is called α-mixing if $\alpha_{\boldsymbol{X}}(k) \to 0$ as $k \to \infty$. Some simple examples are a sequence of independent and identically distributed random variables and an autoregressive and moving average model. We refer to the book by Doukhan (1995) for more examples.

Throughout we let $(\beta_0, \theta_0, \theta_0^{(j)})^T$ denote the true value of $(\beta, \theta, \theta^{(j)})^T$ in (5.2.1) and assume the following regularity conditions:

- A1) As $k \to \infty$, $\alpha_{\boldsymbol{X}}(k) = O(r^k)$ for some $r \in (0, 1)$;

- A2) ψ is a strictly convex function on $[0, \infty]$ with $\psi(0) = 0, \psi(1) = 1, \psi(\infty) = \infty$, and $\psi(t)$ has a continuous second derivative on $(0, \infty)$ with $\psi'(0) = 0$;

- A3) Suppose the density of Y_t is positive at β_0, and there exist $\delta_0 \in (0, 1/2)$ and $\delta_1 > 0$ such that

$$
\begin{cases}
\int_{G(\beta_0)}^1 (1-y)^{\delta_0} |\psi'(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})| \, dG^-(y) < \infty, \\
\sup_{|\beta_0-\beta|+|\theta_0-\theta|\leq\delta_1,\theta>\beta} \int_{G(\beta)}^1 |G^-(y)|\psi''(\frac{G^-(y)-\beta}{\theta-\beta}) \, dy < \infty,
\end{cases}
$$

$$
\begin{cases}
\int_{G(\beta_0)}^1 (1-y)^{\delta_0}\psi''(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})|G^-(y)| \, dG^-(y) < \infty, \\
\sup_{|\beta_0-\beta|+|\theta_0-\theta|\leq\delta_1,\theta>\beta} \int_{G(\beta)}^1 \{G^-(y)\}^2|\psi'(\frac{G^-(y)-\beta}{\theta-\beta})| \, dy < \infty,
\end{cases}
$$

$$
\begin{cases}
\int_{G(\beta_0)}^1 \int_{F_j(0)}^1 \min((1-x)^{\delta_0}, (1-y)^{\delta_0})\psi''(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0}) \, dF_j^-(x)dG^-(y) < \infty, \\
\int_{G(\beta_0)}^1 \int_0^{F_j(0)} \min(x^{\delta_0}, (1-y)^{\delta_0})\psi''(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0}) \, dF_j^-(x)dG^-(y) < \infty, \\
\sup_{|\beta_0-\beta|+|\theta_0-\theta|\leq\delta_1,\theta>\beta} \int_{G(\beta)}^1 \int_0^1 |F_j^-(x)G^-(y)|\psi''(\frac{G^-(y)-\beta}{\theta-\beta}) \, dC(x,y;j) < \infty,
\end{cases}
$$

where $F_j(x)$ and $G(x)$ denote the distribution functions of $X_{j,t}$ and Y_t, respectively.

A key technique in deriving the joint asymptotic distribution of the proposed estimators $\theta(\boldsymbol{a}), \beta(\boldsymbol{a})$ and $\theta^{(j)}(\boldsymbol{a})$ is the approximation of the empirical copula process for an $\alpha-$mixing sequence given in Berghaus, Bücher, and Volgushev (2017), which allows us to express the limit in terms of a common Gaussian process. The technical condition A3) can be ensured by imposing some conditions on the asymptotic behavior of $\psi(x)$ at infinity and moments of $\boldsymbol{X}_t$, and plays an important role in making this approximation of the empirical copula applicable. For example, if $\psi(x) = x^d$ for some $d > 1$, $P(|X_{j,t}| > x)$ and $P(|Y_t| > x)$ are regularly varying with index $-\alpha_1 < 0$ and $-\alpha_2 < 0$, respectively, then the above condition A3) holds when $\delta_0 > d/\alpha_2$, $\delta_0 > d/\alpha_2 - 1/\alpha_1 + 1/\alpha_2$ and

there is an open set Ω including $(\beta_0, \theta_0)^T$ such that

$$\begin{cases} E\left\{\sup_{(\beta,\theta)^T \in \Omega} \psi''(\frac{Y_t-\beta}{\theta-\beta})Y_t^2 I(Y_t > \beta)\right\} < \infty, \\ E\left\{\sup_{(\beta,\theta)^T \in \Omega} |\psi'(\frac{Y_t-\beta}{\theta-\beta})|Y_t^2 I(Y_t > \beta)\right\} < \infty, \\ E\left\{\sup_{(\beta,\theta)^T \in \Omega} \psi''(\frac{Y_t-\beta}{\theta-\beta})|X_{j,t}|Y_t^2 I(Y_t > \beta)\right\} < \infty. \end{cases}$$

Again let $F_j(x)$ and $G(x)$ denote the distribution functions of $X_{j,t}$ and Y_t, respectively. Put $U_{j,t} = F_j(X_{j,t})$ and $V_t = G(Y_t)$ for $t = 1, \cdots, n$. Define $C(u,v;j) = P(U_{j,t} \le u, V_t \le v)$, $C_n(u,v;j) = \frac{1}{n}\sum_{t=1}^n I(U_{j,t} \le u, V_t \le v)$ and $\alpha_n(u,v;j) = \sqrt{n}\left\{C_n(u,v;j) - C(u,v;j)\right\}$. Then, under condition A1), it follows from Theorem 2.2 of Berghaus, Bücher, and Volgushev (2017) that

$$\sup_{c/n \le u,v \le 1-c/n} \frac{|\alpha_n(u,v;j) - W_C(u,v;j)|}{\{\min(u,v,1-u,1-v)\}^\delta} = o_p(1) \tag{5.2.4}$$

for any $c > 0$ and $\delta \in (0, 1/2)$, where $W_C(u,v;j)$ is a tight, centered Gaussian process with covariance matrix

$$\begin{aligned} & E\left\{W_C(u_1,v_1;j)W_C(u_2,v_2;j)\right\} \\ = \ & \sum_{i=-\infty}^{+\infty} Cov\big(I(U_{j,1} \le u_1, V_1 \le v_1), I(U_{j,1+i} \le u_2, V_{1+i} \le v_2)\big). \end{aligned}$$

Using (5.2.4), the following theorem derives the asymptotic distribution of the nonparametric estimators for the H-G risk measure and the sensitivity in terms of the same Gaussian process $W_C(x,y;j)$.

Theorem 5.2.2. *(i) Under conditions A1)–A3), as $n \to \infty$, we have*

$$\sqrt{n}\begin{pmatrix} \hat{\beta}(\boldsymbol{a}) - \beta_0 \\ \hat{\theta}(\boldsymbol{a}) - \theta_0 \end{pmatrix} = \Sigma^{-1}\boldsymbol{Z} + o_p(1),$$

where

$$\boldsymbol{Z} = \begin{pmatrix} \int_{G(\beta_0)}^1 W_C(1,y;j) \frac{1}{\theta_0 - \beta_0} \psi'(\frac{G^-(y)-\beta_0}{\theta_0 - \beta_0}) \, dG^-(y) \\ \int_{G(\beta_0)}^1 W_C(1,y;j) \left\{ \psi''(\frac{G^-(y)-\beta_0}{\theta_0 - \beta_0}) \frac{G^-(y)-\theta_0}{\theta_0 - \beta_0} + \psi'(\frac{G^-(y)-\beta_0}{\theta_0 - \beta_0}) \right\} \, dG^-(y) \end{pmatrix}$$

and $\Sigma = (\sigma_{ij})_{2\times 2}$ with

$$\sigma_{11} = E\left\{ \psi'(\frac{Y_1 - \beta_0}{\theta_0 - \beta_0}) \frac{Y_1 - \theta_0}{(\theta_0 - \beta_0)^2} I(Y_1 > \beta_0) \right\} = 0,$$

$$\sigma_{21} = E\left\{ \psi''(\frac{Y_1 - \beta_0}{\theta_0 - \beta_0}) \frac{(Y_1 - \theta_0)^2}{(\theta_0 - \beta_0)^2} I(Y_1 > \beta_0) \right\},$$

$$\sigma_{12} = -\frac{1}{\theta_0 - \beta_0} E\left\{ \psi'(\frac{Y_1 - \beta_0}{\theta_0 - \beta_0}) I(Y_1 > \beta_0) \right\},$$

$$\sigma_{22} = E\left\{ \left(\psi''(\frac{Y_1 - \beta_0}{\theta_0 - \beta_0}) \frac{(Y_1 - \theta_0)(\beta_0 - Y_1)}{(\theta_0 - \beta_0)^2} - \psi'(\frac{Y_1 - \beta_0}{\theta_0 - \beta_0}) \right) I(Y_1 > \beta_0) \right\}.$$

(ii) Under conditions A1)–A3), as $n \to \infty$, we have

$$\sqrt{n}\left\{ \hat{\theta}^{(j)}(\boldsymbol{a}) - \theta_0^{(j)}(\boldsymbol{a}) \right\} = \frac{b_2 Z_1 - b_1 Z_2}{b_2^2} + \frac{1}{b_2^2}\{b_2 \boldsymbol{A}_1 - b_1 \boldsymbol{A}_2\}^T \Sigma^{-1} \boldsymbol{Z} + o_p(1),$$

where

$$\begin{aligned} Z_1 &= \int_{G(\beta_0)}^1 \int_{F_j(0)}^1 \{W_C(x,y;j) - W_C(x,1;j) - W_C(1,y;j)\} \, dF_j^-(x) d\psi'(\frac{G^-(y)-\beta_0}{\theta_0 - \beta_0}) \\ &\quad + \int_{G(\beta_0)}^1 \int_0^{F_j(0)} \{W_C(x,y;j) - W_C(x,1;j)\} \, dF_j^-(x) d\psi'(\frac{G^-(y)-\beta_0}{\theta_0 - \beta_0}), \end{aligned}$$

$$\boldsymbol{A}_1 = \begin{pmatrix} E\left\{ X_{j,1} \psi''(\frac{Y_1 - \beta_0}{\theta_0 - \beta_0}) \frac{Y_1 - \theta_0}{(\theta_0 - \beta_0)^2} I(Y_1 > \beta_0) \right\} \\ E\left\{ X_{j,1} \psi''(\frac{Y_1 - \beta_0}{\theta_0 - \beta_0}) \frac{\beta_0 - Y_1}{(\theta_0 - \beta_0)^2} I(Y_1 > \beta_0) \right\} \end{pmatrix},$$

$$Z_2 = -\int_{G(\beta_0)}^1 W_C(1,y;j) \, d\psi'(\frac{G^-(y)-\beta_0}{\theta_0 - \beta_0}),$$

$$\boldsymbol{A}_2 = \begin{pmatrix} E\left\{ \psi''(\frac{Y_1 - \beta_0}{\theta_0 - \beta_0}) \frac{Y_1 - \theta_0}{(\theta_0 - \beta_0)^2} I(Y_1 > \beta_0) \right\} \\ E\left\{ \psi''(\frac{Y_1 - \beta_0}{\theta_0 - \beta_0}) \frac{\beta_0 - Y_1}{(\theta_0 - \beta_0)^2} I(Y_1 > \beta_0) \right\} \end{pmatrix},$$

$$b_1 = E\left\{X_{j,1}\psi'(\frac{Y_1 - \beta_0}{\theta_0 - \beta_0})I(Y_1 > \beta_0)\right\} \quad and \quad b_2 = E\left\{\psi'(\frac{Y_1 - \beta_0}{\theta_0 - \beta_0})I(Y_1 > \beta_0)\right\}.$$

It is known that uncertainty quantification is important in risk management. Although the limit in the above theorem is a normal distribution, estimating the asymptotic variance/covariance is highly nontrivial due to dependent data. One may employ the blockwise bootstrap method to construct a confidence interval for $\theta^{(j)}(\boldsymbol{a})$. However, the effective sample size in each block may not be large enough to nonparametrically estimate $\beta(\boldsymbol{a})$ and $\theta(\boldsymbol{a})$ accurately since the level q is always close to one in practice. Alternatively an efficient way to quantify the uncertainty is to model the dependence of each asset returns by a time series model and then to employ a bootstrap method based on residuals to construct intervals/regions. More specifically, we propose to assume each asset return follows the following AR-GARCH model in Chen and Fan (2006):

$$\begin{cases} X_{l,t} = \mu_l + \sum_{i=1}^{P_l} a_{l,i}X_{l,t-i} + e_{l,t}, \\ e_{l,t} = h_{l,t}^{1/2}\eta_{l,t}, \quad h_{l,t} = w_l + \sum_{i=1}^{q_l}\alpha_{l,i}e_{l,t-i}^2 + \sum_{k=1}^{p_l}\beta_{l,k}h_{l,t-k} \end{cases} \tag{5.2.5}$$

for $l = 1, \cdots, d$, where $\{\boldsymbol{\eta}_t := (\eta_{1,t}, \cdots, \eta_{d,t})^T\}$ is a sequence of independent and identically distributed random vectors with means zero and variances one.

A standard inference procedure for the model (5.2.5) is the so-called quasi maximum likelihood estimator, but its asymptotic normality requires finite fourth moments of both $e_{l,t}$ and $\eta_{l,t}$; see Francq and Zakoian (2004). In practice it is quite often that $\sum_{i=1}^{q_l}\alpha_{l,i} + \sum_{k=1}^{p_l}\beta_{l,k}$ is close to one, which implies that assuming $Ee_{l,t}^4 < \infty$ may be problematic. Instead we propose to employ the self-weighted quasi maximum likelihood estimator in Ling (2007) with the weights

$$\delta_{l,t} = \left\{\max\left(1, \frac{1}{C_l}\sum_{k=1}^{t-1}\frac{|X_{t-k}|I(|X_{t-k}| > C_l)}{k^9}\right)\right\}^{-4}$$

to estimate

$$\boldsymbol{\gamma}_l = (\mu_l, a_{l,1}, \cdots, a_{l,P_l}, \alpha_{l,1}, \cdots, \alpha_{l,q_l}, \beta_{l,1}, \cdots, \beta_{l,p_l})^T,$$

say $\hat{\boldsymbol{\gamma}}_l$, where the asymptotic normality only requires $E|e_{l,t}| < 1$ and $E\eta_{l,t}^2 < \infty$. Here C_l is chosen

as the 90% sample quantile of $\{X_{l,t}\}_{t=1}^n$ as suggested by Zhu and Ling (2011).

After obtaining $\hat{\boldsymbol{\gamma}}_1, \cdots, \hat{\boldsymbol{\gamma}}_d$, we get our estimators for $\boldsymbol{\eta}_t$, $t = 1, \cdots, n$, say $\hat{\boldsymbol{\eta}}_t$. Therefore, we resample from $\{\hat{\boldsymbol{\eta}}_t\}_{t=1}^n$ with sample size n and then refit model (5.2.5) to obtain bootstrap sample $X_{l,t}^*$ for $l = 1, \cdots, d$ and $t = 1, \cdots, n$. Based on this bootstrap sample, one can calculate the bootstrapped estimators of $\hat{\beta}(\boldsymbol{a}), \hat{\theta}(\boldsymbol{a})$ and $\hat{\theta}^{(j)}(\boldsymbol{a})$. By repeating this procedure many times, a bootstrap confidence interval for $\theta^{(j)}(\boldsymbol{a})$ based on $\hat{\theta}^{(j)}(\boldsymbol{a})$ and its bootstrapped estimators can be constructed. Similarly, a bootstrap region can be obtained for the nonparametric estimators of the H-G risk measure and the sensitivity.

5.3 Simulation Study

In this section, we carry out a simulation study based on AR-GARCH models. Specifically we simulate data from the following AR(1)-GARCH(1,1) models for $\{X_{1,t}\}$ and $\{X_{2,t}\}$:

$$
\begin{cases}
X_{1,t} = 0.0830 - 0.0390 X_{1,t-1} + e_{1,t}, \quad e_{1,t} = h_{1,t}^{1/2} \eta_{1,t}, \\
\quad h_{1,t} = 0.0548 + 0.0852 e_{1,t-1}^2 + 0.9043 h_{1,t-1}; \\
X_{2,t} = 0.0591 - 0.0676 X_{2,t-1} + e_{2,t}, \quad e_{2,t} = h_{2,t}^{1/2} \eta_{2,t}, \\
\quad h_{2,t} = 0.0259 + 0.1128 e_{2,t-1}^2 + 0.8637 h_{2,t-1},
\end{cases}
\tag{5.3.1}
$$

where $\{\boldsymbol{\eta}_t = (\eta_{1,t}, \eta_{2,t})^T\}$ is a sequence of independent and identically distributed random vectors with marginal t-distributions with degrees of freedom 6.2345 and 5.6237, respectively, and a t-copula with parameters $\rho = 0.7206$ and $\nu = 6.1702$. We consider the return of a portfolio $Y_t = a_1 X_{1,t} + a_2 X_{2,t}$ with $a_1 = a_2 = 0.5$. Note that the chosen parameters in (5.3.1) and the setting for $\boldsymbol{\eta}_t$ come from the fitted values for the returns of Goldman Sachs and S&P 500 index in the real data analysis given in Section 5.4 below.

First we evaluate the true values $\theta_0, \theta_0^{(j)}, j = 1, 2$ of the H-G risk measure and its sensitivity by drawing $100,000$ random samples with sample size $n = 100,000$ from (5.3.1) and considering $q = 0.95, 0.99$, $\psi(x) = x^{1.1}$ and $\psi(x) = x^{1.3}$. For each sample, we compute $\hat{\theta}$ and $\hat{\theta}^{(j)}$. Therefore the true values θ_0 and $\theta_0^{(j)}$ are approximated by the corresponding averages of these computed

$100,000$ estimators and are reported in Tables 5.1 and 5.2.

Next, we draw $1,000$ random samples from model (5.3.1) with sample size $n = 2,000$ or $3,000$ or $5,000$ and compute $\hat{\theta}, \hat{\theta}^{(j)}, j = 1, 2$ for each sample. The corresponding averages and standard errors are reported in Tables 5.1 and 5.2. To examine the efficiency of the proposed bootstrap method in Section 5.2, we also compute the bootstrapped standard deviations for $\hat{\theta}$ and $\hat{\theta}^{(j)}$ by drawing $1,000$ resamples from the estimated residuals in model (5.3.1) with the same sample size n. Since the sum of the two slope parameters in GARCH models of equation (5.3.1) is close to 1 on both margins, we apply the self-weighted quasi maximum likelihood estimator to the resampling procedure of the bootstrap method as stated in Section 5.2.

Our observations from Tables 5.1 and 5.2 are: i) the proposed nonparametric estimators for the H-G risk measure and the sensitivity are close to the true values and the proposed bootstrap method gives a close value to the simulated standard deviation; ii) the nonparametric estimators perform better as n becomes larger; iii) estimators at level $q = 0.99$ have a larger standard deviation than those at level $q = 0.95$; iv) estimators for $\psi(x) = x^{1.3}$ are larger than those for $\psi(x) = x^{1.1}$. In conclusion, the proposed nonparametric estimators have a good finite sample behavior.

5.4 Real Data Analysis

In this section, we apply our estimators to a real data set consisting of the daily log stock returns of Goldman Sachs ($X_{1,t}$) and S&P 500 index ($X_{2,t}$) from January 3, 2005 to December 31, 2016, which are plotted in Figure 5.1. The mean and standard deviation are 0.0321% and 2.3753% for Goldman Sachs, and 0.0203% and 1.2327% for S&P 500. The autocorrelation function plots show that an AR(1)-GARCH(1,1) model fits well to these two returns. The analyzed portfolio contains equal unit of these two securities, that is, the log return of the portfolio is $Y_t = 0.5X_{1,t} + 0.5X_{2,t}$. We scale the data by taking multiplication with 100.

We compute the nonparametric estimates of $\theta, \theta^{(j)}, j = 1, 2$ by solving the two equations (5.2.3) and equation (5.2.2) for $q = 0.95, 0.99$. The obtained estimates are reported in Table 5.3. In order to estimate the standard deviations of the proposed estimators, we apply the proposed bootstrap

method in Section 5.2. More specifically, first we use the self-weighted quasi maximum likelihood estimator in Ling (2007) to fit an AR(1)-GARCH(1,1) model to each asset, where the obtained estimates are summarized in model (5.3.1) in Section 5.3. Second we draw $1,000$ bootstrap samples from the residuals in the AR(1)-GARCH(1,1) models and then compute the corresponding estimates for each bootstrap sample. Therefore the bootstrapped standard deviations are obtained based on these $1,000$ bootstrapped estimates and reported in Table 5.3.

After plotting the autocorrelation functions based on the residuals in Figure 5.2, we confirm that the employed AR(1)-GARCH(1,1) models are reasonable. That is, the proposed bootstrap method for uncertainty quantification is applicable to the investigated data set. From Table 5.3, we observe that the proposed nonparametric estimate and its standard deviation for the sensitivity with respect to Goldman Sachs are much larger than those for the sensitivity with respect to S&P 500 index, and estimates at level 0.99 are larger than those at level 0.95. One possible explanation is that the return of Goldman Sachs has a heavier tail and higher volatility than those of S&P 500 Index which is observed in Figure 5.1. This makes the H-G risk measure of the portfolio more sensitive to the change from Goldman Sachs than S&P 500 Index.

In order to examine the finite sample performance of the proposed estimators and the bootstrap method for quantifying uncertainty under a similar setting with this investigated real data set, we also fit a bivariate parametric distribution with t-copula and marginal t-distributions to the innovations in the AR(1)-GARCH(1,1) models. Results are given in Section 5.3. Note that it has no need to fit a parametric family to the innovations in the employed AR-GARCH models for applying the proposed nonparametric estimators for the H-G risk measure and its sensitivity to a real data set.

5.5 Proofs

Proof of Theorem 5.2.1. We only prove the case of $j = 1$ and $a_1 > 0$ since other cases can be shown in the same way. Put $Z_t = \sum_{i=2}^{d} a_i X_{i,t}$, $x_1(z) = \frac{\beta(\boldsymbol{a}) - z}{a_1}$, and let $G(x, z)$ denote the joint

distribution function of $X_{1,t}$ and Z_t. It follows from (5.2.1) that

$$
\begin{cases}
\int_{-\infty}^{+\infty} \int_{x_1(z)}^{+\infty} \psi\left(\frac{a_1 x + z - \beta(\boldsymbol{a})}{\theta(\boldsymbol{a}) - \beta(\boldsymbol{a})}\right) dG(x, z) = 1 - q, \\
\int_{-\infty}^{+\infty} \int_{x_1(z)}^{+\infty} \psi'\left(\frac{a_1 x + z - \beta(\boldsymbol{a})}{\theta(\boldsymbol{a}) - \beta(\boldsymbol{a})}\right)\left(a_1 x + z - \theta(\boldsymbol{a})\right) dG(x, z) = 0.
\end{cases}
\tag{5.5.1}
$$

By taking derivative with respect to a_1 in the first equation of (5.5.1) and noting that $\psi(0) = 0$, we have

$$
\begin{aligned}
0 &= \int_{-\infty}^{+\infty} \int_{x_1(z)}^{+\infty} \frac{\partial}{\partial a_1}\left(\psi\left(\frac{a_1 x + z - \beta(\boldsymbol{a})}{\theta(\boldsymbol{a}) - \beta(\boldsymbol{a})}\right)\right) dG(x, z) \\
&= \int_{-\infty}^{+\infty} \int_{x_1(z)}^{+\infty} \psi'\left(\frac{a_1 x + z - \beta(\boldsymbol{a})}{\theta(\boldsymbol{a}) - \beta(\boldsymbol{a})}\right) \frac{\left(x - \beta^{(1)}(\boldsymbol{a})\right)\left(\theta(\boldsymbol{a}) - \beta(\boldsymbol{a})\right) - \left(a_1 x + z - \beta(\boldsymbol{a})\right)\left(\theta^{(1)}(\boldsymbol{a}) - \beta^{(1)}(\boldsymbol{a})\right)}{\left(\theta(\boldsymbol{a}) - \beta(\boldsymbol{a})\right)^2} dG(x, z) \\
&= \int_{-\infty}^{+\infty} \int_{x_1(z)}^{+\infty} \psi'\left(\frac{a_1 x + z - \beta(\boldsymbol{a})}{\theta(\boldsymbol{a}) - \beta(\boldsymbol{a})}\right) \frac{x\left(\theta(\boldsymbol{a}) - \beta(\boldsymbol{a})\right) - \theta^{(1)}(\boldsymbol{a})\left(a_1 x + z - \beta(\boldsymbol{a})\right) - \beta^{(1)}(\boldsymbol{a})\left(\theta(\boldsymbol{a}) - a_1 x - z\right)}{\left(\theta(\boldsymbol{a}) - \beta(\boldsymbol{a})\right)^2} dG(x, z) \\
&= E\left\{\psi'\left(\frac{Y_t - \beta(\boldsymbol{a})}{\theta(\boldsymbol{a}) - \beta(\boldsymbol{a})}\right) \frac{X_{1,t}\left(\theta(\boldsymbol{a}) - \beta(\boldsymbol{a})\right) - \theta^{(1)}(\boldsymbol{a})\left(Y_t - \beta(\boldsymbol{a})\right) - \beta^{(1)}(\boldsymbol{a})\left(\theta(\boldsymbol{a}) - Y_t\right)}{\left(\theta(\boldsymbol{a}) - \beta(\boldsymbol{a})\right)^2} I\left(Y_t > \beta(\boldsymbol{a})\right)\right\} \\
&= E\left\{\psi'\left(\frac{Y_t - \beta(\boldsymbol{a})}{\theta(\boldsymbol{a}) - \beta(\boldsymbol{a})}\right) \frac{X_{1,t}}{\theta(\boldsymbol{a}) - \beta(\boldsymbol{a})} I\left(Y_t > \beta(\boldsymbol{a})\right)\right\} \\
&\quad - \theta^{(1)}(\boldsymbol{a}) E\left\{\psi'\left(\frac{Y_t - \beta(\boldsymbol{a})}{\theta(\boldsymbol{a}) - \beta(\boldsymbol{a})}\right) \frac{Y_t - \beta(\boldsymbol{a})}{\left(\theta(\boldsymbol{a}) - \beta(\boldsymbol{a})\right)^2} I\left(Y_t > \beta(\boldsymbol{a})\right)\right\} \\
&\quad - \beta^{(1)}(\boldsymbol{a}) E\left\{\psi'\left(\frac{Y_t - \beta(\boldsymbol{a})}{\theta(\boldsymbol{a}) - \beta(\boldsymbol{a})}\right) \frac{\theta(\boldsymbol{a}) - Y_t}{\left(\theta(\boldsymbol{a}) - \beta(\boldsymbol{a})\right)^2} I\left(Y_t > \beta(\boldsymbol{a})\right)\right\} \\
&= E\left\{\psi'\left(\frac{Y_t - \beta(\boldsymbol{a})}{\theta(\boldsymbol{a}) - \beta(\boldsymbol{a})}\right) \frac{X_{1,t}}{\theta(\boldsymbol{a}) - \beta(\boldsymbol{a})} I\left(Y_t > \beta(\boldsymbol{a})\right)\right\} \\
&\quad - \theta^{(1)}(\boldsymbol{a}) E\left\{\psi'\left(\frac{Y_t - \beta(\boldsymbol{a})}{\theta(\boldsymbol{a}) - \beta(\boldsymbol{a})}\right) \frac{Y_t - \beta(\boldsymbol{a})}{\left(\theta(\boldsymbol{a}) - \beta(\boldsymbol{a})\right)^2} I\left(Y_t > \beta(\boldsymbol{a})\right)\right\}
\end{aligned}
$$

by using the second equation of (5.5.1). That is,

$$
\theta^{(1)}(\boldsymbol{a}) = \frac{E\left\{\psi'\left(\frac{Y_t - \beta(\boldsymbol{a})}{\theta(\boldsymbol{a}) - \beta(\boldsymbol{a})}\right) \frac{X_{1,t}}{\theta(\boldsymbol{a}) - \beta(\boldsymbol{a})} I\left(Y_t > \beta(\boldsymbol{a})\right)\right\}}{E\left\{\psi'\left(\frac{Y_t - \beta(\boldsymbol{a})}{\theta(\boldsymbol{a}) - \beta(\boldsymbol{a})}\right) \frac{Y_t - \beta(\boldsymbol{a})}{\left(\theta(\boldsymbol{a}) - \beta(\boldsymbol{a})\right)^2} I\left(Y_t > \beta(\boldsymbol{a})\right)\right\}}.
\tag{5.5.2}
$$

It follows from the second equation of (5.5.1) that

$$
\begin{aligned}
& E\left\{\psi'\left(\tfrac{Y_t-\beta(\boldsymbol{a})}{\theta(\boldsymbol{a})-\beta(\boldsymbol{a})}\right)\frac{Y_t-\beta(\boldsymbol{a})}{\left(\theta(\boldsymbol{a})-\beta(\boldsymbol{a})\right)^2}I\big(Y_t>\beta(\boldsymbol{a})\big)\right\} \\
=\ & E\left\{\psi'\left(\tfrac{Y_t-\beta(\boldsymbol{a})}{\theta(\boldsymbol{a})-\beta(\boldsymbol{a})}\right)\frac{\left(Y_t-\theta(\boldsymbol{a})\right)+\left(\theta(\boldsymbol{a})-\beta(\boldsymbol{a})\right)}{\left(\theta(\boldsymbol{a})-\beta(\boldsymbol{a})\right)^2}I\big(Y_t>\beta(\boldsymbol{a})\big)\right\} \\
=\ & E\left\{\psi'\left(\tfrac{Y_t-\beta(\boldsymbol{a})}{\theta(\boldsymbol{a})-\beta(\boldsymbol{a})}\right)\frac{1}{\theta(\boldsymbol{a})-\beta(\boldsymbol{a})}I\big(Y_t>\beta(\boldsymbol{a})\big)\right\}.
\end{aligned}
\tag{5.5.3}
$$

Hence (5.2.1) with $j=1$ and $a_1>0$ follows from (5.5.2) and (5.5.3). Other cases can be shown in the same way.

∎

Before proving Theorem 5.2.2, we first list some facts and then show some lemmas.

Under condition A1), it follows from Proposition 4.4 of Berghaus, Bücher, and Volgushev (2017) that for any $\delta \in (0,1/2)$ and positive $\delta_n \to 0$

$$
\sup_{|u_1-u_2|+|v_1-v_2|\le\delta_n}\frac{|\alpha_n(u_1,v_1;j)-\alpha_n(u_2,v_2;j)|}{\max\left(|u_1-u_2|^\delta+|v_1-v_2|^\delta,n^{-\delta}\right)}=o_p(1)
\tag{5.5.4}
$$

and

$$
\sup_{0<u<1}\frac{|\alpha_n(u,1;j)|+|\alpha_n(1,u;j)|}{u^\delta(1-u)^\delta}=O_p(1).
\tag{5.5.5}
$$

Lemma 5.5.1. *Under conditions of Theorem 5.2.2, as $n \to \infty$, we have*

$$
\begin{aligned}
& \sqrt{n}\left\{\tfrac{1}{n}\sum_{t=1}^{n}\psi\big(\tfrac{Y_t-\beta_0}{\theta_0-\beta_0}\big)I(Y_t>\beta_0)-1+q\right\} \\
=\ & -\int_{G(\beta_0)}^{1}W_C(1,y;j)\tfrac{1}{\theta_0-\beta_0}\psi'\big(\tfrac{G^-(y)-\beta_0}{\theta_0-\beta_0}\big)\,dG^-(y)+o_p(1),
\end{aligned}
\tag{5.5.6}
$$

$$
\frac{1}{n}\sum_{t=1}^{n}\psi'\big(\frac{Y_t-\beta_0}{\theta_0-\beta_0}\big)\frac{Y_t-\theta_0}{(\theta_0-\beta_0)^2}I(Y_t>\beta_0)=o_p(1),
\tag{5.5.7}
$$

$$\frac{1}{n}\sum_{t=1}^{n}\psi'(\frac{Y_t-\beta_0}{\theta_0-\beta_0})\frac{\beta_0-Y_t}{(\theta_0-\beta_0)^2}I(Y_t>\beta_0)=-\frac{1}{\theta_0-\beta_0}E\left\{\psi'(\frac{Y_1-\beta_0}{\theta_0-\beta_0})I(Y_1>\beta_0)\right\}+o_p(1).$$

(5.5.8)

Proof. Write

$$
\begin{aligned}
&\frac{1}{n}\sum_{t=1}^{n}\psi(\frac{Y_t-\beta_0}{\theta_0-\beta_0})I(Y_t>\beta_0)-(1-q)\\
&=\int_{G(\beta_0)}^{1}\psi(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})\,d\left\{C_n(1,y;j)-C(1,y;j)\right\}\\
&=-\int_{G(\beta_0)}^{1}\{C_n(1,y;j)-C(1,y;j)\}\frac{1}{\theta_0-\beta_0}\psi'(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})\,dG^-(y)\\
&=-\int_{G(\beta_0)}^{1-1/n}\{C_n(1,y;j)-C(1,y;j)\}\frac{1}{\theta_0-\beta_0}\psi'(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})\,dG^-(y)\\
&\quad-\int_{1-1/n}^{1}\{C_n(1,y;j)-C(1,y;j)\}\frac{1}{\theta_0-\beta_0}\psi'(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})\,dG^-(y)\\
&=-\int_{G(\beta_0)}^{1-1/n}\{C_n(1-\tfrac{1}{n},y;j)-C(1-\tfrac{1}{n},y;j)\}\frac{1}{\theta_0-\beta_0}\psi'(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})\,dG^-(y)\\
&\quad-\int_{G(\beta_0)}^{1-1/n}\{C_n(1,y;j)-C(1,y;j)-C_n(1-\tfrac{1}{n},y;j)+C(1-\tfrac{1}{n},y;j)\}\\
&\quad\times\frac{1}{\theta_0-\beta_0}\psi'(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})\,dG^-(y)\\
&\quad-\int_{1-1/n}^{1}\{C_n(1,y;j)-C(1,y;j)\}\frac{1}{\theta_0-\beta_0}\psi'(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})\,dG^-(y)\\
&=I_1+I_2+I_3.
\end{aligned}
$$

Since $\int_{G(\beta_0)}^{1}(1-y)^{\delta_0}|\psi'(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})|\,dG^-(y)<\infty$, it follows from (5.2.4) and condition A3) that

$$
\begin{aligned}
\sqrt{n}I_1&=-\int_{G(\beta_0)}^{1-1/n}W_C(1-\tfrac{1}{n},y;j)\frac{1}{\theta_0-\beta_0}\psi'(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})\,dG^-(y)+o_p(1)\\
&=-\int_{G(\beta_0)}^{1}W_C(1,y;j)\frac{1}{\theta_0-\beta_0}\psi'(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})\,dG^-(y)+o_p(1).
\end{aligned}
$$

(5.5.9)

Using $\int_{G(\beta_0)}^{1}(1-y)^{\delta_0}|\psi'(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})|\,dG^-(y)<\infty$, it follows from (5.5.4), (5.5.5) and A3) that

$$
\begin{aligned}
\sqrt{n}|I_2|&=o_p\left(\int_{G(\beta_0)}^{1-1/n}n^{-\delta_0}|\psi'(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})|\,dG^-(y)\right)\\
&=o_p\left(\int_{G(\beta_0)}^{1-1/n}(1-y)^{\delta_0}|\psi'(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})|\,dG^-(y)\right)=o_p(1)
\end{aligned}
$$

(5.5.10)

and

$$\sqrt{n}|I_3|=O_p\left(\int_{1-1/n}^{1}(1-y)^{\delta_0}|\psi'(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})|\,dG^-(y)\right)=o_p(1).$$

(5.5.11)

109

Hence (5.5.6) follows from (5.5.9)–(5.5.11). Similarly we can show that

$$\frac{1}{n}\sum_{t=1}^{n}\psi'(\frac{Y_t-\beta_0}{\theta_0-\beta_0})\frac{Y_t-\theta_0}{(\theta_0-\beta_0)^2}I(Y_t>\beta_0) - E\left\{\psi'(\frac{Y_1-\beta_0}{\theta_0-\beta_0})\frac{Y_1-\theta_0}{(\theta_0-\beta_0)^2}I(Y_1>\beta_0)\right\} = o_p(1)$$

and

$$\frac{1}{n}\sum_{t=1}^{n}\psi'(\frac{Y_t-\beta_0}{\theta_0-\beta_0})\frac{\beta_0-Y_t}{(\theta_0-\beta_0)^2}I(Y_t>\beta_0) - E\left\{\psi'(\frac{Y_1-\beta_0}{\theta_0-\beta_0})\frac{\beta_0-Y_1}{(\theta_0-\beta_0)^2}I(Y_1>\beta_0)\right\} = o_p(1).$$

Then the lemma follows by noting that

$$E\left\{\psi'(\frac{Y_1-\beta_0}{\theta_0-\beta_0})\frac{Y_1-\theta_0}{(\theta_0-\beta_0)^2}I(Y_1>\beta_0)\right\} = 0$$

and

$$E\left\{\psi'(\tfrac{Y_1-\beta_0}{\theta_0-\beta_0})\tfrac{\beta_0-Y_1}{(\theta_0-\beta_0)^2}I(Y_1>\beta_0)\right\} = -\tfrac{1}{\theta_0-\beta_0}E\left\{\psi'(\tfrac{Y_1-\beta_0}{\theta_0-\beta_0})I(Y_1>\beta_0)\right\}.$$

∎

Lemma 5.5.2. *Under conditions of Theorem 5.2.2, as $n\to\infty$, we have*

$$\begin{aligned}&\tfrac{1}{\sqrt{n}}\sum_{t=1}^{n}\psi'(\tfrac{Y_t-\beta_0}{\theta_0-\beta_0})(Y_t-\theta_0)I(Y_t>\beta_0)\\ =\ & -\int_{G(\beta_0)}^{1}W_C(1,y;j)\left\{\psi''(\tfrac{G^-(y)-\beta_0}{\theta_0-\beta_0})\tfrac{G^-(y)-\theta_0}{\theta_0-\beta_0}+\psi'(\tfrac{G^-(y)-\beta_0}{\theta_0-\beta_0})\right\}\,dG^-(y)+o_p(1),\end{aligned} \tag{5.5.12}$$

$$\tfrac{1}{n}\sum_{t=1}^{n}\psi''(\tfrac{Y_t-\beta_0}{\theta_0-\beta_0})\tfrac{(Y_t-\theta_0)^2}{(\theta_0-\beta_0)^2}I(Y_t>\beta_0)=E\left\{\psi''(\tfrac{Y_1-\beta_0}{\theta_0-\beta_0})\tfrac{(Y_1-\theta_0)^2}{(\theta_0-\beta_0)^2}I(Y_1>\beta_0)\right\}+o_p(1), \tag{5.5.13}$$

and

$$\begin{aligned}&\tfrac{1}{n}\sum_{t=1}^{n}\left\{\psi''(\tfrac{Y_t-\beta_0}{\theta_0-\beta_0})\tfrac{(\beta_0-Y_t)(Y_t-\theta_0)}{(\theta_0-\beta_0)^2}-\psi'(\tfrac{Y_t-\beta_0}{\theta_0-\beta_0})\right\}I(Y_t>\beta_0)\\ =\ & E\left\{(\psi''(\tfrac{Y_1-\beta_0}{\theta_0-\beta_0})\tfrac{(Y_1-\theta_0)(\beta_0-Y_1)}{(\theta_0-\beta_0)^2}-\psi'(\tfrac{Y_1-\beta_0}{\theta_0-\beta_0}))I(Y_1>\beta_0)\right\}+o_p(1).\end{aligned} \tag{5.5.14}$$

Proof. It can be shown in the same way as the proof of Lemma 5.5.1. ∎

Lemma 5.5.3. *Under conditions of Theorem 5.2.2, as $n \to \infty$, we have*

$$\sqrt{n} \left\{ \frac{1}{n} \sum_{t=1}^{n} X_{j,t} \psi'(\frac{Y_t - \beta_0}{\theta_0 - \beta_0}) I(Y_t > \beta_0) - E\left(X_{j,1} \psi'(\frac{Y_1 - \beta_0}{\theta_0 - \beta_0}) I(Y_1 > \beta_0) \right) \right\}$$

$$= \int_{G(\beta_0)}^{1} \int_{F_j(0)}^{1} \left\{ W_C(x, y; j) - W_C(1, y; j) - W_C(x, 1; j) \right\} dF_j^-(x) d\psi'(\frac{G^-(y) - \beta_0}{\theta_0 - \beta_0}) \quad (5.5.15)$$

$$+ \int_{G(\beta_0)}^{1} \int_{0}^{F_j(0)} \left\{ W_C(x, y; j) - W_C(x, 1; j) \right\} dF_j^-(x) d\psi'(\frac{G^-(y) - \beta_0}{\theta_0 - \beta_0}) + o_p(1),$$

$$\frac{1}{n} \sum_{t=1}^{n} X_{j,t} \psi''(\frac{Y_t - \beta_0}{\theta_0 - \beta_0}) \frac{Y_t - \theta_0}{(\theta_0 - \beta_0)^2} I(Y_t > \beta_0)$$

$$= E\left\{ X_{j,1} \psi''(\frac{Y_1 - \beta_0}{\theta_0 - \beta_0}) \frac{Y_1 - \theta_0}{(\theta_0 - \beta_0)^2} I(Y_1 > \beta_0) \right\} + o_p(1) \quad (5.5.16)$$

and

$$\frac{1}{n} \sum_{t=1}^{n} X_{j,t} \psi''(\frac{Y_t - \beta_0}{\theta_0 - \beta_0}) \frac{\beta_0 - Y_t}{(\theta_0 - \beta_0)^2} I(Y_t > \beta_0)$$

$$= E\left\{ X_{j,1} \psi''(\frac{Y_1 - \beta_0}{\theta_0 - \beta_0}) \frac{\beta_0 - Y_1}{(\theta_0 - \beta_0)^2} I(Y_1 > \beta_0) \right\} + o_p(1). \quad (5.5.17)$$

Proof. Put

$$\Delta_{n1}(x, y) = C_n(x, y; j) - C(x, y; j) - C_n(x, 1; j) + C(x, 1; j) - C_n(1, y; j) + C(1, y; j),$$

$$\tilde{\Delta}_{n1}(x, y) = C_n(x, y; j) - C(x, y; j) - C_n(x, 1 - \tfrac{1}{n}; j) + C(x, 1 - \tfrac{1}{n}; j)$$
$$- C_n(1 - \tfrac{1}{n}, y; j) + C(1 - \tfrac{1}{n}, y; j),$$

$$\Delta_{n2}(x, y) = C_n(x, y; j) - C(x, y; j) - C_n(x, 1; j) + C(x, 1; j),$$

$$\Delta_{n3}(x, y) = C_n(x, y; j) - C(x, y; j) - C_n(1, y; j) + C(1, y; j).$$

111

Use the integration by parts, we have

$$\frac{1}{n}\sum_{t=1}^{n}X_{j,t}\psi'(\frac{Y_t-\beta_0}{\theta_0-\beta_0})I(Y_t>\beta_0)-E\left(X_{j,1}\psi'(\frac{Y_1-\beta_0}{\theta_0-\beta_0})I(Y_1>\beta_0)\right)$$

$$= \int_{G(\beta_0)}^{1}\int_{F_j(0)}^{1}F_j^-(x)\psi'(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})\,d\{C_n(x,y;j)-C(x,y;j)\}$$

$$+ \int_{G(\beta_0)}^{1}\int_{0}^{F_j(0)}F_j^-(x)\psi'(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})\,d\{C_n(x,y;j)-C(x,y;j)\}$$

$$= \int_{G(\beta_0)}^{1}\int_{F_j(0)}^{1}F_j^-(x)\psi'(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})\,d\Delta_{n1}(x,y)$$

$$+ \int_{G(\beta_0)}^{1}\int_{0}^{F_j(0)}F_j^-(x)\psi'(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})\,d\Delta_{n2}(x,y)$$

$$= \int_{G(\beta_0)}^{1}\int_{F_j(0)}^{1}\Delta_{n1}(x,y)\,dF_j^-(x)d\psi'(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})$$

$$+ \int_{G(\beta_0)}^{1}\int_{0}^{F_j(0)}\Delta_{n2}(x,y)\,dF_j^-(x)d\psi'(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})$$

$$= I_1+I_2.$$

Write

$$I_1 = \int_{G(\beta_0)}^{1-1/n}\int_{F_j(0)}^{1-1/n}\Delta_{n1}(x,y)\,dF_j^-(x)d\psi'(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})$$

$$+ \int_{G(\beta_0)}^{1-1/n}\int_{1-1/n}^{1}\Delta_{n1}(x,y)\,dF_j^-(x)d\psi'(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})$$

$$+ \int_{1-1/n}^{1}\int_{F_j(0)}^{1-1/n}\Delta_{n1}(x,y)\,dF_j^-(x)d\psi'(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})$$

$$+ \int_{1-1/n}^{1}\int_{1-1/n}^{1}\Delta_{n1}(x,y)\,dF_j^-(x)d\psi'(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})$$

$$= II_1+II_2+II_3+II_4$$

and

$$II_1 = \int_{G(\beta_0)}^{1-1/n}\int_{F_j(0)}^{1-1/n}\tilde{\Delta}_{n1}(x,y)\,dF_j^-(x)d\psi'(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})$$

$$+ \int_{G(\beta_0)}^{1-1/n}\int_{F_j(0)}^{1-1/n}\left\{\Delta_{n2}(x,1-\frac{1}{n})+\Delta_{n3}(1-\frac{1}{n},y)\right\}\,dF_j^-(x)d\psi'(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})$$

$$= III_1+III_2.$$

It follows from (5.2.4) and A3) that

$$\sqrt{n}III_1$$

$$= \int_{G(\beta_0)}^{1-1/n}\int_{F_j(0)}^{1-1/n}\left\{W_C(x,y;j)-W_C(x,1-\frac{1}{n};j)-W_C(1-\frac{1}{n},y;j)\right\}$$

$$dF_j^-(x)d\psi'(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0})+o_p(1),$$

112

and it follows from (5.5.4) and condition A3 that

$$
\begin{aligned}
&\sqrt{n}|III_2| \\
&= o_p\left(\int_{G(\beta_0)}^{1-1/n}\int_{F_j(0)}^{1-1/n} n^{-\delta_0}\psi''(\tfrac{G^-(y)-\beta_0}{\theta_0-\beta_0})\, dF_j^-(x)dG^-(y)\right) \\
&= o_p\left(\int_{G(\beta_0)}^{1-1/n}\int_{F_j(0)}^{1-1/n} \min\left((1-x)^{\delta_0},(1-y)^{\delta_0}\right)\psi''(\tfrac{G^-(y)-\beta_0}{\theta_0-\beta_0})\, dF_j^-(x)dG^-(y)\right) = o_p(1),
\end{aligned}
$$

which imply that

$$
\begin{aligned}
&\sqrt{n}II_1 \\
&= \int_{G(\beta_0)}^{1}\int_{F_j(0)}^{1}\left\{W_C(x,y;j)-W_C(x,1;j)-W_C(1,y;j)\right\}\, dF_j^-(x)d\psi'(\tfrac{G^-(y)-\beta_0}{\theta_0-\beta_0}) + o_p(1).
\end{aligned}
\tag{5.5.18}
$$

By (5.5.4), (5.5.5) and condition A_3), we have

$$
\begin{aligned}
\sqrt{n}II_2 &= \int_{G(\beta_0)}^{1-1/n}\int_{1-1/n}^{1}\{\alpha_n(x,y;j)-\alpha_n(1,y;j)\}\, dF_j^-(x)d\psi'(\tfrac{G^-(y)-\beta_0}{\theta_0-\beta_0}) \\
&\quad + \int_{G(\beta_0)}^{1-1/n}\int_{1-1/n}^{1}\alpha_n(x,1;j)\, dF_j^-(x)d\psi'(\tfrac{G^-(y)-\beta_0}{\theta_0-\beta_0}) \\
&= o_p\left(\int_{G(\beta_0)}^{1-1/n}\int_{1-1/n}^{1}(1-x)^{\delta}\psi''(\tfrac{G^-(y)-\beta_0}{\theta_0-\beta_0})\, dF_j^-(x)dG^-(y)\right) \\
&= o_p(1).
\end{aligned}
$$

Similarly we can show that

$$
\sqrt{n}II_3 = o_p(1) \quad \text{and} \quad \sqrt{n}II_4 = o_p(1).
$$

Hence we have

$$
\begin{aligned}
&\sqrt{n}I_1 \\
&= \int_{G(\beta_0)}^{1}\int_{F_j(0)}^{1}\left\{W_C(x,y;j)-W_C(1,y;j)-W_C(x,1;j)\right\}\, dF_j^-(x)d\psi'(\tfrac{G^-(y)-\beta_0}{\theta_0-\beta_0}) + o_p(1).
\end{aligned}
\tag{5.5.19}
$$

For I_2, write

$$I_2 = \int_{G(\beta_0)}^{1-1/n} \int_{1/n}^{F_j(0)} \Delta_{n2}(x,y)\, dF_j^-(x) d\psi'(\tfrac{G^-(y)-\beta_0}{\theta_0-\beta_0})$$

$$+ \int_{G(\beta_0)}^{1-1/n} \int_0^{1/n} \Delta_{n2}(x,y)\, dF_j^-(x) d\psi'(\tfrac{G^-(y)-\beta_0}{\theta_0-\beta_0})$$

$$+ \int_{1-1/n}^1 \int_0^{F_j(0)} \Delta_{n2}(x,y)\, dF_j^-(x) d\psi'(\tfrac{G^-(y)-\beta_0}{\theta_0-\beta_0})$$

$$= V_1 + V_2 + V_3.$$

Like the proof of (5.5.19), we can show that

$$\sqrt{n} V_1 = \int_{G(\beta_0)}^1 \int_0^{F_j(0)} \{W_C(x,y;j) - W_C(x,1;j)\}\, dF_j^-(x) d\psi'(\tfrac{G^-(y)-\beta_0}{\theta_0-\beta_0}) + o_p(1). \quad (5.5.20)$$

By (5.5.4), (5.5.5) and condition A_3), we have

$$\sqrt{n} V_2 = \int_{G(\beta_0)}^{1-1/n} \int_0^{1/n} \{\alpha_n(x,y;j) - \alpha_n(0,y;j)\}\, dF_j^-(x) d\psi'(\tfrac{G^-(y)-\beta_0}{\theta_0-\beta_0})$$

$$+ \int_{G(\beta_0)}^{1-1/n} \int_0^{1/n} \alpha_n(x,1;j)\, dF_j^-(x) d\psi'(\tfrac{G^-(y)-\beta_0}{\theta_0-\beta_0}) \quad (5.5.21)$$

$$= O_p\left(\int_{G(\beta_0)}^{1-1/n} \int_0^{1/n} x^\delta\, dF_j^-(x) d\psi'(\tfrac{G^-(y)-\beta_0}{\theta_0-\beta_0}) \right) = o_p(1)$$

and

$$\sqrt{n} V_3 = O_p\left(\int_{1-1/n}^1 \int_0^{F_j(0)} (1-y)^\delta\, dF_j^-(x) d\psi'(\tfrac{G^-(y)-\beta_0}{\theta_0-\beta_0}) \right) = o_p(1). \quad (5.5.22)$$

Therefore we have

$$\sqrt{n} I_2 = \int_{G(\beta_0)}^1 \int_0^{F_j(0)} \{W_C(x,y;j) - W_C(x,1;j)\}\, dF_j^-(x) d\psi'(\tfrac{G^-(y)-\beta_0}{\theta_0-\beta_0}) + o_p(1). \quad (5.5.23)$$

Hence, (5.5.15) follows from (5.5.19) and (5.5.23). Proofs of (5.5.16) and (5.5.17) can be done in the same way. ∎

Lemma 5.5.4. *Under conditions of Theorem 5.2.2, as $n \to \infty$, we have*

$$\sqrt{n}\left\{ \tfrac{1}{n} \sum_{t=1}^n \psi'(\tfrac{Y_t-\beta_0}{\theta_0-\beta_0}) I(Y_t > \beta_0) - E\left(\psi'(\tfrac{Y_1-\beta_0}{\theta_0-\beta_0}) I(Y_1 > \beta_0)\right) \right\}$$

$$= -\int_{G(\beta_0)}^1 W_C(1,y;j)\, d\psi'(\tfrac{G^-(y)-\beta_0}{\theta_0-\beta_0}) + o_p(1), \quad (5.5.24)$$

114

$$\frac{1}{n}\sum_{t=1}^{n}\psi''\Big(\frac{Y_t-\beta_0}{\theta_0-\beta_0}\Big)\frac{Y_t-\theta_0}{(\theta_0-\beta_0)^2}I(Y_t>\beta_0)$$
$$= E\Big\{\psi''\Big(\frac{Y_1-\beta_0}{\theta_0-\beta_0}\Big)\frac{Y_1-\theta_0}{(\theta_0-\beta_0)^2}I(Y_1>\beta_0)\Big\}+o_p(1) \tag{5.5.25}$$

and

$$\frac{1}{n}\sum_{t=1}^{n}\psi''\Big(\frac{Y_t-\beta_0}{\theta_0-\beta_0}\Big)\frac{\beta_0-Y_t}{(\theta_0-\beta_0)^2}I(Y_t>\beta_0)$$
$$= E\Big\{\psi''\Big(\frac{Y_1-\beta_0}{\theta_0-\beta_0}\Big)\frac{\beta_0-Y_1}{(\theta_0-\beta_0)^2}I(Y_1>\beta_0)\Big\}+o_p(1). \tag{5.5.26}$$

Proof. It can be shown in a similar way to the proof of Lemma 5.5.1. ∎

Proof of Theorem 5.2.2. Recall that $\hat\beta(\boldsymbol{a})$ and $\hat\theta(\boldsymbol{a})$ are the solution of estimating equations (5.2.3).
An application of Taylor expansions yields that

$$\frac{1}{n}\sum_{t=1}^{n}\psi\Big(\frac{Y_t-\hat\beta(\boldsymbol{a})}{\hat\theta(\boldsymbol{a})-\hat\beta(\boldsymbol{a})}\Big)I\big(Y_t>\hat\beta(\boldsymbol{a})\big)-(1-q)$$
$$= \frac{1}{n}\sum_{t=1}^{n}\psi\Big(\frac{Y_t-\hat\beta(\boldsymbol{a})}{\hat\theta(\boldsymbol{a})-\hat\beta(\boldsymbol{a})}\Big)I\big(Y_t>\hat\beta(\boldsymbol{a})\big)-\frac{1}{n}\sum_{t=1}^{n}\psi\Big(\frac{Y_t-\beta_0}{\theta_0-\beta_0}\Big)I(Y_t>\beta_0)$$
$$\quad+\frac{1}{n}\sum_{t=1}^{n}\psi\Big(\frac{Y_t-\beta_0}{\theta_0-\beta_0}\Big)I(Y_t>\beta_0)-(1-q)$$
$$= \big(\hat\beta(\boldsymbol{a})-\beta_0\big)\frac{1}{n}\sum_{t=1}^{n}\psi'\Big(\frac{Y_t-\beta_0}{\theta_0-\beta_0}\Big)\frac{Y_t-\theta_0}{(\theta_0-\beta_0)^2}I(Y_t>\beta_0)$$
$$\quad+\big(\hat\theta(\boldsymbol{a})-\theta_0\big)\frac{1}{n}\sum_{t=1}^{n}\psi'\Big(\frac{Y_t-\beta_0}{\theta_0-\beta_0}\Big)\frac{\beta_0-Y_t}{(\theta_0-\beta_0)^2}I(Y_t>\beta_0)$$
$$\quad+\frac{1}{n}\sum_{t=1}^{n}\psi\Big(\frac{Y_t-\beta_0}{\theta_0-\beta_0}\Big)I(Y_t>\beta_0)-(1-q)$$
$$\quad+o_p\big(|\hat\beta(\boldsymbol{a})-\beta_0|+|\hat\theta(\boldsymbol{a})-\theta_0|\big).$$

Then it follows from Lemma 5.5.1 that

$$\sqrt{n}\Big\{\frac{1}{n}\sum_{t=1}^{n}\psi\Big(\frac{Y_t-\hat\beta(\boldsymbol{a})}{\hat\theta(\boldsymbol{a})-\hat\beta(\boldsymbol{a})}\Big)I\big(Y_t>\hat\beta(\boldsymbol{a})\big)-(1-q)\Big\}$$
$$= -\int_{G(\beta_0)}^{1}W_C(1,y;j)\frac{1}{\theta_0-\beta_0}\psi'\Big(\frac{G^-(y)-\beta_0}{\theta_0-\beta_0}\Big)\,dG^-(y)$$
$$\quad+\sqrt{n}\big(\hat\beta(\boldsymbol{a})-\beta_0\big)E\Big\{\psi'\Big(\frac{Y_1-\beta_0}{\theta_0-\beta_0}\Big)\frac{Y_1-\theta_0}{(\theta_0-\beta_0)^2}I(Y_1>\beta_0)\Big\} \tag{5.5.27}$$
$$\quad-\sqrt{n}\big(\hat\theta(\boldsymbol{a})-\theta_0\big)\frac{1}{\theta_0-\beta_0}E\Big\{\psi'\Big(\frac{Y_1-\beta_0}{\theta_0-\beta_0}\Big)I(Y_1>\beta_0)\Big\}$$
$$\quad+o_p\big(\sqrt{n}|\hat\beta(\boldsymbol{a})-\beta_0|+\sqrt{n}|\hat\theta(\boldsymbol{a})-\theta_0|\big).$$

Similarly we can show that

$$
\begin{aligned}
&\frac{1}{\sqrt{n}} \sum_{t=1}^n \psi'\big(\tfrac{Y_t - \hat{\beta}(\boldsymbol{a})}{\hat{\theta}(\boldsymbol{a}) - \hat{\beta}(\boldsymbol{a})}\big)\big(Y_t - \hat{\theta}(\boldsymbol{a})\big) I\big(Y_t > \hat{\beta}(\boldsymbol{a})\big) \\
={}& -\int_{G(\beta_0)}^1 W_C(1, y; j) \Big\{ \psi''\big(\tfrac{G^-(y) - \beta_0}{\theta_0 - \beta_0}\big) \tfrac{G^-(y) - \theta_0}{\theta_0 - \beta_0} + \psi'\big(\tfrac{G^-(y) - \beta_0}{\theta_0 - \beta_0}\big) \Big\} \, dG^-(y) \\
&+ \sqrt{n}\big(\hat{\beta}(\boldsymbol{a}) - \beta_0\big) E\Big\{ \psi''\big(\tfrac{Y_1 - \beta_0}{\theta_0 - \beta_0}\big) \tfrac{(Y_1 - \theta_0)^2}{(\theta_0 - \beta_0)^2} I(Y_1 > \beta_0) \Big\} \\
&+ \sqrt{n}\big(\hat{\theta}(\boldsymbol{a}) - \theta_0\big) E\Big\{ \big(\psi''\big(\tfrac{Y_1 - \beta_0}{\theta_0 - \beta_0}\big) \tfrac{(Y_1 - \theta_0)(\beta_0 - Y_1)}{(\theta_0 - \beta_0)^2} - \psi'\big(\tfrac{Y_1 - \beta_0}{\theta_0 - \beta_0}\big)\big) I(Y_1 > \beta_0) \Big\} \\
&+ o_p\big(\sqrt{n}|\hat{\beta}(\boldsymbol{a}) - \beta_0| + \sqrt{n}|\hat{\theta}(\boldsymbol{a}) - \theta_0|\big).
\end{aligned}
\tag{5.5.28}
$$

Therefore (5.5.27) and (5.5.28) imply

$$
\sqrt{n}\begin{pmatrix} \hat{\beta}(\boldsymbol{a}) - \beta_0 \\ \hat{\theta}(\boldsymbol{a}) - \theta_0 \end{pmatrix} = \Sigma^{-1}\boldsymbol{Z} + o_p(1),
$$

where Σ and $\boldsymbol{Z}$ are given in Theorem 5.2.2. This completes the proof of the first part of the theorem.

For the second part, note that

$$
\begin{aligned}
&\hat{\theta}^{(j)}(\boldsymbol{a}) - \theta_0^{(j)}(\boldsymbol{a}) \\
={}& \frac{\frac{1}{n}\sum_{t=1}^n X_{j,t}\psi'\big(\tfrac{Y_t - \hat{\beta}(\boldsymbol{a})}{\hat{\theta}(\boldsymbol{a}) - \hat{\beta}(\boldsymbol{a})}\big) I\big(Y_t > \hat{\beta}(\boldsymbol{a})\big)}{\frac{1}{n}\sum_{t=1}^n \psi'\big(\tfrac{Y_t - \hat{\beta}(\boldsymbol{a})}{\hat{\theta}(\boldsymbol{a}) - \hat{\beta}(\boldsymbol{a})}\big) I\big(Y_t > \hat{\beta}(\boldsymbol{a})\big)} - \frac{E\big\{ X_{j,1}\psi'(\tfrac{Y_1 - \beta_0}{\theta_0 - \beta_0}) I(Y_1 > \beta_0)\big\}}{E\big\{ \psi'(\tfrac{Y_1 - \beta_0}{\theta_0 - \beta_0}) I(Y_1 > \beta_0)\big\}} \\
={}& \frac{b_2\big\{\frac{1}{n}\sum_{t=1}^n X_{j,t}\psi'\big(\tfrac{Y_t - \hat{\beta}(\boldsymbol{a})}{\hat{\theta}(\boldsymbol{a}) - \hat{\beta}(\boldsymbol{a})}\big) I\big(Y_t > \hat{\beta}(\boldsymbol{a})\big) - b_1\big\} - b_1\big\{\frac{1}{n}\sum_{t=1}^n \psi'\big(\tfrac{Y_t - \hat{\beta}(\boldsymbol{a})}{\hat{\theta}(\boldsymbol{a}) - \hat{\beta}(\boldsymbol{a})}\big) I\big(Y_t > \hat{\beta}(\boldsymbol{a})\big) - b_2\big\}}{b_2 \frac{1}{n}\sum_{t=1}^n \psi'\big(\tfrac{Y_t - \hat{\beta}(\boldsymbol{a})}{\hat{\theta}(\boldsymbol{a}) - \hat{\beta}(\boldsymbol{a})}\big) I\big(Y_t > \hat{\beta}(\boldsymbol{a})\big)},
\end{aligned}
\tag{5.5.29}
$$

where

$$
b_1 = E\Big\{ X_{j,1}\psi'(\tfrac{Y_1 - \beta_0}{\theta_0 - \beta_0}) I(Y_1 > \beta_0) \Big\} \quad \text{and} \quad b_2 = E\Big\{ \psi'(\tfrac{Y_1 - \beta_0}{\theta_0 - \beta_0}) I(Y_1 > \beta_0) \Big\}.
$$

From Lemma 5.5.3, Lemma 5.5.4 and Taylor expansions, we have

$$
\begin{aligned}
&\sqrt{n}\Big\{ \hat{\theta}^{(j)}(\boldsymbol{a}) - \theta_0^{(j)}(\boldsymbol{a}) \Big\} \\
={}& \frac{b_2\{Z_1 + \boldsymbol{A}_1^T\Sigma^{-1}\boldsymbol{Z}\} - b_1\{Z_2 + \boldsymbol{A}_2^T\Sigma^{-1}\boldsymbol{Z}\}}{b_2^2} + o_p(1) \\
={}& \frac{b_2 Z_1 - b_1 Z_2}{b_2^2} + \frac{1}{b_2^2}\{b_2\boldsymbol{A}_1 - b_1\boldsymbol{A}_2\}^T\Sigma^{-1}\boldsymbol{Z} + o_p(1),
\end{aligned}
\tag{5.5.30}
$$

where $Z_1, Z_2, \boldsymbol{A}_1, \boldsymbol{A}_2$ are given in Theorem 5.2.2. Hence the theorem follows. ∎

5.6 Conclusions

Haezendonck-Goovaerts (H-G) risk measure has been studied well in actuarial science. When it is applied to an insurance or a financial portfolio, the sensitivity analysis becomes useful. First we derive an expression for the sensitivity of the H-G risk measure. Second we derive the asymptotic distribution of the nonparametric estimator of the H-G risk measure under the assumption that the returns/losses in the portfolio follow from a strictly stationary α-mixing sequence. This generalizes the result in Ahn and Shyamalkumar (2014) from independent data to dependent data. Third, this chapter proposes a nonparametric estimator for the sensitivity and derives the asymptotic distribution, which use the derived asymptotic distribution of the nonparametric estimator of the H-G risk measure. Since uncertainty quantification is important in risk management and the obtained asymptotic variance of the nonparametric estimator for the sensitivity is quite complicated, we further propose to model each asset/loss by an AR-GARCH model and then employ a bootstrap method, resampling from the residuals in the AR-GARCH models, to construct a confidence interval for the sensitivity. We remark that a blockwise bootstrap confidence interval is not feasible in practice due to the fact that the level q in the H-G risk measure is generally set to be close to one, say 95% or 99%, and so the effective sample size in each block will not be large enough to nonparametrically estimate the H-G risk measure accurately. A simulation study shows that both the proposed nonparametric estimators and the proposed bootstrap method perform quite well. Finally we remark that the obtained results are directly applicable to risk capital allocation in portfolios.

Table 5.1: *Simulation study:* true value and nonparametric estimators of the H-G risk measure at levels 0.95 and 0.99 with $\psi(x) = x^{1.1}$ and its sensitivity are reported with corresponding standard deviation given in the brackets.

(q, n)	θ_0	$\hat{\theta}$	Bootstrap SD
(0.95, 2000)	3.6451	3.5421 (0.8326)	0.7660
(0.95, 3000)	3.6451	3.5544 (0.6666)	0.6665
(0.95, 5000)	3.6451	3.5836 (0.5906)	0.5439
(0.99, 2000)	6.1415	5.6748 (1.8855)	1.7242
(0.99, 3000)	6.1415	5.7735 (1.6126)	1.5900
(0.99, 5000)	6.1415	5.9026 (1.5193)	1.4008
(q, n)	$\theta_0^{(1)}$	$\hat{\theta}^{(1)}$	Bootstrap SD
(0.95, 2000)	5.2225	5.0264 (1.4622)	1.2615
(0.95, 3000)	5.2225	5.0662 (1.1903)	1.1262
(0.95, 5000)	5.2225	5.1314 (1.0857)	0.9533
(0.99, 2000)	8.9861	8.0798 (3.1378)	2.8348
(0.99, 3000)	8.9861	8.2974 (2.8772)	2.6925
(0.99, 5000)	8.9861	8.5769 (2.8467)	2.4850
(q, n)	$\theta_0^{(2)}$	$\hat{\theta}^{(2)}$	Bootstrap SD
(0.95, 2000)	2.0677	2.0581 (0.5223)	0.5317
(0.95, 3000)	2.0677	2.0427 (0.3835)	0.4339
(0.95, 5000)	2.0677	2.0359 (0.3111)	0.3175
(0.99, 2000)	3.2969	3.2714 (1.4598)	1.3760
(0.99, 3000)	3.2969	3.2504 (1.0744)	1.1950
(0.99, 5000)	3.2969	3.2924 (0.9272)	0.9399

Table 5.2: *Simulation study:* true value and nonparametric estimators of the H-G risk measure at levels 0.95 and 0.99 with $\psi(x) = x^{1.3}$ and its sensitivity are reported with corresponding standard deviation given in the brackets.

(q, n)	θ_0	$\hat{\theta}$	Bootstrap SD
(0.95, 2000)	3.7830	3.6496 (0.8810)	0.9342
(0.95, 3000)	3.7830	3.6700 (0.7228)	0.8478
(0.95, 5000)	3.7830	3.7087 (0.6596)	0.6317
(0.99, 2000)	6.3497	5.7926 (1.9537)	2.1686
(0.99, 3000)	6.3497	5.8971 (1.6777)	1.9927
(0.99, 5000)	6.3497	6.0484 (1.5983)	1.6266
(q, n)	$\theta_0^{(1)}$	$\hat{\theta}^{(1)}$	Bootstrap SD
(0.95, 2000)	5.4352	5.1774 (1.5238)	1.5337
(0.95, 3000)	5.4352	5.2360 (1.2926)	1.3784
(0.95, 5000)	5.4352	5.3175 (1.2932)	1.1095
(0.99, 2000)	9.3184	8.2433 (3.2420)	3.5391
(0.99, 3000)	9.3184	8.4766 (2.9879)	3.3885
(0.99, 5000)	9.3184	8.7888 (2.9856)	2.9001
(q, n)	$\theta_0^{(2)}$	$\hat{\theta}^{(2)}$	Bootstrap SD
(0.95, 2000)	2.1308	2.1221 (0.5762)	0.6356
(0.95, 3000)	2.1308	2.1041 (0.4222)	0.5130
(0.95, 5000)	2.1308	2.1000 (0.3626)	0.3703
(0.99, 2000)	3.3845	3.3419 (1.5286)	1.6232
(0.99, 3000)	3.3845	3.3176 (1.1179)	1.3641
(0.99, 5000)	3.3845	3.3079 (0.9939)	1.0253

Table 5.3: *Real data analysis:* nonparametric estimation of H-G risk measure and its sensitivity measures of portifolio (Goldman Sachs, S&P 500 index).

	$\hat{\theta}$	Bootstrap SD	$\hat{\theta}^{(1)}$	Bootstrap SD	$\hat{\theta}^{(2)}$	Bootstrap SD
$q = 0.95$	3.9390	0.5540	5.4174	0.8271	2.4636	0.6751
$q = 0.99$	7.3377	1.3005	10.2146	1.9908	4.4608	1.6542

Figure 5.1: Time series and histograms of Goldman Sachs (left) and S&P 500 Index (right).
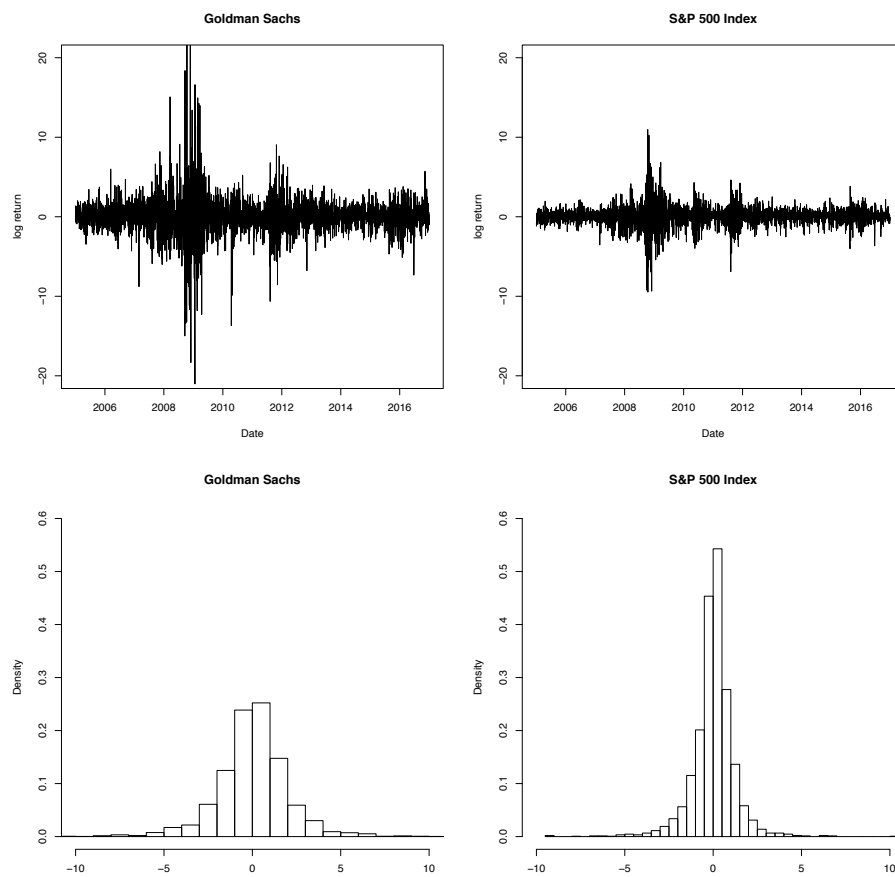
Figure 5.2: Time series of estimated residuals (top) and their autocorrelation functions (bottom) for Goldman Sachs (left) and S&P 500 Index (right).
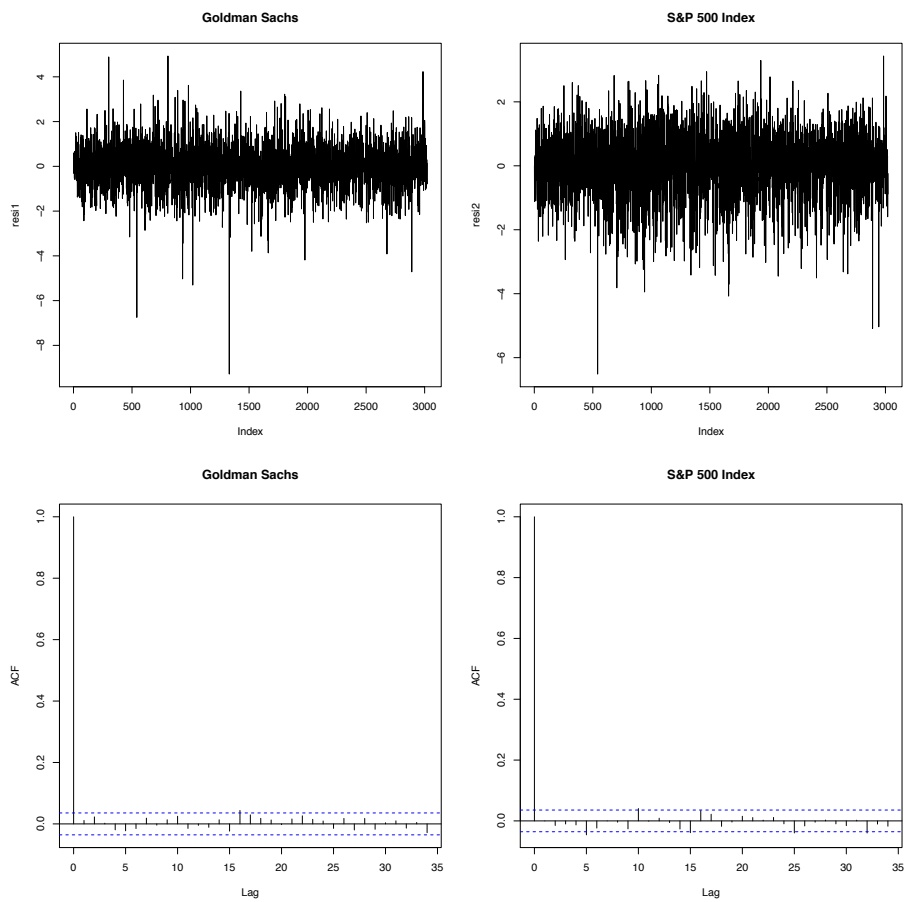
CHAPTER 6

ENDPOINT ESTIMATION FOR OBSERVATIONS WITH NORMAL MEASUREMENT ERRORS

This chapter investigates the estimation of the finite endpoint of a distribution function when the observations are contaminated by normally distributed measurement errors. Under the framework of Extreme Value Theory, we propose a class of estimators for the standard deviation of the measurement errors as well as for the endpoint. Asymptotic theories for the proposed estimators are established while their finite sample performance are demonstrated by simulations. In addition, we apply the proposed methods to the outdoor long jump data to estimate the ultimate limit for human beings in the long jump. The content of this chapter is based on the joint work:

X. Leng, L. Peng, X. Wang and C. Zhou (2018). Endpoint estimation for observations with normal measurement errors. Extremes. To appear.

6.1   Introduction to Endpoint Estimation

For a continuous distribution function F, its right endpoint is defined as $\theta = \sup\{x : F(x) < 1\}$. Estimating the endpoint θ has been applied in various contexts when $\theta < +\infty$. For example, Aarssen and De Haan (1994) estimated the endpoint of the distribution of the life span of human beings, namely, the maximum life span. In productivity analysis, estimating the production frontier can be viewed as estimating the endpoint of the distribution of outputs conditional on the inputs; see e.g. Cazals, Florens, and Simar (2002). Another notable application in Einmahl and Magnus (2008) considers estimating the ultimate world record, in other words, the limit of human being, in a specific sport. For general statistical methods designed for estimating the endpoint, Hall (1982) studied the maximum likelihood estimation, Hall and Wang (1999) proposed a minimum-distance estimation, Hall and Wang (1999) investigated the Bayesian likelihood approach, Girard, Guillou, and Stupfler (2012) used a sequence of high-order moments for endpoint estimation, and Alves and

Neves (2014) studied the finite endpoint estimate with distributions from the Gumbel domain of attraction.

Since the endpoint is only related to the right tail region of the distribution, it is natural to consider Extreme Value Theory (EVT) in endpoint estimation. EVT models the tail region of a distribution function. The endpoint is finite if the distribution function belongs to the domain of attraction of the Weibull distribution. By regarding the endpoint as a high quantile with a probability level tending to one, the limit of the high quantile estimator can be considered as an endpoint estimator. Therefore, one may estimate the endpoint using the estimators on the extreme value index, scale and shift; see e.g. Chapter 4.5 of De Haan and Ferreira (2007).

In reality, data are often contaminated by measurement errors. In other words, instead of observing independent and identically distributed (i.i.d.) sample drawn from the underlying distribution with a finite endpoint, we observe data as the convolution of the initial random variable and an independent measurement error term. Mathematically, suppose $X_1, \ldots, X_n$ are random variables with a continuous distribution function F_X and a finite endpoint θ. Instead of observing $\{X_i\}_{i=1}^n$, we observe $Y_i = X_i + \varepsilon_i, i = 1, 2, \ldots, n$, where $\{\varepsilon_i\}$ are $i.i.d.$ random errors with mean zero, and $\{\varepsilon_i\}$ are independent of $\{X_i\}$. The goal in this chapter is to estimate the endpoint θ based on the observations $\{Y_i\}_{i=1}^n$ when ε_i follows a normal distribution. Notice that in this case, the distribution function of Y_i has an infinite endpoint. That is, existing endpoint estimators designed for the case of no measurement errors are inconsistent under the above setup.

In a broader context, extracting the distribution of X_i based on the observations $\{Y_i\}_{i=1}^n$ is related to the so-called "deconvolution" problem. In the literature of nonparametric deconvolution, kernel estimators for the density function of $\{X_i\}$ were proposed; see, e.g., Carroll and Hall (1988); Stefanski and Carroll (1990); Meister (2006); Meister and Neumann (2010). Based on estimating the density function, Hall and Simar (2002) proposed an estimator of θ assuming that the density f_X is approximated by a flat function in the neighborhood of θ and the variance of the measurement error, $\sigma_n^2 = \text{var}(\varepsilon_i)$, depending on the sample size n, shrinks to zero as $n \to \infty$. By contrast, Kneip, Simar, and Van Keilegom (2015) did not require that σ tends to zero as the sample size

increases while dealing with a constant σ. They proposed a joint estimation of θ and σ when the measurement errors $\{\varepsilon_i\}$ follow a normal distribution $N(0, \sigma^2)$. Nevertheless, both of these two approaches require that $f_X(\theta-) > 0$. Compared to these studies, we aim at estimating the endpoint for a broader class of f_X allowing $f_X(\theta-)$ to be positive, zero and infinite.

Our approach is close to that in Goldenshluger and Tsybakov (2004), which proposed an endpoint estimator for data with measurement error. They modeled f_X near θ by a power function with index α, i.e., $f_X(\theta-) = 0$, and derived that as $n \to \infty$,

$$\frac{\sqrt{\log n}}{\log \log n} \left(\max_{1 \le i \le n} Y_i - \sigma \sqrt{2 \log n} - \theta \right) = O_p(1),$$

Correspondingly, Goldenshluger and Tsybakov (2004) proposed to estimate θ by $\max_{1 \le i \le n} Y_i - \sigma\sqrt{2 \log n}$, where σ is assumed to be known and this estimator has a degenerate limit.

In practice, the estimator proposed in Goldenshluger and Tsybakov (2004) is not applicable because of the unknown σ. This chapter aims at filling this gap by providing an estimator of the endpoint when σ is unknown. More specifically, under an EVT model for F_X, we first estimate σ by using the top k order statistics of $\{Y_i\}$, where k is an intermediate sequence such that $k \to \infty$ and $k/n \to 0$ as $n \to \infty$. With some proper conditions on the intermediate sequence k, the estimator for σ possesses asymptotic normality with a speed of convergence $\sqrt{k}$. Next we estimate the endpoint by $\hat{\theta} = \max_{1 \le i \le n} Y_i - \hat{\sigma}\sqrt{2 \log n}$. This estimator inherits the asymptotic normality of $\hat{\sigma}$, however, with a slightly compromised speed of convergence.

Compared to the existing literature, our approach has two main advantages. Firstly, we do not require that F_X has a density, but only assume a second–order expansion of the survival function $\overline{F}_X(x) = 1 - F_X(x)$ in the neighborhood of θ. Compared to Kneip, Simar, and Van Keilegom (2015), our model assumption allows for a broader class of F_X, which includes the model in Goldenshluger and Tsybakov (2004) as a special case. Secondly, we do not require any additional information such as the variance of the measurement errors. Therefore, our approach is more close to the real situation encountered in applications.

Our estimation procedure and its asymptotic properties are discussed in Section 6.2. We conduct a simulation study in Section 6.3 and then apply our method to sports data in Section 6.4. The proofs are postponed to Section 6.5. Section 6.6 concludes.

6.2   Estimate Endpoints with Observation Errors

Recall that $X_1, \ldots, X_n$ are $i.i.d.$ random variables with a continuous distribution function F_X that has a finite right endpoint $\theta := \sup\{x : F_X(x) < 1\}$. Further, in a neighborhood of θ, we assume the following second–order expansion of the survival function $\overline{F}_X(x) = 1 - F_X(x)$: as $u \to 0^+$,

$$\overline{F}_X(\theta - u) = Lu^\alpha + du^{\alpha+\beta} + o(u^{\alpha+\beta}), \tag{6.2.1}$$

where L, α and β are positive constants, and $d \neq 0$. Under the condition (6.2.1), F_X belongs to the maximum domain of attraction of the Weibull distribution with an extreme value index $\gamma = -1/\alpha$. Suppose we do not observe $\{X_i\}$ directly. Instead, we observe $Y_i = X_i + \varepsilon_i, i = 1, 2, \ldots, n$, where $\{\varepsilon_i\}$ are $i.i.d.$ normally distributed random measurement errors with mean zero and unknown variance σ^2, and $\{\varepsilon_i\}$ are independent of $\{X_i\}$. The question is how to estimate θ based on the observed $\{Y_i\}$.

Recall that with assuming that σ is known, Goldenshluger and Tsybakov (2004) proposed to estimate θ by $\max_{1 \leq i \leq n} Y_i - \sigma\sqrt{2 \log n}$. We will first construct an estimator for σ, and then an estimator for θ. The estimator of the parameter σ is constructed in three steps. Firstly, we investigate the tail expansion of the distribution of Y_i. Secondly, we obtain the approximation of the tail quantile process based on $\{Y_i\}_{i=1}^n$. It will turn out that the parameter σ is a scaling factor in the approximation. Therefore, in the last step we use a weighted moment method to construct an estimator of σ based on inter-quantile ranges.

We first investigate the tail expansion of the distribution function. Without loss of generality, we consider the tail expansion of the distribution of Z_i with $Z_i = Y_i - \theta$. Denote the distribution function and survival function of Z_i as F_Z and $\overline{F}_Z = 1 - F_Z$, respectively. We first show $\overline{F}_Z$ has a

second–order expansion as in the following proposition.

Proposition 6.2.1. *Suppose a random variable X has a distribution function F_X with a finite right endpoint $\theta = \sup\{x : F_X(x) < 1\} < \infty$. Assume that F_X satisfies the condition (6.2.1). ε is a normally distributed random error with mean zero and an unknown variance $\sigma^2 > 0$ and is independent of X. Denote $Z = X + \varepsilon - \theta$. Then, as $t \to \infty$,*

$$-\log \overline{F}_Z(t) = \frac{t^2}{2\sigma^2} + (\alpha + 1)\log t - \log c_1 - \frac{c_2}{c_1}t^{-\beta'}(1 + o(1)),$$

where

$$\beta' = \min(\beta, 2), \ c_1 = \frac{\Gamma(\alpha + 1)L\sigma^{2\alpha+1}}{\sqrt{2\pi}},$$

and

$$c_2 = \frac{1}{\sqrt{2\pi}}\Gamma(\alpha + \beta + 1)d\sigma^{2(\alpha+\beta)+1}1_{\beta'=\beta} - \frac{1}{\sqrt{2\pi}}\left(\frac{\alpha}{2} + 1\right)\Gamma(\alpha + 2)L\sigma^{2\alpha+3}1_{\beta'=2}.$$

Next, the tail property of F_Z provides an approximation for the tail quantile process based on $\{Z_i\}_{i=1}^n$ or $\{Y_i\}_{i=1}^n$ as follows. Denote $Y_{1,n} \leq Y_{2,n} \leq \cdots \leq Y_{n,n}$ as the order statistics of $Y_1, \cdots, Y_n$. Then the order statistics of the unobserved random variables $\{Z_i\}_{i=1}^n$ are $Z_{n-i,n} = Y_{n-i,n} - \theta$. Consider an intermediate sequence $k = k(n)$ such that $k \to \infty$ and $k/n \to 0$ as $n \to \infty$. Since $\overline{F}_Z(Z_{n-i,n}) \approx i/n$, we have approximately from Proposition 6.2.1 that

$$\sqrt{-\log(i/n)} \approx \frac{Y_{n-i,n} - \theta}{\sqrt{2}\sigma},$$

for $i = 1, 2, \ldots, k$.

Finally, we can estimate the parameter σ using a weighted moment method as follows. By considering the inter-quantile range between $Y_{n-i,n}$ and $Y_{n-k,n}$, we get that

$$\frac{Y_{n-i,n} - Y_{n-k,n}}{\sqrt{2}\sigma} \approx \sqrt{-\log(i/n)} - \sqrt{-\log(k/n)} = \frac{\log(k/i)}{\sqrt{\log(n/i)} + \sqrt{\log(n/k)}} \approx \frac{\log(k/i)}{2\sqrt{\log(n/k)}}.$$

Take any positive continuous function g on $(0, 1]$ such that $\int_0^1 g(s)(-\log s)\, ds = 2$. We construct a weighted sum of the differences $\{Y_{n-i,n} - Y_{n-k,n}\}$ using the weights $\{\frac{1}{k} g(i/k)\}$ as

$$\frac{1}{k} \sum_{i=1}^{k-1} g(i/k) \frac{Y_{n-i,n} - Y_{n-k,n}}{\sqrt{2}\sigma} \approx \frac{1}{2\sqrt{\log(n/k)}} \int_{1/k}^1 g(s)(-\log s)\, ds \approx \frac{1}{\sqrt{\log(n/k)}}.$$

Hence, we get an estimator of σ as

$$\hat{\sigma}_g = \frac{\sqrt{\log(n/k)}}{\sqrt{2}k} \sum_{i=1}^{k-1} g(i/k)(Y_{n-i,n} - Y_{n-k,n}). \tag{6.2.2}$$

Replacing σ with $\hat{\sigma}_g$ in $Y_{n,n} - \sigma\sqrt{2 \log n}$, an estimator of θ is then given as

$$\hat{\theta}_g = Y_{n,n} - \hat{\sigma}_g \sqrt{2 \log n}. \tag{6.2.3}$$

Note that $\hat{\theta}_g$ is always smaller than the maximum $Y_{n,n}$ from the positiveness of $\hat{\sigma}_g$. This is different from the endpoint estimator proposed in Alves and Neves (2014), which is based on observations without measurement errors and the assumption that the underlying distribution is in the Gumbel domain of attraction and has a finite endpoint.

To obtain the asymptotic property of the above estimator, we further assume that $\beta > 1$ and the intermediate sequence k satisfies the following condition: as $n \to \infty$,

$$k = k(n) \to \infty, k/n \to 0, \text{ and } \sqrt{k} \left(\log(n/k)\right)^{-\frac{\beta'}{2}} \left(\log\log(n/k)\right)^{2 \cdot 1_{\beta'=2}} = O(1). \tag{6.2.4}$$

We first prove the asymptotic property for $\hat{\sigma}_g$. That of $\hat{\theta}_g$ will then follow as a direct corollary.

The asymptotic normality of $\hat{\sigma}_g$ requires the following conditions on $g(s)$, $s \in (0, 1]$. There exists $\epsilon_0 > 0$, such that

$$\lim_{s \to 0} g(s)\, s^{1/2 - \epsilon_0} = 0, \text{ and } \int_0^1 g(s)(-\log s)\, ds = 2. \tag{6.2.5}$$

In addition, we assume a joint condition on the intermediate sequence k and the function g as

127

follows,

$$\lim_{n \to \infty} \sqrt{k} \sup_{|s-t| \le 1/k, 1/k \le s, t \le 1} \left| \frac{g(s)}{\log s} - \frac{g(t)}{\log t} \right| = 0. \tag{6.2.6}$$

Examples such as $g(s) = -\log s$, and $g(s) = 2(\nu + 1)^2 s^\nu$, $\nu > \frac{1}{2}$, satisfy all conditions.

The following theorem gives the asymptotic normality of $\hat{\sigma}_g$.

Theorem 6.2.2. *Suppose $X_1, \ldots, X_n$ are i.i.d. random variables with a continuous distribution function F_X. Assume that F_X has a finite right endpoint $\theta = \sup\{x : F_X(x) < 1\}$ and satisfies the condition (6.2.1) with $\beta > 1$. Suppose $Y_i = X_i + \varepsilon_i, i = 1, 2, \ldots, n$, where $\{\varepsilon_i\}$ are i.i.d. normally distributed random errors with mean zero and an unknown variance $\sigma^2 > 0$, and $\{\varepsilon_i\}$ are independent of $\{X_i\}$. Assume that $g : (0, 1] \to (0, \infty)$ satisfies (6.2.5) and $k := k(n)$ is an intermediate sequence satisfying (6.2.4) and (6.2.6). Then, as $n \to \infty$, the estimator $\hat{\sigma}_g$ defined in (6.2.2) has the following asymptotic property:*

$$\sqrt{k} \left(\frac{\hat{\sigma}_g}{\sigma} - 1 \right) \xrightarrow{d} N \left(0, \frac{1}{4} \int_0^1 \int_0^1 g(s)g(t) \left(\frac{\min(s, t)}{st} - 1 \right) ds dt \right).$$

Consequently, the estimator $\hat{\theta}_g$ defined in (6.2.3) possesses asymptotic normality as follows.

Corollary 6.2.3. *Under the same conditions as in Theorem 6.2.2, as $n \to \infty$,*

$$\frac{\sqrt{k}}{\sqrt{\log n}}(\hat{\theta}_g - \theta) \xrightarrow{d} N \left(0, \frac{\sigma^2}{2} \int_0^1 \int_0^1 g(s)g(t) \left(\frac{\min(s, t)}{st} - 1 \right) ds dt \right).$$

Remark 6.2.4. *We substitue σ with $\hat{\sigma}_g$ when calculating the confidence interval for $\hat{\theta}_g$ based on Corollary 6.2.3 in Section 6.4.*

Remark 6.2.5. *With $g(s) = -\log s$ and $g(s) = 2(\nu + 1)^2 s^\nu$ for $\nu > 1/2$, Corollary 6.2.3 correspondingly implies*

$$\frac{\sqrt{k}}{\sqrt{\log n}}(\hat{\theta}_g - \theta) \xrightarrow{d} N \left(0, \frac{5}{2}\sigma^2 \right),$$

and

$$\frac{\sqrt{k}}{\sqrt{\log n}}(\hat{\theta}_g - \theta) \xrightarrow{d} N \left(0, \frac{2(\nu + 1)^2}{2\nu + 1}\sigma^2 \right).$$

6.3 Simulations

In this section, we investigate, through simulations, the finite sample behavior of the suggested endpoint estimator. We generate observations $Y_i = X_i + \varepsilon_i, i = 1, 2, \ldots, n$, where $\{X_i\}_{i=1}^n$ and $\{\varepsilon_i\}_{i=1}^n$ are two sets of i.i.d. random variables independently drawn from the following data generating processes. In all three cases, we set the true value of the endpoint for X_i to $\theta = 0$, while setting the distribution of ε_i to a normal distribution $N(0, \sigma^2)$ with two potential levels of σ.

(a) X_i follows a uniform distribution on $[-1, 0]$ and $\sigma = 0.1$ or 0.2. Notice that condition (6.2.1) holds for the uniform distribution with $\alpha = 1$ and $\beta = \infty$.

(b) X_i follows a reversed Burr distribution with the following distribution function

$$F_X(x) = 1 - \left(1 + (-x)^{-2}\right)^{-10}, \ x < 0,{}^{*}$$

and $\sigma = 2$ or 3. Notice that condition (6.2.1) holds for the reversed Burr distribution with $\alpha = 20$ and $\beta = 22$.

(c) X_i follows a shifted Beta distribution with the following probability density function

$$f_X(x) = -42x(1+x)^5, \ -1 < x < 0,$$

and $\sigma = 0.1$ or 0.2. Notice that condition (6.2.1) holds for the shifted Beta distribution with $\alpha = 2$ and $\beta = 1$, which violates our required condition $\beta > 1$ in Theorem 6.2.2.

From each data generating process, we draw $r = 1000$ random samples with a sample size $n = 5000$. Then we estimate the endpoint for each sample using an estimator $\hat{\theta}_{g_0}$ with a specific choice, $g_0(s) = -\log s$ on $(0, 1]$. It is straightforward to check that g_0 satisfies the condition (6.2.5).

During the estimation procedure, we need to determine the value of k, i.e. the number of upper order statistics used. For that purpose, we perform a pre-study as follows. By varying

---

*The quantiles of the reversed Burr distribution are $q_{.10} = -9.72, q_{.25} = -5.85, q_{.75} = -2.59$ and $q_{.90} = 1.97$.

k from 5 to 50, we plot the average estimate[†] $\bar{\theta}_{g_0} := r^{-1} \sum_{i=1}^{r} \hat{\theta}_{g_0,i}$, the standard deviation: $\sqrt{r^{-1} \sum_{i=1}^{r} (\hat{\theta}_{g_0,i} - \bar{\theta}_{g_0})^2}$ and the root mean squared error (RMSE): $\sqrt{r^{-1} \sum_{i=1}^{r} \hat{\theta}_{g_0,i}^2}$ for each data generating process. In Figure 6.1, we demonstrate the results for the two data generating processes in (b). We observe that the optimal k that minimizes the RMSE is achieved at about $k = 10$. We do choose $k = 10$ throughout the simulation study, also for the other data generating processes in (a) and (c).[‡]

Figure 6.1: Endpoint estimation for various k: Reversed Burr Distribution
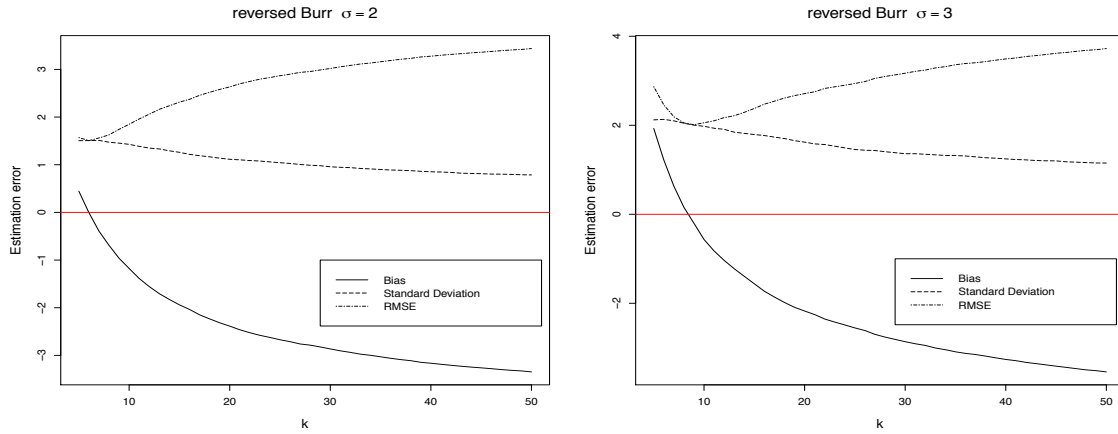
Note: The figure shows the bias, standard deviation and RMSE for the estimates $\hat{\theta}_{g_0}$ across 1000 samples with sample size $n = 5000$. The observations are generated by combining the reversed Burr distribution in (b) with measurement errors following $N(0, \sigma^2)$, $\sigma = 2$ (left) or $\sigma = 3$ (right).

We compare the performance of our suggested endpoint estimator with that of the probability weighted moment (PWM) estimator for the endpoint, the general endpoint estimator proposed in Alves and Neves (2014) and the high-order moments estimator for the endpoint in Girard, Guillou, and Stupfler (2012), respectively. More precisely, the PWM estimator is defined as

$$\hat{\theta}_{PWM} = Y_{n-k,n} - \frac{\hat{a}_{PWM}(n/k)}{\hat{\gamma}_{PWM}}, \tag{6.3.1}$$

[†]Since the true endpoint equals 0, the average estimate can be read as the estimation bias.

[‡]The figures for the other data generating processes are available upon request.

where

$$\hat{\gamma}_{PWM} := \frac{I_1 - 4I_2}{I_1 - 2I_2}, \text{ and } \hat{a}_{PWM}(n/k) := \frac{2I_1 I_2}{I_1 - 2I_2},$$

with the probability weighted moments given by

$$I_j = \frac{1}{k} \sum_{i=1}^{k} \left(\frac{i}{k}\right)^{j-1} (Y_{n-i+1,n} - Y_{n-k,n}), \quad j = 1, 2.$$

Note that $\hat{\theta}_{PWM}$ is an estimator of θ only if there is no measurement error in the observations, i.e. $\varepsilon_i = 0, i = 1, 2, \ldots, n$. The general endpoint estimator is given by

$$\hat{\theta}^F := Y_{n,n} + \sum_{i=0}^{k-1} d_{i,k}(Y_{n-k,n} - Y_{n-k-i,n}), \tag{6.3.2}$$

where $d_{i,k} = -(\log 2)^{-1} \left(\log(k+i) - \log(k+i+1)\right) > 0$, satisfying $\sum_{i=0}^{k-1} d_{i,k} = 1$. Since the weighted spacings in (6.3.2) are non-negative, we observe that $\hat{\theta}^F$ is greater than the sample maximum $Y_{n,n}$ whereas our suggested endpoint estimator in (6.2.3) is less than $Y_{n,n}$. Note that $\hat{\theta}^F$ estimates the endpoint of a distribution from the Gumbel domain of attraction with finite endpoint. The high-order moments estimator for the finite endpoint is defined as

$$\hat{\theta}_M := b\,p_n \left[((b+1)p_n + 1)\frac{\hat{\mu}_{(b+1)p_n}}{\hat{\mu}_{(b+1)p_n+1}} - (p_n + 1)\frac{\hat{\mu}_{p_n}}{\hat{\mu}_{p_n+1}} \right]^{-1} + c_0, \tag{6.3.3}$$

where $b > 0$, $\{p_n\}$ is a nonrandom sequence such that $p_n \to \infty$ and

$$\hat{\mu}_{p_n} = \frac{1}{n} \sum_{i=1}^{n} (Y_i - c_0)^{p_n}$$

with $c_0 = Y_{1,n}$. To ensure that the data used in the endpoint estimation are positive, which is required in Girard, Guillou, and Stupfler (2012), we artificially subtract the sample minimum from the sample, and then add it back in getting $\hat{\theta}_M$. Since the endpoint of each data generating process is infinite, which violates the finite endpoint assumption in both $\hat{\theta}^F$ and $\hat{\theta}_M$, we do not expect $\hat{\theta}^F$ and $\hat{\theta}_M$ perform better than our suggested endpoint estimator in estimating the endpoint for each

simulated case.

To determine the optimal choice of k used in the estimators $\hat{\theta}_{PWM}$ and $\hat{\theta}^F$, respectively, we also conduct a pre-study similar to the aforementioned procedure. We decide to choose $k = 200$ and $k = 70$ for all data generating processes when applying $\hat{\theta}_{PWM}$ and $\hat{\theta}^F$, respectively. Notice that for the PWM estimator, we stop estimating the endpoint if $\hat{\gamma}_{PWM} > 0$ because the PWM estimator is valid only for $\gamma < 0$. In other words, from 1000 samples, we may end up with less than 1000 endpoint estimates when using the PWM estimator. For applying the estimator $\hat{\theta}_M$, we choose $p_n = n^{1/\alpha}/\log\log n$ and b from a set $\mathcal{B} = \{0.2, 0.6, 1.0, \ldots, 21\}$ as suggested in Girard, Guillou, and Stupfler (2012). Also following from a pre-study as before, the optimal choice of b from $\mathcal{B}$ is 20.6.

For each data generating process, we plot the estimated endpoints using the four estimators across all samples with measurement errors in boxplots; see Figure 6.2. We observe that both the PWM estimator $\hat{\theta}_{PWM}$ and the general endpoint estimator $\hat{\theta}^F$ overestimate the true endpoint across all data generating processes, and the high-order moments estimator $\hat{\theta}_M$ underestimates the true endpoint for the cases (a) and (b) and overestimates it in the case (c). In contrast, our estimator $\hat{\theta}_{g_0}$ performs well across all cases, and consistently outperforms all the other estimators. The medians across 1000 simulated samples are close to the true endpoint and the variations are lower.

As our theorems require the existence of measurement errors, we investigate the performance of the proposed endpoint estimator when there is no measurement error in each data generating process, i.e., data are simulated only from the distribution F_X corresponding to the cases (a), (b) and (c). We observe from Figure 6.3 that our estimator $\hat{\theta}_{g_0}$ underestimates the true endpoint across all cases without measurement errors, and is the worst-performing corresponding to the cases (a) and (c). This poses some interesting and challenging questions such that how to test the existence of measurement error and whether there is a consistent endpoint estimator regardless of the presentence of measurement errors.

Figure 6.2: Boxplots of estimated endpoints with measurement errors
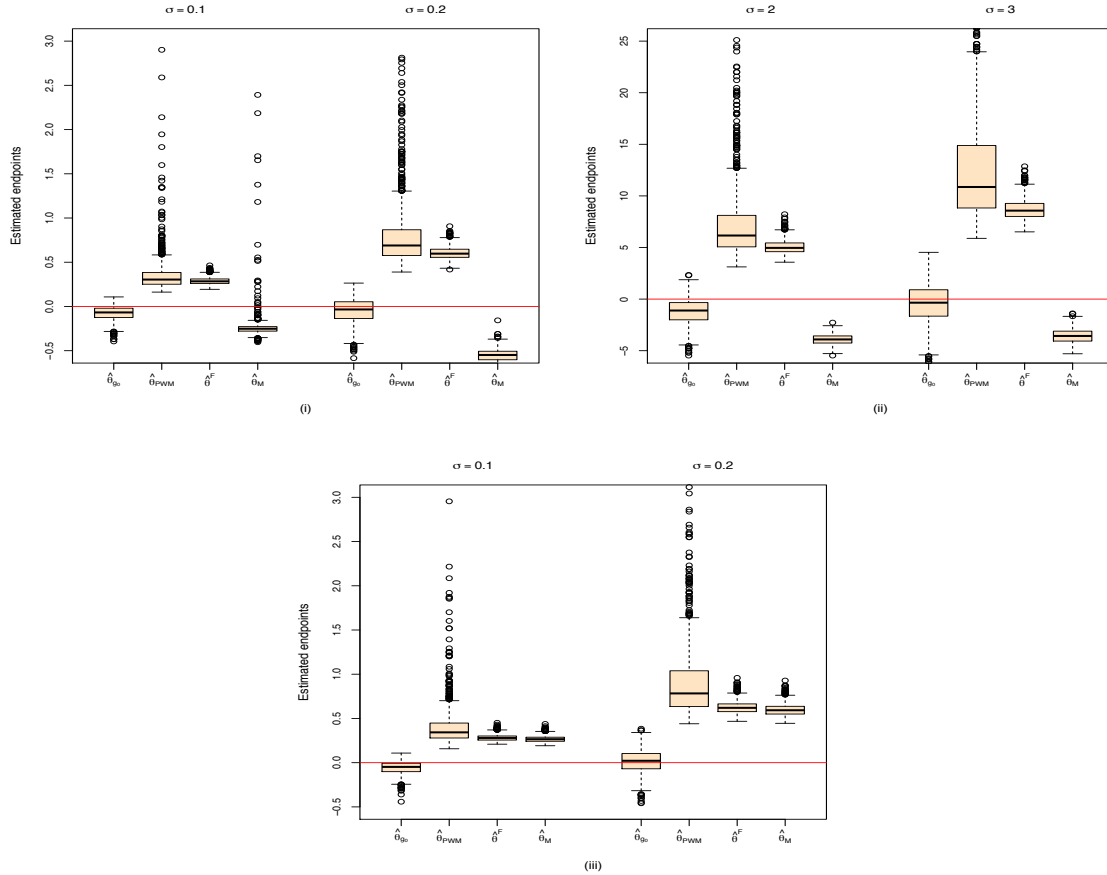
Note: The three plots show the estimated endpoints using our suggested endpoint estimator $\hat{\theta}_{g_0}$, the PWM estimator $\hat{\gamma}_{PWM}$, the general endpoint estimator $\hat{\theta}^F$ and the high-order moments estimator $\hat{\theta}_M$ for six data generating processes. Each plot is based on 1000 samples with sample size n=5000. In the panel (i), the observations are generated by combining the uniform distribution on $[-1, 0]$ with measurement errors following $N(0, \sigma^2)$, $\sigma = 0.1$ (left) or $\sigma = 0.2$ (right). In the panel (ii), the observations are generated by combining the reversed Burr distribution in (b) with measurement errors following $N(0, \sigma^2)$, $\sigma = 2$ (left) or $\sigma = 3$ (right). In the panel (iii), the observations are generated by combining the shifted beta distribution in (c) with measurement errors following $N(0, \sigma^2)$, $\sigma = 0.1$ (left) or $\sigma = 0.2$ (right). The endpoints are estimated by $\hat{\theta}_{g_0}$ with $k = 10$, $\hat{\theta}_{PWM}$ in (6.3.1) with $k = 200$, $\hat{\theta}^F$ in (6.3.2) with $k = 70$ and $\hat{\theta}_M$ in (6.3.3) with $p_n = n^{1/\alpha}/\log\log n$ and $b = 20.6$. Horizontal lines indicate the true endpoints.

Figure 6.3: Boxplots of estimated endpoints without measurement errors
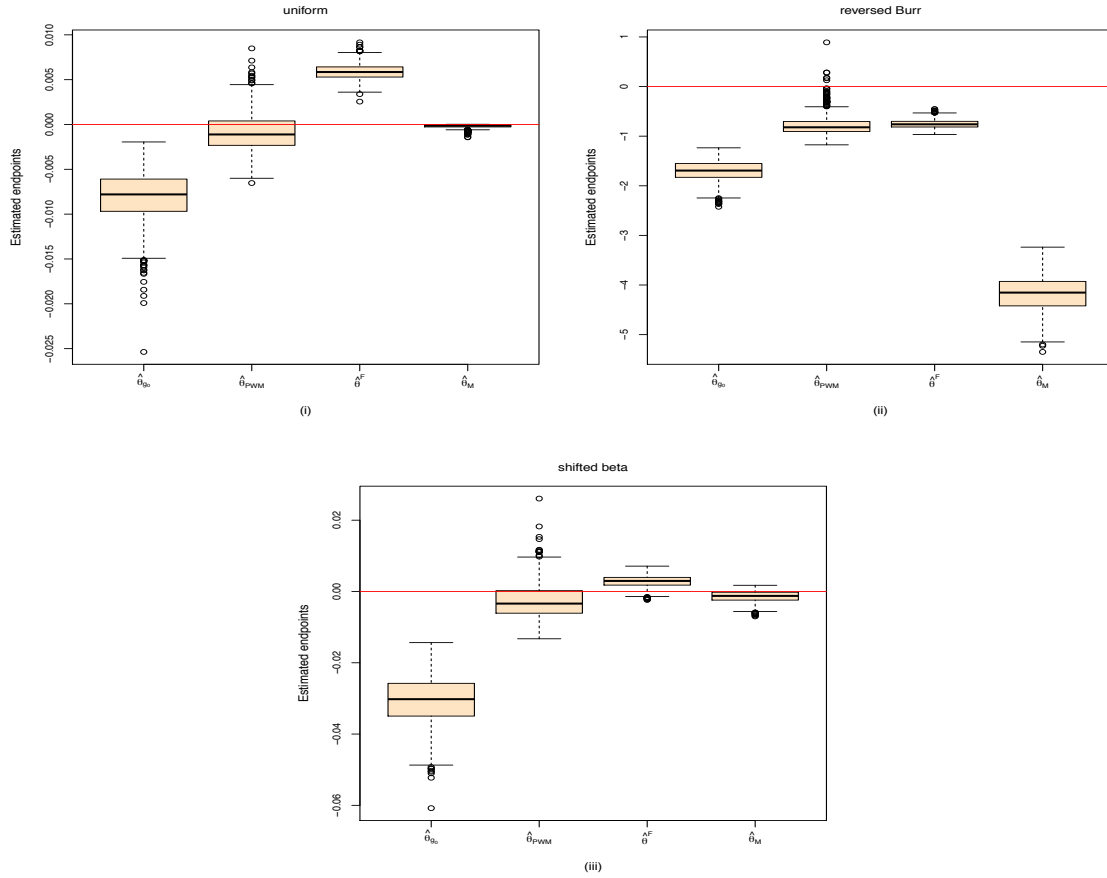
Note: The three plots show the estimated endpoints using our suggested endpoint estimator $\hat{\theta}_{g_0}$, the PWM estimator $\hat{\gamma}_{PWM}$, the general endpoint estimator $\hat{\theta}^F$ and the high-order moments estimator $\hat{\theta}_M$ for six data generating processes. Each plot is based on 1000 samples with sample size n=5000. In the panel (i), the observations are generated by the uniform distribution on $[-1, 0]$ in (a). In the panel (ii), the observations are generated by the reversed Burr distribution in (b). In the panel (iii), the observations are generated by the shifted beta distribution in (c). The endpoints are estimated by $\hat{\theta}_{g_0}$ with $k = 10$, $\hat{\theta}_{PWM}$ in (6.3.1) with $k = 200$, $\hat{\theta}^F$ in (6.3.2) with $k = 70$ and $\hat{\theta}_M$ in (6.3.3) with $p_n = n^{1/\alpha}/\log\log n$ and $b = 20.6$. Horizontal lines indicate the true endpoints.

6.4 Application

In order to investigate the limit of human being in sports, Einmahl and Magnus (2008) applies the endpoint estimation to the training data of top athletes. Initially, they gathered data for 28 types of sports. However, they stopped estimating the endpoint for five out of the 28 sports due to the fact that the estimated extreme value indices, γ, for these five sports are close to 0.[§]

Continuing from their study, we shall apply our estimator on the endpoint to the outdoor long jump data (both men and women) for two reasons. Firstly, from a theoretical perspective, if we assume that the observed training data are contaminated by normally distributed measurement errors, the extreme value index γ for the observations should be equal to 0. This is in line with the empirical observations in Einmahl and Magnus (2008). We will justify this argument by repeating the estimation of the extreme value index γ for the training data and testing whether it is significantly below zero. Different from Einmahl and Magnus (2008), we employ the PWM estimator for this analysis. Secondly, for the outdoor long jump, the presence of wind can be a potential factor generating such a measurement error. To justify this argument, we shall apply our suggested estimator to the training data for the outdoor long jump, while comparing it with applying the PWM endpoint estimator to the training data for the indoor long jump. Assuming that the indoor long jump is much less affected by wind, we expect that the two endpoints estimated from these two different datasets mutually agree with each other.

We collect the data from the official website of the International Association of Athletics Federations (IAAF)[¶] for the indoor and outdoor long jump, both for men and women. In total, we construct four datasets for these four sports. For each sport, the website presents the all time personal bests of the top athletes. In addition, for indoor long jump (both men and women), the website provides the personal bests of the top athletes in each year from 1999 to 2016. Consequently, for each of these two sports, we combine the data across the aforementioned 19 lists. Similarly for men's outdoor long jump, we combine the data across 17 lists because the records of the years

---

[§]The five sports are $10,000$m running and outdoor long jump for both men and women, together with men's 400m running.

[¶]See https://www.iaaf.org/records/toplists

2000 and 2001 are not available. For women's outdoor long jump, 18 lists are combined due to the missing records in 2001. When combining the data for each sport, we keep only the best record for each athlete across the lists. Table 6.1 gives a summary of the number of athletes and the best and worst achievements for each of the four sports. Since there are clusters in the present data, we smooth each dataset using the method suggested in Einmahl and Magnus (2008) as follows. For example, if c athletes share the same personal best, $l = 8.47$ m, we smooth them by

$$l_i = 8.465 + .01\frac{2i - 1}{2c}, \quad i = 1, \ldots, c.$$

Table 6.1: *Data description*

| | men | | | women | | |
Long jump	Number	Best	Worst	Number	Best	Worst
Outdoor	776	8.95	7.80	1176	7.52	6.30
Indoor	623	8.79	7.70	604	7.37	6.30

Note: The table presents the descriptive information regarding the four datasets used in the application. The four datasets correspond to the indoor and outdoor long jump data for both men and women. Each dataset consists of top athletes' personal best in the corresponding sport.

We start with estimating the extreme value index γ by the PWM estimator. Figure 6.4 shows the estimates $\hat{\gamma}_{PWM}$ against various values of k. To balance the estimation bias and variance, we choose k from the first stable region in each plot. Table 6.2 reports the chosen values of k and the corresponding estimated γ and its 95% confidence intervals.

From the confidence intervals, we observe that for indoor long jump, $\gamma = 0$ is rejected for both men and women. Hence, the endpoints of these two distributions exist. By contrast, for outdoor long jump, we cannot reject $\gamma = 0$ for either men or women. These results agree with the finding in Einmahl and Magnus (2008) and support considering the outdoor data as contaminated by measurement errors.

Next, we continue using the PWM method to estimate the endpoints for the indoor long jump data, while using our new estimator, $\hat{\theta}_{g_0}$, to estimate the endpoint for the outdoor long jump data.

We plot the endpoint estimates against various values of k in Figure 6.5. A technical difference between these two estimators is regarding the sample size n. For the PWM estimator, the sample size n is not used whereas for our new method, it is necessary to know the sample size n in advance. Obviously, a lower bound for n is the current sample size, i.e. the number of top athletes included in the IAAF website. We are aware of the caveat that this number may underestimate the true number of top athletes who may potentially produce a similar performance. To address this caveat and test the sensitivity of n, we take an arbitrary value $n = 3000$ which is much higher than the current sample sizes.

Tables 6.2 and 6.3 present the endpoint estimates with their 95% confidence intervals based on the selected values of k. Firstly, we observe that the values of $\hat{\theta}_{g_0}$ based on the outdoor long jump data are not sensitive to the sample size n, particularly after considering the estimation error reflected by the confidence intervals. Secondly, although the point estimates from applying our new estimator to the outdoor long jump data are consistently lower than that from applying the PWM method to the indoor long jump data, the two results agree with each other to certain extent. The point estimates from applying our new estimator to the outdoor long jump data falls into the confidence interval based on applying the PWM method to the indoor long jump data.

Finally, we compare the results to the actual observations in the dataset. Although the PWM estimator based on the indoor long jump data suggested that the endpoints for men's and women's long jump are at 8.790 and 7.394 respectively, there are multiple observations in the outdoor long jump data that exceed those estimated endpoints: for men's long jump there are four observations higher than 8.790, while for women's long jump there are six observations higher than 7.394. Following our assumption, having an observation Y above the endpoint of the distribution of X must be due to a positive value in the measurement error ε. In the context of long jump, it implies that the wind should have helped in delivering these personal bests. The website of IAAF provides the wind speed at the time when the long jump data was recorded: for eight out of the nine cases, the recorded wind speeds were positive. The only exception is the 8.87m set by Carl Lewis in the 1991 Tokyo World Championships, where the wind speed was recorded as -0.2m/s. Aside from

this exceptional case, the other eight cases support the view that the wind can be a potential factor causing measurement errors in the long jump performance.

Figure 6.4: Estimation of the extreme value indices
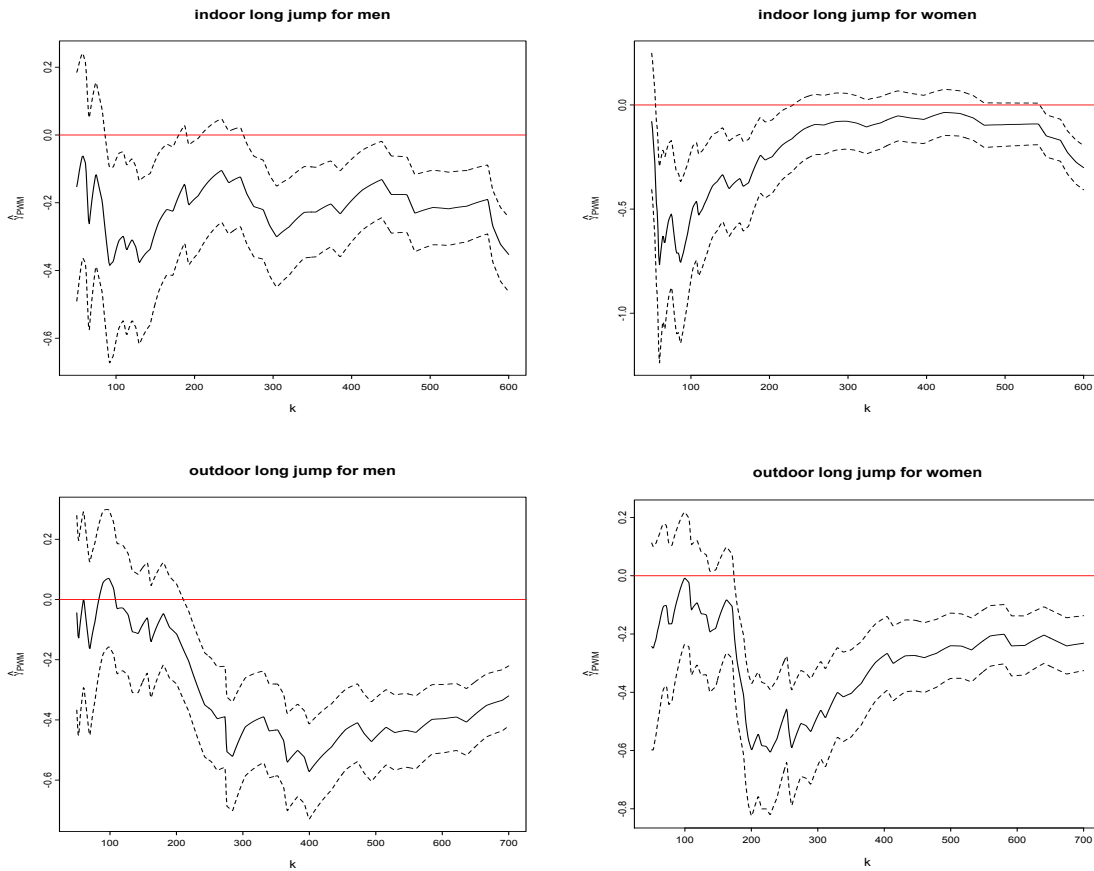
Note: The plots show the estimated extreme value indices with the corresponding 95% confidence intervals for various values of k using the PWM estimator.

6.5 Proofs

Proof of Proposition 6.2.1. Write

$$\overline{F}_Z(x) = \int_{-\infty}^0 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-t)^2}{2\sigma^2}} \overline{F}_X(\theta + t) dt = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \int_{-\infty}^0 \exp\left(\frac{tx}{\sigma^2} - \frac{t^2}{2\sigma^2}\right) \overline{F}_X(\theta + t) dt.$$

Figure 6.5: Estimation of the endpoints
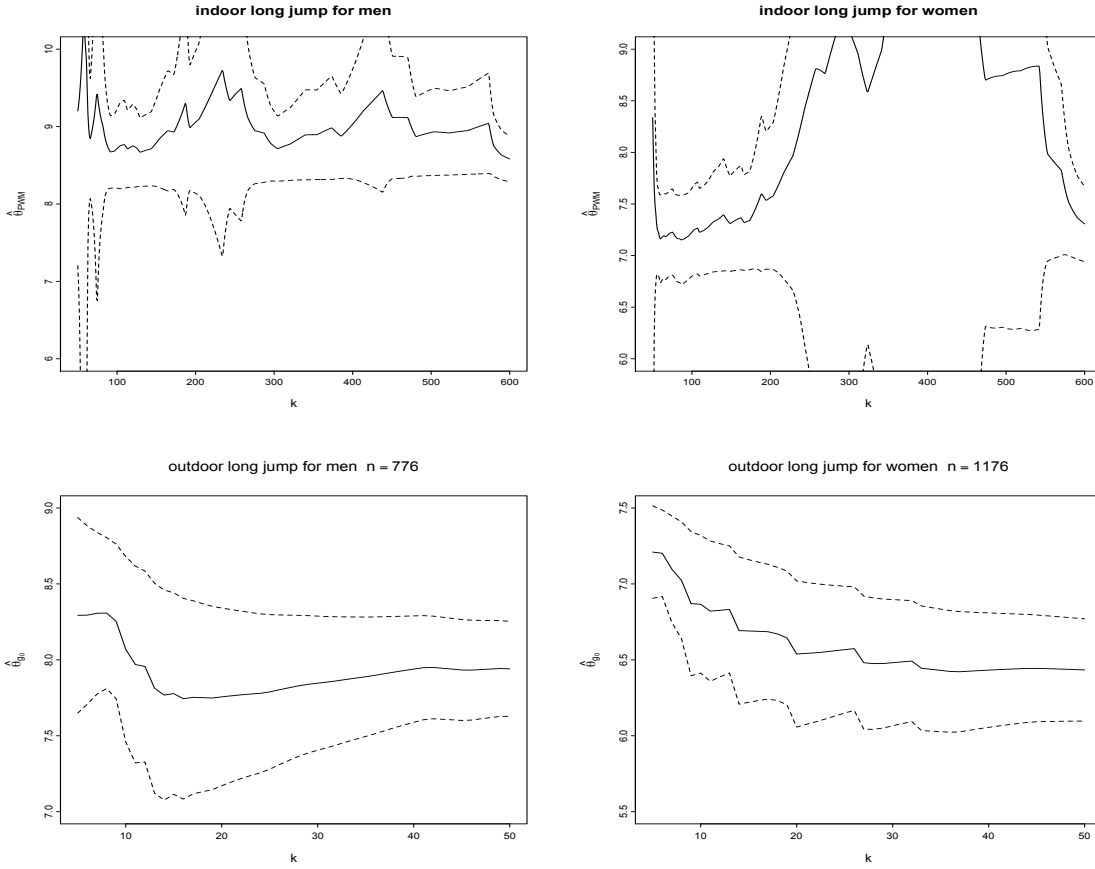
Note: The plots show the estimated endpoints with the corresponding 95% confidence intervals for various values of k. For the indoor long jump data, the endpoints are estimated by the PWM estimator $\hat{\theta}_{PWM}$. For the outdoor long jump data, the endpoints are estimated by the estimator $\hat{\theta}_{g_0}$.

Table 6.2: *Estimation results:* long jump data for men

	n	k	Point Estimate	95% C.I.
Outdoor				
$\hat{\gamma}_{PWM}$	–	180	-0.047	[-0.218, 0.123]
$\hat{\theta}_{g_0}$	776	8	8.308	[7.811, 8.805]
$\hat{\theta}_{g_0}$	3000	8	8.149	[7.528, 8.769]
Indoor				
$\hat{\gamma}_{PWM}$	–	150	-0.286	[-0.075, -0.497]
$\hat{\theta}_{PWM}$	–	150	8.790	[8.226, 9.354]

Note: The table shows the estimation results based on the long jump data for men. For both indoor and outdoor long jump, the estimated extreme value indices using the PWM estimator are presented. For the outdoor long jump data, the endpoints are estimated using the estimator $\hat{\theta}_{g_0}$ with setting the number of athletes n to either the current sample size or 3000. For the indoor long jump data, the endpoints are estimated using the PWM estimator. For all estimates, the corresponding 95% confidence intervals are provided.

Table 6.3: *Estimation results:* long jump data for women

	n	k	Point Estimate	95% C.I.
Outdoor				
$\hat{\gamma}_{PWM}$	–	160	-0.094	[-0.277, 0.090]
$\hat{\theta}_{g_0}$	1176	6	7.202	[6.917, 7.486]
$\hat{\theta}_{g_0}$	3000	6	7.152	[6.824, 7.481]
Indoor				
$\hat{\gamma}_{PWM}$	–	140	-0.335	[-0.560, -0.110]
$\hat{\theta}_{PWM}$	–	140	7.394	[6.850, 7.937]

Note: The table shows the estimation results based on the long jump data for women. For both indoor and outdoor long jump, the estimated extreme value indices using the PWM estimator are presented. For the outdoor long jump data, the endpoints are estimated using the estimator $\hat{\theta}_{g_0}$ with setting the number of athletes n to either the current sample size or 3000. For the indoor long jump data, the endpoints are estimated using the PWM estimator. For all estimates, the corresponding 95% confidence intervals are provided.

The proof is splitted into two parts. First, we show that as $x \to \infty$, the integral above will be dominated by only integrating in the neighborhood of zero, i.e., for some $\epsilon > 0$,

$$\overline{F}_Z(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \int_{-\epsilon}^{0} \exp\left(\frac{tx}{\sigma^2} - \frac{t^2}{2\sigma^2}\right) \overline{F}_X(\theta + t)dt \left[1 + O\left(x^{\alpha+1} e^{-\epsilon x/\sigma^2}\right)\right]. \quad (6.5.1)$$

Then, we will calculate the integral in the right hand side of (6.5.1).

We first handle the equation (6.5.1). Since

$$\int_{-\infty}^{-\epsilon} \exp\left(\frac{tx}{\sigma^2} - \frac{t^2}{2\sigma^2}\right) \overline{F}_X(\theta + t)dt \leq \exp\left(\frac{-\epsilon x}{\sigma^2}\right) \int_{-\infty}^{-\epsilon} \exp\left(-\frac{t^2}{2\sigma^2}\right) dt \leq \sqrt{2\pi}\sigma \exp\left(\frac{-\epsilon x}{\sigma^2}\right),$$

and

$$\int_{-\epsilon}^{0} \exp\left(\frac{tx}{\sigma^2} - \frac{t^2}{2\sigma^2}\right) \overline{F}_X(\theta + t)dt \geq e^{-\frac{\epsilon^2}{2\sigma^2}} \int_{-\epsilon}^{0} \exp\left(\frac{tx}{\sigma^2}\right) \overline{F}_X(\theta + t)dt,$$

(6.5.1) is proved by showing that, as $x \to \infty$,

$$\int_{-\epsilon}^{0} \exp\left(\frac{tx}{\sigma^2}\right) \overline{F}_X(\theta + t) \, dt = O\left(x^{-(\alpha+1)}\right). \quad (6.5.2)$$

Denote $G(t) := \overline{F}_X(\theta - t^{-1})$. Then as $t \to \infty$, $G(t) = Lt^{-\alpha} + dt^{-(\alpha+\beta)} + o\left(t^{-(\alpha+\beta)}\right)$. The left

hand side in (6.5.2) is then calculated as

$$\int_{-\epsilon}^{0} \exp\left(\frac{tx}{\sigma^2}\right) \overline{F}_X(\theta + t)\, dt = \frac{\sigma^2}{x} \int_{\frac{\sigma^2}{\epsilon x}}^{\infty} e^{-1/t} G\left(\frac{tx}{\sigma^2}\right) t^{-2}\, dt$$

$$= \frac{\sigma^2}{x} G\left(\frac{x}{\sigma^2}\right) \left[\int_{\frac{\sigma^2}{\epsilon x}}^{\infty} e^{-1/t} t^{-2} \left(\frac{G\left(tx/\sigma^2\right)}{G\left(x/\sigma^2\right)} - t^{-\alpha} \right) dt + \int_{\frac{\sigma^2}{\epsilon x}}^{\infty} e^{-1/t} t^{-(\alpha+2)}\, dt \right]$$

$$=: \frac{\sigma^2}{x} G\left(\frac{x}{\sigma^2}\right) [I_1 + I_2].$$

To deal with I_1, since the function G is regularly varying, i.e. $\lim_{t \to \infty} \frac{G(tu)}{G(t)} = u^{-\alpha}$, it follows from Proposition B.1.10 in De Haan and Ferreira (2007) that for any $\eta > 0$ there exists $x_0 = x_0(\eta)$ such that if $\frac{x}{\sigma^2} \geq x_0$ and $\frac{tx}{\sigma^2} \geq x_0$,

$$\left| \frac{G\left(tx/\sigma^2\right)}{G\left(x/\sigma^2\right)} - t^{-\alpha} \right| \leq \eta \max\left(t^{-\alpha+\eta}, t^{-\alpha-\eta} \right).$$

By choosing $\epsilon \leq 1/x_0$, we have that for all $t > \frac{\sigma^2}{\epsilon x}$ and sufficiently large x such that $x > \sigma^2 x_0$ the two required conditions hold. Therefore, we can apply the inequality above to obtain that

$$\int_{\frac{\sigma^2}{\epsilon x}}^{\infty} e^{-1/t} t^{-2} \left| \frac{G\left(tx/\sigma^2\right)}{G\left(x/\sigma^2\right)} - t^{-\alpha} \right| dt \leq \eta \int_{0}^{\infty} e^{-1/t} t^{-2-\alpha} \left(t^{\eta} + t^{-\eta} \right) dt < \infty.$$

Thus we can apply the dominated convergence theorem to get that $I_1 \to 0$ as $x \to \infty$.

By verifying that $I_2 \to \Gamma(\alpha + 1)$ as $x \to \infty$ and $\frac{\sigma^2}{x} G\left(\frac{x}{\sigma^2}\right) = O(x^{-\alpha-1})$, we proved (6.5.2) and consequently (6.5.1), for any $\epsilon \leq 1/x_0$.

Next we calculate the integral in the right hand of (6.5.1). We first perform a variable transformation such that the term in the exponential part $\frac{tx}{\sigma^2} - \frac{t^2}{2\sigma^2}$ is replaced by a new term $-s$, i.e. we define $s = \frac{t^2}{2\sigma^2} - \frac{tx}{\sigma^2}$. This transformation is one–to–one for $t \in [-\epsilon, 0]$ with the inverse transformation

$$t = x \left(1 - \sqrt{1 + \frac{2s\sigma^2}{x^2}} \right) =: -\frac{s\sigma^2}{x} \phi(s; x),$$

where

$$\phi(s;x) = 1 - \frac{s\sigma^2}{2x^2}(1 + o(1)), \tag{6.5.3}$$

as $x \to \infty$, and the $o(1)$ term is uniform for all $0 \le s \le \frac{\epsilon(x+\epsilon/2)}{\sigma^2}$. Write

$$\int_{-\epsilon}^{0} \exp\left(\frac{tx}{\sigma^2} - \frac{t^2}{2\sigma^2}\right) \overline{F}_X(\theta+t)dt = \frac{\sigma^2}{x} \int_0^{\frac{\epsilon(x+\epsilon/2)}{\sigma^2}} e^{-s}\left(1 + \frac{2\sigma^2}{x^2}s\right)^{-1/2} G\left(\frac{x}{\sigma^2}\frac{1}{s\phi(s;x)}\right) ds.$$

To calculate this integral, we again need a comparison between $G\left(\frac{x}{\sigma^2}\frac{1}{s\phi(s;x)}\right)$ and $G\left(\frac{x}{\sigma^2}\right)$. However, here we need a second–order expansion. Notice that the function G satisfies the second–order regular variation condition as

$$\lim_{w\to\infty} \frac{\frac{G(wu)}{G(w)} - u^{-\alpha}}{-\frac{d\beta}{L}w^{-\beta}} = u^{-\alpha}\frac{u^{-\beta}-1}{-\beta}.$$

By Theorem B.2.18 in De Haan and Ferreira (2007), for all $\eta^* > 0$, there exists $x_0^* = x_0^*(\eta^*)$, such that for $w, wu \ge x_0^*$,

$$\left|\frac{1}{A(w)}\left(\frac{G(wu)}{G(w)} - u^{-\alpha}\right) - u^{-\alpha}\frac{u^{-\beta}-1}{-\beta}\right| \le \eta^* \max\left(u^{-(\alpha+\beta)+\eta^*}, u^{-(\alpha+\beta)-\eta^*}\right),$$

where $A(w) := \frac{-d\beta}{L}w^{-\beta}$. We intend to apply this inequality with $w = \frac{x}{\sigma^2}$ and $wu = \frac{x}{\sigma^2}\frac{1}{s\phi(s;x)}$. For sufficiently large x, $w = \frac{x}{\sigma^2} > x_0^*$. For wu, notice that $\frac{x}{\sigma^2}\frac{1}{s\phi(s;x)} = -1/t \ge 1/\epsilon$. We get the required condition by choosing ϵ such that $\epsilon < 1/x_0^*$. Hence we obtain from the above inequality that

$$|Z(s;x)| := \left|\frac{1}{A\left(\frac{x}{\sigma^2}\right)}\left(\frac{G\left(\frac{x}{\sigma^2}\frac{1}{s\phi(s;x)}\right)}{G\left(\frac{x}{\sigma^2}\right)} - (s\phi(s;x))^\alpha\right) - (s\phi(s;x))^\alpha\frac{(s\phi(s;x))^\beta-1}{-\beta}\right|$$

$$\le \eta^* \max\left((s\phi(s;x))^{(\alpha+\beta)-\eta^*}, (s\phi(s;x))^{(\alpha+\beta)+\eta^*}\right). \tag{6.5.4}$$

This inequality allows us to write the integral in the right hand of (6.5.1) as

$$\frac{\sigma^2}{x} \int_0^{\frac{\epsilon(x+\epsilon/2)}{\sigma^2}} e^{-s}\left(1 + \frac{2\sigma^2}{x^2}s\right)^{-1/2} G\left(\frac{x}{\sigma^2}\frac{1}{s\phi(s;x)}\right) ds$$

$$= \frac{\sigma^2}{x} G\left(\frac{x}{\sigma^2}\right) \int_0^{\frac{\epsilon(x+\epsilon/2)}{\sigma^2}} e^{-s} \left(1 + \frac{2\sigma^2}{x^2} s\right)^{-1/2} (s\phi(s;x))^\alpha \, ds$$

$$+ \frac{\sigma^2}{x} A\left(\frac{x}{\sigma^2}\right) G\left(\frac{x}{\sigma^2}\right) \int_0^{\frac{\epsilon(x+\epsilon/2)}{\sigma^2}} e^{-s} \left(1 + \frac{2\sigma^2}{x^2} s\right)^{-1/2} (s\phi(s;x))^\alpha \frac{(s\phi(s;x))^\beta - 1}{-\beta} \, ds$$

$$+ \frac{\sigma^2}{x} A\left(\frac{x}{\sigma^2}\right) G\left(\frac{x}{\sigma^2}\right) \int_0^{\frac{\epsilon(x+\epsilon/2)}{\sigma^2}} e^{-s} \left(1 + \frac{2\sigma^2}{x^2} s\right)^{-1/2} Z(s;x), ds$$

$$=: J_1 + J_2 + J_3.$$

In all three terms, we need to deal with integrals in the form

$$I(x;\nu) := \int_0^{\frac{\epsilon(x+\epsilon/2)}{\sigma^2}} e^{-s} \left(1 + \frac{2\sigma^2}{x^2} s\right)^{-1/2} (s\phi(s;x))^\nu \, ds,$$

for $\nu > 0$. Notice that $\frac{2\sigma^2}{x^2} s \to 0$ and $\phi(s;x) \to 1$ as $x \to \infty$ hold uniformly for $s \le \frac{\epsilon(x+\epsilon/2)}{\sigma^2}$. By using dominance convergence theorem, we get that as $x \to \infty$,

$$I(x;\nu) \to \int_0^\infty e^{-s} s^\nu \, ds = \Gamma(\nu + 1). \tag{6.5.5}$$

Further write $\left(1 + \frac{2\sigma^2}{x^2} s\right)^{-1/2} = 1 - \frac{\sigma^2 s}{x^2}(1 + o(1))$, as $x \to \infty$, where the term $o(1)$ is again uniform for all $s \le \frac{\epsilon(x+\epsilon/2)}{\sigma^2}$. Then,

$$I(x;\nu) = \int_0^{\frac{\epsilon(x+\epsilon/2)}{\sigma^2}} e^{-s} (s\phi(s;x))^\nu \, ds - \frac{\sigma^2}{x^2} \int_0^{\frac{\epsilon(x+\epsilon/2)}{\sigma^2}} e^{-s} s (s\phi(s;x))^\nu \, ds(1 + o(1)) =: II_1 - II_2.$$

By applying the dominance convergence theorem again, we obtain that as $x \to \infty$, $x^2 II_2 \to \sigma^2 \Gamma(\nu + 2)$. For II_1, we use the expansion of $\phi(s;x)$ in (6.5.3) to get that

$$II_1 = \int_0^{\frac{\epsilon(x+\epsilon/2)}{\sigma^2}} e^{-s} s^\nu \, ds - \frac{\nu\sigma^2}{2x^2} \int_0^{\frac{\epsilon(x+\epsilon/2)}{\sigma^2}} e^{-s} s^{\nu+1} \, ds(1 + o(1))$$

$$= \Gamma(\nu + 1) + o(x^{-2}) - \frac{\nu\sigma^2}{2x^2} \Gamma(\nu + 2)(1 + o(1)) = \Gamma(\nu + 1) - \frac{\nu\sigma^2}{2x^2} \Gamma(\nu + 2)(1 + o(1)).$$

By combining the two terms we get a second–order expansion of $I(x; \mu)$: as $x \to \infty$,

$$I(x; \nu) = \Gamma(\nu + 1) - \frac{\sigma^2}{x^2} \left(\frac{\nu}{2} + 1 \right) \Gamma(\nu + 2)(1 + o(1)). \qquad (6.5.6)$$

By applying (6.5.6) with $\nu = \alpha$, we have that as $x \to \infty$,

$$J_1 = \frac{\sigma^2}{x} G \left(\frac{x}{\sigma^2} \right) \left(\Gamma(\alpha + 1) - \frac{\sigma^2}{x^2} \left(\frac{\alpha}{2} + 1 \right) \Gamma(\alpha + 2)(1 + o(1)) \right).$$

By applying (6.5.5) with $\nu = \alpha$ and $\nu = \alpha + \beta$, we get that as $x \to \infty$,

$$J_2 = \frac{\sigma^2}{x} A \left(\frac{x}{\sigma^2} \right) G \left(\frac{x}{\sigma^2} \right) \frac{\Gamma(\alpha + \beta + 1) - \Gamma(\alpha + 1)}{-\beta}.$$

Lastly, based on the inequalities (6.5.4) and (6.5.5), we get that $J_3 = o(J_2)$ as $x \to \infty$.

By combining all three terms, we obtain that as $x \to \infty$,

$$
\begin{aligned}
&\int_{-\epsilon}^{0} \exp \left(\frac{tx}{\sigma^2} - \frac{t^2}{2\sigma^2} \right) \overline{F}_X(\theta + t) dt \\
&= \frac{\sigma^2}{x} G \left(\frac{x}{\sigma^2} \right) \left(\Gamma(\alpha + 1) - \frac{\sigma^2}{x^2} \left(\frac{\alpha}{2} + 1 \right) \Gamma(\alpha + 2)(1 + o(1)) \right. \\
&\qquad \left. + A \left(\frac{x}{\sigma^2} \right) \frac{\Gamma(\alpha + \beta + 1) - \Gamma(\alpha + 1)}{-\beta} (1 + o(1)) \right) \\
&= \frac{\sigma^2}{x} G \left(\frac{x}{\sigma^2} \right) \left(\Gamma(\alpha + 1) + \frac{\sqrt{2\pi} c_2}{L \sigma^{2\alpha + 1}} x^{-\beta'} (1 + o(1)) - \frac{d}{L} \Gamma(\alpha + 1) \sigma^{2\beta} x^{-\beta} \right) \\
&= L x^{-\alpha - 1} \sigma^{2\alpha + 2} \left(1 + \frac{d}{L} \sigma^{2\beta} x^{-\beta} (1 + o(1)) \right) \times \\
&\qquad \left(\Gamma(\alpha + 1) + \frac{\sqrt{2\pi} c_2}{L \sigma^{2\alpha + 1}} x^{-\beta'} (1 + o(1)) - \frac{d}{L} \Gamma(\alpha + 1) \sigma^{2\beta} x^{-\beta} \right) \\
&= \sqrt{2\pi} \sigma x^{-\alpha - 1} (c_1 + c_2 x^{-\beta'} (1 + o(1))),
\end{aligned}
$$

where c_1, β' and c_2 are defined in the proposition.

Finally the proposition is proved by substituting this relation to (6.5.1). ∎

Next we prove Theorem 6.2.2. Write the estimator $\hat{\sigma}_g$ in (6.2.2) as

$$\frac{\hat{\sigma}_g}{\sigma} = \log(n/k) \int_{1/k}^{1} g([ks]/k) \frac{Z_{n-[ks],n} - Z_{n-k,n}}{\sigma \sqrt{2\log(n/k)}} \, ds.$$

We firstly establish the asymptotic properties of tail quantile process $\{Z_{n-[ks],n}, \ k^{-1} \leq s \leq 1\}$ as follows.

Proposition 6.5.1. *Assume the same conditions as in Proposition 6.2.1. Then there exists a sequence of standard Brownian motions $\{W_n(s) : s \geq 0\}$ such that as $n \to \infty$,*

$$\frac{Z_{n-[ks],n}}{\sigma\sqrt{2\log(n/k)}} = \psi_{1,n} + \psi_{2,n}(s) + \frac{k^{-1/2}\left(s^{-1}W_n(s) + s^{-1/2-\delta}o_p(1)\right)}{2\log(n/k)} + \frac{q(n/k)}{\log(n/k)}(1 + o_p(1))$$

where

$$\psi_{1,n} = 1 - \frac{(\alpha+1)\log\log(n/k)}{4\log(n/k)} + \frac{c_3}{\log(n/k)},$$

$$c_3 = 2^{-1}\left(\log c_1 - (\alpha+1)\log(\sigma\sqrt{2})\right),$$

$$q(n/k) = \begin{cases} \frac{\sigma^{-\beta}c_2}{2^{1+\beta/2}c_1}\left(\log(n/k)\right)^{-\frac{\beta}{2}} & \beta < 2 \\ -\frac{(\alpha+1)^2}{32}\left(\log(n/k)\right)^{-1}\left(\log\log(n/k)\right)^2 & \beta \geq 2 \end{cases},$$

$$\psi_{2,n}(s) = \frac{-\log s}{2\log(n/k)}\left(1 + \frac{(\alpha+1)\log\log(n/k)}{4\log(n/k)}(1 + o(1))\right),$$

and all terms $o_p(1)$ are uniform for $s \in [k^{-1}, 1]$.

Proof. Denote $U = (-\log \overline{F}_Z)^{\leftarrow}$, where '$\leftarrow$' denotes the left continuous inverse function. Write $Z_i = U(E_i)$ where $\{E_i\}_{i=1}^{n}$ is a sample of *i.i.d.* standard exponential distributed random variables. Let $E_{1,n} \leq E_{2,n} \leq \cdots \leq E_{n,n}$ be the corresponding order statistics. To obtain the asymptotic properties of $\{Z_{n-[ks],n}, \ k^{-1} \leq s \leq 1\}$, we derive the expansion of $U(t)$ and the asymptotic properties of $\{E_{n-[ks],n}, \ k^{-1} \leq s \leq 1\}$.

From the tail expansion of $-\log \bar{F}_Z$ in Proposition 6.2.1, we get that as $t \to \infty$,

$$t = \frac{1}{2\sigma^2}(U(t))^2 + (\alpha+1)\log U(t) - \log c_1 - \frac{c_2}{c_1}(U(t))^{-\beta'}(1+o(1)). \tag{6.5.7}$$

Since $U(t) \to \infty$ as $t \to \infty$, it implies that $U(t) \sim \sigma\sqrt{2t}$. Write $U(t) = \sigma\sqrt{2t} + r(t)$, where $r(t) = o(\sqrt{t})$. We intend to obtain further explicit expression for the $r(t)$ term. Substituting $U(t)$ in (6.5.7) by $\sigma\sqrt{2t} + r(t)$ yields that, as $t \to \infty$,

$$0 = r^2(t) + 2\sigma\sqrt{2t}\, r(t) + \sigma^2(\alpha+1)\log t + 2\sigma^2\left((\alpha+1)\log(\sigma\sqrt{2}) - \log c_1\right)$$

$$+ 2\sigma^2(\alpha+1)\log\left(1 + \frac{r(t)}{\sigma\sqrt{2t}}\right) - \frac{2\sigma^2 c_2}{c_1}\left(\sigma\sqrt{2t}\right)^{-\beta'}(1+o(1)).$$

Since $\frac{r(t)}{\sqrt{t}} \to 0$ as $t \to \infty$, we can solve this equation to obtain that

$$r(t) = -\frac{\sigma(\alpha+1)}{2\sqrt{2}}t^{-1/2}\log t + \frac{\sigma}{\sqrt{2}}\left(\log c_1 - (\alpha+1)\log(\sigma\sqrt{2})\right)t^{-1/2} + t^{-1/2}\tilde{q}(t),$$

where $\tilde{q}(t) = o(1)$ as $t \to \infty$. By reiterating this procedure, we eventually obtain that as $t \to \infty$,

$$U(t) = \sigma\sqrt{2t} - \frac{\sigma(\alpha+1)}{2\sqrt{2}}t^{-1/2}\log t + \sigma\sqrt{2}c_3 t^{-1/2} + t^{-1/2}\tilde{q}(t)(1+o(1)), \tag{6.5.8}$$

where

$$\tilde{q}(t) = \begin{cases} \frac{\sigma^{1-\beta}c_2}{(\sqrt{2})^{1+\beta}c_1}t^{-\frac{\beta}{2}} & \beta < 2 \\ -\frac{\sigma(\alpha+1)^2}{16\sqrt{2}}t^{-1}\log^2 t & \beta \geq 2. \end{cases}$$

Note that $\tilde{q}$ is related to the q function defined in this proposition by $q(t) = \frac{\tilde{q}(\log t)}{\sqrt{2}\sigma}$.

Next we derive the asymptotic properties of $\{E_{n-[ks],n}, k^{-1} \leq s \leq 1\}$ from Theorem 2.4.2 in De Haan and Ferreira (2007) as follows. For any $\delta > 0$, there exists a sequence of standard Brownian motions $\{\{W_n(s)\}_{s\geq 0}\}$ such that as $n \to \infty$,

$$\sup_{k^{-1}\leq s\leq 1} s^{1/2+\delta}\left|\sqrt{k}(E_{n-[ks],n} - \log(n/k) + \log s) - s^{-1}W_n(s)\right| \xrightarrow{p} 0.$$

146

When plugging the asymptotic expansion of $E_{n-[ks],n}$ into $Z_{n-[ks],n} = U(E_{n-[ks],n})$, we observe that the asymptotic property of $\frac{Z_{n-[ks],n}}{\sigma\sqrt{2\log(n/k)}}$ is mainly driven by $\left(\frac{E_{n-[ks],n}}{\log(n/k)}\right)^{1/2}$. In order to make a smooth substitution, we first derive the asymptotic behavior of $\left(\frac{E_{n-[ks],n}}{\log(n/k)}\right)^{\gamma}$ for any given $\gamma \in \Re$.

Write

$$\left(\frac{E_{n-[ks],n}}{\log(n/k)}\right)^{\gamma} = \left[1 + \frac{\log s^{-1}}{\log(n/k)} + \frac{s^{-1}W_n(s) + o_p(1)s^{-1/2-\delta}}{\sqrt{k}\log(n/k)}\right]^{\gamma}$$

$$= 1 + \gamma\frac{\log s^{-1} + k^{-1/2}\left(s^{-1}W_n(s) + o_p(1)s^{-1/2-\delta}\right)}{\log(n/k)}$$

$$+ \theta\left(\frac{\log s^{-1} + k^{-1/2}\left(s^{-1}W_n(s) + o_p(1)s^{-1/2-\delta}\right)}{\log(n/k)}\right)^{2}$$

where $\theta = \frac{1}{2}\frac{\partial^2}{\partial x^2}(1+x)^{\gamma}|_{x=\xi}$ for some ξ between 0 and $\frac{\log s^{-1}+k^{-1/2}\left(s^{-1}W_n(s)+o_p(1)s^{-1/2-\delta}\right)}{\log(n/k)}$. Since as $k \to \infty$, $\frac{\log s^{-1}+k^{-1/2}\left(s^{-1}W_n(s)+o_p(1)s^{-1/2-\delta}\right)}{\log(n/k)} \xrightarrow{p} 0$ uniformly for all $1/k \le s \le 1$ and $\frac{\partial^2}{\partial x^2}(1+x)^{\gamma}$ is bounded in the neighborhood of zero, we get that with probability tending to 1, $|\theta|$ is bounded uniformly for all $1/k \le s \le 1$. By verifying that the quadratic term

$$\left(\frac{\log s^{-1} + k^{-1/2}\left(s^{-1}W_n(s) + o_p(1)s^{-1/2-\delta}\right)}{\log(n/k)}\right)^{2}$$

can be uniformly written as $\frac{1}{\sqrt{k}\log(n/k)}s^{-1/2-\delta}o_p(1)$, we get that as $n \to \infty$

$$\left(\frac{E_{n-[ks],n}}{\log(n/k)}\right)^{\gamma} = 1 + \gamma\frac{\log s^{-1} + k^{-1/2}\left(s^{-1}W_n(s) + o_p(1)s^{-1/2-\delta}\right)}{\log(n/k)},$$

where the $o_p(1)$ term is uniformly for all $1/k \le s \le 1$. For $\gamma = 0$, the relation should be read as

$$\log E_{n-[ks],n} - \log\log(n/k) = \frac{\log s^{-1} + k^{-1/2}\left(s^{-1}W_n(s) + o_p(1)s^{-1/2-\delta}\right)}{\log(n/k)}.$$

Finally, by plugging $E_{n-[ks],n}$ into $Z_{n-[ks],n} = U(E_{n-[ks],n})$, while using the expansion of U in (6.5.8) and the asymptotic expansion of $\left(\frac{E_{n-[ks],n}}{\log(n/k)}\right)^{\gamma}$ for $\gamma = 1/2, -1/2, 0$ and $-(1+\beta)/2$, we obtain the result of Proposition 6.5.1. ∎

Remark 6.5.2. *In the expansion of the tail quantile process, $\psi_{1,n}$ is a deterministic term not depending on s, $\psi_{2,n}$ is a deterministic term depending on s, the third component gives a random term, and finally $\frac{q(n/k)}{\log(n/k)}(1+o_p(1))$ has a uniform approximation independent of s. It will turn out to be clear in the proof of Theorem 6.2.2 that such a detailed expansion is necessary for achieving the intended asymptotic results.*

We use Proposition 6.5.1 to prove the asymptotic normality of $\hat{\sigma}_g$.

Proof of Theorem 6.2.2. Write

$$\frac{\hat{\sigma}_g}{\sigma} = \frac{\log(n/k)}{k} g(1/k) \frac{Z_{n-1,n} - Z_{n-k,n}}{\sigma\sqrt{2\log(n/k)}} + \log(n/k) \int_{2/k}^{1} g([ks]/k) \frac{Z_{n-[ks],n} - Z_{n-k,n}}{\sigma\sqrt{2\log(n/k)}} \, ds$$

$$=: I_1 + I_2.$$

From Proposition 6.5.1, we get that, as $n \to \infty$,

$$\frac{Z_{n-[ks],n} - Z_{n-k,n}}{\sigma\sqrt{2\log(n/k)}} = \frac{-\log s + k^{-1/2}\left(s^{-1}W_n(s) - W_n(1) + o_p(1)s^{-1/2-\delta}\right)}{2\log(n/k)}$$

$$- \frac{(\alpha+1)\log\log(n/k)}{4\log(n/k)} \cdot \frac{\log s(1+o(1))}{2\log(n/k)} + \frac{q(n/k)}{\log(n/k)}o_p(1), \quad (6.5.9)$$

holds uniformly for $1/k \leq s \leq 1$.

First, for I_1, replacing s by k^{-1} in (6.5.9), yields that as $n \to \infty$,

$$\sqrt{k}I_1$$
$$= \frac{\log(n/k)}{k^{1/2}} g(1/k) \left(\frac{\log k + k^{1/2}W_n(1/k) + k^\delta o_p(1)}{2\log(n/k)} + \frac{\log k \log\log(n/k)}{(\log(n/k))^2}O(1) + \frac{q(n/k)}{\log(n/k)}o_p(1) \right).$$

Using the modulus of continuity for $W_n(1/k)$ and note that from the condition (6.2.5), $g(1/k) = o(1)k^{1/2-\epsilon_0}$ for any $\delta < \epsilon_0$, as $n \to \infty$, we then get from the above equation that

$$\sqrt{k}I_1 = k^{-1/2}g(1/k)k^\delta O_p(1) = k^{\delta-\epsilon_0}o_p(1) = o_p(1).$$

148

Next, we deal with I_2. By multiplying both sides of (6.5.9) with $\log(n/k)g(s)$ and taking an integral on the interval $[2/k, 1]$, we get that as $n \to \infty$,

$$
\begin{aligned}
I_2^* &= \log(n/k) \int_{2/k}^1 g(s) \frac{Z_{n-[ks],n} - Z_{n-k,n}}{\sigma\sqrt{2\log(n/k)}} \, ds \\
&= \frac{1}{2} \int_{2/k}^1 g(s)(-\log s) \, ds \\
&\quad + \frac{1}{2\sqrt{k}} \left(\int_{2/k}^1 g(s)\left(s^{-1}W_n(s) - W_n(1)\right) ds + o_p(1) \int_{2/k}^1 g(s)s^{-1/2-\delta} \, ds \right) \\
&\quad + \frac{(\alpha+1)\log\log(n/k)}{4\log(n/k)} \frac{1}{2} \int_{2/k}^1 g(s)(-\log s) \, ds(1+o(1)) + q(n/k)o_p(1) \int_{2/k}^1 g(s) \, ds. \\
&= 1 + \frac{1}{2\sqrt{k}} \left(\int_0^1 g(s)\left(s^{-1}W_n(s) - W_n(1)\right) ds + o_p(1) \right) \\
&\quad + \frac{(\alpha+1)\log\log(n/k)}{4\log(n/k)}(1+o(1)) + q(n/k)o_p(1).
\end{aligned}
$$

To obtain the last equality, we used the condition (6.2.5) to derive the following facts. First, both $g(s)s^{-1/2-\delta}$ and $g(s)$ are integrable on [0,1] for any $\delta < \epsilon_0$. Second, as $n \to \infty$,

$$
\left| \int_0^{2/k} g(s)\left(s^{-1}W_n(s) - W_n(1)\right) ds \right| \le \int_0^{2/k} s^{-1/2+\epsilon_0} \left| s^{-1}W_n(s) - W_n(1) \right| ds = o_p(1).
$$

Lastly, as $k \to \infty$,

$$
\begin{aligned}
\int_0^{2/k} g(s)(-\log s) \, ds &\le \int_0^{2/k} s^{-1/2+\epsilon_0}(-\log s) \, ds = \int_{+\infty}^{\log k - \log 2} e^{(1/2-\epsilon_0)t} t \, de^{-t} \\
&\le \sqrt{\frac{2}{k}} \int_{\log k - \log 2}^{+\infty} e^{-\epsilon_0 t} t \, dt = \frac{1}{\sqrt{k}} o(1).
\end{aligned}
$$

Finally, the condition (6.2.4) implies that $\sqrt{k}\log\log(n/k)/\log(n/k) \to 0$ and $\sqrt{k}q(n/k)$ is bounded as $n \to \infty$, which leads to the asymptotic property of I_2^*: as $n \to \infty$,

$$
\sqrt{k}(I_2^* - 1) - \frac{1}{2} \int_0^1 g(s)\left(s^{-1}W_n(s) - W_n(1)\right) ds \xrightarrow{p} 0.
$$

We remark that the difference between I_2 and I_2^* is of an order $k^{-1/2}o_p(1)$. This is shown by using the condition (6.2.6) as follows. Write

$$\sqrt{k}\,|I_2 - I_2^*|$$

$$\leq \sqrt{k}\log(n/k)\int_{2/k}^1 |g([ks]/k) - g(s)|\,\frac{Z_{n-[ks],n} - Z_{n-k,n}}{\sigma\sqrt{2\log(n/k)}}\,ds$$

$$\leq \sqrt{k}\int_{2/k}^1 |g([ks]/k) - g(s)|\left(\log s^{-1} + k^{-1/2}O_p(1)s^{-1/2-\delta}\right)ds$$

$$\leq \sqrt{k}\int_{2/k}^1 \left|\frac{g([ks]/k)}{\log([ks]/k)} - \frac{g(s)}{\log s}\right||\log s|\left(\log s^{-1} + k^{-1/2}O_p(1)s^{-1/2-\delta}\right)ds$$

$$\quad + \sqrt{k}\int_{2/k}^1 |\log([ks]/k) - \log s|\left|\frac{g([ks]/k)}{\log([ks]/k)}\right|\left(\log s^{-1} + k^{-1/2}O_p(1)s^{-1/2-\delta}\right)ds$$

$$\leq \sqrt{k}\sup_{|s-t|\leq 1/k,\,1/k\leq s,t\leq 1}\left|\frac{g(s)}{\log s} - \frac{g(t)}{\log t}\right|\int_{2/k}^1 (-\log s)\left(\log s^{-1} + k^{-1/2}O_p(1)s^{-1/2-\delta}\right)ds$$

$$\quad + \sqrt{k}\int_{2/k}^1 |\log([ks]/ks)|\left|\frac{g([ks]/k)}{\log([ks]/k)}\right|(-\log s)\,ds$$

$$\quad + O_p(1)\int_{2/k}^1 |\log([ks]/ks)|\left|\frac{g([ks]/k)}{\log([ks]/k)}\right|s^{-1/2-\delta}\,ds$$

$$=: I_{21} + I_{22} + I_{23}.$$

The condition (6.2.6) implies that as $n \to \infty$, $I_{21} = o_p(1)$.

Next, for I_{22}, notice that $|\log([ks]/ks)| = \left|\log\left(1 - \frac{\{ks\}}{ks}\right)\right| \leq \frac{c^*}{ks} < 1$ for $s \geq 2/k$. In addition, the condition (6.2.5) implies that there exists c_1^* such that $g(s) < c_1^* s^{-1/2+\epsilon_0}$ for all $s \in (0,1]$. By applying the condition (6.2.6) again, we get that as $n \to \infty$.

$$I_{22} \leq \sqrt{k}\int_{2/k}^1 |\log([ks]/ks)|\,g(s)\,ds + o(1)\int_{2/k}^1 |\log([ks]/ks)|\,(-\log s)ds$$

$$\leq c_1^*\int_{2/k}^1 \frac{c^*}{k^{1/2}s}s^{-1/2+\epsilon_0}\,ds + o(1)\int_{2/k}^1 \frac{c^*}{ks}(-\log s)ds$$

$$\leq \frac{c_1^* c^*}{2^{1/2-\epsilon_0/2}}\frac{1}{k^{\epsilon_0/2}}\int_{2/k}^1 s^{-1+\epsilon_0/2}ds + o(1)\frac{c^*}{2} \to 0.$$

Lastly, for I_{23}, notice that for all $2/k \leq s < 1$, $|\log([ks]/k)| \geq \max(-\log s, \log(k/(k-1)))$

150

and $\frac{s}{2} < [ks]/k < s$. We have $g([ks]/k) < c_1^*([ks]/k)^{-1/2+\epsilon_0} < c_2^* s^{-1/2+\epsilon_0}$. Hence, as $n \to \infty$

$$I_{23} \le O_p(1) \int_{2/k}^1 \frac{c^*}{ks} c_2^* s^{-1+\epsilon_0-\delta} \frac{1}{\max(-\log s, \log(k/(k-1)))} ds$$

$$\le O_p(1) \left(\int_{2/k}^{1-1/k} \frac{c^*}{ks} c_2^* s^{-1+\epsilon_0-\delta} \frac{1}{-\log s} ds + \frac{1}{\log(k/(k-1))} \int_{1-1/k}^1 \frac{c^*}{ks} c_2^* s^{-1+\epsilon_0-\delta} ds \right).$$

Since as $n \to \infty$, $k \log(k/(k-1)) \to 1$, $\int_{1-1/k}^1 s^{-2+\epsilon_0-\delta} ds \to 0$ and

$$\lim_{k\to\infty} \frac{\int_{2/k}^{1-1/k} \frac{s^{-2+\epsilon_0-\delta}}{-\log s} ds}{k} = \lim_{k\to\infty} \left[\frac{(1-1/k)^{-2+\epsilon_0-\delta}}{-\log(1-1/k)} \frac{1}{k^2} + \frac{(2/k)^{-2+\epsilon_0-\delta}}{-\log(2/k)} \frac{2}{k^2} \right] = 0,$$

we obtain that $I_{23} \to 0$ as $n \to \infty$. Combining the three terms, we have shown that as $n \to \infty$,
$\sqrt{k}(I_2 - I_2^*) = o_p(1)$, which leads to

$$\sqrt{k}(I_2 - 1) - \frac{1}{2} \int_0^1 g(s) \left(\frac{1}{s} W_n(s) - W_n(1) \right) ds \xrightarrow{p} 0.$$

The theorem is thus proved by combining the two parts I_1 and I_2 and further calculating the asymptotic variance. ∎

Proof of Corollary 6.2.3. Write

$$\frac{\sqrt{k}}{\sqrt{\log n}} \left(\hat{\theta}_g - \theta \right) = \sigma\sqrt{2k} \left(\frac{Z_{n,n}}{\sigma\sqrt{2\log n}} - 1 \right) - \sigma\sqrt{2k} \left(\frac{\hat{\sigma}_g}{\sigma} - 1 \right)$$

$$=: I_1 - I_2.$$

Theorem 6.2.2 gives the asymptotic property of I_2. Hence we only need to show that $I_1 \xrightarrow{p} 0$, as $n \to \infty$. Since as $n \to \infty$, $E_{n,n} - \log n \xrightarrow{d} \exp\{-e^{-x}\}$, we get that for $\gamma \ne 0$,

$$\left(\frac{E_{n,n}}{\log(n/k)} \right)^\gamma = \left(\frac{\log n + O_p(1)}{\log(n/k)} \right)^\gamma = 1 + \gamma \frac{\log k}{\log(n/k)} (1 + o_p(1)).$$

151

Plugging this expansion to the U function given in (6.5.8) yields that

$$\frac{Z_{n,n}}{\sigma\sqrt{2\log(n/k)}} = \frac{U(E_{n,n})}{\sigma\sqrt{2\log(n/k)}} = \psi_{1,n} + \frac{\log k}{2\log(n/k)}(1 + o_p(1)),$$

as $n \to \infty$. The corollary is proved by verifying that

$$\sqrt{k}(\psi_{1,n} - 1) \to 0 \quad \text{and} \quad \sqrt{k}\log k/\log(n/k) \to 0,$$

as $n \to \infty$, which are implied by the condition (6.2.6). ∎

6.6 Conclusions

In this chapter, we consider the estimation of the finite endpoint θ of a distribution function F_X. Instead of having observations drawn from F_X, we only observe a contaminated sample $Y_i = X_i + \varepsilon_i$, $i = 1, 2, \cdots, n$, where X_i follows the distribution F_X and ε_i is a measurement error following $N(0, \sigma^2)$ with $\sigma > 0$.

We start with proposing a class of estimators $\hat{\sigma}_g$ for σ, depending on an appropriate weighing function g on $(0, 1]$. Then we suggest an estimator $\hat{\theta}_g = \max_{1 \le i \le n} Y_i - \hat{\sigma}_g\sqrt{2\log n}$ for estimating the endpoint. Both the estimators $\hat{\sigma}_g$ and $\hat{\theta}_g$ possess asymptotic normality.

We demonstrate, by extensive simulation studies, the superior performance of our suggested estimator to that of other three endpoint estimators designed for the case of no measurement errors. In addition, we apply our suggested estimator to resolve the difficulties encountered by Einmahl and Magnus (2008): the estimated extreme value index is close to, and not significantly different from, zero. By assuming the presence of measurement errors stemming from the wind, we apply our suggested endpoint estimator to the outdoor long jump data. The results are comparable with applying the PWM endpoint estimator to the indoor long jump data, for which the impact of wind is negligible and the tail index is estimated to be negative.

REFERENCES

Aarssen, Karin and Laurens De Haan (1994). "On the maximal life span of humans". In: *Mathematical Population Studies* 4(4), pp. 259–281.

Ahn, Jae Youn and Nariankadu D Shyamalkumar (2014). "Asymptotic theory for the empirical Haezendonck–Goovaerts risk measure". In: *Insurance: Mathematics and Economics* 55, pp. 78–90.

Alves, Isabel Fraga and Cláudia Neves (2014). "Estimation of the finite right endpoint in the Gumbel domain". In: *Statistica Sinica*, pp. 1811–1835.

Artzner, Philippe et al. (1999). "Coherent measures of risk". In: *Mathematical finance* 9(3), pp. 203–228.

Baysal, Rafet Evren and Jeremy Staum (2008). "Empirical likelihood for value-at-risk and expected shortfall". In: *The journal of risk* 11(1), p. 3.

Bellini, Fabio and Emanuela Rosazza Gianin (2008a). "On Haezendonck risk measures". In: *Journal of Banking & Finance* 32(6), pp. 986–994.

Bellini, Fabio and Emanuela Rosazza Gianin (2008b). "Optimal portfolios with Haezendonck risk measures". In: *Statistics & Decisions International mathematical journal for stochastic methods and models* 26(2), pp. 89–108.

Bellini, Fabio and Emanuela Rosazza Gianin (2012). "Haezendonck–Goovaerts risk measures and Orlicz quantiles". In: *Insurance: Mathematics and Economics* 51(1), pp. 107–114.

Berghaus, Betina, Axel Bücher, Stanislav Volgushev, et al. (2017). "Weak convergence of the empirical copula process with respect to weighted metrics". In: *Bernoulli* 23(1), pp. 743–772.

Bingham, NH, CM Goldie, and JL Teugels (1987). "Regular Variation (Encyclopedia of Mathematics and its Applications, Vol. 27)". In: *CUP, Cambridge*.

Cao, Xi-Ren and Xiangwei Wan (2017). "Sensitivity analysis of nonlinear behavior with distorted probability". In: *Mathematical Finance* 27(1), pp. 115–150.

Carroll, Raymond J and Peter Hall (1988). "Optimal rates of convergence for deconvolving a density". In: *Journal of the American Statistical Association* 83(404), pp. 1184–1186.

Cazals, Catherine, Jean-Pierre Florens, and Léopold Simar (2002). "Nonparametric frontier estimation: a robust approach". In: *Journal of econometrics* 106(1), pp. 1–25.

Chan, Ngai Hang et al. (2007). "Interval estimation of value-at-risk based on GARCH models with heavy-tailed innovations". In: *Journal of Econometrics* 137(2), pp. 556–576.

Chen, Jian, Liang Peng, and Yichuan Zhao (2009). "Empirical likelihood based confidence intervals for copulas". In: *Journal of Multivariate Analysis* 100(1), pp. 137–151.

Chen, Song Xi and Peter Hall (1993). "Smoothed empirical likelihood confidence intervals for quantiles". In: *The Annals of Statistics*, pp. 1166–1181.

Chen, Xiaohong and Yanqin Fan (2006). "Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification". In: *Journal of econometrics* 135(1-2), pp. 125–154.

Cheung, Ka Chun and Ambrose Lo (2013). "General lower bounds on convex functionals of aggregate sums". In: *Insurance: Mathematics and Economics* 53(3), pp. 884–896.

Csorgo, Miklos et al. (1986). "Weighted empirical and quantile processes". In: *The Annals of Probability*, pp. 31–85.

Danielsson, Jon et al. (2001). "Using a bootstrap method to choose the sample fraction in tail index estimation". In: *Journal of Multivariate analysis* 76(2), pp. 226–248.

De Haan, Laurens and Ana Ferreira (2007). *Extreme value theory: an introduction*. Springer Science & Business Media.

Doukhan, Paul (1995). "Mixing". In: *Mixing*. Springer, pp. 15–23.

Draisma, Gerrit et al. (1999). "A bootstrap-based method to achieve optimality in estimating the extreme-value index". In: *Extremes* 2(4), pp. 367–404.

Drees, Holger and Laurens de Haan (1999). "Conditions for quantile process approximations". In: *Stochastic Models* 15(3), pp. 485–502.

Drees, Holger and Edgar Kaufmann (1998). "Selecting the optimal sample fraction in univariate extreme value estimation". In: *Stochastic Processes and their Applications* 75(2), pp. 149–172.

Einmahl, John HJ (1992). "Limit theorems for tail processes with application to intermediate quantile estimation". In: *Journal of Statistical Planning and Inference* 32(1), pp. 137–145.

Einmahl, John HJ and Jan R Magnus (2008). "Records in athletics through extreme-value theory". In: *Journal of the American Statistical Association* 103(484), pp. 1382–1391.

Fischer, Tom (2003). "Risk capital allocation by coherent risk measures based on one-sided moments". In: *insurance: Mathematics and Economics* 32(1), pp. 135–146.

Fisher, Ronald Aylmer and Leonard Henry Caleb Tippett (1928). "Limiting forms of the frequency distribution of the largest or smallest member of a sample". In: *Mathematical Proceedings of the Cambridge Philosophical Society*. Vol. 24. 2. Cambridge University Press, pp. 180–190.

Francq, Christian, Jean-Michel Zakoian, et al. (2004). "Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes". In: *Bernoulli* 10(4), pp. 605–637.

Frees, Edward W (2009). *Regression modeling with actuarial and financial applications*. Cambridge University Press.

Fu, Michael C, L Jeff Hong, and Jian-Qiang Hu (2009). "Conditional Monte Carlo estimation of quantile sensitivities". In: *Management Science* 55(12), pp. 2019–2027.

Girard, Stéphane, Armelle Guillou, and Gilles Stupfler (2012). "Estimating an endpoint with high-order moments". In: *test* 21(4), pp. 697–729.

Glasserman, Paul (2005). "Measuring marginal risk contributions in credit portfolios". In:

Gnedenko, Boris (1943). "Sur la distribution limite du terme maximum d'une serie aleatoire". In: *Annals of mathematics*, pp. 423–453.

Goldenshluger, A and Alexandre Tsybakov (2004). "Estimating the endpoint of a distribution in the presence of additive observation errors". In: *Statistics & probability letters* 68(1), pp. 39–49.

Goovaerts, Marc et al. (2012). "On the interplay between distortion, mean value and Haezendonck–Goovaerts risk measures". In: *Insurance: Mathematics and Economics* 51(1), pp. 10–18.

Goovaerts, Marc J et al. (2004). "Some new classes of consistent risk measures". In: *Insurance: Mathematics and Economics* 34(3), pp. 505–516.

Gourieroux, Christian, Jean Pierre Laurent, and Olivier Scaillet (2000). "Sensitivity analysis of values at risk". In: *Journal of empirical finance* 7(3-4), pp. 225–245.

Haan, Laurens de and Ulrich Stadtmüller (1996). "Generalized regular variation of second order". In: *Journal of the Australian Mathematical Society* 61(3), pp. 381–395.

Haezendonck, J and Marc Goovaerts (1982). "A new premium calculation principle based on Orlicz norms". In: *Insurance: Mathematics and Economics* 1(1), pp. 41–53.

Hall, Peter (1982). "On estimating the endpoint of a distribution". In: *The Annals of Statistics*, pp. 556–568.

Hall, Peter and Léopold Simar (2002). "Estimating a changepoint, boundary, or frontier in the presence of observation error". In: *Journal of the American statistical Association* 97(458), pp. 523–534.

Hall, Peter, Julian Z Wang, et al. (1999). "Estimating the end-point of a probability distribution using minimum-distance methods". In: *Bernoulli* 5(1), pp. 177–189.

Hill, Bruce M et al. (1975). "A simple general approach to inference about the tail of a distribution". In: *The annals of statistics* 3(5), pp. 1163–1174.

Hong, L Jeff (2009). "Estimating quantile sensitivities". In: *Operations research* 57(1), pp. 118–130.

Hong, L Jeff and Guangwu Liu (2009). "Simulating sensitivities of conditional value at risk". In: *Management Science* 55(2), pp. 281–293.

Jenkinson, Arthur F (1955). "The frequency distribution of the annual maximum (or minimum) values of meteorological elements". In: *Quarterly Journal of the Royal Meteorological Society* 81(348), pp. 158–171.

Jiang, Guangxin and Michael C Fu (2015). "On estimating quantile sensitivities via infinitesimal perturbation analysis". In: *Operations Research* 63(2), pp. 435–441.

Kalkbrener, Michael (2005). "An axiomatic approach to capital allocation". In: *Mathematical Finance* 15(3), pp. 425–437.

Kneip, Alois, Léopold Simar, and Ingrid Van Keilegom (2015). "Frontier estimation in the presence of measurement error with unknown variance". In: *Journal of Econometrics* 184(2), pp. 379–393.

Krasnoselskiǐ, Mark Aleksandrovich and IAkov Bronislavovich Rutitskiǐ (1961). *Convex functions and Orlicz spaces*. P. Noordhoff.

Li, Zhouping, Yun Gong, and Liang Peng (2010). "Empirical likelihood method for intermediate quantiles". In: *Statistics & probability letters* 80(11-12), pp. 1022–1029.

Li, Zhouping and Liang Peng (2012). "Bootstrapping endpoint". In: *Sankhya A* 74(1), pp. 126–140.

Ling, Shiqing (2007). "Self-weighted and local quasi-maximum likelihood estimators for ARMA-GARCH/IGARCH models". In: *Journal of Econometrics* 140(2), pp. 849–873.

Liu, Guangwu and Liu Jeff Hong (2009). "Kernel estimation of quantile sensitivities". In: *Naval Research Logistics (NRL)* 56(6), pp. 511–525.

Mao, Tiantian and Taizhong Hu (2012). "Second-order properties of the Haezendonck–Goovaerts risk measure for extreme risks". In: *Insurance: Mathematics and Economics* 51(2), pp. 333–343.

Matthys, Gunther et al. (2004). "Estimating catastrophic quantile levels for heavy-tailed distributions". In: *Insurance: Mathematics and Economics* 34(3), pp. 517–537.

Meister, Alexander (2006). "Density estimation with normal measurement error with unknown variance". In: *Statistica Sinica*, pp. 195–211.

Meister, Alexander and Michael H Neumann (2010). "Deconvolution from non-standard error densities under replicated measurements". In: *Statistica Sinica*, pp. 1609–1636.

Molanes Lopez, Elisa M, Ingrid Van Keilegom, and Noël Veraverbeke (2009). "Empirical Likelihood for Non-Smooth Criterion Functions". In: *Scandinavian Journal of Statistics* 36(3), pp. 413–432.

Necir, Abdelhakim and Djamel Meraghni (2009). "Empirical estimation of the proportional hazard premium for heavy-tailed claim amounts". In: *Insurance: Mathematics and economics* 45(1), pp. 49–58.

Owen, Art (1990). "Empirical likelihood ratio confidence regions". In: *The Annals of Statistics*, pp. 90–120.

Owen, Art B (2001). *Empirical likelihood*. Chapman and Hall/CRC.

Peng, Liang (2001). "Estimating the mean of a heavy tailed distribution". In: *Statistics & Probability Letters* 52(3), pp. 255–264.

Peng, Liang and Yongcheng Qi (2006). "Confidence regions for high quantiles of a heavy tailed distribution". In: *The Annals of Statistics*, pp. 1964–1986.

Peng, Liang, Xing Wang, and Yanting Zheng (2015). "Empirical likelihood inference for Haezendonck-Goovaerts risk measure". In: *European Actuarial Journal* 5(2), pp. 427–445.

Peng, Liang et al. (2004). "Empirical-likelihood-based confidence interval for the mean with a heavy-tailed distribution". In: *The Annals of Statistics* 32(3), pp. 1192–1214.

Pickands III, James (1975). "Statistical inference using extreme order statistics". In: *the Annals of Statistics*, pp. 119–131.

Qi, Yongcheng (2008). "Bootstrap and empirical likelihood methods in extremes". In: *Extremes* 11(1), pp. 81–97.

Qin, Jin and Jerry Lawless (1994). "Empirical likelihood and general estimating equations". In: *The Annals of Statistics*, pp. 300–325.

Scaillet, Olivier (2004). "Nonparametric estimation and sensitivity analysis of expected shortfall". In: *Mathematical Finance* 14(1), pp. 115–129.

Shao, Jun and Dongsheng Tu (2012). *The jackknife and bootstrap*. Springer Science & Business Media.

Shorack, GR and JA Wellner (1986). "Empirical Processes with Applications to Statistics. 1986". In: *John Wiley&Sons*.

Skii, Ma Krasnosel and Jab Ruttcki (1961). "Convex functions and Orlicz spaces, Translated from the first Russian edition by Leo F". In: *Boron, P. Noordhoff Ltd., Groningen*.

Stefanski, Leonard A and Raymond J Carroll (1990). "Deconvolving kernel density estimators". In: *Statistics* 21(2), pp. 169–184.

Tang, Qihe and Fan Yang (2012). "On the Haezendonck–Goovaerts risk measure for extreme risks". In: *Insurance: Mathematics and Economics* 50(1), pp. 217–227.

Tang, Qihe and Fan Yang (2014). "Extreme value analysis of the Haezendonck–Goovaerts risk measure with a general Young function". In: *Insurance: Mathematics and Economics* 59, pp. 311–320.

Tsanakas, Andreas and Pietro Millossovich (2016). "Sensitivity analysis using risk measures". In: *Risk Analysis* 36(1), pp. 30–48.

Vandewalle, B and Jan Beirlant (2006). "On univariate extreme value statistics and the estimation of reinsurance premiums". In: *Insurance: Mathematics and Economics* 38(3), pp. 441–459.

Von Mises, Richard (1936). "Über allgemeine Quadraturformeln." In: *Journal für die reine und angewandte Mathematik* 174, pp. 56–67.

Wang, Xing and Liang Peng (2016). "Inference for intermediate Haezendonck–Goovaerts risk measure". In: *Insurance: Mathematics and Economics* 68, pp. 231–240.

Zhu, Ke, Shiqing Ling, et al. (2011). "Global self-weighted and local quasi-maximum exponential likelihood estimators for ARMA–GARCH/IGARCH models". In: *The Annals of Statistics* 39(4), pp. 2131–2163.

Zhu, Yunzhou, Lixin Zhang, and Yi Zhang (2013). "Optimal reinsurance under the Haezendonck risk measure". In: *Statistics & Probability Letters* 83(4), pp. 1111–1116.