Distributed Estimator of Market Beta under Extreme Conditions

Suyu Zhu

School of Mathematics and Statistics, Southwest University, Chongqing, China
Email: 906827220@qq.com

Abstract

Market beta is a measure of the volatility or systematic risk of a security or portfolio compared to the market as a whole. This paper considers the distributed estimation of market beta in the case of massive data, and obtains the consistency and asymptotic normality of the estimator. Further, simulations show the finite sample properties of this estimator.

Keywords

Heavy Tail, Tail Dependence, Distributed Statistical Inference, Market Beta

1. Introduction

Distributed statistical inference, as a hot topic and an effective method, has been widely discussed in the past ten years, and a lot of research results have been accumulated. Its representative work are: In theory, Chen and Zhou (Chen and Zhou) [1] put forward the distributed Hill estimator and prove its Oracle property; Volgushev *et al.* (2017) [2] propose distributed inference for quantile regression processes, and propose a method to calculate the efficiency of this inference, which requires almost no additional computational cost. In application, Mohammed *et al.* (2020) [3] propose a technique to divide a Deep Neural Networks (DNN) in multiple partitions, which reduces the total latency for DNN inference; Smith and Hollinger (2018) [4] propose a distributed inference-based multi-robot exploration technique that uses the observed map structure to infer unobserved map features, resulting in a reduction in the cumulative exploration path length in the trial; Ye (2017) [5] started to study the stability of the beta coefficient of the Chinese stock market and found the best beta estimation time. Mitra (2019) [6] uses a smooth linear transfer function to measure the amplitude and direction of market movement, and the proposed classification can better

capture the asymmetric behavior of beta.

Market beta, also known as systematic risk or equity beta, is a measure of a stock's sensitivity to overall market movements. The development of market beta can be traced back to the early 20th century when economists and financial analysts began to understand the importance of systematic risk in determining asset returns. One of the pioneering works in this area is the paper of Markowitz (1952) [7] on portfolio selection, which laid the foundation for modern portfolio theory, and emphasized the importance of diversification. Black and Scholes (1973) [8] introduced the concept of beta to measure the systematic risk of individual securities or stock. Banks with a higher beta are expected to suffer from larger capital losses in the event of an extremely adverse shock in the financial system.

Estimating market beta involves analyzing historical data on a stock's returns and its correlation with the market returns. The usually used econometric model is that the stock's returns are regressed against the market returns over a specific time period, which has been used to evaluate beta of financial returns on commodities, currencies (Atanasov and Nitschka (2014) [9]; Lettau et al. (2013)) [10], stocks (Post and Versijp (2004)) [11], and active trading strategies (Mitchell and Pulvino 2001) [12]. However, in extreme cases, conditional regression is based on a small number of tail observations, which may produce a relatively large variance of the estimator, and the data of the financial market are mostly heavy tail, which may further increase error. To avoid these situations, Oordt and Zhou (2017) [13] proposed a new method to estimate market β.

Let X and Y be continuous random variables with distribution functions F_X and F_Y, respectively. Assume that $1 - F_X$ and $1 - F_Y$ be heavy-tail with tail index α_x and α_y, respectively. This means that

$$1 - F_X(u) = u^{-\alpha_x} l_x(u) \text{ and } 1 - F_Y(u) = u^{-\alpha_y} l_y(u), \tag{1.1}$$

where $l_x(u)$ and $l_y(u)$ are slowly varying functions as $u \to \infty$. Let $Q_X(\overline{p}) = F_X^{\leftarrow}(1 - \overline{p})$ with small $\overline{p}$, relation of X and Y restricted under extreme X is given by

$$Y = \beta X + \varepsilon, \text{ for } X > Q_X(\overline{p}), \tag{1.2}$$

ε is the error term that is assumed to be independent of the X under the condition $X > Q_X(\overline{p})$.

To get estimator under the EVT method, we consider the following tail dependence measure from multivariate EVT (see, e.g, Hult and Lindskog (2002)) [14],

$$\tau := \lim_{p \to 0} \tau(p) = \lim_{p \to 0} \frac{1}{p} P\big(X > Q_X(p), Y > Q_Y(p)\big) \text{ [14]}, \tag{1.3}$$

where $Q_Y(p)$ denotes the quantile function of Y defined as $Q_Y(p) = F_Y^{\leftarrow}(1 - p)$. And we assume that the usual second-order condition (see, e.g, de Haan and Stadtmüller (1996)) [15] for X, which quantifies the speed of convergence in this relation as

$$\lim_{u\to\infty}\frac{\dfrac{P(X>ux)}{P(X>u)}-x^{-\alpha_x}}{a_x(u)}=x^{-\alpha_x}\frac{x^{\rho\alpha_x}-1}{\rho\dfrac{1}{\alpha_x}}\quad[15],\tag{1.4}$$

where $a_x(u)=A_x\left(\dfrac{1}{1-F(u)}\right)$ and $\lim_{u\to\infty}A_x(u)=0$, $\rho\le0$.

Suppose (1.4) hold, under the linear model in (1.2), with $\alpha_y>\alpha_x/2$, $\beta\ge0$, the following conclusion is given in Oordt and Zhou (2017) [13]:

$$\lim_{p\to0}(\tau(p))^{\frac{1}{\alpha_x}}\frac{Q_Y(p)}{Q_X(p)}=\beta\quad[13].\tag{1.5}$$

Naturally, consider independent and identically distributed (i.i.d.) observations $(X_1,Y_1),(X_2,Y_2),\cdots,(X_n,Y_n)$ with the i.i.d. unobserved error terms $\varepsilon_1,\cdots,\varepsilon_n$, we mimic the limit procedure $p\to0$ by considering only the lowest k observations in the tail region, such that $k:=k(n)\to\infty$ and $k/n\to0$ as $n\to\infty$, Oordt and Zhou (2017) [13] gives an estimator of β as

$$\hat{\beta}:=\hat{\tau}(k/n)^{1/\hat{\alpha}_x}\frac{\hat{Q}_Y(k/n)}{\hat{Q}_X(k/n)}\quad[13].\tag{1.6}$$

And to prove asymptotic normality, the second-order condition for Y is given by:

$$\lim_{u\to\infty}\frac{\dfrac{P(Y>uy)}{P(Y>u)}-y^{-\alpha_y}}{a_y(u)}=y^{-\alpha_y}\frac{y^{\rho'\alpha_y}-1}{\rho'\dfrac{1}{\alpha_y}}\quad[15],\tag{1.7}$$

where $a_y(u)=A_y\left(\dfrac{1}{1-F(u)}\right)$ is an eventually positive or negative function, $\lim_{u\to\infty}A_y(u)=0$, $\rho'\le0$. Then, Drees and Huang (1998) [16] define $R(x,y,p)=\dfrac{1}{p}P(X>Q_X(px),Y>Q_Y(py))$. For this dependence structure, we assume that, $R(x,y,p)\to R(x,y)$ as $p\to0$ for some positive function $R(x,y)$, with a speed of convergence as follows: there exists a $\theta>0$ for which, as $p\to0$,

$$R(x,y,p)-R(x,y)=O(p^\theta)\tag{1.8}$$

for all $(x,y)\in[0,1]^2/\{(0,0)\}$. And we can simply get

$$\lim_{p\to0}R(1,1,p)=\tau.\tag{1.9}$$

Under condition (1.4), (1.7) and (1.8) hold, suppose $k=O(n^\zeta)$, where

$$\zeta<\min\left(\frac{2\theta}{1+2\theta},\frac{2\rho}{2\rho+\alpha_x},\frac{2\rho'}{2\rho'+\alpha_y},\frac{3}{2+\alpha_y}\right),$$

Oordt and Zhou (2017) [13] prove the asymptotic normality of $\hat{\beta}$.

This estimation method can be used not only for the assessment of investment risks, but also for banking (Oordt and Zhou (2018)) [17], insurance and other fields. However, due to confidentiality, banks may not share their operating losses with each other, and insurance companies cannot share any observation results with the outside world in order to protect the privacy of customers. Therefore, banks and insurance companies can only make statistics based on their own data and share the results, and cannot re-identify individual data from the shared information. Distributed statistical inference is a good way to deal with these situations, it can analyze data stored in multiple machines, and it usually requires a divide-and-conquer algorithm that estimates the required parameters on each machine, transmits the results to a central machine that combines all the results, usually by simple averaging, to arrive at a computationally feasible estimator.

The objective of this paper is to apply divide-and-conquer idea to estimating market β. Considering independent and identically distributed (i.i.d.) observations $(X_1, Y_1), (X_2, Y_2), \cdots, (X_n, Y_n)$ are distributed across k different machines, each machine has m observations, $n = mk$, and we assume as $n \to \infty$,

$$k := k(n) \to \infty, \ m := m(n) \to \infty, \ \frac{m}{\log k} \to \infty. \tag{1.10}$$

We follow a divide-and-conquer algorithm, first estimating $\hat{\beta}_{n,j}$ in each machine, and then taking the average of k machines as the distributed estimator $\hat{\beta}_D$ for β,

$$\hat{\beta}_D = \frac{1}{k} \sum_{j=1}^{k} \hat{\beta}_{n,j}. \tag{1.11}$$

Sort the observations $X_j^1, X_j^2, \cdots, X_j^m$ in the j-th machine, we get the order statistic $X_j^{(1)} \geq X_j^{(2)} \geq \cdots \geq X_j^{(m)}$, and only the first d are selected to estimate β, where $d := d(n) \to \infty$, $\frac{d}{m} \to 0$ as $n \to \infty$. From (1.5), we have

$$\hat{\beta}_{n,j} = \left(\hat{\tau}_j \left(\frac{d}{m} \right) \right)^{\frac{1}{\hat{\alpha}_j}} \frac{\hat{Q}_j^Y \left(\frac{d}{m} \right)}{\hat{Q}_j^X \left(\frac{d}{m} \right)}, \tag{1.12}$$

where, the tail index is estimated using the Hill estimator given in Hill (1975) [18]:

$$\frac{1}{\hat{\alpha}_j} := \frac{1}{d} \sum_{i=1}^{d} \left(\log X_j^{(i)} - \log X_j^{(d+1)} \right) \quad [18].$$

The estimator of dependence measure is provided by multivariate EVT, see Embrechts *et al.* (2000) [19], that is

$$\hat{\tau}_j \left(\frac{d}{m} \right) = \frac{1}{d} \sum_{t=1}^{m} 1_{\left\{ Y_j^t > Y_j^{(d+1)}, X_j^t > X_j^{(d+1)} \right\}}, \quad [19]$$

where $Y_j^{(d+1)}$ is the $(d+1)$-th highest order statistic of Y_j. Finally,

$$\hat{Q}_j^X\left(\frac{d}{m}\right) = X_j^{(d+1)}, \quad \hat{Q}_j^Y\left(\frac{d}{m}\right) = Y_j^{(d+1)}.$$

When $\limsup\limits_{n\to\infty} \tau\left(\dfrac{d}{m}\right) = 0$, we require some additional conditions to ensure the asymptotic normality of the Hill estimator:

$$\lim_{n\to\infty} \sqrt{d}\, A_x\left(\frac{m}{d}\right) = \lambda < \infty, \frac{\sqrt{d}}{\log m} \to +\infty. \tag{1.13}$$

Suppose there would exist a sequence $p_n \to 0$ as $n \to \infty$ such that, for sufficient large n, we have $p_n < \bar{p}$, which implies that the linear model in (1.2) applies for sufficiently large n.

The remainder of this paper is organized as follows. Section 2 provides the main results; finite behaviors of $\hat{\beta}_D$ are considered in Section 3; all proofs are deferred to Section 4.

2. Main Innovations and Results

The innovations of this paper are:

- Under extreme market conditions, with less data and heavy tails, a new beta estimator is proposed by using the distributed idea.
- In the numerical simulation, the profile of data pollution is considered, and the expected effect is achieved, and the data is more inclusive.

The results of this paper are:

✓ The consistency of $\hat{\beta}_D$.

Theorem 2.1. *Under the linear tail model in (1.2), assume (1.1) and (1.4) hold. (1.13) holds when* $\lim\limits_{p\to 0} \tau(p) = 0$. *Then, as* $n \to \infty$, $\hat{\beta}_D \xrightarrow{P} \beta$.

✓ The asymptotic normality of $\hat{\beta}_D$.

Theorem 2.2. *Assume that the conditions in Theorem 2.1 hold,* $P(Y < -u) = O(P(Y > u))$ *as* $u \to \infty$. *Suppose both (1.7) and (1.8) hold,* $\lim\limits_{p\to 0} \tau(p) = \tau \in (0,1)$, *Further assume that* $kd = O(n^\zeta)$, *where*

$$\zeta < \min\left(\frac{2\theta}{1+2\theta}, \frac{2\rho}{2\rho-1}, \frac{2\rho'}{2\rho'-1}, \frac{2}{3+\alpha_y}\right).$$

Then, as $n \to \infty$,

$$\sqrt{kd}\left(\hat{\beta}_D - \beta\right) \xrightarrow{d} N\left(0, \frac{\beta^2}{\alpha_x^2}\left(\frac{1}{\tau} + 2\log\tau - \frac{2}{\tau}\cdot\log\tau - 1 - 3(\log\tau)^2\right)\right).$$

3. Simulation

We conduct two sets of simulations to demonstrate the finite sample performance of the distributed beta estimator $\hat{\beta}_D$. For each simulation, we consider three linear models, that is, $\beta = 1.5, 1, 0.5$. We generate samples with samples size $n = 10{,}000$. Based on $r = 1000$ repetitions, we obtain the finite sample squared bias, variance and Mean Squared Error (MSE) for our estimator.

3.1. Compare for Different Level of *d*

In the first set of simulations, we vary the level of *d* in the distributed beta estimator to verify the theoretical results on the oracle property. The oracle sample $X_1, \cdots, X_n$ contains *n* = 10,000 observations stored in *k* machines with *m* observations each. We fix *k* = 20 and *m* = 500, compare the finite sample performance of the distributed beta estimator with that of the oracle beta estimator for different values of *d*. Since the Student's *t*-distribution is known to be heavy-tailed with the tail index equal to the degrees of freedom, we perform simulations of *X* and *ε* based on random draws from a Student's *t*-distribution with Four degree of freedom. According to Lemma 1.3.1 in Embrechts *et al.* (1997) [20], the sum of two heavy-tailed random variables is also a heavy-tailed random variable, and the tail index of the sum is controlled by smaller tail index. Then, the observations for *Y* are constructed by aggregating the simulated *X* and *ε*, which could guarantees *Y* is also heavy-tailed and $\alpha_y > \frac{1}{2}\alpha_x$.

The first column of **Figure 1** compares the Mean Square Error of the distributed beta estimator $\hat{\beta}_D$ and the oracle beta estimator $\hat{\beta}$. Firstly, Mean Square Error gradually decreases with the increase of *β*. Theoretically, *τ* increases with the increase of *β*, while Mean Square Error decreases with the increase of
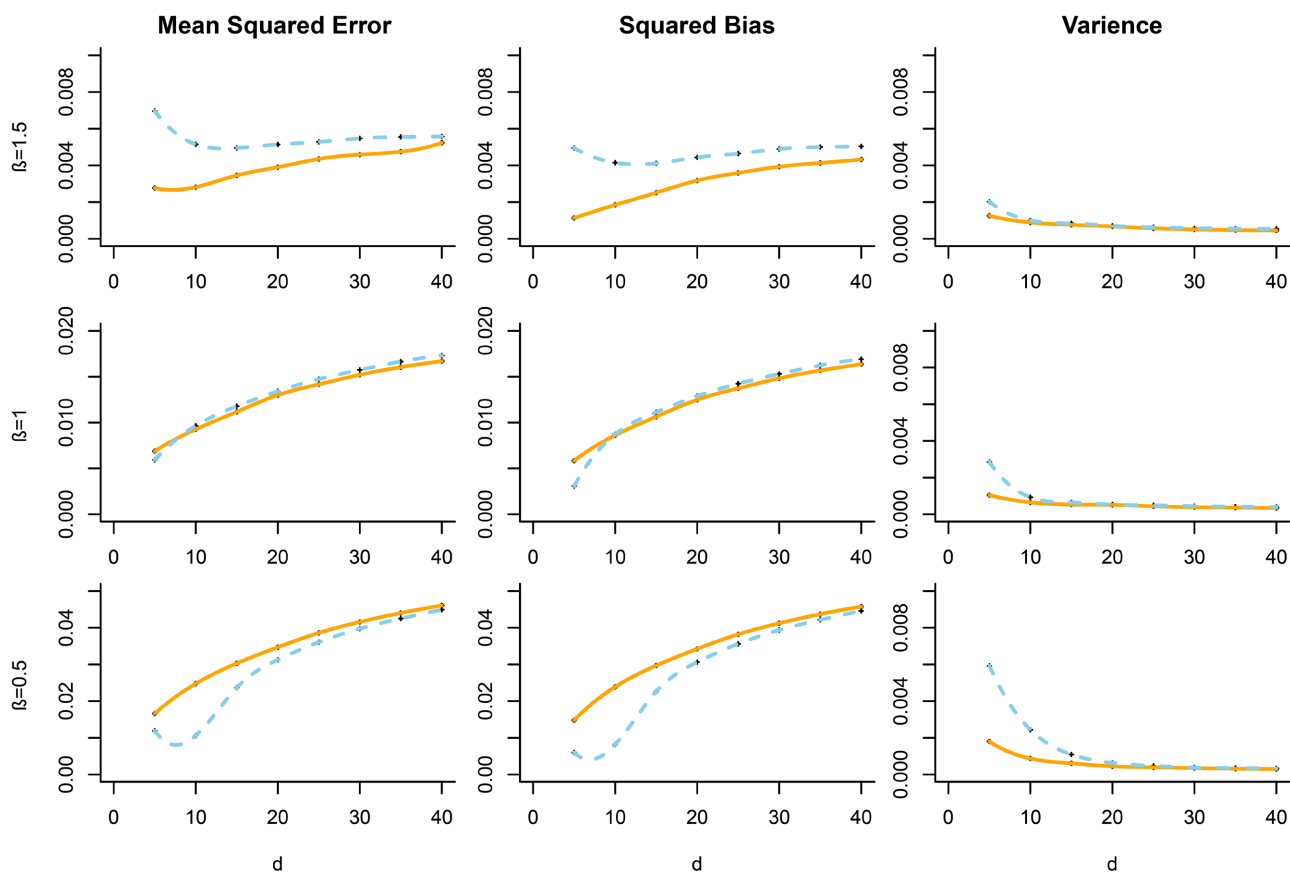

Figure 1. Finite sample performance for the distributed beta estimator and the oracle beta estimator for different levels of *d*. The blue report the simulation results for distributed EVT approach; the yellow lines report those for the EVT approach.

τ. Therefore, the simulation results are in agreement with the theoretical results. Secondly, the second and third columns of **Figure 1** show decomposition of the MSE into squared bias and variance, we observe a trade off between the bias and varience for the both estimators: as d increase, the bias increase while the variance decreases, and when the number of observations is small, the variance of the oracle beta estimator is smaller than that of the distributed beta estimator, and as the number of observations increases, the variance becomes equal, which is in line with the result of Theorem 2.2.

3.2. Data Is Contaminated

In the second set of simulations, we want to know whether distributed estimators have good properties when the data is contaminated. We simulate three cases of X being contaminated, ε being contaminated and both X and ε being contaminated respectively. The total number of observations does not change, that is, $n = 10,000$ is divided into $k = 20$ machines with $m = 500$ observations in each machine.

Figure 2 shows the Mean Square Error, square deviation and variance of the two estimators when X is contaminated. We also model 10,000 observations of ε
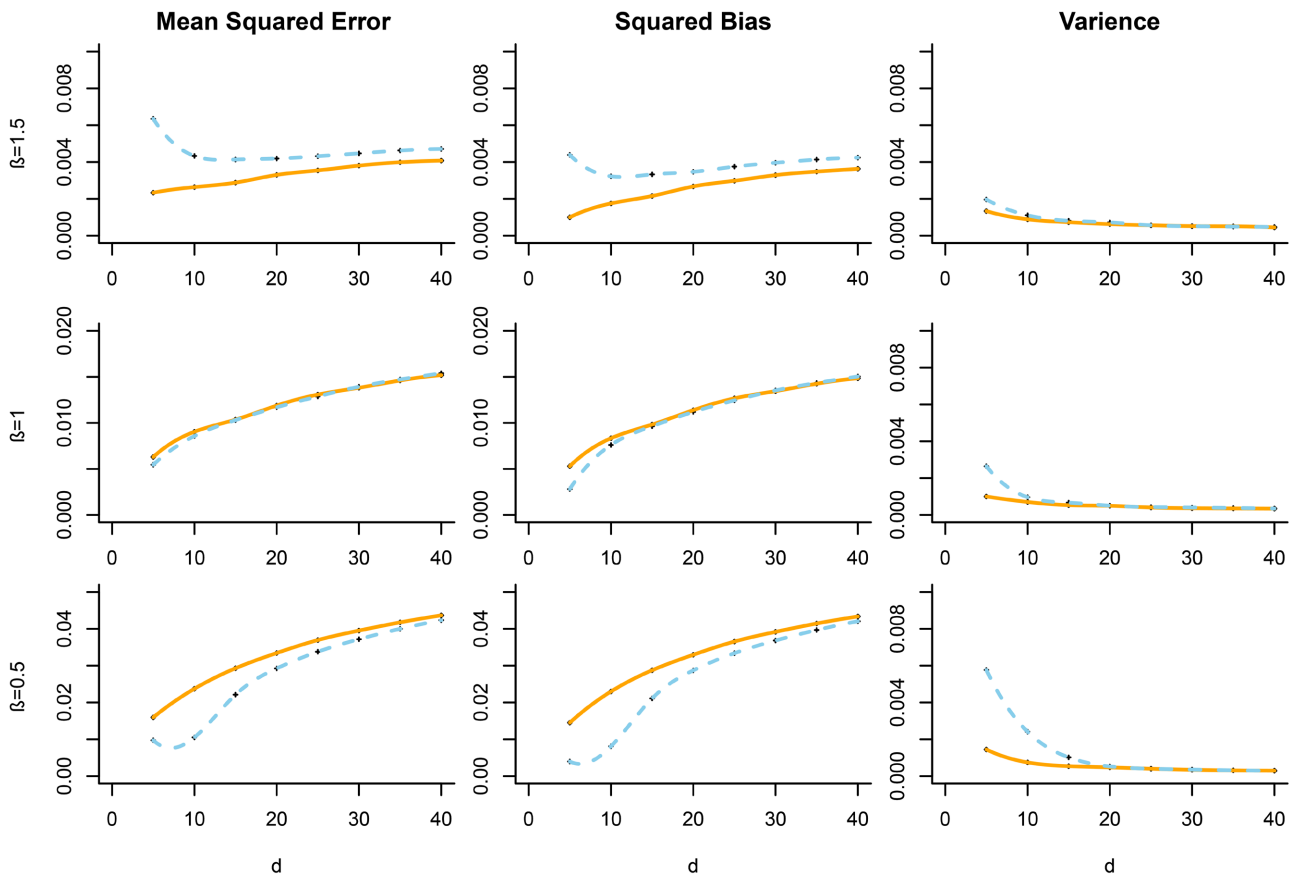

Figure 2. X is contaminated, finite sample performance for the distributed beta estimator and the oracle beta estimator for different levels of p. The blue lines represent the simulation results of the distributed beta estimator, and the red lines represent the corresponding Oracle results.

from a Student's *t*-distribution with 4 degrees of freedom, observations of *X* are drawn from a standard normal distribution with probability 0.1 and a Student's *t*-distribution with 4 degrees of freedom with probability 0.9, this means that 1000 out of 10,000 observations are contaminated. We then sort the observations in each machine and use (1.12) to get $\hat{\beta}_{n,j}$.

The third column in **Figure 2** shows the variance of the two estimators when *d* takes different values, which is almost the same as the result when the observations are not contaminated. When the number of observations is small, the variance of the distributed estimator is larger than that of the Oracle estimator, and with the increase of *d*, the variance is close to zero. Observe the first column, the Mean Square Error is less than 0.05, the estimation effect is good.

Figure 3 shows the Mean Square Error, square deviation and variance of the two estimators when ε is contaminated. We also model 10,000 observations of *X* from a Student's *t*-distribution with 4 degrees of freedom, observations of ε are drawn from a standard normal distribution with probability 0.1 and a Student's *t*-distribution with 4 degrees of freedom with probability 0.9, this means that 1000 out of 10,000 observations are contaminated. We then sort the observations in each machine and use (1.12) to get $\hat{\beta}_{n,j}$. Obviously, **Figure 3** is basically
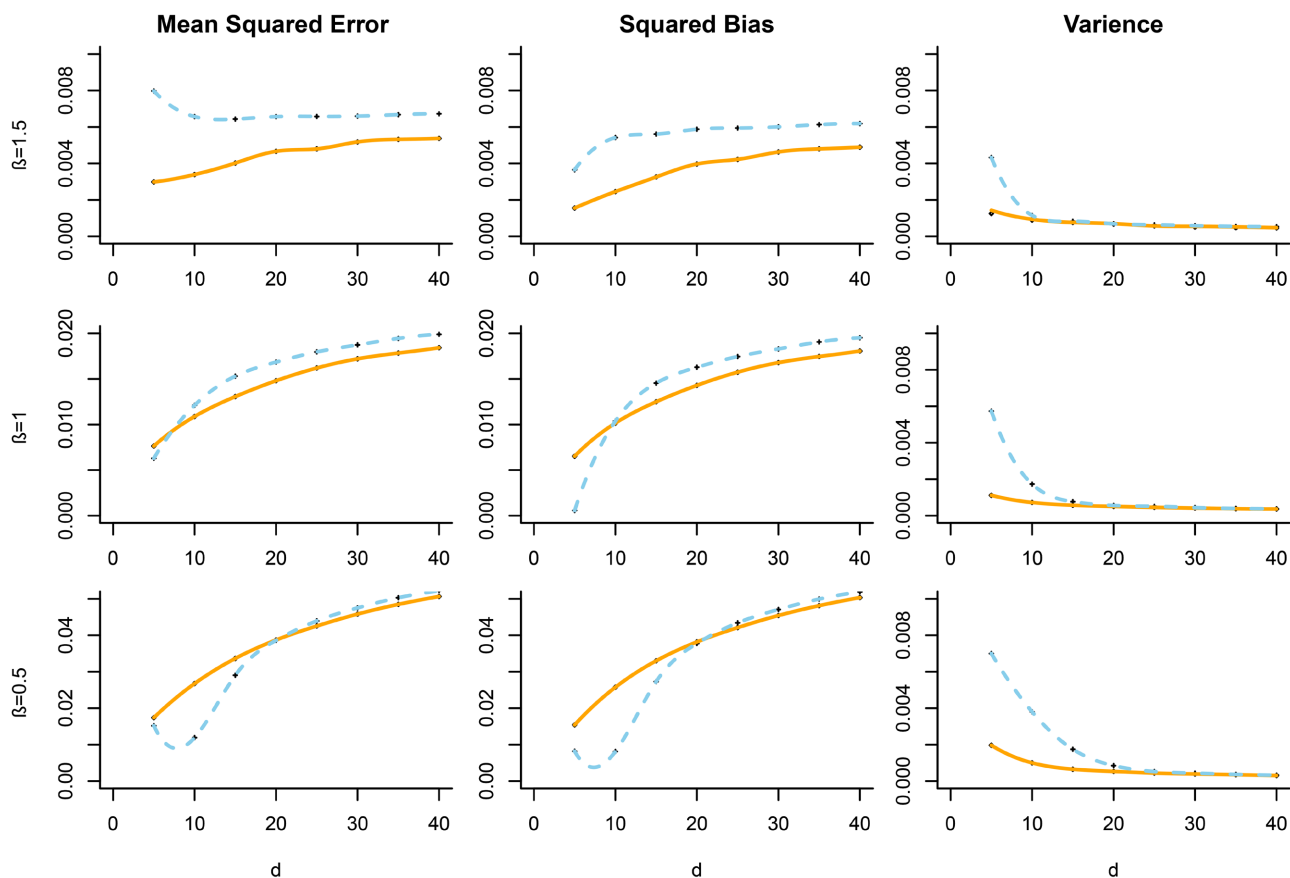

Figure 3. ε is contaminated, finite sample performance for the distributed beta estimator and the oracle beta estimator for different levels of *p*. The blue lines represent the simulation results of the distributed beta estimator, and the red lines represent the corresponding oracle results.

consistent with **Figure 1**, indicating that the selection of ε does not affect the properties of the estimator, which is consistent with the theory that the random error can be thin-tailed.

Figure 4 shows the Mean Square Error, square deviation and variance of the two estimators when both ε and X are contaminated. Observations of ε and X are drawn from a standard normal distribution with probability 0.1 and a Student's t-distribution with 4 degrees of freedom with probability 0.9, this means that 1000 out of 10,000 observations are contaminated. We then sort the observations in each machine and use (1.12) to get $\hat{\beta}_{n,j}$. Similar to **Figure 1**, this is consistent with the theoretical results, indicating that distributed estimators can be treated similarly when the data is contaminated.

4. Proof

In order to prove the main results, we need the following two lemmas.

Lemma 4.1. *Assuming that (1.3) and (1.7) hold,* $kd = O\left(n^{\zeta}\right)$,

$$\zeta < \min\left(\frac{2\rho}{2\rho-1}, \frac{2\rho'}{2\rho'-1}\right), \text{ then as } n \to \infty, \text{ we have}$$

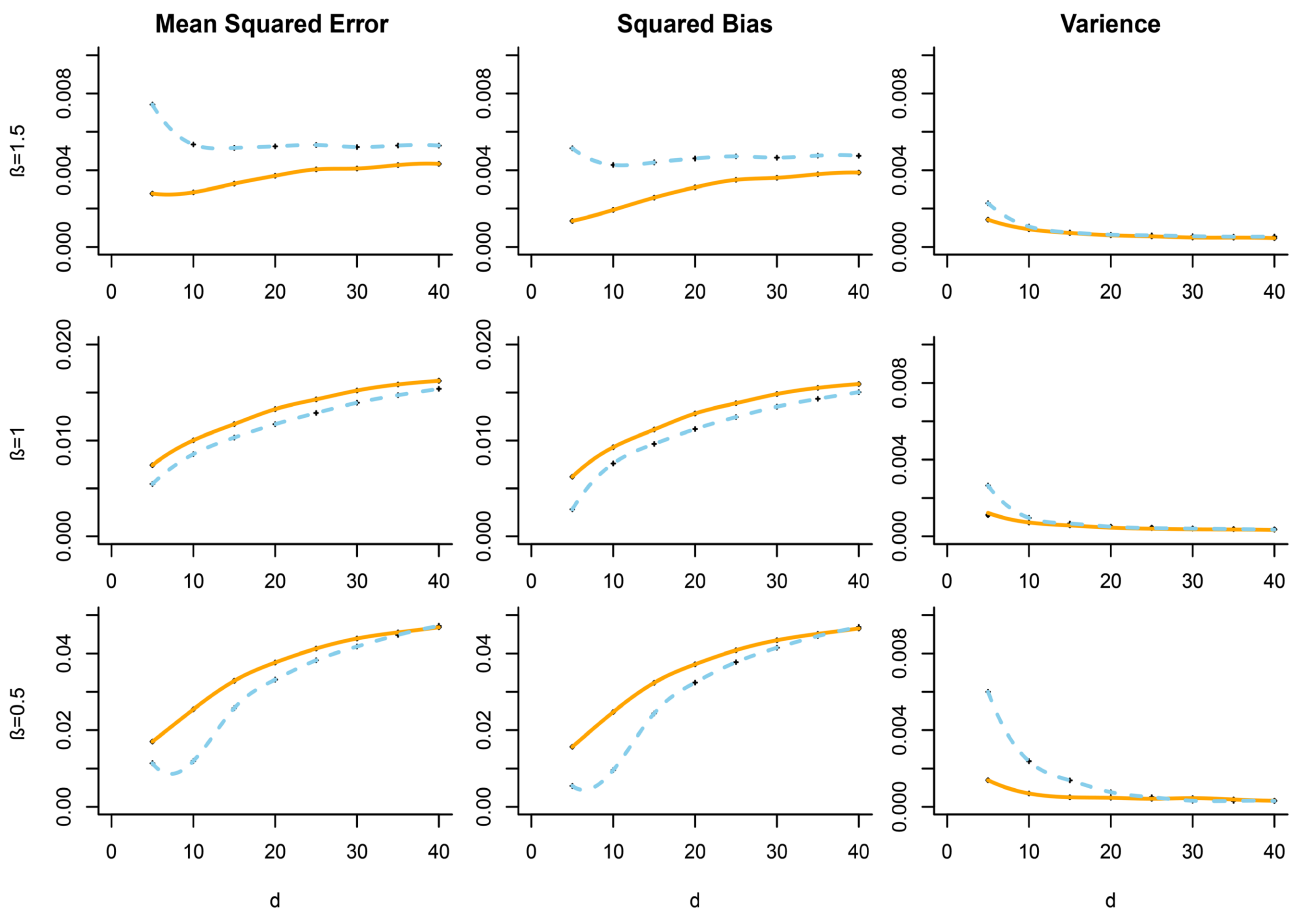**Figure 4.** Both ε and X are contaminated, finite sample performance for the distributed beta estimator and the oracle beta estimator for different levels of p. The blue lines represent the simulation results of the distributed beta estimator, and the red lines represent the corresponding Oracle results.

$$\lim_{n\to\infty}\sqrt{kd}\,A_x\left(\frac{m}{d}\right)=0;\ \lim_{n\to\infty}\sqrt{kd}\,A_y\left(\frac{m}{d}\right)=0.$$

Proof. According to theorem 2.3.3 in de Haan and Ferrira (2006) [21], as $t\to\infty$, $A_x(t)$ is a regular function with parameter $\rho<0$, then

$$\lim_{n\to\infty}\sqrt{kd}\,A_x\left(\frac{m}{d}\right)=\lim_{n\to\infty}\sqrt{kd}\left(\frac{m}{d}\right)^{\rho}\bar{l}_x\left(\frac{m}{d}\right)$$

$$=\lim_{n\to\infty}(kd)^{\frac{1}{2}-\rho}\,n^{\rho}\bar{l}_x\left(\frac{m}{d}\right)$$

$$=\lim_{n\to\infty}n^{\zeta\left(\frac{1}{2}-\rho\right)+\rho}\bar{l}_x\left(\frac{m}{d}\right)$$

$$=\lim_{n\to\infty}(kd)^{\zeta\left(\frac{1}{2}-\rho\right)+\rho}\left(\frac{n}{kd}\right)^{\zeta\left(\frac{1}{2}-\rho\right)+\rho}\bar{l}_x\left(\frac{n}{kd}\right)$$

where $\bar{l}_x$ is a slowly varying function, since $\zeta<\dfrac{2\rho}{2\rho-1}$, we have

$\zeta\left(\dfrac{1}{2}-\rho\right)+\rho<0$, then,

$$\lim_{n\to\infty}\sqrt{kd}\,A_x\left(\frac{m}{d}\right)=0,$$

similarly, $\lim_{n\to\infty}\sqrt{kd}\,A_y\left(\dfrac{m}{d}\right)=0$. $\square$

Let

$$C:=\left\{Y>Q_Y(py),X>Q_X(px)\right\};$$

$$C_1:=\left\{\beta X>Q_Y(py)(1+\delta),X>Q_X(px),\varepsilon>-\delta Q_Y(py)\right\};$$

$$C_{21}:=\left\{\beta X>Q_Y(py)(1-\delta),X>Q_X(px)\right\};$$

$$C_{22}:=\left\{\varepsilon>\delta Q_Y(py),X>Q_X(px)\right\}$$

with $\delta:=\delta(p)>0$ to be specified later. It's clear that for any $0<\delta<1$, $C_1\subset C\subset C_{21}\cup C_{22}$.

Lemma 4.2. *Suppose* $P(Y<-u)=O(P(Y>u))$ *as* $u\to\infty$, *Further assume that (1.4) and (1.7) hold, $kd=O(n^{\zeta})$, where*

$$\zeta<\min\left(\frac{2\theta}{1+2\theta},\frac{2\rho}{2\rho-1},\frac{2\rho'}{2\rho'-1},\frac{2}{3+\alpha_y}\right),$$

then as $n\to\infty$,

$$\sqrt{kd}\,P\left(\varepsilon>\delta_n Q_Y\left(\frac{d}{m}\right)\right)\to0; \tag{4.1}$$

$$\sqrt{kd}\,P\left(\varepsilon\le-\delta_n Q_Y\left(\frac{d}{m}\right)\right)\to0; \tag{4.2}$$

$$\sqrt{kd}\left[\frac{m}{d}P\left(\beta X > Q_Y\left(\frac{d}{m}\right)(1\pm\delta_n)\right) - \left(\frac{\beta Q_X\left(\frac{d}{m}\right)}{Q_Y\left(\frac{d}{m}\right)}\right)^{\alpha_x}\right] = 0. \qquad (4.3)$$

Proof. Let $\delta = \delta_n = (kd)^{-1/2-\kappa}$, $0 < \kappa < 1/2\,\alpha_y$, such that

$$\left(\frac{1}{2}+\kappa\right)\alpha_y + \frac{3}{2} < \frac{1}{\zeta}.$$

Notice that $\zeta < \dfrac{2}{\alpha_y + 3}$, the choice of κ is feasible. And we have that

$\lim\limits_{n\to\infty}\sqrt{kd}\delta_n = 0$ and $\lim\limits_{n\to\infty}\sqrt{kd}\delta_n^{-\alpha_y}\dfrac{d}{m} = 0$. We first prove that as $n\to\infty$,

$$\delta_n Q_Y\left(\frac{d}{m}\right)\to\infty.$$

From the heavy-tailed property of the distribution function of Y in (1), we obtain

that $Q_Y(p) = p^{-\frac{1}{\alpha_y}}\tilde{l}_y(p)$, $\tilde{l}_y$ is a slowly varying function as $p\to 0$, then

$$\lim_{n\to\infty}\delta_n Q_Y\left(\frac{d}{m}\right) = \lim_{n\to\infty}(kd)^{-\frac{1}{2}-\kappa}\left(\frac{kd}{n}\right)^{-\frac{1}{\alpha_y}}\tilde{l}_y\left(\frac{kd}{n}\right)$$

$$= \lim_{n\to\infty}n^{\frac{1}{\alpha_y}-\zeta\left(\frac{1}{2}+\kappa+\frac{1}{\alpha_y}\right)}\tilde{l}_y\left(\frac{kd}{n}\right)$$

$$= \lim_{n\to\infty}(kd)^{\frac{1}{\alpha_y}-\zeta\left(\frac{1}{2}+\kappa+\frac{1}{\alpha_y}\right)}\left(\frac{kd}{n}\right)^{\zeta\left(\frac{1}{2}+\kappa+\frac{1}{\alpha_y}\right)-\frac{1}{\alpha_y}}\tilde{l}_y\left(\frac{kd}{n}\right)$$

Since $\zeta < \dfrac{2}{\alpha_y + 3}$, $0 < \kappa < 1/2\,\alpha_y$, we have $\zeta\left(\dfrac{1}{2}+\kappa+\dfrac{1}{\alpha_y}\right) - \dfrac{1}{\alpha_y} < 0$, together

with $\dfrac{d}{m}\to 0$, as $n\to\infty$, we have

$$\delta_n Q_Y\left(\frac{d}{m}\right)\to\infty. \qquad (4.4)$$

Then we prove (4.1) first: Notice that

$$P\left(\varepsilon > \delta_n Q_Y\left(\frac{d}{m}\right)\right) = \frac{P\left(\varepsilon > \delta_n Q_Y\left(\frac{d}{m}\right), X > Q_X(\bar{p})\right)}{P\left(X > Q_X(\bar{p})\right)}$$

$$\leq \frac{P\left(Y > \delta_n Q_Y\left(\frac{d}{m}\right) + \beta Q_X(\bar{p})\right)}{\bar{p}}$$

$$\leq \frac{P\left(Y > \delta_n Q_Y\left(\frac{d}{m}\right)\right)}{P\left(Y > Q_Y\left(\frac{d}{m}\right)\right)}\cdot\frac{\frac{d}{m}}{\bar{p}}$$

$$\sim \delta_n^{-\alpha_y}\frac{\frac{d}{m}}{\bar{p}}.$$

The penultimate step is based on (4.4). As $n \to \infty$, since $\sqrt{kd}\delta_n^{-\alpha_y}\dfrac{d}{m} \to 0$, then,

$$\sqrt{kd}\,P\left(\varepsilon > \delta_n Q_Y\left(\frac{d}{m}\right)\right) \to 0.$$

Next, we prove (4.2): for some $D > 0$, we write

$$
\begin{aligned}
P\left(\varepsilon \le -\delta_n Q_Y\left(\frac{d}{m}\right)\right) &= \frac{P\left(\varepsilon \le -\delta_n Q_Y\left(\frac{d}{m}\right), Q_X(\overline{p}) \le X \le \frac{1}{\beta+1}\delta_n Q_Y\left(\frac{d}{m}\right)\right)}{P\left(Q_X(\overline{p}) \le X \le \frac{1}{\beta+1}\delta_n Q_Y\left(\frac{d}{m}\right)\right)} \\[2mm]
&\le \frac{P\left(Y \le -\frac{1}{\beta+1}\delta_n Q_Y\left(\frac{d}{m}\right)\right)}{\overline{p} - P\left(X \ge \frac{1}{\beta+1}\delta_n Q_Y\left(\frac{d}{m}\right)\right)} \\[2mm]
&\le D\,\frac{P\left(Y \ge \frac{1}{\beta+1}\delta_n Q_Y\left(\frac{d}{m}\right)\right)}{\overline{p} - P\left(X \ge \frac{1}{\beta+1}\delta_n Q_Y\left(\frac{d}{m}\right)\right)},
\end{aligned}
$$

The last step uses the condition that $P(Y < -u) = O(P(Y > u))$. By (4.4), the denominator converges to $\overline{p}$, which is positive and finite. Same as (4.1), we have

$$\sqrt{kd}\,P\left(\varepsilon \le -\delta_n Q_Y\left(\frac{d}{m}\right)\right) \to 0.$$

Then, we prove (4.3): by Lemme (4.1), we know that $\sqrt{kd}\,a_x\left(Q_X\left(\frac{d}{m}\right)\right) \to 0$. Recalling the second-order condition (1.4), we have

$$\lim_{u \to \infty} \sqrt{kd}\left(\frac{P(X > ux)}{P(X > u)} - x^{-\alpha_x}\right) = 0,$$

substituting ux and u by $Q_Y\left(\dfrac{d}{m}\right)(1 \pm \delta_n)/\beta$ and $Q_X\left(\dfrac{d}{m}\right)$, together with the

fact that $\dfrac{Q_Y\left(\dfrac{d}{m}\right)(1 \pm \delta_n)}{\beta Q_X\left(\dfrac{d}{m}\right)} \to \tau^{-\frac{1}{\alpha_x}}$, we get that

$$\lim_{n \to \infty} \sqrt{kd}\left(\frac{m}{d}P\left(\beta X > Q_Y\left(\frac{d}{m}\right)(1 \pm \delta_n)\right) - \left(\frac{Q_Y\left(\frac{d}{m}\right)(1 \pm \delta_n)}{\beta Q_X\left(\frac{d}{m}\right)}\right)^{-\alpha_x}\right) = 0,$$

Compared with (4.3), since

$$\lim_{n\to\infty}\sqrt{kd}\left(\left(\frac{Q_Y\left(\frac{d}{m}\right)(1\pm\delta_n)}{\beta Q_X\left(\frac{d}{m}\right)}\right)^{-\alpha_x}-\left(\frac{Q_Y\left(\frac{d}{m}\right)}{\beta Q_X\left(\frac{d}{m}\right)}\right)^{-\alpha_x}\right)$$

$$=\lim_{n\to\infty}\left(\frac{Q_Y\left(\frac{d}{m}\right)}{\beta Q_X\left(\frac{d}{m}\right)}\right)^{-\alpha_x}\cdot\sqrt{kd}\left[(1\pm\delta_n)^{-\alpha_x}-1\right]$$

$$=\lim_{n\to\infty}\left(\frac{Q_Y\left(\frac{d}{m}\right)}{\beta Q_X\left(\frac{d}{m}\right)}\right)^{-\alpha_x}\cdot\sqrt{kd}\cdot\delta_n\cdot(-\alpha_x)\cdot x^{-\alpha_x-1}$$

$$=0.$$

The penultimate step uses Lagrange's mean value theorem, where x is between 1 and $1\pm\delta_n$. Hence, (4.3) is proved since $\lim_{n\to\infty}\sqrt{kd}\delta_n=0$. $\square$

Proof of Theorem 2.1. Since

$$\hat{\beta}_D=\frac{1}{k}\sum_{j=1}^{k}\hat{\beta}_{n,j}=\frac{1}{k}\sum_{j=1}^{k}\left[\left(\hat{\tau}_j\left(\frac{d}{m}\right)\right)^{\frac{1}{\hat{\alpha}_j}}\frac{\hat{Q}_j^Y\left(\frac{d}{m}\right)}{\hat{Q}_j^X\left(\frac{d}{m}\right)}\right]$$

$$=\frac{1}{k}\sum_{j=1}^{k}\left[\left(\frac{\hat{\tau}_j\left(\frac{d}{m}\right)}{\tau\left(\frac{d}{m}\right)}\right)^{\frac{1}{\hat{\alpha}_j}}\times\left(\tau\left(\frac{d}{m}\right)\right)^{\frac{1}{\hat{\alpha}_j}-\frac{1}{\alpha_x}}\times\frac{\hat{Q}_j^Y\left(\frac{d}{m}\right)}{Q_Y\left(\frac{d}{m}\right)}\times\frac{Q_X\left(\frac{d}{m}\right)}{\hat{Q}_j^X\left(\frac{d}{m}\right)}\times\left[\left(\tau\left(\frac{d}{m}\right)\right)^{\frac{1}{\alpha_x}}\frac{Q_Y\left(\frac{d}{m}\right)}{Q_X\left(\frac{d}{m}\right)}\right]\right] \quad (4.5)$$

$$:=\frac{1}{k}\sum_{j=1}^{k}\left[I_{j,1}\cdot I_{j,2}\cdot I_{j,3}\cdot I_{j,4}\cdot I_{j,5}\right],$$

where,

$$I_{j,1}=\left(\frac{\hat{\tau}_j\left(\frac{d}{m}\right)}{\tau\left(\frac{d}{m}\right)}\right)^{\frac{1}{\hat{\alpha}_j}};\ I_{j,2}=\left(\tau\left(\frac{d}{m}\right)\right)^{\frac{1}{\hat{\alpha}_j}-\frac{1}{\alpha_x}};\ I_{j,3}=\frac{\hat{Q}_j^Y\left(\frac{d}{m}\right)}{Q_Y\left(\frac{d}{m}\right)};$$

$$I_{j,4}=\frac{Q_X\left(\frac{d}{m}\right)}{\hat{Q}_j^X\left(\frac{d}{m}\right)};\ I_{j,5}=\left(\tau\left(\frac{d}{m}\right)\right)^{\frac{1}{\alpha_x}}\frac{Q_Y\left(\frac{d}{m}\right)}{Q_X\left(\frac{d}{m}\right)},$$

here we show that $I_{j,1}=1+o_p(1)$, $I_{j,2}=1+o_p(1)$, $I_{j,3}=1+o_p(1)$, $I_{j,4}=1+o_p(1)$, $I_{j,5}=\beta+o_p(1)$ uniformly for $j=1,2,\cdots,k$ separately.

We first deal with $I_{j,1}$. For (x,y) in the neighborhood of (1, 1), denote

$$\tilde{\tau}_j(x,y)=\frac{1}{d}\sum_{t=1}^{m}1_{\left\{X_j^t>Q_X\left(\frac{d}{m}x\right),Y_j^t>Q_Y\left(\frac{d}{m}y\right)\right\}},$$

then $\hat{\tau}_j(d/m)$ can be written as

$$\hat{\tau}_j\left(\frac{d}{m}\right) = \tilde{\tau}_j\left(\frac{m}{d}\left(1 - F_X\left(X_j^{(d+1)}\right)\right), \frac{m}{d}\left(1 - F_Y\left(Y_j^{(d+1)}\right)\right)\right),$$

and $\left(\frac{m}{d}\left(1 - F_X\left(X_j^{(d+1)}\right)\right), \frac{m}{d}\left(1 - F_Y\left(Y_j^{(d+1)}\right)\right)\right)$ is in the neighborhood of $(1, 1)$.

According to Corollary 2.2.2 in de Haan and Ferrira (2006) [21], as $n \to \infty$,

$$\sqrt{d}\left(\frac{m}{d}Z_{m-d,m} - 1\right) \sim N(0,1),$$

where $F_Z(a) = 1 - 1/a$, $Z_{1,m} \leq Z_{2,m} \leq \cdots \leq Z_{m,m}$. Combining with

$$Z_{m-d,m} \stackrel{d}{=} \frac{1}{1 - F_X\left(X_j^{(d+1)}\right)},$$ we have $\sqrt{d}\left(\frac{m}{d}\frac{1}{1 - F_X\left(X_j^{(d+1)}\right)} - 1\right) \sim N(0,1)$, then, as

$n \to \infty$,

$$\sqrt{d}\left(\frac{m}{d}\left(1 - F_X\left(X_j^{(d+1)}\right)\right) - 1\right) \sim N(0,1).$$

Hence, for any $\delta > 0$, as $n \to \infty$, we have

$$P\left(\left|\frac{m}{d}\left(1 - F_X\left(X_j^{(d+1)}\right)\right) - 1\right| > d^{-\frac{1}{2}+\delta}\right) \to 0.$$

A similar relation for Y holds. Therefore, in order to prove that $I_{j,1} \stackrel{P}{\to} 1$, we will prove a more general result that

$$\frac{\tilde{\tau}_j(x,y)}{\tau\left(\frac{d}{m}\right)} \stackrel{P}{\to} 1$$

uniformly for all $j = 1, 2, \cdots, k$, $(x, y) \in \left[1 - d^{-1/2+\delta}, 1 + d^{1/2+\delta}\right]$ for some $0 < \delta < 1/2$.

Since the tail dependence function

$$R(x, y, d/m) = \frac{m}{d}P\left(X > Q_X\left(\frac{d}{m}x\right), Y > Q_Y\left(\frac{d}{m}y\right)\right),$$ denote

$$\xi_{n,j}(x,y) = \frac{1}{m}\sum_{t=1}^{m}1_{\left\{X_j^t > Q_X\left(\frac{d}{m}x\right), Y_j^t > Q_Y\left(\frac{d}{m}y\right)\right\}},$$

since the observed values of different machines are independent and identically distributed, we have

$$\frac{\tilde{\tau}_j(x,y)}{R\left(x, y, \frac{d}{m}\right)} = \frac{\frac{1}{d}\sum_{t=1}^{m}1_{\left\{X_j^t > Q_X\left(\frac{d}{m}x\right), Y_j^t > Q_Y\left(\frac{d}{m}y\right)\right\}}}{\frac{m}{d}P\left(X > Q_X\left(\frac{d}{m}x\right), Y > Q_Y\left(\frac{d}{m}y\right)\right)}$$

$$= \frac{\frac{1}{m}\sum_{t=1}^{m}1_{\left\{X_j^t > Q_X\left(\frac{d}{m}x\right), Y_j^t > Q_Y\left(\frac{d}{m}y\right)\right\}}}{P\left(X > Q_X\left(\frac{d}{m}x\right), Y > Q_Y\left(\frac{d}{m}y\right)\right)}$$

$$= \frac{\xi_{n,j}(x,y)}{E\xi_{n,j}(x,y)}.$$

Applying Chebyshev's inequality, as $n \to \infty$, we have

$$P\left(\left|\frac{\xi_{n,j}(x,y)}{E\xi_{n,j}(x,y)} - 1\right| > \varepsilon\right)$$

$$= P\left(\left|\xi_{n,j}(x,y) - E\xi_{n,j}(x,y)\right| > \varepsilon\left|E\xi_{n,j}(x,y)\right|\right) \le \frac{Var\xi_{n,j}(x,y)}{\varepsilon^2\left|E\xi_{n,j}(x,y)\right|^2}$$

$$= \frac{\frac{1}{m}P\left(X > Q_X\left(\frac{d}{m}x\right), Y > Q_Y\left(\frac{d}{m}y\right)\right)\left(1 - P\left(X > Q_X\left(\frac{d}{m}x\right), Y > Q_Y\left(\frac{d}{m}y\right)\right)\right)}{\varepsilon^2 \cdot \left[P\left(X > Q_X\left(\frac{d}{m}x\right), Y > Q_Y\left(\frac{d}{m}y\right)\right)\right]^2}$$

$$= \frac{1 - P\left(X > Q_X\left(\frac{d}{m}x\right), Y > Q_Y\left(\frac{d}{m}y\right)\right)}{d\varepsilon^2 \cdot \frac{m}{d}P\left(X > Q_X\left(\frac{d}{m}x\right), Y > Q_Y\left(\frac{d}{m}y\right)\right)}$$

$$\xrightarrow{P} \frac{1 - P\left(X > Q_X\left(\frac{d}{m}x\right), Y > Q_Y\left(\frac{d}{m}y\right)\right)}{d\varepsilon^2\tau} \xrightarrow{P} 0.$$

The penultimate step using the convergence of $R(x,y,p)$, that is (1.9), then

as $n \to \infty$, $\dfrac{\tilde{\tau}_j(x,y)}{R\left(x,y,\frac{d}{m}\right)} \xrightarrow{P} 1$.

Hence, what remains to be proved is that $\lim\limits_{n\to\infty} \dfrac{R\left(x,y,\frac{d}{m}\right)}{\tau\left(\frac{d}{m}\right)} = 1$ holds uniform-

ly for all $(x,y) \in \left[1 - d^{-1/2+\delta}, 1 + d^{-1/2+\delta}\right]$. If $\beta = 0$, as $n \to \infty$,

$$\lim_{n\to\infty} \frac{R\left(x,y,\frac{d}{m}\right)}{\tau\left(\frac{d}{m}\right)} = \lim_{n\to\infty} \frac{\frac{m}{d} \cdot \frac{d}{m}x \cdot \frac{d}{m}y}{\frac{m}{d} \cdot \frac{d}{m} \cdot \frac{d}{m}} = 1.$$

If $\beta > 0$, applying Lemma 1 in Oordt and Zhou (2017) [13] with $p = d/m$ directly gives that

$$\lim_{n\to\infty} \frac{R\left(x,y,\frac{d}{m}\right)}{\frac{m}{d}P\left(X > \max\left(\frac{Q_Y\left(\frac{d}{m}y\right)}{\beta}, Q_X\left(\frac{d}{m}x\right)\right)\right)} = 1$$

holds uniformly for all $(x,y) \in \left[1 - d^{-1/2+\delta}, 1 + d^{-1/2+\delta}\right]$. We further simplify the denominator as follows: as $n \to \infty$,

$$\frac{m}{d} P\left(X > \max\left(\frac{Q_Y\left(\frac{d}{m}y\right)}{\beta}, Q_X\left(\frac{d}{m}x\right)\right)\right)$$

$$= \frac{P\left(X > \max\left(\frac{Q_Y\left(\frac{d}{m}y\right)}{\beta}, Q_X\left(\frac{d}{m}x\right)\right)\right)}{P\left(X > Q_X\left(\frac{d}{m}\right)\right)}$$

$$\sim \min\left(\left(\frac{\beta Q_X\left(\frac{d}{m}\right)}{Q_Y\left(\frac{d}{m}y\right)}\right)^{\alpha_x}, \left(\frac{Q_X\left(\frac{d}{m}\right)}{Q_X\left(\frac{d}{m}x\right)}\right)^{\alpha_x}\right),$$

the last step uses the second order condition of X, that is (1.4).

From (1.5), as $n \to \infty$, we get that $\left(\frac{\beta Q_X\left(\frac{d}{m}\right)}{Q_Y\left(\frac{d}{m}y\right)}\right)^{\alpha_x} \sim \tau\left(\frac{d}{m}\right)$ holds uniformly

for $|y - 1| \le d^{-1/2+\delta}$. In addition, as $n \to \infty$, $\frac{Q_X\left(\frac{d}{m}\right)}{Q_X\left(\frac{d}{m}x\right)} \xrightarrow{P} 1$ holds uniformly for

$|x - 1| \le d^{-1/2+\delta}$. Combine with $\tau\left(\frac{d}{m}\right) \le 1$, we get that

$$\lim_{n\to\infty} \frac{m}{d} \cdot P\left(X > \max\left(\frac{Q_Y\left(\frac{d}{m}y\right)}{\beta}, Q_X\left(\frac{d}{m}x\right)\right)\right) = \tau\left(\frac{d}{m}\right)$$

holds uniformly for $(x, y) \in \left[1 - d^{-1/2+\delta}, 1 + d^{-1/2+\delta}\right]$, as $n \to \infty$. Hence, we

proved that $\frac{\hat{\tau}_j\left(\frac{d}{m}\right)}{\tau\left(\frac{d}{m}\right)} \xrightarrow{P} 1$, together with the consistency of $1/\hat{\alpha}_j$, we have

$$I_{j,1} = \left(\frac{\hat{\tau}_j\left(\frac{d}{m}\right)}{\tau\left(\frac{d}{m}\right)}\right)^{\frac{1}{\hat{\alpha}_j}} \xrightarrow{P} 1 \text{ uniformly for } j = 1, 2, \cdots, k, \text{ thus, } I_{j,1} = 1 + o_p(1).$$

Next, we deal with $I_{j,2} = \left(\tau\left(\frac{d}{m}\right)\right)^{\frac{1}{\hat{\alpha}_j} - \frac{1}{\alpha_x}}$. Note that the observations of different machines are independently and identically distributed, similar to the proof in Oordt and Zhou (2017) [13], if $\limsup_{n\to\infty} \tau\left(\frac{d}{m}\right) > 0$, then the consistency of

$\hat{\alpha}_j$ leads to $I_{j,2} \xrightarrow{P} 1$, as $n \to \infty$; If $\limsup\limits_{n\to\infty} \tau\left(\dfrac{d}{m}\right) = 0$, to prove that as $n \to \infty$,

there is $I_2 \xrightarrow{P} 1$, equivalent to prove

$$\left(\frac{1}{\hat{\alpha}_j} - \frac{1}{\alpha_x}\right) \log \tau\left(\frac{d}{m}\right) \xrightarrow{P} 0.$$

Theorem 3.2.5 in de Haan and Ferrira (2006) [21] guarantees the asymptotic normality of $\hat{\alpha}_j$ under conditions (1.3) and (1.13): as $n \to \infty$,

$$\sqrt{d}\left(\frac{1}{\hat{\alpha}_j} - \frac{1}{\alpha_x}\right) \xrightarrow{d} N\left(\frac{1}{\alpha_x(1-\rho)}, \left(\frac{1}{\alpha_x}\right)^2\right),$$

that is,

$$\sqrt{d}\left(\frac{1}{\hat{\alpha}_j} - \frac{1}{\alpha_x}\right) \xrightarrow{d} O_p(1),$$

therefore, it only remains to prove that $\log \tau\left(\dfrac{d}{m}\right) = o\left(\sqrt{d}\right)$.

If $\beta = 0$, $\tau\left(\dfrac{d}{m}\right) = \dfrac{d}{m}$, by (1.13), $\dfrac{\sqrt{d}}{\log m} \to +\infty$, then,

$$\log \tau\left(\frac{d}{m}\right) = \log d - \log m$$
$$= \sqrt{d}\left(\frac{\log d}{\sqrt{d}} - \frac{\log m}{\sqrt{d}}\right)$$
$$= o\left(\sqrt{d}\right).$$

If $\beta > 0$, by (1.5), for sufficiently large n, we have

$$\tau\left(\frac{d}{m}\right) \sim \left(\frac{\beta Q_x\left(\frac{d}{m}\right)}{Q_y\left(\frac{d}{m}\right)}\right)^{\alpha_x}$$
$$= \beta^{\alpha_x}\left(\frac{d}{m}\right)^{\frac{\alpha_x}{\alpha_y}-1}\left(\frac{\tilde{l}_x\left(\frac{d}{m}\right)}{\tilde{l}_y\left(\frac{d}{m}\right)}\right)^{\alpha_x}$$
$$> D\left(\frac{d}{m}\right)^{\frac{\alpha_x}{\alpha_y}-1+\delta},$$

for some $D > 0$ and $\delta > 0$, the last step uses the Potter inequality. Therefore, $\log \tau\left(\dfrac{d}{m}\right) = o\left(\sqrt{d}\right)$. Thus, $I_{j,2} \xrightarrow{P} 1$, $I_{j,2} = 1 + o_p(1)$ uniformly for $j = 1, 2, \cdots, k$.

For $I_{j,3}$, according to Theorem 2.2.1 in de Haan and Ferrira (2006) [21], as $n \to \infty$, for $j = 1, 2, \cdots, k$, we have

$$\sqrt{d}\left(\frac{Y_j^{(d+1)}-Q_Y\left(\dfrac{d}{m}\right)}{\dfrac{d}{m}Q_Y'\left(\dfrac{d}{m}\right)}\right)\xrightarrow{d}N(0,1),$$

then

$$\sqrt{d}\left(\frac{Y_j^{(d+1)}-Q_Y\left(\dfrac{d}{m}\right)}{\dfrac{d}{m}Q_Y'\left(\dfrac{d}{m}\right)}\right)=O_p(1).$$

Since $Q_Y(p)=p^{-\frac{1}{\alpha_y}}\tilde{l}_y(p)$,

$$\lim_{p\to0}\frac{pQ_Y'(p)}{Q_Y(p)}=\lim_{p\to0}\frac{p\left(-\dfrac{1}{\alpha_y}\cdot p^{-\frac{1}{\alpha_y}-1}\tilde{l}_y(p)+p^{-\frac{1}{\alpha_y}}\tilde{l}_y'(p)\right)}{p^{-\frac{1}{\alpha_y}}\tilde{l}_y(p)}=\lim_{p\to0}-\frac{1}{\alpha_y}+p\cdot\frac{\tilde{l}_y'(p)}{\tilde{l}_y(p)}=-\frac{1}{\alpha_y}.$$

The last step exploits the properties of slowly varying functions, then we have

$$\frac{Y_j^{(d+1)}}{Q_Y\left(\dfrac{d}{m}\right)}\xrightarrow{P}1-\frac{1}{\sqrt{d}}\frac{1}{\alpha_y}O_p(1),$$

then $I_{j,3}=1-o_p(1)$ uniformly for $j=1,2,\cdots,k$.

For $I_{j,4}$, same as $I_{j,3}$, we know $\dfrac{X_j^{(d+1)}}{Q_X\left(\dfrac{d}{m}\right)}\xrightarrow{P}1-\dfrac{1}{\sqrt{d}}\dfrac{1}{\alpha_x}O_p(1)$, then

$$I_{j,4}=\frac{Q_X\left(\dfrac{d}{m}\right)}{X_j^{(d+1)}}=1-o_p(1).$$

Finally, according to (1.5), $I_{j,5}=\left(\tau\left(\dfrac{d}{m}\right)\right)^{\frac{1}{\alpha_x}}\dfrac{Q_Y\left(\dfrac{d}{m}\right)}{Q_X\left(\dfrac{d}{m}\right)}\to\beta$, then

$I_{j,5}=\beta+o_p(1)$.

From the above analysis,

$$\hat{\beta}_D\xrightarrow{P}\beta.$$

$\square$

Proof of Theorem 2.2. From (4.5), $\hat{\beta}_{n,j}=I_{j,1}\cdot I_{j,2}\cdot I_{j,3}\cdot I_{j,4}\cdot I_{j,5}$, let's analyze $I_{j,1},I_{j,2},I_{j,3},I_{j,4},I_{j,5}$ separately. Deal with $I_{j,1}$ first. By definition, $R(x,y)$ is a homogeneous function of the first degree. According to Lemma 2 in Oordt and Zhou (2017) [13], for $x>0$, we have $R(x,1)=\min(x,\tau)$. Thus, for $\dfrac{x}{y}>\tau$,

$R(x,y) = yR\left(\dfrac{x}{y}, 1\right) = \tau y$; for $\dfrac{x}{y} < \tau$, $R(x,y) = yR\left(\dfrac{x}{y}, 1\right) = x$, hence, the partial derivatives of R at the neighborhood of (1, 1) exist as $R_1(1,1) = 0$ and $R_2(1,1) = \tau$, where R_1, R_2 denotes the partial derivatives of R with respect to x and y, respectively. Due to tail stable dependency function $l(x,y) = x + y - R(x,y)$, Theorem 2 in Section 2 of Huang (1992) [22] gives the asymptotic normality of $\hat{l}(x,y)$ as $m \to \infty$, $d = o\left(m^{\frac{2\theta}{1+2\theta}}\right)$, we have

$$\sqrt{d}\left(\hat{l}(x,y) - l(x,y)\right) \xrightarrow{d} B(x,y),$$

where

$$B(x,y) := W(x,y) - l_1(x,y)W(x,0) - l_2(x,y)W(0,y),$$

$$l_1(x,y) = \frac{\partial}{\partial x}l(x,y), l_2(x,y) = \frac{\partial}{\partial y}l(x,y),$$

$W(x,y)$ is a continuous zero-mean Gaussian process, its covariance is

$$EW(x_1,y_1)W(x_2,y_2)$$
$$= l(x_1 \wedge x_2, y_1) + l(x_1 \wedge x_2, y_2) + l(x_1, y_1 \wedge y_2) + l(x_2, y_1 \wedge y_2)$$
$$- l(x_1, y_2) - l(x_2, y_1) - l(x_1 \wedge x_2, y_1 \wedge y_2),$$

and $l_1(1,1) = 1 - R_1(1,1) = 1$; $l_2(1,1) = 1 - R_2(1,1) = 1 - \tau$; $\tau = R(1,1) = 2 - l(1,1)$, then we have

$$\sqrt{d}\left(\hat{\tau}_j\left(\frac{d}{m}\right) - \tau\left(\frac{d}{m}\right)\right) \xrightarrow{d} (1 - \tau)W_j(0,1) + W_j(1,0) - W_j(1,1).$$

Let

$$S_{j,1} = \sqrt{d}\left(\left(\frac{\hat{\tau}_j\left(\frac{d}{m}\right)}{\tau\left(\frac{d}{m}\right)}\right)^{\frac{1}{\hat{\alpha}_j}} - 1\right),$$

then

$$S_{j,1} = \tau\left(\frac{d}{m}\right)^{-\frac{1}{\hat{\alpha}_j}}\left[\sqrt{d}\left(\hat{\tau}_j\left(\frac{d}{m}\right)^{\frac{1}{\hat{\alpha}_j}} - \tau\left(\frac{d}{m}\right)^{\frac{1}{\hat{\alpha}_j}}\right)\right]$$

$$= \tau\left(\frac{d}{m}\right)^{-\frac{1}{\hat{\alpha}_j}}\frac{1}{\hat{\alpha}_j}\tilde{\tau}\left(\frac{d}{m}\right)^{\frac{1}{\hat{\alpha}_j}-1}\left[\sqrt{d}\left(\hat{\tau}_j\left(\frac{d}{m}\right) - \tau\left(\frac{d}{m}\right)\right)\right]$$

$$\xrightarrow{d} \frac{1}{\alpha_x}\frac{1}{\tau}\left[(1 - \tau)W_j(0,1) + W_j(1,0) - W_j(1,1)\right],$$

the second step uses the Delta method, where $\tilde{\tau}\left(\dfrac{d}{m}\right)$ is between $\tau\left(\dfrac{d}{m}\right)$ and $\hat{\tau}_j\left(\dfrac{d}{m}\right)$. Thus, $I_{j,1} = \dfrac{1}{\sqrt{d}}S_{j,1} + 1$, and

$$S_{j,1} \xrightarrow{d} \frac{1}{\alpha_x} \frac{1}{\tau} \big[(1-\tau) W_j(0,1) + W_j(1,0) - W_j(1,1) \big].$$ (4.6)

Next we deal $I_{j,2} = \left(\tau \left(\frac{d}{m} \right) \right)^{\frac{1}{\hat{\alpha}_j} - \frac{1}{\alpha_x}}$. According to Theorem 3.2.5 in de Haan and Ferrira (2006) [21], we know that Gaussian process can also control the convergence of the tail exponents, *i.e.*,

$$\sqrt{d} \left(\frac{1}{\hat{\alpha}_j} - \frac{1}{\alpha_x} \right) = \frac{1}{\alpha_x} \left(\int_0^1 W_j(s,0) \frac{ds}{s} - W_j(1,0) \right) + \sqrt{d} A_0 \left(\frac{m}{d} \right) \frac{1}{1-\rho} + o_p(1),$$

where $W_j(x,y)$ is the same zero-mean Gaussian process as above, as $n \to \infty$,

$A_x \left(\frac{m}{d} \right) \sim A_0 \left(\frac{m}{d} \right)$. Use Delta method, let $g(x) = \left(\tau \left(\frac{d}{m} \right) \right)^{x - \frac{1}{\alpha_x}}$,

$g'(x) = \log \tau \left(\tau \left(\frac{d}{m} \right) \right)^{x - \frac{1}{\alpha_x}}$, $g' \left(\frac{1}{\alpha_x} \right) = \log \tau$, we have

$$\sqrt{d} \left(\left(\tau \left(\frac{d}{m} \right) \right)^{\frac{1}{\hat{\alpha}_j} - \frac{1}{\alpha_x}} - 1 \right)$$

$$\xrightarrow{P} \left((\log \tau) \frac{1}{\alpha_x} \right) \left(\int_0^1 W_j(s,0) \frac{ds}{s} - W_j(1,0) \right) + \log \tau \sqrt{d} A_0 \left(\frac{m}{d} \right) \frac{1}{1-\rho}.$$

From Lemma 4.1, as $n \to \infty$, $\sqrt{d} A_0 \left(\frac{m}{d} \right) \xrightarrow{P} 0$, then,

$$\sqrt{d} \left(\left(\tau \left(\frac{d}{m} \right) \right)^{\frac{1}{\hat{\alpha}_j} - \frac{1}{\alpha_x}} - 1 \right) \xrightarrow{d} \left((\log \tau) \frac{1}{\alpha_x} \right) \left(\int_0^1 W_j(s,0) \frac{ds}{s} - W_j(1,0) \right).$$

Let

$$S_{j,2} = \sqrt{d} \left(\left(\tau \left(\frac{d}{m} \right) \right)^{\frac{1}{\hat{\alpha}_j} - \frac{1}{\alpha_x}} - 1 \right),$$

thus, $I_{j,2} = \frac{1}{\sqrt{d}} S_{j,2} + 1$ and

$$S_{j,2} \xrightarrow{d} \left((\log \tau) \frac{1}{\alpha_x} \right) \left(\int_0^1 W_j(s,0) \frac{ds}{s} - W_j(1,0) \right).$$ (4.7)

For $I_{j,3} = \frac{Y_j^{(d+1)}}{Q_Y \left(\frac{d}{m} \right)}$. According Theorem 4 in Chen *et al.* (2021) [23], we know that

$$\sqrt{d} \left(\frac{Y_j^{(d+1)}}{Q_Y \left(\frac{d}{m} \right)} - 1 \right) = \frac{1}{\alpha_y} W_j(0,1) + o_p(1),$$

let $S_{j,3} = \sqrt{d}\left(\dfrac{Y_j^{(d+1)}}{Q_Y\left(\dfrac{d}{m}\right)} - 1\right)$, then $I_{j,3} = \dfrac{1}{\sqrt{d}}S_{j,3} + 1$, and

$$S_{j,3} \xrightarrow{d} \frac{1}{\alpha_y}W_j(0,1). \tag{4.8}$$

Finally, we deal with $I_{j,4} = \dfrac{Q_X\left(\dfrac{d}{m}\right)}{X_j^{(d+1)}}$. Similar to $I_{j,3}$, we know that

$$\sqrt{kd}\left(\frac{X_j^{(d+1)}}{Q_X\left(\dfrac{d}{m}\right)} - 1\right) = \sqrt{k}\frac{1}{\alpha_x}W_j(1,0) + o_p(1),$$

use Delta method, let $g(x) = \dfrac{1}{x}$, then $g'(x) = -\dfrac{1}{x^2}$, $g'(1) = -1$,

$$\sqrt{d}\left(\frac{Q_X\left(\dfrac{d}{m}\right)}{X_j^{(d+1)}} - 1\right) \xrightarrow{P} -\frac{1}{\alpha_x}W_j(1,0).$$

Let $S_{j,4} = \sqrt{d}\left(\dfrac{Q_X\left(\dfrac{d}{m}\right)}{X_j^{(d+1)}} - 1\right)$, then $I_{j,4} = \dfrac{1}{\sqrt{d}}S_{j,4} + 1$ and

$$S_{j,4} \xrightarrow{d} -\frac{1}{\alpha_x}W_j(1,0). \tag{4.9}$$

Thus,

$$\sqrt{kd}\left(\frac{\hat{\beta}_D}{\tau^{\frac{1}{\alpha_x}}\dfrac{Q_Y\left(\dfrac{d}{m}\right)}{Q_X\left(\dfrac{d}{m}\right)}} - 1\right) = \frac{1}{\sqrt{k}}\sum_{j=1}^{k}\left[\left(S_{j,1} + S_{j,2} + S_{j,3} + S_{j,4}\right) + o_p(1)\right] = \frac{1}{\sqrt{k}}\sum_{j=1}^{k}S_j.$$

where $S_j = S_{j,1} + S_{j,2} + S_{j,3} + S_{j,4} + o_p(1)$, combine (4.6), (4.7), (4.8) and (4.9), we have

$$S_j \xrightarrow{d} \frac{1}{\alpha_x}\frac{1}{\tau}\left\{(1-\tau)W_j(0,1) + W_j(1,0) - W_j(1,1)\right\}$$
$$+ \frac{\log\tau}{\alpha_x}\left(\int_0^1 W_j(s,0)\frac{\mathrm{d}s}{s} - W_j(1,0)\right) + \frac{1}{\alpha_y}W_j(0,1) - \frac{1}{\alpha_x}W_j(1,0).$$

Based on the expression for $l(x,y)$, we get $ES_j \to 0, VarS_j \to \Sigma$, where

$$\Sigma = \frac{1}{\alpha_x^2}\left(\frac{1}{\tau} + 2\log\tau - \frac{2}{\tau}\cdot\log\tau - 1 - 3\left(\log\tau\right)^2\right).$$

And S_j is independent and identical distributed on different machines, use the Central Limit Theorem,

$$\sqrt{kd}\left(\frac{\hat{\beta}_D}{\tau^{\frac{1}{\alpha_x}}\frac{Q_Y\left(\frac{d}{m}\right)}{Q_X\left(\frac{d}{m}\right)}} - 1\right) \xrightarrow{d} N\left(0,\Sigma\right),$$

where,

$$\Sigma = \frac{1}{\alpha_x^2}\left(\frac{1}{\tau} + 2\log\tau - \frac{2}{\tau}\log\tau - 1 - 3\left(\log\tau\right)^2\right).$$

Therefore, what remains to be proved is the following deterministic relation

$$\lim_{n\to\infty}\sqrt{kd}\left(\frac{\tau^{\frac{1}{\alpha_x}}\frac{Q_Y\left(\frac{d}{m}\right)}{Q_X\left(\frac{d}{m}\right)}}{\beta} - 1\right) = 0.$$

According to (1.5), we know that $\lim_{n\to\infty}\dfrac{Q_y\left(\frac{d}{m}\right)}{Q_x\left(\frac{d}{m}\right)} = \beta\cdot\tau^{-\frac{1}{\alpha_x}} > 0$, use Delta method, the above relation is equivalent to

$$\lim_{n\to\infty}\sqrt{kd}\left(\tau - \left(\beta\frac{Q_X\left(\frac{d}{m}\right)}{Q_Y\left(\frac{d}{m}\right)}\right)^{\alpha_x}\right) = 0.$$

Next, from (1.8) and (1.9), we have $\tau(p) - \tau = O\left(p^\theta\right)$, combine with $kd = O\left(n^\zeta\right)$, $\zeta < 2\theta/1 + 2\theta$, we get that

$$\lim_{n\to\infty}\sqrt{kd}\left(\tau - \tau\left(\frac{d}{m}\right)\right) = \lim_{n\to\infty}\left(kd\right)^{\frac{1}{2}}\left(\frac{d}{m}\right)^\theta = \lim_{n\to\infty}n^{\xi\left(\frac{1}{2}+\theta\right)-\theta} = 0,$$

what remains to be proved is

$$\lim_{n\to\infty}\sqrt{kd}\left(\tau\left(\frac{d}{m}\right) - \left(\beta\frac{Q_X\left(\frac{d}{m}\right)}{Q_Y\left(\frac{d}{m}\right)}\right)^{\alpha_x}\right) = 0. \tag{4.10}$$

Notice that Lemma 1 in Oordt and Zhou (2017) [13] can be written as

$$\lim_{p \to 0} \frac{P\left(X > Q_X\left(\frac{d}{m}\right), Y > Q_Y\left(\frac{d}{m}\right)\right)}{P\left(X > \max\left(\frac{Q_Y\left(\frac{d}{m}\right)}{\beta}, Q_X\left(\frac{d}{m}\right)\right)\right)} = 1, \tag{4.11}$$

when $p = \dfrac{d}{m}$ and $x = y = 1$. And by (1.5), we have

$\lim\limits_{p \to 0} \dfrac{Q_Y(p)}{Q_X(\tau p)} = \lim\limits_{p \to 0} \tau^{\frac{1}{\alpha_x}} \cdot \dfrac{Q_Y(p)}{Q_x(p)} = \beta$, then, for any $\tau < z < 1$, for sufficiently small

p, we have $Q_y(p) \geq \beta \cdot Q_X(zp)$, then, for sufficiently large n,

$Q_X\left(\dfrac{d}{m}\right) < Q_X\left(\dfrac{d}{m} z\right) < Q_Y\left(\dfrac{d}{m}\right)\Big/\beta$. Thus, (4.11) is equal to

$$\lim_{p \to 0} \frac{P\left(X > Q_X\left(\frac{d}{m}\right), Y > Q_Y\left(\frac{d}{m}\right)\right)}{P\left(X > \max\left(\frac{Q_Y\left(\frac{d}{m}\right)}{\beta}, Q_X\left(\frac{d}{m}\right)\right)\right)} = \lim_{p \to 0} \frac{P\left(X > Q_X\left(\frac{d}{m}\right), Y > Q_Y\left(\frac{d}{m}\right)\right)}{P\left(X > \frac{Q_Y\left(\frac{d}{m}\right)}{\beta}\right)}$$

$$= \lim_{p \to 0} \frac{\frac{m}{d} P\left(X > Q_X\left(\frac{d}{m}\right), Y > Q_Y\left(\frac{d}{m}\right)\right)}{\frac{P\left(X > \frac{Q_Y\left(\frac{d}{m}\right)}{\beta}\right)}{P\left(X > Q_X\left(\frac{d}{m}\right)\right)}} = \lim_{p \to 0} \frac{\tau\left(\frac{d}{m}\right)}{\left(\frac{Q_Y\left(\frac{d}{m}\right)}{Q_X\left(\frac{d}{m}\right) \cdot \beta}\right)^{-\alpha_x}} = 1,$$

i.e., as $n \to \infty$, we have

$$\tau\left(\frac{d}{m}\right) - \left(\beta \cdot \frac{Q_X\left(\frac{d}{m}\right)}{Q_Y\left(\frac{d}{m}\right)}\right)^{\alpha_x} \xrightarrow{P} 0.$$

Let's just prove that the convergence rate is $\sqrt{kd}$. By referring to the set $C, C_0, C_1, C_{21}, C_{22}$, without loss of generality, let $p = \dfrac{kd}{n} = \dfrac{d}{m}$, $x = y = 1$, by Lemma 4.2,

We prove (4.10) by dealing with the three sets C_1, C_{21}, C_{22}. The limit relation in (4.3) implies that

$$\lim_{n \to \infty} \sqrt{kd}\left(\frac{m}{d} P(C_{21}) - \left(\frac{\beta Q_X\left(\frac{d}{m}\right)}{Q_Y\left(\frac{d}{m}\right) \cdot (1 \pm \delta_n)}\right)^{\alpha_x}\right) = 0,$$

and the limit relation in (4.1) implies that

$$\lim_{n\to\infty} \sqrt{kd}\, \frac{m}{d} P(C_{22}) = 0.$$

For C_1, due to X and ε independent, we have that

$$P(C_1) = P\left(\varepsilon > -\delta Q_Y\left(\frac{m}{d}\right)\right) P\left(X > Q_X\left(\frac{m}{d}\right), X\beta > (1+\delta)Q_Y\left(\frac{m}{d}\right)\right)$$

$$= P\left(\varepsilon > -\delta Q_Y\left(\frac{m}{d}\right)\right) P\left(X\beta > (1+\delta)Q_Y\left(\frac{m}{d}\right)\right),$$

the limit relation (4.2) implies that $\lim_{n\to\infty} \sqrt{kd}\left(P\left(\varepsilon > -\delta Q_Y\left(\frac{m}{d}\right)\right) - 1\right) = 0$. Together with (4.3), we have

$$\lim_{n\to\infty} \sqrt{kd}\left(\frac{m}{d} P(C_1) - \left(\frac{\beta Q_X\left(\frac{d}{m}\right)}{Q_Y\left(\frac{d}{m}\right)}\right)^{\alpha_x}\right) = 0.$$

Since $C_1 \subset C \subset C_{21} \cup C_{22}$, combining $P(C_1)$, $P(C_{21})$ and $P(C_{22})$, we have

$$\lim_{n\to\infty} \sqrt{kd}\left(\frac{m}{d} P(C) - \left(\frac{\beta Q_X\left(\frac{d}{m}\right)}{Q_Y\left(\frac{d}{m}\right)}\right)^{\alpha_x}\right) = 0,$$

then

$$\lim_{n\to\infty} \sqrt{kd}\left(\tau\left(\frac{d}{m}\right) - \left(\frac{\beta Q_X\left(\frac{d}{m}\right)}{Q_Y\left(\frac{d}{m}\right)}\right)^{\alpha_x}\right) = 0.$$

Therefore,

$$\sqrt{kd}\left(\hat{\beta}_D - \beta\right) \xrightarrow{d} N\left(0, \frac{\beta^2}{\alpha_x^2}\left(\frac{1}{\tau} + 2\log\tau - \frac{2}{\tau}\log\tau - 1 - 3(\log\tau)^2\right)\right).$$

$\square$

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Chen, L., Li, D. and Zhou, C. (2021) Distributed Inference for the Extreme Value Index. *Biometrika*, **109**, 257-264. https://doi.org/10.1093/biomet/asab001

[2] Volgushev, S., Chao, S. and Cheng, G. (2017) Distributed Inference for Quantile Regression Processes. *The Annals of Statistics*, **47**, 1634-1662.

https://doi.org/10.1214/18-AOS1730

[3] Mohammed, T., Joe-Wong, C., Babbar, R. and Di Francesco, M. (2020) Distributed Inference Acceleration with Adaptive DNN Partitioning and Offloading. *IEEE IN-FOCOM* 2020-*IEEE Conference on Computer Communications*, Toronto, 6-9 July 2020, 854-863. https://doi.org/10.1109/INFOCOM41043.2020.9155237

[4] Smith, A.J. and Hollinger, G.A. (2018) Distributed Inference-Based Multi-Robot Exploration. *Autonomous Robots*, **42**, 1651-1668. https://doi.org/10.1007/s10514-018-9708-7

[5] Ye, Y. (2017) The Stability of Beta Coefficient in China's Stock Market. *Journal of Service Science and Management*, **10**, 177-187. https://doi.org/10.4236/jssm.2017.102016

[6] Mitra, S. (2019) Measuring Asymmetric Nature of Beta Using a Smooth Linear Transformation. *Theoretical Economics Letters*, **9**, 2019-2032. https://doi.org/10.4236/tel.2019.96128

[7] Markowitz, H. (1952) Portfolio Selection. *The Journal of Financial*, **7**, 77-91. https://doi.org/10.1111/j.1540-6261.1952.tb01525.x

[8] Black, F. and Scholes, M. (1973) The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, **81**, 637-654. https://doi.org/10.1086/260062

[9] Atanasov, V. and Nitschka, T. (2014) Currency Excess Returns and Global Downside Market Risk. *Journal of International Money and Finance*, **47**, 268-285. https://doi.org/10.1016/j.jimonfin.2014.06.006

[10] Lettau, M., Maggiori, M. and Weber, M. (2013) Conditional Risk Premia in Currency Markets and Other Asset Classes. *Journal of Financial Economics*, **114**, 197-225. https://doi.org/10.1016/j.jfineco.2014.07.001

[11] Post, T. and Versijp, P. (2004) Multivariate Tests for Stochastic Dominance Efficiency of a Given Portfolio. *Journal of Financial and Quantitative Analysis*, **42**, 489-515. https://doi.org/10.1017/S0022109000003367

[12] Mitchell, M. and Pulvino, T. (2001) Characteristics of Risk and Return in Risk Arbitrage. *Journal of Finance*, **56**, 2135-2175. https://doi.org/10.1111/0022-1082.00401

[13] Oordt, M. and Zhou, C. (2017) Estimating Systematic Risk under Extremely Adverse Market Conditions. *Journal of Financial Econometrics*, **17**, 432-461. https://doi.org/10.1093/jjfinec/nbx033

[14] Hult, H., Lindskog, F. (2002) Multivariate Extremes, Aggregation and Dependence in Elliptical Distributions. *Advances in Applied Probability*, **34**, 587-608. https://doi.org/10.1239/aap/1033662167

[15] de Haan, L. and Stadtmüller, U. (1996) Generalized Regular Variation of Second Order. *Journal of the Australian Mathematical Society* (*Series A*), **61**, 381-395. https://doi.org/10.1017/S144678870000046X

[16] Drees, H. and Huang, X. (1998) Best Attainable Rates of Convergence for Estimators of the Stable Tail Dependence Function. *Journal of Multivariate Analysis*, **64**, 25-46. https://doi.org/10.1006/jmva.1997.1708

[17] Oordt, M. and Zhou, C. (2018) Systemic Risk and Bank Business Models. *Journal of Applied Econometrics*, **34**, 365-384. https://doi.org/10.1002/jae.2666

[18] Hill, B.M. (1975) A Simple General Approach to Inference about the Tail of a Distribution. *Annals of Statistics*, **3**, 1163-1174. https://doi.org/10.1214/aos/1176343247

[19] Embrechts, P., de Haan, L. and Huang, X. (2000) Modelling Multivariate Extremes. In: Embrechts, P., Ed., *Extremes and Integrated Risk Management*, Risk Book, London, 59-67.

[20] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997) Modelling Extremal Events. Springer, New York. https://doi.org/10.1007/978-3-642-33483-2

[21] de Haan, L. and Ferreira, A. (2006) Extreme Value Theory. Springer, New York. https://doi.org/10.1007/0-387-34471-3

[22] Huang, X. (1992) Statistics of Bivariate Extreme Values. Doctoral Thesis, Tinbergen Institute Research Series, Amsterdam.

[23] Chen, L., Li, D. and Zhou, C. (2021) Distributed Inference for Tail Empirical and Quantile Processes.