Empirical Likelihood Inference for Haezendonck-Goovaerts Risk Measure

Liang Peng[1], Xing Wang[1] and Yanting Zheng[2]

Abstract. 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 paper, we propose an alternative empirical likelihood inference for this risk measure. A simulation study shows the good performance of the proposed method.

Key words and phrases: Empirical likelihood method, Haezendonck-Goovaerts risk measure, interval estimation

1 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

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

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

$$\theta_q = \inf_{\beta>0}\{\beta + \alpha(\beta)\}. \tag{2}$$

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

[1]Department of Risk Management and Insurance, Georgia State University, Atlanta, GA 30303, USA

[2]Department of Finance, Beijing Technology and Business University, Beijing 100048, China

Recently there has been an increasing interest in studying this risk measure with applications in actuarial science. For example, Goovaerts, Kaas, Dhaene and Tang (2004) showed that this risk measure preserves the convex order property; Bellini and Gianin (2008a, 2012) provided a dual representation for this risk measure; Goovaerts, Linders, Van Weert and Tank (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, 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, Deng, Peng and Xia (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.

We organize this paper as follows. Section 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 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 4. Some conclusions are made in Section 5.

2 Methodology and Main Results

Throughout suppose $X, X_1, \cdots, X_n$ 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 solves the following equation with respect to α for each fixed β:

$$\frac{1}{n}\sum_{i=1}^{n}\psi(\frac{(X_i-\beta)_+}{\alpha})=1-q. \tag{3}$$

This equation is the sample version of equation (1). Denote this solution by $\hat{\alpha}(\beta)$. Next, using (2), Ahn and Shyamalkumar (2014) defined their nonparametric estimator for θ_q as

$$\hat{\theta}_q^{AS}=\inf_{\beta>0}\{\beta+\hat{\alpha}(\beta)\}, \tag{4}$$

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 (1) and (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 1 below.

By taking derivative with respect to β in (1), we obtain

$$E\{\psi'(\frac{X-\beta}{\alpha(\beta)})\frac{-\alpha(\beta)-(X-\beta)\alpha'(\beta)}{\alpha^2(\beta)}I(X>\beta)\}=0. \tag{5}$$

Equation (2) implies that we have to solve the equation $1+\alpha'(\beta)=0$, which, combining with (5), results in the following estimating equation

$$E\{\psi'(\frac{X-\beta}{\alpha(\beta)})(X-\beta-\alpha(\beta))I(X>\beta)\}=0. \tag{6}$$

Hence, it follows from (1) and (6) that $\theta_q(>\beta)$ and β satisfy the following estimating equations:

$$\begin{cases} E\{\psi(\frac{X_i-\beta}{\theta_q-\beta})I(X_i>\beta)\}=1-q, \\ E\{\psi'(\frac{X_i-\beta}{\theta_q-\beta})(X_i-\theta_q)I(X_i>\beta)\}=0. \end{cases} \tag{7}$$

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 (7) 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

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

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

$$l(\theta_q, \beta) := -2 \log L(\theta_q, \beta) = 2 \sum_{i=1}^n \log(1 + \lambda^T Y_i(\theta_q, \beta)), \tag{8}$$

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

$$\sum_{i=1}^n \frac{Y_i(\theta_q, \beta)}{1 + \lambda^T Y_i(\theta_q, \beta)} = 0. \tag{9}$$

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, Van 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:

- C1) 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) F is continuous;

- C3) $E\{\sup_{(\theta_q,\beta)^T\in\Omega} |\psi(\frac{X-\beta}{\theta_q-\beta})|^{2\delta_1} I(X > \beta)\} < \infty$ 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 $\delta_1 > 1$,

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

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

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

  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 (1) and (2).

Theorem 1. Under conditions C1)–C3), 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

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

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

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

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 1. *When Σ_1 has rank 2, then i) in Theorem 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 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.

Remark 2. *Note that we do not assume $\psi'(0+)=0$. Instead we assume F is continuous to ensure (7) holds. So conditions C1) and C2) appear in Tang and Yang (2014). The first two*

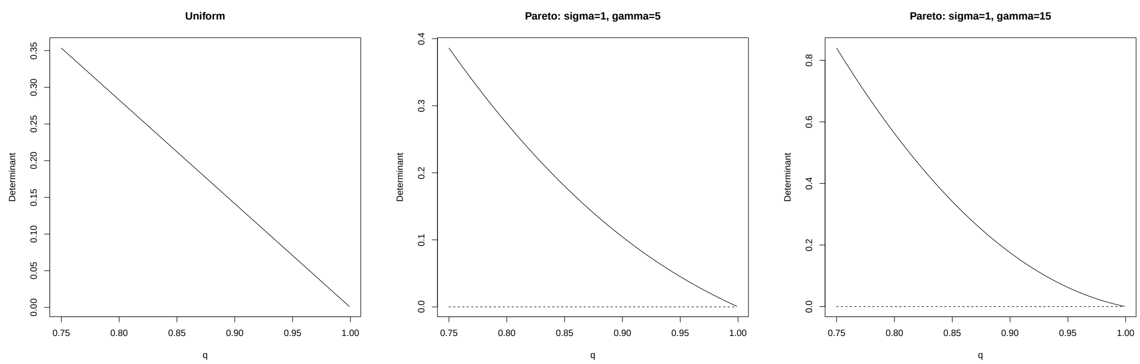

Figure 1: Determinant of Σ_1 in Theorem 1.

inequalities with respect to δ_1 *in C3) ensure Lemmas 2 and 3, which are standard for an empirical likelihood method. The other two inequalities in C3) 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 1. We employ these different conditions due to non-differentiability. All conditions C1)–C3) can be checked straightforward.*

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

$$I_\xi^{EL} = \{\theta_q : l^P(\theta_q) \le \chi^2_{1,\xi}\},$$

where $\chi^2_{1,\xi}$ denotes the $\xi-$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 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.

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

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

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 C1–C3) in Theorem 1 are satisfied. For example, one can choose any $\delta_1 > 1$ and $\delta_2 \in (0, \frac{1}{2})$ in C3) 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 C3) for distribution $F_2(x; \gamma, \sigma)$ when Ω is chosen small enough. In Table 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 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, 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.

4 Proofs

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

$$\sup_{-\infty<x<\infty} |\alpha_n(x) - B(x)| \stackrel{a.s.}{\to} 0 \quad \text{and} \quad \sup_{-\infty<x<\infty} \frac{|\alpha_n(x)|}{F^\nu(x)(1-F(x))^\nu} = O_p(1) \tag{10}$$

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 1. *Under conditions of Theorem 1, when* $|\beta - \beta_0| + |\theta_q - \theta_{0,q}| = \Delta_n \stackrel{p}{\to} 0$ *as* $n \to \infty$, *we have*

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

$$\begin{aligned} & \tfrac{1}{n}\sum_{i=1}^n \psi'(\tfrac{X_i-\beta}{\theta_q-\beta})(X_i-\theta_q)I(X_i > \beta) \\ = & \psi'(0+)(\beta_0-\theta_{0,q})\{F(\beta_0) - F_n(\beta_0)\} \\ & + \int_{\beta_0}^\infty \{F(x) - F_n(x)\}\{\psi''(\tfrac{x-\beta_0}{\theta_{0,q}-\beta_0})\tfrac{x-\theta_{0,q}}{\theta_{0,q}-\beta_0} + \psi'(\tfrac{x-\beta_0}{\theta_{0,q}-\beta_0})\}\,dx \\ & +(\beta-\beta_0)\{-\psi'(0+)(\beta_0-\theta_{0,q}) + \int_{\beta_0}^\infty \psi''(\tfrac{x-\beta_0}{\theta_{0,q}-\beta_0})\tfrac{(x-\theta_{0,q})^2}{(\theta_{0,q}-\beta_0)^2}\,dF(x)\} \\ & +(\theta_q-\theta_{0,q})\int_{\beta_0}^\infty \{\psi''(\tfrac{x-\beta_0}{\theta_{0,q}-\theta_0})\tfrac{(x-\theta_{0,q})(\beta_0-x)}{(\theta_{0,q}-\beta_0)^2} - \psi'(\tfrac{x-\beta_0}{\theta_{0,q}-\beta_0})\}\,dF(x) + o_p(\tfrac{1}{\sqrt{n}} + \Delta_n), \end{aligned}$$

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

$$
\begin{aligned}
&\tfrac{1}{n}\textstyle\sum_{i=1}^n\{\psi'(\frac{X_i-\beta}{\theta_q-\beta})\}^2(X_i-\theta_q)^2 I(X_i>\beta)-\int_{\beta_0}^\infty\{\psi'(\frac{x-\beta_0}{\theta_{0,q}-\beta_0})\}^2(x-\theta_{0,q})^2\,dF(x)\\
=\;&\{F(\beta_0)-F_n(\beta_0)\}\{\psi'(0+)\}^2(\beta_0-\theta_{0,q})^2\\
&+2\int_{\beta_0}^\infty\{F(x)-F_n(x)\}\{\psi'(\tfrac{x-\beta_0}{\theta_{0,q}-\beta_0})\psi''(\tfrac{x-\beta_0}{\theta_{0,q}-\beta_0})\tfrac{(x-\theta_{0,q})^2}{\theta_{0,q}-\beta_0}+(\psi'(\tfrac{x-\beta_0}{\theta_{0,q}-\beta_0}))^2(x-\theta_{0,q})\}\,dx\\
&+(\beta-\beta_0)\{-(\psi'(0+))^2(\beta_0-\theta_{0,q})^2+2\int_{\beta_0}^\infty\psi'(\tfrac{x-\beta_0}{\theta_{0,q}-\beta_0})\psi''(\tfrac{x-\beta_0}{\theta_{0,q}-\beta_0})\tfrac{(x-\theta_{0,q})^3}{(\theta_{0,q}-\beta_0)^2}\}\,dF(x)\\
&+(\theta_q-\theta_{0,q})\int_{\beta_0}^\infty\{2\psi'(\tfrac{x-\beta_0}{\theta_{0,q}-\beta_0})\psi''(\tfrac{x-\beta_0}{\theta_{0,q}-\beta_0})\tfrac{(\beta_0-x)(x-\theta_{0,q})^2}{(\theta_{0,q}-\beta_0)^2}+2(\psi'(\tfrac{x-\beta_0}{\theta_{0,q}-\beta_0}))^2(\theta_{0,q}-x)\}\,dF(x)\\
&+o_p(\tfrac{1}{\sqrt{n}}+\Delta_n)
\end{aligned}
$$

and

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

Proof. It follows from the Taylor expansion that

$$
\begin{aligned}
&\tfrac{1}{n}\textstyle\sum_{i=1}^n\psi(\frac{X_i-\beta}{\theta_q-\beta})I(X_i>\beta)-1+q\\
=\;&\int_\beta^\infty\psi(\tfrac{x-\beta}{\theta_q-\beta})\,dF_n(x)-\int_{\beta_0}^\infty\psi(\tfrac{x-\beta_0}{\theta_{0,q}-\beta_0})\,dF(x)\\
=\;&-\int_\beta^\infty\psi(\tfrac{x-\beta}{\theta_q-\beta})\,d\{F(x)-F_n(x)\}+\int_\beta^\infty\psi(\tfrac{x-\beta}{\theta_q-\beta})\,dF(x)-\int_{\beta_0}^\infty\psi(\tfrac{x-\beta_0}{\theta_{0,q}-\beta_0})\,dF(x)\\
=\;&\int_\beta^\infty\{F(x)-F_n(x)\}\psi'(\tfrac{x-\beta}{\theta_q-\beta})\tfrac{1}{\theta_q-\beta}\,dx+(\beta-\beta_0)\int_{\beta_1}^\infty\psi'(\tfrac{x-\beta_1}{\theta_1-\beta_1})\tfrac{x-\theta_1}{(\theta_1-\beta_1)^2}\,dF(x)\\
&+(\theta-\theta_{0,q})\int_{\beta_1}^\infty\psi'(\tfrac{x-\beta_1}{\theta_1-\beta_1})\tfrac{\beta_1-x}{(\theta_1-\beta_1)^2}\,dF(x)\\
=\;&\int_{\beta_0}^\infty\{F(x)-F_n(x)\}\psi'(\tfrac{x-\beta_0}{\theta_{0,q}-\beta_0})\tfrac{1}{\theta_{0,q}-\beta_0}\,dx+(\beta-\beta_0)\{F_n(\beta_2)-F(\beta_2)\}\psi'(0+)\tfrac{1}{\theta_2-\beta_2}\\
&+(\beta-\beta_0)\int_{\beta_2}^\infty\{F(x)-F_n(x)\}\{\psi''(\tfrac{x-\beta_2}{\theta_2-\beta_2})\tfrac{x-\theta_2}{(\theta_2-\beta_2)^3}+\psi'(\tfrac{x-\beta_2}{\theta_2-\beta_2})\tfrac{1}{(\theta_2-\beta_2)^2}\}\,dx\\
&+(\theta_q-\theta_{0,q})\int_{\beta_2}^\infty\{F(x)-F_n(x)\}\{\psi''(\tfrac{x-\beta_2}{\theta_2-\beta_2})\tfrac{\beta_2-x}{(\theta_2-\beta_2)^2}+\psi'(\tfrac{x-\beta_2}{\theta_2-\beta_2})\tfrac{1}{(\theta_2-\beta_2)^2}\}\,dx\\
&+(\beta-\beta_0)\int_{\beta_1}^\infty\psi'(\tfrac{x-\beta_1}{\theta_1-\beta_1})\tfrac{x-\theta_1}{(\theta_1-\beta_1)^2}\,dF(x)+(\theta-\theta_{0,q})\int_{\beta_1}^\infty\psi'(\tfrac{x-\beta_1}{\theta_1-\beta_1})\tfrac{\beta_1-x}{(\theta_1-\beta_1)^2}\,dF(x)\\
:=\;&I_1+\cdots+I_6,
\end{aligned}
\tag{11}
$$

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 (10) that

$$I_2 = O_p(\frac{1}{\sqrt{n}}\Delta_n) = o_p(\frac{1}{\sqrt{n}} + \Delta_n). \tag{12}$$

By (10) and condition C3), we have

$$I_3 = O_p(\frac{1}{\sqrt{n}}\Delta_n) = o_p(\frac{1}{\sqrt{n}} + \Delta_n) \quad \text{and} \quad I_4 = O_p(\frac{1}{\sqrt{n}}\Delta_n) = o_p(\frac{1}{\sqrt{n}} + \Delta_n). \tag{13}$$

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 C3) implies that

$$\begin{cases} \sup_{(\theta_q,\beta)^T\in\Omega} \int_\beta^\infty |\psi'(\frac{x-\beta}{\theta_q-\beta})|\, dF(x) < \infty \\ \sup_{(\theta_q,\beta)^T\in\Omega} \int_\beta^\infty |\psi'(\frac{x-\beta}{\theta_q-\beta})|x\, dF(x)\} < \infty \end{cases} \tag{14}$$

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 C3) implies that

$$\begin{cases} \sup_{(\theta_q,\beta)^T\in\Omega} \int_\beta^\infty \psi''(\frac{x-\beta}{\theta_q-\beta})x^2\, dF(x) < \infty \\ \sup_{(\theta_q,\beta)^T\in\Omega} \int_\beta^\infty \psi''(\frac{x-\beta}{\theta_q-\beta})x\, dF(x) < \infty \\ \sup_{(\theta_q,\beta)^T\in\Omega} \int_\beta^\infty \psi''(\frac{x-\beta}{\theta_q-\beta})\, dF(x) < \infty. \end{cases} \tag{15}$$

Hence it follows from (14), (15) and the Taylor expansion that

$$\begin{cases} I_5 = (\beta-\beta_0)\int_{\beta_0}^\infty \psi'(\frac{x-\beta_0}{\theta_{0,q}-\beta_0})\frac{x-\theta_{0,q}}{(\theta_{0,q}-\beta_0)^2}\, dF(x) + O_p(\Delta_n^2) \\ I_6 = (\theta_q-\theta_{0,q})\int_{\beta_0}^\infty \psi'(\frac{x-\beta_0}{\theta_{0,q}-\beta_0})\frac{\beta_0-x}{(\theta_{0,q}-\beta_0)^2}\, dF(x) + O_p(\Delta_n^2). \end{cases} \tag{16}$$

Therefore, the first equation in Lemma 1 follows from (11), (12), (13) and (16). The rest can be shown similarly. □

Lemma 2. *Under conditions of Theorem 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 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 1 and (10). 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

$$a\psi(\frac{x-\beta_0}{\theta_{0,q}-\beta_0}) + b\psi'(\frac{x-\beta_0}{\theta_{0,q}-\beta_0})(x-\theta_{0,q}) = c \tag{17}$$

for some constant c and all $x > \beta_0$. Obviously, when $b=0$, (17) can not be true since ψ is a strictly convex function. By assuming $b \neq 0$, it follows from (17) 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 (17) can not be true, i.e., Σ_0 is positive definite. □

Lemma 3. *Under conditions of Theorem 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 1. i) Like the proof of Owen (1990), it follows from Lemmas 1–3 and C3) 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

$$\begin{aligned} l(\theta_q,\beta) &= 2\textstyle\sum_{i=1}^n \lambda^T Y_i(\theta_q,\beta)-\sum_{i=1}^n \lambda^T Y_i(\theta_q,\beta)Y_i^T(\theta_q,\beta)\lambda+o_p(1)\\ &=\{\tfrac{1}{\sqrt{n}}\textstyle\sum_{i=1}^n Y_i(\theta_q,\beta)\}^T\{\tfrac{1}{n}\sum_{i=1}^n Y_i(\theta_q,\beta)Y_i^T(\theta_q,\beta)\}^{-1}\{\tfrac{1}{\sqrt{n}}\sum_{i=1}^n Y_i(\theta_q,\beta)\}+o_p(1)\\ &=\{\tfrac{1}{\sqrt{n}}\textstyle\sum_{i=1}^n Y_i(\theta_q,\beta)\}^T\Sigma_0^{-1}\{\tfrac{1}{\sqrt{n}}\sum_{i=1}^n Y_i(\theta_q,\beta)\}+o_p(1).\end{aligned} \tag{18}$$

Put $\nu/\sqrt{n}=(\beta-\beta_0,\theta_q-\theta_{0,q})^T$. Then it follows from (18) and Lemmas 1–2 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

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

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.,

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

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

Acknowledgments. Peng's research was partly supported by the Simons Foundation and a CEAR small grant in the Robinson College of Business at Georgia State University.

References

J.Y. Ahn and N.D. Shyamalkumar (2014). Asymptotic theory for the empirical Haezendonck-Goovaerts risk measure. *Insurance: Mathematics and Economics 55, 78–90.*

R.E. Baysal and J. Staum (2008). Empirical likelihood for value-at-risk and expected shortfall. *The Journal of Risk 11, 3–32.*

F. Bellini and E.R. Gianin (2008a). On Haezendonck risk measures. *Journal of Banking* & *Finance 32, 986–994.*

F. Bellini and E.R. Gianin (2008b). Optimal portfolios with Haezendonck risk measures. *Statistics* & *Decisions 26, 89–108.*

F. Bellini and E.R. Gianin (2012). Haezendonck-Goovaerts risk measures and Orlicz quantiles. *Insurance: Mathematics and Economics 51, 107–114.*

N.G. Chan, S. Deng, L. Peng and Z. Xia (2007). Interval estimation for the conditional Value-at-Risk based on GARCH models with heavy tailed innovations. *Journal of Econometrics 137, 556–576.*

S.X. Chen and P. Hall (1993). Smoothed empirical likelihood confidence intervals for quantiles. *Annals of Statistics 21, 1166–1181.*

J. Chen, L. Peng and Y. Zhao (2009). Empirical likelihood based confidence intervals for copulas. *Journal of Multivariate Analysis 100, 137 –151.*

K.C. Cheung and A. Lo (2013). General lower bounds on convex functionals of aggregate sums. *Insurance: Mathematics and Economics 53, 884–896.*

M. Goovaerts, R. Kaas, J. Dhaene and Q. Tang (2004). Some new classes of consistent risk measures. *Insurance: Mathematics and Economics 34, 505–516.*

M. Goovaerts, D. Linders, K. Van Weert and F. Tank (2012). On the interplay between distortion, mean value and Haezendonck-Goovaerts risk measures. *Insurance: Mathematics and Economics 51, 10–18.*

J. Haezendonck and M. Goovaerts (1982). A new premium calculation principle based on Orlicz norms. *Insurance: Mathematics and Economics 1, 41 – 53.*

T. Mao and T. Hu (2012). Second-order properties of the Haezendonck-Goovaerts risk measure for extreme risks. *Insurance: Mathematics and Economics 51, 333–343.*

E.M. Molanes Lopez, I. Van Keilegom and N. Veraverbeke (2009). Empirical likelihood for non-smooth criterion functions. *Scandinavian Journal of Statistics 36, 413–432.*

A.B. Owen (1990). Empirical likelihood confidence regions. *Annals of Statistics 18, 90–120.*

A.B. Owen (2001). Empirical Likelihood. *Chapman* & *Hall.*

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

J. Qin and J. Lawless (1994). Empirical likelihood and general estimating equations. *Annals of Statistics 22, 300–325.*

G.R. Shorack and J.A. Wellner (1986). Empirical Processes with Applications to Statistics. *New York: Wiley.*

Q. Tang and F. Yang (2012). On the Haezendonck-Goovaerts risk measure for extreme risks. *Insurance: Mathematics and Economics 50, 217–227.*

Q. Tang and F. Yang (2014). Extreme value analysis of the Haezendonck-Goovaerts risk measure with a general Young function. *Insurance: Mathematics and Economics 59, 311–320.*

Y. Zhu, L. Zhang and Y. Zhang (2013). Optimal reinsurance under Haezendonck risk measure. *Statistics* & *Probability Letters 83, 1111–1116.*

Table 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: 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