Erasmus University Rotterdam

Erasmus School of Economics

Effects of Different QMLEs on the EVT Approach in Estimation of Conditional Tail Risk Measures

Bachelor Thesis BSc^2 Econometrics/Economics

Author:
Haobai Guo

Student number:
477388

Supervisor:
prof.dr. Chen Zhou

Second assessor:
Bram van Os

Abstract

ARMA-GARCH models are commonly used nowadays in modelling returns and volatilities in the financial market. The risk measures are essentially about the tail behavior of the innovations from the ARMA-GARCH models, which can be estimated using the extreme value theory (EVT). By considering consistent non-Gaussian Quasi-Maximum Likelihood Estimator in estimating ARMA-GARCH models, we improve the performance of the EVT approach in conditional Value-at-Risk and conditional expected shortfall estimations. Furthermore, we prove that the gain in performance of the EVT approach can also be achieved when certain inconsistent non-Gaussian Quasi-Maximum Likelihood Estimators are used. We also consider a parametric method with skewed distribution and the ARMA-GJR-GARCH model in this study.

July 4, 2021

ERASMUS UNIVERSITEIT ROTTERDAM

The views stated in this thesis are those of the author and not necessarily those of the supervisor, second assessor, Erasmus School of Economics or Erasmus University Rotterdam.

Acknowledgements

I am grateful to prof. dr. Chen Zhou for the thorough thesis supervision. The helpful suggestions and thoughtful comments from prof. dr. Chen Zhou have inspired me and greatly improved the quality of this thesis.

Contents

1 Introduction

In the past decades, it has always been the primary concern of financial institutions and central banks to be able to accurately capture and forecast market risk. In order to be able to cover unexpected losses, banks are required by the central banks to maintain a sufficient amount of capital against future adverse movements of the market. Financial risk measures are then needed to estimate the risk that a bank is exposed to. The ability of having an idea of future risks is also important to financial institutions themselves as it is crucial for internal risk management. These real-life needs naturally lead to the interests in finding prominent financial risk measures.

Two well-known stylized features of financial returns are heteroskedasticity and heavy-tailed distribution. To take heteroskedasiticy into account, volatility models such as generalized autoregressive conditional heteroskedasticity (GARCH) were developed (Bollerslev, 1986). Similar to the volatility models where the conditional and stationary distribution of the volatility are differentiated, risk measure can be classified into conditional risk measures and stationary risk measures. Though both risk measures provide insightful information on the returns, all the important applications mentioned above are about the conditional risk measures. In the rest of this paper, we will focus on the conditional risk measures.

One of the most straightforward risk measures is Value-at-Risk (VaR). The conditional VaR for a future period of a portfolio is essentially a quantile of the distribution of the return, which is defined over the given period conditioning on all the past information. The distribution of the return on a portfolio is also often called the Profit-and-Loss (P&L) distribution. Another popular risk measure is conditional expected shortfall (ES), which is the expected value of the return given it is smaller or larger than a certain value conditioning on the past information. Many research on estimation methods of these two risk measure have been done in the past years, see, for example, Taylor (2008), Taylor (2007) and Patton et al. (2019). Apart from their popularity in recent research, VaR and ES are also particularly relevant to banks and central banks nowadays. In the most recently revised version of the minimal capital requirements (Basel Committee, 2019) published by Bank for International Settlements, ES is used as a risk measure in the internal model approach and VaR is required in the backtesting for this approach.

Both VaR and ES of a portfolio are directly related to the tail of its P&L distribution. The research on estimating these two risk measure are then essentially about estimating the tail of the P&L distribution. Two of the most basic ways of estimating the P&L distributions are the historical simulation and parametric method based on the volatility models. However, both of these methods suffer from certain drawbacks. In the historical simulation method, we simply use the historical data to calculate the empirical distribution of the returns. Though historical simulation is easy to implement, it naturally imposes the assumption that the distribution of the return on a portfolio is the same over time and hence homoskedastic, which does not fit the stylized features of financial returns. Moreover, historical simulation requires a large number of observations to provide a good estimation of the tail, which is not always feasible in practice. Parametric methods based on volatility models are very straightforward to implement once the models are estimated.

We can forecast the volatility of the return based on the chosen volatility model. The forecast volatility together with the underlying distribution assumed by the model directly lead to the forecast distribution of future returns. This type of methods are expected to perform well given that the assumption of underlying data generating process (DGP) is correct. However, making correct assumptions about the distribution of the innovation can be hard when using real data. And in practice, conditional normality is often assumed, which does not provide a good fit for many financial return series.

Another class of approaches to estimate VaR and ES is based on the extreme value theory (EVT). Some examples of early research in EVT methods are Daníelsson and de Vries (1997) and Embrechts et al. (1999). McNeil and Frey (2000) proposes an EVT method combining volatility models and historical simulation for the estimation of VaR and ES. Their idea is to first fit an AR-GARCH model using the Maximum Likelihood Estimation (MLE) to filter out the residuals, and then estimate the distribution of the residuals using EVT and historical simulation. They focused on a general EVT set-up by assuming that the tail of the distribution can be approximated by a generalized Pareto distribution (GPD). Furthermore, they showed that their method outperforms the unconditional EVT and volatility modelling assuming GARCH and normality. As an addition to their work, Chan et al. (2007) derived the statistical properties of the VaR estimator proposed by McNeil and Frey (2000). They instead considered a semi-parametric method using the Hill estimator (Hill, 1975) that focuses on the situation of heavy-tailed innovations. Moreover, they replaced MLE by the Gaussian Quasi-Maximum Likelihood Estimator (GQMLE) when estimating the GARCH model.

The recent work from Hoga (2019a) generalized the two-step approach from McNeil and Frey (2000). For the estimation of VaR and ES, they considered the EVT method combining with the general ARMA-GARCH model. For financial time series, the ARMA-GARCH model is considered as a benchmark model nowadays. Hoga (2019a) also derived the limit distribution of their estimators and proposed the use of self-normalization in the construction of confidence interval for their estimators. Furthermore, they generalized the non-parametric EVT method in McNeil and Frey (2000) and Chan et al. (2007) by considering other extreme value index estimator. They also considered a different method in selecting the number of order statistics k that are used for tail estimation. For the estimation of ARMA-GARCH model in their simulation study, in addition to the GMQLE, they considered a Laplace Quasi-Maximum Likelihood Estimator which is a special case of the non-Gaussian Quasi-Maximum Likelihood Estimator (NGQMLE) proposed by Berkes and Horváth (2004).

In a simulation study, Hoga (2019a) showed that the data dependent k leads to near-optimal results. However, there is no clear comparison to the k selection method used by Chan et al. (2007). Furthermore, the simulation study also does not show the benefit of considering an ARMA-GARCH model instead of AR-GARCH or GARCH. Hoga (2019a) also used an empirical study to compared the relative performance of their method to the method proposed by Chan et al. (2007) and showed that their own method performs relatively better. However, as showed in the above discussion, there are two major differences between the two methods. Hoga (2019a) considered a general ARMA-GARCH setting with data dependent k, whereas Chan et al. (2007) considered a GARCH model with fixed k given the sample size. Consequently, their result is unclear about what is the exact cause of the improved performance in their method. The first goal of this paper is to construct a more complete simulation and empirical study, where we study alternative choices for each part of the two-step EVT approach in detail.

Next, we consider an improved parametric method for VaR and ES estimation. One major challenge of the parametric method based on the volatility model is to make reasonable assumptions about the distributions of the innovations. McNeil and Frey (2000) suggested that the parametric method with a GARCH model assuming the Student's t-distribution can work quite well when the data is symmetric. However, apart from being heavy-tailed, empirical evidence (Diebold, 2012) has suggested that the innovations from financial data are sometimes asymmetrically distributed. This suggests that neither the Normal distribution nor the Student's t-distribution can be used as a satisfying assumption in practice. In this paper, we consider a parametric method based on an ARMA-GARCH model, where the innovations are assumed to follow the skew t-distribution proposed by Jones and Faddy (2003). This method is then used as the baseline-method when evaluating the performance of the EVT methods.

Thirdly, we further generalize the ARMA-GARCH model from Hoga (2019a) by considering an asymmetric ARMA-GARCH model. Past research have documented asymmetric relation between stock returns and volatility (see, for example, Nelson (1991)), which motivated extended GARCH models that captures this asymmetric relation. In this paper, additional to the standard ARMA-GARCH model, we consider a ARMA-GJR-GARCH model, where the volatility process is modeled using the GJR-GARCH model proposed by Glosten et al. (1993).

Fourthly, we consider a different estimation method for the ARMA-GARCH model. In particular, we consider the three-step NGQMLE approach proposed by Fan et al. (2014). The first-step in the two-step EVT approach requires good estimates of the ARMA-GARCH coefficients. It is then non-trivial to find a consistent and efficient estimation method when estimating the ARMA-GARCH model. As mentioned earlier, McNeil and Frey (2000) considered the MLE, which might be inconsistent when the assumed distribution is different from the true distribution. The first-step estimation procedure was then improved by Chan et al. (2007) by considering the GQMLE, which is proven to be consistent and asymptotically normal when the true innovation distribution has finite fourth moment (Francq and Zakoïan, 2004). Hoga (2019a) further considered the NGQMLE in a simulation study when the true distribution has infinite fourth moment, which was shown by Berkes and Horváth (2004) to be consistent without the finite fourth moment requirement if some alternative moment conditions are met. Another advantage of NGQMLE is that it is more efficient in general comparing to the GQMLE when the true distribution of the innovations is non-Gaussian (Fan et al., 2014). However, the alternative moment conditions required by the consistency of NGQMLE are in general not met by the common GARCH representation. For the consistency of NGQMLE, it is then needed to impose new identifiability constraint based on the chosen non-Gaussian density (see, for example, Francq and Zakoian (2019)). Fan et al. (2014) showed that in general the simple NGQMLE without special moment conditions is inconsistent and they argued that moment conditions should be determined before choosing the estimation method, which is the case for most research. Fan et al. (2014) then proposed a three-step consistent NGQMLE, which is practical, robust and more efficient comparing to GQMLE. Another advantage of the three-step approach is that it can be applied to more general ARMA-GARCH models. We try to study whether a more efficient estimation procedure for ARMA-GARCH can result in better VaR and ES estimates using the two-step EVT methods.

Lastly, we prove that though the simple NGQMLE is in general inconsistent, the conditional VaR and conditional ES estimates using the EVT approach are in fact not affected by the resulting inconsistent coefficient estimates. This result is important as it shows that the inconsistent simple NGQMLE, which is more efficient than GQMLE when innovations are heavy-tailed, can be used as a valid estimator in the two-step EVT approach. More generally, our result shows that the EVT method is unaffected by a class of incorrect estimation methods.

The reminding of this paper is structured as the following. Section 2 contains the main methodology. In Section 3 we first replicate important simulation results from Hoga (2019a), after which we use Monte Carlo simulation to make a thorough comparison among the previously mentioned methods. Section 4 conducts analysis on the real dataset which is also used by Hoga (2019a). Section 5 concludes.

2 Methodology

2.1 The ARMA-GARCH Model

Let X_i be the loss returns, which are the original returns multiplied by -1, on a portfolio. We consider an ARMA(p,q) model

$$X_i = \sum_{j=1}^{p} \phi_j X_{i-j} + \varepsilon_i - \sum_{j=1}^{q} \nu_j \varepsilon_{i-j} \tag{1}$$

with generalized GARCH(h, k) errors:

$$\varepsilon_i = \sigma_i U_i, \quad U_i \sim F.i.i.d.(0,1), \tag{2}$$

$$\sigma_i^2 = g(\varepsilon_{i-1}, ..., \varepsilon_{i-h}, \sigma_{i-1}, ..., \sigma_{i-k}, \rho) \tag{3}$$

where ρ is a parameter vector and F is some continuous distribution with mean zero and unit variance. An example of the function in (3) is:

$$g(\varepsilon_{i-1}, ..., \varepsilon_{i-h}, \sigma_{i-1}, ..., \sigma_{i-k}, \rho) = \omega + \sum_{j=1}^{h} \psi_j \varepsilon_{i-j}^2 + \sum_{j=1}^{k} \beta_j \sigma_{i-j}^2, \text{where} \quad \rho = (\omega, \psi_1, ..., \psi_h, \beta_1, ..., \beta_k),$$

such that (2) and (3) define the standard GARCH(h, k) errors. More complicated GARCH type models can also be represented by $g(\cdot)$, such as the EGARCH model with:

$$g(\varepsilon_{i-1}, ..., \varepsilon_{i-h}, \sigma_{i-1}, ..., \sigma_{i-k}, \rho) = \exp\left(\omega + \sum_{j=1}^{h}\left\{\psi_j(\left|\frac{\varepsilon_{i-j}}{\sigma_{i-j}}\right| - E[\left|\frac{\varepsilon_{i-j}}{\sigma_{i-j}}\right|] + \alpha_j \frac{\varepsilon_{i-j}}{\sigma_{i-j}})\right\} + \sum_{j=1}^{k}\beta_j \log(\sigma_{i-j}^2)\right),$$

where $\rho = (\omega, \psi_1, ..., \psi_h, \alpha_1, ..., \alpha_h, \beta_1, ..., \beta_k)$,

and the GJR-GARCH model with:

$$g(\varepsilon_{i-1}, ..., \varepsilon_{i-h}, \sigma_{i-1}, ..., \sigma_{i-k}, \rho) = \omega + \sum_{j=1}^{h}\psi_j \varepsilon_{i-j}^2 + \sum_{j=1}^{h}\phi_j \varepsilon_{i-j}^2 I(\varepsilon_{i-j} < 0) + \sum_{j=1}^{k}\beta_j \sigma_{i-j}^2,$$

where $I(\cdot)$ is the indicator function and $\rho = (\omega, \psi_1, ..., \psi_h, \phi_1, ..., \phi_h, \beta_1, ..., \beta_k)$.

Following from Hoga (2019a), we impose adjusted standard assumptions on the ARMA-GARCH model as the following:

Assumption 1. The characteristic polynomials of the AR and the MA terms have roots inside the unit circle and do not share common root.

Assumption 2. ρ is restricted such that volatility is guaranteed to be positive and there exist unique stationary to the GARCH equation.

There are two additional mild conditions on the model, for which we refer to Assumption 3 and 4 from Hoga (2019a).

The right-tail one-step ahead 100α ($\alpha \in (0, 1)$) percent conditional VaR can then be defined as

$$x_{\alpha,n} := \inf\{x : P(X_{n+1} \leq x | X_{n+1-k}, k \geq 1) \geq 1 - \alpha, x \in \mathbb{R}\}.$$

Note that the right-tail VaR for loss returns is the left-tail VaR for original returns, which is in the interest of finance. CES is defined as the expected return given the return is in its best (or worst) 100α percent case conditioning on the past information. Using the previous definition on conditional VaR, we can define the right-tail one-step ahead 100α ($\alpha \in (0, 1)$) percent CES as

$$S_{\alpha,n} := E[X_{n+1} | X_{n+1} > x_{\alpha,n}, X_n, X_{n-1}, ..., X_1].$$

The goal is to estimate $x_{\alpha,n}$ and $S_{\alpha,n}$, which can be rewritten into the following as in Hoga (2019a):

$$x_{\alpha,n} = \left[\mu_{n+1} + \sigma_{n+1} x_\alpha^U | X_n, X_{n-1}, ..., X_1\right], \tag{4}$$

$$S_{\alpha,n} = \left[\mu_{n+1} + \sigma_{n+1} S_\alpha^U | X_n, X_{n-1}, ..., X_1\right], \tag{5}$$

where

$$\mu_{n+1} = \sum_{j=1}^{p} \phi_j X_{n+1-j} - \sum_{j=1}^{q} \nu_j \varepsilon_{n+1-j},$$

x_α^U is defined as the $(1-\alpha)$ quantile of U_i and $S_\alpha^U = E[U_i | U_i > x_\alpha^U]$ for any i. Hoga (2019a) showed that μ_{n+1} can be recursively written into a function of $X_1, ..., X_n$ and the ARMA coefficients from (1). We can then consistently estimate μ_{n+1} if we can consistently estimate the ARMA-GARCH model defined in (1), (2) and (3).

We can rewrite (1) and (3) into the following estimators for ε_i and σ_i^2 respectively, $i = 1, ..., n$:

$$\hat{\varepsilon}_i = \tilde{X}_i - \sum_{j=1}^{p} \phi_i \tilde{X}_{i-j} + \sum_{j=1}^{q} \nu_j \hat{\varepsilon}_{i-j} \tag{6}$$

$$\hat{\sigma}_i^2 = g(\hat{\varepsilon}_{i-1}, ..., \hat{\varepsilon}_{i-h}, \hat{\sigma}_{i-1}, ..., \hat{\sigma}_{i-k}, \rho), \tag{7}$$

where $\tilde{X}_i = X_i$ for $i = 1, ..., n$ and $\tilde{X}_0 = \tilde{X}_{-1} = ... = \hat{\varepsilon}_0 = \hat{\varepsilon}_{-1} = ... = \hat{\sigma}_0 = \hat{\sigma}_{-1} = ... = 0$. We then also have an estimator for μ_{n+1}:

$$\hat{\mu}_{n+1} = \sum_{j=1}^{p} \phi_j \tilde{X}_{n+1-j} - \sum_{j=1}^{q} \nu_j \hat{\varepsilon}_{n+1-j},$$

The only reminding part that still need to be estimated in (4) and (5) is then x_α^U and S_α^U. We estimate x_α^U and S_α^U using the parametric and the EVT method which are introduced in later sections.

2.2 EVT Based Estimation of the Tail

We define the residual $\hat{U}_i = \hat{\varepsilon}_i / \hat{\sigma}_i$, where $\hat{\varepsilon}_i$ and $\hat{\sigma}_i$ are defined in (6) and (7). We are interested in estimating x_α^U and S_α^U assuming that the distribution of U_i's is heavy-tailed. As proposed by Hoga (2019a), we assume the following:

$$\lim_{x \to \infty} \frac{U(xy)}{U(x)} = y^\gamma \quad \text{for all} \quad y > 0, \tag{8}$$

where $U(x)$ is the $(1 - 1/x)$-quantile of F, which was defined earlier as the distribution of U_i. We assume $\gamma > 0$ so that F is heavy-tailed. With a sample $U_1, ...U_n$, it is then important to determine which observations are from the tail. We assume that there are k out of the n observations belonging to the right tail, where $k = k_n$. We then have an integer sequence $\{k_n\}$, on which we impose the following assumption as in Hoga (2019a):

$$k = k_n \to \infty \quad \text{with} \quad 1 \leq k < n \quad \text{and} \quad k/n \to 0 \quad \text{as} \quad n \to \infty \tag{9}$$

Now for any $\alpha \ll k/n$, we can estimate x_α^U using the Weissman (1978) estimator. We take the definition of this estimator from Hoga (2019a), which is as presented below:

$$\hat{x}_\alpha^U := U\left(\frac{1}{\alpha}\right) \approx \widehat{U\left(\frac{n}{k}\right)} \left(\frac{n\alpha}{k}\right)^{-\hat{\gamma}}. \tag{10}$$

Let the order statistics of $\{\hat{U}_i\}_{i=1}^n$ be defined as $U_{1:n} \leq U_{2:n} \leq ... \leq U_{n:n}$. We can then estimate $U\left(\frac{n}{k}\right)$ using the $(n-k)$th order statistic of U_n denoted by $U_{n-k:n}$. To estimate γ Hoga (2019a) considered two estimators. The first one considered by him is the Hill (1975) estimator

$$\hat{\gamma}_H := \frac{1}{k}\sum_{i=0}^{k-1} \log\left(\frac{U_{n-i:n}}{U_{n-k:n}}\right), \tag{11}$$

which was also considered in Chan et al. (2007) and McNeil and Frey (2000). The additional estimator for γ employed in Hoga (2019a) is the so-called MR estimator from Danielsson et al. (1996) that can be defined as the following:

$$\hat{\gamma}_{MR} = \frac{1}{2}\frac{\frac{1}{k}\sum_{i=0}^{k-1}\{\log(U_{n-i:n}) - \log(U_{n-k:n})\}^2}{\hat{\gamma}_H} \tag{12}$$

With the assumption in (8) and $0 < \gamma < 1$, Hoga (2019a) suggested the following estimator for S_α^U:

$$\hat{S}_\alpha^U := \frac{\hat{x}_\alpha^U}{1-\hat{\gamma}} \tag{13}$$

The value of k need to be determined for all the above estimators. We want to pick k such that the approximation using GPD fits $U_{n-k+1:n}, U_{n-k+2:n}, ..., U_{n:n}$ reasonably well. Hoga (2019a) proposed the following estimator:

$$k^* := \underset{k=k_{\min},...,k_{\max}}{\arg\min}\left[\sup_{j=1,...,k_{\max}}\left|U_{n-j,n} - U_{n-k,n}\left(\frac{j}{k}\right)^{-\hat{\gamma}}\right|\right]. \tag{14}$$

2.3 Confidence Interval for EVT based conditional VaR and conditional ES

To analyse the asymptotics of $\hat{\gamma}_H$ and $\hat{\gamma}_{MR}$, Hoga (2019a) considered stochastic processes $\hat{\gamma}_H(t)$ and $\hat{\gamma}_{MR}(t)$, which are constructed using a refined sequential tail empirical process (see also Einmahl et al. (2016)). These stochastic processes can be estimated as the following:

$$\hat{\gamma}_H(t) = \frac{1}{\lfloor kt \rfloor}\sum_{i=1}^{\lfloor kt \rfloor} \log\left(\frac{U_k(t,i)}{U_k(t,0)}\right), \tag{15}$$

$$\hat{\gamma}_{MR} = \frac{1}{2}\frac{\frac{1}{\lfloor kt \rfloor}\sum_{i=1}^{\lfloor kt \rfloor}\{\log(U_k(t,i)) - \log(U_k(t,0))\}^2}{\hat{\gamma}_H(t)}, \tag{16}$$

where $U_k(t,i)$ is the $(\lfloor kt \rfloor + 1 - i)$th largest value of $U_{m_n}, ..., U_{\lfloor nt \rfloor}$, $m_n \to \infty$, $m_n < n$ and $m_n = o(\sqrt{k})$ when $n \to \infty$. Instead of using all observations, Hoga (2019a) exclude the first $m_n - 1$ observations, which takes the initialization effect into account.

Hoga (2019a) then showed that, let $\hat{\gamma}(t)$ be either $\hat{\gamma}_H(t)$ or $\hat{\gamma}_{MR}(t)$, (10) and (13) can be adapted into the following:

$$\hat{x}_\alpha^U(t) := \hat{U}_k(t,0)\left(\frac{n\alpha}{k}\right)^{-\hat{\gamma}(t)}, \tag{17}$$

$$\hat{S}_\alpha^U(t) := \frac{\hat{x}_\alpha^U(t)}{1-\hat{\gamma}(t)}, \tag{18}$$

which can then be used with (4) and (5) to obtain the following estimators:

$$\hat{x}_{\alpha,n}(t) = \hat{\mu}_{n+1} + \hat{\sigma}_{n+1} x_\alpha^U(t), \tag{19}$$

$$\hat{S}_{\alpha,n}(t) = \hat{\mu}_{n+1} + \hat{\sigma}_{n+1} S_\alpha^U(t), \tag{20}$$

A note here is that $\hat{x}_{\alpha,n}(1)$ and $\hat{S}_{\alpha,n}(1)$ are essentially the estimators in (11) and (12) after adjusting for the initial effect. In the reminding of this paper we estimate $x_{\alpha,n}$ and $S_{\alpha,n}$ for the whole sample using $\hat{x}_{\alpha,n}(1)$ and $\hat{S}_{\alpha,n}(1)$.

With all the previous assumptions we mentioned in this section and some additional mild conditions (see Hoga (2019a) Section 2.4), Hoga (2019a) showed the following asymptotic distribution of $z \in x, S$:

$$\frac{1}{\hat{\sigma}_{\hat{\gamma},\gamma}} \frac{\sqrt{k}}{\log(k/(n\alpha))} \log\left(\frac{\hat{z}_{\alpha,n}(1)}{z_{\alpha,n}}\right) \xrightarrow[(n\to\infty)]{\mathcal{D}} \mathcal{N}(0,1), \tag{21}$$

$$\frac{\log^2\left(\frac{\hat{z}_{\alpha,n}(1)}{z_{\alpha,n}}\right)}{\int_{t_0}^1 t^2 \log^2\left(\frac{\hat{z}_{\alpha,n}(t)}{\hat{z}_{\alpha,n}(1)}\right) \mathrm{d}t} \xrightarrow[(n\to\infty)]{\mathcal{D}} \frac{W^2(1)}{\int_{t_0}^1 [W(t) - tW(1)]^2 \mathrm{d}t} =: V_{t_0}, \tag{22}$$

where $W(t)$, $0 \le t \le 1$ is the standard Brownian motion, Φ is the cumulative distribution function of a standard normal distribution, $\hat{\sigma}_{\hat{\gamma}_H,\gamma_H} = \hat{\gamma}_H$ and $\hat{\sigma}_{\hat{\gamma}_{MR},\gamma_{MR}} = \sqrt{2}\hat{\gamma}_{MR}$.

Equation (21) made use of normal approximation, which leads to the following asymptotic $1-\tau$-confidence interval for $x_{\alpha,n}$ and $S_{\alpha,n}$ based on normal approximation:

$$I_{\text{na}}^{1-\tau} := \left[\hat{z}_{\alpha,n}(1)\exp\left\{-\Phi(1-\frac{\tau}{2})\hat{\sigma}_{\hat{\gamma},\gamma}\frac{\log(k/(n\alpha))}{\sqrt{k}}\right\}, \hat{z}_{\alpha,n}(1)\exp\left\{\Phi(1-\frac{\tau}{2})\hat{\sigma}_{\hat{\gamma},\gamma}\frac{\log(k/(n\alpha))}{\sqrt{k}}\right\}\right],$$

whereas (22) used the principle of self-normalization, based on which Hoga (2019a) derived another confidence interval based on self-normalization:

$$I_{\text{sn}}^{1-\tau} := \left[\hat{z}_{\alpha,n}(1)\exp\left\{ -\sqrt{V_{t_0,1-\tau}\int_{t_0}^{1} t^2\log^2\left(\frac{\hat{z}_{\alpha,n}(t)}{\hat{z}_{\alpha,n}(1)}\right)\mathrm{d}t} \right\}, \right.$$
$$\left. \hat{z}_{\alpha,n}(1)\exp\left\{ \sqrt{V_{t_0,1-\tau}\int_{t_0}^{1} t^2\log^2\left(\frac{\hat{z}_{\alpha,n}(t)}{\hat{z}_{\alpha,n}(1)}\right)\mathrm{d}t} \right\} \right],$$

where t_0 defines the starting time of the stochastic processes or in other words, the smallest sample, in the approximation of the confidence interval and $V_{t_0,1-\tau}$ is the $(1-\tau)$-quantile of V_{t_0}.

2.4 Parametric method with skew t-distribution

In the previous sections we showed how x_α^U and S_α^U can be estimated using the EVT method. Another approach to estimate x_α^U and S_α^U is the parametric approach, for which we need to assume the exact parametric representation of F. Here we consider the skew t-distribution introduced by Jones and Faddy (2003). The density function of the four-parameter skew t-distribution is as follows,

$$\begin{aligned} f^*(x) &= f(x;a,b,\mu,\sigma) \\ &= C_{a,b}^{-1}\sigma^{-1}\left\{1+\frac{(x-\mu)/\sigma}{(a+b+(x-\mu)^2/\sigma^2)^{1/2}}\right\}^{a+1/2}\left\{1-\frac{(x-\mu)/\sigma}{(a+b+(x-\mu)^2/\sigma^2)^{1/2}}\right\}^{b+1/2}, \end{aligned} \tag{23}$$

where

$$C_{a,b} = 2^{a+b-1}B(a,b)(a+b)^{1/2},$$

$B(a,b)$ is the beta function, μ is the location parameter, σ is the scale parameter, $a>0$ and $b>0$ are used to control skewness and tail-behavior. The corresponding cumulative distribution function is as the following:

$$F^*(x) = I_{\{1+(x-\mu)/\sqrt{(a+b)\sigma^2+(x-\mu)^2}\}/2}(a,b), \tag{24}$$

where $I.(\cdot,\cdot)$ is the incomplete beta function ratio. It is shown by Jones and Faddy (2003) that, given $a>r/2$ and $b>r/2$, the r-th moment of a skew-t random variable T with $\mu=0$, $\sigma=1$ can be written as the following:

$$E(T^r) = \frac{(a+b)^{\frac{r}{2}}}{2^r B(a,b)}\sum_{i=0}^{r}\binom{r}{i}(-1)^i B\left(a+\frac{r}{2}-i, b-\frac{r}{2}+i\right).$$

We then assume F is the standardized skew t-distribution with $\mu=0$, $\sigma=1$, unknown a and b. The unknown parameters can be estimated using MLE. We can then directly obtain an estimator for x_α^U using the left-continuous inverse of F, which is defined as the following :

$$\hat{x}_\alpha^U = F^{\leftarrow}(1-\alpha) \tag{25}$$

and S_α^U can be estimated as

$$\hat{S}_\alpha^U = \int_{\hat{x}_\alpha^U}^{\infty} F(x)\mathrm{d}x. \tag{26}$$

To obtain confidence intervals for the parametric method, we consider the non-parametric bootstrapping method. After noticing that the non-parametric bootstrapping cannot be applied directly to the returns due to their dependencies, we bootstrap the fitted residuals instead. The procedure of constructing the confidence intervals here can then be defined in three steps: 1. We fit the data using a chosen ARMA-GARCH model assuming skew t-distributed innovations and obtain $\hat{x}_\alpha^U$ and $\hat{S}_\alpha^U$. 2. The $\hat{U}_i$'s obtained using the fitted model in the first step are then used to bootstrap the confidence intervals for $\hat{x}_\alpha^U$ and $\hat{S}_\alpha^U$. 3. We calculate the confidence interval of conditional VaR and conditional ES estimates using (4) and (5). Even though this procedure is not ideal as it ignores the uncertainty from model estimation, we prefer it for its simplicity.

2.5 Estimation of the GARCH model

It is important for the finite-sample accuracy of both the EVT and parametric method that we have a consistent and efficient estimating method for the GARCH model. Furthermore, as showed in Hoga (2019b), it is essential to have $\sqrt{n}$-consistent estimator for GARCH to construct the confidence intervals as presented in Section 2.3.

In general, MLE does not lead to consistent estimates if the assumed distribution is misspecified. An example of consistent estimator for ARMA-GARCH is the GQMLE (see Francq and Zakoïan (2004)), which can be defined as the following for the model defined in (1)-(3):

$$\hat{\zeta}_n = \underset{\zeta \in Z}{\arg\min} \left\{ n^{-1} \sum_{i=1}^{n} \ell_i(\zeta) \right\}, \tag{27}$$

where $\zeta = (\phi_1, ..., \phi_p, \nu_1, ..., \nu_q, \rho')'$, Z is set of all possible values for ζ and $\ell_i = \hat{\varepsilon}_i^2/\hat{\sigma}_i^2 + \log\hat{\sigma}_i^2$. The GQMLE is $\sqrt{n}$-consistent if the distribution of the innovations has finite fourth moment.

In this paper, We also consider a more efficient consistent estimator for the ARMA-GARCH model. Specifically, we consider the NGQMLE proposed by Fan et al. (2014). To illustrate the main idea of this estimator, we again consider the model defined in (1)-(3) where we reparameterize (2) and (3) into the following:

$$\varepsilon_i = \kappa\sigma_i U_i, \quad U_i \sim F.i.i.d.(0,1), \tag{28}$$

$$\sigma_i^2 = g(\varepsilon_{i-1}, ..., \varepsilon_{i-h}, \sigma_{i-1}, ..., \sigma_{i-k}, \rho^*), \tag{29}$$

where κ is the scale parameter and ρ^* is the restricted parameter vector such that the intercept is equal to one. As a result, we have the following expanded parameter vector $\zeta = (\phi_1, ..., \phi_p, \nu_1, ..., \nu_q, \kappa, \rho^{*\prime})'$. A note here is that estimators defined in (6) and (7) are not influenced by this reparameterization. Furthermore, it is straightforward how the model can be reparameterized back to the model defined in section 2.1, which allows us to use the previous results.

We use $d(\cdot)$ to denote the density function of a chosen non-Gaussian distribution. Following Fan et al. (2014), we define a parametric family of quasi-likelihood $\{\eta_d : \frac{1}{\eta_d} d(\frac{\cdot}{\eta_d})\}$, where $\eta > 0$. The parameter η is the scaling parameter of the quasi-likelihood, which is used to minimize the discrepancy between the quasi-likelihood family and the true error density f measured by the Kullback-Leibler Information distance (KLID) (Fan et al., 2014). After estimating η_d, the idea is to estimate ζ using the quasi-likelihood function modified by η_d.

We then present the three-step approach (see Fan et al. (2014)) adapted to our generalized ARMA-GARCH model.

The first step is to conduct GQMLE for the reparameterized model:

$$\hat{\zeta}_n = \arg\max_{\zeta \in Z} \frac{1}{n} \left\{ \sum_{i=1}^{n} \left(-\log(\kappa \hat{\sigma}_i) - \frac{\hat{\varepsilon}_i^2}{2\hat{\sigma}_i^2 \kappa^2} \right) \right\}, \tag{30}$$

where $\hat{\sigma}_i$ and $\hat{\varepsilon}_i$ can again be estimated using (6) and (7). After obtaining $\hat{\zeta}_n$, We then estimate η_d by maximizing the following equation:

$$\hat{\eta}_d = \arg\max_{\eta > 0} \frac{1}{n} \sum_{i=1}^{n} \left\{ -\log(\eta) + \log d \left(\frac{\hat{\varepsilon}_i(\hat{\zeta}_n)}{\eta \hat{\kappa} \hat{\sigma}_i(\hat{\zeta}_n)} \right) \right\}, \tag{31}$$

where $\hat{\varepsilon}_i(\hat{\zeta}_n)$ and $\hat{\sigma}_i(\hat{\zeta}_n)$ are the estimator of ε_i and σ_i defined in (6) and (7), where coefficients are replaced by their estimates from $\hat{\zeta}_n$. Finally, we obtain $\tilde{\zeta}_n$, the NGQMLE for ζ, through the following:

$$\tilde{\zeta}_n = \arg\max_{\zeta \in Z} \frac{1}{n} \sum_{i=1}^{n} \left\{ -\log(\hat{\eta}_d \kappa \hat{\sigma}_i) + \log d \left(\frac{\hat{\varepsilon}_i}{\hat{\eta}_d \hat{\sigma}_i \kappa} \right) \right\}. \tag{32}$$

In the reminding of this paper, we refer the NGQMLE constructed using (30)-(32) as 3SNGQMLE.

However, it should be noted that, in general 3SNGQMLE is not always more efficient comparing to GQMLE. Fan et al. (2014) concluded three situations when comparing the relative efficiency of 3NGQMLE and GQMLE: (1) If the true distribution of the innovations is heavier-tailed than the chosen non-Gaussian distribution, 3NGQMLE is more efficient than GQMLE. (2) If true distribution of the innovations has lighter tail comparing to Gaussian distribution, GQMLE is more efficient. (3) If the true distribution of the innovations lies in between the chosen non-Gaussian distribution and Gaussian distribution regarding heaviness of the tail, the relative efficiency of 3SNGQMLE and GQMLE depends on to which distribution the true distribution is closer. As mentioned in section 1, numerous research have shown that financial returns have heavy-tailed innovations. This means that under the context of this paper, it should always be possible to pick a non-Gaussian distribution such that the 3SNGQMLE is more efficient than GQMLE.

To deal with situations where the innovation have infinite fourth moment, Hoga (2019a) consider the NGQMLE with Laplace distribution. However, as shown by Berkes and Horváth (2004), the following moment condition is needed for the Laplace NGQMLE to be consistent:

$$E(|U_i|) = 1.$$

It is straightforward that σ_i^2 in (2) will no longer be the conditional variance with the above condition imposed. Consequently, reparameterization is required for (2) and (3). However, to the best of our knowledge, Hoga (2019a) did not take the above moment condition into account. Hence, the Laplace NGQMLE considered in Hoga (2019a) is not consistent and leads to inconsistent GARCH coefficient estimates.

Even though the GARCH coefficients estimated using simple NGQMLE without additional moment conditions are in general inconsistent, we found that the conditional VaR and conditional ES estimates based on the EVT method are in fact not affect by the this specific kind of inconsistent estimates. We will prove this finding in the reminding of the section.

For simplicity of the notations, here we consider the GARCH(1,1) model as a special case of the model defined in (1)-(3):

$$X_i = \sigma_i U_i, \quad U_i \sim F.i.i.d.(0,1), \tag{33}$$

$$\sigma_i^2 = \omega + \psi_1 X_{i-1}^2 + \beta_1 \sigma_{i-1}^2, \tag{34}$$

with the following reparameterized form:

$$X_i = \kappa v_i U_i, \tag{35}$$

$$v_i^2 = 1 + a_1 X_{i-1}^2 + b_1 v_{i-1}^2, \tag{36}$$

It is straightforward that $\omega = \kappa^2$, $\psi_1 = \kappa^2 a_1$, $\beta_1 = \kappa^2 b_1$.

Fan et al. (2014) showed if the simple NGQMLE with $d(\cdot)$ is used to estimate the model defined in (35) and (36), for the scale parameter the maximum of the likelihood would be achieved at $\hat{\eta}_d \hat{\kappa}$ instead of $\hat{\kappa}$, where $\hat{\kappa}$ and $\hat{\eta}_d$ are the consistent estimators from (30)-(32). Consequently, when we reparameterize the biased estimators for coefficients in (35) and (36) back to coefficients in (33) and (34), we would have the following estimators:

$$\tilde{\omega} = \hat{\eta}_d^2 \hat{\kappa}^2 = \hat{\eta}_d^2 \hat{\omega}, \quad \tilde{\psi}_1 = \hat{\eta}_d^2 \hat{\kappa}^2 \hat{a}_1 = \hat{\eta}_d \hat{\psi}_1, \quad \tilde{\beta}_1 = \hat{\kappa}^2 \hat{b}_1 = \hat{\beta}_1,$$

where $\hat{\omega}$, $\hat{\psi}_1$, $\hat{\beta}_1$, $\hat{a}_1$ and $\hat{b}_1$ are the consistent estimators of parameters in (33)-(36). It is clear that with the simple NGQMLE, we have biased estimates of coefficients of the GARCH model. In the reminding the section, we use the same notations to differentiate the consistent estimators from the (potentially) inconsistent estimators.

We then look into how $x_{\alpha,n}$ and $S_{\alpha,n}$ estimates reacts when the inconsistent NGQMLE estimators were used in the first step of the EVT method. Using the estimators in (7) together with the above inconsistent NGQMLE estimators we have the following inconsistent estimator for σ_i^2, $i = 2, ..., n$:

$$\tilde{\sigma}_i^2 = \tilde{\omega} + \tilde{\psi}_1 X_{i-1}^2 + \tilde{\beta}_1 \tilde{\sigma}_{i-1}^2,$$

where we let $\tilde{\sigma_1}^2 = \tilde{\omega}$ and recursively we have

$$\tilde{\sigma_i}^2 = \hat{\eta}_d^2 \hat{\sigma}_i^2. \tag{37}$$

This directly leads to the biased residual estimates

$$\tilde{U}_i = X_i / \tilde{\sigma}_i = X_i / (\hat{\eta}_d \hat{\sigma}_i) = \hat{U}_i / \hat{\eta}_d. \tag{38}$$

We observe that all the biased residuals are original residuals scaled by $1/\hat{\eta}_d$. Using the invariant property of the Hill and the ES estimator when then have:

$$\tilde{\gamma}_H = \hat{\gamma}_H, \quad \tilde{\gamma}_{MR} = \hat{\gamma}_{MR}. \tag{39}$$

Next we use the positive homogeneity property of coherent risk measures. Notice that the biased residual estimates in (38) can also be seen as unbiased residual estimates for returns obtained by scaling original returns with $1/\hat{\eta}_d$. Then by positive homogeneity we have:

$$\tilde{x}_\alpha^U = \frac{1}{\hat{\eta}_d} \hat{x}_\alpha^U, \quad \tilde{S}_\alpha^U = \frac{1}{\hat{\eta}_d} \hat{S}_\alpha^U. \tag{40}$$

Similar results can be obtained in the same way for estimators defined in (15)-(18). A detailed alternative proof of (39) and (40) are included in appendix A.1. With the above results, we are now ready to derive the $x_{\alpha,n}$ and $S_{\alpha,n}$ estimates based on the inconsistent NGQMLE estimates of GARCH coefficients. Using (19), (20), (37) and (40), we have:

$$\tilde{x}_{\alpha,n}(t) = \tilde{\sigma}_{n+1} \tilde{x}_\alpha^U(t) = \hat{\eta}_d \hat{\sigma}_{n+1} \cdot \frac{1}{\hat{\eta}_d} \hat{x}_\alpha^U(t) = \hat{\sigma}_{n+1} \hat{x}_\alpha^U(t) = \hat{x}_{\alpha,n}(t),$$
$$\tilde{S}_{\alpha,n}(t) = \tilde{\sigma}_{n+1} S_\alpha^U(t) = \hat{\eta}_d \hat{\sigma}_{n+1} \cdot \frac{1}{\hat{\eta}_d} \hat{S}_\alpha^U(t) = \hat{\sigma}_{n+1} \hat{S}_\alpha^U(t) = \hat{S}_{\alpha,n}(t).$$

We proved that using the inconsistent simple NGQMLE instead of the consistent 3SNGMQLES does not change the $x_{\alpha,n}$ and $S_{\alpha,n}$ estimates based on the EVT method. The biasness in coefficients and the resulting residual estimates are canceled out while applying the EVT method. This result of robustness does not only apply to the case where the inconsistent NGQMLE is used. In general, using any incorrect estimation method that only leads to inconsistent κ estimates would not prevent the EVT method from providing correct $x_{\alpha,n}$ and $S_{\alpha,n}$ estimates. Another example here would be imposing incorrect assumptions on the variance of U_i.

Table 1
τ-quantile $V_{t_0,\tau}$ of V_{t_0}

τ	t_0	0.50	0.60	0.70	0.80	0.90	0.95	0.975	0.99	0.995
$V_{t_0,\tau}$	0.3	4.70	7.57	12.46	21.30	41.06	69.40	103.86	150.43	199.03
	0.2	3.93	6.46	10.48	17.79	34.74	54.59	79.38	124.31	157.93
	0.1	3.52	5.82	9.30	15.99	30.34	50.15	76.47	109.86	136.72

3 Monte Carlo Simulation

3.1 Important previous results

In this section we verify the important simulation results from Hoga (2019a). In order to avoid different results from Hoga (2019a) caused by trivial reasons, we adapted part of the code from Hoga (2019b) after verifying its correctness.

The first essential simulation result from Hoga (2019a) is the table of quantile of V_{t_0}, which is defined in (22). As discussed in the previous section, these quantiles are needed when constructing confidence intervals for $x_{\alpha,n}$ and $S_{\alpha,n}$ using the self-normalization method. The quantiles are obtained through fitting the empirical distribution function of V_{t_0}, where we numerically approximate the Brownian motion and the integration. As in Hoga (2019a), we obtained τ-quantials of V_{t0} for $t_0 = [1, 2, 3]$ and a sequence of value for τ between 0.5 and 1. The quantiles are shown in Table 1.

In their simulation study, Hoga (2019a) investigated the relative performances between the Hill estimator and the MR estimator in estimating conditional VaR and conditional ES. They found that in general there is no substantial evidence indicating better performance of the MR estimator. Moreover, they compared the finite sample coverage probabilities of $I_{\text{na}}^{0.95}$ and $I_{\text{sn}}^{0.95}$. They found that $I_{\text{sn}}^{0.95}$ has higher coverage probabilities in most cases at the cost of having a wider confidence interval. To show the above results, Hoga (2019a) considered three different data generating processes (DGPs). Even though the three DGPs are constructed with different underlying models and distributions for the error terms, the conclusions drawn from them by Hoga (2019a) were similar. In this section, we then only consider the first DGP used by Hoga (2019a), which is defined as the following:

$$\begin{aligned} X_i &= \varepsilon_i, \quad \text{where} \quad \varepsilon_i = \sigma_i U_i, \quad U_i \sim st_3(5) \\ \sigma_i^2 &= 0.95 \cdot \frac{20^2}{252} + 0.1 \cdot \varepsilon_{i-1}^2 + 0.85 \cdot \sigma_{i-1}^2, \end{aligned} \tag{41}$$

where $st_a(b)$ denotes the standardized skewed Student's t-distribution introduced by Azzalini and Capitanio (2003). Note that the skewed Student's t-distribution here is defined differently from the skew t-distribution that we mentioned in the Section 2.4. Here, a denotes the degree of freedom, b is the skewness parameter and the standardization is done to ensure the condition in (2) is met.

The rest of the settings are set to be the same as in Hoga (2019a) throughout the reminding of section 3.1. In particular, the ARMA-GARCH coefficients from (33) are estimated by GQMLE. We estimate $x_{\alpha,n}$ and $S_{\alpha,n}$ for $n = 1000$ and $\alpha = 0.025, 0.01, 0.005$. The estimation is done using the EVT method with estimators defined in (19) and (20), where m_n is set to 10 to cancel out initialization effects. In constructing the above estimators, we consider $\hat{\gamma}_H(t)$ and $\hat{\gamma}_{MR}$ as defined in (15) and (16). A note here is that in rare cases $\hat{\gamma}$ is larger than 1, which does not satisfy our previous assumption when estimating $S_{\alpha,n}$. Following the suggestion by Hoga (2019a), we truncate all $\hat{\gamma} > 0.9$ to 0.9 in conditional ES estimations. For the choice k, we consider the estimator in (14) with $k_{\max} = 200$ and $k_{\min} = 50$. The confidence intervals of $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ are constructed using both $I_{\text{na}}^{0.95}$ and $I_{\text{sn}}^{0.95}$, where we take $t_0 = 0.2$. Lastly, all the simulations are run for 10,000 iterations.

Table 2
Average value of k^*, bias, RMSE, coverage probabilities and length of confidence intervals constructed using normal approximation and self-normalization for $x_{\alpha,n}$ and $S_{\alpha,n}$ under three different α with Hill and MR estimators. The considered DGP is as in (41) with $n = 1000$ and 10000 simulation repetitions.

							Coverage		Interval length	
Model	Estimator	k^*	z	α	Bias	RMSE	$I_{na}^{0.95}$	$I_{sn}^{0.95}$	$I_{na}^{0.95}$	$I_{sn}^{0.95}$
(27)	Hill	68	$x_{\alpha,n}$	2.5%	-0.04	0.52	0.49	0.74	0.26	0.54
				1%	-0.11	0.62	0.70	0.80	0.59	0.89
				0.5%	-0.16	0.73	0.76	0.84	0.89	1.28
			$S_{\alpha,n}$	2.5%	-0.09	0.64	0.39	0.82	0.30	1.01
				1%	-0.12	0.77	0.60	0.87	0.68	1.60
				0.5%	-0.10	0.89	0.68	0.89	1.03	2.17
	MR	98	$x_{\alpha,n}$	2.5%	0.06	0.51	0.61	0.75	0.41	0.68
				1%	0.00	0.60	0.81	0.81	0.82	1.07
				0.5%	-0.06	0.70	0.86	0.84	1.20	1.47
			$S_{\alpha,n}$	2.5%	0.03	0.62	0.54	0.82	0.48	1.16
				1%	0.00	0.75	0.71	0.85	0.96	1.71
				0.5%	0.02	0.89	0.76	0.84	1.40	2.26

Table 2 compares performances of the Hill and the MR estimator in estimating $x_{\alpha,n}$ and $S_{\alpha,n}$. We compare the two methods by comparing bias, root mean-squared-error (RMSE) and the coverage probabilities of the confidence intervals. We also compare the length of the confidence interval from the two confidence interval construction methods.

Table 2 show that the MR estimator has lower biases for small α comparing to the Hill estimator. However, this is not the case when we look at the RMSE, from which we cannot draw the conclusion on one estimator outperforming the other. Moreover, the RMSE is higher for conditional ES than for conditional VaR. In Table 2, we can also see that $I_{\text{sn}}^{0.95}$ has higher coverage probability than $I_{\text{na}}^{0.95}$ in almost all situations accompanied by wider confidence intervals. However, confidence intervals constructed using self-normalization still have lower coverage than the nominal level (0.95 in our simulation), which could be caused by the uncertainty from ARMA-GARCH coefficient estimates in finite samples (Hoga, 2019a).

As mentioned in the previous section, $I_{\text{na}}^{0.95}$ and $I_{\text{sn}}^{0.95}$ are constructed using the asymptotic approximations in (21) and (22) respectively. To evaluate the finite-sample validity of these two approximations, Hoga (2019a) compared the distribution of the random variables on the left and on the right side of these equations by probability-probability(PP) plots. Here, we do the same for the DGP from (33) and the results PP plots are presented in Figure 1. We can see from Figure 1 that the approximations based on self-normalization provide better fits comparing to the normal approximations for both the Hill estimator and the MR estimator.

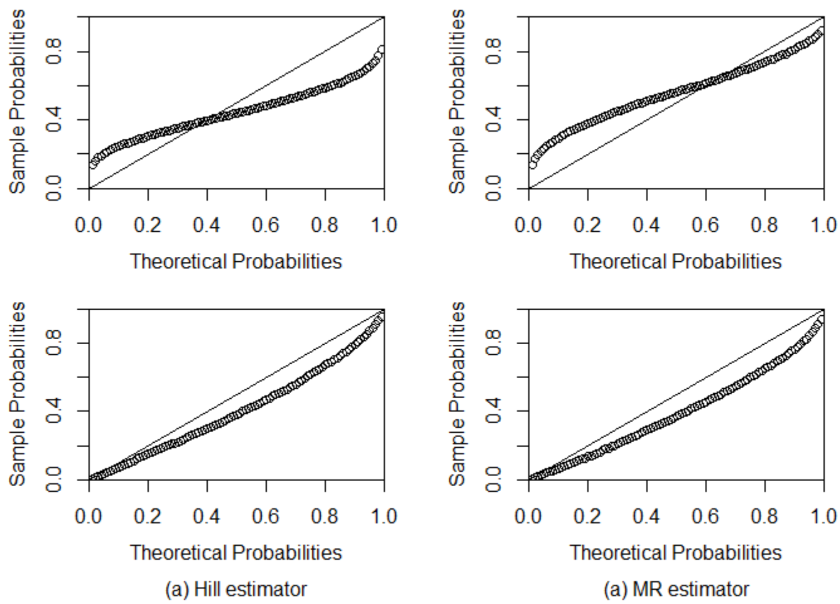


Figure 1: PP plots for the left-hand and right-hand side random variables in (21)(top) and (22)(bottom) for model (33) for left-tail $S_{0.01,1000}$.

In order to evaluate the effect of different choices of k on the results from Table 2, we alter k for a sequence of values between $k_{\min}$ and $k_{\max}$ and check how the evaluation measures from Table 2 change accordingly. Figure 2 presents values of these evaluation measures under different choice of k for left-tail $\hat{S}_{0.01,1000}$ under model (41). The performances of both EVT methods using the Hill estimator and the MR estimator heavily depend on the choice of k. As also documented by Hoga (2019a), the data dependent k^* seems to performs reasonably well since using its average value (in Table 2) attains high coverage probabilities, low biases and low RMSE. However, there is not clear evidence from Figure 2 indicating better performance from the data dependent k^* comparing to the fixed choice $\hat{k} = \lfloor 1.5(\log n)^2 \rfloor = 71$ suggested by Chan et al. (2007). We will further look into the finite sample performance of different choices of k in later sections.

Overall, the conclusions we drew from this section are similar to those in Hoga (2019a). We find that in general $I_{\text{sn}}^{0.95}$ has higher coverage probabilities than $I_{\text{na}}^{0.95}$. This is in line with our finding of the approximation in (22) with self-normalization being more accurate than the approximation in (21) with normal approximation. We also found that the data dependent k^* has desirable performance. However, it is unclear whether the data dependent k^* should be preferred over the fixed choice of k proposed by Chan et al. (2007). We also did not find evidence indicating better performance of the MR estimator comparing to the Hill estimator or vice versa.

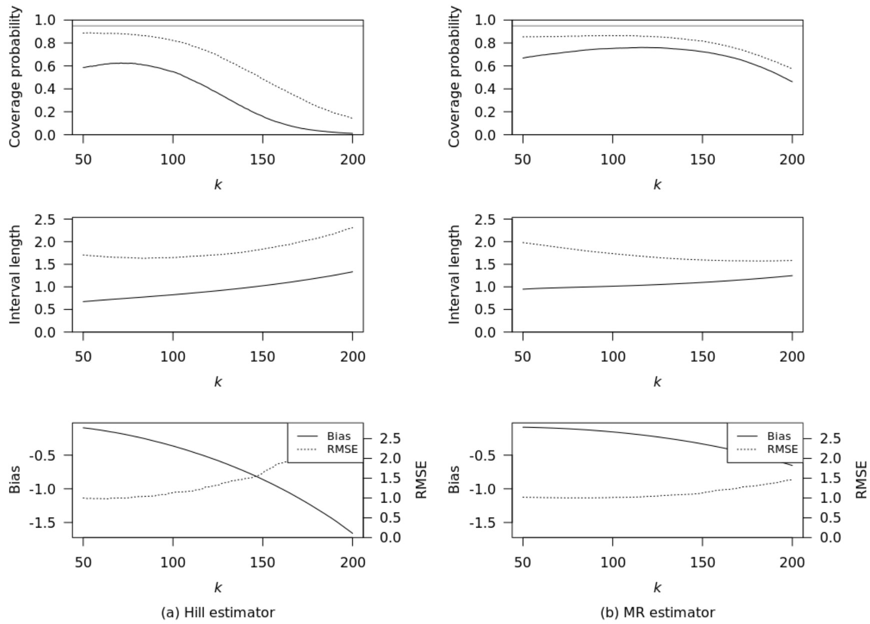


Figure 2: Biases and RMSEs for left-tail $S_{0.01,1000}$ as a function of $k \in [50, 200]$ using the Hill estimator (a) and the MR estimator (b) for model (33), together with the finite sample coverage probabilities and interval lengths of the corresponding $I_{\text{na}}^{0.95}$ (solid) and $I_{\text{sn}}^{0.95}$ (dotted).

3.2 The skew-t Parametric Method

In this section we evaluate the performance of the skew-t parametric method as proposed in section 2.4. Specifically, We compare the performance of the skew-t parametric method in conditional VaR and conditional ES estimation to the EVT method, where the same set-up for the EVT method as in section 3.1 is used. We consider following DGPs with skewed and non-skewed innovation distributions respectively:

$$X_i = \varepsilon_i, \quad \text{where} \quad \varepsilon_i = \sigma_i U_i, \quad U_i \sim st_{4.5}(0), \quad \sigma_i^2 = 1 + 0.2 \cdot \varepsilon_{i-1}^2 + 0.2 \cdot \sigma_{i-1}^2, \tag{42}$$

$$X_i = \varepsilon_i, \quad \text{where} \quad \varepsilon_i = \sigma_i U_i, \quad U_i \sim st_{4.5}(5), \quad \sigma_i^2 = 1 + 0.2 \cdot \varepsilon_{i-1}^2 + 0.2 \cdot \sigma_{i-1}^2, \tag{43}$$

where $st_a(b)$ again denotes the standardized skewed Student's t-distribution from Azzalini and Capitanio (2003). Note that in the DGPs above, the distribution of the innovations has finite fourth moment such that GQMLE is $\sqrt{n}$-consistent (Francq and Zakoïan, 2004). As discussed in section 2.4, the 95% confidence interval of $x_{\alpha,n}$ and $S_{\alpha,n}$ estimated using the parametric method is constructed using the non-parametric bootstrapping method. Due to the fact that bootstrapping methods are very time-consuming, the simulations in this section are all ran for only 1000 iterations. We consider estimation of $x_{\alpha,n}$ and $S_{\alpha,n}$ for $n = 1000$ and $\alpha = 0.01$. The results of the simulations are presented in Table 3-4.

Table 3 presents the results of the model in (34), which is a GARCH(1,1) model with heavy-tailed and standard Student's t-distribution for innovations. We found that in this case the parametric

Table 3

Average value of k, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method and parametric method. The considered EVT methods are based on Hill and MR estimators with model estimated using GQMLE. The considered DGP is as in (34) with $n = 1000$ and 1000 simulation repetitions.

Model	Method	Estimator	k	z	Bias	RMSE	Coverage $I_{na}^{0.95}$	Coverage $I_{sn}^{0.95}$	Interval length $I_{na}^{0.95}$	Interval length $I_{sn}^{0.95}$
(34)	EVT	Hill	$k^* = 58.49$	$x_{\alpha,n}$	0.04	0.48	0.88	0.91	0.97	1.34
				$S_{\alpha,n}$	-0.39	1.00	0.67	0.96	1.45	3.64
		MR	$k^* = 64.60$	$x_{\alpha,n}$	0.28	0.51	0.82	0.76	1.17	1.28
				$S_{\alpha,n}$	-0.03	0.83	0.84	0.93	1.75	3.11
				z	Bias	RMSE	Coverage		Interval length	
	Parametric			$x_{\alpha,n}$	-0.04	0.26	0.75		0.47	
				$S_{\alpha,n}$	0	0.41	0.81		0.94	

Table 4

Average value of k, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method and parametric method. The considered EVT methods are based on Hill and MR estimators with model estimated using GQMLE. The considered DGP is as in (35) with $n = 1000$ and 1000 simulation repetitions.

Model	Method	Estimator	k	z	Bias	RMSE	Coverage $I_{na}^{0.95}$	Coverage $I_{sn}^{0.95}$	Interval length $I_{na}^{0.95}$	Interval length $I_{sn}^{0.95}$
(35)	EVT	Hill	$k^* = 59.57$	$x_{\alpha,n}$	-0.03	0.23	0.72	0.80	0.21	0.29
				$S_{\alpha,n}$	-0.09	0.28	0.57	0.84	0.24	0.54
		MR	$k^* = 87.86$	$x_{\alpha,n}$	0.02	0.22	0.80	0.75	0.28	0.31
				$S_{\alpha,n}$	-0.03	0.26	0.76	0.85	0.33	0.51
				z	Bias	RMSE	Coverage		Interval length	
	Parametric			$x_{\alpha,n}$	-0.13	0.52	0.84		1.59	
				$S_{\alpha,n}$	-0.07	0.62	0.96		2.04	

method has much lower RMSEs for $x_{\alpha,n}$ and $S_{\alpha,n}$ estimation comparing to the EVT method with the Hill or MR estimator, which indicates better performance of the parametric method. The coverage probabilities of the confidence intervals based on the parametric method are too low. Better coverage is achieved by I_{sn} for either the Hill or the MR estimator and I_{na} with the MR estimator. We do, however, observe that the confidence intervals of the EVT method are wider than those of the parametric method. Wider confidence intervals lead to higher coverage probabilities that are closer to the correct level than the shorter ones.

We observe that, however, the performance of the parametric method is not ideal when the innovations are skewed. The results for the GARCH(1,1) model with heavy-tailed and skewed Student's t-distribution for innovations are presented in Table 4. Table 4 shows that, for both $x_{\alpha,n}$ and $S_{\alpha,n}$ estimation, the RMSEs of the parametric method is larger than the RMSEs of the EVT method regardless the choice of the Hill or the MR estimator. Different from what we observed for model (34), here the better coverage is achieved by the parametric method.

To conclude this section, we found that the parametric method has better performance in terms of RMSE when innovations follow the Student's t-distribution, whereas the RMSE indicates better performance of the EVT method when the innovations follow the skewed Student's t-distribution from Azzalini and Capitanio (2003). A possible explanation here is that the skew t-distribution proposed by Jones and Faddy (2003), which is used by our parametric method, does not fit the skewed Student's t-distribution (Azzalini and Capitanio, 2003) well. This does not only lead to undesirable quantile estimation, which is the key of conditional VaR and ES estimation. It also poses concerns on the consistency of the GARCH coefficient estimates. Here we estimate the model using NGQMLE with the skew t-distribution. And as discussed previously, the NGQMLE is in general inconsistent when the underlying distribution is different from the true distribution.

The conclusion regarding coverage probabilities of the confidence intervals is a bit unclear. This is mainly caused by the fact the confidence intervals from both methods suffer from some drawbacks, which are potentially the cause of less ideal coverage. For the parametric method, we do not take the uncertainty of GARCH coefficients estimates into account. This is justified by the fact that the coefficient estimated using GQMLE converges way faster than x_α^U and S_α^U from (4) and (5) asymptotically (Hoga, 2019a). However, in finite sample, the uncertainty of the GARCH coefficient estimates still exist and is not captured by I_{na} or I_{sa}. We will also see in the later sections that better coverage can be achieved when a more efficient estimation method than GQMLE is considered. Since as demonstrated in section 2.4, the confidence interval of the parametric method are estimated using non-parametric bootstrapping of the residuals, where the uncertainty from GARCH coefficient estimates is again ignored.

3.3 Comparison of GQMLE and (3S)NGQMLE

In this section we evaluate the effects of different estimation methods of the ARMA-GARCH model on the two-step EVT method. Specifically, we consider the 3SNGQMLE and GQMLE defined in section 2.5. This section has the following goals: 1. When the innovations have heavy-tailed distribution, we explore whether the more efficient 3SNGQMLE can increase the performance of the EVT method in $x_{\alpha,n}$ and $S_{\alpha,n}$ estimation. 2. We study the relative efficiency of GQMLE and 3SNGQMLE in the EVT method when considering different true distributions for the innovations with varying degrees of heaviness in their tails.

In the first step of the EVT method, we estimate the ARMA-(GJR)-GARCH coefficients with GQMLE and 3SNGQMLE respectively. For 3SNGQMLE, we consider the Student's t-distribution with degree of freedom equal to 5 and 10 as the chosen non-Gaussian distributions. We keep the rest of the settings of the EVT method the same as in section 3.1. We again consider estimation of $x_{\alpha,n}$ and $S_{\alpha,n}$ for $n = 1000$ and $\alpha = 0.01$.

We generate data from GARCH, ARMA-GARCH and ARMA-GJR-GARCH models, where innovations follow the Student's t-distribution and its skewed variant (Azzalini and Capitanio, 2003) with degree of freedom equal to 4.5, 8, 15 and 30. In this section we present results of the GARCH model with the Student's t-distribution with degree of freedom equal to 4.5, 8 and 30. The corresponding DGPs are (42) and the follows:

$$X_i = \varepsilon_i, \quad \text{where} \quad \varepsilon_i = \sigma_i U_i, \quad U_i \sim st_8(0), \quad \sigma_i^2 = 1 + 0.2 \cdot \varepsilon_{i-1}^2 + 0.2 \cdot \sigma_{i-1}^2, \tag{44}$$

$$X_i = \varepsilon_i, \quad \text{where} \quad \varepsilon_i = \sigma_i U_i, \quad U_i \sim st_{30}(0), \quad \sigma_i^2 = 1 + 0.2 \cdot \varepsilon_{i-1}^2 + 0.2 \cdot \sigma_{i-1}^2, \tag{45}$$

Similar conclusions can be drawn from the results of the rest of the models, which are included in Appendix B.1.

Table 5

Average value of k, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method, with sum RMSEs of the coefficient estimates. The considered EVT methods are based on the Hill and MR estimators with two methods in model estimation. The considered DGP is as in (42) with $n = 1000$ and 1000 simulation repetitions.

Model	Coefficient Estimator	RMSE(coefficients)	Estimator	k	z	Bias	RMSE	Coverage $I_{na}^{0.95}$	Coverage $I_{sn}^{0.95}$
(42)	GQMLE	0.578	Hill	$k^* = 58.42$	$x_{\alpha,n}$	0.060	0.334	0.868	0.891
					$S_{\alpha,n}$	-0.354	0.788	0.670	0.932
			MR	$k^* = 63.39$	$x_{\alpha,n}$	0.295	0.431	0.796	0.761
					$S_{\alpha,n}$	-0.004	0.648	0.834	0.922
	3SNGQMLE(t_5)	0.463	Hill	$k^* = 58.98$	$x_{\alpha,n}$	0.059	0.311	0.889	**0.909**
					$S_{\alpha,n}$	-0.364	0.776	0.672	0.948
			MR	$k^* = 64.43$	$x_{\alpha,n}$	0.287	0.414	0.818	0.774
					$S_{\alpha,n}$	-0.023	**0.639**	0.843	0.930
	3SNGQMLE(t_{10})	0.469	Hill	$k^* = 58.54$	$x_{\alpha,n}$	0.062	**0.310**	0.885	0.904
					$S_{\alpha,n}$	-0.356	0.773	0.680	**0.951**
			MR	$k^* = 64.09$	$x_{\alpha,n}$	0.290	0.416	0.810	0.770
					$S_{\alpha,n}$	-0.015	**0.639**	0.840	0.930

Table 5-8 present results obtained using three different coefficient estimators, where both the Hill and MR estimators are considered in the second step of the EVT method. We again evaluate the quality of $x_{\alpha,n}$ and $S_{\alpha,n}$ estimates in terms of bias, RMSE and coverage probabilities. We also report the sum RMSEs of the coefficient estimates from the first step.

Table 5 contains the results for the model with very heavy-tailed innovations. In this case, the distribution of the innovations has heavier tail than both non-Gaussian distributions that we selected for the 3SNGQMLE. We observe that in this case the sum RMSEs of the coefficient estimates of GQMLE is much larger comparing to the two 3SNGQMLEs that are considered. The efficiency gains from the first step coefficient estimates using 3SNGQMLE are also associated with improved $x_{\alpha,n}$ and $S_{\alpha,n}$ estimates. Table 5 shows that the lowest RMSE and highest coverage (in bold) are achieved when the 3SNGQMLE is used. Furthermore, the sum coefficient RMSE of the 3SNGQMLE increases when the Student's t-distribution with a larger degree of freedom is considered. This is to be expected since the distribution in the DGP is heavier than both of the distributions we picked. By considering the Student's t-distribution with higher degree of freedom we are deviating further from the true distribution in terms of heaviness in the tail.

Table 6
Average value of k, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method, with sum RMSEs of the coefficient estimates. The considered EVT methods are based on the Hill and MR estimators with two methods in model estimation. The considered DGP is as in (44) with $n = 1000$ and 1000 simulation repetitions.

Model	Coefficient Estimator	RMSE(coefficients)	Estimator	k	z	Bias	RMSE	Coverage $I_{na}^{0.95}$	Coverage $I_{sn}^{0.95}$
(44)	GQMLE	0.464	Hill	$k^* = 55.40$	$x_{\alpha,n}$	0.048	0.249	0.888	0.922
					$S_{\alpha,n}$	-0.342	0.576	0.658	0.931
			MR	$k^* = 63.62$	$x_{\alpha,n}$	0.287	0.369	0.759	0.693
					$S_{\alpha,n}$	-0.014	0.415	0.881	**0.949**
	3SNGQMLE(t_5)	0.441	Hill	$k^* = 55.63$	$x_{\alpha,n}$	0.048	0.245	0.892	0.928
					$S_{\alpha,n}$	-0.345	0.571	0.654	0.931
			MR	$k^* = 64.03$	$x_{\alpha,n}$	0.284	0.365	0.764	0.717
					$S_{\alpha,n}$	-0.021	0.408	0.888	0.946
	3SNGQMLE(t_{10})	0.435	Hill	$k^* = 55.58$	$x_{\alpha,n}$	0.049	**0.242**	0.895	**0.929**
					$S_{\alpha,n}$	-0.342	0.567	0.659	0.941
			MR	$k^* = 64.10$	$x_{\alpha,n}$	0.286	0.364	0.761	0.707
					$S_{\alpha,n}$	-0.016	**0.405**	0.885	**0.949**

We then consider model (44) where the true distribution has lighter tail comparing to (42). Notice that the true distribution in this situation lies in between the two distributions for 3SNGQMLE in terms of heaviness of the tail. We observe from Table 6 that the 3SNGQMLE(t_{10}) results in better coefficient estimates than the 3SNGQMLE(t_5), while both 3SNGQMLEs outperform the GQMLE. Table 10 further shows that the resulting $x_{\alpha,n}$ and $S_{\alpha,n}$ estimates using 3SNGQMLEs are better the ones obatined using GQMLE. However, the differences in coefficient and risk measure RMSEs between GQMLE and 3SNGQMLE are not as noticeable as the previous situation, where the true distribution has heavier tail.

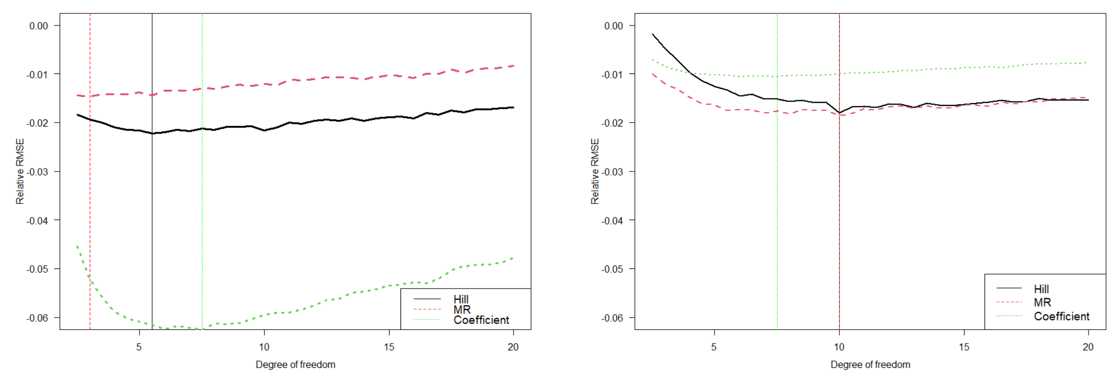
Table 7
Average value of k, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method, with sum RMSEs of the coefficient estimates. The considered EVT methods are based on the Hill and MR estimators with two methods in model estimation. The considered DGP is as in (45) with $n = 1000$ and 1000 simulation repetitions.

Model	Coefficient Estimator	RMSE(coefficients)	Estimator	k	z	Bias	RMSE	Coverage $I_{na}^{0.95}$	Coverage $I_{sn}^{0.95}$
(45)	GQMLE	0.433	Hill	$k^* = 54.79$	$x_{\alpha,n}$	0.043	**0.180**	0.910	0.909
					$S_{\alpha,n}$	-0.317	0.442	0.597	0.906
			MR	$k^* = 67.52$	$x_{\alpha,n}$	0.267	0.311	0.740	0.605
					$S_{\alpha,n}$	-0.028	**0.267**	0.920	0.935
	3SNGQMLE(t_5)	0.443	Hill	$k^* = 54.77$	$x_{\alpha,n}$	0.041	0.184	0.906	0.909
					$S_{\alpha,n}$	-0.324	0.452	0.604	0.896
			MR	$k^* = 67.25$	$x_{\alpha,n}$	0.264	0.311	0.737	0.600
					$S_{\alpha,n}$	-0.035	0.273	0.919	0.933
	3SNGQMLE(t_{10})	0.431	Hill	$k^* = 54.79$	$x_{\alpha,n}$	0.043	0.181	0.910	**0.912**
					$S_{\alpha,n}$	-0.320	0.444	0.612	0.904
			MR	$k^* = 67.37$	$x_{\alpha,n}$	0.266	0.311	0.731	0.599
					$S_{\alpha,n}$	-0.031	0.269	0.921	**0.936**

Lastly we consider the model in (45) where the true innovation distribution is closer to a standard

Gaussian distribution comparing to the previous two models. Table 7 shows that in this situation the GQMLE outperforms 3SNGQMLE(t_5) in terms of both the coefficient and risk measure estimations. The 3SNGQMLE with lighter distribution (t_{10}) results in slightly better coefficient estimates than the GQMLE. However, using the GQMLE in model estimation still leads to better $x_{\alpha,n}$ and $S_{\alpha,n}$ estimates.

Figure 3 shows how RMSE of $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ varies when the Student's t-distribution with different degree of freedom is considered for 3SNGQMLE. The plots provides some intuition on how the chosen non-Gaussian distribution is related to the RMSE of $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$. We observe that, as expected, the closer the chosen distribution to the true distribution the lower the sum coefficient RMSEs. It is further shown by Figure 3 that there is in general a positive relation between coefficient RMSE and RMSE of $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$, except the area close to the optimum. For both $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$, the lowest RMSE is obtained using the distribution different from the one that leads to the lowest sum coefficients RMSEs. Moreover, the lowest RMSEs for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ are achieved with different distributions. We also find that all the relative RMSEs in figure 3 are negative. This indicates that 3SNGQMLE using the Student's t-distribution with any degree of freedom from $[2.5, 20]$ outperforms GQMLE in terms of coefficient, $x_{\alpha,n}$ and $S_{\alpha,n}$ estimates.



$\hat{x}_{\alpha,n}$ $\hat{S}_{\alpha,n}$

Figure 3: Relative sum RMSEs of the GARCH cofficient estimates and RMSE of $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with the model estimated using 3SNGQMLE with the Student's t-distribution with a sequence of degree of freedom from 2.5 to 20, with minimal (vertical lines). The vertical axis indicate the difference from the RMSE using GQMLE. The true model is GARCH(1,1) with $st_6(0)$ and the simulation is done for 10,000 iterations

To sum up, we show that the 3SNGQMLE in general leads to better coefficient, $x_{\alpha,n}$ and $S_{\alpha,n}$ estimates when the innovations are heavy-tailed. The advantage of using 3SNGQMLE with a fixed heavy-tailed distribution in the EVT method is well-established with heavily tailed innovations. However, with the true distribution getting closer to Gaussian, the performance of 3SNGQMLE with a fixed heavy-tailed distribution decreases and even becomes worse than GQMLE. Furthermore, we showed the performance of 3SNGQMLE can be further improved if we are able to select the appropriate non-Gaussian distribution. Finding the best distribution can be tricky, as we have shown that the distribution that leads to the best coefficient estimates does not necessarily lead to the best $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$. In conclusion, given that the returns are heavy-tailed, our results provide evidences indicating that the 3SNGQMLE should be preferred over GQMLE.

Table 8

Average value of k, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method, with sum RMSEs of the coefficient estimates. The considered EVT methods are based on Hill and MR estimators with two methods in model estimation. The considered DGP is as in (46) with $n = 1000$ and 1000 simulation repetitions.

								Coverage	
Model	Coefficient Estimator	RMSE(coefficients)	Estimator	k	z	Bias	RMSE	$I_{na}^{0.95}$	$I_{sn}^{0.95}$
(46)	3SNGQMLE($t_{2.5}$)	0.448	Hill	$k^* = 54.79$	$x_{\alpha,n}$	0.046	0.214	0.901	0.922
					$S_{\alpha,n}$	-0.340	0.519	0.621	0.937
			MR	$k^* = 67.52$	$x_{\alpha,n}$	0.275	0.340	0.771	0.675
					$S_{\alpha,n}$	-0.032	0.350	0.913	0.943
	NGQMLE($t_{2.5}$)	2.367	Hill	$k^* = 54.77$	$x_{\alpha,n}$	0.045	0.211	0.902	0.922
					$S_{\alpha,n}$	-0.341	0.518	0.621	0.937
			MR	$k^* = 67.25$	$x_{\alpha,n}$	0.274	0.337	0.772	0.677
					$S_{\alpha,n}$	-0.033	0.347	0.914	0.943

We prove in section 2.5 that although the simple NGQMLE is inconsistent, the inconsistent estimates do not influence the $x_{\alpha,n}$ and $S_{\alpha,n}$ estimates using the EVT approach. To shed some light on this result, we consider the following DGP:

$$X_i = \varepsilon_i, \quad \text{where} \quad \varepsilon_i = \sigma_i U_i, \quad U_i \sim st_{11}(0), \quad \sigma_i^2 = 1 + 0.2 \cdot \varepsilon_{i-1}^2 + 0.2 \cdot \sigma_{i-1}^2. \tag{46}$$

Now we again apply the EVT approach for estimation of $x_{\alpha,n}$ and $S_{\alpha,n}$, where in the first step we estimate the coefficients using the 3SNGQMLE and the simple NGQMLE respectively. For both non-Gaussian estimation method we consider the Student's t-distribution with degree of freedom equal to 2.5. Table 8 shows that the sum RMSEs of the coefficient estimates using the simple NGQMLE is more than 5 times larger than the sum RMSEs using the 3SNGQMLE, whereas the biases, RMSEs and coverage of the resulting $x_{\alpha,n}$ and $S_{\alpha,n}$ estimates are near identical between two model estimation methods. We notice that the simple NGQMLE results in near negligibly better $x_{\alpha,n}$ and $S_{\alpha,n}$ estimates. One possible explanation is that the 3NGQMLE is exposed to more numerical uncertainty since all three steps of it requires numerical optimization. Table 8 verifies our proof in section 2.5 by showing the inconsistency of the simple NGQMLE and the resulting near identical $x_{\alpha,n}$ and $S_{\alpha,n}$ estimates comparing to the 3SNGQMLE.

3.4 The choice of k

In this section we study whether the data-dependent k^* from Hoga (2019a) leads to better $x_{\alpha,n}$ and $S_{\alpha,n}$ estimates comparing to the fixed sample-size-determined $\hat{k}$ proposed by Chan et al. (2007).

We consider the same settings for the EVT method as in section 3.3 with two changes: 1. The

Table 9
Average value of k^*, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method with k^* and $\hat{k}$. The considered EVT methods are based on Hill and MR estimators with model estimated using GQMLE and 3SNGQMLE. The considered DGP is as in (42) with $n = 1000$ and 1000 simulation repetitions.

Model	Coefficient Estimator	Estimator	k	z	Bias	RMSE	Coverage $I_{na}^{0.95}$	Coverage $I_{sn}^{0.95}$
(42)	GQMLE	Hill	$k^* = 58.42$	$x_{\alpha,n}$	0.060	0.334	0.868	0.891
				$S_{\alpha,n}$	-0.354	0.788	0.670	0.932
			$\hat{k} = 71$	$x_{\alpha,n}$	-0.039	0.343	0.916	0.927
				$S_{\alpha,n}$	-0.791	1.125	0.509	0.879
		MR	$k^* = 63.39$	$x_{\alpha,n}$	0.295	0.431	0.796	0.761
				$S_{\alpha,n}$	-0.004	0.648	0.834	0.922
			$\hat{k} = 71$	$x_{\alpha,n}$	0.281	0.429	0.837	0.765
				$S_{\alpha,n}$	-0.282	0.732	0.855	0.932
	3SNGQMLE(t_{10})	Hill	$k^* = 58.54$	$x_{\alpha,n}$	0.062	**0.310**	0.885	0.904
				$S_{\alpha,n}$	-0.356	0.773	0.680	**0.951**
			$\hat{k} = 71$	$x_{\alpha,n}$	-0.036	0.319	0.933	**0.937**
				$S_{\alpha,n}$	-0.79	1.108	0.513	0.881
		MR	$k^* = 64.09$	$x_{\alpha,n}$	0.290	0.416	0.810	0.770
				$S_{\alpha,n}$	-0.015	**0.639**	0.840	0.930
			$\hat{k} = 71$	$x_{\alpha,n}$	0.276	0.414	0.842	0.777
				$S_{\alpha,n}$	-0.293	0.724	0.854	0.938

model is estimated with GQMLE and 3SNGQMLE with t_{10}. 2. When choosing k, in addition to the data-dependent k^* defined in (14), we also consider $\hat{k} = \lfloor 1.5(logn)^2 \rfloor = 71$ (Chan et al., 2007). The goal is to estimate $x_{\alpha,n}$ and $S_{\alpha,n}$ with $n = 1000$ and $\alpha = 0.01$.

For the DGP, we again consider GARCH, ARMA-GARCH and ARMA-GJR-GARCH models with heavy-tailed, skewed and non-skewed innovations.

Table 9 presents results for the model in (42). We observe that, for both $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$, the lowest RMSE is achieved using the data-dependent k^*. An interesting finding is that the fixed $\hat{k}$ leads to better coverage for conditional VaR with both the Hill and the MR estimator, whereas the highest coverage for conditional ES is still obtained using the data-dependent k^*.

We draw the same conclusion as above from the rest of the DGPs, which are included in Appendix B.2 . Overall, we find that the data-dependent k^* leads to better estimates of $x_{\alpha,n}$ and $S_{\alpha,n}$ when $n = 1000$. However, in this situation, the fixed $\hat{k}$ has acceptable performance and the confidence intervals constructed using it has better coverage for $x_{\alpha,n}$ comparing to the ones constructed using k^*.

4 Empirical Study

The aim of this section is to compare relative performances of all the methods in conditional VaR and conditional ES estimation in an empirical study, where we also demonstrate the empirical relevance of the 3SNGQMLE. In the empirical study we consider data on historic stock index returns. As in

Hoga (2019a), we consider returns of six global indices including DAX 30, CAC 40, DIJA, NASDAQ, Nikkei 225, and HSI. We retrieve 20-year data from 1/1/1997 to 31/12/2016 on the six series from Hoga (2019b).

We again consider the EVT method as well as the parametric method. For both methods, we fit the data using the GARCH(1,1), AR(1)-GARCH(1,1), ARMA(1,1)-GARCH(1,1) and ARMA(1,1)-GJR-GARCH(1,1) models. As mentioned in section 2.4, we consider the parametric method with the skew t-distribution, where the models are then estimated using maximum-likelihood estimation. For the EVT method, we fit the models using GQMLE and 3SNGQMLE with the Student's t-distribution and degree of freedom equal to 10. We again consider the Hill estimator and the MR estimator with fixed and data-dependent k, where we let $k_{\min} = 50$ and $k_{\max} = 200$. And m_n is set to 10 to remove the initial effect. In total we consider 36 combinations of methods.

For all six financial series, we consider one-step-ahead forecast of $x_{\alpha,n}$ and $S_{\alpha,n}$ with the rolling window scheme, where the window size is set to 1000 and α is set to 0.005.

Same as in Hoga (2019a), we check the unconditional and conditional coverage of $x_{\alpha,n}$ estimates using the tests of Kupiec (1995) (UC) and Berkowitz et al. (2011) (CC) respectively. Then, again as suggested by Hoga (2019a), we calculate the quantile score, $Q(X_i, x_i)$, to evaluate the $x_{\alpha,n}$ forecasts and asymmetric Laplace (AL) log score, $LS(x_i, S_i, X_i)$, to jointly evaluate the performance of $S_{\alpha,n}$ and $x_{\alpha,n}$ forecasts, where X_i is the actual return, x_i and S_i are the conditional VaR and ES estimate corresponding the ith observation,

$$Q(X_i, x_i) = (X_i - x_i)(\alpha - I(X_i \leq x_i)) \text{ and}$$
$$LS(x_i, S_i, X_i) = -\log\left(\frac{\alpha - 1}{S_i}\right) - \frac{Q(X_i, x_i)}{\alpha S_i}.$$

For each method, we add up the scores calculated over the entire forecast period, where a smaller score indicates better performance.

We focus on the results regarding NASDAQ and Nikkei 225 in this section, where the results of the other four indices are included in Appendix B.3.

We use UC and CC to evaluate the quality of out-of-sample conditional VaR forecasts, where a good conditional VaR forecast is expected to have correct conditional and unconditional coverage. Table 10 and 11 show that only ARGARCH-3S-k^*-Hill, ARMAGARCH-3S-k^*-Hill, ARMAGJRGARCH-3S-$\hat{k}$-MR, ARMAGJRGARCH-3S-k^*-Hill and ARMAGJRGARCH-3S-k^*-MR are never rejected by both tests at 10%, out of the 36 methods we are considering. We observe that by choosing fixed $\hat{k}$, GQMLE and the GARCH model do not lead to desirable conditional VaR estimates when evaluating with UC and CC. We also find that, when considering the GARCH(1,1) and AR(1)-GARCH(1,1) model, the MR estimator has poor performance. This is in line with the finding of Hoga (2019a). However, our results suggest that the MR estimator has improved performance when the asymmetric GARCH model is considered. The estimates from the parametric method with skew t-distribution have correct UC and CC for Nikkei 225 but not NASDAQ.

Table 10 and 11 present the quantile score and AL log score for the 36 methods considered in

Table 10
P-values of UC and CC test together with quantile score and AL log score for one-step-ahead $x_{\alpha,n}$ and $S_{\alpha,n}$ forecasts of NASDAQ with data from 1/1/1997 to 31/12/2016, $\alpha = 0.005$ and $n = 1000$.

Index	Model	Method			UC	CC	Quantile score	AL log score
NASDAQ	GARCH(1,1)	GQMLE	$\hat{k}$	Hill	0.023**	0.074*	0.963	-8406.65
				MR	0.028**	0.000***	0.952	-8501.9
			k^*	Hill	0.629	0.119	**0.945**	$\mathbf{-8558.712}$
				MR	0.016**	0.000***	0.953	-8505.81
		3SNGQMLE	$\hat{k}$	Hill	0.068*	0.085*	0.979	-8329.79
				MR	0.127	0.000***	0.964	-8386.23
			k^*	Hill	0.48	0.114	0.962	-8433.37
				MR	0.08*	0.000***	0.964	-8377.75
		Parametric			0.007***	0.063*	0.976	-8452.859
	ARMA(1,0)-GARCH(1,1)	GQMLE	$\hat{k}$	Hill	0.023**	0.074*	0.960	-8431.8
				MR	0.028**	0.000***	0.952	-8514.48
			k^*	Hill	0.629	0.119	0.944	-8571.39
				MR	0.016**	0.000***	0.953	-8517.26
		3SNGQMLE	$\hat{k}$	Hill	0.068*	0.085*	0.945	-8536.3
				MR	0.127	0.000***	0.935	-8626.54
			k^*	Hill	0.48	0.114	**0.930**	$\mathbf{-8672.869}$
				MR	0.08*	0.000***	0.935	-8632.46
		Parametric			0.007***	0.063*	0.979	-8431.853
	ARMA(1,1)-GARCH(1,1)	GQMLE	$\hat{k}$	Hill	0.013**	1	0.955	-8447.22
				MR	0.049**	0.000***	0.940	-8556.6
			k^*	Hill	0.48	0.999	0.938	-8589.27
				MR	0.028**	0.000***	0.94	-8565.86
		3SNGQMLE	$\hat{k}$	Hill	0.068*	0.085*	0.944	-8535.43
				MR	0.08*	0.000***	0.938	-8604.95
			k^*	Hill	0.629	0.119	**0.931**	$\mathbf{-8668.357}$
				MR	0.049**	0.000***	0.939	-8606.54
		Parametric			0.007***	0.063*	0.976	-8452.871
	ARMA(1,1)-GJR-GARCH(1,1)	GQMLE	$\hat{k}$	Hill	0.013**	1.000	0.947	-8445.12
				MR	0.193	0.938	0.904	-8638.7
			k^*	Hill	0.068*	1.000	0.925	-8567.03
				MR	0.392	0.965	0.902	-8650.01
		3SNGQMLE	$\hat{k}$	Hill	0.041**	1.000	0.938	-8492.99
				MR	0.682	0.982	**0.895**	-8687.64
			k^*	Hill	0.169	1.000	0.916	-8630.21
				MR	0.682	0.982	0.895	$\mathbf{-8697.105}$
		Parametric			0.007***	1.000	0.961	-8483.126

Note. * indicates significance at 10%, ** indicates significance at 5%, and *** indicates significance at 1%.

this section. For each model the lowest scores are in bold and the lowest scores from all 36 methods are underlined. Three important conclusions can be drawn from Table 10 with respect to the loss scores. First of all, We find that except for the GARCH model, the 3SNGQMLE leads to better $x_{\alpha,n}$ and $S_{\alpha,n}$ estimates. Secondly, the data-dependent k^* performs relatively better than the fixed $\hat{k}$. Lastly, the more complicated models, especially the ARMA(1,1)-GJR-GARCH(1,1) model, have better performance comparing to the simple GARCH(1,1) model. While the last two conclusion can also be drawn from Table 11 for the index Nikkei 225 over our sample, Table 11 indicates that the GQMLE performs better than the 3SNGQMLE in all four models. One possible explanation is that the true distribution of innovations in this case is closer to the standard Gaussian distribution in terms of KLID than the Student's t-distribution with degree of freedom equal to 10. For the parametric method, we focus on its scores in Table 11 as it does not lead to correct UC and CC for

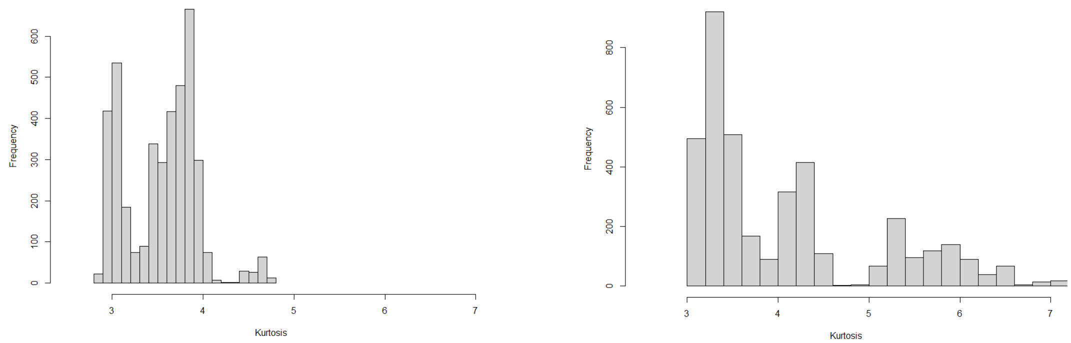
NASDAQ. We observe from Table 11 that it outperforms the EVT method when the GARCH(1,1) model is considered. Moreover, the scores of the parametric-method's estimates are not much higher than the scores from the EVT method for the other three models, where EVT performs better. As it is not feasible to check the significance of difference among scores from 32 method, we perform the Diebold-Mariano (DM) test (Diebold and Mariano, 2002) to test whether the scores from the methods with the lowest scores are significantly lower than the scores from the baseline GARCH-GQMLE-$\hat{K}$-Hill. We find that for both indices, the improvements in scores are significant at 5%. The p-values of the DM tests can be found in Appendix B.3.

Table 11
P-values of UC and CC test together with quantile score and AL log score for one-step-ahead $x_{\alpha,n}$ and $S_{\alpha,n}$ forecasts of Nikkei 225 with data from 1/1/1997 to 31/12/2016, $\alpha = 0.005$ and $n = 1000$.

Index	Model	Method			UC	CC	Quantile score	AL log score
N225	GARCH(1,1)	GQMLE	$\hat{k}$	Hill	0.567	0.003***	**1.114**	-7443.009
				MR	0.090*	0.000***	1.141	-7448.163
			k^*	Hill	0.913	0.000***	1.123	-7473.916
				MR	0.054*	0.000***	1.143	-7425.835
		3SNGQMLE	$\hat{k}$	Hill	0.306	0.093*	1.122	-7411.449
				MR	0.089*	0.000***	1.120	-7463.556
			k^*	Hill	0.426	0.099*	1.117	-7460.401
				MR	0.054*	0.000***	1.127	-7426.960
		Parametric			0.567	0.104	1.113	**−7487.832**
	ARMA(1,0)-GARCH(1,1)	GQMLE	$\hat{k}$	Hill	0.424	0.099*	1.112	-7448.402
				MR	0.141	0.009***	1.116	-7482.260
			k^*	Hill	0.906	0.004***	**1.11**	**−7493.982**
				MR	0.09*	0.010	1.120	-7461.808
		3SNGQMLE	$\hat{k}$	Hill	0.426	0.099*	1.125	-7400.698
				MR	0.054*	0.000***	1.127	-7426.303
			k^*	Hill	0.569	0.104	1.121	-7438.054
				MR	0.017**	0.000***	1.133	-7388.283
		Parametric			0.567	0.104	1.112	-7489.000
	ARMA(1,1)-GARCH(1,1)	GQMLE	$\hat{k}$	Hill	0.424	0.099*	1.112	-7449.081
				MR	0.141	0.009***	1.116	-7482.989
			k^*	Hill	0.906	0.004***	**1.11**	**−7494.619**
				MR	0.09*	0.010	1.120	-7462.281
		3SNGQMLE	$\hat{k}$	Hill	0.426	0.099*	1.124	-7403.153
				MR	0.054*	0.000***	1.126	-7427.061
			k^*	Hill	0.569	0.104	1.120	-7439.910
				MR	0.031**	0.000***	1.132	-7389.377
		Parametric			0.567	0.104	1.112	-7490.047
	ARMA(1,1)-GJR-GARCH(1,1)	GQMLE	$\hat{k}$	Hill	0.305	0.093*	1.063	-7571.569
				MR	0.031**	0.141	1.067	-7627.009
			k^*	Hill	0.424	0.099*	$\underline{\mathbf{1.054}}$	$\underline{\mathbf{-7642.774}}$
				MR	0.031**	0.141	1.070	-7623.713
		3SNGQMLE	$\hat{k}$	Hill	0.306	0.093*	1.067	-7572.645
				MR	0.141	0.137	1.073	-7611.244
			k^*	Hill	0.732	0.109	1.063	-7631.449
				MR	0.213	0.135	1.073	-7612.728
		Parametric			0.424	0.099*	1.064	-7638.153

Note. * indicates significance at 10%, ** indicates significance at 5%, and *** indicates significance at 1%.

The conclusions that we drew above are surprising since Nikkei 225 is known to be volatile with extreme outcomes, for which we then expect the 3SNGQMLE to outperform the GQMLE.



NASDAQ N225

Figure 4: Histogram of sample kurtosis of returns of NASDAQ and Nikkei 225 from 1/1/1997 to 31/12/2016. The returns are fitted using ARMA(1,1)-GJR-GARCH(1,1) and 3SNGQMLE(t_{10}) with the rolling window scheme with size 1000.

Table 12
P-values of UC and CC test together with quantile score and AL log score for one-step-ahead $x_{\alpha,n}$ and $S_{\alpha,n}$ forecasts of Nikkei 225 with filtered data from 1/1/1997 to 31/12/2016, $\alpha = 0.005$ and $n = 1000$.

Index	Model	Method			UC	CC	Quantile score	AL log score
N225	GARCH(1,1)	GQMLE	$\hat{k}$	Hill	0.939	0.999	0.579	-3892.066
				MR	0.034**	0.924	0.614	-3822.094
			k^*	Hill	0.583	0.993	0.594	-3873.816
				MR	0.006***	0.861	0.617	-3795.336
		3SNGQMLE	$\hat{k}$	Hill	0.697	1.000	**0.575**	**−3901.730**
				MR	0.006***	0.000***	0.588	-3855.049
			k^*	Hill	0.939	0.999	0.579	-3889.238
				MR	0.002***	0.000***	0.593	-3820.628

To analyse the differences between two indices, we take a closer look at the residuals from the estimated models. Figure 4 presents the sample kurtosis from our estimation samples for NASDAQ and N225. We observe that the majority of the samples from Nikkei 225 have kurtosis close to 3, whereas the kurtosis for most samples from NASDAQ are centered around 4. We notice that there are some samples with very high kurtosis for both indices, especially Nikkei 225. To study whether the samples with low kurtosis are causing 3SNGQMLE to have poor performance, we remove the observations whose corresponding estimation period has sample kurtosis below 3.7 from the forecast samples of Nikkei 225 and re-evaluate the performance of the EVT methods. In total, we remove 1495 observations from the forecast sample with 3904 observations. Table 12 shows that, as expected, 3SNGQMLE outperforms GQMLE using the filtered sample when the GARCH(1,1) model is considered. The same conclusion is drawn when the ARMA(1,0)-GARCH(1,1), ARMA(1,1)-GARCH(1,1) and ARMA(1,1)-GJR-GARCH(1,1) models are considered, for which the results are included in Appendix B.3.

We notice that most of the reminding samples from Nikkei 225 after filtering overlap with one of the crisis periods in Japan. The beginning of our data are from the periods during which the

Japanese economy was still suffering from the 90's asset price bubble. And most observation from the second half of our data set are from the period affected by the 2008 crisis. While both NASDAQ and Nikkei 225 exhibited high volatility during the crisis periods, our results indicate that innovations of Nikkei 225 have thinner tail than NASDAQ during the more stable periods. One possible explanation for this finding is that NASDAQ mainly consists of stocks from the information technology sector, where there is presence of speculative bubbles that collapse periodically (Anderson et al., 2010). The emergence and collapsing of bubbles could potentially be the cause of heavy-tailed innovations of NASDAQ during non-major-crisis period comparing to Nikkei 225 that consists of stocks from companies in diverse sectors.

A possible way to improve the performance of 3SNGQMLE when estimating samples from Nikkei 225 is to pick a non-Gaussian distribution with lighter tail. However, further research need to be done to determine the best distribution to use in this situation.

In conclusion, we show in this section the advantage of considering the data-dependent k^* and the asymmetric ARMA-GJR-GARCH model when applying the EVT method in conditional VaR and conditional ES estimation. Furthermore, we demonstrate the advantages of considering the 3SNGQMLE instead of the GQMLE in model estimation. We show that using the 3SNGQMLE in the EVT methods leads to better conditional VaR and ES estimates when the innovations are heavy-tailed. From economic aspects, heavy-tailed innovations for stock returns can be associated with characteristics of the sectors and special periods such as the financial crisis periods. We then suggest one to the choose estimation method using the available prior knowledge of the stocks. Evaluating the sample kurtosis of the estimation period could also provide some hints on which estimation method should be used.

5 Discussion and Conclusion

In this article, we extend the EVT method in conditional VaR and ES estimations from Hoga (2019a) by considering a better model estimation method and more general asymmetric ARMA-GARCH models. We also verify the results from Hoga (2019a) and provide a clearer comparison among all models and estimators. Furthermore, we consider an improved parametric method with the skew t-distribution.

First of all, we verify the results from Hoga (2019a) where similar conclusions were drawn. We further explicitly show the advantage of considering the data-dependent k^* and the ARMA-GARCH model in more extensive simulation and empirical studies comparing to Hoga (2019a). In addition, we consider the ARMA-GJR-GARCH model. We verify our simulation results and show that the conclusions overall do not change when the ARMA-GJR-GARCH model is considered. We further show the advantages of considering the ARMA-GJR-GARCH model in an empirical study. Another study that is not included in this article is the robustness check of different models. It is meaningful if one can show which model is the most robust when the data is generated using different models.

Seondly, We show that the 3SNGQMLE can improve the finite-sample performance of the EVT method comparing to the commonly used GQMLE when the innovations are heavy-tailed. We then relax the assumptions from Hoga (2019a) and prove that the inconsistent simple NGQMLE leads to the same conditional VaR and ES estimates as the consistent 3SNGQMLE. Despite the fact that in general the simple NGQMLE is inconsistent and hence is not considered in research or applications, our results show that it is valid when estimating conditional VaR and ES using the EVT method. And the simple NGQMLE should be preferred over GQMLE when we have heavy-tailed innovations. Even though the simple NGQMLE leads to same conditional VaR and ES estimates as the 3SNGQMLE, we nevertheless prefer the latter for two reasons. Firstly, the 3SNGQMLE leads to consistent coefficient estimates while the simple NGQMLE does not. As demonstrated by our results, the bias of the estimates from the simple NGQMLE is high when the picked distribution is far from the true distribution. The biased coefficient estimates loses their interpretability, which is often important in applications. Second of all, the 3SNGQMLE has more room for further improvements. We show that the performance of NGQMLEs can be further improved if the appropriate non-Gaussian distribution is picked. Fan et al. (2014) proposed a way of choosing optimal non-Gaussian distribution when applying the 3SNGQMLE. This method can then be used to improve the efficiency of the 3SNGQMLE in coefficient estimates. However, as shown by our results, the best coefficient estimates do not necessarily lead to the best conditional VaR and ES estimates. It can be interesting for future research to look into the exact relation between efficiency in coefficient estimates and efficiency in conditional VaR and ES estimates in finite sample.

We verify our simulation results on actual data in an empirical study. We show the situation where the 3SNGQMLE outperforms GQMLE and also the case where GQMLE has more desirable performance. We link our findings to the economics environment and characteristics of different stock indices. More importantly, we provide intuitions in the selection between 3SNGQMLE and GQMLE using economic reasoning and statistical features of the sample. However, the intuitions we provide are immature and do not lead to a well-established selecting scheme for model estimation methods, which can be interesting for future research.

Last but not the least, We show in a simulation study that the parametric method provides more accurate conditional VaR and ES estimates when the innovations follow the Student's t-distribution. Whereas the EVT method is more desirable when the innovations follow the skewed Student's t-distribution. Though the performance of the considered parametric method is not ideal and cannot replace the EVT methods, we do see potential in the parametric method as there is plenty room to further improve its performance. Our results also show that the confidence intervals constructed for the parametric method are not ideal. To address this, future research can consider other bootstrapping techniques (e.g. block bootstrapping) instead. Moreover, the bootstrapping in our study was only done with 1000 iterations and better coverage could be achieved if more iterations were used. Overall, our results regarding the parametric method provide some useful insights for future research.

References

Anderson, K., Brooks, C., and Katsaris, A. (2010). Speculative bubbles in the sp 500: Was the tech bubble confined to the tech sector? *Journal of Empirical Finance*, 17(3):345–361.

Azzalini, A. and Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(2):367–389.

Basel Committee (2019). Minimum capital requirements for market risk. `https://www.bis.org/bcbs/publ/d457.pdf`.

Berkes, I. and Horváth, L. (2004). The efficiency of the estimators of the parameters in GARCH processes. *Annals of Statistics*, 32(2):633 – 655.

Berkowitz, J., Christoffersen, P., and Pelletier, D. (2011). Evaluating value-at-risk models with desk-level data. *Management Science*, 57(12):2213–2227.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327.

Chan, N. H., Deng, S.-J., Peng, L., and Xia, Z. (2007). Interval estimation of Value-at-Risk based on GARCH models with heavy-tailed innovations. *Journal of Econometrics*, 137(2):556–576.

Danielsson, J., Jansen, D. W., and de Vries, C. G. (1996). The method of moments ratio estimator for the tail shape parameter. *Communications in Statistics - Theory and Methods*, 25(4):711–720.

Daníelsson, J. and de Vries, C. G. (1997). Tail index and quantile estimation with very high frequency data. *Journal of Empirical Finance*, 4(2):241–257.

Diebold, F. (2012). *Empirical Modeling of Exchange Rate Dynamics*. Lecture Notes in Economics and Mathematical Systems. Springer Berlin Heidelberg.

Diebold, F. X. and Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 20(1):134–144.

Einmahl, J. H., de Haan, L., and Zhou, C. (2016). Statistics of heteroscedastic extremes. *Journal of the Royal Statistical Society: Series B: Statistical Methodology*, pages 31–51.

Embrechts, P., Resnick, S. I., and Samorodnitsky, G. (1999). Extreme value theory as a risk management tool. *North American Actuarial Journal*, 3(2):30–41.

Fan, J., Qi, L., and Xiu, D. (2014). Quasi-maximum likelihood estimation of garch models with heavy-tailed likelihoods. *Journal of Business & Economic Statistics*, 32(2):178–191.

Francq, C. and Zakoian, J.-M. (2019). *GARCH models: structure, statistical inference and financial applications*. John Wiley & Sons.

Francq, C. and Zakoïan, J.-M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli*, 10(4):605 – 637.

Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, 48(5):1779–1801.

Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *Annals of Statistics*, 3(5):1163–1174.

Hoga, Y. (2019a). Confidence intervals for conditional tail risk measures in ARMA–GARCH models. *Journal of Business & Economic Statistics*, 37(4):613–624.

Hoga, Y. (2019b). Supplemental materials for "confidence intervals for conditional tail risk measures in ARMA–GARCH models".

Jones, M. C. and Faddy, M. J. (2003). A skew extension of the t-distribution, with applications. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(1):159–174.

Kupiec, P. (1995). Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives*, 3(2).

McNeil, A. J. and Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of Empirical Finance*, 7(3):271–300. Special issue on Risk Management.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59(2):347–370.

Patton, A. J., Ziegel, J. F., and Chen, R. (2019). Dynamic semiparametric models for expected shortfall (and Value-at-Risk). *Journal of Econometrics*, 211(2):388–413.

Taylor, J. W. (2007). Using Exponentially Weighted Quantile Regression to Estimate Value at Risk and Expected Shortfall. *Journal of Financial Econometrics*, 6(3):382–406.

Taylor, J. W. (2008). Estimating Value-at-Risk and Expected Shortfall Using Expectiles. *Journal of Financial Econometrics*, 6(2):231–252.

Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations. *Journal of the American Statistical Association*, 73(364):812–815.

Appendix A

A.1 Proof

Following from (38), it is clear that the order statistics of $\tilde{U}_i$ is the scaled version of the order statistics of $\hat{U}_i$ without the order being changed . Defining the order statistics of $\tilde{U}_i$ as $\tilde{U}_{1:n} \leq \tilde{U}_{2:n} \leq ... \leq \tilde{U}_{n:n}$, it is easy to see from above that $\tilde{U}_{h:n} = U_{h:n}/\hat{\eta}_d$ for $h = 1, ..., n$. We then see that though $\tilde{U}_i$ and its order statistics are biased, the Hill and the MR estimator remaining the same. From (11) and (12) we have:

$$
\begin{aligned}
\tilde{\gamma}_H &:= \frac{1}{k}\sum_{i=0}^{k-1} \log\left(\frac{\tilde{U}_{n-i:n}}{\tilde{U}_{n-k:n}}\right) = \frac{1}{k}\sum_{i=0}^{k-1} \log\left(\frac{U_{n-i:n}/\hat{\eta}_d}{U_{n-i:n}/\hat{\eta}_d}\right) \\
&= \frac{1}{k}\sum_{i=0}^{k-1} \log\left(\frac{U_{n-i:n}}{U_{n-i:n}}\right) = \hat{\gamma}_H, \\
\tilde{\gamma}_{MR} &= \frac{1}{2}\frac{\frac{1}{k}\sum_{i=0}^{k-1}\{\log(\tilde{U}_{n-i:n}) - \log(\tilde{U}_{n-k:n})\}^2}{\tilde{\gamma}_H} \\
&= \frac{1}{2}\frac{\frac{1}{k}\sum_{i=0}^{k-1}\{\log(U_{n-i:n}/\hat{\eta}_d) - \log(U_{n-i:n}/\hat{\eta}_d)\}^2}{\hat{\gamma}_H} \\
&= \hat{\gamma}_{MR},
\end{aligned}
$$

where the data-dependent estimator for k defined in (14) can be written as the following:

$$
\begin{aligned}
\tilde{k}^* &= \underset{k=k_{\min},\ldots,k_{\max}}{\arg\min}\left[\sup_{j=1,\ldots,k_{\max}}\left|\tilde{U}_{n-j,n} - \tilde{U}_{n-k,n}\left(\frac{j}{k}\right)^{-\tilde{\gamma}}\right|\right] \\
&= \underset{k=k_{\min},\ldots,k_{\max}}{\arg\min}\left[\sup_{j=1,\ldots,k_{\max}}\left|\frac{1}{\hat{\eta}_d}\hat{U}_{n-j:n} - \frac{1}{\hat{\eta}_d}\hat{U}_{n-k:n}\left(\frac{j}{k}\right)^{-\hat{\gamma}}\right|\right] \\
&= \underset{k=k_{\min},\ldots,k_{\max}}{\arg\min}\left[\sup_{j=1,\ldots,k_{\max}}\frac{1}{\hat{\eta}_d}\left|\hat{U}_{n-j:n} - \hat{U}_{n-k:n}\left(\frac{j}{k}\right)^{-\hat{\gamma}}\right|\right] \\
&= k^*
\end{aligned}
$$

Using (10) and (13) we can then write the following:

$$
\begin{aligned}
\tilde{x}_{\alpha}^U &= \widetilde{U\left(\frac{n}{k}\right)}\left(\frac{n\alpha}{k}\right)^{-\tilde{\gamma}} \qquad (47)\\
&= \frac{1}{\hat{\eta}_d}\widehat{U\left(\frac{n}{k}\right)}\left(\frac{n\alpha}{k}\right)^{-\hat{\gamma}} = \frac{1}{\hat{\eta}_d}\hat{x}_{\alpha}^U \\
\tilde{S}_{\alpha}^U &= \frac{\tilde{x}_{\alpha}^U}{1-\tilde{\gamma}} = \frac{\hat{x}_{\alpha}^U/\hat{\eta}_d}{1-\hat{\gamma}} = \hat{S}_{\alpha}^U \qquad (48)
\end{aligned}
$$

A.2 Technical Notes

In this section we discuss the interesting technical problems we ran into during this research. We hope sharing them can provide some help for future research to replicate our results and inspirations.

The first note is regarding the τ-quantile in Table 1. The values in Table 1 are lower comparing to the results from Hoga (2019a). One of the causes of the differences is the sample size used in Brownian motion approximation. After increasing the sample size from $100,000$ to $1000,000$, we obtained larger values for $t_0 = 0.2$ and $t_0 = 0.3$. However, the values for $t_0 = 0.1$ decreased after increasing the sample size.

The second note here is about the starting point for optimizations when applying the 3SNGQMLE. Before applying the 3SNGQMLE in the EVT method, we first tried to replicate part of the results from Fan et al. (2014). We found that the results from Fan et al. (2014) could only be exactly replicated if the true DGP values are used as the starting point for likelihood optimizations. And the performance of 3SNGQMLE gets worse when a random starting point is picked instead. A suggestion by us is to pick a sequence of stating points when applying 3SNGQMLE in empirical studies, which is not implemented in this study due to time constraints.

Thirdly, in the empirical study we find that the coefficients are very close to zero and could not be correctly estimated for some indices. To solve this issue, we multiply the returns by 100 before the estimation and scale the conditional VaR and ES forecasts back to the original level for forecast evaluation.

Lastly, in the empirical study, not all the estimation samples lead to feasible estimates. As this problem doesn't occur often, we drop these samples and take it into account when scaling the final scores.

Appendix B

B.1 Additional Simulation Result for 3.3

This section contains additional simulation results for section 3.3. The additional DGPs considered are as the following:

$$X_i = \varepsilon_i, \quad \text{where} \quad \varepsilon_i = \sigma_i U_i, \quad U_i \sim st_{4.5}(5), \quad \sigma_i^2 = 1 + 0.2 \cdot \varepsilon_{i-1}^2 + 0.2 \cdot \sigma_{i-1}^2, \tag{49}$$

$$X_i = \varepsilon_i, \quad \text{where} \quad \varepsilon_i = \sigma_i U_i, \quad U_i \sim st_{8}(5), \quad \sigma_i^2 = 1 + 0.2 \cdot \varepsilon_{i-1}^2 + 0.2 \cdot \sigma_{i-1}^2, \tag{50}$$

$$X_i = \varepsilon_i, \quad \text{where} \quad \varepsilon_i = \sigma_i U_i, \quad U_i \sim st_{15}(0), \quad \sigma_i^2 = 1 + 0.2 \cdot \varepsilon_{i-1}^2 + 0.2 \cdot \sigma_{i-1}^2, \tag{51}$$

$$X_i = \varepsilon_i, \quad \text{where} \quad \varepsilon_i = \sigma_i U_i, \quad U_i \sim st_{15}(5), \quad \sigma_i^2 = 1 + 0.2 \cdot \varepsilon_{i-1}^2 + 0.2 \cdot \sigma_{i-1}^2, \tag{52}$$

$$X_i = \varepsilon_i, \quad \text{where} \quad \varepsilon_i = \sigma_i U_i, \quad U_i \sim st_{30}(5), \quad \sigma_i^2 = 1 + 0.2 \cdot \varepsilon_{i-1}^2 + 0.2 \cdot \sigma_{i-1}^2, \tag{53}$$

$$X_i = 0.2 \cdot X_{i-1} + 0.2 \cdot \varepsilon_{i-1} + \varepsilon_i, \quad \text{where} \quad \varepsilon_i = \sigma_i U_i, \quad U_i \sim st_{4.5}(0)$$
$$\sigma_i^2 = 1 + 0.2 \cdot \varepsilon_{i-1}^2 + 0.2 \cdot \sigma_{i-1}^2, \tag{54}$$

$$X_i = 0.2 \cdot X_{i-1} + 0.2 \cdot \varepsilon_{i-1} + \varepsilon_i, \quad \text{where} \quad \varepsilon_i = \sigma_i U_i, \quad U_i \sim st_{4.5}(5)$$
$$\sigma_i^2 = 1 + 0.2 \cdot \varepsilon_{i-1}^2 + 0.2 \cdot \sigma_{i-1}^2, \tag{55}$$

$$X_i = 0.2 \cdot X_{i-1} + 0.2 \cdot \varepsilon_{i-1} + \varepsilon_i, \quad \text{where} \quad \varepsilon_i = \sigma_i U_i, \quad U_i \sim st_{8}(0)$$
$$\sigma_i^2 = 1 + 0.2 \cdot \varepsilon_{i-1}^2 + 0.2 \cdot \sigma_{i-1}^2, \tag{56}$$

$$X_i = 0.2 \cdot X_{i-1} + 0.2 \cdot \varepsilon_{i-1} + \varepsilon_i, \quad \text{where} \quad \varepsilon_i = \sigma_i U_i, \quad U_i \sim st_{8}(5)$$
$$\sigma_i^2 = 1 + 0.2 \cdot \varepsilon_{i-1}^2 + 0.2 \cdot \sigma_{i-1}^2, \tag{57}$$

$$X_i = 0.2 \cdot X_{i-1} + 0.2 \cdot \varepsilon_{i-1} + \varepsilon_i, \quad \text{where} \quad \varepsilon_i = \sigma_i U_i, \quad U_i \sim st_{15}(0)$$
$$\sigma_i^2 = 1 + 0.2 \cdot \varepsilon_{i-1}^2 + 0.2 \cdot \sigma_{i-1}^2, \tag{58}$$

$$X_i = 0.2 \cdot X_{i-1} + 0.2 \cdot \varepsilon_{i-1} + \varepsilon_i, \quad \text{where} \quad \varepsilon_i = \sigma_i U_i, \quad U_i \sim st_{15}(5)$$
$$\sigma_i^2 = 1 + 0.2 \cdot \varepsilon_{i-1}^2 + 0.2 \cdot \sigma_{i-1}^2, \tag{59}$$

$$X_i = 0.2 \cdot X_{i-1} + 0.2 \cdot \varepsilon_{i-1} + \varepsilon_i, \quad \text{where} \quad \varepsilon_i = \sigma_i U_i, \quad U_i \sim st_{30}(0)$$
$$\sigma_i^2 = 1 + 0.2 \cdot \varepsilon_{i-1}^2 + 0.2 \cdot \sigma_{i-1}^2, \tag{60}$$

$$X_i = 0.2 \cdot X_{i-1} + 0.2 \cdot \varepsilon_{i-1} + \varepsilon_i, \quad \text{where} \quad \varepsilon_i = \sigma_i U_i, \quad U_i \sim st_{30}(5)$$
$$\sigma_i^2 = 1 + 0.2 \cdot \varepsilon_{i-1}^2 + 0.2 \cdot \sigma_{i-1}^2, \tag{61}$$

$$X_i = 0.2 \cdot X_{i-1} + 0.2 \cdot \varepsilon_{i-1} + \varepsilon_i, \quad \text{where} \quad \varepsilon_i = \sigma_i U_i, \quad U_i \sim st_{4.5}(0)$$
$$\sigma_i^2 = 1 + 0.2 \cdot \varepsilon_{i-1}^2 + 0.1 \cdot \varepsilon_{i-1}^2 \cdot I(\varepsilon_{i-1} < 0) + 0.2 \cdot \sigma_{i-1}^2, \tag{62}$$

$$X_i = 0.2 \cdot X_{i-1} + 0.2 \cdot \varepsilon_{i-1} + \varepsilon_i, \quad \text{where} \quad \varepsilon_i = \sigma_i U_i, \quad U_i \sim st_{4.5}(5)$$
$$\sigma_i^2 = 1 + 0.2 \cdot \varepsilon_{i-1}^2 + 0.1 \cdot \varepsilon_{i-1}^2 \cdot I(\varepsilon_{i-1} < 0) + 0.2 \cdot \sigma_{i-1}^2, \tag{63}$$

$$X_i = 0.2 \cdot X_{i-1} + 0.2 \cdot \varepsilon_{i-1} + \varepsilon_i, \quad \text{where} \quad \varepsilon_i = \sigma_i U_i, \quad U_i \sim st_{8}(0)$$
$$\sigma_i^2 = 1 + 0.2 \cdot \varepsilon_{i-1}^2 + 0.1 \cdot \varepsilon_{i-1}^2 \cdot I(\varepsilon_{i-1} < 0) + 0.2 \cdot \sigma_{i-1}^2, \tag{64}$$

$$X_i = 0.2 \cdot X_{i-1} + 0.2 \cdot \varepsilon_{i-1} + \varepsilon_i, \quad \text{where} \quad \varepsilon_i = \sigma_i U_i, \quad U_i \sim st_{8}(5)$$
$$\sigma_i^2 = 1 + 0.2 \cdot \varepsilon_{i-1}^2 + 0.1 \cdot \varepsilon_{i-1}^2 \cdot I(\varepsilon_{i-1} < 0) + 0.2 \cdot \sigma_{i-1}^2, \tag{65}$$

$$X_i = 0.2 \cdot X_{i-1} + 0.2 \cdot \varepsilon_{i-1} + \varepsilon_i, \quad \text{where} \quad \varepsilon_i = \sigma_i U_i, \quad U_i \sim st_{15}(0)$$
$$\sigma_i^2 = 1 + 0.2 \cdot \varepsilon_{i-1}^2 + 0.1 \cdot \varepsilon_{i-1}^2 \cdot I(\varepsilon_{i-1} < 0) + 0.2 \cdot \sigma_{i-1}^2, \tag{66}$$
$$X_i = 0.2 \cdot X_{i-1} + 0.2 \cdot \varepsilon_{i-1} + \varepsilon_i, \quad \text{where} \quad \varepsilon_i = \sigma_i U_i, \quad U_i \sim st_{15}(5)$$
$$\sigma_i^2 = 1 + 0.2 \cdot \varepsilon_{i-1}^2 + 0.1 \cdot \varepsilon_{i-1}^2 \cdot I(\varepsilon_{i-1} < 0) + 0.2 \cdot \sigma_{i-1}^2, \tag{67}$$
$$X_i = 0.2 \cdot X_{i-1} + 0.2 \cdot \varepsilon_{i-1} + \varepsilon_i, \quad \text{where} \quad \varepsilon_i = \sigma_i U_i, \quad U_i \sim st_{30}(0)$$
$$\sigma_i^2 = 1 + 0.2 \cdot \varepsilon_{i-1}^2 + 0.1 \cdot \varepsilon_{i-1}^2 \cdot I(\varepsilon_{i-1} < 0) + 0.2 \cdot \sigma_{i-1}^2, \tag{68}$$
$$X_i = 0.2 \cdot X_{i-1} + 0.2 \cdot \varepsilon_{i-1} + \varepsilon_i, \quad \text{where} \quad \varepsilon_i = \sigma_i U_i, \quad U_i \sim st_{30}(5)$$
$$\sigma_i^2 = 1 + 0.2 \cdot \varepsilon_{i-1}^2 + 0.1 \cdot \varepsilon_{i-1}^2 \cdot I(\varepsilon_{i-1} < 0) + 0.2 \cdot \sigma_{i-1}^2. \tag{69}$$

Table B1.1

Average value of k, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method, with sum RMSEs of the coefficient estimates. The considered EVT methods are based on Hill and MR estimators with two methods in model estimation. The considered DGP is as in (43) with $n = 1000$ and 1000 simulation repetitions.

								Coverage	
Model	Coefficient Estimator	RMSE(coefficients)	Estimator	k	z	Bias	RMSE	$I_{na}^{0.95}$	$I_{sn}^{0.95}$
(43)	GQMLE	0.588	Hill	$k^* = 57.45$	$x_{\alpha,n}$	0.026	0.332	0.850	0.887
					$S_{\alpha,n}$	-0.265	0.658	0.685	**0.948**
			MR	$k^* = 65.40$	$x_{\alpha,n}$	0.222	0.389	0.790	0.744
					$S_{\alpha,n}$	0.014	0.570	0.842	0.917
	3SNGQMLE(t_5)	0.468	Hill	$k^* = 58.36$	$x_{\alpha,n}$	0.032	**0.278**	0.886	**0.911**
					$S_{\alpha,n}$	-0.261	0.598	0.691	0.957
			MR	$k^* = 66.07$	$x_{\alpha,n}$	0.225	0.346	0.805	0.760
					$S_{\alpha,n}$	0.014	0.502	0.851	0.928
	3SNGQMLE(t_{10})	0.482	Hill	$k^* = 57.86$	$x_{\alpha,n}$	0.034	**0.278**	0.885	0.910
					$S_{\alpha,n}$	-0.255	0.594	0.694	0.958
			MR	$k^* = 66.10$	$x_{\alpha,n}$	0.227	0.346	0.809	0.762
					$S_{\alpha,n}$	0.019	**0.501**	0.848	0.928

Table B1.2

Average value of k, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method, with sum RMSEs of the coefficient estimates. The considered EVT methods are based on Hill and MR estimators with two methods in model estimation. The considered DGP is as in (50) with $n = 1000$ and 1000 simulation repetitions.

								Coverage	
Model	Coefficient Estimator	RMSE(coefficients)	Estimator	k	z	Bias	RMSE	$I_{na}^{0.95}$	$I_{sn}^{0.95}$
(50)	GQMLE	0.473	Hill	$k^* = 55.83$	$x_{\alpha,n}$	0.037	0.220	0.880	0.896
					$S_{\alpha,n}$	-0.304	0.504	0.624	0.941
			MR	$k^* = 65.011$	$x_{\alpha,n}$	0.251	0.321	0.785	0.691
					$S_{\alpha,n}$	-0.016	0.355	0.882	0.939
	3SNGQMLE(t_5)	0.438	Hill	$k^* = 56.08$	$x_{\alpha,n}$	0.041	**0.210**	0.884	0.906
					$S_{\alpha,n}$	-0.304	0.496	0.644	0.941
			MR	$k^* = 65.09$	$x_{\alpha,n}$	0.254	0.318	0.793	0.683
					$S_{\alpha,n}$	-0.018	**0.346**	0.890	**0.946**
	3SNGQMLE(t_{10})	0.439	Hill	$k^* = 56.03$	$x_{\alpha,n}$	0.040	**0.210**	0.883	**0.907**
					$S_{\alpha,n}$	-0.304	0.495	0.640	0.940
			MR	$k^* = 65.07$	$x_{\alpha,n}$	0.253	0.317	0.789	0.682
					$S_{\alpha,n}$	-0.017	**0.346**	0.893	0.940

Table B1.3

Average value of k, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method, with sum RMSEs of the coefficient estimates. The considered EVT methods are based on Hill and MR estimators with two methods in model estimation. The considered DGP is as in (51) with $n = 1000$ and 1000 simulation repetitions.

Model	Coefficient Estimator	RMSE(coefficients)	Estimator	k	z	Bias	RMSE	Coverage $I_{na}^{0.95}$	Coverage $I_{sn}^{0.95}$
(51)	GQMLE	0.458	Hill	$k^* = 54.88$	$x_{\alpha,n}$	0.032	0.206	0.890	0.911
					$S_{\alpha,n}$	-0.354	0.513	0.587	0.906
			MR	$k^* = 65.02$	$x_{\alpha,n}$	0.267	0.326	0.754	0.628
					$S_{\alpha,n}$	-0.042	0.325	0.910	0.944
	3SNGQMLE(t_5)	0.458	Hill	$k^* = 55.12$	$x_{\alpha,n}$	0.03	0.206	0.898	0.918
					$S_{\alpha,n}$	-0.361	0.518	0.587	0.908
			MR	$k^* = 64.95$	$x_{\alpha,n}$	0.265	0.324	0.757	0.627
					$S_{\alpha,n}$	-0.05	0.326	0.91	0.942
	3SNGQMLE(t_{10})	0.453	Hill	$k^* = 54.94$	$x_{\alpha,n}$	0.031	**0.205**	0.902	**0.919**
					$S_{\alpha,n}$	-0.358	0.516	0.594	0.912
			MR	$k^* = 65.09$	$x_{\alpha,n}$	0.267	0.325	0.762	0.638
					$S_{\alpha,n}$	-0.046	**0.324**	0.913	**0.95**

Table B1.4

Average value of k, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method, with sum RMSEs of the coefficient estimates. The considered EVT methods are based on Hill and MR estimators with two methods in model estimation. The considered DGP is as in (52) with $n = 1000$ and 1000 simulation repetitions.

Model	Coefficient Estimator	RMSE(coefficients)	Estimator	k	z	Bias	RMSE	Coverage $I_{na}^{0.95}$	Coverage $I_{sn}^{0.95}$
(52)	GQMLE	0.431	Hill	$k^* = 54.58$	$x_{\alpha,n}$	0.039	0.179	0.917	0.925
					$S_{\alpha,n}$	-0.308	0.438	0.616	0.911
			MR	$k^* = 66.22$	$x_{\alpha,n}$	0.258	0.305	0.752	0.620
					$S_{\alpha,n}$	-0.022	**0.271**	0.939	**0.947**
	3SNGQMLE(t_5)	0.434	Hill	$k^* = 54.78$	$x_{\alpha,n}$	0.037	0.178	0.914	**0.929**
					$S_{\alpha,n}$	-0.314	0.441	0.616	0.913
			MR	$k^* = 66.07$	$x_{\alpha,n}$	0.255	0.303	0.763	0.625
					$S_{\alpha,n}$	-0.028	0.273	0.946	**0.947**
	3SNGQMLE(t_{10})	0.429	Hill	$k^* = 54.71$	$x_{\alpha,n}$	0.038	**0.177**	0.916	0.924
					$S_{\alpha,n}$	-0.309	0.437	0.617	0.919
			MR	$k^* = 66.15$	$x_{\alpha,n}$	0.256	0.303	0.758	0.623
					$S_{\alpha,n}$	-0.025	**0.271**	0.941	0.944

Table B1.5
Average value of k, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method, with sum RMSEs of the coefficient estimates. The considered EVT methods are based on Hill and MR estimators with two methods in model estimation. The considered DGP is as in (53) with $n = 1000$ and 1000 simulation repetitions.

Model	Coefficient Estimator	RMSE(coefficients)	Estimator	k	z	Bias	RMSE	Coverage $I_{na}^{0.95}$	Coverage $I_{sn}^{0.95}$
(53)	GQMLE	0.418	Hill	$k^* = 54.34$	$\mathrm{x}_{\alpha,n}$	0.032	0.171	0.911	0.916
					$\mathrm{S}_{\alpha,n}$	-0.322	0.436	0.571	0.896
			MR	$k^* = 64.60$	$\mathrm{x}_{\alpha,n}$	0.254	0.297	0.739	0.610
					$\mathrm{S}_{\alpha,n}$	-0.036	0.250	0.931	0.963
	3SNGQMLE(t_5)	0.435	Hill	$k^* = 54.53$	$\mathrm{x}_{\alpha,n}$	0.030	0.172	0.904	0.915
					$\mathrm{S}_{\alpha,n}$	-0.328	0.442	0.575	0.893
			MR	$k^* = 67.62$	$\mathrm{x}_{\alpha,n}$	0.250	0.295	0.749	0.616
					$\mathrm{S}_{\alpha,n}$	-0.044	0.257	0.932	0.960
	3SNGQMLE(t_{10})	0.423	Hill	$k^* = 54.46$	$\mathrm{x}_{\alpha,n}$	0.031	0.171	0.911	0.920
					$\mathrm{S}_{\alpha,n}$	-0.325	0.438	0.577	0.886
			MR	$k^* = 67.63$	$\mathrm{x}_{\alpha,n}$	0.252	0.296	0.754	0.622
					$\mathrm{S}_{\alpha,n}$	-0.040	0.253	0.930	0.958

Table B1.6
Average value of k, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method, with sum RMSEs of the coefficient estimates. The considered EVT methods are based on Hill and MR estimators with two methods in model estimation. The considered DGP is as in (54) with $n = 1000$ and 1000 simulation repetitions.

Model	Coefficient Estimator	RMSE(coefficients)	Estimator	k	z	Bias	RMSE	Coverage $I_{na}^{0.95}$	Coverage $I_{sn}^{0.95}$
(54)	GQMLE	0.763	Hill	$k^* = 57.93$	$x_{\alpha,n}$	0.046	0.358	0.870	0.911
					$S_{\alpha,n}$	-0.372	0.818	0.679	0.954
			MR	$k^* = 64.509$	$x_{\alpha,n}$	0.282	0.447	0.805	0.764
					$S_{\alpha,n}$	-0.021	0.672	0.840	0.940
	3SNGQMLE(t_5)	0.618	Hill	$k^* = 58.50$	$x_{\alpha,n}$	0.041	**0.317**	0.891	**0.931**
					$S_{\alpha,n}$	-0.397	0.798	0.689	**0.953**
			MR	$k^* = 65.16$	$x_{\alpha,n}$	0.272	0.422	0.823	0.790
					$S_{\alpha,n}$	-0.053	0.659	0.864	0.944
	3SNGQMLE(t_{10})	0.631	Hill	$k^* = 58.31$	$x_{\alpha,n}$	0.045	0.320	0.886	0.927
					$S_{\alpha,n}$	-0.383	0.791	0.688	0.955
			MR	$k^* = 65.19$	$x_{\alpha,n}$	0.276	0.424	0.816	0.781
					$S_{\alpha,n}$	-0.040	**0.654**	0.862	0.942

Table B1.7

Average value of k, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method, with sum RMSEs of the coefficient estimates. The considered EVT methods are based on Hill and MR estimators with two methods in model estimation. The considered DGP is as in (55) with $n = 1000$ and 1000 simulation repetitions.

Model	Coefficient Estimator	RMSE(coefficients)	Estimator	k	z	Bias	RMSE	Coverage $I_{na}^{0.95}$	Coverage $I_{sn}^{0.95}$
(55)	GQMLE	0.780	Hill	$k^* = 58.72$	$x_{\alpha,n}$	0.017	0.352	0.826	0.877
					$S_{\alpha,n}$	-0.312	0.748	0.644	0.932
			MR	$k^* = 64.66$	$x_{\alpha,n}$	0.217	0.398	0.785	0.761
					$S_{\alpha,n}$	-0.025	**0.629**	0.794	0.916
	3SNGQMLE(t_5)	0.615	Hill	$k^* = 59.40$	$x_{\alpha,n}$	0.021	**0.346**	0.865	0.901
					$S_{\alpha,n}$	-0.319	0.798	0.659	**0.935**
			MR	$k^* = 65.24$	$x_{\alpha,n}$	0.218	0.387	0.806	0.785
					$S_{\alpha,n}$	-0.035	0.669	0.831	0.919
	3SNGQMLE(t_{10})	0.630	Hill	$k^* = 59.53$	$x_{\alpha,n}$	0.024	0.357	0.856	**0.903**
					$S_{\alpha,n}$	-0.310	0.796	0.652	0.934
			MR	$k^* = 65.11$	$x_{\alpha,n}$	0.221	0.399	0.802	0.774
					$S_{\alpha,n}$	-0.028	0.674	0.825	0.913

Table B1.8

Average value of k, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method, with sum RMSEs of the coefficient estimates. The considered EVT methods are based on Hill and MR estimators with two methods in model estimation. The considered DGP is as in (56) with $n = 1000$ and 1000 simulation repetitions.

Model	Coefficient Estimator	RMSE(coefficients)	Estimator	k	z	Bias	RMSE	Coverage $I_{na}^{0.95}$	Coverage $I_{sn}^{0.95}$
(56)	GQMLE	0.640	Hill	$k^* = 55.79$	$x_{\alpha,n}$	0.059	0.249	0.876	0.906
					$S_{\alpha,n}$	-0.319	0.555	0.658	0.939
			MR	$k^* = 64.13$	$x_{\alpha,n}$	0.290	0.373	0.738	0.665
					$S_{\alpha,n}$	-0.002	**0.413**	0.878	**0.941**
	3SNGQMLE(t_5)	0.609	Hill	$k^* = 56.00$	$x_{\alpha,n}$	0.052	**0.248**	0.880	**0.916**
					$S_{\alpha,n}$	-0.340	0.581	0.660	0.937
			MR	$k^* = 64.18$	$x_{\alpha,n}$	0.282	0.365	0.753	0.680
					$S_{\alpha,n}$	-0.024	0.427	0.871	0.939
	3SNGQMLE(t_{10})	0.607	Hill	$k^* = 55.87$	$x_{\alpha,n}$	0.056	**0.248**	0.884	0.908
					$S_{\alpha,n}$	-0.332	0.574	0.675	0.936
			MR	$k^* = 64.43$	$x_{\alpha,n}$	0.286	0.367	0.745	0.671
					$S_{\alpha,n}$	-0.015	0.423	0.880	0.936

Table B1.9

Average value of k, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method, with sum RMSEs of the coefficient estimates. The considered EVT methods are based on Hill and MR estimators with two methods in model estimation. The considered DGP is as in (57) with $n = 1000$ and 1000 simulation repetitions.

Model	Coefficient Estimator	RMSE(coefficients)	Estimator	k	z	Bias	RMSE	Coverage	
								$I_{na}^{0.95}$	$I_{sn}^{0.95}$
(57)	GQMLE	0.653	Hill	$k^* = 55.66$	$x_{\alpha,n}$	0.033	0.237	0.857	0.898
					$S_{\alpha,n}$	-0.303	0.513	0.645	0.916
			MR	$k^* = 65.59$	$x_{\alpha,n}$	0.245	0.332	0.762	0.707
					$S_{\alpha,n}$	-0.018	0.377	0.860	0.932
	3SNGQMLE(t_5)	0.611	Hill	$k^* = 55.88$	$x_{\alpha,n}$	0.033	0.232	0.866	0.902
					$S_{\alpha,n}$	-0.313	0.518	0.641	0.919
			MR	$k^* = 65.53$	$x_{\alpha,n}$	0.243	0.328	0.773	0.727
					$S_{\alpha,n}$	-0.030	0.376	0.859	**0.933**
	3SNGQMLE(t_{10})	0.611	Hill	$k^* = 55.90$	$x_{\alpha,n}$	0.034	**0.230**	0.867	**0.908**
					$S_{\alpha,n}$	-0.307	0.511	0.648	0.924
			MR	$k^* = 65.58$	$x_{\alpha,n}$	0.244	0.327	0.767	0.725
					$S_{\alpha,n}$	-0.025	**0.372**	0.865	0.930

Table B1.10

Average value of k, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method, with sum RMSEs of the coefficient estimates. The considered EVT methods are based on Hill and MR estimators with two methods in model estimation. The considered DGP is as in (58) with $n = 1000$ and 1000 simulation repetitions.

Model	Coefficient Estimator	RMSE(coefficients)	Estimator	k	z	Bias	RMSE	Coverage	
								$I_{na}^{0.95}$	$I_{sn}^{0.95}$
(58)	GQMLE	0.610	Hill	$k^* = 55.02$	$x_{\alpha,n}$	0.045	0.204	0.879	0.900
					$S_{\alpha,n}$	-0.324	0.475	0.614	0.918
			MR	$k^* = 65.55$	$x_{\alpha,n}$	0.275	0.330	0.748	0.642
					$S_{\alpha,n}$	-0.021	**0.300**	0.906	0.932
	3SNGQMLE(t_5)	0.611	Hill	$k^* = 55.17$	$x_{\alpha,n}$	0.040	0.202	0.871	0.908
					$S_{\alpha,n}$	-0.339	0.488	0.595	0.906
			MR	$k^* = 65.45$	$x_{\alpha,n}$	0.270	0.325	0.745	0.657
					$S_{\alpha,n}$	-0.036	0.304	0.896	**0.945**
	3SNGQMLE(t_{10})	0.600	Hill	$k^* = 55.19$	$x_{\alpha,n}$	0.043	**0.200**	0.883	**0.909**
					$S_{\alpha,n}$	-0.332	0.481	0.601	0.902
			MR	$k^* = 65.47$	$x_{\alpha,n}$	0.272	0.326	0.743	0.655
					$S_{\alpha,n}$	-0.030	0.301	0.904	0.939

Table B1.11

Average value of k, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method, with sum RMSEs of the coefficient estimates. The considered EVT methods are based on Hill and MR estimators with two methods in model estimation. The considered DGP is as in (59) with $n = 1000$ and 1000 simulation repetitions.

Model	Coefficient Estimator	RMSE(coefficients)	Estimator	k	z	Bias	RMSE	Coverage $I_{na}^{0.95}$	$I_{sn}^{0.95}$
(59)	GQMLE	0.621	Hill	$k^* = 54.90$	$x_{\alpha,n}$	0.048	0.213	0.888	0.901
					$S_{\alpha,n}$	-0.294	0.450	0.626	0.928
			MR	$k^* = 67.00$	$x_{\alpha,n}$	0.267	0.334	0.723	0.620
					$S_{\alpha,n}$	-0.008	**0.303**	0.913	0.944
	3SNGQMLE(t_5)	0.616	Hill	$k^* = 55.05$	$x_{\alpha,n}$	0.044	0.213	0.885	0.907
					$S_{\alpha,n}$	-0.305	0.461	0.601	0.921
			MR	$k^* = 66.89$	$x_{\alpha,n}$	0.263	0.330	0.730	0.635
					$S_{\alpha,n}$	-0.018	0.306	0.911	0.945
	3SNGQMLE(t_{10})	0.611	Hill	$k^* = 54.97$	$x_{\alpha,n}$	0.046	**0.212**	0.885	**0.908**
					$S_{\alpha,n}$	-0.300	0.455	0.612	0.923
			MR	$k^* = 66.99$	$x_{\alpha,n}$	0.266	0.331	0.722	0.621
					$S_{\alpha,n}$	-0.013	**0.303**	0.915	**0.951**

Table B1.12

Average value of k, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method, with sum RMSEs of the coefficient estimates. The considered EVT methods are based on Hill and MR estimators with two methods in model estimation. The considered DGP is as in (60) with $n = 1000$ and 1000 simulation repetitions.

Model	Coefficient Estimator	RMSE(coefficients)	Estimator	k	z	Bias	RMSE	Coverage $I_{na}^{0.95}$	$I_{sn}^{0.95}$
(60)	GQMLE	0.594	Hill	$k^* = 54.70$	$\mathrm{x}_{\alpha,n}$	0.026	0.190	0.899	0.906
					$\mathrm{S}_{\alpha,n}$	-0.347	0.475	0.549	0.881
			MR	$k^* = 67.91$	$\mathrm{x}_{\alpha,n}$	0.257	0.308	0.734	0.616
					$\mathrm{S}_{\alpha,n}$	-0.048	0.280	0.918	0.954
	3SNGQMLE(t_5)	0.611	Hill	$k^* = 54.52$	$\mathrm{x}_{\alpha,n}$	0.021	0.191	0.894	0.905
					$\mathrm{S}_{\alpha,n}$	-0.359	0.484	0.535	0.894
			MR	$k^* = 67.80$	$\mathrm{x}_{\alpha,n}$	0.251	0.305	0.738	0.619
					$\mathrm{S}_{\alpha,n}$	-0.061	0.285	0.928	0.949
	3SNGQMLE(t_{10})	0.596	Hill	$k^* = 54.69$	$\mathrm{x}_{\alpha,n}$	0.023	0.189	0.895	0.911
					$\mathrm{S}_{\alpha,n}$	-0.353	0.478	0.543	0.883
			MR	$k^* = 67.81$	$\mathrm{x}_{\alpha,n}$	0.253	0.306	0.737	0.621
					$\mathrm{S}_{\alpha,n}$	-0.055	0.281	0.921	0.952

Table B1.13
Average value of k, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method, with sum RMSEs of the coefficient estimates. The considered EVT methods are based on Hill and MR estimators with two methods in model estimation. The considered DGP is as in (61) with $n = 1000$ and 1000 simulation repetitions.

Model	Coefficient Estimator	RMSE(coefficients)	Estimator	k	z	Bias	RMSE	Coverage	
								$I_{na}^{0.95}$	$I_{sn}^{0.95}$
(61)	GQMLE	0.593	Hill	$k^* = 54.57$	$\mathrm{x}_{\alpha,n}$	0.032	0.200	0.876	0.895
					$\mathrm{S}_{\alpha,n}$	-0.321	0.452	0.577	0.887
			MR	$k^* = 67.43$	$\mathrm{x}_{\alpha,n}$	0.256	0.315	0.688	0.592
					$\mathrm{S}_{\alpha,n}$	-0.033	0.279	0.911	0.941
	3SNGQMLE(t_5)	0.604	Hill	$k^* = 54.93$	$\mathrm{x}_{\alpha,n}$	0.030	0.192	0.876	0.902
					$\mathrm{S}_{\alpha,n}$	-0.330	0.454	0.558	0.895
			MR	$k^* = 67.23$	$\mathrm{x}_{\alpha,n}$	0.253	0.310	0.707	0.598
					$\mathrm{S}_{\alpha,n}$	-0.042	0.275	0.914	0.941
	3SNGQMLE(t_{10})	0.592	Hill	$k^* = 54.83$	$\mathrm{x}_{\alpha,n}$	0.032	0.194	0.878	0.900
					$\mathrm{S}_{\alpha,n}$	-0.325	0.452	0.571	0.892
			MR	$k^* = 67.35$	$\mathrm{x}_{\alpha,n}$	0.255	0.311	0.699	0.591
					$\mathrm{S}_{\alpha,n}$	-0.038	0.276	0.913	0.945

Table B1.14
Average value of k, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method, with sum RMSEs of the coefficient estimates. The considered EVT methods are based on Hill and MR estimators with two methods in model estimation. The considered DGP is as in (62) with $n = 1000$ and 1000 simulation repetitions.

Model	Coefficient Estimator	RMSE(coefficients)	Estimator	k	z	Bias	RMSE	Coverage	
								$I_{na}^{0.95}$	$I_{sn}^{0.95}$
(62)	GQMLE	0.861	Hill	$k^* = 58.06$	$x_{\alpha,n}$	-0.210	0.460	0.793	0.860
					$S_{\alpha,n}$	-0.752	1.109	0.454	0.872
			MR	$k^* = 64.49$	$x_{\alpha,n}$	0.043	0.386	0.909	0.876
					$S_{\alpha,n}$	-0.374	0.829	0.784	0.915
	3SNGQMLE(t_5)	0.694	Hill	$k^* = 58.40$	$x_{\alpha,n}$	-0.206	0.406	0.817	0.880
					$S_{\alpha,n}$	-0.749	1.055	0.472	0.869
			MR	$k^* = 64.73$	$x_{\alpha,n}$	0.042	**0.333**	**0.932**	0.893
					$S_{\alpha,n}$	-0.378	**0.773**	0.786	**0.931**
	3SNGQMLE(t_{10})	0.709	Hill	$k^* = 58.27$	$x_{\alpha,n}$	-0.203	0.412	0.816	0.881
					$S_{\alpha,n}$	-0.743	1.058	0.474	0.874
			MR	$k^* = 64.90$	$x_{\alpha,n}$	0.046	0.341	0.928	0.887
					$S_{\alpha,n}$	-0.370	0.778	0.782	0.925

Table B1.15

Average value of k, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method, with sum RMSEs of the coefficient estimates. The considered EVT methods are based on Hill and MR estimators with two methods in model estimation. The considered DGP is as in (63) with $n = 1000$ and 1000 simulation repetitions.

								Coverage	
Model	Coefficient Estimator	RMSE(coefficients)	Estimator	k	z	Bias	RMSE	$I_{na}^{0.95}$	$I_{sn}^{0.95}$
(63)	GQMLE	0.867	Hill	$k^* = 58.45$	$x_{\alpha,n}$	-0.194	0.412	0.782	0.861
					$S_{\alpha,n}$	-0.589	0.893	0.466	0.884
			MR	$k^* = 64.51$	$x_{\alpha,n}$	0.011	0.349	0.895	0.886
					$S_{\alpha,n}$	-0.295	0.685	0.777	0.923
	3SNGQMLE(t_5)	0.697	Hill	$k^* = 58.61$	$x_{\alpha,n}$	-0.179	0.350	0.810	0.889
					$S_{\alpha,n}$	-0.572	0.836	0.462	0.895
			MR	$k^* = 64.83$	$x_{\alpha,n}$	0.023	0.286	0.920	0.906
					$S_{\alpha,n}$	-0.281	0.615	0.788	**0.937**
	3SNGQMLE(t_{10})	0.713	Hill	$k^* = 58.27$	$x_{\alpha,n}$	-0.178	0.347	0.808	0.896
					$S_{\alpha,n}$	-0.567	0.824	0.475	0.899
			MR	$k^* = 64.90$	$x_{\alpha,n}$	0.024	**0.285**	**0.923**	0.902
					$S_{\alpha,n}$	-0.277	**0.607**	0.796	0.935

Table B1.16

Average value of k, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method, with sum RMSEs of the coefficient estimates. The considered EVT methods are based on Hill and MR estimators with two methods in model estimation. The considered DGP is as in (64) with $n = 1000$ and 1000 simulation repetitions.

								Coverage	
Model	Coefficient Estimator	RMSE(coefficients)	Estimator	k	z	Bias	RMSE	$I_{na}^{0.95}$	$I_{sn}^{0.95}$
(64)	GQMLE	0.719	Hill	$k^* = 55.78$	$x_{\alpha,n}$	-0.220	0.365	0.762	0.845
					$S_{\alpha,n}$	-0.714	0.889	0.317	0.800
			MR	$k^* = 63.81$	$x_{\alpha,n}$	0.026	0.267	**0.947**	0.884
					$S_{\alpha,n}$	-0.374	**0.596**	0.753	0.903
	3SNGQMLE(t_5)	0.667	Hill	$k^* = 55.55$	$x_{\alpha,n}$	-0.223	0.368	0.762	0.862
					$S_{\alpha,n}$	-0.722	0.903	0.305	0.803
			MR	$k^* = 63.86$	$x_{\alpha,n}$	0.022	0.264	**0.947**	0.893
					$S_{\alpha,n}$	-0.384	0.609	0.752	0.901
	3SNGQMLE(t_{10})	0.668	Hill	$k^* = 55.67$	$x_{\alpha,n}$	-0.222	0.365	0.772	0.860
					$S_{\alpha,n}$	-0.718	0.897	0.304	0.792
			MR	$k^* = 63.86$	$x_{\alpha,n}$	0.024	**0.262**	**0.953**	0.894
					$S_{\alpha,n}$	-0.380	0.602	0.768	**0.906**

Table B1.17

Average value of k, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method, with sum RMSEs of the coefficient estimates. The considered EVT methods are based on Hill and MR estimators with two methods in model estimation. The considered DGP is as in (65) with $n = 1000$ and 1000 simulation repetitions.

								Coverage	
Model	Coefficient Estimator	RMSE(coefficients)	Estimator	k	z	Bias	RMSE	$I_{na}^{0.95}$	$I_{sn}^{0.95}$
(65)	GQMLE	0.741	Hill	$k^* = 55.80$	$x_{\alpha,n}$	-0.207	0.378	0.726	0.812
					$S_{\alpha,n}$	-0.634	0.813	0.290	0.769
			MR	$k^* = 64.40$	$x_{\alpha,n}$	0.016	0.299	0.928	0.882
					$S_{\alpha,n}$	-0.332	0.570	0.730	0.885
	3SNGQMLE(t_5)	0.681	Hill	$k^* = 55.85$	$x_{\alpha,n}$	-0.212	0.359	0.739	0.822
					$S_{\alpha,n}$	-0.642	0.803	0.269	0.768
			MR	$k^* = 64.44$	$x_{\alpha,n}$	0.012	**0.273**	**0.941**	0.905
					$S_{\alpha,n}$	-0.340	**0.553**	0.733	0.886
	3SNGQMLE(t_{10})	0.684	Hill	$k^* = 55.83$	$x_{\alpha,n}$	-0.209	0.362	0.728	0.823
					$S_{\alpha,n}$	-0.637	0.803	0.282	0.768
			MR	$k^* = 64.35$	$x_{\alpha,n}$	0.014	0.279	0.939	0.898
					$S_{\alpha,n}$	-0.336	0.555	0.736	**0.887**

Table B1.18

Average value of k, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method, with sum RMSEs of the coefficient estimates. The considered EVT methods are based on Hill and MR estimators with two methods in model estimation. The considered DGP is as in (66) with $n = 1000$ and 1000 simulation repetitions.

								Coverage	
Model	Coefficient Estimator	RMSE(coefficients)	Estimator	k	z	Bias	RMSE	$I_{na}^{0.95}$	$I_{sn}^{0.95}$
(66)	GQMLE	0.664	Hill	$k^* = 54.98$	$x_{\alpha,n}$	-0.210	0.330	0.728	0.824
					$S_{\alpha,n}$	-0.664	0.794	0.239	0.735
			MR	$k^* = 65.13$	$x_{\alpha,n}$	0.035	**0.236**	0.928	0.878
					$S_{\alpha,n}$	-0.339	**0.510**	0.701	0.870
	3SNGQMLE(t_5)	0.662	Hill	$k^* = 55.01$	$x_{\alpha,n}$	-0.214	0.338	0.712	0.807
					$S_{\alpha,n}$	-0.672	0.805	0.216	0.727
			MR	$k^* = 64.87$	$x_{\alpha,n}$	0.031	0.242	**0.933**	0.875
					$S_{\alpha,n}$	-0.347	0.520	0.708	0.873
	3SNGQMLE(t_{10})	0.655	Hill	$k^* = 55.12$	$x_{\alpha,n}$	-0.212	0.336	0.720	0.820
					$S_{\alpha,n}$	-0.668	0.801	0.214	0.733
			MR	$k^* = 64.91$	$x_{\alpha,n}$	0.034	0.242	**0.933**	0.869
					$S_{\alpha,n}$	-0.343	0.517	0.706	**0.881**

Table B1.19

Average value of k, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method, with sum RMSEs of the coefficient estimates. The considered EVT methods are based on Hill and MR estimators with two methods in model estimation. The considered DGP is as in (67) with $n = 1000$ and 1000 simulation repetitions.

								Coverage	
Model	Coefficient Estimator	RMSE(coefficients)	Estimator	k	z	Bias	RMSE	$I_{na}^{0.95}$	$I_{sn}^{0.95}$
(67)	GQMLE	0.677	Hill	$k^* = 55.13$	$x_{\alpha,n}$	-0.200	0.319	0.708	0.780
					$S_{\alpha,n}$	-0.621	0.741	0.226	0.700
			MR	$k^* = 66.42$	$x_{\alpha,n}$	0.033	0.233	**0.932**	0.878
					$S_{\alpha,n}$	-0.315	**0.475**	0.718	0.835
	3SNGQMLE(t_5)	0.679	Hill	$k^* = 55.00$	$x_{\alpha,n}$	-0.202	0.319	0.709	0.780
					$S_{\alpha,n}$	-0.627	0.750	0.220	0.691
			MR	$k^* = 66.21$	$x_{\alpha,n}$	0.030	**0.227**	0.927	0.880
					$S_{\alpha,n}$	-0.322	0.482	0.713	0.840
	3SNGQMLE(t_{10})	0.669	Hill	$k^* = 55.17$	$x_{\alpha,n}$	-0.201	0.317	0.701	0.772
					$S_{\alpha,n}$	-0.623	0.744	0.224	0.693
			MR	$k^* = 66.33$	$x_{\alpha,n}$	0.031	**0.227**	0.930	0.878
					$S_{\alpha,n}$	-0.319	0.477	0.713	**0.842**

Table B1.20

Average value of k, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method, with sum RMSEs of the coefficient estimates. The considered EVT methods are based on Hill and MR estimators with two methods in model estimation. The considered DGP is as in (68) with $n = 1000$ and 1000 simulation repetitions.

								Coverage	
Model	Coefficient Estimator	RMSE(coefficients)	Estimator	k	z	Bias	RMSE	$I_{na}^{0.95}$	$I_{sn}^{0.95}$
(68)	GQMLE	0.657	Hill	$k^* = 54.72$	$\mathrm{x}_{\alpha,n}$	-0.217	0.33	0.698	0.793
					$\mathrm{S}_{\alpha,n}$	-0.658	0.774	0.161	0.655
			MR	$k^* = 67.23$	$\mathrm{x}_{\alpha,n}$	0.021	0.224	0.932	0.887
					$\mathrm{S}_{\alpha,n}$	-0.349	0.494	0.69	0.824
	3SNGQMLE(t_5)	0.663	Hill	$k^* = 54.64$	$\mathrm{x}_{\alpha,n}$	-0.219	0.338	0.701	0.787
					$\mathrm{S}_{\alpha,n}$	-0.664	0.783	0.149	0.635
			MR	$k^* = 67.20$	$\mathrm{x}_{\alpha,n}$	0.019	0.234	0.939	0.884
					$\mathrm{S}_{\alpha,n}$	-0.354	0.503	0.693	0.828
	3SNGQMLE(t_{10})	0.649	Hill	$k^* = 54.88$	$\mathrm{x}_{\alpha,n}$	-0.218	0.333	0.707	0.791
					$\mathrm{S}_{\alpha,n}$	-0.661	0.777	0.154	0.641
			MR	$k^* = 67.18$	$\mathrm{x}_{\alpha,n}$	0.02	0.229	0.935	0.891
					$\mathrm{S}_{\alpha,n}$	-0.352	0.498	0.699	0.824

Table B1.21

Average value of k, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method, with sum RMSEs of the coefficient estimates. The considered EVT methods are based on Hill and MR estimators with two methods in model estimation. The considered DGP is as in (69) with $n = 1000$ and 1000 simulation repetitions.

Model	Coefficient Estimator	RMSE(coefficients)	Estimator	k	z	Bias	RMSE	Coverage $I_{na}^{0.95}$	$I_{sn}^{0.95}$
(69)	GQMLE	0.653	Hill	$k^* = 54.66$	$x_{\alpha,n}$	-0.205	0.293	0.701	0.764
					$S_{\alpha,n}$	-0.627	0.720	0.172	0.612
			MR	$k^* = 67.39$	$x_{\alpha,n}$	0.027	0.185	0.945	0.874
					$S_{\alpha,n}$	-0.327	0.440	0.671	0.822
	3SNGQMLE(t_5)	0.669	Hill	$k^* = 54.58$	$x_{\alpha,n}$	-0.209	0.298	0.687	0.766
					$S_{\alpha,n}$	-0.634	0.731	0.164	0.602
			MR	$k^* = 67.55$	$x_{\alpha,n}$	0.023	0.188	0.944	0.870
					$S_{\alpha,n}$	-0.334	0.451	0.664	0.822
	3SNGQMLE(t_{10})	0.654	Hill	$k^* = 54.55$	$x_{\alpha,n}$	-0.206	0.294	0.688	0.760
					$S_{\alpha,n}$	-0.630	0.725	0.167	0.610
			MR	$k^* = 67.35$	$x_{\alpha,n}$	0.025	0.185	0.945	0.875
					$S_{\alpha,n}$	-0.330	0.444	0.666	0.817

B.2 Additional Simulation Result for 3.4

Table B2.1

Average value of k^*, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method with k^* and $\hat{k}$. The considered EVT methods are based on Hill and MR estimators with model estimated using GQMLE and 3SNGQMLE. The considered DGP is as in (42) with $n = 1000$ and 1000 simulation repetitions.

Model	Coefficient Estimator	Estimator	k	z	Bias	RMSE	Coverage $I_{na}^{0.95}$	Coverage $I_{sn}^{0.95}$
(43)	GQMLE	Hill	$k^* = 57.45$	$x_{\alpha,n}$	0.026	0.332	0.850	0.887
				$S_{\alpha,n}$	-0.265	0.658	0.685	**0.948**
			$\hat{k} = 71$	$x_{\alpha,n}$	-0.061	0.358	0.916	0.918
				$S_{\alpha,n}$	-0.618	0.937	0.547	0.890
		MR	$k^* = 65.40$	$x_{\alpha,n}$	0.222	0.389	0.790	0.744
				$S_{\alpha,n}$	0.014	0.570	0.842	0.917
			$\hat{k} = 71$	$x_{\alpha,n}$	0.207	0.387	0.821	0.745
				$S_{\alpha,n}$	-0.213	0.656	0.861	0.920
	3SNGQMLE(t_{10})	Hill	$k^* = 57.86$	$x_{\alpha,n}$	0.034	**0.278**	0.885	0.910
				$S_{\alpha,n}$	-0.255	0.594	0.694	0.958
			$\hat{k} = 71$	$x_{\alpha,n}$	-0.052	0.297	**0.935**	0.923
				$S_{\alpha,n}$	-0.606	0.865	0.553	0.894
		MR	$k^* = 66.10$	$x_{\alpha,n}$	0.227	0.346	0.809	0.762
				$S_{\alpha,n}$	0.019	**0.501**	0.848	0.928
			$\hat{k} = 71$	$x_{\alpha,n}$	0.211	0.344	0.837	0.763
				$S_{\alpha,n}$	-0.207	0.571	0.873	0.927

Table B2.2

Average value of k^*, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method with k^* and $\hat{k}$. The considered EVT methods are based on Hill and MR estimators with model estimated using GQMLE and 3SNGQMLE. The considered DGP is as in (42) with $n = 1000$ and 1000 simulation repetitions.

Model	Coefficient Estimator	Estimator	k	z	Bias	RMSE	Coverage $I_{na}^{0.95}$	Coverage $I_{sn}^{0.95}$
(54)	GQMLE	Hill	$k^* = 57.93$	$x_{\alpha,n}$	0.046	0.358	0.870	0.911
				$S_{\alpha,n}$	-0.372	0.818	0.679	**0.954**
			$\hat{k} = 71$	$x_{\alpha,n}$	-0.057	0.381	0.917	0.928
				$S_{\alpha,n}$	-0.827	1.199	0.530	0.891
		MR	$k^* = 64.51$	$x_{\alpha,n}$	0.282	0.447	0.805	0.764
				$S_{\alpha,n}$	-0.021	0.672	0.840	0.940
			$\hat{k} = 71$	$x_{\alpha,n}$	0.270	0.443	0.828	0.765
				$S_{\alpha,n}$	-0.306	0.779	0.856	0.938
	3SNGQMLE(t_{10})	Hill	$k^* = 58.31$	$x_{\alpha,n}$	0.045	**0.320**	0.886	0.927
				$S_{\alpha,n}$	-0.383	0.791	0.688	0.955
			$\hat{k} = 71$	$x_{\alpha,n}$	-0.055	0.338	0.935	**0.932**
				$S_{\alpha,n}$	-0.827	1.147	0.526	0.894
		MR	$k^* = 65.19$	$x_{\alpha,n}$	0.276	0.424	0.816	0.781
				$S_{\alpha,n}$	-0.040	**0.654**	0.862	0.942
			$\hat{k} = 71$	$x_{\alpha,n}$	0.261	0.421	0.845	0.783
				$S_{\alpha,n}$	-0.323	0.753	0.865	0.940

Table B2.3
Average value of k^*, bias, RMSE, coverage probabilities and length of confidence intervals for $\hat{x}_{\alpha,n}$ and $\hat{S}_{\alpha,n}$ with $\alpha = 0.01$, estimated using EVT method with k^* and $\hat{k}$. The considered EVT methods are based on Hill and MR estimators with model estimated using GQMLE and 3SNGQMLE. The considered DGP is as in (42) with $n = 1000$ and 1000 simulation repetitions.

Model	Coefficient Estimator	Estimator	k	z	Bias	RMSE	Coverage $I_{na}^{0.95}$	$I_{sn}^{0.95}$
(62)	GQMLE	Hill	$k^* = 57.47$	$x_{\alpha,n}$	0.055	0.434	0.837	0.881
				$S_{\alpha,n}$	-0.362	0.889	0.637	0.933
			$\hat{k} = 71$	$x_{\alpha,n}$	-0.054	0.452	0.880	0.901
				$S_{\alpha,n}$	-0.844	1.257	0.490	0.859
		MR	$k^* = 62.84$	$x_{\alpha,n}$	0.306	0.514	0.765	0.746
				$S_{\alpha,n}$	0.011	0.749	0.808	0.931
			$\hat{k} = 71$	$x_{\alpha,n}$	0.294	0.511	0.799	0.753
				$S_{\alpha,n}$	-0.290	0.847	0.826	0.918
	3SNGQMLE(t_{10})	Hill	$k^* = 57.63$	$x_{\alpha,n}$	0.062	**0.379**	0.863	0.903
				$S_{\alpha,n}$	-0.364	0.831	0.659	**0.940**
			$\hat{k} = 71$	$x_{\alpha,n}$	-0.045	0.392	0.906	**0.916**
				$S_{\alpha,n}$	-0.840	1.209	0.497	0.882
		MR	$k^* = 64.30$	$x_{\alpha,n}$	0.305	0.473	0.780	0.766
				$S_{\alpha,n}$	-0.001	**0.691**	0.823	0.926
			$\hat{k} = 71$	$x_{\alpha,n}$	0.291	0.472	0.813	0.778
				$S_{\alpha,n}$	-0.303	0.801	0.831	0.930

B.3 Additional Empirical Results

Table B3.1

P-values of UC and CC test and quantile score for one-step-ahead $x_{\alpha,n}$ forecast together with AL log score for one-step-ahead $S_{\alpha,n}$ forecast of CAC40, where $\alpha = 0.005$

Index	Model	Method			UC	CC	Quantile score	AL log score
CAC40	GARCH(1,1)	GQMLE	$\hat{k}$	Hill	0.061*	1	1.015	-8349.89
				MR	0.22	0.011**	<u>**0.993**</u>	-8460.04
			k^*	Hill	0.225	0.106	0.997	**−8486.993**
				MR	0.148	0.012**	0.996	-8460.27
		3SNGQMLE	$\hat{k}$	Hill	0.152	1	1.018	-8354.16
				MR	0.094*	0.014**	1.009	-8395.74
			k^*	Hill	0.904	0.138	1.002	-8473.46
				MR	0.058*	0.015**	1.013	-8386.99
		Parametric			0.061	0.088	1.023	-8380.962
	ARMA(1,0)-GARCH(1,1)	GQMLE	$\hat{k}$	Hill	0.061*	1	1.015	-8350.67
				MR	0.434	0.15	<u>**0.993**</u>	-8458.63
			k^*	Hill	0.225	0.106	0.996	-8485.09
				MR	0.315	0.153	0.995	-8463.1
		3SNGQMLE	$\hat{k}$	Hill	0.061*	1	1.013	-8367.75
				MR	0.314	0.01	0.994	-8448.94
			k^*	Hill	0.747	0.128	0.993	<u>**−8501.952**</u>
				MR	0.094*	0.014**	0.997	-8442.38
		Parametric			0.061	0.088	1.023	-8379.010
	ARMA(1,1)-GARCH(1,1)	GQMLE	$\hat{k}$	Hill	0.061*	1	1.015	-8348.91
				MR	0.315	0.01	**0.994**	-8455.19
			k^*	Hill	0.225	0.106	0.997	**−8483.78**
				MR	0.148	0.012**	0.996	-8453.77
		3SNGQMLE	$\hat{k}$	Hill	0.098*	1	1.017	-8348.53
				MR	0.219	0.011**	1.004	-8401.19
			k^*	Hill	0.904	0.138	1	-8470.61
				MR	0.094*	0.014**	1.008	-8391.28
		Parametric			0.061	0.088	1.024	-8370.274
	ARMA(1,1)-GJR-GARCH(1,1)	GQMLE	$\hat{k}$	Hill	0.098*	1	1.035	-8353.94
				MR	0.001***	0.154	1.035	-8331.32
			k^*	Hill	0.917	0.134	1.027	-8387.12
				MR	0.001***	0.154	1.041	-8274.36
		3SNGQMLE	$\hat{k}$	Hill	0.227	1	1.025	-8393.86
				MR	0.034**	0.16	1.023	-8380.73
			k^*	Hill	0.573	0.146	**1.019**	**−8441.622**
				MR	0.034**	0.16	1.026	-8339.99
		Parametric			0.442	0.999	1.034	-8411.463

Note. * indicates significance at 10%, ** indicates significance at 5%, and *** indicates significance at 1%.

Table B3.2

P-values of UC and CC test and quantile score for one-step-ahead $x_{\alpha,n}$ forecast together with AL log score for one-step-ahead $S_{\alpha,n}$ forecast of DJI, where $\alpha = 0.005$

Index	Model	Method			UC	CC	Quantile score	AL log score
DJI	GARCH(1,1)	GQMLE	$\hat{k}$	Hill	0.352	0.108	0.79	-9109.87
				MR	0.127	0.001***	**0.771**	$\mathbf{-9269.175}$
			k^*	Hill	0.682	0.007***	0.779	-9225.62
				MR	0.049**	0.001***	0.774	-9252.25
		3SNGQMLE	$\hat{k}$	Hill	0.353	0.999	0.787	-9139.52
				MR	0.028**	0.001***	0.773	-9243.92
			k^*	Hill	0.85	0.133	0.777	-9232.51
				MR	0.028**	0.001***	0.777	-9228.1
		Parametric			0.110	1.000	0.802	-9133.778
	ARMA(1,0)-GARCH(1,1)	GQMLE	$\hat{k}$	Hill	0.249	0.102	0.791	-9101.96
				MR	0.127	0.001***	0.771	-9265.09
			k^*	Hill	0.682	0.007***	0.779	-9225.4
				MR	0.049**	0.001***	0.774	-9250.24
		3SNGQMLE	$\hat{k}$	Hill	0.25	1	0.78	-9173.13
				MR	0.048**	0.014**	**0.758**	$\mathbf{-9317.754}$
			k^*	Hill	0.972	0.992	0.767	-9274.91
				MR	0.028**	0.001***	0.763	-9299.87
		Parametric			0.110	1.000	0.802	-9135.885
	ARMA(1,1)-GARCH(1,1)	GQMLE	$\hat{k}$	Hill	0.249	0.102	0.791	-9104.42
				MR	0.127	0.001***	**0.771**	$\mathbf{-9267.057}$
			k^*	Hill	0.682	0.007***	0.778	-9228.15
				MR	0.049**	0.001***	0.774	-9253.3
		3SNGQMLE	$\hat{k}$	Hill	0.353	0.999	0.787	-9143.71
				MR	0.028**	0.001***	0.773	-9251.33
			k^*	Hill	0.85	0.133	0.777	-9232.51
				MR	0.028**	0.001***	0.777	-9229.91
		Parametric			0.110	1.000	0.802	-9135.885
	ARMA(1,1)-GJR-GARCH(1,1)	GQMLE	$\hat{k}$	Hill	0.11	1	0.763	-9243.85
				MR	0.127	0.92	0.73	-9450.32
			k^*	Hill	0.48	0.999	0.745	-9363.46
				MR	0.028**	0.85	0.731	-9468.62
		3SNGQMLE	$\hat{k}$	Hill	0.041**	1	0.761	-9272.1
				MR	0.279	0.952	$\underline{\mathbf{0.727}}$	-9461.87
			k^*	Hill	0.25	1	0.741	-9397.64
				MR	0.08*	0.899	0.728	$\underline{\mathbf{-9471.748}}$
		Parametric			0.041	1	0.781	-9229.8

Note. * indicates significance at 10%, ** indicates significance at 5%, and *** indicates significance at 1%.

Table B3.3

P-values of UC and CC test and quantile score for one-step-ahead $x_{\alpha,n}$ forecast together with AL log score for one-step-ahead $S_{\alpha,n}$ forecast of GDAXI, where $\alpha = 0.005$

Index	Model	Method			UC	CC	Quantile score	AL log score
GDAXI	GARCH(1,1)	GQMLE	$\hat{k}$	Hill	0.064*	1.000	0.999	-8276.641
				MR	0.032**	0.016**	0.973	-8429.394
			k^*	Hill	0.603	0.998	**0.969**	$\mathbf{-8478.759}$
				MR	0.018**	0.017**	0.973	-8453.724
		3SNGQMLE	$\hat{k}$	Hill	0.064*	1.000	0.997	-8310.101
				MR	0.032**	0.158	0.975	-8410.189
			k^*	Hill	0.458	0.999	0.969	-8471.466
				MR	0.032**	0.158	0.972	-8432.074
		Parametric			0.022	1.000	1.016	-8306.475
	ARMA(1,0)-GARCH(1,1)	GQMLE	$\hat{k}$	Hill	0.064*	1.000	0.997	-8284.672
				MR	0.054*	0.015**	0.973	-8434.676
			k^*	Hill	0.603	0.998	**0.968**	$\mathbf{-8485.091}$
				MR	0.032**	0.016**	0.972	-8458.212
		3SNGQMLE	$\hat{k}$	Hill	0.064*	1.000	1.008	-8258.094
				MR	0.089*	0.157	0.979	-8362.935
			k^*	Hill	0.458	0.999	0.979	-8407.418
				MR	0.089*	0.157	0.977	-8378.788
		Parametric			0.064	1.000	1.024	-8274.403
	ARMA(1,1)-GARCH(1,1)	GQMLE	$\hat{k}$	Hill	0.064*	1.000	0.997	-8285.406
				MR	0.054*	0.015**	0.972	-8436.013
			k^*	Hill	0.603	0.998	<u>**0.967**</u>	$\underline{\mathbf{-8487.041}}$
				MR	0.032**	0.016**	0.972	-8458.860
		3SNGQMLE	$\hat{k}$	Hill	0.159	0.099*	1.031	-8194.249
				MR	0.032**	0.016**	1.007	-8273.620
			k^*	Hill	0.766	0.126	1.002	-8341.301
				MR	0.032**	0.016**	1.006	-8278.335
		Parametric			0.038	1.000	1.018	-8294.144
	ARMA(1,1)-GJR-GARCH(1,1)	GQMLE	$\hat{k}$	Hill	0.335	0.999	1.001	-8329.239
				MR	0***	0.026**	1.003	-8284.454
			k^*	Hill	0.416	0.147	0.988	-8381.558
				MR	0***	0.026**	1.003	-8278.463
		3SNGQMLE	$\hat{k}$	Hill	0.766	0.996	1.001	-8347.151
				MR	0.001***	0.151	0.997	-8297.357
			k^*	Hill	0.208	0.153	**0.987**	$\mathbf{-8418.035}$
				MR	0.001***	0.151	0.999	-8280.203
		Parametric			0.159	1.000	1.021	-8332.851

Note. * indicates significance at 10%, ** indicates significance at 5%, and *** indicates significance at 1%.

Table B3.4

P-values of UC and CC test and quantile score for one-step-ahead $x_{\alpha,n}$ forecast together with AL log score for one-step-ahead $S_{\alpha,n}$ forecast of HSI, where $\alpha = 0.005$

Index	Model	Method			UC	CC	Quantile score	AL log score
HSI	GARCH(1,1)	GQMLE	$\hat{k}$	Hill	0.050	0.073*	0.985	-7940.409
				MR	0.097*	0.001***	0.983	-8039.546
			k^*	Hill	0.083*	0.001***	0.970	-8078.663
				MR	0.097*	0***	0.985	-8033.145
		3SNGQMLE	$\hat{k}$	Hill	0.016**	1.000	0.961	-8043.929
				MR	0.151	0***	0.946	-8199.277
			k^*	Hill	0.029**	0.068*	**0.938**	**−8213.388**
				MR	0.034**	0***	0.954	-8184.611
		Parametric			0.008	1.000	0.993	-8042.463
	ARMA(1,0)-GARCH(1,1)	GQMLE	$\hat{k}$	Hill	0.050	0.073*	0.988	-7933.981
				MR	0.152	0***	0.983	-8031.099
			k^*	Hill	0.132	0.002***	**0.966**	**−8097.73**
				MR	0.152	0***	0.988	-8022.184
		3SNGQMLE	$\hat{k}$	Hill	0.133	0***	0.986	-7976.857
				MR	0.059*	0***	0.991	-8035.187
			k^*	Hill	0.293	0***	0.981	-8082.382
				MR	0.019**	0***	0.997	-8011.871
		Parametric			0.008	1.000	0.988	-8064.060
	ARMA(1,1)-GARCH(1,1)	GQMLE	$\hat{k}$	Hill	0.050	0.073*	0.988	-7936.318
				MR	0.152	0***	0.983	-8030.647
			k^*	Hill	0.132	0.002***	**0.966**	-8096.370
				MR	0.152	0***	0.987	-8024.243
		3SNGQMLE	$\hat{k}$	Hill	0.084*	0.078*	0.979	-7990.634
				MR	0.096*	0.001***	0.973	-8091.992
			k^*	Hill	0.202	0.002***	0.970	**−8114.44**
				MR	0.034**	0***	0.978	-8076.125
		Parametric			0.008	1.000	0.985	-8077.279
	ARMA(1,1)-GJR-GARCH(1,1)	GQMLE	$\hat{k}$	Hill	0.029**	1.000	0.949	-8079.471
				MR	0.152	0.009***	0.927	-8247.380
			k^*	Hill	0.050	1.000	0.927	-8249.356
				MR	0.059*	0.011**	0.931	-8255.223
		3SNGQMLE	$\hat{k}$	Hill	0.016**	1.000	0.938	-8147.449
				MR	0.325	0.134	**<u>0.916</u>**	-8316.625
			k^*	Hill	0.202	1.000	0.920	-8298.768
				MR	0.226	0.008***	0.918	**<u>−8329.556</u>**
		Parametric			0.008	1.000	0.951	-8201.880

Note. * indicates significance at 10%, ** indicates significance at 5%, and *** indicates significance at 1%.

Table B3.5

P-values of UC and CC test together with quantile score and AL log score for one-step-ahead $x_{\alpha,n}$ and $S_{\alpha,n}$ forecasts of Nikkei 225 with filtered data from 1/1/1997 to 31/12/2016, $\alpha = 0.005$ and $n = 1000$.

Index	Model	Method			UC	CC	Quantile score	AL log score
N225	ARMA(1,0)-GARCH(1,1)	GQMLE	$\hat{k}$	Hill	0.697	1.000	0.577	-3895.662
				MR	0.071*	0.947	0.590	-3854.531
			k^*	Hill	0.813	0.997	0.581	-3890.553
				MR	0.015**	0.896	0.594	-3830.416
		3SNGQMLE	$\hat{k}$	Hill	0.697	1.000	**0.573**	**−3904.435**
				MR	0.006***	0***	0.587	-3856.574
			k^*	Hill	0.939	0.999	0.578	-3893.909
				MR	0.002***	0***	0.591	-3829.911

Table B3.6

P-values of UC and CC test together with quantile score and AL log score for one-step-ahead $x_{\alpha,n}$ and $S_{\alpha,n}$ forecasts of Nikkei 225 with filtered data from 1/1/1997 to 31/12/2016, $\alpha = 0.005$ and $n = 1000$.

Index	Model	Method			UC	CC	Quantile score	AL log score
N225	ARMA(1,1)-GARCH(1,1)	GQMLE	$\hat{k}$	Hill	0.697	1.000	0.577	-3896.705
				MR	0.071*	0.947	0.590	-3855.512
			k^*	Hill	0.813	0.997	0.581	-3891.261
				MR	0.015**	0.896	0.594	-3831.205
		3SNGQMLE	$\hat{k}$	Hill	0.697	1.000	**0.575**	$\mathbf{-3900.817}$
				MR	0.006***	0***	0.589	-3851.606
			k^*	Hill	0.939	0.999	0.579	-3888.824
				MR	0.002***	0***	0.593	-3820.122

Table B3.7

P-values of UC and CC test together with quantile score and AL log score for one-step-ahead $x_{\alpha,n}$ and $S_{\alpha,n}$ forecasts of Nikkei 225 with filtered data from 1/1/1997 to 31/12/2016, $\alpha = 0.005$ and $n = 1000$.

Index	Model	Method			UC	CC	Quantile score	AL log score
N225	ARMA(1,1)-GJR-GARCH(1,1)	GQMLE	$\hat{k}$	Hill	0.309	1.000	0.571	-3913.687
				MR	0.034**	0.924	0.570	-3944.222
			k^*	Hill	0.309	1.000	0.563	-3963.633
				MR	0.034**	0.924	0.575	-3930.322
		3SNGQMLE	$\hat{k}$	Hill	0.309	1.000	0.563	-3946.524
				MR	0.388	0.987	0.565	-3971.853
			k^*	Hill	0.697	1.000	**0.56**	$\mathbf{-3980.805}$
				MR	0.583	0.993	0.565	-3972.351

Table B3.8

P-values of the one-sided DM test-statistics for the quantile score and AL log score

	Quantile score	AL log score
NASDAQ	0.0023	0.026
N225	0.005	0.03