Ann Inst Stat Math (2016) 68:237–267
DOI 10.1007/s10463-014-0495-4

On the tail index inference for heavy-tailed GARCH-type innovations

Moosup Kim · **Sangyeol Lee**

Received: 6 February 2014 / Revised: 9 September 2014 / Published online: 5 November 2014

Abstract In this study, we investigate the smoothing Hill plot and change point test for the tail index of power-transformed and threshold generalized autoregressive conditional heteroscedasticity (PTTGARCH) and autoregressive and moving average (ARMA)–GARCH innovations. It is shown that their asymptotic properties are the same as those in the i.i.d. sample case. For illustration, we provide a simulation study and real data analysis.

Keywords Tail index · Hill's estimator · Power-transformed and threshold GARCH model · ARMA–GARCH model · Residuals · Smoothing Hill plot · Change point test

1 Introduction

Tail index estimation has long been a core issue in extreme value theory and diverse research fields. For instance, managing the risk of extreme events is a crucial task in finance and financial asset returns often follow heavy-tailed distributions. Since the seminal paper of Hill (1975), Hill's estimator has been extensively studied by many authors: for example, see Hall (1982), Mason (1982), Hall and Welsh (1985) and Drees et al. (2000), who focus on i.i.d. samples, and Hsing (1991), Resnick and Stărică (1998), Drees (2003) and Hill (2010), who focus on dependent data. For relevant studies, we can also refer to Csörgő et al. (1985), Feuerverger and Hall (1999), Gomes et al. (2008), and Kim and Lee (2008). In financial time series, obtaining information on

Electronic supplementary material The online version of this article (doi:10.1007/s10463-014-0495-4) contains supplementary material, which is available to authorized users.

M. Kim · S. Lee (✉)
Department of Statistics, Seoul National University, Seoul 151-747, Korea
e-mail: sylee@stats.snu.ac.kr

the tail behavior of model innovations rather than the time series itself often plays an important role. For example, when the one-step ahead conditional value-at-risk of a financial asset is calculated, the tail index of innovations should be estimated based on the residuals [cf. Chan et al. (2007)]. Furthermore, in heavy-tailed autoregressive and moving average (ARMA) models, estimating the tail index of i.i.d. innovations enhances the accuracy of estimation [cf. Ling and Peng (2004)].

In this study, among financial time series models, we pay special attention to the power-transformed and threshold generalized autoregressive conditional heteroscedasticity (PTTGARCH) and ARMA–GARCH models, considering their popularity and importance in theories and applications. Fitting procedures in these models are properly implemented with statistical softwares such as R-package 'fGarch'. For the relevant references for the former, we refer to Lee and Lee (2012) and the papers cited therein. In these two models, we focus on the issues of the smoothing Hill plot and tail index change test: the parameter estimation procedure in those models is well established under mild moment conditions on i.i.d. innovations [cf. Pan et al. (2008) and Zhu and Ling (2011)]. It is well known that Hill's estimators fluctuate according to the choice of tail sample fractions, which makes the task of selecting an appropriate estimate difficult to accomplish in practice. As a remedy to overcome this difficulty, here we consider the smoothing Hill plot proposed by Resnick and Stărică (1997b). Meanwhile, the change point test for time series models has long been a core issue in the time series context since time series often undergo parameter changes in their underlying models due to critical social events and changes in monetary policies: see, for instance, Kim and Lee (2009) and the papers cited therein and Quintos et al. (2001) who propose several tests for examining tail index changes. In this study, the smoothing method and change point test are all designed based on residuals. It will be shown that the same asymptotic properties as in i.i.d. samples are also obtained in PTTGARCH and ARMA–GARCH models.

The remainder of this paper is organized as follows. In Sect. 2, we present the main results of this study: in Sect. 2.1, we review the asymptotic properties of Hill's estimator in i.i.d. samples; in Sects. 2.2 and 2.3, we investigate Hill's estimators based on residuals; in Sect. 2.4, we introduce the smoothing method and change point test for the tail index. In Sects. 3 and 4, we provide a simulation study and real data analysis. In Sect. 5, we provide the proofs of the theorems in Sect. 2.

2 Main results

2.1 A review of the asymptotic results of Hill's estimator

In this subsection, we review the asymptotic properties of Hill's estimator in independent identically distributed random variables (i.i.d. r.v.s). In what follows, $\{U_i\}$ denotes a sequence of i.i.d. r.v.s defined on a probability space $(\Omega, \mathcal{F}, P)$ and F denotes the common distribution of $\{U_i\}$. We assume that there exist $\alpha > 0$ and a measurable function, ℓ, such that

$$1 - F(x) = x^{-\alpha}\ell(x), \quad \lim_{x\to\infty} \frac{\ell(\lambda x)}{\ell(x)} = 1 \quad \text{for each } \lambda > 0, \tag{1}$$

where α is the tail index of F. Define

$$H_n(\rho, s, t) := \frac{1}{\lfloor \rho k \rfloor} \sum_{i=\lfloor ns \rfloor+1}^{\lfloor nt \rfloor} \left(\log U_i - \log U_{(\rho,s,t)}\right)_+, \quad \rho > 0, \quad k \in \mathbb{N}, \tag{2}$$

where $U_{(\rho,s,t)}$ indicates the $(\lfloor \rho k(t-s) \rfloor + 1)$-th largest order statistic in $U_{\lfloor ns \rfloor+1}, \ldots,$ $U_{\lfloor nt \rfloor}$. It is assumed that $k = k_n$ varies as $n \to \infty$, such that for some $\nu > 0$,

$$n^\nu = o(k) \quad \text{and} \quad k = o(n) \quad \text{as}\, n \to \infty. \tag{3}$$

Let $b(x) = \inf \left\{y : F(y) \geq 1 - x^{-1}\right\}$. In addition, we provide the following regularity conditions.

Regularity conditions
There exist $C > 0$, $\gamma < 0$, and non-zero $D \in \mathbb{R}$, such that

$$\ell(x) = C \left(1 + Dx^\gamma + x^\gamma \Delta(x)\right), \tag{4}$$

where $\Delta(x)$ is differentiable in terms of x, with

$$\Delta(x) \to 0 \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}x}\Delta(x) = o\left(x^{-1}\right) \quad \text{as}\, x \to \infty. \tag{5}$$

Furthermore, there exists a finite number, $M \geq 0$, such that

$$\lim_{n\to\infty} \sqrt{k}(b(n/k))^\gamma = M. \tag{6}$$

Remark 1 Many articles assume (4) [see Hall (1982), Hall and Welsh (1985), and Feuerverger and Hall (1999)]. In fact, this is an example of a second-order regular variation. In particular, we have

$$\int_0^\infty \mathrm{e}^{-\alpha u} \frac{\ell(ux)}{\ell(x)} \mathrm{d}u - \frac{1}{\alpha} \sim \frac{\gamma D x^\gamma}{\alpha(\alpha - \gamma)} \quad \text{as}\, x \to \infty \tag{7}$$

[cf. Goldie and Smith (1987)]. The assumption of second-order regular variations is necessary to investigate the asymptotic properties of several tail index estimators [see Hsing (1991), Feuerverger and Hall (1999), and Resnick and Stărică (1997a)].

Remark 2 Under (4), we have

$$b(x) = C^{1/\alpha} x^{1/\alpha} \left(1 + \frac{C^{\gamma/\alpha} D}{\alpha} x^{\gamma/\alpha} + o\left(x^{\gamma/\alpha}\right)\right) \quad \text{as}\, x \to \infty. \tag{8}$$

Let $\nu_\circ = -\gamma/\alpha$ and $\nu_* = \frac{2\nu_\circ}{2\nu_\circ+1}$. The constant M in (6) is strictly positive and finite if and only if $k \sim \delta n^{\nu_*}$ $(\delta > 0)$ with $M = \delta^{1/2+\nu_\circ} C^{-\nu_\circ}$. If $k = o(n^{\nu_*})$, $M = 0$.

Theorem 1 *Let $0 < \rho_\circ < \rho_* < \infty$ and $t_0 \in \left(0, \frac{1}{2}\right)$. Suppose that $\{U_i\}$ is i.i.d. and* (1)*,* (3)*,* (4) *and* (6) *hold. Then,*

$$\rho t\sqrt{k}\left(H_n(\rho, 0, t) - \frac{1}{\alpha} - \frac{\gamma D M\rho^{-\gamma/\alpha}}{\sqrt{k}\alpha(\alpha-\gamma)}\right) \xrightarrow{d} \frac{1}{\alpha}B(\rho, t) \quad \text{in } D([\rho_\circ, \rho_*]\times[t_0, 1]), \tag{9}$$

$$\rho\sqrt{k}\begin{bmatrix} t\left(H_n(\rho, 0, t) - \frac{1}{\alpha} - \frac{\gamma D M\rho^{-\gamma/\alpha}}{\sqrt{k}\alpha(\alpha-\gamma)}\right) \\ (1-t)\left(H_n(\rho, t, 1) - \frac{1}{\alpha} - \frac{\gamma D M\rho^{-\gamma/\alpha}}{\sqrt{k}\alpha(\alpha-\gamma)}\right)\end{bmatrix} \xrightarrow{d} \frac{1}{\alpha}\begin{pmatrix} B(\rho, t) \\ B(\rho, 1) - B(\rho, t)\end{pmatrix} \tag{10}$$

in $D^2([\rho_\circ, \rho_*]\times[t_0, 1-t_0])$, where B denotes a standard Brownian sheet.

A standard Brownian sheet B is a continuous Gaussian process with two-dimensional parameter $(\rho, t) \in \mathbb{R}_+^2$ such that

$$EB(\rho_1, t_2) = 0, \quad \mathrm{Cov}(B(\rho_1, t_1), B(\rho_2, t_2)) = (\rho_1 \wedge \rho_2)\cdot(t_1 \wedge t_2)$$

for all (ρ_1, t_1) and (ρ_2, t_2). The results of Theorem 1 are a sort of combinations of Proposition 3.1 of Resnick and Stărică (1997b) and Theorem 1 of Quintos et al. (2001). The former deals with the case of $t = 1$ and the latter handles the case of fixed $\rho > 0$.

Below, we verify that the above theorem still holds for the residuals from PTTGARCH and ARMA–GARCH models. We then apply it to the smoothing Hill plot and tail index change test. For this task, we need a consistent estimator of the model parameters that converges to the true parameter at the rate of n^κ, $\kappa > 0$. In PTTGARCH and ARMA–GARCH models, $\sqrt{n}$-consistent estimators are established under mild moment conditions on innovations. For PTTGARCH models, n^κ-consistent estimators with $\kappa \in \left(0, \frac{1}{2}\right]$ can be employed. In comparison, for ARMA–GARCH models, the $\sqrt{n}$-consistency is required to include the case of ARMA–IGARCH models.

2.2 PTTGARCH models

In this subsection, we consider the power-transformed and threshold generalized autoregressive conditional heteroscedasticity (PTTGARCH) models. Let $\{\varepsilon_i\}$ be a strictly stationary PTTGARCH(p, q) process satisfying the equation:

$$\begin{aligned}\varepsilon_i &= \sigma_i U_i,\\ \sigma_i^{2\delta^\circ} &= \omega^\circ + \sum_{j=1}^{p}\left\{\psi_{1,j}^\circ(\varepsilon_{i-j})_+^{2\delta^\circ} + \psi_{2,j}^\circ(\varepsilon_{i-j})_-^{2\delta^\circ}\right\} + \sum_{j=1}^{q}\beta_j^\circ\sigma_{i-j}^{2\delta^\circ},\end{aligned} \tag{11}$$

where

$$p, q \in \mathbb{N}, \quad \min\{\delta^\circ, \omega^\circ, \psi_{1,1}^\circ, \psi_{2,1}^\circ, \ldots, \psi_{1,p}^\circ, \psi_{2,p}^\circ, \beta_1^\circ, \ldots, \beta_q^\circ\} > 0, \quad \sum_{i=1}^q \beta_i^\circ < 1. \tag{12}$$

Note that (11) can be represented as a stochastic recurrence equation as in Pan et al. (2008), (A.1–A.3). We assume that

$$E|U_0|^\nu < \infty \text{ for some } \nu > 0 \text{ and the top Lyapunov exponent is strictly negative.} \tag{13}$$

Then, $\{\varepsilon_i\}$ is the unique non-anticipative strictly stationary solution of (11) [cf. Bougerol and Picard (1992)]: see also the appendix of Pan et al. (2008) for more details.

PTTGARCH models include diverse variants of GARCH models. If $\psi_{1,i}^\circ \neq \psi_{2,i}^\circ$ for some $i \in \{1, \ldots, p\}$, the model becomes an asymmetric GARCH model that accommodates leverage effects; δ° indicates the order of the power transformation in the conditional variance equation, which makes the PTTGARCH models include nonlinear ARCH models with $\delta^\circ > 0$ not being equal to 1 [cf. Higgins and Bera (1992)]. Further, $\psi_{1,j}^\circ$ and $\psi_{2,j}^\circ$ are the coefficients depending upon the sign of ϵ_{i-j}. If

$$\psi_{1,j}^\circ = \psi_{2,j}^\circ =: \psi_j^\circ \quad \text{for each } j = 1, \ldots, p, \quad \delta^\circ = 1, \tag{14}$$

model (11) is reduced to a simple GARCH(p, q) process.

Suppose that $\varepsilon_1, \ldots, \varepsilon_n$, $n \in \mathbb{N}$, are observed. Let $\boldsymbol{\theta}_1^\circ := (\delta^\circ, \omega^\circ, \psi_{1,1}^\circ, \psi_{2,1}^\circ, \ldots, \psi_{1,p}^\circ, \psi_{2,p}^\circ, \beta_1^\circ, \ldots, \beta_q^\circ)'$ be the parameter vector of (11) and let $\hat{\boldsymbol{\theta}}_1$ be its estimator based on $\varepsilon_1, \ldots, \varepsilon_n$. We assume that

$$\text{there exists } \kappa > 0 \text{ such that } n^\kappa |\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1^\circ| = O_P(1) \text{ and } \sqrt{k} = o(n^{\kappa-\nu}) \text{ for some } \nu > 0, \tag{15}$$

where $|\cdot|$ denotes the Euclidean norm. This condition implicitly assumes that a suitable identifiability condition is satisfied [see assumptions (A1), (A3) and (A4) in Pan et al. (2008)].

Now, we define Hill's estimator based on the residuals for this model. For a given

$$\boldsymbol{\theta}_1 = (\delta, \omega, \psi_{1,1}, \psi_{2,1}, \ldots, \psi_{1,p}, \psi_{2,p}, \beta_1, \ldots, \beta_q)',$$

we recursively obtain

$$\hat{h}_i(\boldsymbol{\theta}_1) = \omega + \sum_{j=1}^p \left\{\psi_{1,j}(\tilde{\varepsilon}_{i-j})_+^{2\delta} + \psi_{2,j}(\tilde{\varepsilon}_{i-j})_-^{2\delta}\right\} + \sum_{j=1}^q \beta_j \hat{h}_{i-j}(\boldsymbol{\theta}_1) \quad \text{for } i = 1, \ldots, n, \tag{16}$$

where $\tilde{\varepsilon}_i = \varepsilon_i$ $(i = 1, \ldots, n)$ and $\tilde{\varepsilon}_0 = \cdots = \tilde{\varepsilon}_{1-p} = \hat{h}_0(\boldsymbol{\theta}_1) = \cdots = \hat{h}_{1-q}(\boldsymbol{\theta}_1) = 0$. Letting $\hat{h}_i = \hat{h}_i(\hat{\boldsymbol{\theta}}_1)$, we define $\tilde{U}_i = \varepsilon_i/(\hat{h}_i)^{1/(2\hat{\delta})}$ and $\tilde{H}_n$ in the same fashion as in (2), with U_i replaced by $\tilde{U}_i$.

Theorem 2 *Assume that the conditions in Theorem 1 hold. Moreover, suppose that (11)–(13) and (15) hold. Then, (9)–(10) with H_n replaced by $\tilde{H}_n$ are fulfilled.*

2.3 ARMA–GARCH models

In this subsection, we deal with ARMA–GARCH models. Let $\{\varepsilon_i\}$ be a strictly stationary sequence satisfying (11–13) and let $\{X_i\}$ be a strictly stationary ARMA$(\bar{p}, \bar{q})$ process, where $\bar{p}, \bar{q}$ are non-negative integers, with possibly infinite variance error process $\{\varepsilon_i\}$ in (11) as follows:

$$X_i = \sum_{j=1}^{\bar{p}} \phi_j^\circ X_{i-j} + \varepsilon_i - \sum_{j=1}^{\bar{q}} \vartheta_j^\circ \varepsilon_{i-j} \quad \text{for every } i \in \mathbb{Z}, \tag{17}$$

where $(\phi_1^\circ, \ldots, \phi_{\bar{p}}^\circ)$ and $(\vartheta_1^\circ, \ldots, \vartheta_{\bar{q}}^\circ)$ satisfy the following:

$$\text{if } 1 - \phi_1^\circ z - \cdots - \phi_{\bar{p}}^\circ z^{\bar{p}} \neq 0 \text{ and } 1 - \vartheta_1^\circ z - \cdots - \vartheta_{\bar{q}}^\circ z^{\bar{q}} \neq 0 \text{ for all } |z| \leq 1. \tag{18}$$

Note that $\{X_i\}$ is causal and invertible. Moreover, we assume (14) with

$$EU_0 = 0, \quad EU_0^2 < \infty, \quad EU_0^2 \cdot \sum_{j=1}^{p} \psi_j^\circ + \sum_{j=1}^{q} \beta_j^\circ \leq 1 \tag{19}$$

to cover the case of ARMA–IGARCH models, where 'I' means 'integrated'.

Suppose that $X_1, \ldots, X_n$ are observed. Let $\boldsymbol{\theta}_2^\circ := (\phi_1^\circ, \ldots, \phi_{\bar{p}}^\circ, \vartheta_1^\circ, \ldots, \vartheta_{\bar{q}}^\circ, \omega^\circ, \psi_1^\circ, \ldots, \psi_p^\circ, \beta_1^\circ, \ldots, \beta_q^\circ)'$ be the parameter vector of (17) and let $\hat{\boldsymbol{\theta}}_2$ be its estimator based on $X_1, \ldots, X_n$. It is assumed that

$$\sqrt{n}|\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_2^\circ| = O_P(1) \tag{20}$$

under a suitable identifiability condition [cf. Zhu and Ling (2011)]. The residual-based estimator is defined in the following way: for a given

$$\boldsymbol{\theta}_2 = (\phi_1, \ldots, \phi_{\bar{p}}, \vartheta_1, \ldots, \vartheta_{\bar{q}}, \omega, \psi_1, \ldots, \psi_p, \beta_1, \ldots, \beta_q)',$$

we recursively obtain

$$\hat{\varepsilon}_i(\boldsymbol{\theta}_2) = \tilde{X}_i - \sum_{j=1}^{\bar{p}} \phi_j \tilde{X}_{i-j} + \sum_{j=1}^{\bar{q}} \vartheta_j \hat{\varepsilon}_{i-j}(\boldsymbol{\theta}_2), \tag{21}$$

where $\tilde{X}_i = X_i, i = 1, \ldots, n$, and $\tilde{X}_0 = \cdots = \tilde{X}_{1-\bar{p}} = \hat{\varepsilon}_0(\boldsymbol{\theta}_2) = \cdots = \hat{\varepsilon}_{1-\bar{q}}(\boldsymbol{\theta}_2) = 0$. Then, for $\omega > 0$ and $\boldsymbol{\theta}_2$, we define

$$\hat{\sigma}_i^2(\boldsymbol{\theta}_2) = \omega + \sum_{j=1}^{p} \psi_j \hat{\varepsilon}_{i-j}^2(\boldsymbol{\theta}_2) + \sum_{j=1}^{q} \beta_j \hat{\sigma}_{i-j}^2(\boldsymbol{\theta}_2), \tag{22}$$

where $\hat{\varepsilon}_0^2(\boldsymbol{\theta}_2) = \cdots = \hat{\varepsilon}_{1-p}^2(\boldsymbol{\theta}_2) = \hat{\sigma}_0^2(\boldsymbol{\theta}_2) = \cdots = \hat{\sigma}_{1-q}^2(\boldsymbol{\theta}_2) = 0$. Define $\hat{\varepsilon}_i = \hat{\varepsilon}_i(\hat{\boldsymbol{\theta}}_2)$, $\hat{\sigma}_i^2 = \hat{\sigma}_i^2(\hat{\boldsymbol{\theta}}_2)$, and $\bar{U}_i = \hat{\varepsilon}_i/\hat{\sigma}_i$ as the residuals, and $\bar{H}_n$ in the same way as in (2), with U_i replaced by $\bar{U}_i$.

Theorem 3 *Assume that the conditions in Theorem 1 hold. Moreover, suppose that (11)–(14) and (17)–(20) hold. Then, (9)–(10) with H_n replaced by $\bar{H}_n$ are satisfied.*

Remark 3 Theorems 2 and 3 still hold when the residuals are replaced by their absolute values, provided that (1) and (4)–(6) hold for the distribution of $|U_0|$. In this case, the proofs are modified slightly.

2.4 Applications of asymptotic results

In this subsection, we apply the asymptotic results from the previous subsections to the smoothing Hill plot and tail index change test.

Smoothing Hill's estimator

Let

$$H_n^*(m) = H_n^*\left(k^{-1}m, 0, 1\right), \quad \text{where } m \in \mathbb{N} \text{ and } H_n^* \text{ stands for } \tilde{H}_n \text{ or } \bar{H}_n.$$

Then, $\{(m, H_n^*(m)) : m = 1, \ldots, m_0\}$ ($m_0 \in \mathbb{N}$ is chosen to be suitably large) is called the Hill plot and is used to make a proper estimate [cf. Drees et al. (2000)]. However, it is not easy to implement when the Hill plot fluctuates significantly. In this case, we consider a local average of Hill's estimators. That is, we use

$$\mathrm{av}H_n^*(k\,;\rho_0,\rho_1) = \frac{1}{k(\rho_1-\rho_0)} \sum_{m=\lfloor\rho_0 k\rfloor+1}^{\lfloor\rho_1 k\rfloor} H_n^*(m), \quad 0<\rho_0<\rho_1<\infty, \tag{23}$$

$$\mathrm{av}H_n^*(m_0,m_1) = \frac{1}{m_1-m_0} \sum_{m=m_0+1}^{m_1} H_n^*(m), \quad m_0,m_1\in\mathbb{N},\ m_0<m_1. \tag{24}$$

The following is the asymptotic law of the local averages. The proof is essentially the same as that of Proposition 4.1 of Resnick and Stărică (1997b).

Corollary 1 *Assume that* (9) *holds with* H_n *replaced by* H_n^*. *Let* $\nu_\circ = -\gamma/\alpha$ *and* $\nu_* = \frac{2\nu_\circ}{2\nu_\circ+1}$. *Suppose that* $k = n^{\nu_*}$. *Then, for* $0 < \rho_0 < \rho_1 < \infty$, *we have*

$$\alpha\sqrt{k}\left(\mathrm{av}H_n^*(k\,;\rho_0,\rho_1) - \frac{1}{\alpha}\right) \xrightarrow{d} N\left(-\frac{\nu_\circ D}{C^{\nu_\circ}(1+\nu_\circ)^2}\cdot\frac{\rho_1^{1+\nu_\circ}-\rho_0^{1+\nu_\circ}}{\rho_1-\rho_0}, \frac{2}{\rho_1-\rho_0}\left(1-\frac{\rho_0}{\rho_1-\rho_0}\log\frac{\rho_1}{\rho_0}\right)\right). \tag{25}$$

Furthermore, if $k = o(n^{\nu_*})$ *as* $n \to \infty$, (25) *still holds with the asymptotic mean 0.*

Change point test for tail index

Suppose we need to test the following hypothesis:

$$\mathcal{H}_0 : \alpha \text{ remains constant in } i = 1, \ldots, n, \quad \text{vs.} \quad \mathcal{H}_1 : \text{not } \mathcal{H}_0.$$

For $\rho > 0$ and $t_0 \in \left(0, \frac{1}{2}\right)$, we define

$$\begin{aligned}
Q_{\mathrm{rec}}(\rho) &:= \sup_{t_0\le t\le 1}\left\{t\sqrt{k\rho}\left(\frac{H_n^*(\rho,0,t)}{H_n^*(\rho,0,1)}-1\right)\right\}^2,\\
Q^\star_{\mathrm{rec}}(\rho) &:= \sup_{0\le t\le 1-t_0}\left\{(1-t)\sqrt{k\rho}\left(\frac{H_n^*(\rho,t,1)}{H_n^*(\rho,0,1)}-1\right)\right\}^2,\\
Q_{\mathrm{rol}}(\rho) &:= \sup_{t_0\le t\le 1}\left\{t_0\sqrt{k\rho}\left(\frac{H_n^*(\rho,t-t_0,t)}{H_n^*(\rho,0,1)}-1\right)\right\}^2,\\
Q_{\mathrm{seq}}(\rho) &:= \sup_{t_0\le t\le 1-t_0} t(1-t)\left\{\sqrt{k\rho}\left(\frac{H_n^*(\rho,0,t)}{H_n^*(\rho,t,1)}-1\right)\right\}^2,\\
Q^\star_{\mathrm{seq}}(\rho) &:= \sup_{t_0\le t\le 1-t_0} t(1-t)\left\{\sqrt{k\rho}\left(\frac{H_n^*(\rho,t,1)}{H_n^*(\rho,0,t)}-1\right)\right\}^2,
\end{aligned}$$

which are referred to as the recursive (rec), rolling (rol), and sequential (seq) tests, respectively. For a motivation of the above test statistics, see Quintos et al. (2001). All the three tests measure the discrepancy between Hill's estimators based on sub-samples in different sub-periods and reject $\mathcal{H}_0$ when the discrepancy is sufficiently large. The following corollary provides the asymptotic distributions of the above statistics under $\mathcal{H}_0$. Since these can be easily proven using Theorems 2 and 3 and a mapping theorem [cf. Billingsley (1999)], we omit the detailed proofs for brevity.

Corollary 2 *Assume that* (9)–(10) *with* H_n *replaced by* H_n^* *hold. Then, for* $\rho > 0$ *and* $t_0 \in \left(0, \frac{1}{2}\right)$,

$$Q_{\mathrm{rec}}(\rho) \xrightarrow{d} \sup_{t_0 \le t \le 1} \left\{B^\circ(t)\right\}^2, \quad Q^\star_{\mathrm{rec}}(\rho) \xrightarrow{d} \sup_{t_0 \le t \le 1} \left\{B^\circ(t)\right\}^2,$$

$$Q_{\mathrm{seq}}(\rho) \xrightarrow{d} \sup_{t_0 \le t \le 1-t_0} \frac{\{B^\circ(t)\}^2}{t(1-t)}, \quad Q^\star_{\mathrm{seq}}(\rho) \xrightarrow{d} \sup_{t_0 \le t \le 1-t_0} \frac{\{B^\circ(t)\}^2}{t(1-t)},$$

$$Q_{\mathrm{rol}}(\rho) \xrightarrow{d} \sup_{t_0 \le t \le 1} \{(B(t) - B(t-t_0)) - t_0 B(1)\}^2,$$

where B *denotes a standard Brownian motion and* $B^\circ(t) = B(t) - tB(1)$.

The above tests rely on the choice of ρ. Thus, we consider smoothing the test statistics with respect to ρ. That is, for $0 < \rho_0 < \rho_1 < \infty$,

$$\mathrm{av}Q_{\mathrm{rec}}(\rho_0, \rho_1) := \frac{1}{\rho_1 - \rho_0} \int_{\rho_0}^{\rho_1} Q_{\mathrm{rec}}(\rho)\mathrm{d}\rho,$$

$$\mathrm{av}Q^\star_{\mathrm{rec}}(\rho_0, \rho_1) := \frac{1}{\rho_1 - \rho_0} \int_{\rho_0}^{\rho_1} Q^\star_{\mathrm{rec}}(\rho)\mathrm{d}\rho,$$

$$\mathrm{av}Q_{\mathrm{seq}}(\rho_0, \rho_1) := \frac{1}{\rho_1 - \rho_0} \int_{\rho_0}^{\rho_1} \sup_{t_0 \le t \le 1-t_0} \left\{ t(1-t)\sqrt{k\rho} \left(\frac{H_n^*(\rho, 0, t)}{H_n^*(\rho, t, 1)} - 1 \right) \right\}^2 \mathrm{d}\rho,$$

$$\mathrm{av}Q^\star_{\mathrm{seq}}(\rho_0, \rho_1) := \frac{1}{\rho_1 - \rho_0} \int_{\rho_0}^{\rho_1} \sup_{t_0 \le t \le 1-t_0} \left\{ t(1-t)\sqrt{k\rho} \left(\frac{H_n^*(\rho, t, 1)}{H_n^*(\rho, 0, t)} - 1 \right) \right\}^2 \mathrm{d}\rho.$$

Corollary 3 *Assume that* (9)–(10) *with* H_n *replaced by* H_n^* *hold. Then, for* $0 < \rho_0 < \rho_1 < \infty$ *and* $t_0 \in \left(0, \frac{1}{2}\right)$,

$$\mathrm{av}Q_{\mathrm{rec}}(\rho_0, \rho_1) \xrightarrow{d} \frac{1}{\rho_1 - \rho_0} \int_{\rho_0}^{\rho_1} \sup_{t_0 \le t \le 1} \{B^\circ(\rho, t)\}^2 \frac{\mathrm{d}\rho}{\rho},$$

$$\mathrm{av}Q^\star_{\mathrm{rec}}(\rho_0, \rho_1) \xrightarrow{d} \frac{1}{\rho_1 - \rho_0} \int_{\rho_0}^{\rho_1} \sup_{1 \le t \le 1-t_0} \{B^\circ(\rho, t)\}^2 \frac{\mathrm{d}\rho}{\rho},$$

$$\mathrm{av}Q_{\mathrm{seq}}(\rho_0, \rho_1) \xrightarrow{d} \frac{1}{\rho_1 - \rho_0} \int_{\rho_0}^{\rho_1} \sup_{t_0 \le t \le 1-t_0} \{B^\circ(\rho, t)\}^2 \frac{\mathrm{d}\rho}{\rho},$$

$$\mathrm{av}Q^\star_{\mathrm{seq}}(\rho_0, \rho_1) \xrightarrow{d} \frac{1}{\rho_1 - \rho_0} \int_{\rho_0}^{\rho_1} \sup_{t_0 \le t \le 1-t_0} \{B^\circ(\rho, t)\}^2 \frac{\mathrm{d}\rho}{\rho},$$

where B *denotes a standard Brownian sheet and* $B^\circ(\rho, t) = B(\rho, t) - tB(\rho, 1)$.

Remark 4 Since the asymptotic laws of recursive and sequential tests are well known, the critical values can be found in several studies [e.g., Quintos et al. (2001), Appendix, and Csörgő and Horváth (1997), p. 25]. The critical values of the rolling test are provided in Quintos et al. (2001). However, the critical values of the local averages,

Table 1 The critical values of the local average av$Q$s

(ρ_0, ρ_1)	t_0	Nominal level		
		10 %	5 %	1 %
(0.5, 1)	0.1	1.39	1.71	2.45
(0.5, 1)	0.2	1.38	1.68	2.40
(0.25, 1)	0.1	1.29	1.54	2.34
(0.25, 1)	0.2	1.32	1.60	2.25

Table 2 The performance of the estimators with $\alpha = 2.5$ in the PTTGARCH case

		m						
		10	20	30	40	50	60	70
$H_n^*(m)$	MSE	1.065	0.349	0.203	0.167	0.125	0.114	0.118
	Coverage rate	0.950	0.913	0.963	0.941	0.968	0.931	0.903
av$H_n^*(10, 70)$	MSE	0.138						
	Coverage rate	0.961						

av$Q$s, are unknown. Therefore, we obtain these by means of Monte Carlo simulations. The critical values are listed in Table 1. Note that both the smoothing method and change point tests are still valid when we use the absolute values of the residuals, as long as (1) and (4–6) hold for the distribution of $|U_0|$ (cf. Remark 3).

3 Simulation study

In this section, we conduct a simulation study to evaluate the performance of the proposed methods.

Smoothing Hill's estimator

In this study, we consider the PTTGARCH(1,1) model with parameters $(\delta^\circ, \omega^\circ, \psi_{1,1}^\circ, \psi_{2,1}^\circ, \beta_1^\circ) = (0.8, 0.2, 0.2, 0.1, 0.4)$ and the AR(1)-GARCH(1,1) model with $(\phi_1^\circ, \omega^\circ, \psi_1^\circ, \beta_1^\circ) = (0.8, 0.1, 0.1, 0.5)$. The innovations follow a t-distribution with α degrees of freedom. In this case, α is the tail index of the innovations. For each case, 1,000 series are generated. A PTTGARCH(1,1) model is fitted to each series and the least absolute deviations estimation (LADE) residuals are obtained through the recursion in (16) [cf. Pan et al. (2008)]. Furthermore, in the AR-GARCH case, the AR(1)-GARCH(1,1) model is fitted and the global self-weighted quasi-maximum exponential likelihood estimator (QMELE) residuals are obtained using (21–22) [cf. Zhu and Ling (2011)]. Using the obtained absolute residuals, we estimate the tail index of the innovations based on both Hill's estimator and the smoothing method. Tables 2, 3, 4, 5 present the MSEs and coverage rates of 95 % confidence intervals. Overall, the results look fairly reasonable and are similar to those in the i.i.d. sample case. Note that the local averages are more accurate and yield good coverage rates without a bias correction.

Table 3 The performance of the estimators with $\alpha = 3.5$ in the PTTGARCH case

		m						
		10	15	20	25	30	35	40
$H_n^*(m)$	MSE	1.329	0.690	0.450	0.431	0.393	0.391	0.370
	Coverage rate	0.971	0.957	0.990	0.953	0.920	0.901	0.903
$\text{av}H_n^*(10, 40)$	MSE	0.378						
	Coverage rate	0.959						

Table 4 The performance of the estimators with $\alpha = 2.5$ in the AR-GARACH case

		m						
		10	20	30	40	50	60	70
$H_n^*(m)$	MSE	1.026	0.366	0.230	0.164	0.129	0.117	0.111
	Coverage rate	0.942	0.950	0.944	0.938	0.934	0.924	0.896
$\text{av}H_n^*(10, 70)$	MSE	0.133						
	Coverage rate	0.949						

Table 5 The performance of the estimators with $\alpha = 3.5$ in the AR-GARACH case

		m						
		10	15	20	25	30	35	40
$H_n^*(m)$	MSE	1.369	0.784	0.610	0.502	0.450	0.447	0.430
	Coverage rate	0.958	0.958	0.948	0.920	0.916	0.886	0.850
$\text{av}H_n^*(10, 40)$	MSE	0.401						
	Coverage rate	0.950						

Change point test for tail index

Here, we only consider the GARCH(1,1) model case as follows: $\boldsymbol{\theta}^\circ = (\omega^\circ, \psi_1^\circ, \beta_1^\circ) = (0.01, 0.05, 0.5)$ and $0 < \tau \leq 1$,

$$\varepsilon_i = \begin{cases} \varepsilon_{i,1}, & i \leq [n\tau], \\ \varepsilon_{i,2}, & i > [n\tau], \end{cases} \qquad \begin{cases} \varepsilon_{i,j} = \sigma_{i,j} \cdot U_{i,j}, \\ \sigma_{i,j}^2 = \omega^\circ + \psi_1^\circ \varepsilon_{i-1,j}^2 + \beta_1^\circ \sigma_{i-1,j}^2, \end{cases} \quad \text{for } j = 1, 2,$$

where $\{U_{i,1}\}$ and $\{U_{i,2}\}$ follow a scaled t-distribution with degrees of freedom α_1 and α_2, respectively, satisfying

$$E\frac{|U_{0,1}|}{1 + |U_{0,1}|} = E\frac{|U_{0,2}|}{1 + |U_{0,2}|} = \frac{1}{2}. \tag{26}$$

We fit the GARCH(1,1) model [i.e., (14) holds] to $\varepsilon_1, \ldots, \varepsilon_n$ using the quasi-maximum likelihood estimator (QMLE), based on the density function $h(x) = 0.5/(1 + |x|)^2$ instead of the standard normal density, as follows:

Table 6 The size of change point tests for tail index at nominal level 5 % ($\tau = 1$ and $\alpha_1 = 2.1$)

n	Type	k											
		40	50	60	70	80	90	100	110	120	130	140	150
1,000	Q_{rec}	0.080	0.078	0.060	0.064	0.056	0.058	0.082	0.052	0.074	0.064	0.054	0.060
	$Q^{\star}_{\text{rec}}$	0.090	0.082	0.074	0.086	0.056	0.076	0.068	0.058	0.062	0.058	0.070	0.060
	Q_{rol}	0.058	0.048	0.050	0.050	0.036	0.048	0.052	0.058	0.056	0.058	0.032	0.038
	Q_{seq}	0.118	0.104	0.096	0.068	0.066	0.060	0.046	0.048	0.040	0.056	0.060	0.056
	$Q^{\star}_{\text{seq}}$	0.162	0.118	0.104	0.088	0.068	0.056	0.048	0.058	0.056	0.056	0.060	0.048
2,000	Q_{rec}	0.114	0.090	0.102	0.084	0.094	0.068	0.058	0.064	0.068	0.060	0.046	0.062
	$Q^{\star}_{\text{rec}}$	0.090	0.082	0.098	0.088	0.076	0.066	0.064	0.062	0.046	0.060	0.054	0.068
	Q_{rol}	0.054	0.038	0.052	0.078	0.068	0.056	0.046	0.052	0.044	0.044	0.042	0.044
	Q_{seq}	0.132	0.100	0.088	0.086	0.090	0.062	0.062	0.040	0.044	0.034	0.032	0.042
	$Q^{\star}_{\text{seq}}$	0.146	0.116	0.116	0.098	0.086	0.082	0.080	0.076	0.074	0.070	0.066	0.066

We set $t_0 = 0.2$ in recursive and sequential tests, $t_0 = 0.4$ in rolling test

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \sum_{i=1}^{n} \left\{ 2\log\left(1 + \frac{|\varepsilon_i|}{\hat{\sigma}_i(\boldsymbol{\theta})}\right) + \log\hat{\sigma}_i(\boldsymbol{\theta}) \right\}, \quad \hat{\sigma}_i(\boldsymbol{\theta}) = \{\hat{h}_i(\boldsymbol{\theta})\}^{1/2}$$

[cf. Berkes and Horváth (2004), Example 2.3]. When $\{\varepsilon_i\}$ is stationary (equivalently, $\tau = 1$), $\sqrt{n}|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^{\circ}| = O_P(1)$ as $n \to \infty$ by (26). Note that $h(x)$ is more suitable to deal with heavy-tailed distributions since a moment condition such as $E|U_{0,j}|^4 < \infty$, $j = 1, 2$, is not required: the $\sqrt{n}$-consistency may not be guaranteed when the standard normal density is used unless the 4th moment condition is imposed. We obtain the residuals using the recursion in (16) with $\hat{\psi}_{1,1} = \hat{\psi}_{2,1} = \hat{\psi}_1$, $\hat{\delta} = 1$, and $(\hat{\boldsymbol{\theta}} = (\hat{\omega}, \hat{\psi}_1, \hat{\beta}_1)')$ and implement a test. Table 6 lists the sizes of the change point tests when the tail of innovations is thick with $\alpha_1 = 2.1$ and $\tau = 1$. In all cases, the sizes are fairly acceptable when k is selected to be moderate. On the other hand, Table 7 presents the sizes when the tail is relatively thin, say, with $\alpha_1 = 7$. Here, the sizes are slightly larger than the nominal level of 5 % when $n = 1,000$, which is improved as n increases to 2,000. Table 9 shows their powers when the tail of innovations becomes thinner with $\alpha_1 = 2.1$, $\alpha_2 = 7$ and $\tau = 0.5$. Among the tests, $Q^{\star}_{\text{rec}}$ and Q_{seq} appear to be the most powerful. On the other hand, Table 10 lists the powers when the tail becomes thicker with $\alpha_1 = 7$ and $\alpha_2 = 2.1$. In this case, Q_{rec} and $Q^{\star}_{\text{seq}}$ are the most powerful. These results show that the recursive and sequential tests have asymmetric powers, whereas the rolling test has symmetric power similarly to the i.i.d. case [cf. Quintos et al. (2001)].

We also investigate the performance of the local averages. Table 8 illustrates their sizes at the nominal level of 5 %. They are seemingly acceptable although both avQ_{seq} and av$Q^{\star}_{\text{seq}}$ have sizes slightly larger than 5 %. Table 11 shows their powers in the same cases as in the previous study. The results show that the local averages perform better regardless of ρ. The performance of the tail index estimator, local average method,

Table 7 The size of change point tests for tail index at nominal level 5 % ($\tau = 1$ and $\alpha_1 = 7$)

n	Type	k											
		40	50	60	70	80	90	100	110	120	130	140	150
1,000	Q_{rec}	0.080	0.076	0.082	0.086	0.086	0.098	0.070	0.098	0.088	0.082	0.090	0.094
	$Q^{\star}_{\text{rec}}$	0.094	0.096	0.088	0.096	0.082	0.084	0.088	0.092	0.082	0.078	0.080	0.078
	Q_{rol}	0.070	0.056	0.074	0.054	0.078	0.078	0.076	0.072	0.070	0.082	0.086	0.098
	Q_{seq}	0.084	0.074	0.048	0.054	0.058	0.050	0.054	0.050	0.058	0.048	0.058	0.048
	$Q^{\star}_{\text{seq}}$	0.128	0.090	0.096	0.100	0.090	0.074	0.066	0.062	0.068	0.060	0.054	0.064
2,000	Q_{rec}	0.078	0.074	0.044	0.052	0.050	0.044	0.066	0.074	0.062	0.058	0.044	0.032
	$Q^{\star}_{\text{rec}}$	0.072	0.056	0.046	0.050	0.070	0.052	0.050	0.058	0.058	0.048	0.042	0.040
	Q_{rol}	0.032	0.038	0.036	0.048	0.042	0.050	0.044	0.062	0.056	0.040	0.038	0.044
	Q_{seq}	0.096	0.060	0.068	0.050	0.048	0.038	0.032	0.058	0.038	0.038	0.030	0.024
	$Q^{\star}_{\text{seq}}$	0.144	0.108	0.082	0.080	0.066	0.066	0.044	0.044	0.050	0.040	0.036	0.036

We set $t_0 = 0.2$ in recursive and sequential tests, $t_0 = 0.4$ in rolling test

Table 8 The size of local averages at nominal level 5 % ($\tau = 1$)

n	α_1	av Q_{rec}	av $Q^{\star}_{\text{rec}}$	av Q_{seq}	av $Q^{\star}_{\text{seq}}$
1,000	7	0.036	0.030	0.070	0.072
	2.1	0.040	0.044	0.088	0.078
2,000	7	0.036	0.032	0.064	0.078
	2.1	0.062	0.048	0.088	0.084

We set $k = 120$, $\rho_0 = 0.5$, $\rho_1 = 1$ and $t_0 = 0.2$

Table 9 The power of change point tests for tail index at nominal level 5 % ($\tau = 0.5, \alpha_1 = 2.1, \alpha_2 = 7$).

n	Type	k											
		40	50	60	70	80	90	100	110	120	130	140	150
1,000	Q_{rec}	0.29	0.36	0.40	0.49	0.50	0.58	0.53	0.59	0.63	0.65	0.63	0.64
	$Q^{\star}_{\text{rec}}$	0.66	0.70	0.63	0.66	0.66	0.70	0.73	0.67	0.63	0.64	0.66	0.65
	Q_{rol}	0.23	0.30	0.33	0.42	0.38	0.38	0.42	0.47	0.44	0.44	0.43	0.52
	Q_{seq}	0.70	0.79	0.75	0.76	0.71	0.76	0.78	0.69	0.68	0.71	0.68	0.69
	$Q^{\star}_{\text{seq}}$	0.02	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.03	0.02	0.01	0.01
2,000	Q_{rec}	0.16	0.25	0.27	0.38	0.46	0.49	0.60	0.57	0.62	0.66	0.69	0.75
	$Q^{\star}_{\text{rec}}$	0.90	0.89	0.94	0.95	0.93	0.93	0.91	0.89	0.95	0.93	0.91	0.92
	Q_{rol}	0.48	0.55	0.59	0.63	0.68	0.72	0.82	0.82	0.80	0.80	0.80	0.80
	Q_{seq}	0.89	0.89	0.89	0.92	0.94	0.95	0.92	0.93	0.93	0.91	0.91	0.90
	$Q^{\star}_{\text{seq}}$	0.03	0.03	0.01	0.01	0.03	0.05	0.06	0.08	0.11	0.15	0.19	0.25

We set $t_0 = 0.2$ in recursive and sequential tests, $t_0 = 0.4$ in rolling test

and change point test might depend upon the choice of GARCH parameter estimators. However, our past experience indicates that the result is not much affected as far as

Table 10 The power of change point tests for tail index at nominal level 5% ($\tau = 0.5, \alpha_1 = 7, \alpha_2 = 2.1$)

n	Type	k											
		40	50	60	70	80	90	100	110	120	130	140	150
1,000	Q_{rec}	0.55	0.66	0.66	0.61	0.61	0.65	0.71	0.59	0.61	0.61	0.59	0.59
	$Q^{\star}_{\text{rec}}$	0.23	0.27	0.37	0.47	0.43	0.52	0.54	0.52	0.60	0.59	0.62	0.57
	Q_{rol}	0.19	0.32	0.37	0.41	0.40	0.45	0.47	0.50	0.53	0.54	0.55	0.50
	Q_{seq}	0.02	0.02	0.00	0.01	0.00	0.00	0.00	0.00	0.02	0.02	0.03	0.05
	$Q^{\star}_{\text{seq}}$	0.76	0.74	0.74	0.73	0.76	0.73	0.74	0.72	0.69	0.71	0.71	0.66
2,000	Q_{rec}	0.85	0.81	0.88	0.90	0.91	0.93	0.94	0.93	0.94	0.95	0.92	0.92
	$Q^{\star}_{\text{rec}}$	0.21	0.19	0.29	0.34	0.38	0.56	0.52	0.60	0.65	0.67	0.77	0.75
	Q_{rol}	0.43	0.51	0.62	0.70	0.69	0.72	0.70	0.77	0.76	0.79	0.82	0.81
	Q_{seq}	0.03	0.01	0.01	0.01	0.01	0.00	0.04	0.07	0.09	0.10	0.13	0.21
	$Q^{\star}_{\text{seq}}$	0.81	0.82	0.85	0.87	0.91	0.90	0.95	0.95	0.94	0.94	0.95	0.95

We set $t_0 = 0.2$ in recursive and sequential tests, $t_0 = 0.4$ in rolling test

Table 11 The power of local averages at nominal level 5 % ($\tau = 0.5$)

n	α_1	α_2	avQ_{rec}	av$Q^{\star}_{\text{rec}}$	avQ_{seq}	av$Q^{\star}_{\text{seq}}$
1,000	7	2.1	0.67	0.57	0.09	0.86
	2.1	7	0.57	0.67	0.83	0.09
2,000	7	2.1	0.94	0.70	0.37	0.97
	2.1	7	0.53	0.92	0.93	0.27

We set $k = 120, \rho_0 = 0.5, \rho_1 = 1$, and $t_0 = 0.2$

the estimators behave regularly ($\sqrt{n}$-consistent). Overall, our simulation confirms the validity of our methods in both the smoothing Hill plot and change point test.

4 Real data analysis

Smoothing Hill Plot for Hang Seng Index

We analyze the data set $\varepsilon_1, \ldots, \varepsilon_n$ of the daily Hang Seng Index from January 2, 2001 to December 31, 2003 ($n = 739$). Pan et al. (2008) fitted a PTTGARCH(1,1) model to the same data and estimated the tail index of the absolute innovations using the absolute values of the residuals. However, since the reciprocals of Hill's estimates fluctuate from 3.5 to 5 as k changes from 1 to 100, it is difficult to select an appropriate estimate [see Figure 6 in Pan et al. (2008)]. Therefore, we apply the smoothing method. Table 12 list the standard errors and 95 % confidence intervals obtained from (25) with assuming $M = 0$ in (6). The results show that the innovation has a finite third moment and an infinite sixth moment.

Change point analysis of tail behavior for Kuala Lumpur composite index

We next perform a change point test for the data set $\varepsilon_1, \ldots, \varepsilon_n$ of the daily Kuala Lumpur Composite Index from January 2, 1995 to October 16, 1998. Here, the sample size is $n = 935$. Quintos et al. (2001) analyzed this data set and demonstrated that the tail index of daily returns significantly decreases around the Asian financial crisis

Table 12 Smoothing Hill's estimates for the real data (* confidence level: 95 %)

	Estimate	Standard error	Confidence interval*
$1/\text{av}H_n^*(10, 70)$	4.467	0.670	(3.329, 5.995)
$1/\text{av}H_n^*(20, 80)$	4.575	0.613	(3.519, 5.948)
$1/\text{av}H_n^*(30, 90)$	4.446	0.546	(3.497, 5.653)
$1/\text{av}H_n^*(40, 100)$	4.281	0.488	(3.424, 5.351)

Table 13 The values of the test statistics for the real data (bold figures indicate significance at level 5 %)

	Type	k					
		40	50	60	70	80	90
Upper tail	Q_{rec}	0.772	1.428	0.568	0.776	1.197	1.270
	$Q^{\star}_{\text{rec}}$	0.650	0.830	1.164	0.997	1.628	**5.614**
	Q_{rol}	**4.392**	0.966	1.473	0.946	1.681	1.226
	Q_{seq}	0.985	1.948	1.459	2.354	2.688	5.510
	$Q^{\star}_{\text{seq}}$	2.219	5.466	3.224	6.049	7.229	**23.265**
Lower tail	Q_{rec}	1.459	0.642	0.376	0.478	0.607	0.379
	$Q^{\star}_{\text{rec}}$	0.789	1.140	1.111	0.858	0.178	0.291
	Q_{rol}	**2.386**	1.389	1.142	0.825	1.667	**2.990**
	Q_{seq}	7.685	8.655	4.315	2.194	1.060	1.011
	$Q^{\star}_{\text{seq}}$	2.150	2.217	1.734	1.162	1.100	1.568
Tail of absolute residuals	Q_{rec}	0.321	0.177	0.264	0.632	0.818	1.032
	$Q^{\star}_{\text{rec}}$	1.124	**2.407**	1.338	0.279	0.544	0.633
	Q_{rol}	0.631	1.246	1.058	1.015	1.093	**2.338**
	Q_{seq}	3.595	3.755	1.936	0.570	2.278	4.032
	$Q^{\star}_{\text{seq}}$	1.405	1.559	1.043	0.408	1.157	1.769

$t_0 = 0.214$ in recursive and sequential tests, $t_0 = 0.428$ in rolling test

Table 14 The values of the local averages ($k = 90$, $\rho_0 = 0.5$, $\rho_1 = 1$, $t_0 = 0.214$)

	$\text{av}Q_{\text{rec}}$	$\text{av}Q^{\star}_{\text{rec}}$	$\text{av}Q_{\text{seq}}$	$\text{av}Q^{\star}_{\text{seq}}$
Upper tail	0.756	1.436	0.554	1.579
Lower tail	0.631	0.794	0.714	0.315
Tail of absolute residuals	0.821	0.555	0.464	0.228

period in 1997, that is, the tail gets thicker after the crisis. To see the tail index change of innovations, a GARCH(1,1) model is fitted to the returns and the QMLE is obtained based on the power density $h(x) = 0.5/(1 + |x|)^2$ [cf. Berkes and Horváth (2004), Example 2.3], that is, $(\hat{\omega}, \hat{\psi}_1, \hat{\beta}_1) = (0.0118, 0.0534, 0.8012)$. Then, the residuals are obtained using the recursion in (16) with $\hat{\psi}_{1,1} = \hat{\psi}_{2,1} = \hat{\psi}_1$ and $\hat{\delta} = 1$. Table 13 reveals that the test statistics and significance at the 5 % level depend upon both k and the type of tests. Therefore, we employ the smoothing method and obtain the results in Table 14. These results indicate that none of them reject $\mathcal{H}_0$ at the 5 % level.

Since this confirms the constancy of the tail index, there are other possible reasons for the increased volatility at the Asian financial crisis period. First, maybe the returns follow an IGARCH(1,1) model: note that the second moment of the innovations is estimated to be 4.05 and $\widehat{EU_0^2} \cdot \hat{\psi}_1 + \hat{\beta}_1 = 4.05 \cdot 0.0534 + 0.801 \approx 1$. Second, the GARCH parameters may experience a change and cause a spurious integration [cf. Hillebrand (2005)], since the tail index of GARCH(1,1) models is determined by both the innovation distribution and the GARCH parameters [cf. Mikosch and Stărică (2000), Theorem 2.1]. To get a conclusive answer, a more detailed analysis would be required. Since this is beyond the scope of this study, we leave it as a task of our future study.

5 Proofs

In what follows, K denotes a universal positive constant. Let $\mathcal{F}_i = \sigma\{\ldots, U_{i-1}, U_i\}$ and let $E_i(\cdot) = E(\cdot|\mathcal{F}_i)$ be the conditional expectation with respect to $\mathcal{F}_i$. The proofs of all the lemmas in this section are provided in the supplementary material.

5.1 Preliminary results

In this subsection, we assume that $\{U_i\}$ is a sequence of i.i.d. r.v.s and (1–6) hold. Further, N_m ($m \in \mathbb{N}$) and B denote an m-variate normal distribution and a standard Brownian sheet, respectively. Let $G(y,t)$ be a continuous Gaussian process defined on $(y,t) \in [0,\infty) \times [0,1]$ such that

- For $0 \le s < t \le 1$ and $0 = y_0 < y_1 < \cdots, y_m < \infty$ ($m \in \mathbb{N}$),

$$(G(y_1,t) - G(y_0,t), \ldots, G(y_m,t) - G(y_{m-1},t))' \sim N_m(\mathbf{0}, t\alpha^{-2}\Upsilon), \quad (27)$$

  where

$$\Upsilon := \Upsilon(y_1,\ldots,y_m) = \begin{pmatrix} \tau_{1,1} & \cdots & \tau_{1,m} \\ \vdots & \ddots & \vdots \\ \tau_{1,m} & \cdots & \tau_{m,m} \end{pmatrix}$$

  with $\tau_{i,i} := 2\left(y_i - y_{i-1} - y_{i-1}\log\frac{y_i}{y_{i-1}}\right)$ and $\tau_{i,j} := (y_i - y_{i-1})\log\frac{y_j}{y_{j-1}}$ (conventionally, $0 \cdot \log(\infty) = 0$),
- For $0 \le s < t \le 1$, $A(t) - A(s)$ and $A(s)$ are independent, where

$$A(u) = (G(y_1,u), G(y_2,u), \ldots, G(y_m,u)),\ u \in [0,1].$$

In this subsection, we provide three propositions that play a prominent role in proving Theorem 1. In what follows, let $y_* > 1$ and

$$M_n^*(y,t) := \frac{1}{\sqrt{k}} \sum_{i=1}^{\lfloor nt \rfloor} \left\{ I\left(U_i \geq y^{-1/\alpha} b(n/k)\right) - P\left(U_i \geq y^{-1/\alpha} b(n/k)\right) \right\}$$

for $0 \leq y \leq y_*$ and $t \in [0, 1]$. The following lemma is useful to verify Proposition 1 below.

Lemma 1

$$\lim_{n\to\infty} \sup_{0 \leq y_1 < y \leq y_*} \left| \frac{n}{k} \frac{P\left(y^{-1/\alpha} \leq U_1/b(n/k) < y_1^{-1/\alpha}\right)}{y - y_1} - 1 \right| = 0.$$

In what follows, we use the symbols:

$$M_n^*(B) = M_n^*(y_2, t_2) - M_n^*(y_1, t_2) - M_n^*(y_2, t_1) + M_n^*(y_1, t_1)$$

with $B = (y_1, y_2] \times (t_1, t_2]$ and $T_n = \{i/n : i = 1, 2, \ldots, n\}$.

Proposition 1 *There exist $K_1 = K_1(y^*) > 0$ and $n_0 \in \mathbb{N}$ such that for every $n \geq n_0$,*

$$E\left(\left\{M_n^*(B_1)\right\}^2 \left\{M_n^*(B_2)\right\}^2\right) \leq K_1(y_2 - y)(y - y_1)(t_2 - t_1)^2, \tag{28}$$

$$E\left(\left\{M_n^*(B_3)\right\}^2 \left\{M_n^*(B_4)\right\}^2\right) \leq K_1(y_2 - y_1)^2(t - t_1)(t_2 - t) \tag{29}$$

with

$$B_1 = (y_1, y] \times (t_1, t_2], \quad B_2 = (y, y_2] \times (t_1, t_2],$$
$$B_3 = (y_1, y_2] \times (t_1, t], \quad B_4 = (y_1, y_2] \times (t, t_2],$$

where $0 \leq y_1 < y < y_2 \leq y_*$ and $t_1 < t < t_2$ lie in T_n. Further, for $t \in [0, 1]$ and $0 = y_0 < y_1 < \cdots < y_m \leq y_*$ ($m \in \mathbb{N}$),

$$\begin{pmatrix} M_n^*(y_1, t) - M_n^*(y_0, t) \\ \vdots \\ M_n^*(y_m, t) - M_n^*(y_{m-1}, t) \end{pmatrix} \xrightarrow{d} N_m\left(\mathbf{0}\ , t \cdot \mathrm{diag}\{y_1 - y_0, \cdots, y_m - y_{m-1}\}\right). \tag{30}$$

Hence,

$$M_n^*(y,t) \xrightarrow{d} B(y,t) \quad \text{in } D([0, y^*] \times [0, 1]). \tag{31}$$

Proof Following the arguments in the proof of Proposition 2.1 of Resnick and Stărică (1997b) and using Lemma 1, we can easily see that (28) holds. Further, (29) is easy to check owing to the independence. This indicates that $\{M_n^*\}$ is tight [cf. Bickel and Wichura (1971), Theorem 3, and the appended comment]. Since (30) is easy to check using standard arguments, the proposition is validated. □

For $y \geq 0$, set

$$Z_{ni}(y) := \left(\log U_i - \log b(n/k) + \frac{1}{\alpha}\log y\right)_+ \quad \text{and}$$
$$L_n^*(y,t) := \frac{1}{\sqrt{k}} \sum_{i=1}^{\lfloor nt \rfloor} \{Z_{ni}(y) - EZ_{ni}(y)\}. \tag{32}$$

The following results are useful to prove Proposition 2 below:

Lemma 2 *Let $y_* > 1$. Then, as $n \to \infty$,*

$$E\{Z_{ni}(y_2) - Z_{ni}(y_1)\} \sim \frac{k}{\alpha n}(y_2 - y_1), \tag{33}$$
$$E\{Z_{ni}(y_2) - Z_{ni}(y_1)\}^2 \sim \frac{2k}{\alpha^2 n}\left(y_2 - y_1 - y_1 \log\frac{y_2}{y_1}\right), \tag{34}$$
$$E\{Z_{ni}(y_2) - Z_{ni}(y_1)\}^3 \sim \frac{3k}{\alpha^3 n}\left(2y_2 - y_1 - y_1\left(\log\frac{y_2}{y_1} + 1\right)^2\right), \tag{35}$$
$$E\{Z_{ni}(y_4) - Z_{ni}(y_3)\}\{Z_{ni}(y_2) - Z_{ni}(y_1)\} \sim \frac{k}{\alpha^2 n}\left(\log\frac{y_4}{y_3}\right)(y_2 - y_1), \tag{36}$$
$$E\{Z_{ni}(y_4) - Z_{ni}(y_3)\}^2\{Z_{ni}(y_2) - Z_{ni}(y_1)\}^2 \sim \frac{2k}{\alpha^2 n}\left(\log\frac{y_4}{y_3}\right)^2 \times \left(y_2 - y_1 - y_1\log\frac{y_2}{y_1}\right) \tag{37}$$

uniformly in $0 \leq y_1 < y_2 \leq y_3 < y_4 < y_*$.

We set

$$L_n^*(B) = L_n^*(y_2, t_2) - L_n^*(y_1, t_2) - L_n^*(y_2, t_1) + L_n^*(y_1, t_1)$$

with $B = (y_1, y_2] \times (t_1, t_2]$.

Proposition 2 *Let $0 < y_\circ < y_* < \infty$. Then, there exist $K_1 = K_1(y_\circ, y_*) > 0$ and $n_0 \in \mathbb{N}$, such that for every $n \geq n_0$,*

$$E\left(\left\{L_n^*(B_1)\right\}^2\left\{L_n^*(B_2)\right\}^2\right) \leq K_1(y_2 - y)(y - y_1)(t_2 - t_1)^2, \tag{38}$$
$$E\left(\left\{L_n^*(B_3)\right\}^2\left\{L_n^*(B_4)\right\}^2\right) \leq K_1(y_2 - y_2)^2(t - t_1)(t_2 - t) \tag{39}$$

for

$$B_1 = (y_1, y] \times (t_1, t_2], \quad B_2 = (y, y_2] \times (t_1, t_2],$$
$$B_3 = (y_1, y_2] \times (t_1, t], \quad B_4 = (y_1, y_2] \times (t, t_2],$$

where $0 \le y_1 < y < y_2 \le y_*$ *and* $t_1 < t < t_2$ *lie in* T_n. *Further, for* $t \in [0, 1]$ *and* $0 = y_0 < y_1 < \cdots < y_m \le y_*$ $(m \in \mathbb{N})$,

$$\begin{aligned}&\left(L_n^*(y_1, t) - L_n^*(y_0, t), \ldots, L_n^*(y_m, t) - L_n^*(y_{m-1}, t)\right)\\ &\xrightarrow{d} N_m\left(\mathbf{0}\,, t\alpha^{-2}\Upsilon(y_1, \ldots, y_m)\right). \end{aligned} \tag{40}$$

Hence,

$$L_n^*(y, t) \xrightarrow{d} G(y, t) \quad in\, D\left([y_\circ, y_*] \times [0, 1]\right). \tag{41}$$

Proof The proof is omitted since it is essentially the same as that of Proposition 1. □

Proposition 3 *Let* $0 < y_\circ < y_* < \infty$. *Then,*

$$L_n^*(y, t) - \alpha^{-1}M_n^*(y, t) \xrightarrow{d} \alpha^{-1}B(y, t) \quad in\, D\left([y_\circ, y_*] \times [0, 1]\right). \tag{42}$$

Proof Since $\{L_n^*(y, t) - \alpha^{-1}M_n^*(y, t)\}$ is tight, it suffices to show that every finite-dimensional distribution of $L_n^*(y, t) - \alpha^{-1}M_n^*(y, t)$ converges weakly to the corresponding finite-dimensional distribution of $\alpha^{-1}B$. Let $t \in [0, 1]$, $0 = y_0 < y_1 < \cdots, y_m \le y_*$ $(m \in \mathbb{N})$, and let

$$\begin{aligned}\boldsymbol{Y}_i &= \left(L_n^*(y_i, t) - L_n^*(y_{i-1}, t), \alpha^{-1}\left\{M_n^*(y_i, t) - M_n^*(y_{i-1}, t)\right\}\right)', \\ R_i &= \boldsymbol{v}'\boldsymbol{Y}_i, \ \boldsymbol{v} = (1, -1)'.\end{aligned}$$

Then, we can easily check that

$$\begin{pmatrix}\boldsymbol{Y}_1\\ \vdots \\ \boldsymbol{Y}_m\end{pmatrix} \xrightarrow{d} N_{2m}\left(\mathbf{0}, \begin{pmatrix}\Gamma_1 & \cdots & \Gamma_{1,m}\\ \vdots & \ddots & \vdots\\ \Gamma_{1,m}{}' & \cdots & \Gamma_m\end{pmatrix}\right),$$

where

$$\begin{aligned}\Gamma_i &:= \frac{t}{\alpha^2}\begin{pmatrix} 2\left(y_i - y_{i-1} - y_{i-1}\log\frac{y_i}{y_{i-1}}\right) & y_i - y_{i-1} - y_{i-1}\log\frac{y_i}{y_{i-1}} \\ y_i - y_{i-1} - y_{i-1}\log\frac{y_i}{y_{i-1}} & y_i - y_{i-1}\end{pmatrix}, \\ \Gamma_{i,j} &:= \frac{t}{\alpha^2}\begin{pmatrix}(y_i - y_{i-1})\log\frac{y_j}{y_{j-1}} & 0\\ (y_i - y_{i-1})\log\frac{y_j}{y_{j-1}} & 0\end{pmatrix}.\end{aligned}$$

Thus,

$$(R_1, \ldots, R_m)' \xrightarrow{d} N_m\left(\mathbf{0}, t\alpha^{-2}\cdot \mathrm{diag}\{y_1 - y_0, \cdots, y_m - y_{m-1}\}\right).$$

Further, we also have that the limit process of

$$A_n(s) = \left(L_n^*(y_1, s) - \alpha^{-1} M_n^*(y_1, s), \ldots, L_n^*(y_m, s) - \alpha^{-1} M_n^*(y_m, s)\right)$$

has stationary and independent increments. Hence, the proposition is established. □

5.2 The proof of theorem 1

In this subsection, we assume that (1)-(6), (31), (41), and (42) hold. Further, it is assumed that $M \in [0, \infty)$ in (6). Here, we do not require the independence assumption on $\{U_i\}$.

For $\zeta \in \mathbb{R}$, $\rho > 0$, and $0 \leq s < t \leq 1$, set

$$\begin{aligned} L_n(\rho, \zeta, s, t) &= \frac{1}{\sqrt{k}} \sum_{i=\lfloor ns \rfloor + 1}^{\lfloor nt \rfloor} \left\{ \left(\log U_i - \log b\left(\frac{n}{\rho k}\right) - \frac{\zeta}{\sqrt{k}} \right)_+ \right. \\ &\quad \left. - E\left(\log U_i - \log b\left(\frac{n}{\rho k}\right) - \frac{\zeta}{\sqrt{k}} \right)_+ \right\}, \\ M_n(\rho, \zeta, s, t) &= \frac{1}{\alpha\sqrt{k}} \sum_{i=\lfloor ns \rfloor + 1}^{\lfloor nt \rfloor} \left\{ I\left(U_i \geq e^{\zeta/\sqrt{k}} b\left(\frac{n}{\rho k}\right)\right) \right. \\ &\quad \left. - P\left(U_i \geq e^{\zeta/\sqrt{k}} b\left(\frac{n}{\rho k}\right)\right) \right\}, \\ W_n(\rho, s, t) &= \sqrt{k} \left\{ \log U_{(\rho, s, t)} - \log b\left(\frac{n}{\rho k}\right) \right\}. \end{aligned}$$

Below, we state a series of standard lemmas useful to prove Theorem 1: Lemmas 3, 4 and 5, and Lemmas 6 and 7 are used to prove Lemmas 4, 7, and Theorem 1, respectively.

Lemma 3 *Let $\rho_* > 0$. Then,*

$$\log b\left(\frac{n}{\rho k}\right) = \log b(n/k) - \frac{1}{\alpha} \log \rho + \frac{MD}{\alpha\sqrt{k}} (\rho^{-\gamma/\alpha} - 1) + o\left(\frac{1}{\sqrt{k}}\right)$$

uniformly in $0 < \rho \leq \rho_*$.

Lemma 4 *Let $\zeta \in \mathbb{R}$ and $0 < \rho_\circ < \rho_* < \infty$. Then,*

$$W_n(\rho, s, t) \geq \zeta \quad \textit{if and only if} \quad M_n(\rho, \zeta, s, t) \geq \zeta\rho(t - s) + o(1) \tag{43}$$

uniformly in $\rho_\circ < \rho < \rho_*$ *and* $0 \le s < t \le 1$ *with* $t - s > 1/\sqrt{k}$. *Further, for any* $K > 0$ *and* $\epsilon > 0$,

$$\lim_{n\to\infty} P\left(\sup_{\rho_\circ\le\rho\le\rho_*}\ \sup_{-K\le\zeta\le K}\ \sup_{t\in[0,1]}\left|L_n(\rho,\zeta,0,t) - L_n^*(\rho,t)\right| > \epsilon\right) = 0, \tag{44}$$

$$\lim_{n\to\infty} P\left(\sup_{\rho_\circ\le\rho\le\rho_*}\ \sup_{-K\le\zeta\le K}\ \sup_{t\in[0,1]}\left|M_n(\rho,\zeta,0,t) - \alpha^{-1}M_n^*(\rho,t)\right| > \epsilon\right) = 0, \tag{45}$$

and

$$\lim_{\delta\to 0}\limsup_{n\to\infty} P\left(\sup_{\delta}|L_n(\rho_1,0,0,s) - L_n(\rho_2,0,0,t)| > \epsilon\right) = 0, \tag{46}$$

where the supremum is taken over $|(\rho_1,s)-(\rho_2,t)| \le \delta$ *with* $\rho_\circ < \rho_1 \le \rho_2 < \rho_*$. *Hence,* $\{L_n(\rho,0,0,t) : (\rho,t) \in [\rho_\circ,\rho_*]\times[0,1]\}$ *is tight.*

Lemma 5 *Let* $0 < \rho_\circ < \rho_* < \infty$ *and* $t_0 \in (0,1)$. *Then,* $\{\rho t W_n(\rho,0,t) : (\rho,t) \in [\rho_\circ,\rho_*]\times[t_0,1]\}$ *and* $\{\rho(1-t)W_n(\rho,t,1) : (\rho,t) \in [\rho_\circ,\rho_*]\times[0,1-t_0]\}$ *are tight.*

Lemma 6 *Let* $0 < \rho_\circ < \rho_* < \infty$ *and* $K > 0$. *Then,*

$$E\left(\log U_i - \log b\left(\frac{n}{\rho k}\right) - \frac{\zeta}{\sqrt{k}}\right)_+ = \frac{\rho k}{n}\left\{\frac{1}{\alpha} - \frac{\zeta}{\sqrt{k}} + \frac{\gamma DM}{\sqrt{k}\alpha(\alpha-\gamma)}\rho^{-\gamma/\alpha} + o\left(\frac{1}{\sqrt{k}}\right)\right\} \tag{47}$$

uniformly in $\zeta \in [-K,K]$ *and* $\rho_\circ \le \rho \le \rho_*$.

Lemma 7 *Let* $0 < \rho_\circ < \rho_* < \infty$. *Then,*

$$L_n(\rho,0,0,t) - \rho t W_n(\rho,0,t) \xrightarrow{d} \frac{1}{\alpha}B(\rho,t) \quad \text{in } D([\rho_\circ,\rho_*]\times[t_0,1]),$$

$$\begin{pmatrix} L_n(\rho,0,0,t) - \rho t W_n(\rho,0,t) \\ L_n(\rho,0,t,1) - \rho(1-t)W_n(\rho,t,1)\end{pmatrix} \xrightarrow{d} \frac{1}{\alpha}\begin{pmatrix} B(\rho,t) \\ B(\rho,1) - B(\rho,t)\end{pmatrix}, \quad \text{in } D^2([\rho_\circ,\rho_*]\times[t_0,1]).$$

Now, we are ready to prove Theorem 1.

Proof of Theorem 1. We only prove (9) since (10) can be proven similarly. From Lemmas 5 and 6, we have

$$L_n(\rho, W_n(\rho,0,t),0,t) = \rho t\sqrt{k}\left(H_n(\rho,0,t) - \frac{1}{\alpha} + \frac{W_n(\rho,0,t)}{\sqrt{k}} - \frac{\gamma DM}{\sqrt{k}\alpha(\alpha-\gamma)}\rho^{-\gamma/\alpha} + o_P\left(\frac{1}{\sqrt{k}}\right)\right)$$

uniformly in $\rho_\circ < \rho < \rho_*$ and $t \in [t_0, 1]$. Thus, from (44) and Lemma 5, we get

$$\begin{aligned}
&\rho t\sqrt{k}\left(H_n(\rho, 0, t) - \frac{1}{\alpha} - \frac{\gamma D M \rho^{-\gamma/\alpha}}{\sqrt{k}\alpha(\alpha-\gamma)}\right) \\
&\quad = L_n(\rho, W_n(\rho, 0, t), 0, t) - \rho t W_n(\rho, 0, t) + o_P(1) \\
&\quad = L_n(\rho, 0, 0, t) - \rho t W_n(\rho, 0, t) + o_P(1)
\end{aligned}$$

uniformly in $\rho_\circ < \rho < \rho_*$ and $t \in [t_0, 1]$. Hence, the lemma is asserted by Lemma 7. □

5.3 Proof of theorem 2

In this subsection, we assume that (1–6), (11–13), and (15) hold. We first verify the following proposition.

Proposition 4 *Let* $0 < \underline{y} < \bar{y} < \infty$. *Then,*

$$\sup_{\underline{y}<y<\bar{y}} \sup_{0\le t\le 1} \frac{1}{\sqrt{k}} \left| \sum_{i=1}^{\lfloor nt \rfloor} \left\{ I\left(U_i \ge y^{-\frac{1}{\alpha}} b(n/k)\right) - I\left(\tilde{U}_i \ge y^{-\frac{1}{\alpha}} b(n/k)\right) \right\} \right| = o_P(1), \tag{48}$$

$$\begin{aligned}
&\sup_{\underline{y}<y\le\bar{y}} \sup_{0\le t\le 1} \frac{1}{\sqrt{k}} \left| \sum_{i=1}^{\lfloor nt \rfloor} \left\{ \left(\log U_i - \log y^{-\frac{1}{\alpha}} b(n/k)\right)_+ \right.\right. \\
&\quad \left.\left. - \left(\log \tilde{U}_i - \log y^{-\frac{1}{\alpha}} b(n/k)\right)_+ \right\} \right| = o_P(1).
\end{aligned} \tag{49}$$

Note that Proposition 4 implies Theorem 2 under (1–6). Letting

$$\begin{aligned}
\tilde{M}_n^*(y, t) &:= \frac{1}{\sqrt{k}} \sum_{i=1}^{\lfloor nt \rfloor} \left\{ I\left(\tilde{U}_i \ge y^{-1/\alpha} b(n/k)\right) - P\left(U_i \ge y^{-1/\alpha} b(n/k)\right) \right\}, \\
\tilde{L}_n^*(y, t) &:= \frac{1}{\sqrt{k}} \sum_{i=1}^{\lfloor nt \rfloor} \left\{ \left(\log \tilde{U}_i - \log y^{-1/\alpha} b(n/k)\right)_+ - E\left(\log U_i - \log y^{-1/\alpha} b(n/k)\right)_+ \right\},
\end{aligned}$$

owing to (31), (41), (42), and Proposition 4, we can see that

$$\begin{aligned}
&\tilde{M}_n^*(y, t) \xrightarrow{d} B(y, t), \quad \tilde{L}_n^*(y, t) \xrightarrow{d} G(y, t) \quad \text{in } D([\underline{y}, \bar{y}] \times [0, 1]), \\
&\tilde{L}_n^*(y, t) - \alpha^{-1} \tilde{M}_n^*(y, t) \xrightarrow{d} \alpha^{-1} B(y, t) \quad \text{in } D\left([\underline{y}, \bar{y}] \times [0, 1]\right).
\end{aligned}$$

Thus, Theorem 2 can be established by following the lines in the proof of Theorem 1 with M_n^* and L_n^* replaced with $\tilde{M}_n^*$ and $\tilde{L}_n^*$, respectively.

Below, we only prove (48) since (49) can be proved in essentially the same way. For brevity, we remove the subscript from $\boldsymbol{\theta}_1$ and $\hat{\boldsymbol{\theta}}_1$, that is, $\boldsymbol{\theta}$ and $\hat{\boldsymbol{\theta}}$, respectively. Let

$$B = B(\boldsymbol{\theta}) = \begin{bmatrix} \boldsymbol{\beta}_{1\times(q-1)} & \beta_q \\ \mathrm{I}_{(q-1)\times(q-1)} & \mathbf{0}_{(q-1)\times 1} \end{bmatrix},$$

where $\boldsymbol{\beta}_{1\times(q-1)} = (\beta_1, \cdots, \beta_{q-1})$, $\mathrm{I}_{(q-1)\times(q-1)}$ denotes the $(q-1)\times(q-1)$ identity matrix, and $\mathbf{0}_{(q-1)\times 1}$ denotes the $(q-1)\times 1$ zero matrix. Set $B_\circ = B(\boldsymbol{\theta}^\circ)$, and let $\{h_i(\boldsymbol{\theta}) : i \in \mathbb{Z}\}$ be the unique strictly stationary sequence satisfying

$$\begin{bmatrix} h_i(\boldsymbol{\theta}) \\ \vdots \\ h_{i-q+1}(\boldsymbol{\theta}) \end{bmatrix} = B \begin{bmatrix} h_{i-1}(\boldsymbol{\theta}) \\ \vdots \\ h_{i-q}(\boldsymbol{\theta}) \end{bmatrix} + \begin{bmatrix} \omega + \sum_{j=1}^{p} f_j(\varepsilon_{i-j}) \\ \mathbf{0} \end{bmatrix},$$

where $f_j(x) = \psi_{1,j}(x)_+^{2\delta} + \psi_{2,j}(x)_-^{2\delta}$. From this, we get

$$h_i(\boldsymbol{\theta}) = \omega \sum_{j=0}^{\infty} B^j(1,1) + \sum_{j=0}^{\infty} B^j(1,1) \sum_{l=1}^{p} f_l(\varepsilon_{i-j-l}),$$

where $B^j(a,b)$ denotes the (a,b)-th entry of B^j. Finally, we set

$$N_n(\eta) = \left\{\boldsymbol{\theta} : |\boldsymbol{\theta} - \boldsymbol{\theta}^\circ| \le \frac{\eta}{n^\kappa}\right\} \quad \text{and} \quad N_n^-(\eta) = N_n(\eta) - \{\boldsymbol{\theta}^\circ\}, \quad \eta > 0,$$

which is a shrinking neighborhood of $\boldsymbol{\theta}^\circ$ of order $n^{-\kappa}$.

The following two standard lemmas (Lemmas 9 and 10) are useful to prove Lemma 11 below and can be verified using (4.3), the arguments up to (4.6) in Francq and Zakoïan (2004), and Lemma 8:

Lemma 8 *There exists $s > 0$ such that $E|\varepsilon_0|^s < \infty$.*

Lemma 9 *Let $\eta > 0$. Then, there exist $r \in (0,1)$ and $K > 0$, such that for large n,*

$$\sup_{\boldsymbol{\theta}\in N_n(\eta)} B^j(1,b) \le K r^j \quad \text{for } j \in \mathbb{N} \text{ and } b = 1, 2, \ldots, q, \tag{50}$$

and there exists a sequence of positive $\mathcal{F}_i$-measurable r.v.s $\{V_i\}$ such that $EV_i^\nu < \infty$ for some $\nu > 0$, and

$$\sup_{\boldsymbol{\theta}\in N_n(\eta)} h_i(\boldsymbol{\theta}) \le V_i. \tag{51}$$

Lemma 10 *There exists $r \in (0,1)$ and a $\mathcal{F}_0$-measurable r.v. $V \ge 0$, such that $EV^\nu < \infty$ for some $\nu > 0$ and for large n,*

$$\sup_{\boldsymbol{\theta}\in N_n(\eta)} \left|\hat{h}_i(\boldsymbol{\theta}) - h_i(\boldsymbol{\theta})\right| \le r^i V.$$

Let $h_i = h_i(\boldsymbol{\theta}^\circ)$. Then, we can express

$$h_i - h_i(\boldsymbol{\theta}) = \omega^\circ \sum_{j=0}^{\infty} B_\circ^j(1,1) - \omega \sum_{j=0}^{\infty} B^j(1,1) + \sum_{j=0}^{\infty}\sum_{l=1}^{p} B_\circ^j(1,1)\left\{f_l^\circ(\varepsilon_{i-j-l}) - f_l(\varepsilon_{i-j-l})\right\} + \sum_{j=0}^{\infty}\sum_{l=1}^{p}\left\{B_\circ^j(1,1) - B^j(1,1)\right\} f_l(\varepsilon_{i-j-l}),$$

where $f_j^\circ(x) = \psi_{1,j}^\circ(x)_+^{2\delta_\circ} + \psi_{2,j}^\circ(x)_-^{2\delta_\circ}$. Lemma 12 below plays a crucial role to allow an approximation of $h_1(\boldsymbol{\theta})$ to h_1 while $\boldsymbol{\theta}$ stays in a shrinking neighborhood of $\boldsymbol{\theta}^\circ$. The following lemma is useful to verify Lemma 12 that is used to establish Lemma 13 below.

Lemma 11 *Let $\epsilon > 0$. For each $j \in \mathbb{N}$,*

$$\{B((1+\epsilon)\boldsymbol{\theta}^\circ)\}^j(1,1) \le (1+\epsilon)^j B_\circ^j(1,1), \tag{52}$$

and there exists $n_0 \in \mathbb{N}$ such that for $n \ge n_0$ and $j \in \mathbb{N}$,

$$B_\circ^j(1,1) \le (1+\epsilon)^j \inf_{\boldsymbol{\theta}\in N_n(\eta)} B^j(1,1). \tag{53}$$

Lemma 12 *Let $\eta > 0$. Then, for every $w > 0$,*

$$\limsup_{n\to\infty} E\left|\sup_{\boldsymbol{\theta}\in N_n^-(\eta)} \frac{|h_1 - h_1(\boldsymbol{\theta})|}{|\boldsymbol{\theta}^\circ - \boldsymbol{\theta}|h_1(\boldsymbol{\theta})}\right|^w < \infty.$$

Lemma 15 below plays a main role to prove Proposition 4. Its proof is rather standard and similar to that of Lemma 3.1 in Ling and Peng (2004). Set $b_n = b(n/k)$,

$$\tilde{U}_i(\boldsymbol{\theta}) := U_i\left\{1+\frac{h_i(\boldsymbol{\theta}) - \hat{h}_i(\boldsymbol{\theta})}{\hat{h}_i(\boldsymbol{\theta})}\right\}^{1/(2\delta)}\left\{1+\frac{h_i - h_i(\boldsymbol{\theta})}{h_i(\boldsymbol{\theta})}\right\}^{1/(2\delta)}\{h_i\}^{1/(2\delta_\circ)-1/(2\delta)},$$

$$A_i(y,\boldsymbol{\theta}) := I\left(\tilde{U}_i(\boldsymbol{\theta}) > y^{-1/\alpha}b_n\right), \quad A_i(y) := I\left(U_i > y^{-1/\alpha}b_n\right).$$

Then, $\tilde{U}_i(\hat{\boldsymbol{\theta}}) = \tilde{U}_i$. The following two lemmas are useful to verify Lemma 15.

Lemma 13 *Let $\eta > 0$ and $m_n \to \infty$ with $m_n < n$ as $n \to \infty$. Then, there exist*

$$r_0 \in [0,1),\ \nu_0 > 0,\ a\ \mathcal{F}_{i-1}\text{-measurable r.v. } V_i \ge 0,\ \text{and } \Delta_{n,i} = \Delta_{n,i}(\eta) \ge 0 \tag{54}$$

such that $\sup_{i\in\mathbb{N}} E|V_i|^{\nu_0} < \infty$ *and*

$$\frac{1}{n}\sum_{i=1}^{n} E\Delta_{n,i} = O(1), \quad \max_{1\le i\le n} \frac{\Delta_{n,i}}{n^{\kappa}} = o_P(1) \quad as\, n \to \infty. \tag{55}$$

Moreover, there exists $\eta_0 > 0$*, such that with probability tending to 1 as* $n \to \infty$*,*

$$A_i(y,\eta,-\eta_0) \le A_i(y,\boldsymbol{\theta}) \le A_i(y,\eta,\eta_0) \quad for\ each\, \boldsymbol{\theta} \in N_n(\eta)\ and\, i = m_n,\dots,n, \tag{56}$$

where

$$A_i(y,\eta,\eta_0) := I\left(U_i\left\{1 + r_0^i\eta_0 V_i\right\}\left\{1 + \frac{\eta_0\Delta_{n,i}}{n^{\kappa}}\right\}\left\{1 + \frac{\eta_0 h_i^{\eta_0/n^{\kappa}}|\log h_i|}{n^{\kappa}}\right\} > y^{-\frac{1}{\alpha}}b_n\right). \tag{57}$$

Lemma 14 *Let* $\eta > 0$*,* $\epsilon_0 \in (0,1)$*, and* $r_1 \in (r_0, 1)$*, where* r_0 *is the one in (54). If* $m_n < n$ *and* $m_n \to \infty$ *as* $n \to \infty$ *and* $\eta_0 \in \mathbb{R}$*, we have that for* $i = m_n,\dots,n$*,*

$$\frac{n}{k} w_i \left|E_{i-1}\left\{A_i(y,\eta,\eta_0) - A_i(y)\right\}\right| \le K y \max\left\{r_1^i, \frac{|\eta_0|\Delta_{n,i}}{n^{\kappa}}, \frac{|\eta_0| h_i^{\eta_0/n^{\kappa}}|\log h_i|}{n^{\kappa}}\right\},$$

where

$$w_i = I\left(|\eta_0|\max\left\{\frac{\Delta_{n,i}}{n^{\kappa}}, \frac{h_i^{\eta_0/n^{\kappa}}|\log h_i|}{n^{\kappa}}\right\} \le \epsilon_0,\ r_0^i|\eta_0|V_i < r_1^i\right). \tag{58}$$

Lemma 15 *Let* $\eta > 0$*,* $\eta_0 \in \mathbb{R}$*,* $0 < \underline{y} < \bar{y} < \infty$ *and* $B_i(y,\eta,\eta_0) = A_i(y,\eta,\eta_0) - A_i(y)$*. Then,*

$$\sup_{\underline{y}\le y\le\bar{y}}\ \sup_{0\le t\le 1}\left|\frac{1}{\sqrt{k}}\sum_{i=1}^{\lfloor nt\rfloor} B_i(y,\eta,\eta_0)\right| = o_P(1).$$

The Proof of (48) in Proposition 4. Let $\eta > 0$ and let $\{m_n\}$ be a sequence of positive integers such that $m_n \to \infty$ and $m_n = o(\sqrt{k})$ as $n \to \infty$. Then, due to (56), there exists $\eta_0 > 0$ such that with probability tending to 1,

$$A_i(y,\eta,-\eta_0) - A_i(y) \le A_i(y,\boldsymbol{\theta}) - A_i(y) \le A_i(y,\eta,\eta_0) - A_i(y) \tag{59}$$

for $\boldsymbol{\theta} \in N_n(\eta)$ and $i = m_n, \ldots, n$, (cf. Lemma 13). Thus, from Lemma 15, we have that for every $\eta > 0$,

$$\sup_{\boldsymbol{\theta}\in N_n(\eta)} \sup_{\underline{y}<y<\bar{y}} \sup_{0\le t\le 1} \frac{1}{\sqrt{k}} \left| \sum_{i=1}^{\lfloor nt \rfloor} \{A_i(y,\boldsymbol{\theta}) - A_i(y)\} \right| = o_P(1).$$

Owing to this and (15), we assert the proposition. □

5.4 The proof of theorem 3

In this subsection, we assume that (1), (3), (4), (6), (17)–(20) hold.

Proposition 5 *Let* $0 < \underline{y} < \bar{y} < \infty$. *Then,*

$$\sup_{\underline{y}<y<\bar{y}} \sup_{0\le t\le 1} \frac{1}{\sqrt{k}} \left| \sum_{i=1}^{\lfloor nt \rfloor} \left\{ I\left(\bar{U}_i \ge y^{-\frac{1}{\alpha}} b(n/k)\right) - I\left(U_i \ge y^{-\frac{1}{\alpha}} b(n/k)\right) \right\} \right| = o_P(1), \tag{60}$$

$$\sup_{\underline{y}<y\le\bar{y}} \sup_{0\le t\le 1} \frac{1}{\sqrt{k}} \left| \sum_{i=1}^{\lfloor nt \rfloor} \left\{ \left(\log \bar{U}_i - \log y^{-\frac{1}{\alpha}} b(n/k)\right)_+ - \left(\log U_i - \log y^{-\frac{1}{\alpha}} b(n/k)\right)_+ \right\} \right| = o_P(1). \tag{61}$$

Below, we only prove (60) since (61) can be proven in essentially the same manner. We denote $\boldsymbol{\theta}_2$ and $\hat{\boldsymbol{\theta}}_2$ by $\boldsymbol{\theta}$ and $\hat{\boldsymbol{\theta}}$, respectively. Let $\{\varepsilon_i(\boldsymbol{\theta})\}$ and $\{\sigma_i^2(\boldsymbol{\theta})\}$ be the strictly stationary sequences satisfying

$$\begin{bmatrix} \varepsilon_i(\boldsymbol{\theta}) \\ \vdots \\ \varepsilon_{i-\bar{q}+1}(\boldsymbol{\theta}) \end{bmatrix} = B_1 \begin{bmatrix} \varepsilon_{i-1}(\boldsymbol{\theta}) \\ \vdots \\ \varepsilon_{i-\bar{q}}(\boldsymbol{\theta}) \end{bmatrix} + \begin{bmatrix} X_i - \sum_{j=1}^{\bar{p}} \phi_j X_{i-j} \\ \mathbf{0} \end{bmatrix},$$

and

$$\begin{bmatrix} \sigma_i^2(\boldsymbol{\theta}) \\ \vdots \\ \sigma_{i-q+1}^2(\boldsymbol{\theta}) \end{bmatrix} = B \begin{bmatrix} \sigma_{i-1}^2(\boldsymbol{\theta}) \\ \vdots \\ \sigma_{i-q}^2(\boldsymbol{\theta}) \end{bmatrix} + \begin{bmatrix} \omega + \sum_{j=1}^{p} \psi_j \varepsilon_{i-j}^2(\boldsymbol{\theta}) \\ \mathbf{0} \end{bmatrix},$$

respectively, where

$$B_1 = B_1(\boldsymbol{\theta}) = \begin{bmatrix} \vartheta_1 \ \cdots \ \vartheta_{\bar{q}-1} & \vartheta_{\bar{q}} \\ \mathrm{I}_{(\bar{q}-1)\times(\bar{q}-1)} & \mathbf{0}_{(\bar{q}-1)\times 1} \end{bmatrix}.$$

From this, we obtain

$$\varepsilon_i(\boldsymbol{\theta}) = \sum_{j=0}^{\infty} B_1^j(1,1) \left\{ X_{i-j} - \sum_{l=1}^{\bar{p}} \phi_j X_{i-j-l} \right\},$$
$$\sigma_i^2(\boldsymbol{\theta}) = \omega \sum_{j=0}^{\infty} B^j(1,1) + \sum_{j=0}^{\infty} B^j(1,1) \sum_{l=1}^{p} \psi_l \varepsilon_{i-j-l}^2(\boldsymbol{\theta}),$$

where $B_1^j(a,b)$ denotes the (a,b)-th entry of B_1^j. Let

$$\tilde{\sigma}_i^2(\boldsymbol{\theta}) = \omega^\circ \sum_{j=0}^{\infty} B_\circ^j(1,1) + \sum_{j=0}^{\infty} B_\circ^j(1,1) \sum_{l=1}^{p} \psi_l^\circ \varepsilon_{i-j-l}^2(\boldsymbol{\theta}).$$

As in the previous subsection, we set

$$N_n(\eta) = \left\{ \boldsymbol{\theta} : |\boldsymbol{\theta} - \boldsymbol{\theta}^\circ| \le \frac{\eta}{\sqrt{n}} \right\}, \quad N_n^-(\eta) = N_n(\eta) - \{\boldsymbol{\theta}^\circ\}, \quad \eta > 0.$$

The following two lemmas are due to Theorem 1 of Ling (2007) and the arguments up to (4.38) and (4.39) of Francq and Zakoïan (2004), respectively.

Lemma 16 $E|\varepsilon_0|^\nu < \infty$ *for every* $\nu \in (0,2)$.

Lemma 17 *There exists* $r \in (0,1)$ *and* $\mathcal{F}_{i-1}$*-measurable r.v.s* $V_i \ge 0$*, such that* $EV_i^\nu < \infty$ *for some* $\nu > 0$*, and*

$$\sup_{\boldsymbol{\theta} \in N_n(\eta)} \left| \hat{\varepsilon}_i(\boldsymbol{\theta}) - \varepsilon_i(\boldsymbol{\theta}) \right| \vee \left| \hat{\sigma}_i^2(\boldsymbol{\theta}) - \sigma_i^2(\boldsymbol{\theta}) \right| \le r^i V_i \quad \textit{for large } n \textit{ and each } i \in \mathbb{N}.$$

Lemmas 16, 17 and 18 below are used to verify some arguments useful to verify Proposition 5.

Lemma 18 *Let* $\eta > 0$ *and* $\nu \in (0,2)$*. Then, for every* $w > 0$*,*

$$\limsup_{n\to\infty} E \left\{ \sup_{\boldsymbol{\theta} \in N_n^-(\eta)} \frac{|\tilde{\sigma}_1^2(\boldsymbol{\theta}) - \sigma_1^2(\boldsymbol{\theta})|}{|\boldsymbol{\theta} - \boldsymbol{\theta}^\circ| \sigma_1^2(\boldsymbol{\theta})} \right\}^w < \infty, \tag{62}$$

and there exist $\Pi_{1,n,i} = \Pi_{1,n,i}(\eta, \upsilon) \ge 0$ *and* $\Pi_{2,n,i} = \Pi_{2,n,i}(\eta, \upsilon) \ge 0$*, such that*

$$\frac{|\sigma_i^2 - \tilde{\sigma}_i^2(\boldsymbol{\theta})|}{\tilde{\sigma}_i^2(\boldsymbol{\theta})} \le |\boldsymbol{\theta} - \boldsymbol{\theta}^\circ| \Pi_{1,n,i} + |\boldsymbol{\theta} - \boldsymbol{\theta}^\circ|^2 \Pi_{2,n,i}, \quad \textit{when } \boldsymbol{\theta} \in N_n(\eta), \tag{63}$$

and $\limsup_n E\{\Pi_{1,n,1}\}^\upsilon < \infty$ *and* $\limsup_n E\{\Pi_{2,n,1}\}^{\upsilon/2} < \infty$.

Now, let $\eta > 0$ and $v_0 \in (1, 2)$ such that

$$n^{1-v_0/2} = o(\sqrt{k}) \quad \text{as } n \to \infty \tag{64}$$

[cf. (3)]. Define

$$\bar{U}_i(\boldsymbol{\theta}) := U_i \left\{1 + \frac{\sigma_i^2(\boldsymbol{\theta}) - \hat{\sigma}_i^2(\boldsymbol{\theta})}{\hat{\sigma}_i^2(\boldsymbol{\theta})}\right\}^{\frac{1}{2}} \left\{1 + \frac{\sigma_i^2 - \sigma_i^2(\boldsymbol{\theta})}{\sigma_i^2(\boldsymbol{\theta})}\right\}^{\frac{1}{2}} + \frac{\hat{\varepsilon}_i(\boldsymbol{\theta}) - \varepsilon_i}{\hat{\sigma}_i(\boldsymbol{\theta})},$$

$$\Pi_{3,n,i} := \Pi_{3,n,i}(\eta) := \sup_{\boldsymbol{\theta} \in N_n^-(\eta)} \frac{1}{|\boldsymbol{\theta} - \boldsymbol{\theta}^\circ|} \frac{|\varepsilon_i(\boldsymbol{\theta}) - \varepsilon_i|}{\hat{\sigma}_i(\boldsymbol{\theta})} \in \mathcal{F}_{i-1}.$$

Then,

$$\limsup_n E\Pi_{3,n,1}^{v_0} < \infty, \tag{65}$$

(cf. Lemma 16) and $\bar{U}_i = \bar{U}_i(\hat{\boldsymbol{\theta}})$. Moreover, we have from Lemma 18 that when $\boldsymbol{\theta} \in N_n(\eta)$,

$$\begin{aligned}
\frac{\left|\sigma_i^2 - \sigma_i^2(\boldsymbol{\theta})\right|}{\sigma_i^2(\boldsymbol{\theta})} &\le \frac{\left|\sigma_i^2 - \tilde{\sigma}_i^2(\boldsymbol{\theta})\right|}{\tilde{\sigma}_i^2(\boldsymbol{\theta})} \cdot \frac{\tilde{\sigma}_i^2(\boldsymbol{\theta})}{\sigma_i^2(\boldsymbol{\theta})} + \frac{\left|\tilde{\sigma}_i^2(\boldsymbol{\theta}) - \sigma_i^2(\boldsymbol{\theta})\right|}{\sigma_i^2(\boldsymbol{\theta})} \\
&\le \left\{|\boldsymbol{\theta} - \boldsymbol{\theta}^\circ|\Pi_{1,n,i} + |\boldsymbol{\theta} - \boldsymbol{\theta}^\circ|^2\Pi_{2,n,i}\right\} \frac{\tilde{\sigma}_i^2(\boldsymbol{\theta})}{\sigma_i^2(\boldsymbol{\theta})} + \frac{\left|\tilde{\sigma}_i^2(\boldsymbol{\theta}) - \sigma_i^2(\boldsymbol{\theta})\right|}{\sigma_i^2(\boldsymbol{\theta})} \\
&= |\boldsymbol{\theta} - \boldsymbol{\theta}^\circ| \left\{\Pi_{1,n,i} \frac{\tilde{\sigma}_i^2(\boldsymbol{\theta})}{\sigma_i^2(\boldsymbol{\theta})} + \frac{\left|\tilde{\sigma}_i^2(\boldsymbol{\theta}) - \sigma_i^2(\boldsymbol{\theta})\right|}{|\boldsymbol{\theta} - \boldsymbol{\theta}^\circ|\sigma_i^2(\boldsymbol{\theta})}\right\} + |\boldsymbol{\theta} - \boldsymbol{\theta}^\circ|^2 \Pi_{2,n,i} \frac{\tilde{\sigma}_i^2(\boldsymbol{\theta})}{\sigma_i^2(\boldsymbol{\theta})}.
\end{aligned}$$

Further, owing to Lemma 18, we have

$$\Pi_{1,n,i}^* := \sup_{\boldsymbol{\theta} \in N_n^-(\eta)} \left\{\Pi_{1,n,i} \frac{\tilde{\sigma}_i^2(\boldsymbol{\theta})}{\sigma_i^2(\boldsymbol{\theta})} + \frac{\left|\tilde{\sigma}_i^2(\boldsymbol{\theta}) - \sigma_i^2(\boldsymbol{\theta})\right|}{|\boldsymbol{\theta} - \boldsymbol{\theta}^\circ|\sigma_i^2(\boldsymbol{\theta})}\right\}, \quad \Pi_{2,n,i}^* := \sup_{\boldsymbol{\theta} \in N_n(\eta)} \Pi_{2,n,i} \frac{\tilde{\sigma}_i^2(\boldsymbol{\theta})}{\sigma_i^2(\boldsymbol{\theta})},$$

so that

$$E|\Pi_{1,n,1}^*|^{v_0} < \infty, \quad E|\Pi_{2,n,1}^*|^{v_0/2} < \infty \quad \text{for large } n. \tag{66}$$

Note that due to (64), it follows from Lemma 18 that for every $\epsilon > 0$,

$$\frac{1}{\sqrt{k}} \sum_{i=1}^n I\left(\max\left\{\frac{|\eta_0|\Pi_{1,n,i}^*}{\sqrt{n}}, \frac{|\eta_0|^2\Pi_{2,n,i}^*}{n}, \frac{|\eta_0|\Pi_{3,n,i}}{\sqrt{n}b_n}\right\} > \epsilon\right) = o_P(1). \tag{67}$$

Now, let $A_i(y, \boldsymbol{\theta}) := I(\bar{U}_i(\boldsymbol{\theta}) > y^{-\frac{1}{\alpha}} b_n)$, $A_i(y) := I(U_i > y^{-\frac{1}{\alpha}} b_n)$ and

$$\begin{aligned}
A_i(y, \eta, \eta_0) := I\Bigg(&U_i \left\{1 + r_0^i \eta_0 V_i\right\} \left\{1 + \frac{\eta_0 \Pi_{1,n,i}^*}{\sqrt{n}} + \operatorname{sgn}(\eta_0) \frac{\eta_0^2 \Pi_{2,n,i}^*}{n}\right\} \\
&+ \frac{\eta_0 \Pi_{3,n,i}}{\sqrt{n}} + r_0^i \eta_0 V_i > y^{-\frac{1}{\alpha}} b_n\Bigg),
\end{aligned}$$

for $y \in [\underline{y}, \bar{y}]$ and $\eta_0 \in \mathbb{R}$, where $r_0 \in [0, 1)$ and the $\mathcal{F}_{i-1}$-measurable $V_i \geq 0$ are obtained from Lemma 17. Using the arguments between (64) and (67), we can obtain Lemmas 19 and 20: the former is used to verify the latter.

Lemma 19 *Let $\eta > 0$, $\epsilon_0 \in (0, 0.1)$, $r_1 \in (r_0, 1)$, and $\{m_n\}$ be a sequence such that $m_n < n$ and $m_n \to \infty$ as $n \to \infty$. Then, there exists $\eta_0 > 0$ such that when $\boldsymbol{\theta} \in N_n(\eta)$,*

$$w_i A_i(y, \eta, -\eta_0) \leq w_i A_i(y, \boldsymbol{\theta}) \leq w_i A_i(y, \eta, \eta_0) \quad \textit{for each } i = m_n, \ldots, n,$$

where

$$w_i := I\left(\max\left\{\frac{|\eta_0|\Pi^*_{1,n,i}}{\sqrt{n}}, \frac{|\eta_0|^2\Pi^*_{2,n,i}}{n}, \frac{|\eta_0|\Pi_{3,n,i}}{\sqrt{n}b_n}\right\} < \epsilon_0,\ r_0^i|\eta_0|V_i < r_1^i\right) \tag{68}$$

and

$$\begin{aligned} &\frac{n}{k} w_i |E_{i-1}\{A_i(y, \eta, \eta_0) - A_i(y)\}| \\ &\quad \leq K w_i \max\left\{r_1^i, \frac{|\eta_0|\Pi^*_{1,n,i}}{\sqrt{n}}, \frac{|\eta_0|^2\Pi^*_{2,n,i}}{n}, \frac{|\eta_0|\Pi_{3,n,i}}{\sqrt{n}b_n}\right\}. \end{aligned} \tag{69}$$

In particular,

$$\frac{1}{\sqrt{k}} \sum_{i=m_n}^{n} (1 - w_i) = o_P(1), \tag{70}$$

$$\sqrt{k}\log n \cdot E\frac{|\eta_0|^2\Pi^*_{2,n,1}}{n} I\left(\frac{|\eta_0|^2\Pi^*_{2,n,1}}{n} < \epsilon_0\right) = o(1) \quad \textit{as } n \to \infty. \tag{71}$$

Lemma 20 *Let $\eta > 0$, $\eta_0 \in \mathbb{R}$, $0 < \underline{y} < \bar{y} < \infty$, and $B_i(y, \eta, \eta_0) = A_i(y, \eta, \eta_0) - A_i(y)$. Then, we have*

$$\sup_{\underline{y} \leq y \leq \bar{y}} \sup_{0 \leq t \leq 1} \left|\frac{1}{\sqrt{k}} \sum_{i=1}^{\lfloor nt \rfloor} B_i(y, \eta, \eta_0)\right| = o_P(1).$$

The Proof of (60) in Proposition 5. The proposition can be verified using Lemmas 19 and 20 and following the lines in the proof of (48). □

Acknowledgments We thank the editor, an AE and two anonymous referees for their valuable comments and suggestion. This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MEST) (No. 2012R1A2A2A01046092) (S. Lee) and by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (Grant Number: NRF-2012R1A6A3A01039754) (M. Kim).

References

Berkes, I., Horváth, L. (2004). The efficiency of the estimators of the parameters in GARCH processes. *Annals of Statististics*, *32*, 633–655.

Bickel, P.J., Wichura, M.J. (1971). Convergence criteria for multiparameter stochastic processes and some applications. *Annals of Mathematical Statististics*, *42*, 1656–1670.

Billingsley, P. (1999). *Convergence of probability measures. 2nd ed. Wiley Series in Probability and Statistics: Probability and Statistics*. New York: Wiley

Bougerol, P., Picard, N. (1992). Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics*, *52*, 115–127.

Chan, N.H., Deng, S.-J., Peng, L., Xia, Z. (2007). Interval estimation of value-at-risk based on GARCH models with heavy-tailed innovations. *Journal of Econometrics*, *137*, 556–576.

Csörgő, M., Horváth, L. (1997). *Limit theorems in change-point analysis. Wiley Series in Probability and Statistics*. Chichester: Wiley (With a foreword by David Kendall).

Csörgő, S., Deheuvels, P., Mason, D. (1985). Kernel estimates of the tail index of a distribution. *Annals of Statistics*, *13*, 1050–1077.

Drees, H. (2003). Extreme quantile estimation for dependent data, with applications to finance. *Bernoulli*, *9*, 617–657.

Drees, H., de Haan, L., Resnick, S. (2000). How to make a Hill plot. *Annals of Statististics*, *28*, 254–274.

Feuerverger, A., Hall, P. (1999). Estimating a tail exponent by modelling departure from a Pareto distribution. *Annals of Statistics*, *27*, 760–781.

Francq, C., Zakoïan, J.-M. (2004). Maximum likelihood estimation of pure GARCH and ARMA–GARCH processes. *Bernoulli*, *10*, 605–637.

Goldie, C. M., Smith, R. L. (1987). Slow variation with remainder: theory and applications. *Quarterly Journal of Mathematics: Oxford Series (2)*, *38*, 45–71.

Gomes, M.I., de Haan, L., Rodrigues, L.H. (2008). Tail index estimation for heavy-tailed models: accommodation of bias in weighted log-excesses. *Journal of Royal Statistical Society, Series B, Statistical Methodology*, *70*, 31–52.

Hall, P. (1982). On some simple estimates of an exponent of regular variation. *Journal of Royal Statistical Society, Series B, Statistical Methodology*, *44*, 37–42.

Hall, P., Welsh, A.H. (1985). Adaptive estimates of parameters of regular variation. *Annals of Statistics*, *13*, 331–341.

Higgins, M.L., Bera, A.K. (1992). A class of nonlinear arch models. *International Economic Review*, *33*, 137–158.

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

Hill, J.B. (2010). On tail index estimation for dependent, heterogeneous data. *Econometric Theory*, *26*, 1398–1436.

Hillebrand, E. (2005). Neglecting parameter changes in GARCH models. *Journal of Econometrics*, *129*, 121–138.

Hsing, T. (1991). On tail index estimation using dependent data. *Annals of Statistics*, *19*, 1547–1569.

Kim, M., Lee, S. (2008). Estimation of a tail index based on minimum density power divergence. *Journal of Multivariate Analysis*, *99*, 2453–2471.

Kim, M., Lee, S. (2009). Test for tail index change in stationary time series with Pareto-type marginal distribution. *Bernoulli*, *15*, 325–356.

Lee, S., Lee, T. (2012). Inference for box–cox transformed threshold garch models with nuisance parameters. *Scandinavian Journal of Statistics*, *39*, 568–589.

Ling, S. (2007). Self-weighted and local quasi-maximum likelihood estimators for ARMA–GARCH/IGARCH models. *Journal of Econometrics*, *140*, 849–873.

Ling, S., Peng, L. (2004). Hill's estimator for the tail index of an ARMA model. *Journal of Statistical Planning and Inference*, *123*, 279–293.

Mason, D.M. (1982). Laws of large numbers for sums of extreme values. *Annals of Probability*, *10*, 754–764.

Mikosch, T., Stărică, C. (2000). Limit theory for the sample autocorrelations and extremes of a GARCH (1, 1) process. *Annals of Statistics*, *28*, 1427–1451.

Pan, J., Wang, H., Tong, H. (2008). Estimation and tests for power-transformed and threshold GARCH models. *Journal of Econometrics*, *142*, 352–378.

Quintos, C., Fan, Z., Phillips, P.C.B. (2001). Structural change tests in tail behaviour and the Asian crisis. *The Review of Economics Studies*, *68*, 633–663.
Resnick, S., Stărică, C. (1997a). Asymptotic behavior of Hill's estimator for autoregressive data. *Communications in Statistics, Stochastic Models*, *13*, 703–721.
Resnick, S., Stărică, C. (1997b). Smoothing the Hill estimator. *Advances in Applied Probability*, *29*, 271–293.
Resnick, S., Stărică, C. (1998). Tail index estimation for dependent data. *Annals of Applied Probability*, *8*, 1156–1183.
Zhu, K., Ling, S. (2011). Global self-weighted and local quasi-maximum exponential likelihood estimators for ARMA–GARCH/IGARCH models. *Annals of Statistics*, *39*, 2131–2163.