European Journal of Operational Research 312 (2024) 1168–1177

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Innovative Applications of O.R.

Kernel quantile estimators for nested simulation with application to portfolio value-at-risk measurement

Xiaoyu Liu[a], Xing Yan[b], Kun Zhang[b,*]

[a] Department of Service Research and Development, Inspur Cloud, Jinan, China
[b] Institute of Statistics and Big Data, Renmin University of China, No. 59 Zhongguancun Street, Haidian District, Beijing, China

ARTICLE INFO

Article history:
Received 13 December 2022
Accepted 30 July 2023
Available online 6 August 2023

Keywords:
Simulation
Value-at-risk
Kernel quantile estimator
Bandwidth selection
Budget allocation

ABSTRACT

Nested simulation has been widely used in portfolio risk measurement in recent years. We focus on one risk measure, value at risk (VaR), and study a kernel quantile estimator (KQE) for nested simulation to estimate this risk. We analyze the bias, variance, and mean squared error (MSE), based on which we show that the variance is reduced in the lower-order terms, while in some cases bias could be reduced in the dominant term. For practical implementation, we propose an efficient bootstrap-based algorithm to guide kernel bandwidth selection and budget allocation in nested simulation. We also conduct numerical experiments to show that KQE works quite well at different significance levels compared with the widely used sample quantile.

1. Introduction

Nested estimation involves estimating the functional of a conditional expectation without analytical expression, formulated as

$$\theta = \rho(\mathrm{E}[X|\xi]),$$

where ξ is a random vector in $\mathbb{R}^d$ with $d \geq 1$, X is a random variable in $\mathbb{R}^1$, and the functional ρ is a mapping from a random variable to a real number. Typically, the analytical expression of a conditional expectation is unknown, and the functional is nonlinear. Therefore, a nested simulation is required for the estimation, which is conducted at two levels. At the outer level, several samples of ξ are simulated, and at the inner level, several samples of X are simulated given each outer level sample. Because of this two-level framework, the simulation budget is prohibitive. This study improves nested simulation estimators, especially their finite sample performance.

Nested simulation has important applications in portfolio risk measurement, where ξ denotes the risk factors at some future time (known as the risk horizon), $\mathrm{E}[X|\xi]$ denotes the portfolio loss whose distribution is unknown, and the functional ρ is the risk measure. Often a portfolio contains tens of thousands of financial instruments, including many sophisticated financial derivatives, so its loss–the conditional expectation $\mathrm{E}[X|\xi]$–does not have a closed form. Therefore, a nested simulation must be used to estimate the portfolio loss over the risk horizon. Additionally, different functionals lead to different risk measures. For example, Zhang, Liu, & Wang (2022b) and Wang, Wang, & Zhang (2022b) studied five functionals, or five risk measures.

We consider a widely used risk measure, value at risk (VaR), which is defined as the worst expected loss over a target horizon given a significance level. That is, the functional ρ is a quantile operator–$\rho(Y) = \inf\{y : F_Y(y) \geq \alpha\}$–where F_Y denotes the cumulative distribution function of Y and α is the fixed significance level. In addition to being used as a measure of financial risk, VaR has been applied to input uncertainty quantification for stochastic simulation (Xie, Nelson, & Barton, 2014; Zhu, Liu, & Zhou, 2020).

The estimation of portfolio VaR and other risk measures of interest within a nested simulation framework has received a lot of attention in the simulation community in recent years and several methods have been proposed in the literature. The estimation of portfolio VaR via nested simulation was first considered by Lee (1998), who assumed that the conditional expectation follows a normal distribution. It was then studied by Gordy & Juneja (2010) under a different but more reasonable set of assumptions. Gordy & Juneja (2010) examined the budget allocation problem for nested simulation and provided the asymptotic sample sizes at the optimal outer and inner levels under the criterion of minimizing the asymptotic mean squared error (MSE). However, the coefficients in optimal parameter formulations are normally unknown and involve quantities, such as the density of the conditional expectation and its derivatives, which are arguably more difficult to

* Corresponding author.
E-mail addresses: liuxiaoyu12@inspur.com (X. Liu), xingyan@ruc.edu.cn (X. Yan), kunzhang@ruc.edu.cn (K. Zhang).

https://doi.org/10.1016/j.ejor.2023.07.040
0377-2217/

estimate than the risk measure, the quantity of interest. Therefore, to tackle this issue, Zhang et al. (2022b) used bootstrap sampling and presented a sample-driven budget allocation algorithm for practical use. Sun, Apley, & Staum (2011) and Goda (2017) constructed unbiased estimators for the variance of the conditional expectation, so that its estimation only requires a finite number of inner-level samples that guarantee the convergence of the nested simulation estimator as the total simulation budget goes to infinity. Furthermore, Cheng & Zhang (2021) proposed three non-nested estimators for the higher-order central moments of a conditional expectation and proved the asymptotic bias, variance, and MSE, and the central limit theorem. In addition to point estimation, Cheng, Liu, & Zhang (2022) set unified confidence intervals (CIs) for nested simulation, in the sense that both the MSE of the point estimator and the CI width attain the optimal convergence rate.

Some researchers have used more efficient sampling strategies by allocating the simulation budget non-uniformly across outer-level scenarios. For example, Liu, Nelson, & Staum (2010) adaptively allocated the inner-level samples to each outer-level scenario based on ranking and selection tools when estimating conditional VaR (CVaR). Similarly, Lan, Nelson, & Staum (2015) combined tools from the ranking and selection literature and statistical theory of empirical likelihood to construct CIs for CVaR. Furthermore, Broadie, Du, & Moallemi (2011a) proposed a sequential method to allocate different inner-level sample sizes to different outer-level scenarios and reduce bias and MSE when estimating the probability of large losses. To estimate the same risk measure as in Broadie et al. (2011a), Wang, Xu, Hu, & Chen (2022a) combined the well-known optimal computing budget allocation method, which allowed them to allocate inner-level samples to each outer-level scenario in a sequential manner.

Gordy & Juneja (2010) showed that the squared asymptotic bias of the nested simulation estimator is not negligible compared with the asymptotic variance, such that the optimal MSE is the result of the trade-off between bias and variance. Therefore, most researchers have focused on reducing bias and proposed various parametric or nonparametric models to fit the inner conditional expectation as a surface. Broadie, Du, & Moallemi (2011b, 2015) used a group of basis functions to fit the conditional expectation via least squares regression, which uses information from all outer-level scenarios in fitting. However, in practice, basis functions cannot span the true conditional expectation in the $\mathcal{L}^2$ space, leading to a regression error in portfolio risk estimation, such that the MSE of the risk estimator converges to $\Gamma^{-1+\delta}$ until reaching this error, for all positive δ, where Γ denotes the simulation budget. Fort, Gobet, & Moulines (2017) designed and analyzed an algorithm based on least squares regression and Markov chain Monte Carlo sampling to estimate risk measures when the outer expectation is related to a rare event, and the convergence results are similar to those of Broadie, Du, & Moallemi (2015) but $\delta = 0$. To eliminate the regression error, Wang et al. (2022b) considered the fit of the conditional expectation in a subspace of $\mathcal{L}^2$, the reproducing kernel Hilbert space, and analyzed the relationship between the smoothness of the portfolio loss density function and the convergence rate of the MSE of the risk estimator.

Nonparametric methods have also been used in risk estimation. For example, Hong & Juneja (2009) used the kernel smoothing method within a nested simulation framework to fit the conditional expectation, an asymptotically unbiased estimator, showing an improvement in the convergence rate of the MSE for low-dimensional problems (the risk factors are of low dimension). Based on this, Hong, Juneja, & Liu (2017) proposed a decomposition technique for high-dimensional portfolio risk measurement problems, which allows an efficient application in practice. Liu & Staum (2010) used the metamodeling approach, such as stochastic kriging, to fit the conditional expectation in CVaR estimation. Motivated by importance sampling, Zhang, Feng, Liu, & Wang (2022a) constructed a weighted-average estimator of the conditional expectation via likelihood ratios, which is unbiased, leading to portfolio risk estimators with a reduction in bias to the order Γ^{-1}. Moreover, the asymptotic variance is of the order Γ^{-1}, as is the MSE, which is the same as the convergence order of the non-nested simulation estimators. However, their method requires known likelihood ratios of portfolio assets.

Almost all studies on nested simulation have focused on bias reduction. In this paper, however, we focus on variance reduction, because from the standpoint of finite sample performance, both bias and variance reductions are equally important. This is illustrated by a numerical example of portfolio VaR estimation in Section 6, and the squared biases and variances under estimated optimal allocation rules with different budgets over the significance level $\alpha = 0.95$ are shown in Fig. 1. It can be seen that the variances are much larger than the squared biases, so variance reduction is at least as important as bias reduction. Therefore, we focus mainly on variance reduction for portfolio VaR estimation.

Note that the studies we summarize above have all used a simple sample quantile for VaR estimation. However, Navruz & Özdemir (2020) showed that the sample quantile has long been considered inefficient because of the variability of individual order statistics and that a better quantile estimate should take advantage of all order statistics. The question of whether we can replace the sample quantile with a more advanced quantile estimator to improve efficiency certainly deserves further exploration. Therefore, we investigate a popular class of quantile estimators known as kernel quantile estimator (KQE) in the nested simulation framework for portfolio VaR estimation and include its asymptotic theoretical properties and experimental performance. Unlike studies that have generally improved the performance of nested simulation estimators by reducing their bias, in this study KQE uses the appropriate kernel weighting functions to average all order statistics (Al-Kenani & Yu, 2012; Cheng & Sun, 2006; Falk, 1984; McCune & McCune, 1991; Parzen, 1979; Sheather & Marron, 1990), which alleviates the variability of a single-order statistic to some extent. The asymptotic theoretical results of this study demonstrate that KQE achieves variance reduction for the portfolio VaR measurement, which represents a significant improvement over finite sample simulation studies. This study makes the following contributions to the literature.

1. *We use KQE to estimate portfolio VaR.* We study KQE, which takes advantage of all order statistics for the estimation of portfolio VaR. To the best of our knowledge, our study is the first to use this kernel-type approach to measure portfolio VaR.
2. *We provide a theoretical analysis of KQE within the nested simulation framework.* The asymptotic bias and variance of KQE for nested simulation are established. Compared with the sample quantile, KQE achieves a reduction in variance, which is of lower-order but shows up in a significant improvement over finite sample studies. Besides, under certain conditions, the bias of KQE may be reduced to the order of $N^{-3/2}$ while that of the sample quantile is N^{-1}. Next, we provide the asymptotic MSE for KQE and discuss its convergence rate. Based on this, we derive the optimal kernel bandwidth. We demonstrate that the MSE of KQE can converge at least as fast as the MSE of the sample quantile approach within the nested simulation framework.
3. *We develop an algorithm for the practical implementation of nested KQE.* An algorithm is proposed based on the bootstrap method to allocate budgets and determine appropriate bandwidths. Using the bootstrap technique, all unknown parameters of the budget allocation and bandwidth selection are estimated in light of our theoretical results, from which the optimal band-

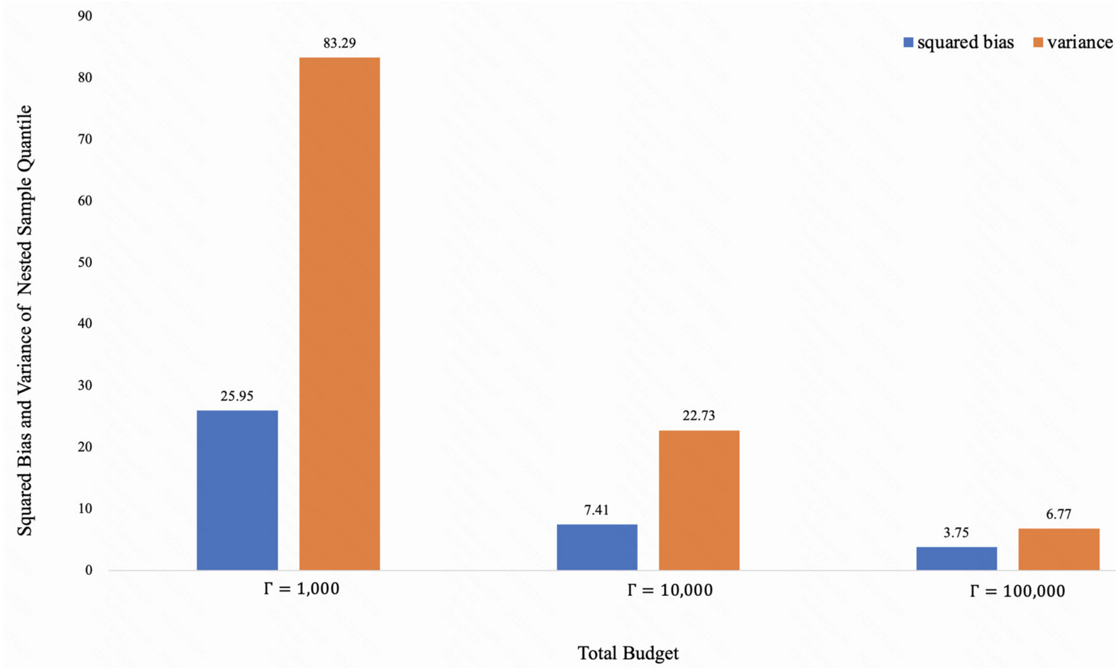

Fig. 1. Squared bias and variance for different budgets.

width choice, the outer- and inner-level sample sizes can be computed.

4. *We demonstrate the practical benefits of our proposed estimator through numerical experiments.* We compare the root mean squared error (RMSE) of KQE with that of the sample quantile in a simulation study, and the numerical results verify the effectiveness of our proposed estimator. First, our approach works quite well at different significance levels and for different total budgets, which suggests that KQE may be a viable tool for estimating portfolio VaR. Second, the lower variance and proportion of variance of the proposed estimator compared with sample quantile justify our original intention to reduce the variability of single-order statistics by formulating a weighted-average estimator using appropriate kernel functions.

The rest of this paper is organized as follows. We introduce the background of nested simulation in Section 2 and formulate the nested KQE for portfolio VaR estimation in Section 3. In Section 4, we present our theoretical results, including the asymptotic bias and variance, based on which we further investigate the convergence rate of the MSE. We propose an efficient algorithm for practical implementation in Section 5 and present our numerical results in Section 6, before concluding the study in Section 7. For the sake of clarity, we present all the proofs in the appendix.

2. Background

We introduce the nested simulation approach of Lee (1998) and Gordy & Juneja (2010) and the asymptotic results of the nested sample quantile. For notational convenience, let $Y(\xi) = \mathrm{E}[X|\xi]$. When estimating portfolio VaR, nested simulation is performed at two levels, which is illustrated in Fig. 2.

(i) *Outer-level scenarios.* We draw L independent and identically distributed (i.i.d.) scenarios of ξ according to the physical (or real-world) distribution of ξ, denoted by $\{\xi_i, i = 1, \ldots, L\}$.
(ii) *Inner-level samples.* The inner simulation takes place under the risk-neutral distribution, conditioned on the scenario ξ_i. For each ξ_i, we generate N i.i.d. samples $\{X_{ij}, j = 1, \ldots, N\}$.

Then $Y(\xi_i)$ can be approximated by

$$\widetilde{Y}(\xi_i) = \frac{1}{N}\sum_{j=1}^{N} X_{ij}. \tag{1}$$

That is, we obtain a group of nested simulation estimators of portfolio losses for different scenarios. Let $\widetilde{Y}(\xi_1), \ldots, \widetilde{Y}(\xi_L)$ and $\widetilde{Y}_{[1]} \geq \cdots \geq \widetilde{Y}_{[L]}$ be the corresponding order statistics, then the estimator of portfolio VaR $y_\alpha = \inf\{y : F_Y(y) \geq \alpha\}$ is the nested sample quantile $\widetilde{Y}_{\lceil \alpha L \rceil}$, where the operator $\lceil x \rceil$ denotes the smallest integer that is not less than x.

Gordy & Juneja (2010) established the asymptotic bias and variance for the nested sample quantile $\widetilde{Y}_{\lceil \alpha L \rceil}$ under a set of regularity conditions. Let f and F denote the density and cumulative distribution function (CDF) of Y. The asymptotic bias and variance of $\widetilde{Y}_{\lceil \alpha L \rceil}$ can be expressed as

$$\mathrm{E}\left[\widetilde{Y}_{\lceil \alpha L \rceil}\right] - y_\alpha = W_\alpha N^{-1} + o_N(1/N) + O_L(1/L) + o_N(1)O_L(1/L) \tag{2}$$

$$V\left[\widetilde{Y}_{\lceil \alpha L \rceil}\right] = C_\alpha (L+2)^{-1} + O_L(1/L^2) + O_N(1/\sqrt{N})O_L(1/L), \tag{3}$$

where $W_\alpha = \theta_\alpha / f(y_\alpha)$, $C_\alpha = \alpha(1-\alpha)[f(y_\alpha)]^{-2}$, $\theta_\alpha = -\Theta'(y_\alpha)$, $\Theta(u) = \frac{1}{2} f(u)\mathrm{E}[\sigma_\xi^2 | Y(\xi) = u]$, σ_ξ^2 denotes the conditional variance of $\sqrt{n}(Y(\xi) - Y(\xi))$ (conditioned on ξ), $a_N = O_N(b_N)$ if $\lim_{N\to\infty} |a_N/b_N|$ is bounded, and $a_N = o_N(b_N)$ if $\lim_{N\to\infty} |a_N/b_N| \to 0$.

The asymptotic results in Eqs. (2) and (3) serve as the basis for the budget allocation under the criterion of minimizing the asymptotic MSE. Suppose that the average simulation effort to generate an outer-level sample ξ_i and an inner-level sample X_{ij} is γ_0 and γ_1, respectively; thus, the total simulation budget is $\Gamma = L(N\gamma_1 + \gamma_0)$. Without loss of generality, we adopt the convention that $\gamma_0 = 0$ and $\gamma_1 = 1$, as γ_0 is significantly smaller than γ_1 by an order of magnitude in many practical applications. Therefore, throughout the paper for ease of presentation, we assume that the overall simulation budget is set to $\Gamma = NL$. The MSE of the nested sample quantile $\widetilde{Y}_{\lceil \alpha L \rceil}$ is then verified to achieve the convergence rate of order $\Gamma^{-2/3}$ when the inner-level sample size N and the outer-level

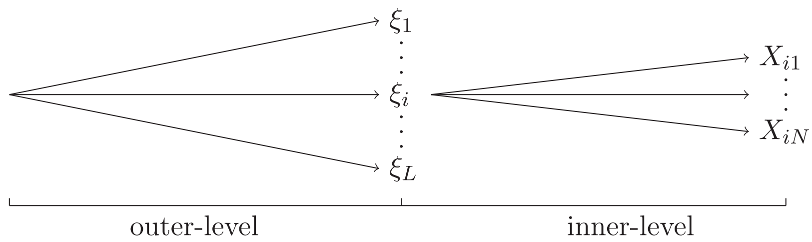

Fig. 2. Illustration of the Nested Simulation.

sample size L are chosen to be

$$N^* = \left(\frac{2W_\alpha^2}{C_\alpha}\right)^{1/3}\Gamma^{1/3} + o_\Gamma(\Gamma^{1/3}),$$
$$L^* = \left(\frac{C_\alpha}{2W_\alpha^2}\right)^{1/3}\Gamma^{2/3} + o_\Gamma(\Gamma^{2/3}), \tag{4}$$

respectively (See Section 3.2 in Gordy & Juneja, 2010).

3. KQE for nested simulation

KQE is a popular class of L-estimators that improve the efficiency of sample quantiles by taking a weighted average of all order statistics. If the portfolio loss of each scenario $\{Y(\xi_1),\dots,Y(\xi_L)\}$ is known, and letting $Y_{[1]} \geq \cdots \geq Y_{[L]}$ be the corresponding order statistics, then KQE for portfolio VaR can be formulated as

$$KQ_\alpha = \sum_{i=1}^{L}\left[\int_{(i-1)/L}^{i/L} k_h(t-\alpha)\ dt\right]Y_{[i]}, \tag{5}$$

where $k_h(\cdot) = k(\cdot/h)/h$, k is a kernel function that is a density function symmetric about 0 on $\mathbb{R}^1$, and h is a bandwidth parameter from $h \to 0$ to $L \to \infty$. The bandwidth h has long been considered crucial to the implementation of KQE, as there is considerable empirical evidence to show that different bandwidths may lead to substantially different MSEs (Sheather & Marron, 1990).

Many kernels can be used in the calculation of KQE (5). A simple example is the uniform kernel $k(x) = \mathbf{1}\{-\frac{1}{2} \leq x \leq \frac{1}{2}\}, x \in \mathbb{R}^1$ (Hong et al., 2017). For more kernels, such as Gaussian, tri-cube, and Epanechnikov kernels, see Chapter 6 in Hastie, Tibshirani, & Friedman (2009). Note that the choice of the kernel function k is not critical (see, e.g., Section 4.5 of Härdle, 1990). Conversely, it can be easily verified that KQE assigns greater weights to those close to the sample $L\cdot\alpha$th quantile. In fact, if we let k be the uniform kernel, then by the mean value theorem,

$$\int_{(i-1)/L}^{i/L} k_h(t-\alpha)\ dt = \mathbf{1}\left\{-\frac{1}{2} \leq \eta \leq \frac{1}{2}\right\}\cdot\frac{1}{hL},$$

where $\eta \in \left[\left(\frac{i-1}{L}-\alpha\right)/h, \left(\frac{i}{L}-\alpha\right)/h\right]$. Therefore, as i/L is close to α, η is close to 0, and so $\mathbf{1}\left\{-\frac{1}{2} \leq \eta \leq \frac{1}{2}\right\}$ is close to its maximum. However, it can be seen that the total weights $\sum_{i=1}^{L}\left[\int_{(i-1)/L}^{i/L} k_h(t-\alpha)\ dt\right] = \int_{-\alpha/h}^{(1-\alpha)/h}\mathbf{1}\left\{-\frac{1}{2} \leq x \leq \frac{1}{2}\right\}dx = 1$ if h is close to 0.

Recall that f and F denote the density and CDF of Y, respectively. Let $G = F^{-1}$ be the inverse CDF of Y and G' and G'' denote its first and second derivatives. Sheather & Marron (1990) established the asymptotic bias and variance of KQ_α. Under certain assumptions, for all fixed significance levels $\alpha \in (0,1)$, aside from $\alpha = 0.5$ when f is symmetric, we have

$$\begin{aligned}\mathrm{E}(KQ_\alpha) - y_\alpha =& A_\alpha h^2 + o(h^2) + O(L^{-1}),\\ V(KQ_\alpha) =& C_\alpha L^{-1} + D_\alpha L^{-1}h + o(L^{-1}h),\end{aligned}$$

where $K^{(-1)}$ is the antiderivative of k, $A_\alpha = \frac{1}{2}G''(1-\alpha)\int_{-\infty}^{+\infty}u^2k(u)du$, $C_\alpha = \alpha(1-\alpha)[f(y_\alpha)]^{-2}$ and $D_\alpha = -[f(y_\alpha)]^{-2}\int_{-\infty}^{+\infty}uk(u)k^{(-1)}(u)\,du$ are constants depending only on α. By minimizing the asymptotic MSE with regard to h, we obtain the following asymptotically optimal bandwidth

$$h_{\text{opt}} = c\cdot L^{-1/3},$$

where

$$c = \left(\frac{-D_\alpha}{4A_\alpha^2}\right)^{1/3} = \left[2\int_{-\infty}^{\infty}uk(u)K^{(-1)}(u)\,du\cdot\left[\int_{-\infty}^{\infty}u^2k(u)\,du\right]^{-2}\right]^{1/3} \cdot\left[\frac{G'(1-\alpha)}{G''(1-\alpha)}\right]^{2/3}. \tag{6}$$

Attempts to calculate the optimal bandwidth h_{opt} reveal that the unknown parameter c is infinitely more complicated to estimate than the y_α of interest. For example, c involves the first and second derivatives of G, which may be more difficult to estimate than $G(1-\alpha)$. Moreover, sophisticated integrals of the kernel function k are required to calculate c. This has motivated researchers to investigate how to select appropriate bandwidths for KQ_α (Cheng & Sun, 2006; Sheather & Marron, 1990). In this paper, we also focus on proposing an efficient algorithm to select the bandwidth.

As in Section 1, the analytical expression of portfolio loss $Y(\xi)$ is typically unknown; therefore, nested simulation estimators $\widetilde{Y}(\xi_1),\dots,\widetilde{Y}(\xi_L)$ have to be constructed (see (1)) for $Y(\xi)$ in each scenario. Let $\widetilde{Y}_{[1]} \geq \cdots \geq \widetilde{Y}_{[L]}$ be the corresponding order statistics. Suppose that k is an mth kernel, which is a density fulfilling $\int_{-\infty}^{\infty}u^ik(u)\,du = 0$ for $i = 1,\dots,m$. Then, the *nested KQE* for portfolio VaR is given by

$$\widetilde{KQ}_\alpha = \sum_{i=1}^{L}\left[\int_{(i-1)/L}^{i/L} k_h(t-\alpha)\,dt\right]\widetilde{Y}_{[i]}. \tag{7}$$

Note that to calculate the nested KQE, we must determine three parameters: the bandwidth h, the outer-level sample size L, and the inner-level sample size N. The principle is to minimize the asymptotic MSE of the nested KQE. Therefore, in the next section, we provide the asymptotic properties of the nested KQE, including its asymptotic bias, variance, and MSE. We demonstrate that compared with the nested sample quantile, the variance of the nested KQE is lower, and under certain conditions, so is its bias.

4. Theoretical results

We need fundamental assumptions to obtain the theoretical results of the nested KQE. Let $\bar{Z}^N = Y - \widetilde{Y}$ and $\widetilde{Z}_N = \bar{Z}^N\sqrt{N}$ such that $\widetilde{Z}_N$ has a nontrivial limiting distribution as $N \to \infty$. Our asymptotic analysis is based on the following assumption, which is also provided in Gordy & Juneja (2010).

Assumption 1. *(i)* The joint density $g_N(\cdot,\cdot)$ of Y and $\widetilde{Z}_N$ and its partial derivatives

$$\frac{\partial^i}{\partial y^i}\, g_N(y,z), \quad i=1,2,3,$$

exist for each N and for all (y,z).

(ii) For $N\geq 1$, there exist nonnegative functions $p_{0,N}(\cdot)$ and $p_{i,N}(\cdot)$ $(i=1,2,3)$, such that

$$g_N(y,z)\leq p_{0,N}(z), \qquad \left|\frac{\partial^i}{\partial y^i}\, g_N(y,z)\right| \leq p_{i,N}(z),$$

for all y,z. In addition,

$$\sup_N \int_{-\infty}^{+\infty} |z|^r p_{i,N}(z)\,dz < \infty \tag{8}$$

for $i=0,1,2,3$ and $0\leq r\leq 4$.

The rationality of Assumption 1 was illustrated by Gordy & Juneja (2010). Similar assumptions have also been widely used in the field of nested simulation (Hong et al., 2017; Zhang et al., 2022a; Zhang et al., 2022b). Assumption 1 concerns the joint density function of $(Y,\widetilde{Z}_N)$, which ensures that the higher-order terms of its Taylor series expansion can be ignored.

4.1. Asymptotic bias of the nested KQE

The asymptotic bias of the nested KQE is summarized in the following theorem.

Theorem 1. *Suppose that k is a first-order kernel (i.e., $\int_{-\infty}^{+\infty} uk(u)du=0$) and that Assumption 1 holds. Then, for all fixed $\alpha\in(0,1)$, apart from $\alpha=0.5$ when f is symmetric, we have*

$$\begin{aligned}\mathrm{E}\left[\widetilde{KQ}_\alpha\right]-y_\alpha &= A_\alpha h^2+W_\alpha N^{-1}+Q_\alpha h^3+H_\alpha h^2N^{-1}+J_\alpha N^{-2}\\ &\quad + o_h(h^3)+O_L(1/L)+o_N(1/N^2)+O_N(1/N)o_h(h^2),\end{aligned} \tag{9}$$

where $A_\alpha=\frac{1}{2}G''(1-\alpha)\int_{-\infty}^{+\infty}u^2k(u)du$, $W_\alpha=\theta_\alpha/f(y_\alpha)$, $Q_\alpha=\frac{1}{6}G^{(3)}(1-\alpha)\int_{-\infty}^{+\infty}u^3k(u)du$, $H_\alpha=\frac{1}{2}[\int_{-\infty}^{+\infty}u^2k(u)\,du]W''_\alpha$ and $J_\alpha=\frac{1}{6}\frac{d^2}{du^2}f(u)\mathrm{E}[Z_1^3\,|\,Y=y_\alpha]/f(y_\alpha)$ are constants, depending only on α.

To obtain the asymptotic bias (9), we followed a three-step process. First, we expand $\mathrm{P}(\widetilde{Y}>u)-\mathrm{P}(Y>u)$ into higher-order terms, as shown in Lemma 2 in the Appendix A. Based on this, we next analyze the weighted average of order statistics, and finally prove the converging limit of the weighted average. Notably, since the order statistics are with respect to nested simulation losses $\widetilde{Y}(\xi_1),\dots,\widetilde{Y}(\xi_L)$, the proof of Theorem 1 are novel in both literatures of nested simulation and KQE. For a more detailed analysis, please see Appendix A.

Theorem 1 provides the asymptotic bias of nested KQE, and for k being a general m-order kernel (*i.e.,* $\int_{-\infty}^{+\infty}u^ik(u)du=0$ for $i=1,\dots,m$), we discuss the corresponding result in Appendix B. Theorem 1 demonstrates that the bias of $\widetilde{KQ}_\alpha$ consists of two dominant terms: $A_\alpha h^2$ and $W_\alpha N^{-1}$. We summarize some of the related useful discussions here, containing an investigation of these two components, as well as the possibility of reducing bias.

(i) $W_\alpha N^{-1}$. This term is exactly equal to the dominant term of the bias of the nested sample quantile $\widetilde{Y}_{\lceil\alpha L\rceil}$, which is established in Eq. (2). The sign of W_α depends on that of θ_α. Gordy & Juneja (2010) found that under mild conditions, there exists u^* such that $-\Theta'(u)>0$ for all $u<u^*$. That is, the bias term $W_\alpha N^{-1}$ tends to be positive (i.e. $\theta_\alpha>0$) for some large VaR values y_α that will most likely happen when the extremely small α is of interest in risk management.

(ii) $A_\alpha h^2$. This term originates from KQE itself. The bandwidth h controls the kernel weight assigned to the order statistic $\widetilde{Y}_{[i]}$ and is a crucial parameter in kernel smoothing methods. The coefficient A_α is either positive or negative, which depends on that of $G''(1-\alpha)$.

(iii) *Bias reduction.* We now put the two dominant terms together and turn to a brief discussion of bias. Their signs are the key to reduce bias. If $A_\alpha\cdot W_\alpha<0$, then the two dominant terms will disappear by taking $h^2=-W_\alpha A_\alpha^{-1}N^{-1}$, leading to a much faster convergence rate of the bias of the nested KQE. In contrast, if $A_\alpha\cdot W_\alpha>0$, we simply choose $h=o_N(N^{-1/2})$, and $A_\alpha h^2$ turns out to be an ignorant term with respect to $W_\alpha N^{-1}$, so the convergence rate of the asymptotic bias of the nested KQE is equal to that of the sample quantile for nested simulation.

4.2. Asymptotic variance of the nested KQE

The asymptotic variance of the nested KQE is expressed as follows in Theorem 2.

Theorem 2. *Let k be a compactly supported kernel and $k^{(-1)}$ be the antiderivative of k. Under Assumption 1, for all fixed $\alpha\in(0,1)$, apart from $\alpha=0.5$ when f is symmetric, we have*

$$V[\widetilde{KQ}_\alpha]=C_\alpha L^{-1}+D_\alpha L^{-1}h+O_L(L^{-1})o_h(h)+O_L(L^{-1})O_N(1/\sqrt{N}), \tag{10}$$

where $C_\alpha=\alpha(1-\alpha)[f(y_\alpha)]^{-2}$ and $D_\alpha=-[f(y_\alpha)]^{-2}\int_{-\infty}^{+\infty}uk(u)k^{(-1)}(u)\,du$ are constants depending only on α.

The proof of Theorem 2 is based on the asymptotic variance of the KQE in Sheather & Marron (1990). However, since the nested KQE is a weighted average of nested sample quantiles (i.e., order statistics with respect to nested simulation losses), it is necessary to carefully examine the variance of the nested KQE by combining the variance of the KQE and the asymptotic properties of nested sample quantiles. The detailed proof is provided in Appendix A.

The asymptotic result given in Theorem 2 is obtained under the compactly supported kernel assumption. Sheather & Marron (1990) illustrated that the same conclusion can be shown to hold for normal and other reasonable infinite-support positive kernels, using a straightforward but tedious truncation argument. Therefore, we omit the study of infinite-support kernels. Also, we discuss the corresponding result in Appendix B, if k is a m-order $(m\geq 1)$ kernel. We now turn to a brief discussion of the conclusion and applicable condition of Theorem 2.

(i) The dominant term in the asymptotic variance of the nested KQE is $C_\alpha L^{-1}$, which is the same as that of the nested sample quantile (see Eq. (3)). However, $D_\alpha<0$ is in (10), so the asymptotic variance of the nested KQE reduces by $D_\alpha L^{-1}h$ compared with that of the nested sample quantile. Despite that, if $h\to 0$, $D_\alpha L^{-1}h$ is a smaller-order term of $C_\alpha L^{-1}$ and thus the improvement only appears in the lower-order terms (this phenomenon is called deficiency, see e.g., Sheather & Marron, 1990). We emphasize that with finite samples in practice, variance reduction yields a significant effect.

(ii) We focus on the case where f is not symmetric or where f is symmetric but $\alpha\neq 0.5$, which will cause no loss of generality in the scientific field of VaR estimation.

4.3. Asymptotic MSE of the nested KQE

Based on the asymptotic bias and variance presented in Theorems 1 and 2, respectively, we demonstrate that the convergence rate of the MSE of the nested KQE is always at least as fast as that of the nested sample quantile. As analyzed in Section 4.1, the convergence rate of the bias of the nested KQE depends on two cases: $A_\alpha \cdot W_\alpha > 0$ and $A_\alpha \cdot W_\alpha < 0$. Therefore, we also explore the convergence rate of the MSE of the nested KQE in these two cases.

4.3.1. The case where $A_\alpha \cdot W_\alpha > 0$

In Section 4.1, we demonstrate that when $A_\alpha \cdot W_\alpha > 0$,

$$h_1^* = o_N(N^{-1/2}), \tag{11}$$

is an appropriate choice for minimizing bias. In this case, we just have to determine the values of L and N given a budget Γ. Note that the dominant term of the MSE is $W_\alpha^2 N^{-2} + C_\alpha L^{-1}$. To minimize this, we consider the following optimization problem:

$$\min_{L,N} \; W_\alpha^2 \, N^{-2} + C_\alpha L^{-1}$$

subject to $LN = \Gamma$.

It is easy to obtain the solution, i.e., the budget allocation.

$$N_1^* = \left(\frac{2W_\alpha^2}{C_\alpha}\right)^{1/3} \Gamma^{1/3} + o_\Gamma(\Gamma^{1/3}), \quad L_1^* = \left(\frac{C_\alpha}{2W_\alpha^2}\right)^{1/3} \Gamma^{2/3} + o_\Gamma(\Gamma^{2/3}), \tag{12}$$

which are the best choices for L and N. In this case, the asymptotic MSE of the nested KQE is of order $\Gamma^{-\frac{2}{3}}$ and is the same as that of the nested sample quantile.

4.3.2. The case where $A_\alpha \cdot W_\alpha < 0$

In this case, minimizing the two dominant terms of the bias of the nested KQE (Eq. (9)) with regard to h leads to the following optimal bandwidth choice:

$$h_2^* = \left[-W_\alpha A_\alpha^{-1} N^{-1}\right]^{1/2}, \tag{13}$$

which consequently leads to the following higher-order asymptotic bias:

$$\mathrm{E}\left[\widetilde{KQ}_\alpha\right] - y_\alpha = E_\alpha N^{-3/2} + o_h(h^3) + O_L(1/L) + o_N(1/N^{3/2}) + O_N(1/N) o_h(h^2), \tag{14}$$

where $E_\alpha = Q_\alpha\left[-W_\alpha A_\alpha^{-1}\right]^{3/2}$.

When combined with Eq. (10), the dominant terms of the asymptotic MSE are $E_\alpha N^{-3/2} + C_\alpha L^{-1}$. To minimize the MSE, we consider the following optimization problem:

$$\min_{L,N} \; E_\alpha N^{-3/2} + C_\alpha L^{-1}$$

subject to $LN = \Gamma$,

which shows that the optimal outer-level sample size N_2^* and inner-level sample size L_2^* are given by

$$N_2^* = \left(\frac{C_\alpha}{3E_\alpha^2}\right)^{-\frac{1}{4}} \Gamma^{\frac{1}{4}} + o_\Gamma(\Gamma^{\frac{1}{4}}), \quad L_2^* = \left(\frac{C_\alpha}{3E_\alpha^2}\right)^{\frac{1}{4}} \Gamma^{\frac{3}{4}} + o_\Gamma(\Gamma^{\frac{3}{4}}). \tag{15}$$

With the optimal bandwidth and budget allocation appropriately selected under the condition $A_\alpha \cdot W_\alpha < 0$, we obtain the asymptotic MSE of the nested KQE of order $\Gamma^{-\frac{3}{4}}$, which is faster than that of the nested sample quantile.

Remark 1. Comparing the analyses of nested sample quantiles and nested KQEs, the latter requires to identify the bandwidth and budget allocation in cases of $A_\alpha \cdot W_\alpha < 0$ and $A_\alpha \cdot W_\alpha > 0$. This leads to that it not only has a smaller variance but also has the potential to exhibit smaller bias. In other words, the nested KQEs possess this advantage that neither the nested sample quantile nor the non-nested KQE does.

5. Implementation

To calculate the nested KQE with the optimal MSE, we first have to determine the optimal bandwidth h_i^* and the budget allocation rule N_i^* and L_i^* $(i = 1, 2)$. By their definitions (see (12), (13) and (15)), the key issue is to estimate the parameters A_α, W_α, C_α, and E_α.

5.1. Estimating the parameter A_α

Following Zhang et al. (2022b) we use a bootstrap-based method to estimate the parameter A_α. Eq. (9) shows that given N and L, the bias of the nested KQE is a linear function of h^2, and A_α is the slope. Therefore, if we can construct a rough estimator of this bias for the given N and L and a different h's, then regressing this bias on h^2 leads to an estimator of A_α. However, "bias in particular is notoriously difficult to estimate" (Broadie et al., 2011a), so a rough estimator of this bias may require a large simulation budget, which would be prohibitive. To overcome this issue, we use the bootstrap method, a popular resampling technique in statistics to estimate bias.

Suppose that the initial samples of $\{(\xi_l, X_{li}), l = 1, \ldots, L_0; i = 1, \ldots, N_0\}$ are generated, where L_0 and N_0 are relatively small ($L_0 \cdot N_0 \ll \Gamma$), leading to loss estimators $\widetilde{\boldsymbol{Y}}_{\boldsymbol{L_0}} = \{\widetilde{Y}_1, \ldots, \widetilde{Y}_{L_0}\}$. Let $\{\widetilde{Y}_1^*, \ldots, \widetilde{Y}_{L_0}^*\}$ be the bootstrap sample from $\widetilde{\boldsymbol{Y}}_{\boldsymbol{L_0}}$ referring to the sample obtained by independently and randomly drawing replacement L_0 times from $\widetilde{\boldsymbol{Y}}_{\boldsymbol{L_0}}$. Let $\{\widetilde{Y}_{[1]}^*, \ldots, \widetilde{Y}_{[L_0]}^*\}$ be the corresponding order statistics. For any $h_m, m = 1, \ldots, M$, the nested KQE based on the bootstrap sample is

$$\widetilde{KQ}_\alpha^{(1)}(h_m) = \sum_{i=1}^{L_0}\left[\int_{(i-1)/L_0}^{i/L_0} k_{h_m}(t - \alpha)\, dt\right] \widetilde{Y}_{[i]}^*.$$

By repeating the bootstrap procedure B times, we obtain B estimators $\left\{\widetilde{KQ}_\alpha^{(2,b)}(h_m),\ b = 1, \ldots, B\right\}$, which immediately yield the bootstrap mean as

$$\widetilde{KQ}_\alpha^{(1)}(h_m) = \frac{1}{B}\sum_{b=1}^{B} \widetilde{KQ}_\alpha^{(1,b)}(h_m).$$

We regress the bootstrap mean $\left(\widetilde{KQ}_\alpha^{(1)}(h_1), \ldots, \widetilde{KQ}_\alpha^{(1)}(h_M)\right)^T$ on $(h_1^2, \ldots, h_M^2)^T$, and the estimated slope of the regression $\hat{A}_\alpha$ is an appropriate approximation of A_α.

During implementation, the initial sample sizes N_0 and L_0 are usually set to be small. With a small N_0 and L_0, the implementation of the bootstrap procedure is usually fast, and the corresponding computation time is often negligible compared with the overall simulation time of the samples of (ξ, X).

5.2. Estimating the parameters W_α and C_α

The results of Eqs. (9) and (10) demonstrate that given h, the bias and variance are linear functions of $1/N$ and $1/L$ and W_α and C_α are the slopes, respectively. Therefore, to estimate the parameters W_α and C_α, we apply the same method as in the previous subsection to estimate the bias and variance.

Given the initial samples $\{(\xi_l, X_{li}), l = 1, \ldots, L_0; i = 1, \ldots, N_0\}$, let $\{\xi_1^*, \ldots, \xi_{L'}^*\}$ be the bootstrap sample from $\{\xi_1, \ldots, \xi_{L_0}\}$ with size L'. Conditional on a bootstrap observation ξ_l^*, we can similarly obtain a bootstrap sample with size N' from the inner-level samples corresponding to ξ_l^*, denoted by $\{X_{l1}^{**}, \ldots, X_{lN'}^{**}\}$. Let $\widetilde{Y}_l^{**}(N') = \frac{1}{N'}\sum_{i=1}^{N'} X_{li}^{**}$ and $\widetilde{Y}_{[1]}^{**}(N') \geq \cdots \geq \widetilde{Y}_{[L']}^{**}(N')$ denote the corresponding order statistics. Furthermore, let $h(L') = [\alpha(1-\alpha)(L'+1)]^{-1/2}$, then the nested KQE based on the bootstrap sample is

$$\widetilde{KQ}_{\alpha}^{(2)}(L',N') = \sum_{i=1}^{L'}\left[\int_{(i-1)/L'}^{i/L'} k_{h(L')}(t-\alpha)\,dt\right]\widetilde{Y}_{[i]}^{**}(N').$$

By repeating the bootstrap procedure B times, we obtain B estimators $\{\widetilde{KQ}_{\alpha}^{(2,b)}(L',N'),\ b=1,\dots,B\}$, which immediately give the bootstrap mean and variance as

$$\widetilde{KQ}_{\alpha}^{(2)}(L_0,N') = \frac{1}{B}\sum_{b=1}^{B}\widetilde{KQ}_{\alpha}^{(2,b)}(L_0,N') \tag{16}$$

and

$$\widetilde{s}_{\alpha}^{(2)}(L',N_0) = \frac{1}{B}\sum_{b=1}^{B}\left[\widetilde{KQ}_{\alpha}^{(2,b)}(L',N_0) - \frac{1}{B}\sum_{b=1}^{B}\widetilde{KQ}_{\alpha}^{(2,b)}(L',N_0)\right]^2. \tag{17}$$

We choose one set of K_1 integer values $N_1 < N_2 < \cdots < N_{K_1} = N_0$ and the other set of K_2 integer values $L_1 < L_2 < \cdots < L_{K_2} = L_0$, regress the bootstrap mean and variance vector $\left(\widetilde{KQ}_{\alpha}^{(2)}(L_0,N_1), \widetilde{KQ}_{\alpha}^{(2)}(L_0,N_2),\dots,\widetilde{KQ}_{\alpha}^{(2)}(L_0,N_{K_1})\right)^T$ and $\left(\widetilde{s}_{\alpha}^{(2)}(L_1,N_0),\dots,\widetilde{s}_{\alpha}^{(2)}(L_{K_2},N_0)\right)^T$ on $(1/N_1,\dots,1/N_{K_1})^T$ and $(1/L_1,\dots,1/L_{K_2})^T$, respectively, and set the slopes of the regression, $\hat{B}_\alpha$ and $\hat{C}_\alpha$, as the estimators of W_α and C_α.

5.3. Estimating the parameter E_α

If $\hat{A}_\alpha \cdot \hat{B}_\alpha < 0$, then the asymptotic bias is given in Eq. (14), so the parameter E_α should be estimated. Similarly, given the initial samples, let $\widetilde{Y}_l^{*}(N')$ and $\widetilde{Y}_{[l]}^{*}(N')$ be the quantities defined as in Section 5.2. Taking the bandwidth as $h(N') = \left[-\hat{B}_\alpha\hat{A}_\alpha^{-1}N'^{-1}\right]^{1/2}$, the nested KQE-based bootstrap sample is

$$\widetilde{KQ}_{\alpha}^{(3)}(L_0,N') = \sum_{i=1}^{L_0}\left[\int_{(i-1)/L_0}^{i/L_0} k_{h(N')}(t-\alpha)\,dt\right]\widetilde{Y}_{[i]}^{*}(N'),$$

By repeating this procedure B times, we obtain B estimators $\{\widetilde{KQ}_{\alpha}^{(3,b)}(L_0,N'),\ b=1,\dots,B\}$, which immediately yield the bootstrap mean as

$$\widetilde{KQ}_{\alpha}^{(3)}(L_0,N') = \frac{1}{B}\sum_{b=1}^{B}\widetilde{KQ}_{\alpha}^{(3,b)}(L_0,N').$$

For $N_1 < N_2 < \cdots < N_{K_1} = N_0$, we regress the bootstrap mean vector $\left(\widetilde{KQ}_{\alpha}^{(3)}(L_0,N_1),\dots,\ \widetilde{KQ}_{\alpha}^{(3)}(L_0,N_{N_{K_1}})\right)^T$ on $\left(1/N_1^{3/2},\dots,1/N_{K_1}^{3/2}\right)^T$, and set the slope of the regression $\hat{E}_\alpha$ to be the estimator of E_α.

5.4. Calculation of the nested KQE

After obtaining the estimators of A_α, W_α, C_α, and E_α, we establish the nested KQE for portfolio VaR. As discussed in Section 4.3, if $\hat{A}_\alpha \cdot \hat{B}_\alpha > 0$, we set $\hat{h}_1 = [\alpha(1-\alpha)(L+1)]^{-1/2}$ and the budget allocation is

$$\hat{N}_1 = (2\hat{B}_\alpha^2/\hat{C}_\alpha)\Gamma^{1/3} \text{ and } \hat{L}_1 = (\hat{C}_\alpha/(2\hat{B}_\alpha^2))\Gamma^{2/3}. \tag{18}$$

Based on Eq. (18), we generate the loss estimators $\widetilde{Y}_i, i=1,\dots,\hat{L}_1$ and compute the nested KQE as follows:

$$\widetilde{KQ}_{\alpha,1} = \sum_{i=1}^{\hat{L}_1}\left[\int_{(i-1)/\hat{L}_1}^{i/\hat{L}_1} k_{\hat{h}_1}(t-\alpha)\,dt\right]\widetilde{Y}_{[i]}.$$

If $\hat{A}_\alpha \cdot \hat{B}_\alpha < 0$, then we set the budget allocation to be

$$\hat{N}_2 = \left(\hat{C}_\alpha/3\hat{E}_\alpha^2\right)^{-\frac{1}{4}}\Gamma^{\frac{1}{4}} \text{ and } \hat{L}_2 = \left(\hat{C}_\alpha/3\hat{E}_\alpha^2\right)^{\frac{1}{4}}\Gamma^{\frac{3}{4}}, \tag{19}$$

and the bandwidth $\hat{h}_2 = \left[-\hat{B}_\alpha\hat{A}_\alpha^{-1}\hat{N}_2^{-1}\right]^{1/2}$. Based on Eq. (19), we generate the loss estimators $\widetilde{Y}_i, i=1,\dots,\hat{L}_2$ and compute the nested KQE as follows:

$$\widetilde{KQ}_{\alpha,2} = \sum_{i=1}^{\hat{L}_2}\left[\int_{(i-1)/\hat{L}_2}^{i/\hat{L}_2} k_{\hat{h}_2}(t-\alpha)\,dt\right]\widetilde{Y}_{[i]}.$$

Having fully discussed the computation of nested KQE, for the ease of presentation, we provides the implementation procedure in the following Algorithm 1.

Algorithm 1: Nested KQEs.

Input: Γ, K_1 integer values $N_1 < N_2 < \cdots < N_{K_1} = N_0$, K_2 integer values $L_1 < L_2 < \cdots < L_{K_2} = L_0$, M bandwidths $h_1,\dots,h_M$;Generate initial samples $\{(\xi_l, X_{li}), l=1,\dots,L_0; i=1,\dots,N_0\}$.

Step 1. Estimate the parameters A_α, W_α, C_α and E_α;Regress $\left(\widetilde{KQ}_{\alpha}^{(1)}(h_1),\dots,\widetilde{KQ}_{\alpha}^{(1)}(h_M)\right)^T$ on $(h_1^2,\dots,h_M^2)^T$;Regress $\left(\widetilde{KQ}_{\alpha}^{(2)}(L_0,N_1),\dots,\widetilde{KQ}_{\alpha}^{(2)}(L_0,N_{K_1})\right)^T$ on $(\mathbf{1}; 1/N_1,\dots,1/N_{K_1})^T$;Regress $\left(\widetilde{s}_{\alpha}^{(2)}(L_1,N_0),\dots,\widetilde{s}_{\alpha}^{(2)}(L_{K_2},N_0)\right)^T$ on $(1/L_1,\dots,1/L_{K_2})^T$;If $\hat{A}_\alpha \cdot \hat{B}_\alpha < 0$, then regress $\left(\widetilde{KQ}_{\alpha}^{(3)}(L_0,N_1),\dots,\widetilde{KQ}_{\alpha}^{(3)}(L_0,N_{N_{K_1}})\right)^T$ on $\left(1/N_1^{3/2},\dots,1/N_{K_1}^{3/2}\right)^T$;Set the slopes of the regression $\hat{A}_\alpha$, $\hat{B}_\alpha$, $\hat{C}_\alpha$ and $\hat{E}_\alpha$, as the estimators of A_α, W_α, C_α and E_α, respectively;

Step 2. Compute nested KQE;**if** $\hat{A}_\alpha \cdot \hat{B}_\alpha > 0$ **then**
  $\hat{N}_1 = (2\hat{B}_\alpha^2/\hat{C}_\alpha)\Gamma^{1/3}$, $\hat{L}_1 = (\hat{C}_\alpha/(2\hat{B}_\alpha^2))\Gamma^{2/3}$,
  $\hat{h}_1 = [\alpha(1-\alpha)(\hat{L}_1+1)]^{-1/2}$;Compute $\widetilde{KQ}_{\alpha,1}$;
else if $\hat{A}_\alpha \cdot \hat{B}_\alpha < 0$ **then**
  $\hat{N}_2 = \left(\hat{C}_\alpha/3\hat{E}_\alpha^2\right)^{-\frac{1}{4}}\Gamma^{\frac{1}{4}}$, $\hat{L}_2 = \left(\hat{C}_\alpha/3\hat{E}_\alpha^2\right)^{\frac{1}{4}}\Gamma^{\frac{3}{4}}$,
  $\hat{h}_2 = \left[-\hat{B}_\alpha\hat{A}_\alpha^{-1}\hat{L}_2^{-1}\right]^{1/2}$.Compute $\widetilde{KQ}_{\alpha,2}$;
end

Output: The nested KQE is $\widetilde{KQ}_{\alpha,1}$ or $\widetilde{KQ}_{\alpha,2}$;

6. Numerical experiments

In this section, we present the computational experiments to validate our theoretical results as well as explore the effectiveness of KQE for nested portfolio VaR estimation. We first construct a stylized example in Section 6.1, to ensure that the constants in bias, variance and bandwidth can be calculated.

KQE equipped with the theoretically optimal bandwidth h_2^* defined as in (13), and the widely used VaR estimator sample quantile are compared over the proposed data set to justify the asymptotic results of KQE for nested simulation. A risk measurement study is carried out in Section 6.2 to further justify the effectiveness of the nested KQE for portfolio VaR estimation.

6.1. A stylized example

We validate our theoretical results with the following parametric example. Distributions for loss Y and the inner step pric-

Table 1
Performance of KQEs v.s. sample quantiles for fixed budget allocations.

Fixed Budget ($L \times N$)	Bias		Variance		RMSE	
	$\widetilde{KQ}_\alpha(h_2^*)$	$\widetilde{Y}_{\lceil \alpha L \rceil}$	$\widetilde{KQ}_\alpha(h_2^*)$	$\widetilde{Y}_{\lceil \alpha L \rceil}$	$\widetilde{KQ}_\alpha(h_2^*)$	$\widetilde{Y}_{\lceil \alpha L \rceil}$
10×20	**−0.202**	0.342	**0.953**	2.065	**0.997**	1.477
10×30	**−0.148**	0.221	**0.922**	1.570	**0.971**	1.272
50×20	**−0.039**	0.314	**0.202**	0.337	**0.451**	0.660
100×20	**−0.027**	0.213	**0.100**	0.144	**0.317**	0.435

* The results with better performance are marked in **bold**.

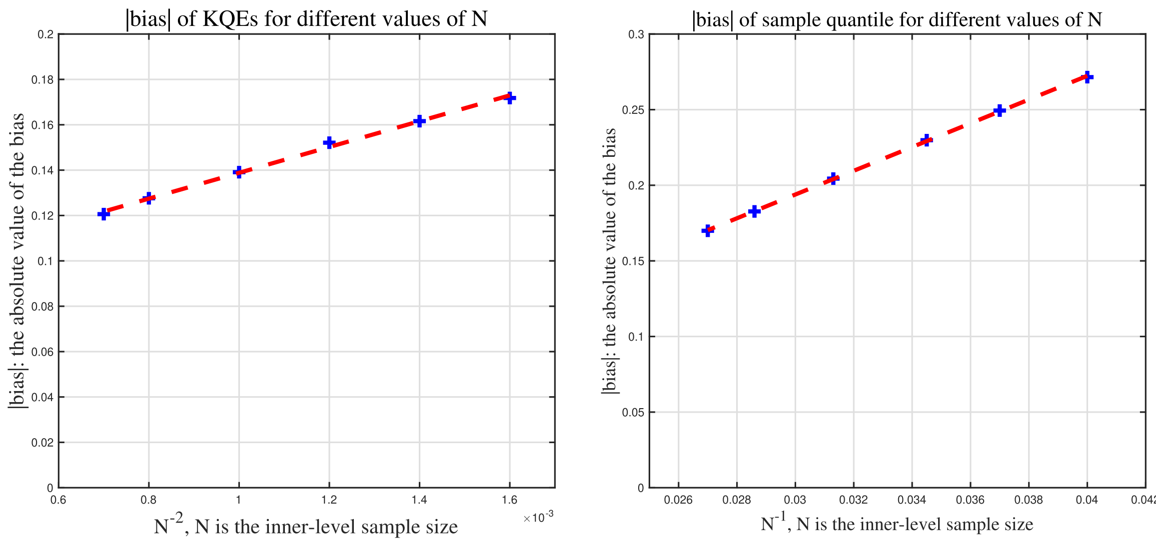

Fig. 3. Abstract biases of KQEs and sample quantiles for different N's when $L = 10$.

ing errors are specified to ensure that the bias and variance of our simulation estimators are in closed-form. Therefore, it is feasible to calculate the budget allocation directly without employing the sample-driven Algorithm 1. Note that although the example is highly stylized, it allows us to compare the nested KQE with the sample quantile under different budget allocations.

Let the exact distribution for portfolio loss $Y(\xi) = \xi$ be $N(0,1)$, and the distribution of $\widetilde{Z}_N$ conditional on Y be $N\big(0, \exp(Y)\big)$. The loss estimator is $\widetilde{Y} = Y + \widetilde{Z}_N/\sqrt{N}$. We consider the VaR estimation at the significance level $\alpha = 0.05$. In this example, k is set to be the Gaussian kernel. According to Theorems 1 and 2, it can be calculated that $A_\alpha = 77.3$ and $W_\alpha = -4.8$, which indicates that the following analysis falls within the case of $A_\alpha \cdot W_\alpha < 0$.

We compare the performance of the most frequently used sample quantile and our KQE equipped with the theoretically optimal bandwidth $h_2^* = [-W_\alpha A_\alpha^{-1} N^{-1}]^{1/2}$. We carry out 100,000 runs for each setting, and utilize the RMSE as the final criteria. Table 1 summarizes the biases, variances and RMSEs of KQEs and sample quantiles for different setting. It can be seen from Table 1 that KQEs with the theoretically optimal bandwidth give VaR estimators with smaller RMSEs, which demonstrates the effectiveness of KQE for nested simulation. We note that the results are robust for different budget allocations $L \times N$. In particular, the biases of KQEs turn out to be smaller than those of sample quantiles. This indicates the potential bias reduction of KQE within the case $A_\alpha \cdot W_\alpha < 0$. Besides, $\widetilde{KQ}_\alpha(h_2^*)$ also gives smaller variances than sample quantile, justifying the variance reduction by averaging individual order statistics with finite samples in practice.

Fig. 3 verifies the convergence rate of bias of KQE when $A_\alpha \cdot W_\alpha < 0$. Note that when k is the Gaussian kernel, we have $Q_\alpha = 0$ and the theoretical bias of KQE shows up as $O(N^{-2})$. Therefore, we plots the abstract value of the estimated bias with respect to N^{-2} and N^{-1} for KQE and sample quantile, respectively, with $L = 10$. It is shown in the figure that the abstract bias of KQEs equipped with the theoretically optimal bandwidth h_2^* is linear with respect to N^{-2} while that of sample quantile is linear to N^{-1}, confirming the theoretical conclusions verified in Section 4.3.2 of this paper.

6.2. Portfolio VaR estimation

In this subsection, we conduct a risk measurement experiment to provide the comparison between the KQE and the sample quantile in order to examine the performances of the proposed algorithm. More specifically, we consider a portfolio that consists of options written on d stocks, and use the risk measure VaR to quantify its risk. Before proceeding, we assume that all stock prices follow the Black-Scholes model. For simplicity, assume that stock returns are the same, denoted by μ, whereas risk-free interest rate is r. Price dynamics of the stocks $\boldsymbol{S}_t = (S_t^1, \ldots, S_t^d)^T \in \mathrm{R}^d$ follow the multidimensional geometric Brownian motion, that is

$$S_t^k = \mu' S_t^k \, dt + \sum_{l=1}^{q} \sigma_{kl} S_t^k \, dB_t^l, \quad k = 1, \ldots, d, \tag{20}$$

where μ' is chosen to be μ under the real-world probability measure, whereas it is chosen to be r under the risk neutral probability measure. Here $\boldsymbol{B}_t = (B_t^1, \ldots, B_t^q)$ is a standard q-dimensional Brownian motion, and without loss of generality we let $\Sigma = (\sigma_{kl})$ be a subtriangular matrix.

Proceeding further in this direction, we assume that the options in the portfolios have the common maturity date, denoted by T, and we want to measure the portfolio risk at a future time t_τ $(t_\tau < T)$. During implementation, we work with a discretized version of $\boldsymbol{S}_t$ valued at a sequence of time points $0 = t_0 < t_1 < \cdots < t_n = T$. For notational simplicity, we write $\boldsymbol{S}_{t_\nu}$ as $\boldsymbol{S}_\nu$ for $\nu = 0, 1, \ldots, n$, and without loss of generality assume that τ takes value in $\{1, \ldots, n-1\}$. The portfolio consists of three geometric Asian

Table 2
Performance of VaR estimates for different α and Γ.

Quantile Level	VaR	Budget	Bias		Variance		RMSE		AT (s)	
			$\widetilde{KQ}_\alpha$	$\widetilde{Y}_{\lceil \alpha L\rceil}$	$\widetilde{KQ}_\alpha$	$\widetilde{Y}_{\lceil \alpha L\rceil}$	$\widetilde{KQ}_\alpha$	$\widetilde{Y}_{\lceil \alpha L\rceil}$	$\widetilde{KQ}_\alpha$	$\widetilde{Y}_{\lceil \alpha L\rceil}$
$\alpha = 0.1$	$y_\alpha = 42.5$	$\Gamma = 10^5$	**1.42**	1.53	**3.76**	5.08	**2.41**	2.72	62.0	49.4
		$\Gamma = 10^6$	**0.72**	0.88	**0.89**	1.90	**1.19**	1.63	239.0	202.8
$\alpha = 0.05$	$y_\alpha = 53.7$	$\Gamma = 10^5$	**1.52**	1.94	**3.35**	6.77	**2.38**	3.24	63.2	50.4
		$\Gamma = 10^6$	**0.85**	1.01	**0.96**	1.60	**1.29**	1.62	243.6	205.7
$\alpha = 0.01$	$y_\alpha = 73.6$	$\Gamma = 10^5$	**2.35**	2.39	**13.77**	17.21	**4.39**	4.79	61.7	48.7
		$\Gamma = 10^6$	**1.04**	1.25	**1.76**	3.12	**1.69**	2.16	251.0	216.3

* The results with better performance are marked in **bold**.

call options and three European call options with three strikes K_1, K_2 and K_3, each of which is written on 10 assets, that is, $d = 10$ in Eq. (20). To fit into our framework, the random loss of the portfolio can be written as

$$Y(\xi) = \mathrm{E}\big[X|\xi\big],$$

where

$$X = V_0 - \sum_{k=1}^{10}\sum_{l=1}^{3} e^{-r(T-t_\tau)}\left(S_n^k - K_l\right) - \sum_{k=1}^{10}\sum_{l=1}^{3} e^{-r(T-t_\tau)}\left(\prod_{\nu=1}^{n} S_{\nu,i}^k - K_l\right)^+,$$

$$\xi = [S_\tau^k, \prod_{\nu=1}^{\tau} S_\nu^k],\ k = 1, \ldots, 10,$$

and V_0 is the current value of the portfolio and is a known constant.

Parameters about the portfolios are set as follows: $(S_0^1, \ldots, S_0^d) = 100$, T is 1/12 year, $t_\tau = 1/52$, $\mu = 8\%$, $r = 5\%$, $\sigma = 20\%$, $K_1 = 90$, $K_2 = 100$, and $K_3 = 110$. To measure the performance of the estimators based on our allocation rule, we need the true value of y_α as a benchmark. Note that in this example, the analytical expression of $Y(\xi)$ can be derived using the Black-Scholes formula. We generate 10^9 samples of ξ, and calculate the corresponding values of $Y(\xi)$, which can be used to accurately approximate y_α. We then use this accurate estimate as a benchmark to examine the performance of our proposed nested KQE and the conventional nested sample quantile. In order to ensure both the accuracy and the stability of this experimental result, we carry out 1,000 replications, and utilize the RMSE as the final criteria.

In the experiments, we set the total simulation budget to be 10^5 and 10^6, and adopt the Gaussian kernel. When $\Gamma = 10^5$, we set $N_0 = L_0 = 100$ and $(N_1, \ldots, N_{K_1}) = (L_1, \ldots, L_{K_2}) = (50, 55, 60, \ldots, 100)$, and when $\Gamma = 10^6$, set $N_0 = L_0 = 200$ and $(N_1, \ldots, N_{K_1}) = (L_1, \ldots, L_{K_2}) = (50, 60, \ldots, 200)$. We carry out 1,000 runs for each parameter setting, and the number bootstrap sampling within the sample-driven budget allocation algorithm is set to be $B = 200$.

Numerical results including bias, variance, RMSE, and AT (average time) required to perform one replication of VaR estimates are summarized in Table 2, from which we have the following observations.

(i) On the whole, compared with the conventional nested sample quantile, our proposed nested KQE gives a VaR estimator with smaller variances. This suggests that substituting the single order statistic into an appropriate average of all order statistics lead to a more stable VaR approximation within the nested simulation framework. We note that this result is robust across several values of α, and for different cases of total budget Γ.

(ii) A generally accepted point of view is that accurate quantile estimates are easier to obtain if α is close to 0.5. It can be seen from Table 2 that as α approaches 0, namely from 0.1 to 0.01, RMSE of nested KQEs and nested sample quantiles appear bigger and bigger when the budget is as large as 10^6. In fact, both bias and variance of two estimates show a monotonically increasing trend.

(iii) From Table 2, we can see that biases of KQE and sample quantile are comparable, indicating that $A_\alpha \cdot W_\alpha > 0$ in this example, and this coincides with the discussion in Section 4.3.1. More precisely, however, we find that the biases of KQE are all a little smaller than those of sample quantile. This may be because given a finite simulation budget, when we trade off the bias and variance, the estimator with smaller variance encourages to allocate more samples to the inner level in order to reduce the bias.

(iv) Because KQE has smaller biases and variances than sample quantile, its RMSE must be smaller, as shown in Table 2. This again demonstrates that our proposed nested KQE outperforms nested sample quantile in the estimation of portfolio VaR.

(v) The average computation times of both methods are displayed in the last two columns of Table 2. As shown, nested KQEs take slightly longer to compute than nested sample quantiles. This is mainly due to the fact that KQEs require the calculation of the integral of kernel functions and the weighted average of all order statistics (see the definition of nested KQE (7)), in contrast to individual order statistics used in sample quantiles. This additional computation leads to more training time being required in exchange for higher precision.

In addition, to verify the accuracy of the nested KQE compared with nested sample quantile, we construct the 95% confidence intervals for biases by 1000 replications. The detailed results are presented in Appendix C.

7. Conclusion

We propose the KQE in nested simulation framework for estimating portfolio VaR, and investigate its theoretical properties, including the asymptotic bias and variance. We show that the variance is reduced in the lower-order terms, while in some cases bias could be reduced in the dominant term, and the convergence rate of MSE is established in a case-by-case way. Also, we establish a bootstrap-based algorithm to estimate the unknown parameters in nested KQE, so that this estimator can be implemented in practice. Numerical results show that nested KQE works quite well for portfolio VaR estimation with different quantile levels and different budgets.

The KQE can also be combined with previous literatures about parametric or nonparametric approaches to achieve better performance. For example, when the portfolio loss estimators are obtained by such approaches, substituting them into the KQE for portfolio VaR estimation would lead to variance reducation. Furthermore, based on this, the corresponding CVaR estimator may be more efficient. This remains an open area for future research.

Acknowledgments

This research was supported by National Natural Science Foundation of China (NNSFC) grants 72101260, 62206300 and Public Computing Cloud, Renmin University of China.

Declaration of generative AI and AI-assisted technologies in the writing process.

During the preparation of this work the author(s) used ChatGPT in order to improve language and readability. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ejor.2023.07.040.

References

Al-Kenani, A., & Yu, K. (2012). New bandwidth selection for kernel quantile estimators. *Journal of Probability and Statistics*, Article 138450. 18.
Broadie, M., Du, Y., & Moallemi, C. C. (2011). Efficient risk estimation via nested sequential simulation. *Management Science, 57*(6), 1172–1194.
Broadie, M., Du, Y., & Moallemi, C. C. (2011). Risk estimation via weighted regression. In *Proceedings of the 2011 winter simulation conference (WSC)* (pp. 3854–3865).
Broadie, M., Du, Y., & Moallemi, C. C. (2015). Risk estimation via regression. *Operations Research, 63*(5), 1077–1097.
Cheng, H. F., Liu, X., & Zhang, K. (2022). Constructing confidence intervals for nested simulation. *Naval Research Logistics, 69*, 1138–1149.
Cheng, H. F., & Zhang, K. (2021). Non-nested estimators for the central moments of a conditional expectation and their convergence properties. *Operations Research Letters, 49*(5), 625–632.
Cheng, M. Y., Sun, S., et al., (2006). Bandwidth selection for kernel quantile estimation. *Journal of the Chinese Statistical Association, 44*, 271–295.
Falk, M. (1984). Relative deficiency of kernel type estimators of quantiles. *The Annals of Statistics, 12*, 261–268.
Fort, G., Gobet, E., & Moulines, E. (2017). MCMC design-based non-parametric regression for rare event. application to nested risk computations. *Monte Carlo Methods and Applications, 23*(1), 21–42.
Goda, T. (2017). Computing the variance of a conditional expectation via non-nested monte carlo. *Operations Research Letters, 45*(1), 63–67.
Gordy, M. B., & Juneja, S. (2010). Nested simulation in portfolio risk measurement. *Management Science, 56*, 1833–1848.
Härdle, W. (1990). *Applied nonparametric regression*. Cambridge university press.
Hastie, T., Tibshirani, R., & Friedman, J. H. (2009). *The elements of statistical learning: Data mining, inference, and prediction*. New York: Springer. 2
Hong, L. J., & Juneja, S. (2009). Estimating the mean of a non-linear function of conditional expectation. In *Proceedings of the 2009 winter simulation conference (WSC)* (pp. 1223–1236).
Hong, L. J., Juneja, S., & Liu, G. (2017). Kernel smoothing for nested estimation with application to portfolio risk measurement. *Operations Research, 65*(3), 657–673.
Lan, H., Nelson, B. L., & Staum, J. (2015). A confidence interval procedure for expected shortfall risk measurement via two-level simulation. *Operations Research, 58*(5), 1481–1490.
Lee, S. H. (1998). *Monte Carlo computation of conditional expectation quantiles*. Stanford University Ph.D. thesis.
Liu, G., & Staum, J. (2010). Stochastic kriging for efficient nested simulation of expected shortfall. *Journal of Risk, 12*(3), 3–27.
Liu, M., Nelson, B. L., & Staum, J. (2010). An efficient simulation procedure for point estimation of expected shortfall. In *Proceedings of the 2010 winter simulation conference (WSC)* (pp. 2821–2831).
McCune, E. D., & McCune, S. L. (1991). Jackknifed kernel quantile estimators. *Communications in Statistics - Theory and Methods, 20*, 2719–2725.
Navruz, G., & Özdemir, A. F. (2020). A new quantile estimator with weights based on a subsampling approach. *British Journal of Mathematical and Statistical Psychology, 73*, 506–521.
Parzen, E. (1979). Nonparametric statistical data modeling. *Journal of the American Statistical Association, 74*, 105–131.
Sheather, S. J., & Marron, J. S. (1990). Kernel quantile estimators. *Journal of the American Statistical Association, 85*, 410–416.
Sun, Y., Apley, D. W., & Staum, J. (2011). Efficient nested simulation for estimating the variance of a conditional expectation. *Operations Research, 59*(4), 998–1007.
Wang, T., Xu, J., Hu, J.-Q., & Chen, C. H. (2022). Efficient estimation of a risk measure requiring two-stage simulation optimization. *European Journal of Operational Research, 305*(3), 1355–1365.
Wang, W., Wang, Y., & Zhang, X. (2022). Smooth nested simulation: Bridging cubic and square root convergence rates in high dimensions. arXiv preprint arXiv:2201.02958.
Xie, W., Nelson, B. L., & Barton, R. R. (2014). A bayesian framework for quantifying uncertainty in stochastic simulation. *Operations Research, 62*(6), 1439–1452.
Zhang, K., Feng, B. M., Liu, G., & Wang, S. (2022). Sample recycling for nested simulation with application in portfolio risk measurement. arXiv preprint arXiv:2203.15929.
Zhang, K., Liu, G., & Wang, S. (2022). Bootstrap-based budget allocation for nested simulation. *Operations Research, 70*(2), 1128–1142.
Zhu, H., Liu, T., & Zhou, E. (2020). Risk quantification in stochastic simulation under input uncertainty. *ACM Transactions on Modeling and Computer Simulation (TOMACS), 30*(1), 1–24.