Acta Mathematica Sinica, English Series
Published online: June 10, 2020
https://doi.org/10.1007/s10114-020-9100-3
http://www.ActaMath.com

Acta Mathematica Sinica, English Series

A New Robust Risk Measure: Quantile Shortfall

You Li CHEN
Law School, Wuhan University, Wuhan 430072*, P. R. China*
E-mail: 39298915*@qq.com*

Yan Yan LIU[1)]
School of Mathematics and Statistics, Wuhan University, Wuhan 430072*, P. R. China*
E-mail: *liuyy@whu.edu.cn*

Guang Cai MAO
School of Mathematics and Statistics, Central China Normal University,
Wuhan 430079*, P. R. China*
E-mail: *maoguangcai@mail.ccnu.edu.cn*

Yuan Shan WU
School of Statistics and Mathematics, Zhongnan University of Economics and Law,
Wuhan 430073*, P. R. China*
E-mail: *wu@zuel.edu.cn*

Fei YAN
School of Mathematics and Statistics, Wuhan University, Wuhan 430072*, P. R. China*
E-mail: 275785436*@qq.com*

Abstract Among recent measures for risk management, value at risk (VaR) has been criticized because it is not coherent and expected shortfall (ES) has been criticized because it is not robust to outliers. Recently, [*Math. Oper. Res.*, **38**, 393–417 (2013)] proposed a risk measure called median shortfall (MS) which is distributional robust and easy to implement. In this paper, we propose a more generalized risk measure called quantile shortfall (QS) which includes MS as a special case. QS measures the conditional quantile loss of the tail risk and inherits the merits of MS. We construct an estimator of the QS and establish the asymptotic normality behavior of the estimator. Our simulation shows that the newly proposed measures compare favorably in robustness with other widely used measures such as ES and VaR.

Keywords Nonparametric estimation, quantile shortfall, risk measure, robust

MR(2010) Subject Classification 62G08, 62G35

Received March 28, 2019, revised October 16, 2019, accepted December 5, 2019
Supported by the National Natural Science Foundation of China (Grant No. 11571263) and Fundamental Research Funds for the Central Universities (Grant No. 2042018kf0243)
1) Corresponding author

1 Introduction

The extreme risk exists in various areas of financial investment, credit and insurance. For investors and risk managers, predicting the probability of the extreme risk loss is an important task. A common risk quantification index named as Value at Risk (VaR), measures the maximum potential loss of a given portfolio over a prescribed holding period at a given confidence level e. g. 1% or 5%. Assessing VaR amounts to estimating the tail quantiles of the conditional distribution of financial returns. The parametric, semiparametric and nonparametric models have been widely used to compute the VaR [9]. Although VaR has become one of the standard measure of financial market risk, it has been criticized for not being sub-additive [4]. This means that diversification does not necessarily reduce VaR therefore it is contrast to the framework of modern portfolio theory. In addition, VaR ignores the statistical properties of significant loss beyond the quantile point of interest [1–3].

To overcome the shortcomings of VaR, [2] proposed an alternative risk measure named as Expected Shortfall (ES), which is defined as the conditional expected return given that it exceeds the VaR. In general, ES has the coherence property except when the underlying loss distributions have discontinuities. Since the concept of ES was put forward, it has become an important tool in financial market risk prediction. However, the parametric and semiparametric methods in the VaR theory provide only estimation for VaR and it is not clear how to calculate the corresponding ES [10]. The most widely used nonparametric VaR method is historical simulation, which estimates the VaR as the quantile of the empirical distribution of historical returns from a moving window of the most recent periods. By this approach, the ES can be estimated as the mean of the returns, in the moving window, that exceed the VaR estimate. However, it is difficult to make choice of how many past periods should be included in the moving window. Using too few observations will lead to large sampling error, while including too many will result in estimates that are slow to react to changes in the true distribution. Based on the fact that there exists a one-to-one mapping between expectiles and quantiles, [10] proposed to estimate ES by using asymmetric least squares methods. [6] proposed the dynamic additive quantile (DAQ) model for calculating conditional quantiles. However, since ES is defined as the mean of the tail loss distribution, the ES estimators are not robust [12]. Therefore, [8] proposed to use the median of shortfall (MS) to measure the risk. Taking into account that the median estimation is more robust than the mean estimation and is less susceptible to outliers, MS has a desirable property of distributional robustness with respect to model misspecification in the sense that a small deviation of the model only results in a small change in the risk measurement [7].

In practical studies, not only the conditional median but also other conditional quantities is of interest. It motivates a more general risk measure, referred as Quantile Shortfall (QS), which measures the conditional quantile loss of the tail risk, in this paper. Compared with MS, instead of looking only at local properties of specified quantile, QS takes into account both global model coherence and local approximation. We propose an estimation procedure of QS and unravel its asymptotic property, which makes statistical inference for the QS feasible.

The paper is organized as follows. In Section 2, we introduce the definition of QS and propose an estimate procedure for QS. The asymptotic properties of the proposed estimator is described. In Section 3, we conduct simulation studies to examine the finite sample performance of the proposed estimator. Section 4 employs six stocks closing price to evaluate the advantages of our proposed estimator. Some remarks are concluded in Section 5.

2 Notation and Method

Denote the loss return on the equity of a financial institution as X and that of the entire market as Y. In practice, we are often interested in understanding the τ $(0 < \tau < 1)$ quantile of the equity of the financial institution X conditioning on the event that the entire market equity Y is beyond a given cutpoint b. To address this issue, we propose the quantile shortfall (QS), which is defined as

$$\xi_\tau(b) = \inf\{x : G(x; b) \geq \tau\},$$

where $G(x; b) = P(X \leq x | Y > b)$ is the conditional distribution of X given Y is greater than b. Obviously, the QS is more robust to some extreme observations than ES. Furthermore, QS provides a complete description of the equity of the financial institution by varying the quantile levels. In particular, if $\tau = 0.5$, the median shortfall $\xi_{0.5}(b)$ can be considered as a comparable counterpart of the ES. As usual, b is chosen to be the higher quantile level of Y.

For $i = 1, \ldots, n$, (X_i, Y_i) are assumed to be independent and identically distributed copies of (X, Y). We can estimate $G(x; b)$ by its plug-in version

$$\widehat{G}_n(x; \widehat{b}) = \frac{\sum_{i=1}^n I(X_i \leq x, Y_i > \widehat{b})}{\sum_{i=1}^n I(Y_i > \widehat{b})}, \tag{2.1}$$

where $\widehat{b}$, as the estimator of b, is the $(1-p)$ quantile of observed sample $Y_1, \ldots, Y_n$. Then, $\xi_\tau(b)$ can be estimated by

$$\widehat{\xi}_\tau(\widehat{b}) = \inf\{x : \widehat{G}_n(x; \widehat{b}) \geq \tau\}.$$

Remark 2.1 For any $\boldsymbol{a} = (a_1, \ldots, a_d)^{\mathrm{T}} \in \mathbb{R}^d$ and $\boldsymbol{b} = (b_1, \ldots, b_d)^{\mathrm{T}} \in \mathbb{R}^d$, denote $\boldsymbol{a} \leq \boldsymbol{b}$ if $a_j \leq b_j$ for $j = 1, \ldots, d$. For d-dimensional variable $\boldsymbol{Y}$, the QS can also be defined as $\xi_\tau(\boldsymbol{b}) = \inf\{x : G(x; \boldsymbol{b}) \geq \tau\}$, where $G(x; \boldsymbol{b}) = P(X \leq x | \boldsymbol{Y} > \boldsymbol{b})$.

Theorem 2.2 *For $P(Y > b) > 0$, $\widehat{\xi}_\tau(\widehat{b})$ is a consistent estimator of $\xi_\tau(b)$.*

Theorem 2.3 *Suppose that $P(Y > b) > 0$, then we have*

$$\sqrt{n}(\widehat{\xi}_\tau(\widehat{b}) - \xi_\tau(b)) \xrightarrow{d} N(0, \sigma_b^2),$$

where $\sigma_b^2 = \frac{\tau - \tau^2 P(Y>b)}{P(Y>b) f(\xi_\tau(b) \mid Y>b)^2}$ and $f(x \mid Y > b)$ is the conditional density function of X given $Y > b$.

Although the asymptotic properties of $\widehat{\xi}_\tau(\widehat{b})$ have been established, we do not apply these theorems to construct the confidence interval of $\xi_\tau(b)$. Suggested by [11], in practice, the $(1-\alpha)$ confidence interval can be constructed by following procedure. Let $(X_{(1)}, Y_{(1)}), \ldots, (X_{(m)}, Y_{(m)})$ be the sub-sample of the observed sample $(X_1, Y_1), \ldots, (X_n, Y_n)$ such that $Y_{(i)} > \widehat{b}$ for $i = 1, \ldots, m$. Denote the α_1 and α_2 quantile among $X_{(1)}, \ldots, X_{(m)}$ by $\xi^*_{\alpha_1}$ and $\xi^*_{\alpha_2}$ respectively, where $\alpha_1 = \tau - u_{1-\alpha/2}\sqrt{\frac{\tau(1-\tau)}{m}}$, $\alpha_2 = \tau + u_{1-\alpha/2}\sqrt{\frac{\tau(1-\tau)}{m}}$ and $u_{1-\alpha/2}$ denote the $1 - \alpha/2$ quantile of the standard normal distribution. Then, the $(1 - \alpha)$ confidence interval of $\xi_\tau(b)$ is given by $[\xi^*_{\alpha_1}, \xi^*_{\alpha_2}]$.

3 Simulation Studies

We conduct some simulation studies to evaluate the finite sample performance of the proposed method.

Example 3.1 We first generate (X, Y) from bivariate normal distribution $N(\mu, A)$ and considered the following cases of (μ, A).

(a) $\mu = \binom{0}{0}$, $A = \begin{pmatrix} 2 & 0.4 \\ 0.4 & 2 \end{pmatrix}$, where X and Y are positively correlated.

(b) $\mu = \binom{0}{0}$, $A = \begin{pmatrix} 2 & -0.4 \\ -0.4 & 2 \end{pmatrix}$, where X and Y are negatively correlated.

(c) $\mu = \binom{0}{0}$, $A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$, where X and Y are uncorrelated.

We consider the sample size $n = 400$ or 500 and $\tau = 0.2$, 0.5 or 0.7. b is set as the 0.95 quantile of the distribution Y. In each scenario, we repeat 1000 simulations.

Table 1 summarizes the simulation results under each setup, including the bias, the standard error and the mean square error of the proposed estimator, the coverage probability and the average interval length of 95% confidence intervals, also the simulation results of ES and VaR. It can be seen that all the bias, standard error and mean square error of the proposed estimator and the average interval length decrease when the sample size increases from 400 to 500. Furthermore, the coverage probabilities of 95% confidence intervals are around the ideal level. Consequently, our method performs well under the finite sample settings.

Case	Statistics	$n = 400$					$n = 500$				
		Bias	SE	MSE	CP	LCI	Bias	SE	MSE	CP	LCI
(a)	QS($\tau = 0.2$)	0.082	0.415	0.179	0.949	1.856	0.052	0.368	0.138	0.951	1.632
	QS($\tau = 0.5$)	−0.017	0.380	0.144	0.948	1.504	−0.004	0.334	0.112	0.957	1.358
	QS($\tau = 0.7$)	−0.060	0.382	0.150	0.945	1.620	−0.043	0.348	0.123	0.954	1.437
	ES	−0.008	0.310	0.096	–	–	−0.011	0.274	0.075	–	–
	VaR	−0.003	0.106	0.011	–	–	−0.003	0.097	0.009	–	–
(b)	QS($\tau = 0.2$)	0.074	0.412	0.175	0.949	1.874	0.058	0.391	0.156	0.939	1.646
	QS($\tau = 0.5$)	−0.020	0.369	0.137	0.949	1.502	−0.024	0.355	0.127	0.944	1.350
	QS($\tau = 0.7$)	−0.071	0.389	0.157	0.946	1.606	−0.041	0.368	0.137	0.948	1.460
	ES	−0.019	0.305	0.094	–	–	−0.008	0.287	0.082	–	–
	VaR	−0.005	0.107	0.012	–	–	−0.003	0.097	0.009	–	–
(c)	QS($\tau = 0.2$)	0.096	0.428	0.193	0.960	1.958	0.079	0.389	0.158	0.953	1.650
	QS($\tau = 0.5$)	0.019	0.382	0.146	0.946	1.518	−0.001	0.347	0.120	0.946	1.368
	QS($\tau = 0.7$)	−0.040	0.401	0.162	0.936	1.634	−0.047	0.364	0.134	0.952	1.464
	ES	0.009	0.314	0.098	–	–	−0.002	0.282	0.080	–	–
	VaR	−0.003	0.110	0.012	–	–	0.000	0.101	0.010	–	–

Bias, SE and MSE: the bias, standard error and mean square error of the proposed estimator; CP: the coverage probability of 95% confidence intervals; LCI: the average interval length of 95% confidence intervals

Table 1 Simulation results of Example 1

Example 3.2 We further generate (X, Y) from non-normal distribution and considered the joint density function of (X, Y) taking the following forms.

(a) $f(x, y) = 4e^{-(2x+2y)}I(x > 0, y > 0)$, where X and Y are independent.

(b) $f(x, y) = (x + y)I(0 \leq x \leq 1, 0 \leq y \leq 1)$, where X and Y are nonindependent.

The simulation results are summarized in Table 2, from which we can draw the similar conclusions.

Case	τ	$n = 400$					$n = 500$				
		Bias	SE	MSE	CP	LCI	Bias	SE	MSE	CP	LCI
	QS($\tau = 0.2$)	0.021	0.058	0.004	0.962	0.216	0.020	0.053	0.003	0.946	0.195
	QS($\tau = 0.5$)	0.012	0.111	0.013	0.940	0.436	0.009	0.102	0.011	0.942	0.396
(a)	QS($\tau = 0.7$)	−0.011	0.164	0.027	0.941	0.706	0.001	0.147	0.022	0.950	0.644
	ES	−0.005	0.110	0.012	–	–	0.002	0.099	0.010	–	–
	VaR	−0.176	0.076	0.037	–	–	−0.178	0.072	0.037	–	–
	QS($\tau = 0.2$)	0.009	0.053	0.003	0.957	0.202	0.006	0.047	0.002	0.945	0.184
	QS($\tau = 0.5$)	−0.001	0.053	0.003	0.953	0.204	−0.001	0.047	0.002	0.948	0.182
(b)	QS($\tau = 0.7$)	−0.002	0.041	0.002	0.958	0.167	−0.004	0.038	0.001	0.951	0.149
	ES	0.001	0.031	0.001	–	–	0.000	0.029	0.001	–	–
	VaR	0.178	0.015	0.032	–	–	0.178	0.013	0.032	–	–

Bias, SE and MSE: the bias, standard error and mean square error of the proposed estimator; CP: the coverage probability of 95% confidence intervals; LCI: the average interval length of 95% confidence intervals

Table 2 Simulation results of Example 2

4 Real Data Analysis

In this section, we use MS and ES as risk measures to analyze data include six stocks daily closing prices of 000625.SZ, 300002.SZ, 601898.SH, 002146.SZ, 002554.SZ, 600030.SH, from January 4, 2014 to December 31, 2015. The total number of observations is 489 and the amount of the initial investment is supposed to be $1,000$ million yuan.

We first plot the daily closing price movements of our chosen stock in Figure 1. It can be seen that the stock closing price has the nature of the peak which is better for us to compare the robustness of ES and MS.

Further the VaR estimate is calculated based on the daily return rate of each stock. There are three common models for calculating VaR, namely Historical Simulation, Variance-Covariance Method and Monte Carlo Simulation. We apply the Variance-Covariance Method which is mainly used by the Riskmetrics System to get the VaR estimate. The confidence level is chosen as 95%. Then we calculate ES value, which is the mean of the daily return rate beyond VaR at 95% level, and, simultaneously, MS which is the median of the daily return rate beyond VaR at 95% level. We are also inspired to compute the τ-th quantile of the daily return rate beyond VaR which is defined as the τ-th quantile shortfall (QS_τ). The trend of each stock return rate and the comparation between ES value and MS value are drawn in Figures 2–7.

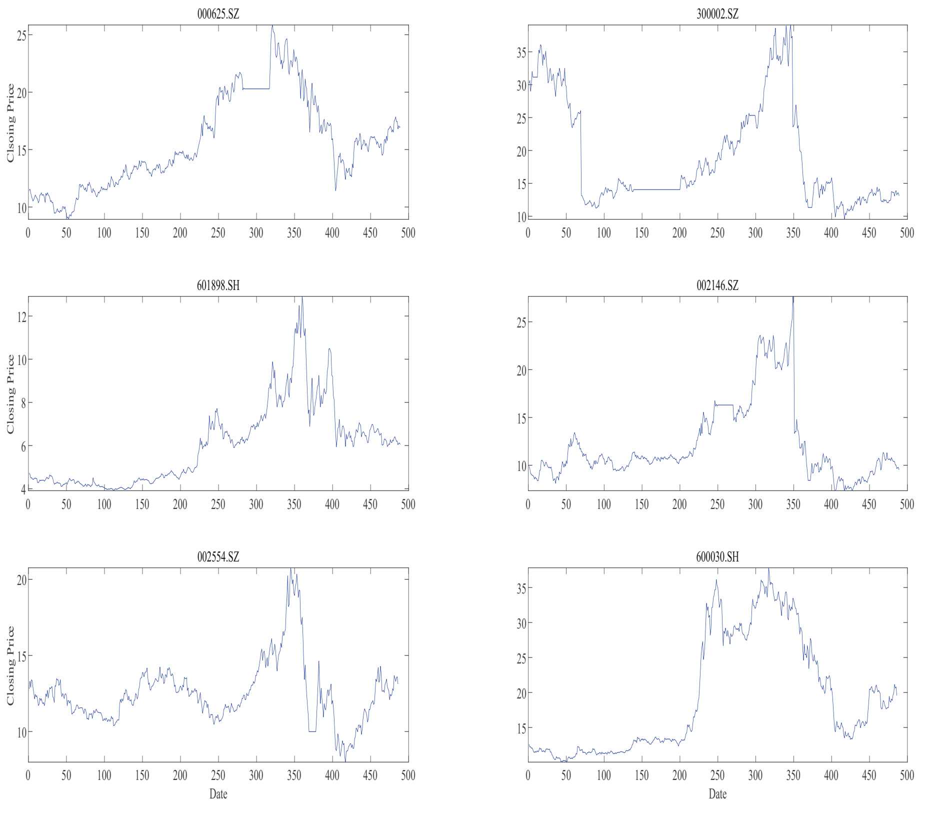

Figure 1 The x-coordinate corresponds to the 489 days from January 4, 2014 to December 31, 2015. The y-coordinate indicates the closing price of each stock along with time variation. We can see that the stock price has many peaks and is unstable

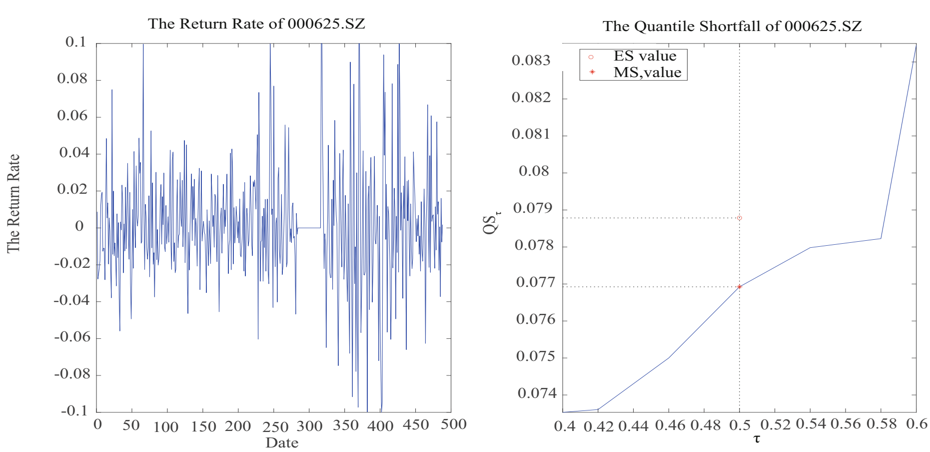

Figure 2 The yield curve of 000625.SZ is haphazard, and the differences between the ES value and MS values result from the presence of some yield extreme values

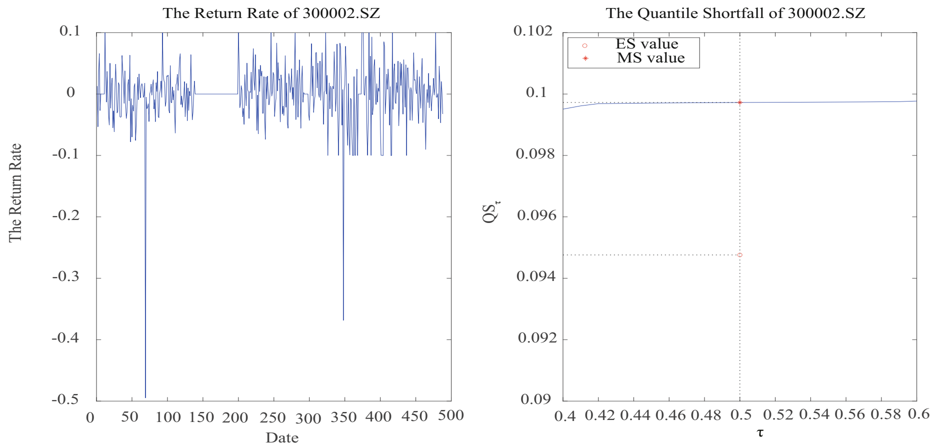

Figure 3 That the MS value is greater than the ES value is obvious because of the presence of the minimum rate of return, and the average level is reduced

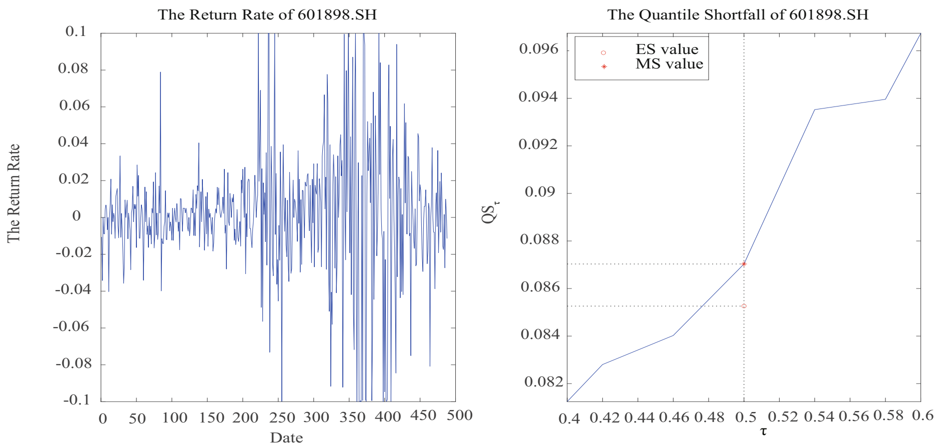

Figure 4 The yield curve of 601898.SH is also haphazard, and the differences between the ES value and MS values result from the presence of some extreme values

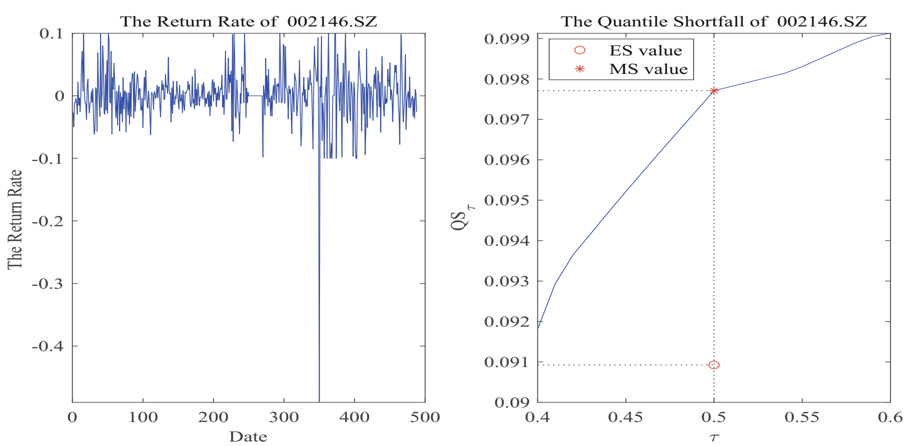

Figure 5 It is obvious that the reason why the ES value is much smaller than MS value is the existence of the unusually small yield values

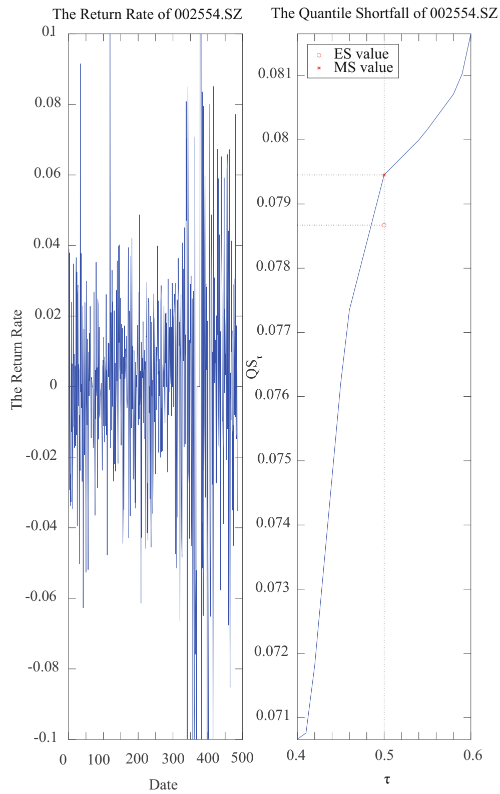

Figure 6 The yield curve of 002554.SZ is also haphazard, and the differences between the ES value and MS values result from the presence of some extreme values

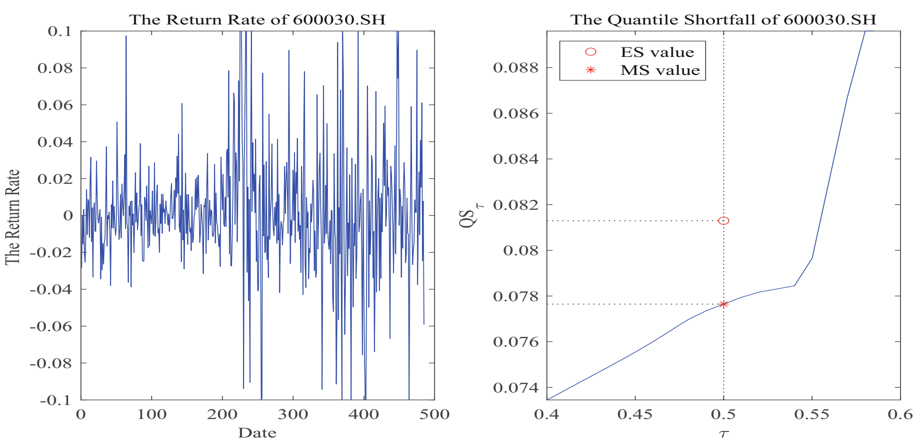

Figure 7 The yield curve of 600030.SH is also haphazard, and the differences between the ES value and MS values result from the presence of some extreme values

We can see that the volatility of the stock return rate has the characteristics of time-varying and some peaks. The average will be too high or too low because of the outlier observations, while the median leads to less sensitivity to extreme outlier observations. The estimators of the real data are listed in Table 3.

Stock	VaR estimate	ES estimate	MS estimate
000625.SZ	0.051	0.079	0.078
300002.SZ	0.079	0.095	0.099
601898.SH	0.058	0.085	0.087
002146.SZ	0.072	0.091	0.098
002554.SZ	0.055	0.077	0.079
600030.SH	0.056	0.081	0.078

Table 3 The estimators of the real data

Having known the calculated values of each QS_τ with τ varying from 0.1 to 0.9, now we concentrate on the confidence intervals corresponding to each QS_τ when $\tau = 0.2$, $\tau = 0.5$ and $\tau = 0.7$. The confidence level is $\alpha = 0.95$. Figure 8 shows the results.

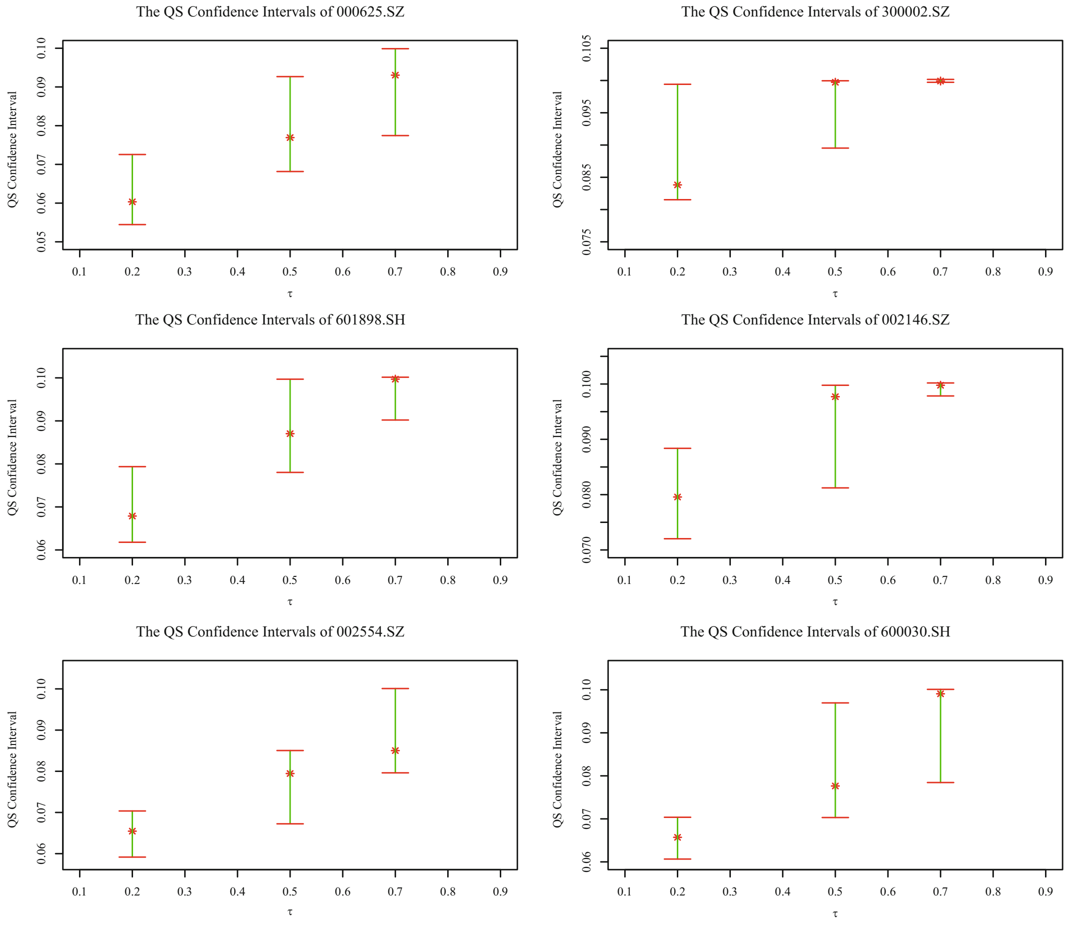

Figure 8 The confidence intervals corresponding to each QS_τ when $\tau = 0.2$, $\tau = 0.5$ and $\tau = 0.7$ of each stock

5 Concluding Remarks

In this paper, we propose a more generalized risk measure Quantile Shortfall (QS) which is an extension of MS proposed by [7]. Compared with the MS, QS can capture the global properties of conditional risk loss while MS only focuses on the central trend. We further propose an easy implement estimator of QS and study the asymptotic properties of proposed QS estimators. Simulation studies and real data analysis show the robustness of our proposed method.

Acknowledgements We thank the referees for their time and comments.

References

[1] Acerbi, C., Tasche, D.: On the coherence of expected shortfall. *J. Bank Financ.*, **26**, 1487–1503 (2002)
[2] Artzner, P., Delbaen, F., Eber, J. M., et al.: Thinking coherently. *Risk*, **10**, 68–71 (1997)
[3] Artzner, P., Delbaen, F., Eber, J. M., et al.: Coherent measures of risk. *Math. Financ.*, **9**, 203–228 (1999)
[4] Carlo, A., Dirk, T.: On the coherence of expected shortfall. *J. Bank Financ.*, **26**, 1487–1503 (2002)
[5] Doss, H., Gill, R. D.: An elementary approach to weak convergence for quantile processes, with applications to censored survival data. *J. Am. Stat. Assoc.*, **87**, 869–877 (1992)
[6] Gourieroux, C., Jasiak, J.: Dynamic quantile models. *J. Econom.*, **147**, 198–205 (2008)
[7] Kou, S., Peng, X.: Expected shortfall or median shortfall. *J. Financ. Eng.*, **1**, 1–6 (2014)
[8] Kou, S., Peng, X., Heyde, C. C.: External risk measures and Basel accords. *Math. Oper. Res.*, **38**, 393–417 (2013)
[9] Linsmeier, T. J., Pearson, N. D.: Risk Management: An introduction to Value at Risk. Working paper, University of Illinois at Urbana-Champaign, 1996
[10] Taylor, J. W.: Estimating value at risk and expected shortfall using expectiles. *J. Financ. Econom.*, **6**, 231–252 (2008)
[11] Van der Vaart, A. W.: Asymptotic Statistics, Cambridge University Press, Cambridge, 1998
[12] Yamai, Y., Yoshiba, T.: Comparative analyses of expected shortfall and value-at-risk: Their estimation error, decomposition, and optimization. *Monetary and Econ. Stud.*, **20**, 181–237 (2002)

Appendix

Proof of Theorem 2.2 To prove the consistency of $\widehat{\xi}_\tau(b)$, we just need to show that $\widehat{G}_n(x;\widehat{b})$ converges to $G(x;b)$ in probability for all x. It follows from [11, Lemma 21.7] that $|\widehat{b}-b| = o_p(1)$. On the other hand, by Glivenko–Cantelli theorem, we have

$$\sup_{-\infty<t<\infty}\left|\frac{1}{n}\sum_{i=1}^{n} I(Y_i > t) - P(Y > t)\right| = o_p(1).$$

As a consequence,

$$\left|\frac{1}{n}\sum_{i=1}^{n} I(Y_i > \widehat{b}) - P(Y > b)\right| = o_p(1). \tag{5.1}$$

Utilizing the analogous arguments, we can also conclude that

$$\sup_{-\infty<x<\infty}\left|\frac{1}{n}\sum_{i=1}^{n} I(X_i \le x, Y_i > \widehat{b}) - P(X \le x, Y > b)\right| = o_p(1).$$

Following Slutsky theorem, we have

$$\sup_{-\infty<x<\infty}|\widehat{G}(x;\widehat{b}) - G(x;b)| = o_p(1),$$

which proves Theorem 2.2.

Proof of Theorem 2.3 Following [11, Lemma 19.24] and combining the fact $|\widehat{b} - b| = o_p(1)$ shown in the proof of Theorem 2.2, we have

$$\sqrt{n}\bigg(\frac{1}{n}\sum_{i=1}^{n} I(X_i \le x, Y_i > \widehat{b}) - P(X \le x, Y > b)\bigg) \xrightarrow{d} N(0, \sigma^2),$$

where

$$\sigma^2 = P(X \le x, Y > b) - \{P(X \le x, Y > b)\}^2.$$

By (5.1) and Slutsky theorem, we conclude that

$$\sqrt{n}(\widehat{G}_n(x; \widehat{b}) - G(x; b)) \xrightarrow{d} N\bigg(0, \frac{\sigma^2}{\{P(Y > b)\}^2}\bigg).$$

Utilizing Vervaat's Method in quantile regression [5], we have

$$\sqrt{n}(\widehat{\xi_\tau}(\widehat{b}) - \xi_\tau(b)) \xrightarrow{d} N(0, \sigma_b^2).$$

Hence, we complete the proof of Theorem 2.3.