Statistica Sinica **31** (2021), 547-569
doi:https://doi.org/10.5705/ss.202018.0304

NON-PARAMETRIC ESTIMATION OF CONDITIONAL TAIL EXPECTATION FOR LONG-HORIZON RETURNS

Hwai-Chung Ho[1,2], Hung-Yin Chen[3] and Henghsiu Tsai[1]

[1]*Academia Sinica*, [2]*National Taiwan University, and* [3]*Chung Yuan Christian University*

Abstract: When evaluating the tail risk of stock portfolio returns, providing statistically sound solutions for long return horizons is important, but difficult. Furthermore, there are drawbacks to using traditional parametric methods that rely on strong model assumptions or simulations. This study investigates the problem by focusing on an important risk measure, the conditional tail expectation (CTE), under a general multivariate stochastic volatility model. To overcome the estimation difficulties caused by the long period, we derive an asymptotic formula to approximate the CTE. Based on this formula, we propose a simple nonparametric estimate of the unconditional CTE, and show that it is both consistent and asymptotically normal. Next, we forecast the CTE using a modified form of the nonparametric estimator. With the help of the asymptotic formula, we evaluate the accuracy of the CTE predictor by treating it as an interval forecast for furure returns. Simulation studies demonstrate the applicability of our approach. Lastly, we apply the proposed estimation and predictor to daily S&P 500 index returns.

Key words and phrases: Asymptotic normality, conditional tail expectation, integrated process, interval forecast, long-horizon returns, stochastic volatility model.

1. Introduction

Quantifying extreme, high-impact events that have a low probability of occurring is becoming increasingly important. Today, many studies use the functionals derived from the distribution quantile of a random variable to achieve this goal. For example, the quantile itself, is used to define one of the most popular risk measures, namely, the value at risk (VaR). Motivated by the need to assess the market risk of long-term investments, we examine inferences on a less researched quantile functional under a framework of long-horizon returns. We focus on a risk measure called the conditional tail expectation (CTE), which is

Corresponding author: Henghsiu Tsai, Institute of Statistical Science, Academia Sinica, Taipei, Taiwan, R.O.C. E-mail: htsai@stat.sinica.edu.tw.

defined as the expected value of the loss, conditional on the loss being greater than a given quantile or VaR. This measure is also referred to as the conditional VaR (Du and Escanciano (2015)), tail VaR (Artzner et al. (1999)), or expected shortfall (Hardy (2003, p.158)).

Artzner et al. (1999) proposed the CTE to address the incoherence inherent in the VaR. Since the Great Recession of 2008, regulators have become increasingly concerned about tail events that create severe adverse efferts and, thus, are now more inclined to apply excessive prudence in captial reserve requirements (Asimit et al. (2011)). As a more conservative risk measure than the VaR, the CTE is receiving increasing attention from practitioners and academics, and has replaced the VaR in the regulatory requirements of Canada, Israel, and Switzerland (Asimit et al. (2011)). Furthermore, the Basel Committee has explicitly discussed the prospect of phasing out the VaR and replacing it with the CTE (Basel Committee on Banking Supervision (2012)).

The body of research on long-horizon returns is extensive in the literature on financial economics, but tends to focus on the forecastability of future returns; here, recent works include those of Boudoukh, Richardson and Whitelaw (2008); Neuberger (2012); Rapach and Zhou (2013); Fama and French (2018). In contrast, we address this issue from the perspective of risk management. Institutional investors, such as insurance companies, pension funds, and sovereign funds, usually hold their portfolios for a long time, possibly several years, depending on the nature of the funds. Moreover, long-term equity-linked financial instruments exist, such as the Long-term Equity Anticipation Securities issued by the Chicago Board Options Exchange (CBOE), with maturity from one to five years. Evaluating the risk exposure of such contracts, requires an accurate estimate of a tail risk measure (e.g., the CTE) of their returns.

In general, there are two approaches to estimating the CTE and, thus, assessing the capital reserve required for a long-term equity portfolio or financial product. The first comprises approaches in which both the marginal distribution and the dependence structures of the return sequence are fully speified, shcuh as the Gaussian framework (Lai and Xing (2008, Chap.12)). As a result, a closed-form solution of the estimator and its limiting distribution can be derived. However, these approaches have been criticized for their overly restrictive assumptions. The second type of approach includes two estimation procedures, both of which rely on generating a sample. The first procedure employs subsampling (Politis, Romano and Wolf (1999)) to estimate the CTE of the accumulated returns for a given long horizon, and then makes inferences using the bootstrap distribution based on these estimates. The second procedure imposes a dynamic parametric model, such as the regime-switching lognormal model on the daily returns. Then, it estimates the tail risk (i.e., the CTE, in this study) from a large number of accumulated returns simulated by the model, with parameters estimated from historical observations (Hardy (2003); Hardy, Freeland and Till (2006)). By repeating the procedure to produce sufficiently many estimates, we can derive their distribution and, thus, perform inferences. These two sample-generation methods exhibit significant estimation bias, especially when the return horizon is long; see the finite-sample comparative analysis presented in Section S3 of the online Supplementary Material.

To avoid the shortcomings of the aforementioned methods, we adopt a new approach in which we treat the long-horizon portfolio returns as an integrated process of daily returns that follows a general multivariate stochastic volatility (GMSV) model. Then, we derive an asymptotic formula to approximate the long-term CTE, with the integration length tending to infinity. Based on this formula, we propose a nonparametric estimate that is easy to implement. Lastly, we verify that the proposed estimate is consistent and asymptotically normal. The GMSV model assumed for the daily portfolio returns generalizes several stochastic volatility (SV) models popular in the literature (Taylor (1986); Hamilton (1994); Hardy (2001)). The advantage of this model generality lies in the model-free nature of our approach, because we do not need to estimate the underlying dynamic SV system. As a result, our estimator is free of the constraints of nonidentifiability and is less vulnerable to model misspecification. The asymptotic normality of the normalized estimate is instrumental in rigorous inferences on the CTE, and helps to verify that, for accumulated returns of middle to long horizons, the coverage ratio of our estimate is significantly more accurate than that of the simulation-based method. As discussed part (a) of Remark 2 in Subsection 3.1 and demonstrated in Subsection 4.1, the proposed estimate works very well in finite-sample analyses, even when the sample size is not greater than the return horizon. The asymptotic formula we derive for the CTE of the integrated process, which forms the basis of our estimator, is itself a result of independent interest.

Given the wide application of forecasting tail risk, we modify the nonparametric estimator to predict future CTE. Here, we treat the CTE predictor as an interval forecast for future returns. As such, we test the accuracy of the predictor using a t-test based on the asymptotic formula.

The rest of the paper is organized as follows. In Section 2, we introduce a general multivariate version of the SV process to model the daily returns of a portfolio's component stocks. Subsection 3.1 discusses the method proposed to estimate the CTE; this includes our main theoretical findings, stated as Proposition 1, where we establish the asymptotic normality and consistency of the estimate. Our approach has advantages related to both estimation and prediction. Although the α-level CTE of the long-horizon returns diverges as the horizon increases, the rule that governs the asymptotic behavior turns out to have an analytic form. This enables us to construct a consistent and asymptotically normal estimate. The unconditional confidence intervals based on the asymptotic normality can then be used to perform inferences on the CTE of interest. In Subsection 3.2, we consider dynamic predictions of the conditional CTE. Section 4 presents a finite-sample analysis of our approach, and verifies the accuracy of the proposed predictor using a formal test. Section 5 provides empirical results of our CTE estimates and predictions for long-horizon S&P 500 returns, and Section 6 concludes the paper. The proof of Proposition 1 is provided in Section S2 of the online Supplementary Material, and Section S3 shows that the coverage ratio of the proposed method is superior to those of two traditional sample-generation methods.

2. A GMSV Model

For a security portfolio comprising m component assets, we consider the following GMSV model for its return vector $\tilde{r}_t = (r_{1,t}, \ldots, r_{m,t})'$ at time t:

$$\tilde{r}_t = \tilde{\mu} + V_t U_t, \tag{2.1}$$

where $\tilde{\mu} = (\mu_1, \ldots, \mu_m)'$ is the mean vector of $\tilde{r}_t$, $V_t = \mathrm{diag}(v_{1,t}, \ldots, v_{m,t})$, with the diagonal matrix representing the volatility component, and $U_t = (u_{1,t}, \ldots, u_{m,t})'$ is a sequence of independent and identically distributed (i.i.d.) shocks with mean zero and a positive-definite covariance matrix $\Sigma_U = [\rho_{U,ij}]$. Each component $v_{i,t}$ of V_t is defined by

$$v_{i,t} = h_i(Z_{i,t}), \quad i = 1, \ldots, m, \tag{2.2}$$

where $h_i(\cdot)$ is a positive functional satisfying certain regularity conditions (specified in Section **S1** of the online Supplementary Material; see **Remark 1**), and $\tilde{Z}_t = (Z_{1,t}, \ldots, Z_{m,t})'$ is an m-dimensional stationary linear process, defined as

$$\tilde{Z}_t = \tilde{\mu}_z + \sum_{s=0}^{\infty} A_s \eta_{t-s}, \tag{2.3}$$

where $\tilde{\mu}_z = (\mu_{z,1}, \ldots, \mu_{z,m})'$, $A_s = [A_{ij}^{(s)}]_{i,j=1}^m$, and $\eta_t = (\eta_{1,t}, \ldots, \eta_{m,t})'$ is a sequence of i.i.d. zero-mean random vectors with a positive-definite covariance matrix $\Sigma_\eta = [\rho_{\eta,ij}]$, and is independent of $\{U_t\}$. In (2.3), the component random variable $Z_{i,t}$ of $\tilde{Z}_t$ is $Z_{i,t} = \mu_{z,i} + \sum_{s=0}^{\infty} \tilde{A}_i^{(s)} \star \eta_{t-s}$, for $1 \le i \le m$, where $\tilde{A}_i^{(s)} = (A_{i,1}^{(s)}, \ldots, A_{i,m}^{(s)})$ is the ith row of A_s, and $\star$ denotes the inner product. We assume that $\{\tilde{Z}_t\}$ is short memory in the sense that $\{A_s\}_{s=0}^{\infty}$ is absolutely summable; that is, $\sum_{s=0}^{\infty} \left| A_{ij}^{(s)} \right| < \infty$, for $i, j = 1, 2, \ldots, m$ (see, e.g., Chapter 10 of Hamilton (1994) for further details). Several multivariate SV models have been considered in the literature, including those of Harvey, Ruiz and Shephard (1994); Robinson (2001); Asai, McAleer and Yu (2006); Yu and Meyer (2006); Ho, Chen and Tsai (2016), among others. In particular, Robinson (2001) considers an SV model with general transformations in the volatility component, but requires that $\tilde{Z}_t$ be Gaussian. Note that the models mentioned avove are not nested.

3. Goals and Main Findings

When assessing the quantitative risk of a portfolio, it is common to focus on the weighted returns of all component assets. Suppose the weight allocated to each asset during the investment horizon is fixed. Then, the weighted return of the portfolio is given by $r_t = \sum_{i=1}^m w_i r_{i,t} = \mu + \sum_{i=1}^m w_i v_{i,t} u_{i,t}$, with mean $\mu = \sum_{i=1}^m w_i \mu_i$ and variance $\sigma^2 = E(r_t - \mu)^2$. The weights $\{w_i : 1 \le i \le m\}$ satisfy $w_i > 0$ and $\sum_{i=1}^m w_i = 1$.

For a fixed T, we consider the stationary sequence $\{R_{T,s} : s = 1, 2, \ldots\}$ of integrated returns for horizon T (i.e., $R_{T,s} = \sum_{j=s}^{s+T-1} r_j$, for $s = 1, 2, \ldots$), and denote the distribution of $R_{T,s}$ by $F_T(\cdot)$ (assumed continuous). Let $q_\alpha(T)$ be the αth quantile of $F_T(\cdot)$; that is, $\alpha = F_T(q_\alpha(T))$. Then, the αth CTE of F_T is defined as

$$C_\alpha^T = \alpha^{-1} \int_{-\infty}^{q_\alpha(T)} x dF_T(x). \tag{3.1}$$

If we set $R_T = R_{T,1}$, then (3.1) is equivalent to

$$C_\alpha^T = E\left(R_T \,|R_T < q_\alpha(T)\right), \tag{3.2}$$

which is also commonly used to define the CTE.

Given a sample $\{r_t : t = 1, 2, \ldots, n\}$ of the portfolio's daily returns, our goal is to estimate the CTE of $F_T(\cdot)$ for a small α and large T. For the CTE of the right tail, equation (3.2) becomes $C_\alpha^T = E\left(R_T \,|R_T > q_\alpha(T)\right)$; however, we do not consider this here, because our only concern is the loss. Instead of T being fixed, we allow T to increase to infinity. Our approach is particularly suitable for returns of middle to long horizons; Section S3 of the Spplementary Material demonstrates that the finite-sample performance of the proposed approach is superior to that of two traditional methods.

Let Z_α be the αth CTE of a standard normal Z with distribution function $\Phi(\cdot)$; that is, $Z_\alpha = E(Z|Z < \Phi^{-1}(\alpha)) = -\phi(\Phi^{-1}(\alpha))/\alpha$. One of our main findings is that C_α^T is asymptotically

$$C_\alpha^T = T\mu + \sqrt{T}\sigma Z_\alpha + O\left(\frac{1}{\sqrt{T}}\right), \tag{3.3}$$

as shown in (S2.19) in the online Supplementary Material. Equation (3.3) leads to a natural nonparametric estimate

$$\widehat{C}_\alpha^T = T\mu^* + \sqrt{T}\hat{\sigma}_n Z_\alpha, \tag{3.4}$$

for C_α^T, where $\hat{\sigma}_n = (n^{-1}\sum_{t=1}^n (r_t - \mu^*)^2)^{1/2}$, with $\mu^* = \mu$ if μ is known, and $\mu^* = \hat{\mu}_n = n^{-1}\sum_{t=1}^n r_t$ otherwise.

In the next two subsections, we show how to use the nonparametric estimate (3.4) to make conditional and unconditional inferences. For the latter, we derive the asymptotic normality of $\widehat{C}_\alpha^T$, which we use to determine confidence intervals for the location of the true value of C_α^T from the full sample. For the former, we use (3.4), with adaptive estimates of μ and σ, conditional on the returns $\{r_1, r_2, \ldots, r_{t-1}\}$, to dynamically predict the CTE of future returns $R_{T,t} = \sum_{s=t}^{t+T-1} r_s$ for horizon T.

3.1. Unconditional confidence intervals

Because C_α^T diverges as T increases, it is important to know whether $\widehat{C}_\alpha^T$ is close to C_α^T when both n and T are large. To this end, in Proposition 1, we establish the asymptotic normality and consistency of $\widehat{C}_\alpha^T$, assuming certain conditions on the sample size n and the integration length T are satisfied. The proposition relies on applying the central limit theorem to $\hat{\mu}_n = n^{-1}\sum_{t=1}^n r_t$ and $\hat{\sigma}_n^2 = n^{-1}\sum_{t=1}^n (r_t - \hat{\mu}_n)^2$; both are given in the first lemma in the online Supplementary Material.

Remark 1. To apply the central limit theorem to $\hat{\sigma}_n^2 = n^{-1}\sum_{t=1}^n (r_t - \hat{\mu}_n)^2$, certain technical conditions are required on the functions h_i given in (2.2); see the online Supplementary Material. These conditions are satisfied by many commonly used functions, including the exponential function, absolute value function, and positive polynomials.

In the following, we use $a_n \ll b_n$ to denote $a_n = o(b_n)$.

Proposition 1. *Assume the GMSV model specified in* (2.1)*,* (2.2)*, and* (2.3) *satisfies* $Ev_{i,1}^4 < \infty$ *for* $1 \le i \le m$*, and that* $\{U_t\}$ *is a sequence of i.i.d.* m*-variate normal vectors that is independent of* $\{V_t\}$ *and has mean zero and covariance matrix* Σ_U*. Let* $N = n/T$*, where* n *and* T *denote the sample size and the length of integration, respectively.*

(i) μ *is unknown: If* N *and* T *are such that* $T \ll N \ll T^2$*, then*

$$\sqrt{\frac{N}{T}}(\widehat{C}_\alpha^T - C_\alpha^T) \xrightarrow{d} N(0, \sigma^2) \tag{3.5}$$

as T *tends to infinity.*

(ii) μ *is known: Under condition* J*, given in the online Supplementary Material, if* N *and* T *are such that* $N \ll T$*, then*

$$\sqrt{N}(\widehat{C}_\alpha^T - C_\alpha^T) \xrightarrow{d} N\left(0, \left(\frac{gZ_\alpha}{2\sigma}\right)^2\right) \tag{3.6}$$

as N *tends to infinity;* g^2 *is defined in Lemma* 1 *in the Supplementary Material.*

Remark 2. (a) An interesting aspect of the asymptotic behavior of $\widehat{C}_\alpha^T - C_\alpha^T$ is the case in which the number of blocks N is a constant that does not increase with T or n. For example, consider part (i) of Proposition 1. With only T tending to infinity, we have

$$\sqrt{\frac{N}{T}}\left(\widehat{C}_\alpha^T - C_\alpha^T\right) = \sqrt{n}\,(\hat{\mu}_n - \mu) + O_p\left(\frac{1}{\sqrt{T}}\right), \tag{3.7}$$

implying that $\sqrt{N/T}\left(\widehat{C}_\alpha^T - C_\alpha^T\right)$ is approximately distributed as $N(0, \sigma^2)$, even for moderately large T. As a result, we can use $\widehat{C}_\alpha^T$ to infer the location of C_α^T, even though the estimation error $\widehat{C}_\alpha^T - C_\alpha^T$ does not vanish asymptotically. A similar analysis can be applied to part (ii) of Proposition 1. This type of asymptotic normality with a small N performs quite well in finite-sample analyses (see Table 3 of Subection 4.1). This is particularly useful in circumstances similar to that described in the example in Subsection 5.1. (b) Note that in Proposition 1, no parametric assumptions are imposed on the dynamic model of V_t or on the covariance matrix of the normal U_t. Because the estimator $\widehat{C}_\alpha^T$ proposed in (3.4) only needs to estimate the mean and the variance of the return distribution, it is free of the constraints of nonidentifiability, and is less vulnerable to model misspecification. (c) Ho, Chen and Tsai (2016) also studied the tail risk of integrated returns, but focused on the VaR, not the CTE. Other major distinctions between our study and that of Ho, Chen and Tsai (2016) are as follows. First, we obtain an approximation error of order $O(1/T)$ for the VaR, shown in Lemma 2 of the Supplementary Material, which serves as a preparatory step toward proving Proposition 1; this is much sharper than that of order $O(1/\sqrt{T})$ derived in Ho, Chen and Tsai (2016). Second, the SV model considered by Ho, Chen and Tsai (2016) is a special case of the GSVM introduced in Section 2. Third, Ho, Chen and Tsai (2016) based their asymptotic analysis on a setting in which the return horizon is required to be equal to the sample size. As such, they were unable to address the issue of consistency in their proposed VaR estimate.

The next proposition is based on Proposition 1, and provides confidence intervals for the diverging CTE, C_α^T, of the distribution $F_T(\cdot)$ of R_T.

Proposition 2. *(i) If μ is unknown, the $100\,(1-\beta)\,\%$ confidence interval for C_α^T is*

$$\widehat{C}_\alpha^T - \sqrt{\frac{T}{N}}\hat{\sigma}_n U \le C_\alpha^T \le \widehat{C}_\alpha^T - \sqrt{\frac{T}{N}}\hat{\sigma}_n L, \tag{3.8}$$

where L and U are the $100\,(\beta/2)\,\%$ and $100\,(1-\beta/2)\,\%$ quantiles, respectively, of $N\,(0,1)$.
(ii) If μ is known, the $100\,(1-\beta)\,\%$ confidence interval for C_α^T is

$$\widehat{C}_\alpha^T - \frac{hU}{\sqrt{N}} \le C_\alpha^T \le \widehat{C}_\alpha^T - \frac{hL}{\sqrt{N}}, \tag{3.9}$$

where $h = gZ_\alpha/(2\sigma)$.

The main task when creating confidence intervals that agree with (3.9) is to estimate the limiting standard deviation g of (3.6) contained in h. To achieve this, we follow a well-established resampling scheme, called the sampling window method, for dependent data; see Politis, Romano and Wolf (1999), and the references therein, for a comprehensive survey on the topic. The resampling method is given as follows. Divide the whole sample into $n-k+1$ subsamples, each of size k, denoted by $B_i = \left(r_i, \ldots, r_{i+k-1}\right)$. Here $k = \lambda n^{1/3}$, for some $\lambda \geq 1$. For the subsample B_i, let $\hat{\mu}_i = k^{-1} \sum_{t=i}^{i+k-1} r_t$ and $\hat{\sigma}_{k,i}^2 = \sum_{t=i}^{i+k-1} \left(r_t - \hat{\mu}_i\right)^2 / (k-1)$ be its sample mean and sample variance, respectively. Because g^2 is the limiting variance of $\sqrt{n}(\hat{\sigma}_n^2 - \sigma^2)$ (see (S2.4) and (S2.5) of Lemma 1 in the Supplementary Material),

$$\frac{1}{n-k+1} \sum_{i=1}^{n-k+1} \left\{ \sqrt{k}(\hat{\sigma}_{k,i}^2 - \hat{\sigma}_n^2) \right\}^2 \tag{3.10}$$

is a consistent estimate of g^2.

3.2. Conditional interval forecasts

To transform $\widehat{C}_\alpha^T$ in (3.4) into a dynamic predictor for the CTE of future returns, a natural method is to replace μ^* and $\hat{\sigma}_n$ in (3.4) with certain adaptive estimates. Focusing on the case of unknown μ, which is more realistic in practice, we use the equal-weighted moving averages,

$$\hat{\mu}_{t-1}^* = W_T^{-1} \sum_{s=t-W_T}^{t-1} r_s \quad \text{and} \quad \hat{\sigma}_{t-1}^* = \left(W_T^{-1} \sum_{s=t-W_T}^{t-1} (r_s - \hat{\mu}_{t-1}^*)^2 \right)^{1/2}, \tag{3.11}$$

and consider the conditional forecast

$$\widehat{C}_{\alpha,t}^T = T\hat{\mu}_{t-1}^* + \sqrt{T}\hat{\sigma}_{t-1}^* Z_\alpha \tag{3.12}$$

of the level-α CTE of the future return $R_{T,t} = \sum_{s=t}^{t+T-1} r_s$. If we treat $\widehat{C}_{\alpha,t}^T$ as an interval forecast $(-\infty,\ \widehat{C}_{\alpha,t}^T)$ for the future return $R_{T,t}$ of horizon T, then the following is a reasonable criterion that a good predictor $\widehat{C}_{\alpha,t}^T$ needs to satisfy: for each t,

$$\left| P(R_{T,t} < \widehat{C}_{\alpha,t}^T) - P(R_{T,t} < C_\alpha^T) \right| = o(1) \tag{3.13}$$

when T is large. To ensure (3.13), we assume the window length W_T diverges faster than T; that is, $T = o(W_T)$. The assumption $T = o(W_T)$, (3.3), and the central limit theorem togehter imply that, as $T \to \infty$,

$$P(R_{T,t} < \hat{C}_{\alpha,t}^T) = P\left(\frac{R_{T,t} - T\mu}{\sqrt{T}\sigma} < Z_\alpha + o_p(1) \right) \longrightarrow \Phi(Z_\alpha). \tag{3.14}$$

Then, (3.13) follows from

$$P(R_{T,t} < C_\alpha^T) = P\left(\frac{R_{T,t} - T\mu}{\sqrt{T}\sigma} < Z_\alpha + O\left(\frac{1}{\sqrt{T}}\right)\right) = \Phi(Z_\alpha) + O\left(\frac{1}{\sqrt{T}}\right) \quad (3.15)$$

according to (3.3) and the Berry–Esseen theorem. To evaluate the prediction accuracy of $\widehat{C}_{\alpha,t}^T$, we introduce the indicator function $I_t^T(\alpha) = I(R_{T,t} < C_\alpha^T)$, and its conditional version

$$\hat{I}_t^T(\alpha) = I(R_{T,t} < \hat{C}_{\alpha,t}^T), \quad (3.16)$$

based on the past returns $\{r_{t-W_T}, \ldots, r_{t-1}\}$. Rewrite (3.14) as

$$P(R_{T,t} < \widehat{C}_{\alpha,t}^T) \approx \Phi(Z_\alpha). \quad (3.17)$$

From (3.17), the accuracy of the forecast $\widehat{C}_{\alpha,t}^T$ can be evaluated by testing how close $P(R_{T,t} < \widehat{C}_{\alpha,t}^T)$ is to its stationary limit $\Phi(Z_\alpha)$. To achieve this, we employ a simple t-test that uses the sample average,

$$\hat{\pi}_n(\alpha) = \sum_{t=t_0+1}^{n-T} \frac{\hat{I}_t^T(\alpha)}{n^*},$$

as the test statistic, where n is the sample size and $n^* = n - T - t_0$ for some positive t_0. To derive the critical values using the correlated sequence $\{\hat{I}_t^T(\alpha)\}$ resulting from overlapping returns $R_{T,t}$, we adopt the subsample method (Politis, Romano and Wolf (1999, Chap. 3)). This is similar in spirit to the method we employed to find the limiting variance g^2 (cf., Equation (3.10)) for the case of a known mean. Denote by k the size of the subsamples, and by $\widetilde{B}_i = \{\hat{I}_t^T(\alpha), t = t_0 + i, \ldots, t_0 + i + k - 1\}$ the ith subsample of the conditional indicator functions defined in (3.16). For $i = 1, \ldots, n^* - k + 1$, let $\hat{\pi}_{i,k}(\alpha) = \sum_{t=t_0+i}^{t_0+i+k-1} \hat{I}_t^T(\alpha)/k$ be the average over the ith subsample $\widetilde{B}_i$. Then, the $1 - \beta$ acceptance region of a size β test for testing (3.17) is

$$(\hat{\pi}_n(\alpha) - U_{\beta/2,k}(n^*)^{-1/2},\ \hat{\pi}_n(\alpha) - L_{\beta/2,k}(n^*)^{-1/2}), \quad (3.18)$$

where $L_{\beta/2,k}$ and $U_{\beta/2,k}$ are the $(\beta/2)$th and $(1-\beta/2)$th quantiles, respectively, of $\{\sqrt{k}(\hat{\pi}_{i,k}(\alpha) - \hat{\pi}_n(\alpha)), i = 1, \ldots, n^* - k + 1\}$.

4. Finite-Sample Analysis

4.1. Unconditional coverage ratios

The numerical studies discussed in this subsection concentrate on the empirical coverage ratios of the confidence intervals (3.8) and (3.9). Because our

SV model for the return vectors $\tilde{r}_t = (r_{1,t}, \ldots, r_{m,t})'$ allows multivariate cases, we focus on $m = 2$ and $m = 10$.

The m-dimensional return vector $\tilde{r}_t = (r_{i,t}, \ldots, r_{m,t})'$ is modeled by $\tilde{r}_t = \tilde{\mu} + V_t U_t$, $\tilde{Z}_t = \Phi \tilde{Z}_{t-1} + \epsilon_t$, where $\tilde{\mu} = (0.0003, \ldots, 0.0003)'$, U_t and ϵ_t are independent m-dimensional multivariate normal vectors with zero mean, Φ is an $m \times m$ diagonal matrix with diagonal entries equal to ϕ, and the weight for each component return $r_{i,t}$ is $w_i = 1/m$. For $v_{i,t}$, we consider two cases: (i) $v_{i,t} = \bar{\sigma} \exp(Z_{i,t}/2)$, and (ii) $v_{i,t} = \bar{d}|Z_{i,t} + 2 \log \bar{\sigma}|$, both for $i = 1, \ldots, m$. We set $\bar{\sigma} = 0.0099$, $\bar{d} = 0.001115$, and $\phi = 0.5$. Here, we calibrate the values of $\bar{\sigma}$ and $\bar{d}$ such that the variances of $v_{i,t}$ in cases (i) and (ii) are close.

For the case of $m = 2$, let $U_t = (u_{1,t}, u_{2,t})'$ and $\epsilon_t = (\epsilon_{1,t}, \epsilon_{2,t})$, and denote by ρ_u and ρ_ϵ the correlations between the two marginals of U_t and ϵ_t, respectively. To include the null, positive, and negative correlations between the component random variables of U_t and ϵ_t, three combinations for (ρ_u, ρ_ϵ) are considered: (ρ_u, ρ_ϵ)=$(-0.5, 0.5)$, $(0.5, 0.5)$, and $(-0.5, -0.5)$. The covariance matrices of U_t and ϵ_t are $\Sigma_U = [\sigma_{U,ij}]$ and $\Sigma_\epsilon = [\sigma_{\epsilon,ij}]$, respectively, where $\sigma_{U,ij} = 1$ if $i = j$, and $\sigma_{U,ij} = \rho_u$ otherwise. Similarly, $\sigma_{\epsilon,ij} = \bar{c}$ if $i = j$, and $\sigma_{\epsilon,ij} = \bar{c}\rho_\epsilon$ otherwise, where $\bar{c} = \bar{\beta}^2(1 - \phi^2)$ with $\bar{\beta} = 0.4$.

For $m = 10$, we configure the correlations between the marginals of U_t and ϵ_t similarly to the case of $m = 2$, where we consider independence and positive and negative correlations. We choose the following five representative combinations to depict the correlation structures for U_t and ϵ_t. Specifically, the covariance matrix $\Sigma_U = [\Sigma_{U,ij}]$ of U_t is defined as follows: (i) if $i = j$, then $\Sigma_{U,ij} = 1$; (ii) if $i \neq j$, $i \wedge j = 1$, and $i - j$ is odd, then $\Sigma_{U,ij} = \rho_{U,1}$; (iii) if $i \neq j$, $i \wedge j = 1$, and $i - j$ is even, then $\Sigma_{U,ij} = \rho_{U,2}$; (iv) if $i \neq j$, $i \wedge j \geq 2$, and $i - j$ is odd, then $\Sigma_{U,ij} = \rho_{U,3}$; and (v) if $i \neq j$, $i \wedge j \geq 2$, and $i - j$ is even, then $\Sigma_{U,ij} = \rho_{U,4}$, where $i \wedge j = \min(i, j)$. The covariance matrix Σ_ϵ of ϵ_t is defined similarly. Let $\rho_U = (\rho_{U,1}, \ldots, \rho_{U,4})$ and $\rho_\epsilon = (\rho_{\epsilon,1}, \ldots, \rho_{\epsilon,4})$. To incorporate correlation diversity, we consider $(\rho_U, \rho_\epsilon) = (\tilde{a}, \tilde{a})$, $(\tilde{b}, \tilde{b})$, $(\tilde{b}, \tilde{a})$, and $(\tilde{a}, \tilde{b})$, where $\tilde{a} = (0.5, 0.5, 0.5, 0.5)$ and $\tilde{b} = (-0.5, 0.5, -0.5, 0.5)$.

Table 1 reports the coverage ratios of the 95% confidence intervals for C_α^T, based on 1,000 simulated stochastic volatility series with $T = 84, 105$, and 126, $N = \lceil c_2 T^\delta \rceil$, $c_2 = 0.8, 1.0$, and 1.2, $\delta = 0.8$ and 1.2, and $\alpha = 0.01$, for case (i) $v_{i,t} = \bar{\sigma} \exp(Z_{i,t}/2)$, for $i = 1, \ldots, m$. Those for case (ii) $v_{i,t} = \bar{d}|Z_{i,t} + 2 \log \bar{\sigma}|$ are reported in Table 2. The notation $\lceil x \rceil$ denotes the lowest integer greater than or equal to x. Because a closed-form solution to the true value of C_α^T is intractable, we compute C_α^T using the Monte Carlo method, based on 10^6 price paths generated by the given GMSV model. As mentioned earlier, we use the sampling window method to estimate the limiting variance g^2 for the case of a known mean; for the block size $k = \lambda n^{1/3}$ (cf., Equation (3.10)), we choose $\lambda = 3$.

Overall, the finite-sample performance for both unknown μ and known μ are equally good. For case (i) $v_{i,t} = \bar{\sigma}\exp(Z_{i,t}/2)$, the empirical coverage ratios for known μ and unknown μ fall within the ranges $(0.916, 0.963)$ and $(0.925, 0.960)$, respectively; those for case (ii) $v_{i,t} = \bar{d}|Z_{i,t} + 2\log\bar{\sigma}|$ are $(0.931, 0.963)$ and $(0.930, 0.966)$, respectively. The simulation scheme we adopt to produce Tables 1 and 2 intentionally includes a moderate $T = 84$ and combinations of (N, T) that include $T < N$ and $T > N$ for both an unknown and a known mean. The purpose of the design is to demonstrate that our estimator performs well even when T is not large and the two conditions assumed on T and N in parts (i) and (ii) of Proposition 1 are not strictly met. In the first subsection of Section 5, we encounter a situation in which N is small owing to the choice of large T.

To determine whether confidence intervals (3.8) and (3.9) in Proposition 2 still work well for small N, we conduct a simulation with $N = 0.8, 1, 2$ and T given as in Tables 1 and 2. Moreover, to highlight the robustness of our approach to a departure from normality, as well as the seven normal models considered in Tables 1 and 2, we consider an additional six nonnormal univariate (i.e., $m = 1$) models. Specifically, for $m = 1$, the returns are generated from $r_t = \mu + V_t U_t$, where $V_t = \bar{\sigma}\exp(Z_t/2)$. Here, $\{Z_t = Z_{1,t}\}$ is the Gaussian AR(1) process determined by the dynamic equation $Z_t = \phi Z_{t-1} + \epsilon_t$, where $\varepsilon_t \overset{i.i.d.}{\sim} N\left(0, \bar{\beta}^2\left(1-\phi^2\right)\right)$. The six univariate nonnormal distributions we consider for $\{U_t = u_{1,t}\}$ are the following: a generalized error distribution (GED) with mean=0, sd=1, $\nu = 1$ and 1.5, and a skew-normal (SN) distribution with (ξ, ω, α) =(-0.68, 1.21, 1),(-1.22, 1.58, 4), $(1.22, 1.58, -4)$, and $(0.68, 1.21, -1)$. (The values of ξ and ω ensure that the mean and variance of U_t are zero and one, respectively).

Because the sample size n is much smaller than those used for Tables 1 and 2, the tuning parameter λ that controls the block size k for the mean-known case reduces to one. The results are summarized in Table 3. The overall performance is reasonably good, because the empirical coverage ratios lie in the range $(0.858, 0.974)$ and most are above 0.90; the exceptions include a few instances when U_t is GED(0,1,1) or SN(-1.22,1.58,4). These results demonstrate two points. First, despite the normal assumption imposed on the shock sequence $\{U_t\}$ in Proposition 1, our approach is robust to nonnormal shocks. Second, they provide strong evidence supporting part (a) of **Remark 2** that, even for small N $(= 0.8, 1.0, 2.0)$ such that the sample size is no greater than the horizon (i.e.,

Table 1. Coverage ratios of the 95% confidence intervals for C_α^T based on stochastic volatility sequences. The results are based on 1,000 replicates, and the true C_α^T is computed by simulating 10^6 price paths from the true model, with $v_{i,t} = \bar{\sigma}\exp(Z_{i,t}/2)$, for $i = 1, \ldots, m$, $N = \lceil c_2 T^\delta \rceil$, and $\alpha = 0.01$.

c_2	μ	known						unknown					
	T	84		105		126		84		105		126	
	δ	0.8	1.2	0.8	1.2	0.8	1.2	0.8	1.2	0.8	1.2	0.8	1.2
	$U_t \setminus N$	28	164	34	214	39	266	28	164	34	214	39	266
	(a)	0.935	0.938	0.944	0.940	0.948	0.934	0.948	0.944	0.945	0.947	0.952	0.938
	(b)	0.931	0.921	0.937	0.946	0.952	0.960	0.950	0.946	0.947	0.953	0.940	0.960
	(c)	0.948	0.962	0.958	0.955	0.950	0.933	0.945	0.940	0.942	0.952	0.956	0.933
0.8	(d)	0.941	0.941	0.952	0.959	0.947	0.940	0.947	0.949	0.934	0.956	0.950	0.947
	(e)	0.934	0.952	0.941	0.950	0.957	0.952	0.957	0.948	0.934	0.945	0.954	0.944
	(f)	0.942	0.946	0.952	0.954	0.952	0.948	0.955	0.937	0.937	0.951	0.954	0.949
	(g)	0.937	0.950	0.946	0.953	0.960	0.950	0.939	0.948	0.943	0.960	0.954	0.940
	$U_t \setminus N$	35	204	42	267	48	332	35	204	42	267	48	332
	(a)	0.949	0.932	0.948	0.926	0.941	0.945	0.958	0.941	0.952	0.944	0.956	0.957
	(b)	0.939	0.916	0.959	0.929	0.952	0.941	0.931	0.945	0.945	0.944	0.940	0.925
	(c)	0.950	0.945	0.951	0.949	0.934	0.944	0.950	0.938	0.939	0.945	0.941	0.936
1.0	(d)	0.949	0.951	0.951	0.958	0.942	0.946	0.940	0.940	0.948	0.946	0.944	0.936
	(e)	0.955	0.948	0.943	0.943	0.942	0.951	0.946	0.952	0.933	0.955	0.939	0.949
	(f)	0.957	0.948	0.943	0.952	0.945	0.933	0.944	0.944	0.935	0.944	0.925	0.955
	(g)	0.946	0.954	0.944	0.956	0.943	0.952	0.941	0.937	0.947	0.941	0.932	0.932
	$U_t \setminus N$	42	245	50	320	58	398	42	245	50	320	58	398
	(a)	0.940	0.935	0.958	0.921	0.934	0.923	0.953	0.940	0.949	0.935	0.950	0.941
	(b)	0.938	0.925	0.938	0.936	0.956	0.953	0.939	0.943	0.941	0.943	0.956	0.939
	(c)	0.955	0.939	0.946	0.963	0.944	0.939	0.944	0.941	0.951	0.931	0.941	0.936
1.2	(d)	0.951	0.931	0.942	0.944	0.943	0.945	0.948	0.950	0.926	0.949	0.939	0.955
	(e)	0.947	0.960	0.952	0.942	0.951	0.946	0.938	0.935	0.950	0.949	0.942	0.942
	(f)	0.950	0.963	0.946	0.948	0.945	0.948	0.947	0.933	0.955	0.953	0.940	0.943
	(g)	0.952	0.946	0.948	0.948	0.947	0.957	0.954	0.953	0.936	0.951	0.938	0.953

(a) MVN_2(-0.5,0.5); (b) MVN_2(0.5,0.5); (c) MVN_2(-0.5,-0.5); (d) $\mathrm{MVN}_{10}(\tilde{a},\tilde{a})$; (e) $\mathrm{MVN}_{10}(\tilde{b},\tilde{b})$; (f) $\mathrm{MVN}_{10}(\tilde{b},\tilde{a})$; (g) $\mathrm{MVN}_{10}(\tilde{a},\tilde{b})$, where $\tilde{a} = (0.5,\ldots,0.5)$ and $\tilde{b} = (-0.5, 0.5, -0.5, 0.5)$.

$n = NT \le T$), $\widehat{C}_\alpha^T$ can be used effectively to locate C_α^T.

4.2. Performance of the predicted CTE for future returns

Following the method outlined in Subsection 3.2, we conduct a simulation study to evaluate the size of the $1-\beta$ acceptance region $(\hat{\pi}_n(\alpha) - U_{\beta/2,k}(n^*)^{-1/2}$, $\hat{\pi}_n(\alpha) - L_{\beta/2,k}(n^*)^{-1/2})$ given in (3.18), for a value of $\Phi(Z_\alpha)$. We consider $n =$

Table 2. Coverage ratios of the 95% confidence intervals for C_α^T based on stochastic volatility sequences. The results are based on 1,000 replicates, and the true C_α^T is computed by simulating 10^6 price paths from the true model, with $v_{i,t} = \bar{d}|Z_{i,t} + 2\log\bar{\sigma}|$, for $i = 1, \ldots, m$, $N = \lceil c_2 T^\delta \rceil$, and $\alpha = 0.01$.

c_2	μ	known						unknown					
	T	84		105		126		84		105		126	
	δ	0.8	1.2	0.8	1.2	0.8	1.2	0.8	1.2	0.8	1.2	0.8	1.2
	$U_t \setminus N$	28	164	34	214	39	266	28	164	34	214	39	266
	(a)	0.956	0.938	0.946	0.942	0.942	0.951	0.943	0.947	0.956	0.935	0.942	0.949
	(b)	0.940	0.961	0.956	0.957	0.935	0.932	0.943	0.961	0.942	0.943	0.937	0.942
	(c)	0.954	0.940	0.947	0.951	0.955	0.963	0.946	0.940	0.947	0.951	0.931	0.955
0.8	(d)	0.948	0.940	0.952	0.952	0.943	0.942	0.935	0.966	0.930	0.949	0.951	0.944
	(e)	0.955	0.953	0.952	0.949	0.947	0.940	0.937	0.944	0.948	0.938	0.954	0.949
	(f)	0.954	0.939	0.951	0.950	0.938	0.942	0.934	0.967	0.934	0.949	0.950	0.945
	(g)	0.952	0.953	0.954	0.950	0.949	0.943	0.942	0.945	0.947	0.940	0.952	0.946
	$U_t \setminus N$	35	204	42	267	48	332	35	204	42	267	48	332
	(a)	0.938	0.940	0.945	0.952	0.950	0.950	0.933	0.948	0.941	0.945	0.961	0.945
	(b)	0.942	0.950	0.952	0.929	0.937	0.936	0.956	0.935	0.941	0.951	0.951	0.945
	(c)	0.940	0.948	0.942	0.962	0.946	0.945	0.951	0.948	0.961	0.944	0.945	0.951
1.0	(d)	0.941	0.964	0.952	0.950	0.959	0.949	0.940	0.946	0.947	0.950	0.953	0.952
	(e)	0.938	0.950	0.941	0.949	0.951	0.952	0.939	0.947	0.931	0.945	0.946	0.963
	(f)	0.941	0.963	0.944	0.952	0.965	0.946	0.939	0.943	0.943	0.951	0.952	0.948
	(g)	0.942	0.950	0.942	0.948	0.948	0.956	0.938	0.948	0.932	0.944	0.947	0.966
	$U_t \setminus N$	42	245	50	320	58	398	42	245	50	320	58	398
	(a)	0.956	0.950	0.955	0.955	0.952	0.949	0.947	0.937	0.936	0.955	0.942	0.950
	(b)	0.955	0.956	0.949	0.940	0.944	0.931	0.939	0.945	0.946	0.941	0.945	0.953
	(c)	0.949	0.949	0.945	0.965	0.949	0.947	0.946	0.948	0.935	0.949	0.934	0.947
1.2	(d)	0.954	0.942	0.960	0.949	0.949	0.950	0.941	0.950	0.943	0.951	0.953	0.949
	(e)	0.953	0.959	0.948	0.948	0.939	0.943	0.930	0.954	0.944	0.939	0.942	0.947
	(f)	0.955	0.942	0.960	0.951	0.947	0.945	0.940	0.953	0.941	0.952	0.952	0.949
	(g)	0.956	0.960	0.948	0.949	0.938	0.939	0.935	0.952	0.947	0.939	0.940	0.947

(a) MVN_2(-0.5,0.5); (b) MVN_2(0.5,0.5); (c) MVN_2(-0.5,-0.5); (d) $\text{MVN}_{10}(\tilde{a}, \tilde{a})$; (e) $\text{MVN}_{10}(\tilde{b}, \tilde{b})$; (f) $\text{MVN}_{10}(\tilde{b}, \tilde{a})$; (g) $\text{MVN}_{10}(\tilde{a}, \tilde{b})$, where $\tilde{a} = (0.5, \ldots, 0.5)$ and $\tilde{b} = (-0.5, 0.5, -0.5, 0.5)$.

15,500; $t_0 = 1{,}000$; $T = 120, 180$, and 250, corresponding roughly to six months, nine months, and one year, respectively; and window size $W_T = T, 2T$, and $3T$. Similarly to Subsection 4.1, we set $k = \lambda n^{1/3}$, where $\lambda = 2$, 3, and 4. The true data-generating process is the univariate version (i.e., $m = 1$) of the standard SV model employed in case (i) $v_{i,t} = \bar{\sigma}\exp(Z_{i,t}/2)$ of Subsection 4.1, with $U_t \sim N(0, 1)$. The empirical acceptance rates of the 95% confidence intervals

Table 3. Coverage ratios of the 95% confidence intervals for C_α^T. The results are based on 1,000 replicates, and the true C_α^T is computed by simulating 10^6 price paths from the true model, with (i) $v_{i,t} = \bar{\sigma}\exp(Z_{i,t}/2)$, and (ii) $v_{i,t} = \bar{d}|Z_{i,t} + 2\log\bar{\sigma}|$, for $i = 1, \ldots, m$, $N = 0.8$, 1, and 2, and $\alpha = 0.01$.

	$(i) v_{i,t} = \bar{\sigma}\exp(Z_{i,t}/2)$																	
	$N=0.8$						$N=1$						$N=2$					
μ	known			unknown			known			unknown			known			unknown		
$U_t \setminus T$	84	105	126	84	105	126	84	105	126	84	105	126	84	105	126	84	105	126
(1)	0.872	0.884	0.897	0.913	0.917	0.939	0.884	0.883	0.902	0.934	0.936	0.927	0.889	0.919	0.922	0.917	0.934	0.930
(2)	0.901	0.928	0.915	0.924	0.933	0.953	0.930	0.919	0.920	0.950	0.943	0.933	0.931	0.931	0.923	0.943	0.933	0.933
(3)	0.914	0.934	0.908	0.915	0.932	0.925	0.931	0.928	0.943	0.933	0.940	0.931	0.914	0.924	0.923	0.921	0.928	0.939
(4)	0.858	0.875	0.863	0.894	0.905	0.894	0.880	0.888	0.868	0.892	0.906	0.905	0.881	0.858	0.881	0.900	0.910	0.927
(5)	0.928	0.933	0.932	0.959	0.974	0.965	0.947	0.935	0.938	0.965	0.963	0.958	0.935	0.942	0.933	0.969	0.956	0.961
(6)	0.929	0.935	0.942	0.947	0.943	0.944	0.939	0.923	0.944	0.945	0.933	0.951	0.938	0.953	0.949	0.949	0.945	0.941
(7)	0.925	0.934	0.951	0.930	0.949	0.941	0.930	0.932	0.928	0.943	0.954	0.934	0.938	0.943	0.949	0.942	0.947	0.943
(8)	0.929	0.933	0.934	0.939	0.952	0.944	0.946	0.939	0.938	0.943	0.933	0.951	0.948	0.957	0.942	0.940	0.937	0.944
(9)	0.946	0.939	0.947	0.933	0.917	0.942	0.938	0.940	0.946	0.931	0.939	0.947	0.950	0.953	0.954	0.957	0.926	0.938
(10)	0.941	0.946	0.935	0.942	0.941	0.943	0.951	0.943	0.957	0.934	0.935	0.946	0.948	0.952	0.932	0.943	0.946	0.951
(11)	0.948	0.958	0.944	0.943	0.936	0.945	0.954	0.948	0.949	0.918	0.942	0.944	0.958	0.947	0.950	0.932	0.941	0.946
(12)	0.946	0.951	0.946	0.952	0.936	0.943	0.957	0.950	0.954	0.939	0.943	0.946	0.955	0.960	0.952	0.940	0.943	0.950
(13)	0.951	0.950	0.936	0.938	0.936	0.934	0.939	0.930	0.951	0.922	0.939	0.952	0.957	0.952	0.957	0.943	0.943	0.957
	$(ii) v_{i,t} = \bar{d}\lvert Z_{i,t} + 2\log\bar{\sigma}\rvert$																	
	$N=0.8$						$N=1$						$N=2$					
μ	known			unknown			known			unknown			known			unknown		
$U_t \setminus T$	84	105	126	84	105	126	84	105	126	84	105	126	84	105	126	84	105	126
(1)	0.903	0.895	0.922	0.940	0.939	0.927	0.900	0.920	0.918	0.931	0.933	0.937	0.918	0.917	0.916	0.923	0.935	0.936
(2)	0.941	0.927	0.932	0.921	0.947	0.936	0.935	0.947	0.950	0.934	0.950	0.937	0.947	0.945	0.944	0.948	0.944	0.953
(3)	0.936	0.953	0.949	0.932	0.950	0.930	0.948	0.952	0.944	0.936	0.946	0.940	0.942	0.954	0.941	0.932	0.939	0.947
(4)	0.915	0.916	0.925	0.903	0.917	0.915	0.898	0.891	0.898	0.908	0.902	0.905	0.896	0.902	0.901	0.909	0.915	0.927
(5)	0.945	0.938	0.938	0.961	0.962	0.958	0.943	0.941	0.940	0.970	0.964	0.961	0.939	0.942	0.940	0.959	0.971	0.966
(6)	0.949	0.961	0.945	0.942	0.948	0.952	0.946	0.950	0.960	0.956	0.936	0.948	0.956	0.960	0.955	0.944	0.948	0.957
(7)	0.939	0.958	0.945	0.943	0.933	0.950	0.961	0.950	0.943	0.937	0.943	0.945	0.958	0.966	0.959	0.941	0.936	0.941
(8)	0.945	0.948	0.956	0.941	0.943	0.950	0.943	0.935	0.943	0.949	0.925	0.944	0.957	0.955	0.959	0.951	0.947	0.940
(9)	0.955	0.959	0.954	0.934	0.940	0.935	0.944	0.955	0.958	0.952	0.936	0.950	0.955	0.950	0.957	0.955	0.932	0.954
(10)	0.952	0.957	0.948	0.943	0.937	0.942	0.952	0.951	0.954	0.940	0.923	0.943	0.951	0.950	0.960	0.949	0.942	0.945
(11)	0.959	0.959	0.933	0.944	0.942	0.936	0.960	0.948	0.944	0.944	0.953	0.937	0.948	0.964	0.959	0.937	0.951	0.959
(12)	0.954	0.956	0.951	0.941	0.936	0.945	0.949	0.952	0.956	0.939	0.922	0.942	0.952	0.946	0.958	0.948	0.943	0.942
(13)	0.960	0.958	0.931	0.941	0.939	0.936	0.962	0.950	0.947	0.943	0.953	0.941	0.952	0.953	0.954	0.937	0.951	0.962

(1) GED(0,1,1); (2) GED(0,1,1.5); (3) SN(-0.68,1.21,1); (4) SN(-1.22,1.58,4); (5) SN(1.22,1.58,-4); (6) SN(0.68,1.21,-1); (7) MVN_2(-0.5,0.5); (8) MVN_2(0.5,0.5); (9) MVN_2(-0.5,-0.5); (10) $\mathrm{MVN}_{10}(\tilde{a}, \tilde{a})$; (11) $\mathrm{MVN}_{10}(\tilde{b}, \tilde{b})$; (12) $\mathrm{MVN}_{10}(\tilde{b}, \tilde{a})$; (13) $\mathrm{MVN}_{10}(\tilde{a}, \tilde{b})$, where $\tilde{a} = (0.5, \ldots, 0.5)$ and $\tilde{b} = (-0.5, 0.5, -0.5, 0.5)$.

for $\Phi(Z_\alpha)$, where α ranges from 0.01 to 0.2 with increments of 0.01, based on 1,000 independent replicates, are summarized in Table 4. In general, $W_T = 3T$ and $\lambda = 4$ yield the best results in terms of the closeness between the empirical and nominal coverage ratios. The results further confirm the appropriateness of our method. Although we assume $T = o(W_T)$ in Subsection 3.2 to ensure (3.17), a mild $W_T = 3T$ is sufficient to warrant satisfactory rates. In addition, for $W_T = 3T$ and $\lambda = 4$, a few acceptance rates are less than 0.900 for $\alpha \leq 0.03$. This is expected, because the corresponding $\Phi(Z_\alpha)$ (not greater than 0.012) represents an extreme tail event that is prone to relatively large estimation errors.

5. Applications

To demonstrate the potential applications of our CTE estimator and predictor, we conduct two experiments using data on the S&P 500 index (United States). From Yahoo Finance (`http://finance.yahoo.com/`), we obtained adjusted daily prices $\{P_t\}$ of the index from January 3, 1950, to January 8, 2019, yielding 17,365 log-returns $r_t = \log(P_t/P_{t-1})$. Similarly to Subsections 4.1 and 4.2, we present the applications for the unconditional coverage ratios and the accuracy of the conditional forecast in separate subsections.

5.1. Unconditional coverage ratios

In this subsection our goal is to compute the empirical coverage ratios for the CTE of integrated S&P 500 returns, based on the confidence intervals given in (3.8). To this end, we need to overcome two disadvantages: there is no analytic solution to the true value of the CTE, and we have only a single series of the index. We do so using the rolling window method, from which two families of subsamples are generated. These subsamples enable us to approximate the true CTE and create sufficient replicas to mimic the procedure in a simulation experiment. The steps of our experiment are as follows. First, we generate two families of subsamples: $\{\mathcal{R}_i\}$, with $\mathcal{R}_i = \{r_{1+(i-1)20}, r_{2+(i-1)20}, \ldots, r_{T+(i-1)20}\}$, for $i = 1, \ldots, \lfloor (H-T)/20 + 1 \rfloor$, and $\{\mathcal{S}_j\}$, with $\mathcal{S}_j = \{r_{1+(j-1)20}, r_{2+(j-1)20}, \ldots, r_{n+(j-1)20}\}$, for $j = 1, \ldots, \lfloor (H-n)/20 + 1 \rfloor$. Here, $H =$ 17,365 denotes the total number of daily returns $\{r_t\}_{t=1,\ldots,H}$; $T = 252 \cdot M$, where $M = 1, 2, \ldots, 8$, and 10; and $n =$ 2,520 and 3,000 are the length of the integration and the sample size of each replica, respectively. The lag of 20 corresponds roughly to the number of trading days in a month. This design is intended to have sufficient subsamples that do not overlap too heavily. Second, we compute the integrated

Table 4. Coverage ratios of the approximate 95% confidence intervals $(\hat{\pi}_n(\alpha) - U_{\beta/2,k}(n^*)^{-1/2}, \hat{\pi}_n(\alpha) - L_{\beta/2,k}(n^*)^{-1/2})$ for $\Phi(Z_\alpha)$. The results are based on 1,000 replicates.

			$T=120$			$T=180$			$T=250$		
λ	α	$\Phi(Z_\alpha)$	$W_T=T$	$W_T=2T$	$W_T=3T$	$W_T=T$	$W_T=2T$	$W_T=3T$	$W_T=T$	$W_T=2T$	$W_T=3T$
2	0.01	0.004	0.136	0.602	0.443	0.428	0.526	0.261	0.576	0.409	0.204
	0.02	0.008	0.083	0.934	0.930	0.378	0.921	0.824	0.628	0.858	0.721
	0.03	0.012	0.060	0.920	0.962	0.303	0.960	0.916	0.542	0.930	0.872
	0.04	0.016	0.047	0.887	0.966	0.241	0.945	0.949	0.428	0.939	0.907
	0.05	0.020	0.028	0.864	0.969	0.210	0.920	0.947	0.380	0.903	0.892
	0.06	0.024	0.028	0.824	0.970	0.159	0.882	0.930	0.324	0.868	0.870
	0.07	0.028	0.021	0.808	0.966	0.123	0.895	0.944	0.278	0.877	0.901
	0.08	0.032	0.020	0.767	0.946	0.144	0.858	0.923	0.264	0.843	0.877
	0.09	0.036	0.016	0.782	0.948	0.109	0.818	0.927	0.210	0.808	0.869
	0.10	0.040	0.018	0.749	0.944	0.100	0.796	0.914	0.185	0.771	0.862
	0.11	0.044	0.009	0.740	0.940	0.083	0.768	0.910	0.178	0.754	0.840
	0.12	0.048	0.017	0.726	0.929	0.093	0.767	0.895	0.177	0.725	0.818
	0.13	0.052	0.017	0.699	0.931	0.090	0.749	0.894	0.141	0.715	0.835
	0.14	0.056	0.007	0.679	0.914	0.056	0.739	0.902	0.129	0.713	0.834
	0.15	0.060	0.014	0.649	0.927	0.051	0.681	0.863	0.103	0.658	0.801
	0.16	0.064	0.009	0.652	0.907	0.057	0.661	0.855	0.108	0.637	0.780
	0.17	0.068	0.009	0.647	0.904	0.051	0.642	0.850	0.094	0.601	0.779
	0.18	0.072	0.008	0.606	0.880	0.040	0.625	0.825	0.085	0.616	0.773
	0.19	0.077	0.006	0.599	0.896	0.041	0.628	0.840	0.092	0.629	0.777
	0.20	0.081	0.011	0.598	0.892	0.044	0.592	0.809	0.082	0.587	0.763
3	0.01	0.004	0.369	0.911	0.805	0.723	0.820	0.593	0.873	0.672	0.448
	0.02	0.008	0.187	0.980	0.957	0.560	0.965	0.887	0.806	0.924	0.809
	0.03	0.012	0.103	0.953	0.969	0.451	0.976	0.920	0.686	0.946	0.883
	0.04	0.016	0.070	0.938	0.972	0.353	0.976	0.953	0.599	0.969	0.914
	0.05	0.020	0.057	0.927	0.973	0.324	0.964	0.951	0.544	0.955	0.907
	0.06	0.024	0.045	0.893	0.979	0.257	0.942	0.952	0.468	0.926	0.893
	0.07	0.028	0.029	0.864	0.978	0.225	0.947	0.960	0.434	0.941	0.935
	0.08	0.032	0.033	0.848	0.969	0.227	0.926	0.948	0.382	0.930	0.913
	0.09	0.036	0.025	0.844	0.975	0.190	0.898	0.955	0.342	0.914	0.921
	0.10	0.040	0.027	0.838	0.976	0.167	0.903	0.945	0.322	0.889	0.926
	0.11	0.044	0.014	0.833	0.969	0.149	0.894	0.946	0.329	0.873	0.910
	0.12	0.048	0.024	0.820	0.961	0.155	0.872	0.942	0.308	0.852	0.884
	0.13	0.052	0.028	0.794	0.964	0.149	0.873	0.945	0.264	0.856	0.899
	0.14	0.056	0.014	0.792	0.950	0.118	0.852	0.945	0.237	0.847	0.904
	0.15	0.060	0.017	0.770	0.971	0.115	0.842	0.938	0.209	0.814	0.897
	0.16	0.064	0.011	0.759	0.954	0.115	0.826	0.921	0.209	0.800	0.877
	0.17	0.068	0.018	0.775	0.955	0.111	0.818	0.932	0.208	0.788	0.893
	0.18	0.072	0.016	0.741	0.943	0.103	0.788	0.910	0.198	0.773	0.869
	0.19	0.077	0.011	0.749	0.946	0.107	0.793	0.931	0.197	0.807	0.879
	0.20	0.081	0.022	0.744	0.951	0.099	0.791	0.911	0.179	0.758	0.860
4	0.01	0.004	0.484	0.966	0.893	0.808	0.910	0.748	0.934	0.802	0.617
	0.02	0.008	0.234	0.985	0.959	0.638	0.976	0.898	0.866	0.939	0.838
	0.03	0.012	0.138	0.957	0.969	0.526	0.981	0.924	0.753	0.951	0.887
	0.04	0.016	0.089	0.956	0.977	0.421	0.984	0.958	0.690	0.976	0.922
	0.05	0.020	0.070	0.946	0.980	0.387	0.972	0.957	0.641	0.963	0.912
	0.06	0.024	0.056	0.909	0.981	0.302	0.962	0.956	0.559	0.949	0.901
	0.07	0.028	0.041	0.889	0.983	0.262	0.958	0.963	0.512	0.964	0.940
	0.08	0.032	0.042	0.869	0.975	0.257	0.954	0.955	0.468	0.965	0.923
	0.09	0.036	0.026	0.866	0.982	0.236	0.945	0.967	0.449	0.943	0.932
	0.10	0.040	0.030	0.856	0.981	0.207	0.934	0.962	0.418	0.929	0.939
	0.11	0.044	0.015	0.853	0.967	0.204	0.930	0.962	0.412	0.923	0.929
	0.12	0.048	0.027	0.845	0.965	0.188	0.915	0.958	0.399	0.906	0.922
	0.13	0.052	0.037	0.825	0.975	0.185	0.923	0.960	0.350	0.919	0.932
	0.14	0.056	0.012	0.818	0.961	0.174	0.913	0.967	0.328	0.914	0.930
	0.15	0.060	0.020	0.795	0.972	0.153	0.902	0.965	0.317	0.894	0.938
	0.16	0.064	0.013	0.798	0.970	0.165	0.892	0.946	0.314	0.875	0.915
	0.17	0.068	0.022	0.807	0.970	0.148	0.890	0.958	0.318	0.870	0.931
	0.18	0.072	0.018	0.785	0.969	0.142	0.879	0.951	0.289	0.853	0.912
	0.19	0.077	0.013	0.789	0.959	0.150	0.884	0.958	0.299	0.876	0.924
	0.20	0.081	0.023	0.786	0.972	0.159	0.864	0.947	0.275	0.836	0.916

return $R_T^{(i)} = \sum_{t=1+(i-1)20}^{T+(i-1)20} r_t$ using the subsample $\mathcal{R}_i$, for each i ranging from 1 to $\lfloor (H-T)/20)+1 \rfloor$. Third, we rank these $R_T^{(i)}$ to compute the αth CTE, which we treat as the true C_α^T with $\alpha = 0.01$, and denote it by $\widetilde{C}_\alpha^T$. Fourth, we compute the sample mean $\hat{\mu}_n^{(j)}$ and the sample standard deviation $\hat{\sigma}_n^{(j)}$ of the subsample $\mathcal{S}_j$, for $j = 1, \ldots, \lfloor (H-n)/20) + 1 \rfloor$. Then, we use these values to derive the confidence interval, following (3.4) and (3.8), to determine whether it covers $\widetilde{C}_\alpha^T$ obtained in the third step. Finally, we compute the empirical coverage ratios for $\widetilde{C}_\alpha^T$ based on the $\lfloor (H-n)/20+1 \rfloor$ confidence intervals.

Some of the horizons, such as $T = 2{,}016$ and 2,520, are much greater than those considered in the simulation studies carried out in Subsection 4.1. This is because the performance of the coverage ratios depends on three factors: (i) the term $\mathcal{A}_T \equiv \sqrt{N/T}(\widehat{C}_\alpha^T - C_\alpha^T) - \sqrt{n}\,(\hat{\mu}_n - \mu)$, shown in (3.7) for fixed N, which is $O_p(1/\sqrt{T})$; (ii) the convergence rate $\mathcal{N}_n$ of $\sqrt{n}(\hat{\mu}_n - \mu)/\sigma$ to $N(0,1)$, which is $O(1/\sqrt{n})$, where n denotes the size of the replica; and (iii) how quickly the average obtained in the final step described above converges to the nominal level. For factor (iii), the probability error $\mathcal{E}_{H,T}$ of the convergence is $O_p(1/\sqrt{H^*})$, where $H^* = H - n$ represents the number of replicas. If the replicas are independent, such as those simulated for the finite-sample analysis in Section 4, the probability order of $\mathcal{E}_{H,T}$ is $\sqrt{c/H^*}$. Here, c is close to 0.95×0.05, according to the central limit theorem, and thus is negligible when sufficient replicas can be generated. However, in the current experiment, the speed of $\mathcal{E}_{H,T}$ is considerably slower, because a limited number of replicas can be created from one common S&P 500 series, and all are dependent. This may result in the orders of $\mathcal{A}_T$ and $\mathcal{E}_{H,T}$ being of equal magnitude, even though H^* is much greater than T. In this case, we obtain good coverage ratios by using the large T to reduce the order $O_p(1/\sqrt{T})$ of $\mathcal{A}_T$ in factor (i). Although this increases the order of $\mathcal{E}_{H,T}$, the effect is much less than that on the order of $\mathcal{A}_T$. The choice of large T means the ratio $N = n/T$ is necessarily restricted in order to produce sufficient subsample replicas of size n. Therefore, the order of $\mathcal{N}_n$ in factor (ii) is also $O(1/\sqrt{T})$, and thus shrinks as T increases.

The results in Table 5 show that the coverage ratios increase with T, and are close to the nominal level of 0.95 only when T (2,016 and 2,052) is fairly large, as explained above. For these two cases of T, a few remarks are worth mentioning. Because the difference $|\widetilde{C}_\alpha^T - C_\alpha^T|$ between $\widetilde{C}_\alpha^T$ and C_α^T is of probability order $O_p(1/\sqrt{H^*})$, the coverage ratios in Table 5 for the approximated $\widetilde{C}_\alpha^T$ are close to those obtained using the same procedures and the true, but unknown C_α^T.

Table 5. Coverage ratios of the 95% confidence intervals of C_α^T with $\alpha = 0.01$ for S&P 500 daily returns.

	R_T		coverage ratios ($N = n/T$)	
T	mean	median	$n = 2{,}520$	$n = 3{,}000$
252	0.0750	0.1001	0.2651 (10.0)	0.1961 (11.9)
504	0.1441	0.1599	0.5976 (5.0)	0.5758 (6.0)
756	0.2150	0.2352	0.7470 (3.3)	0.7538 (4.0)
1,008	0.2838	0.3033	0.8681 (2.5)	0.8818 (3.0)
1,260	0.3551	0.3904	0.7456 (2.0)	0.7622 (2.4)
1,512	0.4177	0.4585	0.8869 (1.7)	0.8873 (2.0)
1,764	0.4805	0.4742	0.8600 (1.4)	0.8679 (1.7)
2,016	0.5452	0.5574	0.9354 (1.3)	0.9138 (1.5)
2,520	0.6653	0.7016	0.9435 (1.0)	0.9207 (1.2)

As pointed out in part (a) of **Remark 2**, the confidence intervals constructed from (3.8) capture the true value $\widetilde{C}_\alpha^T$, even when the value of n/T (displayed in parentheses in Table 5) is small.

5.2. Conditional forecasts

To apply the conditional forecast $\widehat{C}_{\alpha,t}^T$ to the data on S&P 500 index returns, two concerns need to be addressed. First, it is impractical to choose an 8- or 10-year horizon, as we did in the previous subsection. Second, each condition indicator function $\hat{I}_t^T(\alpha)$ defined in (3.16) involves W_T past returns and T future returns, creating persistent dependence between these functions. Consequently, the subsample method used in Subsection 5.1 to produce replicates as the basis for an inference is not suitable. Therefore, rather than conducting a formal test, we use estimation biases to assess the accuracy of the predictor $\widehat{C}_{\alpha,t}^T$. Our approach follows the framework formulated in Subsection 3.2. We reduce the estimation bias caused by the dependence between $R_{T,t}$s by adopting the same method as in the previous subsection, where any pair of consecutive integrated returns are Δ trading days apart. Let $J_\Delta = \{W_T + (i-1)\Delta + 1 : i = 1, 2, \ldots, \lfloor (H - W_T - T)/\Delta + 1 \rfloor\}$ be the set of all t such that $R_{T,t}$ is included. Here, W_T denotes the window length used to compute $\hat{\mu}_{t-1}^*$ and $\hat{\sigma}_{t-1}^*$ (see (3.11)). Then, we use

$$\hat{\pi}(\alpha) = \sum_{j \in J_\Delta} \frac{\hat{I}_j^T(\alpha)}{|J_\Delta|}$$

to estimate $P(R_{T,t} < \widehat{C}_{\alpha,t}^T)$; we report the biases $\hat{\pi}(\alpha) - \Phi(Z_\alpha)$ in Table 6. The values of the design parameters are as follows: the return horizon T is

Table 6. Biases $\hat{\pi}(\alpha) - \Phi(Z_\alpha)$ of estimated probabilities for CTE forecasts

				$T = 500$			$T = 600$			$T = 700$		
Δ	α	$\Phi(Z_\alpha)$	$W_T =$	1,000	1,250	1,500	1,200	1,500	1,800	1,400	1,750	2,100
20	0.10	0.0396		0.0699	0.0755	0.0592	0.0810	0.0585	0.0351	0.0598	0.0353	0.0399
	0.11	0.0437		0.0671	0.0752	0.0564	0.0795	0.0584	0.0337	0.0597	0.0326	0.0414
	0.12	0.0478		0.0643	0.0724	0.0537	0.0780	0.0583	0.0364	0.0609	0.0326	0.0414
	0.13	0.0518		0.0603	0.0709	0.0496	0.0765	0.0607	0.0363	0.0620	0.0298	0.0442
	0.14	0.0559		0.0612	0.0745	0.0481	0.0750	0.0592	0.0335	0.0619	0.0297	0.0442
	0.15	0.0600		0.0634	0.0768	0.0440	0.0773	0.0578	0.0307	0.0617	0.0310	0.0428
	0.16	0.0642		0.0618	0.0778	0.0451	0.0745	0.0576	0.0280	0.0615	0.0309	0.0442
	0.17	0.0683		0.0589	0.0762	0.0474	0.0716	0.0560	0.0238	0.0587	0.0267	0.0428
	0.18	0.0725		0.0611	0.0772	0.0446	0.0739	0.0558	0.0223	0.0584	0.0266	0.0400
	0.19	0.0766		0.0581	0.0743	0.0430	0.0710	0.0543	0.0195	0.0556	0.0265	0.0372
	0.20	0.0808		0.0552	0.0765	0.0402	0.0707	0.0527	0.0180	0.0580	0.0236	0.0344
40	0.10	0.0396		0.0662	0.0755	0.0591	0.0783	0.0598	0.0324	0.0598	0.0379	0.0398
	0.11	0.0437		0.0621	0.0740	0.0576	0.0794	0.0610	0.0310	0.0610	0.0339	0.0412
	0.12	0.0478		0.0606	0.0699	0.0535	0.0779	0.0596	0.0349	0.0648	0.0325	0.0427
	0.13	0.0518		0.0565	0.0709	0.0495	0.0764	0.0607	0.0335	0.0660	0.0310	0.0441
	0.14	0.0559		0.0599	0.0668	0.0480	0.0748	0.0592	0.0294	0.0619	0.0296	0.0427
	0.15	0.0600		0.0634	0.0729	0.0439	0.0784	0.0578	0.0280	0.0604	0.0335	0.0441
	0.16	0.0642		0.0618	0.0739	0.0475	0.0769	0.0589	0.0238	0.0615	0.0321	0.0454
	0.17	0.0683		0.0576	0.0724	0.0434	0.0727	0.0600	0.0197	0.0574	0.0280	0.0440
	0.18	0.0725		0.0611	0.0759	0.0418	0.0711	0.0611	0.0182	0.0584	0.0265	0.0399
	0.19	0.0766		0.0594	0.0717	0.0429	0.0695	0.0595	0.0141	0.0543	0.0303	0.0357
	0.20	0.0808		0.0552	0.0727	0.0413	0.0679	0.0580	0.0099	0.0580	0.0262	0.0343

$500, 600, 700$; the window length W_T, used to compute the adaptive estimates, is $W_T = 2T, 2.5T, 3T$; the distance Δ that separates consecutive returns is $20, 40$; and the level of the CTE ranges from 0.1 to 0.2, in increments of 0.01.

Table 6 shows that $W_T = 3T$ and $T = 700$ achieve the best results. This is expected, because the larger T and W_T help to reduce the estimation errors, according to properties (3.15) and (3.13). In addition, most of the biases reported in $W_T = 3T$ are close to or within the magnitude of the error bound $O(1/\sqrt{T})$ given in (3.15).

6. Conclusion

We estimate a distribution's conditional tail expectation for long return horizons. This study contributes to the literatures in several ways. Even though the target parameter diverges as the horizon increases, we derive a simple nonparametric estimate, and show that it is asymptotically normal under a GMSV model. Furthermore, by using adaptive estimates, the estimator can be transformed into a conditional predictor of the CTE for future returns of long horizons. Treating the predictor as an interval forecast, we developed a t-test to evaluate the accuracy of the predictor. The results of Monte Carlo experiments show that the proposed estimate outperforms traditional simulation-based approaches. Furthermore, we show that the modified predictor performs well, and is consistent with the theoretical results. We demonstrate the usefulness of our findings by applying the estimator and the predictor to data on the S&P 500 index. Several challenges remain. First, it is worth studying models that allow for structural breaks and/or long-range dependence in the volatility process $\{V_t\}$. This theoretically challenging extension is also of practical interest in terms of long-horizon returns. Second, our GMSV model does not include the popular ARCH-type process (Engle (1982); Bollerslev (1986)). As a result, the argument (i.e., being conditional on the volatility component) that we use to build our theory for the SV process may not work, and a completely new approach may be needed. Thus, it is worth extending our findings to include the class of ARCH processes.

Supplementary Material

The online Supplementary Material contains three sections. The regularity conditions used to define the class of functions h are stated in Section S1. Section S2 presents the proof of Proposition 1. Section S3 provides numerical evidence that the coverage ratio of our approach is superior to those of two traditional sample-generation methods.

Acknowledgments

We thank the co-deitor, associate editor, and two anonymous referees for their valuable comments and suggestions. Ho's research was partially supported by the Ministry of Science and Technology (MOST 106-2118-M-001-008) of the Republic of China, and Tsai's research was supported by Academia Sinica and the MOST (MOST 106-2118-M-001-003-MY2). The computational assistance provided by Miao-Chen Chiang is also deeply appreciated.

References

Artzner, P., Delbaen, F., Eber, J-M. and Heath, D.(1999). Coherent measures of risk. *Mathematical Finance* **9**, 203–228.

Asai, M., McAleer, M. and Yu, J. (2006). Multivariate Stochastic Volatility: A Review. *Econometric Reviews* **25**, 145–175.

Asimit, A. V., Furman, E., Tang, Q. and Vernic, R. (2011). Asymptotics for risk capital allocations based on conditional tail expectation. *Insurance: Mathematics and Economics* **49**, 310–324.

Basel Committee on Banking Supervision (2012). *Consultative Document: Fundamental Review of the Trading Book*. Bank for International Settlements Communications, Basel. Available at: `http://www.bis.org/publ/bcbs219.pdf`.

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

Boudoukh, J., Richardson, M. and Whitelaw, R. F. (2008). The myth of long-horizon predictability. *Review of Financial Studies* **21**, 1577–1605.

Du, Z. and Escanciano, J. C. (2015). *Backtesting Expected Shortfall: Accounting for Tail Risk*. CAEPR Working, 2015–001. Available at: `http://goo.gl/xTYKgM`.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflations. *Econometrica* **50**, 987–1007.

Fama, E. F and French, K. R. (2018). Long-horizon returns. *Review of Asset Pricing Studies* **8**, 232–252.

Hamilton, J. D. (1994). *Time Series Analysis*. University Press, New Jersey.

Hardy, M. R. (2001). A regime-switching model of long-term stock returns. *North American Actuarial Journal* **5**, 41–53.

Hardy, M. (2003). *Investment Guarantees: Modeling and Risk Management for Equity-Linked Life Insurance*. John Wiley & Sons, Inc.

Hardy, M, Freeland, R. K. and Till, M.C. (2006). Validation of long-term equity return models for equity-linked guarantees. *North American Actuarial Journal* **10**, 28–47.

Harvey, A., Ruiz, E. and Shephard, N. (1994). Multivariate stochastic variance models. *Review of Economic Studies* **61**, 247–264.

Ho, H.-C., Chen, H.-Y. and Tsai, H. (2016). Value at risk for integrated returns and its applications to equity portfolios. *Statistica Sinica* **26**, 1631–1648.

Lai, T. L. and Xing, H. (2008). *Statistical Models and Methods for Financial Markets*. Springer, New York.

Neuberger, A. (2012). Realized skewness. *Review of Financial Studies* **25**, 3423–3455.

Politis, D. N., Romano, J. P. and Wolf, M. (1999). *Subsampling*. Springer, New York.

Rapach, D. and Zhou, G. (2013). Chapter 6: Forecasting Stock Returns. In *Handbook of Economic Forecasting* **2A** (Edited by G. Elliott, A. Timmermann), 328–383. Elsevier B.V.

Robinson, P. M. (2001). The memory of stochastic volatility models. *Journal of Econometrics* **101**, 195–218.

Taylor, S. (1986). *Modelling Financial Time Series*. John Wiley & Sons, New York.

Yu, J. and Meyer, R. (2006). Multivariate stochastic volatility models: Bayesian estimation and model comparison. *Econometric Reviews* **25**, 361–384.

Hwai-Chung Ho

Institute of Statistical Science, Academia Sinica, Taipei, Taiwan, R.O.C.

E-mail: hcho@stat.sinica.edu.tw

Hung-Yin Chen

Department of Accounting, Chung Yuan Christian University, Taoyuan City, Taiwan, R.O.C.

E-mail: hungyin@cycu.edu.tw
Henghsiu Tsai
Institute of Statistical Science, Academia Sinica, Taipei, Taiwan, R.O.C.
E-mail: htsai@stat.sinica.edu.tw

(Received August 2018; accepted May 2019)