Online supplementary material for "Estimating the VaR-induced Euler allocation rule": lemmas and proofs

N.V. Gribkova [*,1,2], J. Su [†,3,4], and R. Zitikis [‡,4,5]

[1]Saint Petersburg State University, Saint Petersburg, 199034 Russia

[2]Saint Petersburg State Transport University, Saint Petersburg, 190031 Russia

[3]Purdue University, West Lafayette, Indiana 47907, United States

[4]Risk and Insurance Studies Centre, York University, Toronto, Ontario M3J 1P3, Canada

[5]Western University, London, Ontario N6A 5B7, Canada

Abstract. We present proofs of the results formulated and discussed in the main body of the article, together with accompanying technical lemmas and their proofs.

S.1 Proofs

To prove Theorem 2.1, we need a lemma.

Lemma S.1.1. *Suppose that conditions* (C_1) *and* (D_1) *are satisfied. Then for* $k_{1,n}$ *and* $k_{2,n}$ *defined by equations* (2.4), *we have*

$$\mathbb{P}\big(Y_{k_{j,n}:n} \notin W_\varepsilon\big) \to 0, \quad j \in \{1,2\}, \tag{S.1}$$

when $n \to \infty$, *where*

$$W_\varepsilon = \big(\mathrm{VaR}_{p-\varepsilon}(Y), \mathrm{V@R}_{p+\varepsilon}(Y)\big). \tag{S.2}$$

[*]e-mail: n.gribkova@spbu.ru

[†]Corresponding author; e-mail: jianxi@purdue.edu

[‡]e-mail: rzitikis@uwo.ca

Proof. We only prove statement (S.1) for $k_{2,n}$ because analogous arguments work for $k_{1,n}$ as well. We start with the equation

$$\mathbb{P}\big(Y_{k_{2,n}:n} \notin W_\varepsilon\big) = \mathbb{P}\big(Y_{k_{2,n}:n} \le \mathrm{VaR}_{p-\varepsilon}(Y)\big) + \mathbb{P}\big(Y_{k_{2,n}:n} \ge \mathrm{V@R}_{p+\varepsilon}(Y)\big). \tag{S.3}$$

We shall next show that

$$\mathbb{P}\big(Y_{k_{2,n}:n} \le \mathrm{VaR}_{p-\varepsilon}(Y)\big) \to 0 \tag{S.4}$$

when $n \to \infty$, and analogous arguments can be used to show that the right-most probability in equation (S.3) also converges to 0. Hence, we shall only prove statement (S.4) and start with

$$\begin{aligned}
\mathbb{P}\big(Y_{k_{2,n}:n} \le \mathrm{VaR}_{p-\varepsilon}(Y)\big) &= \mathbb{P}\big(\#\{i : Y_i \le \mathrm{VaR}_{p-\varepsilon}(Y)\} \ge k_{2,n}\big) \\
&= \mathbb{P}\big(nG_n(\mathrm{VaR}_{p-\varepsilon}(Y)) \ge [n(p+\Delta_{2,n})]\big) \\
&\le \mathbb{P}\big(nG_n(\mathrm{VaR}_{p-\varepsilon}(Y)) \ge n(p+\Delta_{2,n}) - 1\big) \\
&= \mathbb{P}\big(G_n(\mathrm{VaR}_{p-\varepsilon}(Y)) \ge (p+\Delta_{2,n}) - 1/n\big) \\
&\le \mathbb{P}\big(G(\mathrm{VaR}_{p-\varepsilon}(Y)) \ge (p+\Delta_{2,n}) - \kappa_n - 1/n\big),
\end{aligned} \tag{S.5}$$

where

$$\kappa_n = \sup_{y \in \mathbb{R}} |G_n(y) - G(y)|.$$

By the Glivenko-Cantelli theorem, κ_n converges to 0 almost surely, and hence in probability. By condition (D_1), $\Delta_{2,n}$ converges to 0. Consequently, the right-hand side of bound (S.5) converges to 0 because $G(\mathrm{VaR}_{p-\varepsilon}(Y)) < p$ due to G being strictly increasing to the left of $\mathrm{VaR}_p(Y)$ (see Figure 2.1). This finishes the proof of Lemma S.1.1. □

Proof of Theorem 2.1. For any $\delta > 0$, Lemma S.1.1 implies that, when $n \to \infty$,

$$\mathbb{P}\Big(\big|\mathrm{EAR}_{p,n} - \mathrm{EAR}_p(X \mid Y)\big| > \delta\Big) = \iint_{W_\varepsilon \times W_\epsilon} \Pi(y_1, y_2)\mathrm{d}F_{k_{1,n},k_{2,n}}(y_1, y_2) + o(1) \tag{S.6}$$

where $F_{k_{1,n},k_{2,n}}$ denotes the joint cdf of the order statistics $Y_{k_{1,n}-1:n}$ and $Y_{k_{2,n}+1:n}$, and

$$\Pi(y_1, y_2) = \mathbb{P}\Big(\big|\mathrm{EAR}_{p,n} - \mathrm{EAR}_p(X \mid Y)\big| > \delta \mid Y_{k_{1,n}-1:n} = y_1, Y_{k_{2,n}+1:n} = y_2\Big).$$

We have the bound

$$\Pi(y_1,y_2) \le \mathbb{P}\Big(\big|\mathrm{EAR}_{p,n} - \mathbb{E}\big(X \mid Y \in (y_1,y_2]\big)\big| > \delta/2 \mid Y_{k_{1,n}-1:n} = y_1, Y_{k_{2,n}+1:n} = y_2\Big)$$
$$+\, \mathbb{P}\Big(\big|\mathbb{E}\big(X \mid Y \in (y_1,y_2]\big) - \mathrm{EAR}_p(X \mid Y)\big| > \delta/2 \mid Y_{k_{1,n}-1:n} = y_1, Y_{k_{2,n}+1:n} = y_2\Big). \tag{S.7}$$

No matter what $y_1, y_2 \in W_\varepsilon$ are, as long as $\varepsilon > 0$ is sufficiently small, the event inside the right-most probability is impossible due to the continuity of the regression function g in a neighborhood of $\mathrm{VaR}_p(Y)$. Hence, the aforementioned probability vanishes. We are left to show that the penultimate probability on the right-hand side of bound (S.7) converges to 0 for all $y_1, y_2 \in W_\varepsilon$. The Lebesgue dominated convergence theorem will then finish the proof that the integral on the right-hand side of equation (S.6) converges to 0.

To better understand the event inside the penultimate probability on the right-hand side of bound (S.7), we write

$$\mathrm{EAR}_{p,n} - \mathbb{E}\big(X \mid Y \in (y_1,y_2]\big) = \frac{1}{k_{2,n}-k_{1,n}+1} \sum_{i=k_{1,n}}^{k_{2,n}} \Big(X_{i,n} - \mathbb{E}\big(X \mid Y \in (y_1,y_2]\big)\Big). \tag{S.8}$$

Conditionally on $Y_{k_{1,n}-1:n} = y_1$ and $Y_{k_{2,n}+1:n} = y_2$, the order statistics $Y_{k_{1,n}:n}, \dots, Y_{k_{2,n}:n}$ are distributed (e.g., Arnold et al., 2008; David and Nagaraja, 2003) as the order statistics $\widetilde{Y}_{1:k_{2,n}-k_{1,n}+1}, \dots, \widetilde{Y}_{k_{2,n}-k_{1,n}+1:k_{2,n}-k_{1,n}+1}$ of a sample $\widetilde{Y}_1, \dots, \widetilde{Y}_{k_{2,n}-k_{1,n}+1}$ of size $(k_{2,n}-k_{1,n}+1)$ from the distribution

$$\widetilde{G}(y) = \begin{cases} 0, & y < y_1, \\ \frac{G(y)-G(y_1)}{G(y_2)-G(y_1)}, & y_1 \le y < y_2, \\ 1, & y \ge y_2. \end{cases} \tag{S.9}$$

The following two notes give technical insights into the definition of $\widetilde{G}$ and in this way facilitate further steps that will complete the proof of Theorem 2.1.

Note S.1.1 (Ignoring the gap)**.** It is possible that there can be a region inside the interval W_ε (recall its definition (S.2)) where the cdf G is horizontal (see Figure 2.1). If both y_1 and y_2 fall into the region, then $G(y_2) - G(y_1) = 0$ and thus an issue with the definition of $\widetilde{G}$ arises. However, this is not a problem because when $G(y_2) - G(y_1) = 0$, the random variable Y does not (almost surely) place any points between y_1 and y_2, and such cases can therefore be ignored. Indeed, conditional expectations are always defined up to sets of measure 0.

Note S.1.2 (Sewing the gap)**.** In the previous note we argued that the gap $\big(\mathrm{VaR}_p(Y), \mathrm{V@R}_p(Y)\big)$ in the distribution of Y-values can be ignored. An alternative approach would be to "sew" the gap, which means replacing the original random variables Y_i by

$$Y_i^{\text{sewn}} = Y_i \mathbb{1}_{\{Y_i \le \mathrm{VaR}_p(Y)\}} + (Y_i - A_p)\mathbb{1}_{\{Y_i > \mathrm{V@R}_p(Y)\}},$$

where the size of the gap A_p is defined by

$$A_p = \mathrm{V@R}_p(Y) - \mathrm{VaR}_p(Y).$$

Obviously, if $A_p = 0$, then $Y_i^{\text{sewn}} = Y_i$, but irrespective of whether $A_p = 0$ or $A_p > 0$, the cdf of Y_i^{sewn} is strictly increasing in a neighbourhood of $\mathrm{VaR}_p(Y)$. It remains to observe that the order statistics arising from Y_i^{sewn}'s are the same as those arising from Y_i's, and so when instead of the original pairs (X_i, Y_i) we work with the pairs (X_i, Y_i^{sewn}), $i = 1, \dots, n$, the gap-related issues disappear, because the gap vanishes, while all the concomitant-based estimators remain the same.

Continuing the proof of Theorem 2.1, we note that the random variables $X_{k_{1,n},n}, \dots, X_{k_{2,n},n}$ are distributed as concomitants corresponding to the order statistics $Y_{k_{1,n}:n}, \dots, Y_{k_{2,n}:n}$. Hence, conditionally on $Y_{k_{1,n}-1:n} = y_1$ and $Y_{k_{2,n}+1:n} = y_2$, the right-hand side of equation (S.8) is distributed as the average of centered i.i.d. random variables. We can now streamline the rest of the proof.

Let $Z_{k_{1,n}}, \dots, Z_{k_{2,n}}$ be independent copies of a random variable Z whose cdf $F_Z(x) = \mathbb{P}(Z \le x)$ is given by

$$F_Z(x) = \mathbb{P}\big(X \le x \mid Y \in (y_1, y_2]\big). \tag{S.10}$$

For any fixed $y_1, y_2 \in W_\varepsilon$ and conditionally on $Y_{k_{1,n}-1:n} = y_1$ and $Y_{k_{2,n}+1:n} = y_2$, the right-hand side of equation (S.8) has the same distribution as

$$\zeta_n := \sum_{i=k_{1,n}}^{k_{2,n}} \xi_{i,n},$$

where

$$\xi_{i,n} = \frac{Z_i - \mathbb{E}(Z_i)}{k_{2,n} - k_{1,n} + 1}.$$

We have an array $\xi_{k_{1,n},n}, \dots, \xi_{k_{2,n},n}$, $n \ge 1$, of random variables, which are independent for every given n, and they are all identically distributed irrespective of n. Hence, when n changes, only the number $k_{2,n} - k_{1,n} + 1$ of random variables changes, but not their distributions. Note that by condition (D$_2$), we have

$$\begin{aligned} k_{2,n} - k_{1,n} + 1 &= [n(p + \Delta_{2,n})] - [n(p - \Delta_{1,n})] + 1 \\ &\sim n\big(\Delta_{2,n} + \Delta_{1,n}\big) \to \infty \end{aligned}$$

when $n \to \infty$. Our task becomes to show that, when $n \to \infty$,

$$\zeta_n \xrightarrow{\mathbb{P}} 0. \tag{S.11}$$

For this, we follow the arguments of Borovkov (1988, pp. 217–219).

We first split each summand of ζ_n into two parts: $\xi_{i,n} = \xi'_{i,n} + \xi''_{i,n}$, where $\xi'_{i,n} = \xi_{i,n}\mathbb{1}_{\{|\xi_{i,n}| \le 1\}}$ and $\xi''_{i,n} = \xi_{i,n}\mathbb{1}_{\{|\xi_{i,n}| > 1\}}$. Hence,

$$\zeta_n = \zeta'_n + \zeta''_n := \sum_{i=k_{1,n}}^{k_{2,n}} \xi'_{i,n} + \sum_{i=k_{1,n}}^{k_{2,n}} \xi''_{i,n}$$

and so

$$\begin{aligned} \mathbb{E}(|\zeta_n|) &\le \mathbb{E}(|\zeta'_n - \mathbb{E}\zeta'_n|) + \mathbb{E}(|\zeta''_n - \mathbb{E}\zeta''_n|) \\ &\le \sqrt{\mathbb{V}(\zeta'_n)} + 2\mathbb{E}(|\zeta''_n|) \\ &\le \sqrt{\sum_{i=k_{1,n}}^{k_{2,n}} \mathbb{E}((\xi'_{i,n})^2)} + 2\sum_{i=k_{1,n}}^{k_{2,n}} \mathbb{E}(|\xi''_{i,n}|) \\ &\le \sqrt{D_n} + 2D_n, \end{aligned} \tag{S.12}$$

where

$$\begin{aligned} D_n &= \sum_{i=k_{1,n}}^{k_{2,n}} \mathbb{E}\big(|\xi'_{i,n}|\big) + \sum_{i=k_{1,n}}^{k_{2,n}} \mathbb{E}\big((\xi''_{i,n})^2\big) \\ &= \sum_{i=k_{1,n}}^{k_{2,n}} \mathbb{E}\Big(\min\big\{|\xi_{i,n}|, \xi_{i,n}^2\big\}\Big). \end{aligned}$$

For any $0 < \lambda \le 1$, we write

$$D_n \le \sum_{i=k_{1,n}}^{k_{2,n}} \mathbb{E}\Big(|\xi_{i,n}|\mathbb{1}_{\{|\xi_{i,n}|>\lambda\}}\Big) + \sum_{i=k_{1,n}}^{k_{2,n}} \mathbb{E}\Big(\xi_{i,n}^2\mathbb{1}_{\{|\xi_{i,n}|\le\lambda\}}\Big)$$
$$\le \sum_{i=k_{1,n}}^{k_{2,n}} \mathbb{E}\Big(|\xi_{i,n}|\mathbb{1}_{\{|\xi_{i,n}|>\lambda\}}\Big) + \lambda\sum_{i=k_{1,n}}^{k_{2,n}} \mathbb{E}\Big(|\xi_{i,n}|\mathbb{1}_{\{|\xi_{i,n}|\le\lambda\}}\Big). \tag{S.13}$$

Taking into account that $\xi_{k_{1,n},n},\dots,\xi_{k_{2,n},n}$ are identically distributed, we estimate the second sum on the right-hand side of bound (S.13) as follows:

$$\begin{aligned}\sum_{i=k_{1,n}}^{k_{2,n}} \mathbb{E}\Big(|\xi_{i,n}|\mathbb{1}_{\{|\xi_{i,n}|\le\lambda\}}\Big) &\le (k_{2,n}-k_{1,n}+1)\mathbb{E}(|\xi_{k_{1,n},n}|)\\ &= \mathbb{E}(|Z-\mathbb{E}Z|)\\ &\le 2\mathbb{E}(|Z|)\\ &= 2\mathbb{E}\big(X \mid Y\in(y_1,y_2]\big).\end{aligned} \tag{S.14}$$

Since $y_1,y_2\in W_\varepsilon$, condition (C_2) implies that the right-hand side of (S.14) does not exceed a constant that does not depend on the individual pairs (y_1,y_2), although possibly depend on the neighbourhood W_ε. Since $\lambda>0$ can be made as small as needed, we can make the second sum on the right-hand side of bound (S.13) as small as needed.

As for the first sum on the right-hand side of bound (S.13), we write the equation

$$\sum_{i=k_{1,n}}^{k_{2,n}} \mathbb{E}\Big(|\xi_{i,n}|\mathbb{1}_{\{|\xi_{i,n}|>\lambda\}}\Big) = \mathbb{E}\Big(|Z-\mathbb{E}Z|\mathbb{1}_{\{|Z-\mathbb{E}Z|>(k_{2,n}-k_{1,n}+1)\lambda\}}\Big). \tag{S.15}$$

Since $\mathbb{E}(|Z|)$ is bounded by a finite constant that does not depend on the points $(y_1,y_2)\in W_\varepsilon$ due to continuity of g in the neighbourhood W_ε, and since $(k_{2,n}-k_{1,n}+1)$ tends to ∞, the expectation on the right-hand side of equation (S.15) converges to 0 when $n\to\infty$.

Hence, by letting $n\to\infty$, the first sum on the right-hand side of bound (S.13) asymptotically vanishes, and then by letting $0<\lambda\to 0$, the second sum on the right-hand side of bound (S.13) asymptotically vanishes as well. Consequently, $D_n\to 0$ when $n\to\infty$, and so bound (S.12) implies $\mathbb{E}(|\zeta_n|)\to 0$ when $n\to\infty$. We can now use the Markov inequality to conclude the proof of statement (S.11) and in this way finish the proof of Theorem 2.1. $\square$

Before we commence the proof of Theorem 2.2, we need a lemma.

Lemma S.1.2. *Suppose that conditions* (C$_1$) *and* (D$_1$)–(D$_2$) *are satisfied. Then for every* $\varepsilon > 0$ *we can find a constant* $C > 0$ *such that*

$$\mathbb{P}\big(Y_{k_{j,n}:n} \notin W_{C\Delta_n}\big) \le \varepsilon, \quad j \in \{1,2\}, \tag{S.16}$$

for all sufficiently large n*, where* $k_{1,n}$ *and* $k_{2,n}$ *are defined by equations* (2.4)*, and the interval* $W_{C\Delta_n}$ *is given by equation* (S.2) *but with* ε *replaced by* $C\Delta_n$*, where*

$$\Delta_n = \Delta_{1,n} + \Delta_{2,n}$$

Proof. We shall only prove statement (S.16) for $j = 1$, as the proof for $j = 2$ is analogous. We have

$$\mathbb{P}\big(Y_{k_{1,n}:n} \notin W_{C\Delta_n}\big) = \mathbb{P}\big(Y_{k_{1,n}:n} \le \mathrm{VaR}_{p-C\Delta_n}(Y)\big) + \mathbb{P}\big(Y_{k_{1,n}:n} \ge \mathrm{V@R}_{p+C\Delta_n}(Y)\big). \tag{S.17}$$

We shall only prove that the first probability on the right-hand side of equation (S.17) does not exceed ε for all sufficiently large constants C, as the same holds for the second probability.

We have

$$\mathbb{P}\big(Y_{k_{1,n}:n} \le \mathrm{VaR}_{p-C\Delta_n}(Y)\big) = \mathbb{P}\big(U_{k_{1,n}:n} \le G(\mathrm{VaR}_{p-C\Delta_n}(Y))\big), \tag{S.18}$$

where $U_{k_{1,n}:n}$ is the uniform on $[0,1]$ order statistic such that $Y_{k_{1,n}:n}$ on the original probability space $(\Omega, \mathcal{A}, \mathbb{P})$ has the same distribution as $G^{-1}(U_{k_{1,n}:n})$ on a possibly different probability space $(\Omega^*, \mathcal{A}^*, \mathbb{P}^*)$, but we skip the asterisk from $\mathbb{P}^*$ for simplicity. Due to condition (C$_1$), we have $G(\mathrm{VaR}_{p-C\Delta_n}(Y)) = p - C\Delta_n$ for all sufficiently large n. We are thus left to verify

$$\mathbb{P}\big(U_{k_{1,n}:n} \le p - C\Delta_n\big) \le \varepsilon. \tag{S.19}$$

With $p_{1,n} := \mathbb{E}(U_{k_{1,n}:n})$, which is equal to $k_{1,n}/(n+1)$, we have

$$\mathbb{P}(U_{k_{1,n}:n} \le p - C\Delta_n) = \mathbb{P}(U_{k_{1,n}:n} - p_{1,n} \le p - p_{1,n} - C\Delta_n).$$

Since

$$\begin{aligned}0 < p - p_{1,n} &= p - \frac{[n(p-\Delta_{1,n})]}{n+1}\\ &\le p - \frac{n(p-\Delta_{1,n})-1}{n+1}\\ &= \Delta_{1,n} + \frac{p-\Delta_{1,n}+1}{n+1}\\ &\le \Delta_n + 2n^{-1}\\ &\le 3\Delta_n\end{aligned}$$

for all sufficiently large n, we have the bound

$$\mathbb{P}(U_{k_{1,n}:n} \le p - C\Delta_n) \le \mathbb{P}(U_{k_{1,n}:n} - p_{1,n} \le -(C-3)\Delta_n).$$

Since without loss of generality we can choose $C > 3$, with the help of the Chebyshev inequality and the formula

$$\mathbb{V}\big(U_{k_{1,n}:n}\big) = \frac{p_{1,n}(1-p_{1,n})}{n+2}$$

for the variance of the order statistic $U_{k_{1,n}:n}$, we obtain

$$\begin{aligned}\mathbb{P}(U_{k_{1,n}:n} \le p - C\Delta_n) &\le \frac{p_{1,n}(1-p_{1,n})}{(n+2)(C-3)^2\Delta_n^2}\\ &\le \frac{p_{1,n}(1-p_{1,n})}{(C-3)^2(\liminf\sqrt{n}\Delta_n)^2}\\ &= \frac{p(1-p)}{(C-3)^2(\liminf\sqrt{n}\Delta_n)^2} + o(1)\end{aligned} \tag{S.20}$$

when $n \to \infty$. Given condition (D_2), by choosing a sufficiently large C we can make the ratio on the right-hand side of equation (S.20) smaller than ε. With this, we complete the proof of Lemma S.1.2. □

Proof of Theorem 2.2. The asymptotic behaviors of the estimators defined by (2.1) and (2.3) are the same, and so we only work with $\widehat{\mathrm{EAR}}_{p,n}$. Denote

$$T_n = \sqrt{N_n}\Big(\widehat{\mathrm{EAR}}_{p,n} - \mathrm{EAR}_p(X \mid Y)\Big). \tag{S.21}$$

Theorem 2.2 follows if, for every $t \in \mathbb{R}$,

$$\mathbb{E}\big(\exp\{itT_n\}\big) \to \exp\big\{-\sigma^2t^2/2\big\}, \tag{S.22}$$

where σ^2 is defined by equation (2.6). To prove statement (S.22), we fix any $\epsilon > 0$ and let $C > 0$ be such that statement (S.16) holds. We have the following asymptotic representation for the characteristic function of T_n:

$$\iint_{W_{C\Delta_n} \times W_{C\Delta_n}} \mathbb{E}\Big(\exp\big\{itT_n\big\} \mid Y_{k_{1,n}-1:n} = y_1, Y_{k_{2,n}+1:n} = y_2 \Big) \mathrm{d}F_{k_{1,n},k_{2,n}}(y_1, y_2) + r_n, \tag{S.23}$$

where $F_{k_{1,n},k_{2,n}}$ is the joint cdf of $Y_{k_{1,n}-1:n}$ and $Y_{k_{2,n}+1:n}$, and the remainder term r_n satisfies the bound

$$|r_n| \le \mathbb{P}\Big((Y_{k_{1,n}:n}, Y_{k_{2,n}:n}) \notin W_{C\Delta_n} \times W_{C\Delta_n} \Big) \le 2\epsilon$$

for all sufficiently large n. As for the conditional expectation under the integral sign in expression (S.23), we have

$$\begin{aligned}
&\mathbb{E}\Big(\exp\big\{itT_n\big\} \mid Y_{k_{1,n}-1:n} = y_1, Y_{k_{2,n}+1:n} = y_2 \Big) \\
&\quad = \mathbb{E}\left(\exp\left\{ it\sqrt{N_n}\left(\widehat{\mathrm{EAR}}_{p,n} - \mathrm{EAR}_p(X \mid Y) \right) \right\} \mid Y_{k_{1,n}-1:n} = y_1, Y_{k_{2,n}+1:n} = y_2 \right) \\
&\quad = \mathbb{E}\left(\exp\left\{ it\sqrt{N_n}\left(\widehat{\mathrm{EAR}}_{p,n} - \mathbb{E}\big(X \mid Y \in (y_1, y_2]\big) \right) \right\} \mid Y_{k_{1,n}-1:n} = y_1, Y_{k_{2,n}+1:n} = y_2 \right) \\
&\qquad \times \exp\left\{ it\sqrt{N_n}\Big(\mathbb{E}\big(X \mid Y \in (y_1, y_2]\big) - \mathrm{EAR}_p(X \mid Y) \Big) \right\}.
\end{aligned} \tag{S.24}$$

For the right-most difference of two expectations, we have

$$\begin{aligned}
\mathbb{E}\big(X \mid Y \in (y_1, y_2]\big) - \mathrm{EAR}_p(X \mid Y) &= \frac{1}{G(y_2) - G(y_1)} \mathbb{E}\big(g(Y)\mathbb{1}_{\{Y \in (y_1, y_2]\}}\big) - \mathrm{EAR}_p(X \mid Y) \\
&= \frac{1}{G(y_2) - G(y_1)} \int_{G(y_1)}^{G(y_2)} \Big(\mathrm{EAR}_\tau(X \mid Y) - \mathrm{EAR}_p(X \mid Y) \Big) \mathrm{d}\tau.
\end{aligned}$$

Since $y_1, y_2 \in W_{C\Delta_n}$, the integration interval is covered by the interval

$$\Big[G\big(\mathrm{VaR}_{p-C\Delta_n}(Y)\big), G\big(\mathrm{VaR}_{p+C\Delta_n}(Y)\big) \Big] \tag{S.25}$$

whose length, for all sufficiently large n, is equal to $2C\Delta_n$ due to condition (C_1). Using condition (C_3), we obtain

$$\begin{aligned}
\sqrt{N_n}\Big(\mathbb{E}\big(X \mid Y \in (y_1, y_2]\big) - \mathrm{EAR}_p(X \mid Y) \Big) &\le c\sqrt{n\Delta_n}\, \Delta_n^{\alpha} \\
&= c\sqrt{n\Delta_n^{2\alpha+1}}
\end{aligned}$$

for a constant c that does not depend on n. By condition (D$_3$), we have

$$\Delta_n = o\big(n^{-1/(2\alpha+1)}\big),$$

and so the exponent on the right side of equation (S.24) converges to 1 when $n \to \infty$.

Consequently, to complete the proof of statement (S.22), we need to show that, for all $y_1, y_2 \in W_{C\Delta_n}$, we have

$$\mathbb{E}\left(\exp\left\{it\sqrt{N_n}\left(\widehat{\mathrm{EAR}}_{p,n} - \mathbb{E}\big(X \mid Y \in (y_1, y_2]\big)\right)\right\} \mid Y_{k_{1,n}-1:n} = y_1, Y_{k_{2,n}+1:n} = y_2\right) \to \exp\big\{-\sigma^2 t^2/2\big\} \quad \text{(S.26)}$$

when $n \to \infty$. For this, we first note the asymptotic equivalence of

$$\sqrt{N_n}\left(\widehat{\mathrm{EAR}}_{p,n} - \mathbb{E}\big(X \mid Y \in (y_1, y_2]\big)\right)$$

and

$$\frac{1}{\sqrt{k_{2,n} - k_{1,n} + 1}} \sum_{i=k_{1,n}}^{k_{2,n}} \Big(X_{i,n} - \mathbb{E}\big[X \big| Y \in (y_1, y_2]\big]\Big). \quad \text{(S.27)}$$

Conditionally on the events $Y_{k_{1,n}-1:n} = y_1$ and $Y_{k_{2,n}+1:n} = y_2$, quantity (S.27) has the same distribution as

$$S_n := \frac{1}{\sqrt{k_{2,n} - k_{1,n} + 1}} \sum_{i=k_{1,n}}^{k_{2,n}} \big(Z_i - \mathbb{E}(Z)\big),$$

where Z_i, $i = k_{1,n}, \ldots k_{2,n}$, are independent copies of a random variable Z whose cdf is given by equation (S.10). We shall next show that S_n converges to a centered normal random variable with the variance $\sigma^2 = \mathbb{V}\big(X \mid Y = \mathrm{VaR}_p(Y)\big)$.

Note that the cdf F_Z depends on the points y_1 and y_2, which are in the neighbourhood $W_{C\Delta_n}$ that depends on n. Hence, some care is needed when showing that the aforementioned asymptotic normality of S_n holds. We use the approach based on the characteristic function, an thus start with the equation

$$\mathbb{E}\big(\exp\{itS_n\}\big) = \left(\varphi_Z\left(\frac{t}{\sqrt{k_{2,n} - k_{1,n} + 1}}\right)\right)^{k_{2,n}-k_{1,n}+1}, \quad \text{(S.28)}$$

where

$$\varphi_Z(t) = \mathbb{E}\big(\exp\{it(Z - \mathbb{E}Z)\}\big).$$

Hence,

$$\begin{aligned}\log\Big(\mathbb{E}\exp\big(itS_n\big)\Big) &= (k_{2,n}-k_{1,n}+1)\log\left(1-\frac{t^2\mathbb{V}(Z)}{k_{2,n}-k_{1,n}+1}+o\left(\frac{t^2}{k_{2,n}-k_{1,n}+1}\right)\right)\\ &= (k_{2,n}-k_{1,n}+1)\left(-\frac{t^2\mathbb{V}(Z)}{2(k_{2,n}-k_{1,n}+1)}+o\left(\frac{t^2}{k_{2,n}-k_{1,n}+1}\right)\right)\\ &= -\frac{t^2\sigma^2}{2}+o(1), \qquad \text{(S.29)}\end{aligned}$$

provided that $\mathbb{V}(Z)\to\sigma^2$ when $n\to\infty$, which we check next.

Using condition (C$_2$) and in particular the continuity of $\tau\mapsto g(\mathrm{VaR}_\tau(Y))=\mathrm{EAR}_\tau(X\mid Y)$, we have

$$\begin{aligned}\mathbb{E}Z=\mathbb{E}\big(X\mid Y\in(y_1,y_2]\big) &= \frac{1}{G(y_2)-G(y_1)}\mathbb{E}\big(g(Y)\mathbb{1}_{\{Y\in(y_1,y_2]\}}\big)\\ &= \frac{1}{G(y_2)-G(y_1)}\int_{G(y_1)}^{G(y_2)}\mathrm{EAR}_\tau(X\mid Y)\mathrm{d}\tau\\ &= \mathrm{EAR}_p(X\mid Y)+o(1),\end{aligned}$$

where the right-most asymptotic equation holds because the integration interval is covered by interval (S.25) whose length, due to condition (C$_1$), is equal $2C\Delta_n$ for all sufficiently large n and thus converges to 0 when $n\to\infty$.

Likewise, using condition (C$_4$) and in particular the continuity of $\tau\mapsto g_2(\mathrm{VaR}_\tau(Y))=\mathbb{E}(X^2\mid Y=\mathrm{VaR}_\tau(Y))$, we have

$$\begin{aligned}\mathbb{E}Z^2=\mathbb{E}\big(X^2\mid Y\in(y_1,y_2]\big) &= \frac{1}{G(y_2)-G(y_1)}\mathbb{E}\big(g_2(Y)\mathbb{1}_{\{Y\in(y_1,y_2]\}}\big)\\ &= \frac{1}{G(y_2)-G(y_1)}\int_{G(y_1)}^{G(y_2)}g_2(\mathrm{VaR}_\tau(Y))\mathrm{d}\tau\\ &= \mathbb{E}\big(X^2\mid Y=\mathrm{VaR}_p(Y)\big)+o(1).\end{aligned}$$

Combining the two asymptotic expressions that we have just derived for the first two moments of Z, we arrive at the following asymptotic expression for the variance

$$\begin{aligned}\mathbb{V}(Z) &= \mathbb{E}\big(X^2\mid Y=\mathrm{VaR}_p(Y)\big)-\Big(\mathbb{E}\big(X\mid Y=\mathrm{VaR}_p(Y)\big)\Big)^2+o(1)\\ &= \sigma^2+o(1)\end{aligned}$$

when $n\to\infty$. This establishes statement (S.26) and concludes the proof of Theorem 2.2. $\square$

Proof of Theorem 2.3. The proof of statement (2.7) follows the same arguments as the proof of consistency of $\widehat{\mathrm{EAR}}_{p,n}$ given in Theorem 2.1. $\square$

References

Arnold, B.C., Balakrishnan, N. and Nagaraja, H.N. (2008). *A First Course in Order Statistics.* (Classic edition.) Society for Industrial and Applied Mathematics, Philadelphia.

Borovkov, A.A. (1998). *Mathematical Statistics.* Routledge, London.

David, H.A. and Nagaraja, H.N. (2003). *Order Statistics.* (Third edition.) Wiley, New York