Supplementary Document

for

ESTIMATION AND PREDICTION OF CONDITIONAL TAIL EXPECTATION FOR LONG-HORIZON RETURNS

Hwai-Chung Ho[a,b], Hung-Yin Chen[c], and Henghsiu Tsai[a]

[a]Academia Sinica, [c]Chung Yuan Christian University

and [b]National Taiwan University

Supplementary Material

S1 Regularity conditions

In order to prove the central limit theorem for $\hat{\sigma}_n^2 = n^{-1}\sum_{t=1}^{n}(r_t - \hat{\mu}_n)^2$ as specified below in (S2.4) of Lemma 1, certain regularity conditions on the functions h_i's given in (2.2) are required. To specify these conditions, which are collectively called condition J, we adopt the same approach as by Ho and Hsing (1997). For $\ell \geq 1$, define

$$Z_{i,t,\ell} = \mu_{z,i} + \sum_{s=0}^{\ell} \tilde{A}_i^{(s)} \star \eta_{t-s},\ 1 \leq i \leq m, \tag{S1.1}$$

and denote by F_i and $F_{i,\ell}$ the distribution functions of $Z_{i,t}$ and $Z_{i,t,\ell}$, respectively. For each h_i, define

$$h_{i,j}(x) = \int h_i(x+y)dF_{i,j}(y) \quad \text{and} \quad h_{i,j,\lambda}^{(k)}(x) = \sup_{|y|\leq\lambda} |h_{i,j}^{(k)}(x+y)|, \quad \lambda \geq 0, \tag{S1.2}$$

provided that the k-th derivative $h_{i,j}^{(k)}$ of $h_{i,j}$ exists. Let k be nonnegative integers and λ a nonnegative real number. We say that h satisfies condition J, if the following properties hold.

Condition J.

1. For $i = 1, 2, \ldots, m$ and $k = 0, 1$, $h_{i,1}^{(k)}(x)$ is continuous at all x, and

$$\sup_{I \subset \{0,1,2,\ldots\}} E\Big[h_{i,1,\lambda}^{(k)}\Big(x + \sum_{s \in I} \tilde{A}_i^{(s)} \star \eta_s\Big)\Big]^4 < \infty,$$

   where the sup is taken over all subsets I of $\{0, 1, 2, \ldots\}$.

2. $E(h_i(Z_{i,1}) - h_i(Z_{i,1,\ell}))^2 = o(1)$ as $\ell \to \infty$.

3. $E\eta_{i,1}^4 < \infty$ for $1 \le i \le m$.

S2 Proof of Theorem 1

Throughout the proofs given below, C denotes a generic positive constant whose value may vary from place to place. For proving Proposition 1, the following two lemmas serve as a preparatory step.

Lemma 1. Assume that the GMSV model defined by (2.1), (2.2) and (2.3) satisfies the moment condition that $Ev_{i,1}^4 < \infty$ for $1 \le i \le m$, and that the conditions for $\{U_t\}$ and $\{V_t\}$ stated in Proposition 1 hold. Then, as $n \to \infty$,

$$\sqrt{n}\,(\hat{\mu}_n - \mu) \xrightarrow{d} N(0, \sigma^2). \tag{S2.3}$$

Furthermore, if the condition J holds, then

$$\sqrt{n}\left(\hat{\sigma}_n^2 - \sigma^2\right) \xrightarrow{d} N\left(0, g^2\right). \tag{S2.4}$$

If $\{\eta_t\} \sim N(0, \Sigma_\eta)$, then

$$
\begin{aligned}
g^2 \quad = \quad & \sum_{i,j,k,l=1}^{m} w_i w_j w_k w_l \left(\sigma_{U,ik}\sigma_{U,jl} + \sigma_{U,il}\sigma_{U,jk}\right) e^{J_4' \mu_z(i,j,k,l) + \frac{1}{2} J_4' \Sigma_Z(i,j,k,l) J_4} \\
& + \sum_{i,j,k,l=1}^{m} w_i w_j w_k w_l e^{J_4' \mu_z(i,j,k,l) + \frac{1}{2} J_2' \{\Sigma_Z(i,j) + \Sigma_Z(k,l)\} J_2} \sigma_{U,ij}\sigma_{U,kl} \left(e^{\sigma_{Z,ik} + \sigma_{Z,il} + \sigma_{Z,jk} + \sigma_{Z,jl}} - 1\right) \\
& \times \left\{ 1 + \frac{2}{e^{\sigma_{Z,ik} + \sigma_{Z,il} + \sigma_{Z,jk} + \sigma_{Z,jl}} - 1} \sum_{u=1}^{\infty} \left(e^{J_2' \{\Sigma_{Z,\{(i,j),(k,l)\}}(-u)\} J_2} - 1 \right) \right\}, \qquad \text{(S2.5)}
\end{aligned}
$$

where $J_2 = (1,1)'$, $J_4 = (1,1,1,1)'$, $\mu_z(i,j,k,l) = (\mu_{z,i}, \mu_{z,j}, \mu_{z,k}, \mu_{z,l})'$,

$$
\Sigma_Z(i,j) = \begin{pmatrix} \sigma_{Z,ii} & \sigma_{Z,ij} \\ \sigma_{Z,ji} & \sigma_{Z,jj} \end{pmatrix},
$$

$$
\Sigma_Z(i,j,k,l) = \begin{pmatrix} \sigma_{Z,ii} & \sigma_{Z,ij} & \sigma_{Z,ik} & \sigma_{Z,il} \\ \sigma_{Z,ji} & \sigma_{Z,jj} & \sigma_{Z,jk} & \sigma_{Z,jl} \\ \sigma_{Z,ki} & \sigma_{Z,kj} & \sigma_{Z,kk} & \sigma_{Z,kl} \\ \sigma_{Z,li} & \sigma_{Z,lj} & \sigma_{Z,lk} & \sigma_{Z,ll} \end{pmatrix},
$$

$$
\Sigma_Z = [\sigma_{Z,ij}] = E\left[(Z_t - \mu_z)(Z_t - \mu_z)'\right],
$$

$$
\Sigma_Z(r) = [\sigma_{Z,ij}(r)] = E\left[(Z_t - \mu_z)(Z_{t-r} - \mu_z)'\right],
$$

and

$$
\Sigma_{Z,\{(i,j),(k,l)\}}(-u) = \begin{pmatrix} \sigma_{Z,ik}(-u) & \sigma_{Z,il}(-u) \\ \sigma_{Z,jk}(-u) & \sigma_{Z,jl}(-u) \end{pmatrix}.
$$

The proofs of (S2.3) and (S2.4) are sketched briefly as follows. Because r_t is a stationary sequence of martingale differences in the form of the weighted sum of products of two independent processes , (S2.3) follows from the martingale central limit theorem (cf. Theorem 3.2 of Hall and Heyde (1980)). For (S2.4), recall from Section 3 that $r_t = \mu + \sum_{i=1}^m w_i v_{i,t} u_{i,t}$, and define $\rho_{V,ij} = E v_{i,t} v_{j,t}$, $\rho_{U,ij} = E u_{i,t} u_{j,t}$,

$$
W_{t,1} = \sum_{i=1}^{m} w_i^2 \left(v_{i,t}^2 - \rho_{V,ii}\right) + 2 \sum_{1 \le i < j \le m} w_i w_j \rho_{U,ij} \left(v_{i,t} v_{j,t} - \rho_{V,ij}\right),
$$

and

$$W_{t,2} = \sum_{i=1}^{m} w_i^2 v_{i,t}^2 \left(u_{i,t}^2 - 1\right) + 2 \sum_{1 \le i < j \le m} w_i w_j v_{i,t} v_{j,t} \left(u_{i,t} u_{j,t} - \rho_{U,ij}\right).$$

We can then write

$$\begin{aligned} \sqrt{n}\left(\hat{\sigma}_n^2 - \sigma^2\right) &= \sum_{t=1}^{n} W_{t,1}/\sqrt{n} + \sum_{t=1}^{n} W_{t,2}/\sqrt{n} + O_p\left(1/\sqrt{n}\right) \\ &\equiv \Pi_{n,1} + \Pi_{n,2} + O_p\left(1/\sqrt{n}\right). \end{aligned}$$

To prove (S2.4), we first need to show the marginal central limit theorem for both $\Pi_{n,1}$ and $\Pi_{n,2}$. The former is straightforward since it is also the sum of stationary martingale differences. For the latter, we employ the same ℓ-dependence method as used in Ho and Hsing (1997) and approximate $W_{t,2}$ by $W_{t,2,\ell}$ of which the terms $v_{i,t} = h_i(Z_{i,t})$'s are replaced by $v_{i,t,\ell} = h_i(Z_{i,t,\ell})$ (cf. (2.2)). Then the asymptotic normality of $\Pi_{n,2}$ follows since $\Pi_{n,2,\ell} \equiv \sum_{t=1}^{n} W_{t,2,\ell}/\sqrt{n}$ is asymptotically normal due to ℓ-dependence and the fact that $\Pi_{n,2}$ and $\Pi_{n,2,\ell}$ are asymptotically close in the L_2 sense of

$$\lim_{\ell \to \infty} \limsup_{n \to \infty} E\left(\Pi_{n,2} - \Pi_{n,2,\ell}\right)^2 = 0$$

(cf. Theorem 3.2. of Ho and Hsing (1997)), which requires the use of condition J. Second, with regard to the joint weak convergence of $(\Pi_{n,1}, \Pi_{n,2})$, applying the Cramér-Wold device will do the job since $\Pi_{n,1}$ and $\Pi_{n,2}$ are uncorrelated.

Lemma 2. Under the same assumptions as in Lemma 1, for $0 < \alpha < 1$,

$$q_\alpha(T) = \sqrt{T}\sigma\Phi^{-1}\left(\alpha + O(T^{-1})\right) + T\mu \tag{S2.6}$$

as $T \to \infty$.

Proof of Lemma 2. Recall that $R_T = R_{T,0}$. Let $\sigma_T^* = (T^{-1}\sum_{t=1}^{T}\sum_{i,j=1}^{m} w_i w_j v_{i,t} v_{j,t} \sigma_{U,ij})^{1/2}$. $A_T = (q_\alpha(T) - T\mu)\big/\sqrt{T}$, $R_T^* = (R_T - T\mu)\big/(\sqrt{T}\sigma_T^*)$, and $\mathcal{V}_T$ be the information σ-field generated by $\{Z_1, Z_2, ..., Z_T\}$. Since $\alpha = P\left(R_T^* < A_T/\sigma_T^*\right)$, the central limit theorem of (S2.3) and σ_T^* converging to σ with probability one imply that $\lim_{T\to\infty} A_T = \Phi^{-1}(\alpha)\sigma$. Then, for any positive ε such that $\varepsilon < \sigma^2$,

$$\begin{aligned}
\alpha &= E\left[P\left(R_T^* < A_T/\sigma_T^* \middle| \mathcal{V}_T\right)\right] \\
&= \Phi\left(\frac{A_T}{\sigma}\right) + E\left[\Phi\left(\frac{A_T}{\sigma_T^*}\right) - \Phi\left(\frac{A_T}{\sigma}\right)\right] \\
&= \Phi\left(\frac{A_T}{\sigma}\right) + E\left[\left(\Phi\left(\frac{A_T}{\sigma_T^*}\right) - \Phi\left(\frac{A_T}{\sigma}\right)\right) I\left(\left|\sigma_T^{*2} - \sigma^2\right| > \varepsilon\right)\right] \\
&\quad + E\left[\left(\Phi\left(\frac{A_T}{\sigma_T^*}\right) - \Phi\left(\frac{A_T}{\sigma}\right)\right) I\left(\left|\sigma_T^{*2} - \sigma^2\right| \le \varepsilon\right)\right] \\
&= \Phi\left(\frac{A_T}{\sigma}\right) + E_{T,1} + E_{T,2}.
\end{aligned} \tag{S2.7}$$

where $I(S)$ is the indicator function of the event S. Using $|\Phi(x) - \Phi(y)| \le 1$ and the Chebyshev inequality, we have

$$|E_{T,1}| \le \varepsilon^{-2} E\left(\sigma_T^{*2} - \sigma^2\right)^2 = O(T^{-1}). \tag{S2.8}$$

Set $K(x) = \Phi(A_T/\sqrt{x})$. Taylor's expansion of function $K(\cdot)$ with respect to $(\sigma_T^{*2} - \sigma^2)$ enables us to express $E_{T,2}$ as

$$\begin{aligned}
E_{T,2} &= \left(-A_T\phi(A_T/\sigma)/(2\sigma^3)\right) \cdot E\left[\left(\sigma_T^{*2} - \sigma^2\right) + \left(\sigma_T^{*2} - \sigma^2\right)\left(I\left(\left|\sigma_T^{*2} - \sigma^2\right| \le \varepsilon\right) - 1\right)\right] \\
&\quad + (1/2)E\left[\left(\sigma_T^{*2} - \sigma^2\right)^2 K''(A^*) I\left(\left|\sigma_T^{*2} - \sigma^2\right| \le \varepsilon\right)\right] \\
&= \left(A_T\phi(A_T/\sigma)/(2\sigma^3)\right) \cdot E\left[\left(\sigma_T^{*2} - \sigma^2\right) I\left(\left|\sigma_T^{*2} - \sigma^2\right| > \varepsilon\right)\right] \\
&\quad + (1/2)E\left[\left(\sigma_T^{*2} - \sigma^2\right)^2 K''(A^*) I\left(\left|\sigma_T^{*2} - \sigma^2\right| \le \varepsilon\right)\right] \\
&= E_{T,2,1} + E_{T,2,2},
\end{aligned}$$

where A^* lies between σ^2 and σ_T^{*2}. Now

$$|E_{T,2,1}| \le C \cdot \left(E\left(\sigma_T^{*2} - \sigma^2\right)^2\right)^{1/2} P^{1/2}\left(\left|\sigma_T^{*2} - \sigma^2\right| > \varepsilon\right) = O(T^{-1}),$$

and, because $K''(A^*) I\left(\left|\sigma_T^{*2} - \sigma^2\right| \le \varepsilon\right)$ is bounded,

$$|E_{T,2,2}| \le C \cdot E\left(\sigma_T^{*2} - \sigma^2\right)^2 = O(T^{-1}).$$

Hence

$$|E_{T,2}| = O(T^{-1}). \tag{S2.9}$$

Then (S2.6) follows from (S2.7), (S2.8) and (S2.9).

Proof of Proposition 1. Let $U_T = A_T/\sigma_T^*$, where A_T is defined previously as $(q_\alpha(T) - T\mu)\Big/\sqrt{T}$. Then (S2.6) implies

$$U_T = \Phi^{-1}(\alpha + O(T^{-1}))\frac{\sigma}{\sigma_T^*}. \tag{S2.10}$$

Define

$$\begin{aligned} H(u) &= (\Phi(u))^{-1}\int_{-\infty}^{u} x\phi(x)dx \\ &= (\Phi(u))^{-1}(-\phi(u)), \end{aligned} \tag{S2.11}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the pdf and cdf of the standard normal distribution, respectively. It can be seen that $H\left(\Phi^{-1}(\alpha)\right)$ is precisely the α-th CTE of standard normal Z defined by

$$Z_\alpha = E\left(Z\middle|Z < \Phi^{-1}(\alpha)\right) = \frac{-\phi(\Phi^{-1}(\alpha))}{\alpha}.$$

Let $G_T(\alpha) = A_T/\sigma$ which is $\Phi^{-1}(\alpha + O(T^{-1}))$ according to Lemma 2. By the same argumet of conditional on the σ-field $\mathcal{V}_T$ as employed in proving (S2.7), the α-th CTE of R_T is

$$\begin{aligned} C_\alpha^T &= E(R_T|R_T < q_\alpha(T)) \\ &= T\mu + \sqrt{T}\sigma Z_\alpha + E\left\{\sqrt{T}\left[\sigma_T^* H(U_T) - \sigma H(G_T(\alpha))\right]\right\} \\ &\quad +\sqrt{T}\sigma\left[H(G_T(\alpha)) - H(\Phi^{-1}(\alpha))\right] \\ &= T\mu + \sqrt{T}\sigma Z_\alpha + B_T + B_T'. \end{aligned} \tag{S2.12}$$

For any positive $\varepsilon < \sigma^2$, write

$$\begin{aligned} B_T &= E\left\{\sqrt{T}\left[\sigma_T^* H(U_T) - \sigma H(G_T(\alpha))\right]\cdot\left(I\left(\sigma_T^{*2} > \sigma^2 + \varepsilon\right) + I\left(\sigma_T^{*2} < \sigma^2 - \varepsilon\right) + I\left(\left|\sigma_T^{*2} - \sigma^2\right| \le \varepsilon\right)\right)\right\} \\ &= B_{T,1} + B_{T,2} + B_{T,3}. \end{aligned}$$

We further express $B_{T,1}$ as

$$
\begin{aligned}
B_{T,1} &= E\sqrt{T}(\sigma_T^* - \sigma)H(U_T) \cdot I\left(\sigma_T^{*2} > \sigma^2 + \varepsilon\right) + E\sqrt{T}\sigma\left[H(U_T) - H(G_T(\alpha))\right] \cdot I\left(\sigma_T^{*2} > \sigma^2 + \varepsilon\right) \\
&= B_{T,1,1} + B_{T,1,2}
\end{aligned}
$$

Since $H(U_T) \cdot I\left(\sigma_T^{*2} > \sigma^2 + \varepsilon\right)$ is bounded,

$$
\begin{aligned}
|B_{T,1,1}| &\leq CE\sqrt{T}\,|\sigma_T^* - \sigma| \cdot I\left(\sigma_T^{*2} > \sigma^2 + \varepsilon\right) \\
&\leq C\left\{E\left[\sqrt{T}(\sigma_T^* - \sigma)\right]^2\right\}^{1/2} \cdot P^{1/2}\left(\left|\sigma_T^{*2} - \sigma^2\right| > \varepsilon\right) \\
&= O(1/\sqrt{T}). \qquad \text{(S2.13)}
\end{aligned}
$$

Because $|H(U_T) - H(G_T(\alpha))| \cdot I\left(\sigma_T^{*2} > \sigma^2 + \varepsilon\right)$ is also bounded, we have

$$
\begin{aligned}
|B_{T,1,2}| &\leq CE\sqrt{T} \cdot I\left(\sigma_T^{*2} > \sigma^2 + \varepsilon\right) \\
&\leq C\sqrt{T}P\left(\left|\sigma_T^{*2} - \sigma^2\right| > \varepsilon\right) \\
&= O(1/\sqrt{T}). \qquad \text{(S2.14)}
\end{aligned}
$$

To handle $B_{T,2}$, because $\sigma H(G_T(\alpha))$ is clearly bounded, we focus on $\sigma_T^* H(U_T) \cdot I\left(\sigma_T^{*2} < \sigma^2 - \varepsilon\right)$. We first note that for $c > 0$

$$
\lim_{x \to 0} xH(c/x) = c, \qquad \text{(S2.15)}
$$

which is due to the fact that $\Phi(z) \approx \phi(z)/|z|$ as $z \to -\infty$ (Chapter 7 of Abramowitz and Stegun, 1972). Then

$$
\sigma_T^* H(U_T) \cdot I\left(\sigma_T^{*2} < \sigma^2 - \varepsilon\right) = \sigma_T^* H\left(\sigma\Phi^{-1}(\alpha + O(T^{-1}))/\sigma_T^*\right) I\left(\sigma_T^{*2} < \sigma^2 - \varepsilon\right)
$$

which by (S2.15) is bounded with probability one. Therefore

$$
\begin{aligned}
|B_{T,2}| &\le C\sqrt{T}P\left(\sigma_T^{*2} < \sigma^2 - \varepsilon\right) \\
&\le C\sqrt{T}P\left(|\sigma_T^{*2} - \sigma^2| > \varepsilon\right) \\
&\le C\sqrt{T}E(\sigma_T^{*2} - \sigma^2)^2 \\
&= O(1/\sqrt{T}). \qquad \text{(S2.16)}
\end{aligned}
$$

To deal with $B_{T,3}$, let $\tilde{H}(x) = \sqrt{x}H(\Phi^{-1}(\alpha + O(T^{-1}))/\sqrt{x})$ and expand $B_{T,3}$ as

$$
\begin{aligned}
B_{T,3} &= \sqrt{T}E\left[\tilde{H}(\sigma_T^{*2}) - \tilde{H}(\sigma^2)\right] \cdot I\left(\left|\sigma_T^{*2} - \sigma^2\right| \le \varepsilon\right) \\
&= \sqrt{T}\tilde{H}'(\sigma^2)E\left[\left(\sigma^{*2} - \sigma^2\right) + \left(\sigma^{*2} - \sigma^2\right)\left(I\left(\left|\sigma_T^{*2} - \sigma^2\right| \le \varepsilon\right) - 1\right)\right] \\
&\quad + (\sqrt{T}/2)E\left[\left(\sigma_T^{*2} - \sigma^2\right)^2 \tilde{H}''(U^*)I\left(\left|\sigma_T^{*2} - \sigma^2\right| \le \varepsilon\right)\right] \\
&= \sqrt{T}\tilde{H}'(\sigma^2)E\left[-\left(\sigma_T^{*2} - \sigma^2\right)I\left(\left|\sigma_T^{*2} - \sigma^2\right| > \varepsilon\right)\right] \\
&\quad + (\sqrt{T}/2)E\left[\left(\sigma_T^{*2} - \sigma^2\right)^2 \tilde{H}''(U^*)I\left(\left|\sigma_T^{*2} - \sigma^2\right| \le \varepsilon\right)\right],
\end{aligned}
$$

where U^* lies between σ^2 and σ_T^{*2}. Using Holder's and Chebyshev's inequalities and the bounded $\tilde{H}''(U^*)I\left(\left|\sigma_T^{*2} - \sigma^2\right| \le \varepsilon\right)$, we can bound $B_{T,3}$ by

$$
\begin{aligned}
|B_{T,3}| &\le C\left(P^{1/2}\left(\left|\sigma_T^{*2} - \sigma^2\right| > \varepsilon\right) + \sqrt{T}E\left(\sigma_T^{*2} - \sigma^2\right)^2\right) \\
&= O(1/\sqrt{T}). \qquad \text{(S2.17)}
\end{aligned}
$$

According to the mean value theorem, we have $H(G_T(\alpha)) - H(\Phi^{-1}(\alpha)) = O(T^{-1})$, which yields

$$
\left|B_T'\right| = O(1/\sqrt{T}). \qquad \text{(S2.18)}
$$

Combining (S2.12), (S2.13), (S2.14), (S2.16), (S2.17) and (S2.18) gives

$$
C_\alpha^T = T\mu + \sqrt{T}\sigma Z_\alpha + O(1/\sqrt{T}). \qquad \text{(S2.19)}
$$

Since $n = N \cdot T$, the estimation error of the non-parametric estimate $\widehat{C}_\alpha^T = T\mu^* + \sqrt{T}\hat{\sigma}_n Z_\alpha$

proposed in (3.4) for C_α^T is, by (S2.19),

$$\hat{C}_\alpha^T - C_\alpha^T = \begin{cases} \sqrt{T/N}\sqrt{n}(\hat{\mu}_n - \mu) + N^{-1/2}\sqrt{n}(\hat{\sigma}_n - \sigma) + O(T^{-1/2}) & \text{if } \mu \text{ is unknown,} \\ N^{-1/2}\sqrt{n}(\hat{\sigma}_n - \sigma) + O(T^{-1/2}) & \text{if } \mu \text{ is known.} \end{cases}$$

Equations (3.5) and (3.6) then follow from Lemma 1 and the conditions imposed on T and N immediately.

S3 Comparison with two sample generation methods

To demonstrate the superiority of our approach, we choose two sample generation methods, the Markov-chain simulation (MCS) method and the sampling window (SW) method, for the performance comparison in terms of empirical coverage ratios. The MCS and SW represent two typical methods that form the category of the second approach mentioned in the Introduction. Due to the massive amount of computation time needed for carrying out the MCS, we restrict our attention to the standard univariate SV model and only a moderate set of representative cases on (T, N) that suffices to show the advantages of the estimator we propose. The standard univariate SV model is defined as $r_t = \bar{\sigma}\exp\{Z_t/2\}u_t$, where Z_t is AR(1) normal and u_t is iid N(0,1), and the parameters of the AR(1) model are the same as those of the marginal return components used in Subection 4.1. The specific procedures of the two methods, the MCS and the SW, are as follows.

For the MCS method, we first simulate $\{r_t\}_{t=1,\dots,n}$ from the true data generating process and then follow the Markov-chain simulation procedures outlined by Jacquier, Polson and Rossi (1994) to estimate the parameters as they demonstrate that the MCS method is the best as compared to the quasi-maximum likelihood and the method of moments. Furthermore, we use the refined Bayesian MCMC sampler from Kastner and Frühwirth-Schnatter (2014) which is implemented in the R package 'stochvol' (Kastner (2016)) and is available at https://CRAN.R-project.org/package=stochvol. Second, based on these estimated parameters, we simulate $\{r_t\}_{t=1,\ldots,T}$ 1000 times to compute 1000 integrated returns $\{R_T^1, \ldots, R_T^{1000}\}$ from which one $\widehat{C}_\alpha^T$ is derived. Third, repeat the last step for 1000 times to generate 1000 $\widehat{C}_\alpha^T$'s and then use the 0.025 and 0.975 percentiles of these 1000 $\widehat{C}_\alpha^T$'s to form a 95% confidence interval. Finally, repeat the above three steps to build 1000 confidence intervals to compute the empirical coverage ratio.

As an alternative to the model-dependent approach illustrated above, we consider the sampling window (SW) method described in Subection 3.1 following the statement of Proposition 2, which is model-free, also relies on generation of samples, and is widely used in handling dependent data (Politis et al., 1999). We break the procedure into five steps. First, recalling that $n = NT$, simulate $\{r_t\}_{t=1,\ldots,n}$ from the true data generating process. Second, divide the sample into overlapping blocks each consisting of T returns and every two consecutive blocks are 10 observations apart. The overlapping of observations is designed to meet the purpose of generating sufficiently many blocks of returns to be integrated. Then, we compute the i-th integrated return $R_T^{(i)}$ from the i-th block, $\{r_{1+(i-1)10},\ r_{2+(i-1)10},\ \ldots,\ r_{T+(i-1)10}\}$, where i ranges from 1 to $M = \lfloor (N-1)T/10 + 1 \rfloor$. Here $\lfloor x \rfloor$ stands for the greatest integer less than or equal to x. Third, treat the M integrated returns $\{R_T^{(1)}, \ldots, R_T^{(M)}\}$ as a sample of stationary observations from which $M - M_4 + 1$ subsamples of integrated returns are generated by the SW method with window size of $M_4 = \lfloor M/4 \rfloor$, that is, the list of all the subsamples is $\{R_T^{(j)}, \ldots, R_T^{(j-1+M_4)}\}$ with j ranging from 1 to $M - M_4 + 1$. Then we construct $M - M_4 + 1$ CTE's of which the j-th CTE is based on its corresponding j-th subsample $\{R_T^{(j)}, \ldots, R_T^{(j-1+M_4)}\}$. Fourth, similar to the MCS method, the 0.025 and 0.975 percentiles of the $M - M_4 + 1$ CTE's are respectively used as the left and right limits of the 95% confidence interval. Finally, repeat the above four steps 1000 times to build 1000 confidence intervals, and calculate the coverage ratios. Only the case of unknown mean is reported because it is more realistic in practice to assume that the mean is unknown. To better illustrate the strength of our approach, we choose $T = 168, 210$, and 252 for the horizon, which are longer than those in Subection 4.1 and correspond to eight, ten and twelve months of trading days, respectively. For each T, the number N of blocks is then determined by a pre-given sample size $n = NT$; we consider $n = 4200, 6300, 8400$ and 12600. The true C_α^T is computed by simulating 10^6 price paths from the true model and $\alpha = 0.01$

The results that are summarized in Table 1 exhibit some notable features. First of all, our approach (labeled as E) clearly outperforms the other two methods, the MCS and the SW. Across all the different choices of n, T and N, the coverage ratios of the E do not vary much and remain close to the nominal level. However, those of the MCS and the SW are significantly less than .95 except for a few cases, and the SW in particular suffers huge under coverage. The increasing trend of the coverage ratios with n for both the MCS and the SW indicates that the two sample-generation methods require very large sample sizes in order to overcome the too strong dependence created by the aggregation. The MCS can achieve reasonably good coverage only when $n = 12600$. It is also important to note that for a fixed sample size n, the coverage decreases as the integration length increases. For example, when $n = 6300$ (or 4200), the coverage ratios of the MCS drop from 0.913 (or 0.829) to 0.846 (or 0.724) as the return horizon increases from 168 to 252 days. This serves as further evidence that our approach is more suitable than the two sample-generation methods for estimating the CTE of integrated returns. It is also worth mentioning that the computing time the MSC method needs is about 700 times more than our non-parametric method.

Table 1: Comparison of the coverage ratios of 95% confidence intervals for C_α^T based on equation (3.10)(labeled as E) with two sample-generation method: the sampling window method (SW) and the Markov-chain simulation (MCS) method, with the parameters being estimated by the Bayesian appraoch of Jacquier, Polson and Rossi (1994). The results are based on 1000 replicates, and the true C_α^T is computed by simulating 10^6 price paths from the true model. $T = 168$, 210, and 252, and $\alpha = 0.01$.

	T		168			210			252	
n		E	SW	MCS	E	SW	MCS	E	SW	MCS
	N		25			20			17	
4200		0.944	0.196	0.829	0.952	0.170	0.804	0.942	0.144	0.724
	N		38			30			25	
6300		0.947	0.349	0.913	0.939	0.342	0.875	0.945	0.325	0.846
	N		50			40			33	
8400		0.954	0.448	0.932	0.937	0.411	0.926	0.958	0.365	0.893
	N		75			60			50	
12600		0.944	0.619	0.990	0.946	0.541	0.966	0.939	0.539	0.955

Bibliography

Abramowitz, M. and Stegun, I.A. (1972). *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables (10th printing) 55*. National Bureau of Standards, Applied Mathematics Series.

Hall, P. and Heyde, C. C. (1980). *Martingale limit theory and its application*. Academic Press, New York.

Ho, H.-C. and Hsing, T. (1997). Limit theorems for functionals of moving averages. *The Annals of Probability 25*, 1636-1669.

Jacquier, E., Polson, N. G. and Rossi, P. E. (1994). Bayesian analysis of stochastic volatility models. *Jounal of Business & Economic Statistics 12*, 371-389.

Kastner, G. (2016). Dealing with Stochastic Volatility in Time Series Using the R Package stochvol. *Journal of Statistical Software 69*, 1-30.

Kastner, G. and Frühwirth-Schnatter, S. (2014). Ancillarity-sufficiency interweaving strategy (ASIS) for boosting MCMC estimation of stochastic volatility models. *Computational Statistics & Data Analysis 76*, 408-423.

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