risks

MDPI

Article

Bootstrapping Average Value at Risk of Single and Collective Risks

Eric Beutner [1] and Henryk Zähle [2,*]

[1] Department of Quantitative Economics, Maastricht University, 6200 MD Maastricht, The Netherlands; e.beutner@maastrichtuniversity.nl
[2] Department of Mathematics, Saarland University, 66123 Saarbrücken, Germany
* Correspondence: zaehle@math.uni-sb.de; Tel.: +49-681-302-4383

Received: 1 August 2018; Accepted: 7 September 2018; Published: 12 September 2018

Abstract: Almost sure bootstrap consistency of the blockwise bootstrap for the Average Value at Risk of single risks is established for strictly stationary β-mixing observations. Moreover, almost sure bootstrap consistency of a multiplier bootstrap for the Average Value at Risk of collective risks is established for independent observations. The main results rely on a new functional delta-method for the almost sure bootstrap of uniformly quasi-Hadamard differentiable statistical functionals, to be presented here. The latter seems to be interesting in its own right.

Keywords: Average Value at Risk; compound distribution; nonparametric estimation; multiplier bootstrap; blockwise bootstrap; functional delta-method; uniform quasi-Hadamard differentiability; chain rule

1. Introduction

One of the most popular risk measures in practice is the so-called Average Value at Risk which is also referred to as Expected Shortfall (see Acerbi and Szekely (2014); Acerbi and Tasche (2002a, 2002b); Emmer et al. (2015) and references therein). For a fixed level $\alpha \in (0,1)$, the corresponding Average Value at Risk is the map $\mathrm{AV@R}_\alpha : L^1 \to \mathbb{R}$ defined by $\mathrm{AV@R}_\alpha(X) := \mathcal{R}_\alpha(F_X)$, where F_X refers to the distribution function of X, L^1 is the usual L^1-space associated with some atomless probability space, and

$$\mathcal{R}_\alpha(F) := \int_0^1 F^{\leftarrow}(s)\, dg_\alpha(s) = -\int_{-\infty}^0 g_\alpha(F(x))\, dx + \int_0^\infty \big(1 - g_\alpha(F(x))\big)\, dx \tag{1}$$

for any $F \in \mathbf{F}_1$ with $\mathbf{F}_1$ the set of the distribution functions F_X of all $X \in L^1$. Here, $g_\alpha(t) := \frac{1}{1-\alpha}\max\{t-\alpha; 0\}$ and $F^{\leftarrow}(s) := \inf\{x \in \mathbb{R} : F(x) \geq s\}$ denotes the left-continuous inverse of F. The statistical functional $\mathcal{R}_\alpha : \mathbf{F}_1 \to \mathbb{R}$ is sometimes referred to as risk functional associated with $\mathrm{AV@R}_\alpha$. Note that $\mathrm{AV@R}_\alpha(X) = \mathbb{E}[X | X \geq F_X^{\leftarrow}(\alpha)]$ when F_X is continuous at $F_X^{\leftarrow}(\alpha)$.

In this article, we mainly focus on bootstrap methods for the Average Value at Risk. Before doing so, we briefly review nonparametric estimation techniques and asymptotic results for the Average Value at Risk. Given identically distributed observations $X_1, \ldots, X_n$ $(, X_{n+1}, \ldots)$ on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with unknown marginal distribution $F \in \mathbf{F}_1$, a natural estimator for $\mathcal{R}_\alpha(F)$ is the empirical plug-in estimator

$$\mathcal{R}_\alpha(\widehat{F}_n) = \int_0^1 \widehat{F}_n^{\leftarrow}(s)\, dg_\alpha(s) = \sum_{i=1}^n \big\{g_\alpha\big(\tfrac{i}{n}\big) - g_\alpha\big(\tfrac{i-1}{n}\big)\big\} X_{i:n}, \tag{2}$$

where $\widehat{F}_n := \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{[X_i,\infty)}$ is the empirical distribution function of $X_1, \ldots, X_n$ and $X_{1:n}, \ldots, X_{n:n}$ refer to the order statistics of $X_1, \ldots, X_n$. The second representation in Equation (2) shows that $\mathcal{R}_\alpha(\widehat{F}_n)$ is a

specific L-statistic which was already mentioned in Acerbi (2002); Acerbi and Tasche (2002a); Jones and Zitikis (2003).

In particular, if the underlying sequence $(X_i)_{i\in\mathbb{N}}$ is strictly stationary and ergodic, classical results of van Zwet (1980) and Gilat and Helmers (1997) show that $\mathcal{R}_\alpha(\widehat{F}_n)$ converges $\mathbb{P}$-almost surely to $\mathcal{R}_\alpha(F)$ as $n \to \infty$, i.e., that strong consistency holds. If $X_1, X_2, \ldots$ are i.i.d. and F has a finite second moment and takes the value α only once, then a result of Stigler ((Stigler 1974, Theorems 1–2)) yields the asymptotic distribution of the estimation error:

$$\sqrt{n}\big(\mathcal{R}_\alpha(\widehat{F}_n) - \mathcal{R}_\alpha(F)\big) \rightsquigarrow Z \sim \mathrm{N}_{0,\sigma_F^2}, \tag{3}$$

where $\sigma_F^2 := \iint g'_\alpha(F(x_0))\Gamma(x_0,x_1)g'_\alpha(F(x_1))\,dx_0\,dx_1$ with $\Gamma(x_0,x_1) := F(x_0 \wedge x_1)(1 - F(x_0 \vee x_1)) + \sum_{i=0}^{1}\sum_{k=2}^{\infty}\mathbb{C}\mathrm{ov}(\mathbb{1}_{\{X_1\le x_i\}}, \mathbb{1}_{\{X_k\le x_{1-i}\}})$, $g'_\alpha := \frac{1}{1-\alpha}\mathbb{1}_{(\alpha,1]}$, and $\rightsquigarrow$ refers to convergence in distribution (see also Shorack 1972; Shorack and Wellner 1986). In fact, for independent $X_1, X_2, \ldots$ the second summand in the definition of $\Gamma(x_0,x_1)$ vanishes. Results of Beutner and Zähle (2010) show that Equation (3) still holds if $(X_i)_{i\in\mathbb{N}}$ is strictly stationary and α-mixing with mixing coefficients $\alpha(i) = \mathcal{O}(i^{-\theta})$ and $\lim_{x\to\infty}(1-F(x))x^{2\theta/(\theta-1)} < \infty$ for some $\theta > 1+\sqrt{2}$. Tsukahara (2013) obtained the same result. A similar result can also be derived from an earlier work by Mehra and Rao (1975), but under a faster decay of the mixing coefficients and under an additional assumption on the dependence structure. We emphasize that the method of proof proposed by Beutner and Zähle is rather flexible, because it easily extends to other weak and strong dependence concepts and other risk measures (see Beutner et al. 2012; Beutner and Zähle 2010, 2016; Krätschmer et al. 2013; Krätschmer and Zähle 2017).

Even in the i.i.d. case the asymptotic variance σ_F^2 depends on F in a fairly complex way. For the approximation of the distribution of $\sqrt{n}(\mathcal{R}_\alpha(\widehat{F}_n) - \mathcal{R}_\alpha(F))$, bootstrap methods should thus be superior to the method of estimating σ_F^2. However, to the best of our knowledge, theoretical investigations of the bootstrap for the Average Value at Risk seem to be rare. According to Gribkova (2016), a result of Gribkova (2002) yields bootstrap consistency for Efron's bootstrap when $X_1, X_2, \ldots$ are i.i.d, while Theorem 3 of Helmers et al. (1990) seems not to cover the Average Value at Risk, because there the function J (which plays the role of g'_α) is assumed to be Lipschitz continuous. In these articles, bootstrap consistency is typically proved by first proving consistency of the bootstrap variance and then using this result by showing that upper bounds for the difference between the sampling distribution and the bootstrap distribution converge to zero. Employing different techniques, Beutner and Zähle (2016) established bootstrap consistency in probability for the multiplier bootstrap when $X_1, X_2, \ldots$ are i.i.d. as well as bootstrap consistency in probability for the circular bootstrap when $X_1, X_2, \ldots$ are strictly stationary and β-mixing with mixing coefficients $\beta(i) = \mathcal{O}(i^{-b})$ and $\int |x|^p dF(x) < \infty$ for some $p > 2$ and $b > p/(p-2)$. Recently, Sun and Cheng (2018) established bootstrap consistency in probability for the moving blocks bootstrap when $X_1, X_2, \ldots$ are strictly stationary and α-mixing with mixing coefficients $\alpha(i) \le c\delta^i$ and $\int |x|^p dF(x) < \infty$ for some $p > 4$, $c > 0$ and $\delta \in (0,1)$. Strictly speaking, Sun and Cheng did not consider the Average Value at Risk (Expected Shortfall) but the Tail Conditional Expectation in the sense of Acerbi and Tasche (2002a, 2002b).

The contribution of the article at hand is twofold. First, we extend the results of Beutner and Zähle (2016) on the Average Value at Risk from bootstrap consistency *in probability* to bootstrap consistency *almost surely*. Second, we establish bootstrap consistency for the Average Value at Risk of *collective* risks, i.e., for $\mathcal{R}_\alpha(F^{*m})$ and more general expressions.

The rest of the article is organized as follows. In Section 2, we present and illustrate our main results which are proved in Section 3. Section 3 is followed by the conclusions. The proofs of Section 3 rely on a new functional delta-method for the almost sure bootstrap which seems to be interesting in its own right and which is presented in Appendix B. Roughly speaking, the (functional) delta method studies properties of particular estimators for quantities of the form $H(\theta)$. Here, H is a known functional, such as the Average Value at Risk functional, and θ is a possibly infinite dimensional parameter, such as an unknown distribution function. The particular estimators covered by the

(functional) delta method are of the form $H(\widehat{T}_n)$ where $\widehat{T}_n$ is an estimator for θ. In general and in the particular application considered here, the appeal of the (functional) delta method lies in the fact that, once "differentiability" of H (here, the Average Value at Risk functional) is established, the asymptotic error distribution of $H(\widehat{T}_n)$ can immediately be derived from the asymptotic error distribution of $\widehat{T}_n$ (here $\widehat{F}_n$). This also applies to the (functional) delta method for the bootstrap where bootstrap consistency of the bootstrapped version of $H(\widehat{T}_n)$ will follow from the respective property of the bootstrapped version of $\widehat{T}_n$ (here $\widehat{F}_n$). Thus, if in financial or actuarial applications the data show dependencies for which the asymptotic error distribution and/or bootstrap consistency of plug-in estimators for the Average Value at Risk have not been established yet, it would be enough to check if for these dependencies the asymptotic error distribution and/or bootstrap consistency of $\widehat{F}_n$ is known; thanks to the (functional) delta method the Average Value at Risk functional would inherit these properties. In Appendix A.1, we give results on convergence in distribution for the open-ball σ-algebra which are needed for the main results, and in Appendix A.2 we prove a delta-method for uniformly quasi-Hadamard differentiable maps that is the basis for the method of Appendix B. Readers interested in these methods used to prove the main results might wish to first work through Appendices A and B before reading Sections 2 and 3.

2. Main Results

2.1. The Case of i.i.d. Observations

Keep the notation of Section 1. Assume that $(X_i)_{i\in\mathbb{N}}$ is a sequence of i.i.d. real-valued random variables on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with distribution function F. Let $\widehat{F}_n := \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{[X_i,\infty)}$ and (W_{ni}) be a triangular array of nonnegative real-valued random variables on another probability space $(\Omega', \mathcal{F}', \mathbb{P}')$ such that one of the following two settings is met.

S1. The random vector $(W_{n1}, \ldots, W_{nn})$ is multinomially distributed according to the parameters n and $p_1 = \cdots = p_n := 1/n$ for every $n \in \mathbb{N}$.
S2. $W_{ni} = Y_i/\overline{Y}_n$ for every $i = 1, \ldots, n$ and $n \in \mathbb{N}$, where $\overline{Y}_n := \frac{1}{n}\sum_{j=1}^n Y_j$ and (Y_j) is any sequence of nonnegative i.i.d. random variables on $(\Omega', \mathcal{F}', \mathbb{P}')$ with $\int_0^\infty \mathbb{P}'[Y_1 > t]^{1/2}\, dt < \infty$ and $\mathbb{V}\mathrm{ar}'[Y_1]^{1/2} = \mathbb{E}'[Y_1] > 0$.

Let $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}}) := (\Omega \times \Omega', \mathcal{F} \otimes \mathcal{F}', \mathbb{P} \otimes \mathbb{P}')$ and $\widehat{F}_n^*(\omega, \omega') := \frac{1}{n}\sum_{i=1}^n W_{ni}(\omega')\, \mathbb{1}_{[X_i(\omega),\infty)}$. Setting S1. is nothing but *Efron's boostrap* (Efron 1979). If in Setting S2. the distribution of Y_1 is the exponential distribution with parameter 1, then the resulting scheme is in line with the *Bayesian bootstrap* of Rubin (1981). Let $\sigma_F^2 := \iint g'_\alpha(F(x_0))\Gamma(x_0, x_1) g'_\alpha(F(x_1))\, dx_0\, dx_1$ with $\Gamma(x_0, x_1) := F(x_0 \wedge x_1)(1 - F(x_0 \vee x_1))$.

Theorem 1. *In the setting above assume that $\int \phi^2\, dF < \infty$ for some continuous function $\phi : \mathbb{R} \to [1, \infty)$ with $\int 1/\phi(x)\, dx < \infty$ (in particular $F \in \mathbf{F}_1$), and that F takes the value α only once. Then*

$$\sqrt{n}\big(\mathcal{R}_\alpha(\widehat{F}_n) - \mathcal{R}_\alpha(F)\big) \rightsquigarrow Z \sim \mathrm{N}_{0,\sigma_F^2} \tag{4}$$

and

$$\sqrt{n}\big(\mathcal{R}_\alpha(\widehat{F}_n^*(\omega, \cdot)) - \mathcal{R}_\alpha(\widehat{F}_n(\omega))\big) \rightsquigarrow Z \sim \mathrm{N}_{0,\sigma_F^2}, \qquad \mathbb{P}\text{-a.e. } \omega. \tag{5}$$

Theorem 1 is a special case of Corollary 1 below. For the bootstrap Scheme S1. the result of Theorem 1 can be also deduced from Theorem 7 in Gribkova (2002). According to Gribkova (2016), Condition (1) of this theorem is satisfied if there are $0 = a_0 < a_1 < \cdots < a_k = 1$ for some $k \in \mathbb{N}$ such that J is Hölder continuous on each interval (a_{i-1}, a_i), $1 \le i \le k$, and the measure dF^{-1} has no mass at the points $a_1, \ldots, a_{k-1}$. For the bootstrap Scheme S2. the result seems to be new.

We now consider the collective risk model. Let $(X_i)_{i\in\mathbb{N}}$ and $\widehat{F}_n$ be as above, and let $p = (p_k)_{k\in\mathbb{N}_0}$ be the counting density of a distribution on $\mathbb{N}_0$. Let $\mathbf{F}$ denote the set of all distribution functions on

$\mathbb{R}$, and consider the functional $\mathcal{C}_p : \mathbf{F} \to \mathbf{F}$ defined by $\mathcal{C}_p(F) := \sum_{k=0}^{\infty} p_k F^{*k}$, where F^{*k} refers to the k-fold convolution of F, i.e., $F^{*0} := \mathbb{1}_{[0,\infty)}$ and $F^{*k}(x) := \int F(x - x_{k-1}) dF^{*(k-1)}(x_{k-1}) = \int \cdots \int F(x - x_{k-1} - \cdots - x_1) dF(x_1) \cdots dF(x_{k-1})$ for $k \in \mathbb{N}$. If $p_m = 1$ for some $m \in \mathbb{N}_0$, then $\mathcal{C}_p(F) = F^{*m}$. Let $\sigma^2_{p,F} := \iiiint g'_\alpha(F(x_0))\Gamma(x_0 - y_0, x_1 - y_1) g'_\alpha(F(x_1))\, dH_{p,F}(y_0) dH_{p,F}(y_1)\, dx_0\, dx_1$ with $\Gamma(z_0, z_1) := F(z_0 \wedge z_1)(1 - F(z_0 \vee z_1))$ and $H_{p,F} := \sum_{k=1}^{\infty} k p_k F^{*(k-1)}$.

Theorem 2. *In the setting above assume that $\int |x|^{2\lambda}\, dF(x) < \infty$ for some $\lambda > 1$ (in particular $F \in \mathbf{F}_1$) and $\sum_{k=1}^{\infty} p_k\, k^{1+\lambda} < \infty$, and that $\mathcal{C}_p(F)$ takes the value α only once. Then,*

$$\sqrt{n}\big(\mathcal{R}_\alpha(\mathcal{C}_p(\widehat{F}_n)) - \mathcal{R}_\alpha(\mathcal{C}_p(F))\big) \rightsquigarrow Z \sim \mathrm{N}_{0,\sigma^2_{p,F}} \tag{6}$$

and

$$\sqrt{n}\big(\mathcal{R}_\alpha(\mathcal{C}_p(\widehat{F}^*_n(\omega,\cdot))) - \mathcal{R}_\alpha(\mathcal{C}_p(\widehat{F}_n(\omega)))\big) \rightsquigarrow Z \sim \mathrm{N}_{0,\sigma^2_{p,F}}, \qquad \mathbb{P}\text{-a.e. } \omega. \tag{7}$$

Theorem 2 is a special case of Corollary 4 below. Lauer and Zähle (2015, 2017) derive the asymptotic distribution as well as almost sure bootstrap consistency for the Average Value at Risk (and more general risk measures) of F^{*m_n} when m_n/n is asymptotically constant, but we do not know any result in the existing literature which is comparable to that of Theorem 2.

2.2. The Case of β-Mixing Observations

Keep the notation of Section 1. Assume that $(X_i)_{i\in\mathbb{N}}$ is a strictly stationary sequence of β-mixing random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ with distribution function F. As before let $\widehat{F}_n := \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{[X_i,\infty)}$. Let (ℓ_n) be a sequence of integers such that $\ell_n \nearrow \infty$ as $n \to \infty$, and $\ell_n < n$ for all $n \in \mathbb{N}$. Set $k_n := \lceil n/\ell_n \rceil$ for all $n \in \mathbb{N}$. Let $(I_{nj})_{n\in\mathbb{N}, 1\le j\le k_n}$ be a triangular array of random variables on $(\Omega', \mathcal{F}', \mathbb{P}')$ such that $I_{n1}, \ldots, I_{nk_n}$ are i.i.d. according to the uniform distribution on $\{1, \ldots, n - \ell_n + 1\}$ for every $n \in \mathbb{N}$. Let $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}}) := (\Omega \times \Omega', \mathcal{F} \otimes \mathcal{F}', \mathbb{P} \otimes \mathbb{P}')$ and $\widehat{F}^*_n(\omega, \omega') := \frac{1}{n}\sum_{i=1}^n W_{ni}(\omega')\, \mathbb{1}_{[X_i(\omega),\infty)}$ with

$$W_{ni}(\omega') := \sum_{j=1}^{k_n-1} \mathbb{1}_{\{I_{nj} \le i \le I_{nj}+\ell_n-1\}}(\omega') + \mathbb{1}_{\{I_{nk_n} \le i \le I_{nk_n}+(n-(k_n-1)\ell_n)-1\}}(\omega'). \tag{8}$$

Note that the sequence (X_i) and the triangular array (W_{ni}) regarded as families of random variables on the product space $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}}) := (\Omega \times \Omega', \mathcal{F} \otimes \mathcal{F}', \mathbb{P} \otimes \mathbb{P}')$ are independent. At an informal level, this means that, given a sample $X_1, \ldots, X_n$, we pick $k_n - 1$ blocks of length ℓ_n and one block of length $n - (k_n - 1)\ell_n$ in the sample $X_1, \ldots, X_n$, where the start indices $I_{n1}, I_{n2}, \ldots, I_{nk_n}$ are chosen independently and uniformly in the set of indices $\{1, \ldots, n - \ell_n + 1\}$:

$$\begin{array}{ll}
\text{block 1:} & X_{I_{n1}}, X_{I_{n1}+1}, \ldots, X_{I_{n1}+\ell_n-1} \\
\text{block 2:} & X_{I_{n2}}, X_{I_{n2}+1}, \ldots, X_{I_{n2}+\ell_n-1} \\
 & \vdots \\
\text{block } k_n - 1: & X_{I_{n(k_n-1)}}, X_{I_{n(k_n-1)}+1}, \ldots, X_{I_{n(k_n-1)}+\ell_n-1} \\
\text{block } k_n: & X_{I_{nk_n}}, X_{I_{nk_n}+1}, \ldots, X_{I_{nk_n}+(n-(k_n-1)\ell_n)-1}.
\end{array}$$

The bootstrapped empirical distribution function $\widehat{F}^*_n$ is then defined to be the distribution function of the discrete probability measure with atoms $X_1, \ldots, X_n$ carrying masses $W_{n1}, \ldots, W_{nn}$, respectively, where W_{ni} specifies the number of blocks which contain X_i. This is known as the *blockwise bootstrap* (see, e.g., Bühlmann (1994, 1995) and references therein). Assume that the following assertions hold:

A1. $\int \phi^p\, dF < \infty$ for some $p > 4$ (in particular $F \in \mathbf{F}_1$).
A2. The sequence of random variables (X_i) is strictly stationary and β-mixing with mixing coefficients (β_i) satisfying $\beta_i \le c\delta^i$ for some constants $c > 0$ and $\delta \in (0, 1)$.
A3. The block length ℓ_n satisfies $\ell_n = \mathcal{O}(n^\gamma)$ for some $\gamma \in (0, 1/2)$.

Let $\widehat{C}_n := \mathbb{E}'[\widehat{F}_n^*] = \frac{1}{n}\sum_{i=1}^n w_{ni}\mathbb{1}_{[X_i,\infty)}$ with $w_{ni} := \mathbb{E}'[W_{ni}]$, and note that

$$w_{ni} = \begin{cases} k_n \frac{i}{n-\ell_n+1} & , \quad i = 1,\dots,n-(k_n-1)\ell_n \\ (k_n-1)\frac{i}{n-\ell_n+1} + \frac{n-(k_n-1)\ell_n}{n-\ell_n+1} & , \quad i = n-(k_n-1)\ell_n+1,\dots,\ell_n \\ (k_n-1)\frac{\ell_n}{n-\ell_n+1} + \frac{n-(k_n-1)\ell_n}{n-\ell_n+1} = \frac{n}{n-\ell_n+1} & , \quad i = \ell_n+1,\dots,n-\ell_n \\ (k_n-1)\frac{n-i+1}{n-\ell_n+1} + \frac{2n-k_n\ell_n-i+1}{n-\ell_n+1} & , \quad i = n-\ell_n+1,\dots,n-(k_n\ell_n-n) \\ (k_n-1)\frac{n-i+1}{n-\ell_n+1} & , \quad i = n-(k_n\ell_n-n)+1,\dots,n \end{cases} \tag{9}$$

which can be verified easily. Let $\sigma_F^2 := \iint g'_\alpha(F(x_0))\Gamma(x_0,x_1)g'_\alpha(F(x_1))\,dx_0\,dx_1$ with $\Gamma(x_0,x_1) := F(x_0\wedge x_1)(1-F(x_0\vee x_1)) + \sum_{i=0}^1\sum_{k=2}^\infty \mathbb{C}\mathrm{ov}(\mathbb{1}_{\{X_1\le x_i\}},\mathbb{1}_{\{X_k\le x_{1-i}\}})$.

Theorem 3. *In the setting above (in particular under A1.–A3.) assume that F takes the value α only once. Then, we have*

$$\sqrt{n}\big(\mathcal{R}_\alpha(\widehat{F}_n) - \mathcal{R}_\alpha(F)\big) \rightsquigarrow Z \sim \mathrm{N}_{0,\sigma_F^2} \tag{10}$$

and

$$\sqrt{n}\big(\mathcal{R}_\alpha(\widehat{F}_n^*(\omega,\cdot)) - \mathcal{R}_\alpha(\widehat{C}_n(\omega))\big) \rightsquigarrow Z \sim \mathrm{N}_{0,\sigma_F^2}, \qquad \mathbb{P}\text{-a.e. } \omega. \tag{11}$$

Theorem 3 is a special case of Corollary 1 below. To the best of our knowledge, there does not yet exist any result on *almost sure* bootstrap consistency for the Average Value at Risk when the underlying data are dependent.

2.3. Applications

2.3.1. Bootstrapping the Down Side Risk of an Asset Price

Let $(A_i)_{i\in\mathbb{N}_0}$ be the price process of an asset. Let us assume that it is induced by an initial state $A_0 \in \mathbb{R}_+$ and a sequence of $\mathbb{R}_+$-valued i.i.d. random variables $(R_i)_{i\in\mathbb{N}}$ via $A_i := R_iA_{i-1}$, $i \in \mathbb{N}$. Here, R_i is the return of the asset in between time $i-1$ and time i. For instance, if $A_0, A_1, A_2, \dots$ are the observations of a time-continuous Black–Scholes–Merton model with drift μ and volatility σ at the points of the time grid $\{0, h, 2h, \dots\}$, then the distribution of R_1 is the log-normal distribution with parameters $(\mu - \sigma^2/2)h$ and $\sigma^2 h$. However, the adequacy of a specific parametric model is usually hard to verify. For this reason, we do not restrict ourselves to any particular parametric structure for the dynamics of $(R_i)_{i\in\mathbb{N}}$.

Let us assume that we can observe the asset prices $A_0,\dots,A_n$ up to time n, and that we are interested in the Average Value at Risk at level α of the negative price change $A_n - A_{n+1}$ (which specifies the *down side risk* of the asset) in between time n and $n+1$. That is, since for any $a_0,\dots,a_n \in \mathbb{R}_+$ the unconditional distribution of $(1-R_{n+1})a_n$ coincides with the factorized conditional distribution of $A_n - A_{n+1} = (1-R_{n+1})A_n$ given $(A_0,\dots,A_n) = (a_0,\dots,a_n)$, we are in fact interested in $\mathcal{R}_\alpha(F) = \mathrm{AV@R}_\alpha(X)$ for the distribution function F of $X := (1-R_{n+1})a$ for any fixed $a \in \mathbb{R}_+$. As the random variables $X_1 := (1-R_1)a,\dots,X_n := (1-R_n)a$ are i.i.d. copies of X, we can use $\mathcal{R}_\alpha(\widehat{F}_n)$ as an estimator for $\mathcal{R}_\alpha(F)$ and derive from Equation (4) an asymptotic confidence interval at a given level $\tau \in (0,1)$ for $\mathcal{R}_\alpha(F)$ where one has to estimate σ_F^2 by $\iint g'_\alpha(\widehat{F}_n(x_0))\widehat{F}_n(x_0\wedge x_1)(1-\widehat{F}_n(x_0\vee x_1))g'_\alpha(\widehat{F}_n(x_1))\,dx_0\,dx_1$. As the estimator for σ_F^2 depends on $\widehat{F}_n$ in a somewhat complex way, the bootstrap confidence interval

$$\big[\mathcal{R}_\alpha(\widehat{F}_n(\omega)) - \tfrac{1}{\sqrt{n}}\widehat{q}^*_{1-\tau/2}(\omega)\,, \mathcal{R}_\alpha(\widehat{F}_n(\omega)) - \tfrac{1}{\sqrt{n}}\widehat{q}^*_{\tau/2}(\omega)\big] \tag{12}$$

at level τ derived from Equations (4) and (5) is supposed to have a slightly better performance. Here, $\widehat{q}^*_t(\omega)$ denotes a t-quantile of (a Monte Carlo approximation of) the distribution of the left-hand side in Equation (5) for fixed ω. For Equations (4) and (5) it suffices to assume that $\mathbb{E}[|R_1|^{2+\varepsilon}] < \infty$ for some arbitrarily small $\varepsilon > 0$.

2.3.2. Bootstrapping the Total Risk Premium in Insurance Models

In actuarial mathematics, the collective risk model is frequently used for modeling the total claim distribution of an insurance collective. If the counting density $p = (p_k)_{k\in\mathbb{N}_0}$ corresponds to the distribution of the random number N of claims caused by the whole collective within one insurance period, and if $X_1, \ldots, X_N$ $(, X_{N+1}, \ldots)$ denote the i.i.d. sizes of the corresponding claims with marginal distribution F, then $\mathcal{C}_p(F)$ is the distribution of the total claim $\sum_{i=1}^N X_i$ (the latter sum is set to 0 if $N = 0$). Now, $\mathcal{R}_\alpha(\mathcal{C}_p(F))$ is a suitable insurance premium for the whole collective when the Average Value at Risk at level α is considered to be a suitable premium principle.

Assume that p is known, for instance $p_m = 1$ for some fixed $m \in \mathbb{N}$, and let $X_1, \ldots, X_n$ be observed historical (i.i.d.) claims with n large. On the one hand, the construction of an exact confidence interval for $\mathcal{R}_\alpha(\mathcal{C}_p(F))$ at level $\tau \in (0,1)$ based on $X_1, \ldots, X_n$ is hardly possible. Likewise, the performance of an asymptotic confidence interval at level τ derived from Equation (6) with (nonparametrically) estimated $\sigma^2_{p,F}$ is typically only moderate. Take into account that $\sigma^2_{p,F}$ depends on the unknown F in a fairly complex way. On the other hand, the bootstrap confidence interval

$$\left[\mathcal{R}_\alpha\left(\mathcal{C}_p(\widehat{F}_n(\omega))\right) - \tfrac{1}{\sqrt{n}}\widehat{q}^*_{1-\tau/2}(\omega)\,, \mathcal{R}_\alpha\left(\mathcal{C}_p(\widehat{F}_n(\omega))\right) - \tfrac{1}{\sqrt{n}}\widehat{q}^*_{\tau/2}(\omega)\right]$$

at level τ derived from Equation (7) should have a better performance. Here, $\widehat{q}^*_t(\omega)$ denotes a t-quantile of (a Monte Carlo approximation of) the distribution of the left-hand side in Equation (7) for fixed ω.

Note that Theorem 2 ensures that Equations (6) and (7) hold true when the marginal distribution F of the X_i is any log-normal distribution, any Gamma distribution, any Pareto distribution with tail index greater than 2, or any convex combination of one of these distributions with the Dirac measure δ_0, and the counting density p corresponds to any Dirac measure with atom in $\mathbb{N}$, any binomial distribution, any Poisson distribution, or any geometric distribution. The former distributions are classical examples for the single claim distribution and the latter distributions are classical examples for the claim number distribution.

3. Proofs of Main Results

Here, we prove the results of Section 2. In fact, Theorems 1–3 are special cases of Corollaries 1 and 4. The latter corollaries are proved with the help of the technique introduced in Appendix B.2, which in turn avails the concept of uniform quasi-Hadamard differentiability (see Definition A1 in Appendix B.1).

Keep the notation introduced in Section 1. Let $\mathbf{D}$ be the space of all cádlág functions v on $\mathbb{R}$ with finite sup-norm $\|v\|_\infty := \sup_{t\in\mathbb{R}} |v(t)|$, and $\mathcal{D}$ be the σ-algebra on $\mathbf{D}$ generated by the one-dimensional coordinate projections π_t, $t \in \mathbb{R}$, given by $\pi_t(v) := v(t)$. Let $\phi : \mathbb{R} \to [1, \infty)$ be a weight function, i.e., a continuous function being non-increasing on $(-\infty, 0]$ and non-decreasing on $[0, \infty)$. Let $\mathbf{D}_\phi$ be the subspace of $\mathbf{D}$ consisting of all $x \in \mathbf{D}$ satisfying $\|v\|_\phi := \|v\phi\|_\infty < \infty$ and $\lim_{|t|\to\infty} |v(t)| = 0$. The latter condition automatically holds when $\lim_{|t|\to\infty} \phi(t) = \infty$. We equip $\mathbf{D}_\phi$ with the trace σ-algebra of $\mathcal{D}$, and note that this σ-algebra coincides with the σ-algebra $\mathcal{B}^\circ_\phi$ on $\mathbf{D}_\phi$ generated by the $\|\cdot\|_\phi$-open balls (see Lemma 4.1 in Beutner and Zähle (2016)).

3.1. Average Value at Risk functional

Using the terminology of Part (i) of Definition A1, we obtain the following result.

Proposition 1. *Let $F \in \mathbf{F}_1$ and assume that F takes the value α only once. Let $\mathcal{S}$ be the set of all sequences $(G_n) \subseteq \mathbf{F}_1$ with $G_n \to F$ pointwise. Moreover, assume that $\int 1/\phi(x)\,dx < \infty$. Then, the map $\mathcal{R}_\alpha : \mathbf{F}_1\ (\subseteq \mathbf{D}) \to \mathbb{R}$ is uniformly quasi-Hadamard differentiable with respect to $\mathcal{S}$ tangentially to $\mathbf{D}_\phi\langle\mathbf{D}_\phi\rangle$, and the uniform quasi-Hadamard derivative $\dot{\mathcal{R}}_{\alpha;F} : \mathbf{D}_\phi \to \mathbb{R}$ is given by*

$$\dot{\mathcal{R}}_{\alpha;F}(v) := -\int g'_\alpha(F(x))v(x)\,dx, \tag{13}$$

where as before $g'_\alpha := \frac{1}{1-\alpha}\mathbb{1}_{(\alpha,1]}$.

Proposition 1 shows in particular that for any $F \in \mathbf{F}_1$ which takes the value α only once, the map $\mathcal{R}_\alpha : \mathbf{F}_1\,(\subseteq \mathbf{D}) \to \mathbb{R}$ is uniformly quasi-Hadamard differentiable at F tangentially to $\mathbf{D}_\phi\langle\mathbf{D}_\phi\rangle$ (in the sense of Part (ii) of Definition A1) with uniform quasi-Hadamard derivative given by Equation (13).

Proof. (of Proposition 1) First, note that the map $\dot{\mathcal{R}}_{\alpha;F}$ defined in Equation (13) is continuous with respect to $\|\cdot\|_\phi$, because

$$|\dot{\mathcal{R}}_{\alpha;F}(v_1) - \dot{\mathcal{R}}_{\alpha;F}(v_2)| \le \int \frac{1}{1-\alpha}\,|v_1(x) - v_2(x)|\,dx \le \Big(\frac{1}{1-\alpha}\int 1/\phi(x)\,dx\Big)\|v_1 - v_2\|_\phi$$

holds for every $v_1, v_2 \in \mathbf{D}_\phi$.

Now, let $((F_n), v, (v_n), (\varepsilon_n))$ be a quadruple with $(F_n) \subseteq \mathbf{F}_1$ satisfying $F_n \to F$ pointwise, $v \in \mathbf{D}_\phi$, $(v_n) \subseteq \mathbf{D}_\phi$ satisfying $\|v_n - v\|_\phi \to 0$ and $(F_n + \varepsilon_n v_n) \subseteq \mathbf{F}_1$, and $(\varepsilon_n) \subseteq (0,\infty)$ satisfying $\varepsilon_n \to 0$. It remains to show that

$$\lim_{n\to\infty}\Big|\frac{\mathcal{R}_\alpha(F_n + \varepsilon_n v_n) - \mathcal{R}_\alpha(F_n)}{\varepsilon_n} - \dot{\mathcal{R}}_{\alpha;F}(v)\Big| = 0,$$

that is,

$$\lim_{n\to\infty}\Big|\int\Big(\frac{g_\alpha\big(F_n(x)\big) - g_\alpha\big((F_n + \varepsilon_n v_n)(x)\big)}{\varepsilon_n} - \big(-g'_\alpha(F(x))v(x)\big)\Big)\,dx\Big| = 0. \tag{14}$$

Let us denote the integrand of the integral in Equation (14) by $I_n(x)$. In virtue of $F_n \to F$ pointwise, $\|v_n - v\|_\phi \to 0$, $\varepsilon_n \to 0$, and

$$|(F_n + \varepsilon_n v_n)(x) - F(x)| \le |F_n(x) - F(x)| + \varepsilon_n|v_n(x) - v(x)| + \varepsilon_n|v(x)|,$$

we have $\lim_{n\to\infty} F_n(x) = F(x)$ and $\lim_{n\to\infty}(F_n(x) + \varepsilon_n v_n(x)) = F(x)$ for every $x \in \mathbb{R}$. Thus, for every $x \in \mathbb{R}$ with $F(x) < alpha$, we obtain $g'_\alpha(F(x))v(x) = 0$ and

$$\frac{g_\alpha\big(F_n(x)\big) - g_\alpha\big((F_n + \varepsilon_n v_n)(x)\big)}{\varepsilon_n} = 0 \qquad \text{for sufficiently large } n,$$

i.e., $\lim_{n\to\infty} I_n(x) = 0$. Moreover, for every $x \in \mathbb{R}$ with $F(x) > \alpha$, we obtain $g'_\alpha(F(x))v(x) = \frac{1}{1-\alpha}v(x)$ and

$$\frac{g_\alpha\big(F_n(x)\big) - g_\alpha\big((F_n + \varepsilon_n v_n)(x)\big)}{\varepsilon_n} = -\frac{v_n(x)}{1-\alpha} \qquad \text{for sufficiently large } n,$$

i.e., $\lim_{n\to\infty} I_n(x) = 0$. Since we assumed that F takes the value α only once, we can conclude that $\lim_{n\to\infty} I_n(x) = 0$ for Lebesgue-a.e. $x \in \mathbb{R}$. Moreover, by the Lipschitz continuity of g_α with Lipschitz constant $\frac{1}{1-\alpha}$ we have

$$\begin{aligned}
|I_n(x)| &= |I_n(x)|\,\phi(x)\,\phi(x)^{-1} \\
&= \Big|\frac{g_\alpha\big(F_n(x)\big) - g_\alpha\big((F_n + \varepsilon_n v_n)(x)\big)}{\varepsilon_n} + g'_\alpha(F(x))v(x)\Big|\,\phi(x)\,\phi(x)^{-1} \\
&\le \frac{1}{1-\alpha}\big(\|v_n\|_\phi + \|v\|_\phi\big)\,\phi(x)^{-1} \\
&\le \frac{1}{1-\alpha}\big(\sup_{n\in\mathbb{N}}\|v_n\|_\phi + \|v\|_\phi\big)\,\phi(x)^{-1}.
\end{aligned}$$

Since $\sup_{n\in\mathbb{N}}\|v_n\|_\phi < \infty$ (recall $\|v_n - v\|_\phi \to 0$), the assumption $\int 1/\phi(x)\,dx < \infty$ ensures that the latter expression provides a Borel measurable majorant of I_n. Now, the Dominated Convergence theorem implies Equation (14). □

As an immediate consequence of Corollary A4, Examples A1 and A2, and Proposition 1, we obtain the following corollary.

Corollary 1. *Let F, $\widehat{F}_n$, $\widehat{F}_n^*$, $\widehat{C}_n$, and B_F be as in Example A1 (S1. or S2.) or as in Example A2 respectively, and assume that the assumptions discussed in Example A1 or in Example A2 respectively are fulfilled for some weight function ϕ with $\int 1/\phi(x)\,dx < \infty$ (in particular $F \in \mathbf{F}_1$). Moreover, assume that F takes the value α only once. Then,*

$$\sqrt{n}\big(\mathcal{R}_\alpha(\widehat{F}_n) - \mathcal{R}_\alpha(F)\big) \rightsquigarrow \dot{\mathcal{R}}_{\alpha;F}(B_F) \qquad \text{in } (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

and

$$\sqrt{n}\big(\mathcal{R}_\alpha(\widehat{F}_n^*(\omega,\cdot)) - \mathcal{R}_\alpha(\widehat{C}_n(\omega))\big) \rightsquigarrow \dot{\mathcal{R}}_{\alpha;F}(B_F) \qquad \text{in } (\mathbb{R}, \mathcal{B}(\mathbb{R})), \qquad \mathbb{P}\text{-a.e. } \omega.$$

3.2. Compound Distribution Functional

Let $\mathcal{C}_p : \mathbf{F} \to \mathbf{F}$ be the compound distribution functional introduced in Section 2.1. For any $\lambda \geq 0$, let the function $\phi_\lambda : \mathbb{R} \to [1, \infty)$ be defined by $\phi_\lambda(x) := (1 + |x|)^\lambda$ and denote by $\mathbf{F}_{\phi_\lambda}$ the set of all distribution functions F that satisfy $\int \phi_\lambda(x)\,dF(x) < \infty$. Using the terminology of Part (ii) of Definition A1, we obtain the following Proposition 2. In the proposition, the functional $\mathcal{C}_p$ is restricted to the domain $\mathbf{F}_{\phi_\lambda}$ in order to obtain $\mathbf{D}_{\phi_{\lambda'}}$ as the corresponding trace. The latter will be important for Corollary 3.

Proposition 2. *Let $\lambda > \lambda' \geq 0$ and $F \in \mathbf{F}_{\phi_\lambda}$. Assume that $\sum_{k=1}^\infty p_k\, k^{(1+\lambda)\vee 2} < \infty$. Then, the map $\mathcal{C}_p : \mathbf{F}_{\phi_\lambda}(\subseteq \mathbf{D}) \to \mathbf{F}(\subseteq \mathbf{D})$ is uniformly quasi-Hadamard differentiable at F tangentially to $\mathbf{D}_{\phi_\lambda}\langle \mathbf{D}_{\phi_\lambda}\rangle$ with trace $\mathbf{D}_{\phi_{\lambda'}}$. Moreover, the uniform quasi-Hadamard derivative $\dot{\mathcal{C}}_{p;F} : \mathbf{D}_{\phi_\lambda} \to \mathbf{D}_{\phi_{\lambda'}}$ is given by*

$$\dot{\mathcal{C}}_{p;F}(v)(\cdot) := v * H_{p,F}(\,\cdot\,) := \int v(\,\cdot\, - x)\,dH_{p,F}(x), \tag{15}$$

where as before $H_{p,F} := \sum_{k=1}^\infty k\, p_k F^{*(k-1)}$. In particular, if $p_m = 1$ for some $m \in \mathbb{N}$, then

$$\dot{\mathcal{C}}_{p;F}(v)(\cdot) = m \int v(\,\cdot\, - x)\,dF^{*(m-1)}(x).$$

Proposition 2 extends Proposition 4.1 of Pitts (1994). Before we prove the proposition, we note that the proposition together with Corollary A4 and Examples A1 and A2 yields the following corollary.

Corollary 2. *Let F, $\widehat{F}_n$, $\widehat{F}_n^*$, $\widehat{C}_n$, and B_F be as in Example A1 (S1. or S2.) or as in Example A2 respectively, and assume that the assumptions discussed in Example A1 or in Example A2 respectively are fulfilled for $\phi = \phi_\lambda$ for some $\lambda > 0$. Then, for $\lambda' \in [0, \lambda)$*

$$\sqrt{n}\big(\mathcal{C}_p(\widehat{F}_n) - \mathcal{C}_p(F)\big) \rightsquigarrow^\circ \dot{\mathcal{C}}_{p;F}(B_F) \qquad \text{in } (\mathbf{D}_{\phi'_\lambda}, \mathcal{B}^\circ_{\phi'_\lambda}, \|\cdot\|_{\phi_{\lambda'}})$$

and

$$\sqrt{n}\big(\mathcal{C}_p(\widehat{F}_n^*(\omega,\cdot)) - \mathcal{C}_p(\widehat{C}_n(\omega))\big) \rightsquigarrow^\circ \dot{\mathcal{C}}_{p;F}(B_F) \qquad \text{in } (\mathbf{D}_{\phi'_\lambda}, \mathcal{B}^\circ_{\phi'_\lambda}, \|\cdot\|_{\phi_{\lambda'}}), \qquad \mathbb{P}\text{-a.e. } \omega.$$

To ease the exposition of the proof of Proposition 2, we first state a lemma that follows from results given in Pitts (1994). In the sequel we use $f * H$ to denote the function defined by $f * H(\cdot) := \int v(\,\cdot\, - x)\,dH(x)$ for any measurable function f and any distribution function H of a finite (not necessarily probability) Borel measure on $\mathbb{R}$ for which $f * H(\cdot)$ is well defined on $\mathbb{R}$.

Lemma 1. *Let $\lambda > \lambda' \geq 0$, and $(F_n) \subseteq \mathbf{F}_{\phi_\lambda}$ and $(G_n) \subseteq \mathbf{F}_{\phi_\lambda}$ be any sequences such that $\|F_n - F\|_{\phi_\lambda} \to 0$ and $\|G_n - G\|_{\phi_\lambda} \to 0$ for some $F, G \in \mathbf{F}_{\phi_\lambda}$. Then, the following two assertions hold.*

(i) *There exists a constant $C_1 > 0$ such that for every $k, n \in \mathbb{N}$*

$$\|\mathbb{1}_{[0,\infty)} - F_n^{*k}\|_{\phi_{\lambda'}} \leq (2^{\lambda'-1} \vee 1)(1 + k^{\lambda' \vee 1} C_1).$$

(ii) *For every $v \in \mathbf{D}_{\phi_{\lambda'}}$ there exists a constant $C_2 > 0$ such that for every $k, \ell, n \in \mathbb{N}$*

$$\|v * (F_n^{*k} * G_n^{*\ell})\|_{\phi_{\lambda'}} \leq 2^{\lambda'}\big(1 + 2^{\lambda'}(2^{\lambda'-1} \vee 1)(2 + (k+\ell)^{\lambda' \vee 1} C_2)\big)\|v\|_{\phi_{\lambda'}}.$$

Proof. (i): From Equation (2.4) in Pitts (1994) we have

$$\|\mathbb{1}_{[0,\infty)} - F_n^{*k}\|_{\phi_{\lambda'}} \leq (2^{\lambda'-1} \vee 1)\Big(1 + k^{\lambda' \vee 1} \int |x|^{\lambda'} \, dF_n(x)\Big),$$

so that it remains to show that $\int |x|^{\lambda'} \, dF_n(x)$ is bounded above uniformly in $n \in \mathbb{N}$. The functions $\mathbb{1}_{[0,\infty)} - F_n$ and $\mathbb{1}_{[0,\infty)} - F$ both lie in $\mathbf{D}_{\phi_\lambda}$, because $F_n, F \in \mathbf{F}_{\phi_\lambda}$. Along with $\|F_n - F\|_{\phi_\lambda} \to 0$, this implies $\int |x|^{\lambda'} \, dF_n(x) \to \int |x|^{\lambda'} \, dF(x)$ (see Lemma 2.1 in Pitts (1994)). Therefore, $\int |x|^{\lambda'} \, dF_n(x) \leq C_1$ for some suitable finite constant $C_1 > 0$ and all $n \in \mathbb{N}$.

(ii): With the help of Lemma 2.3 of Pitts (1994) (along with $\|F_n^{*k} * G_n^{*\ell}\|_\infty = 1$), Lemma 2.4 of Pitts (1994), and Equation (2.4) in Pitts (1994), we obtain

$$\begin{aligned}
&\|v * (F_n^{*k} * G_n^{*\ell})\|_{\phi_{\lambda'}} \\
&\leq \; 2^{\lambda'}\|v\|_{\phi_{\lambda'}}\big(1 + \|\mathbb{1}_{[0,\infty)} - F_n^{*k} * G_n^{*\ell}\|_{\phi_{\lambda'}}\big) \\
&\leq \; 2^{\lambda'}\|v\|_{\phi_{\lambda'}}\big(1 + 2^{\lambda'}\big(\|\mathbb{1}_{[0,\infty)} - F_n^{*k}\|_{\phi_{\lambda'}} + \|\mathbb{1}_{(0,\infty)-} G_n^{*\ell}\|_{\phi_{\lambda'}}\big)\big) \\
&\leq \; 2^{\lambda'}\|v\|_{\phi_{\lambda'}}\Big(1 + 2^{\lambda'}(2^{\lambda'-1} \vee 1)\Big(1 + k^{\lambda' \vee 1} \int |x|^{\lambda'} \, dF_n(x) + 1 + \ell^{\lambda' \vee 1} \int |x|^{\lambda'} \, dG_n(x)\Big)\Big).
\end{aligned}$$

It hence remains to show that $\int |x|^{\lambda'} \, dF_n(x)$ and $\int |x|^{\lambda'} \, dG_n(x)$ are bounded above uniformly in $n \in \mathbb{N}$. However, this was already done in the proof of Part (i). □

Proof. Proof of Proposition 2. First, note that for $G_1, G_2 \in \mathbf{F}_{\phi_\lambda}$, we have

$$\begin{aligned}
\|\mathcal{C}_p(G_1) - \mathcal{C}_p(G_2)\|_{\phi_{\lambda'}} &\leq \|\mathcal{C}_p(G_1) - \mathbb{1}_{[0,\infty)}\|_{\phi_{\lambda'}} + \|\mathbb{1}_{[0,\infty)} - \mathcal{C}_p(G_2)\|_{\phi_{\lambda'}} \\
&\leq \int (1 + |x|)^{\lambda'} \, d\mathcal{C}_p(G_1)(x) + \int (1 + |x|)^{\lambda'} \, d\mathcal{C}_p(G_2)(x)
\end{aligned}$$

by Equation (2.1) in Pitts (1994). Moreover, according to Lemma 2.2 in Pitts (1994), we have that the integrals $\int |x|^{\lambda'} d\mathcal{C}_p(F)(x)$ and $\int |x|^{\lambda'} d\mathcal{C}_p(G)(x)$ are finite under the assumptions of the proposition. Hence, $\mathbf{D}_{\phi_{\lambda'}}$ can indeed be seen as the trace.

Second, we show $(\|\cdot\|_{\phi_\lambda}, \|\cdot\|_{\phi_{\lambda'}})$-continuity of the map $\dot{\mathcal{C}}_{p;F} : \mathbf{D}_{\phi_\lambda} \to \mathbf{D}_{\phi_{\lambda'}}$. To this end let $v \in \mathbf{D}_{\phi_\lambda}$ and $(v_n) \subseteq \mathbf{D}_{\phi_\lambda}$ such that $\|v_n - v\|_{\phi_\lambda} \to 0$. For every $k \in \mathbb{N}$, we have

$$\begin{aligned}
&\|p_k k(v_n - v) * F^{*(k-1)}\|_{\phi_{\lambda'}} \\
&\leq \; 2^{\lambda'}\|v_n - v\|_{\phi_{\lambda'}} \, p_k \, k\big(\|\mathbb{1}_{[0,\infty)} \, \|F^{*(k-1)}\|_\infty - F^{*(k-1)}\|_{\phi_{\lambda'}} + \|F^{*(k-1)}\|_\infty\big) \\
&= \; 2^{\lambda'}\|v_n - v\|_{\phi_{\lambda'}} \, p_k \, k\big(\|\mathbb{1}_{[0,\infty)} - F^{*(k-1)}\|_{\phi_{\lambda'}} + 1\big) \\
&\leq \; 2^{\lambda'}\|v_n - v\|_{\phi_{\lambda'}} \, p_k \, k\Big((2^{\lambda'-1} \vee 1)\Big(1 + (k-1)^{\lambda' \vee 1} \int |x|^{\lambda'} \, dF(x)\Big) + 1\Big),
\end{aligned}$$

where the first and the second inequality follow from Lemma 2.3 and Equation (2.4) in Pitts (1994) respectively. Hence,

$$\begin{aligned}\|\dot{\mathcal{C}}_{p;F}(v_n) - \dot{\mathcal{C}}_{p;F}(v)\|_{\phi_{\lambda'}} &= \|v_n * H_{p,F} - v * H_{p,F}\|_{\phi_{\lambda'}} \\ &\leq 2^{\lambda'}\|v_n - v\|_{\phi_{\lambda'}} \sum_{k=1}^{\infty} p_k\, k\Big((2^{\lambda'-1} \vee 1)\Big(1 + (k-1)^{\lambda' \vee 1} \int |x|^{\lambda'}\, dF(x)\Big) + 1\Big).\end{aligned}$$

Now, the series converges due to the assumptions, and $\|v_n - v\|_{\phi_\lambda} \to 0$ implies $\|v_n - v\|_{\phi_{\lambda'}} \to 0$. Thus, $\|\dot{\mathcal{C}}_{p;F}(v_n) - \dot{\mathcal{C}}_{p;F}(v)\|_{\phi_{\lambda'}} \to 0$, which proves continuity.

Third, let $((F_n), v, (v_n), (\varepsilon_n))$ be a quadruple with $(F_n) \subseteq \mathbf{F}_{\phi_\lambda}$ satisfying $\|F_n - F\|_{\phi_\lambda} \to 0$, $v \in \mathbf{D}_{\phi_\lambda}$, $(v_n) \subseteq \mathbf{D}_{\phi_\lambda}$ satisfying $\|v_n - v\|_{\phi_\lambda} \to 0$ and $(F_n + \varepsilon_n v_n) \subseteq \mathbf{F}_{\phi_\lambda}$, and $(\varepsilon_n) \subseteq (0, \infty)$ satisfying $\varepsilon_n \to 0$. It remains to show that

$$\lim_{n\to\infty} \Big\| \frac{\mathcal{C}_p(F_n + \varepsilon_n v_n) - \mathcal{C}_p(F_n)}{\varepsilon_n} - \dot{\mathcal{C}}_{p;F}(v) \Big\|_{\phi_{\lambda'}} = 0.$$

To do so, define for $k \in \mathbb{N}_0$ a map $H_k : \mathbf{F} \times \mathbf{F} :\to \mathbf{F}$ by

$$H_k(G_1, G_2) := \sum_{j=0}^{k-1} G_1^{*(k-1-j)} * G_2^{*j}$$

with the usual convention that the sum over the empty set equals zero. We find that for every $M \in \mathbb{N}$

$$\begin{aligned}&\Big\| \frac{\mathcal{C}_p(F_n + \varepsilon_n v_n) - \mathcal{C}_p(F_n)}{\varepsilon_n} - \dot{\mathcal{C}}_{p;F}(v) \Big\|_{\phi_{\lambda'}} \\ &= \Big\| \frac{1}{\varepsilon_n}\Big(\sum_{k=0}^{\infty} p_k (F_n + \varepsilon_n v_n)^{*k} - \sum_{k=0}^{\infty} p_k F_n^{*k} \Big) - \dot{\mathcal{C}}_{p;F}(v) \Big\|_{\phi_{\lambda'}} \\ &= \Big\| \frac{1}{\varepsilon_n}\Big(\sum_{k=1}^{\infty} \big(p_k (F_n + \varepsilon_n v_n)^{*k} - p_k F_n^{*k} \big) \Big) - \dot{\mathcal{C}}_{p;F}(v) \Big\|_{\phi_{\lambda'}} \\ &= \Big\| \sum_{k=1}^{\infty} p_k v_n * H_k(F_n + \varepsilon_n v_n, F_n) - \dot{\mathcal{C}}_{p;F}(v) \Big\|_{\phi_{\lambda'}} \\ &\leq \Big\| \sum_{k=M+1}^{\infty} p_k v_n * H_k(F_n + \varepsilon_n v_n, F_n) \Big\|_{\phi_{\lambda'}} + \Big\| \sum_{k=1}^{M} p_k (v_n - v) * H_k(F_n + \varepsilon_n v_n, F_n) \Big\|_{\phi_{\lambda'}} \\ &\quad + \Big\| v * \sum_{k=M+1}^{\infty} k p_k F^{*(k-1)} \Big\|_{\phi_{\lambda'}} + \Big\| \sum_{k=1}^{M} p_k v * H_k(F_n + \varepsilon_n v_n, F_n) - k p_k v * F^{*(k-1)} \Big\|_{\phi_{\lambda'}} \\ &=: S_1(n, M) + S_2(n, M) + S_3(M) + S_4(n, M),\end{aligned}$$

where for the third "=" we use the fact that for $G_1, G_2 \in \mathbf{F}$

$$(G_1 - G_2) * H_k(G_1, G_2) = G_1^{*k} - G_2^{*k}. \tag{16}$$

By Part (ii) of Lemma reflemma preceding qHD of compound (this lemma can be applied since $\|F_n + \varepsilon_n v_n - F\|_{\phi_\lambda} \to 0$) there exists a constant $C_2 > 0$ such that for all $n \in \mathbb{N}$

$$\begin{aligned}S_1(n, M) &= \Big\| \sum_{k=M+1}^{\infty} p_k v_n * H_k(F_n + \varepsilon_n v_n, F_n) \Big\|_{\phi_{\lambda'}} \\ &\leq 2^{\lambda'} \|v_n\|_{\phi_{\lambda'}} \sum_{k=M+1}^{\infty} p_k\, k\big(1 + 2^{\lambda'}(2^{\lambda'-1} \vee 1)\big(2 + (k-1)^{\lambda' \vee 1} C_2\big)\big).\end{aligned} \tag{17}$$

Since $\lambda' < \lambda$ and $\|v_n - v\|_{\phi_\lambda} \to 0$, we have $\|v_n\|_{\phi_{\lambda'}} \leq K_1$ for some finite constant $K_1 > 0$ and all $n \in \mathbb{N}$. Hence, the right-hand side of Equation (17) can be made arbitrarily small by choosing M large enough. That is, $S_1(n, M)$ can be made arbitrarily small uniformly in $n \in \mathbb{N}$ by choosing M large enough.

Furthermore, it is demonstrated in the proof of Proposition 4.1 of Pitts (1994) that $S_3(M)$ can be made arbitrarily small by choosing M large enough.

Next, applying again Part (ii) of Lemma 1, we obtain

$$\begin{aligned} S_2(n, M) &= \Big\| \sum_{k=1}^{M} p_k (v_n - v) * H_k(F_n + \varepsilon_n v_n, F_n) \Big\|_{\phi_{\lambda'}} \\ &\leq 2^{\lambda'} \sum_{k=1}^{M} p_k \, k \, \|v_n - v\|_{\phi_{\lambda'}} \big(1 + 2^{\lambda'} (2^{\lambda'-1} \vee 1) \big(2 + (k-1)^{\lambda' \vee 1} C_2\big)\big). \end{aligned}$$

Using $\|v_n - v\|_{\phi_{\lambda'}} \leq \|v_n - v\|_{\phi_\lambda} \to 0$, this term tends to zero as $n \to \infty$ for a given M.

It remains to consider the summand

$$\begin{aligned} S_4(n, M) &= \Big\| \sum_{k=1}^{M} p_k v * H_k(F_n + \varepsilon_n v_n, F_n) - k p_k v * F^{*(k-1)} \Big\|_{\phi_{\lambda'}} \\ &= \Big\| \sum_{k=1}^{M} p_k \sum_{\ell=0}^{k-1} \Big(v * (F_n + \varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell} - v * F^{*(k-1)} \Big) \Big\|_{\phi_{\lambda'}}. \end{aligned}$$

We show that for M fixed this term can be made arbitrarily small by letting $n \to \infty$. This would follow if for every given $k \in \{1, \ldots, M\}$ and $\ell \in \{0, \ldots, k-1\}$ the expression

$$\|v * (F_n + \varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell} - v * F^{*(k-1)}\|_{\phi_{\lambda'}}$$

could be made arbitrarily small by letting $n \to \infty$. For every such k and ℓ we can find a linear combination of indicator functions of the form $\mathbb{1}_{[a,b)}$, $-\infty < a < b < \infty$, which we denote by $\widetilde{v}$, such that

$$\begin{aligned} &\|v * (F_n + \varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell} - v * F^{*(k-1)}\|_{\phi_{\lambda'}} \\ &\quad \leq \|v * (F_n + \varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell} - \widetilde{v} * (F_n + \varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell}\|_{\phi_{\lambda'}} \\ &\qquad + \|\widetilde{v} * (F_n + \varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell} - \widetilde{v} * F^{*(k-1)}\|_{\phi_{\lambda'}} \\ &\qquad + \|\widetilde{v} * F^{*(k-1)} - v * F^{*(k-1)}\|_{\phi_{\lambda'}} \\ &\quad \leq 2^{\lambda'} \|\widetilde{v} - v\|_{\phi_{\lambda'}} \big(\|\mathbb{1}_{[0,\infty)} - (F_n + \varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell}\|_{\phi_{\lambda'}} + 1\big) \\ &\qquad + c(\lambda', \widetilde{v}) \, \|(F_n + \varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell} - F^{*(k-1)}\|_{\phi_{\lambda'}} \\ &\qquad + 2^{\lambda'} \|\widetilde{v} - v\|_{\phi_{\lambda'}} \big(\|\mathbb{1}_{[0,\infty)} - F^{*(k-1)}\|_{\phi_{\lambda'}} + 1\big) \end{aligned} \tag{18}$$

for some suitable finite constant $c(\lambda', \widetilde{v}) > 0$ depending only on λ' and $\widetilde{v}$. The first inequality in Equation (18) is obvious (and holds for any $\widetilde{v} \in \mathbf{D}_{\phi_{\lambda'}}$). The second inequality in Equation (18) is obtained by applying Lemma 2.3 of Pitts (1994) to the first summand (noting that $\|(F_n + \varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell}\|_\infty = 1$; recall $F_n + \varepsilon_n v_n \in \mathbf{F}$), by applying Lemma 4.3 of Pitts (1994) to the second summand (which requires that $\widetilde{v}$ is as described above), and by applying Lemma 2.3 of Pitts (1994) to the third summand.

We now consider the three summands on the right-hand side of Equation (18) separately. We start with the third term. Since $v \in \mathbf{D}_{\phi_\lambda}$, Lemma 4.2 of Pitts (1994) ensures that we may assume that $\widetilde{v}$ is chosen such that $\|\widetilde{v} - v\|_{\phi_{\lambda'}}$ is arbitrarily small. Hence, for fixed M the third summand in Equation (18) can be made arbitrarily small.

We next consider the the second summand in Equation (18). Obviously,

$$\begin{aligned}
&\|(F_n+\varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell} - F^{*(k-1)}\|_{\phi_{\lambda'}} \\
&= \|(F_n+\varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell} - F_n^{*(k-1)} + F_n^{*(k-1)} - F^{*(k-1)}\|_{\phi_{\lambda'}} \\
&\leq \left\|\left((F_n+\varepsilon_n v_n)^{*(k-1-\ell)} - F_n^{*(k-1-\ell)}\right) * F_n^{*\ell}\right\|_{\phi_{\lambda'}} + \|F_n^{*(k-1)} - F^{*(k-1)}\|_{\phi_{\lambda'}}.
\end{aligned} \tag{19}$$

We start by considering the first summand in Equation (19). In view of Equation (16), it can be written as

$$\begin{aligned}
&\left\|\left((F_n+\varepsilon_n v_n)^{*(k-1-\ell)} - F_n^{*(k-1-\ell)}\right) * F_n^{*\ell}\right\|_{\phi_{\lambda'}} \\
&= \left\|\left((F_n+\varepsilon_n v_n - F_n) * H_{k-1-\ell}(F_n+\varepsilon_n v_n, F_n)\right) * F_n^{*\ell}\right\|_{\phi_{\lambda'}} \\
&= \left\|\left(\varepsilon_n v_n * H_{k-1-\ell}(F_n+\varepsilon_n v_n, F_n)\right) * F_n^{*\ell}\right\|_{\phi_{\lambda'}}.
\end{aligned}$$

Applying Lemma 2.3 of Pitts (1994) with $f = \varepsilon_n v_n * H_{k-\ell-1}(F_n+\varepsilon_n v_n, F_n)$ and $H = F_n^{*\ell}$ we obtain

$$\begin{aligned}
&\left\|\left(\varepsilon_n v_n * H_{k-1-\ell}(F_n+\varepsilon_n v_n, F_n)\right) * F_n^{*\ell}\right\|_{\phi_{\lambda'}} \\
&\leq 2^{\lambda'}\left\|\left(\varepsilon_n v_n * H_{k-\ell-1}(F_n+\varepsilon_n v_n, F_n)\right)\right\|_{\phi_{\lambda'}}\left(\|\mathbb{1}_{[0,\infty)}\|F_n^{*\ell}\|_\infty - F_n^{*\ell}\|_{\phi_{\lambda'}} + \|F_n^{*\ell}\|_\infty\right) \\
&= 2^{\lambda'}\left\|\left(\varepsilon_n v_n * H_{k-\ell-1}(F_n+\varepsilon_n v_n, F_n)\right)\right\|_{\phi_{\lambda'}}\left(\|\mathbb{1}_{[0,\infty)} - F_n^{*\ell}\|_{\phi_{\lambda'}} + 1\right) \\
&\leq 2^{\lambda'}\left\|\left(\varepsilon_n v_n * H_{k-\ell-1}(F_n+\varepsilon_n v_n, F_n)\right)\right\|_{\phi_{\lambda'}}\left\{(2^{\lambda'-1}\vee 1)\left(1+\ell^{\lambda'\vee 1}C_1\right)+1\right\},
\end{aligned} \tag{20}$$

where we applied Part (i) of Lemma 1 to $\|\mathbb{1}_{[0,\infty)} - F_n^{*\ell}\|_{\phi_{\lambda'}}$ to obtain the last inequality. Hence, for the left-hand side of Equation (20) to go to zero as $n \to \infty$ it suffices to show that $\|(\varepsilon_n v_n * H_{k-\ell-1}(F_n + \varepsilon_n v_n, F_n))\|_{\phi_{\lambda'}} \to 0$ as $n \to \infty$. The latter follows from

$$\begin{aligned}
&\left\|\left(\varepsilon_n v_n * H_{k-\ell-1}(F_n+\varepsilon_n v_n, F_n)\right)\right\|_{\phi_{\lambda'}} \\
&\leq 2^{\lambda'}(k-\ell-1)\varepsilon_n\|v_n\|_{\phi_{\lambda'}}\left(1+2^{\lambda'}(2^{\lambda'-1}\vee 1)\left(2+((k-\ell-2))^{\lambda'\vee 1}C_2\right)\right),
\end{aligned} \tag{21}$$

where we applied Part (ii) of Lemma 1 with $v = \varepsilon_n v_n$ to all summands in $H_{k-\ell-1}(F_n + \varepsilon_n v_n, F_n)$. For every k and $\ell \in \{0, \ldots, k-1\}$ this expression goes indeed to zero as $n \to \infty$, because, as mentioned before, $\|v_n\|_{\phi_{\lambda'}}$ is uniformly bounded in $n \in \mathbb{N}$, and we have $\varepsilon_n \to 0$. Next, we consider the second summand in Equation (19). Applying Equation (16) to $F_n^{*(k-1)}$ and $F^{*(k-1)}$ and subsequently Part (ii) of Lemma 1 to the summands in $H_{k-1}(F_n, F)$, we have

$$\|F_n^{*(k-1)} - F^{*(k-1)}\|_{\phi_{\lambda'}} \leq 2^{\lambda'}(k-1)\|F_n - F\|_{\phi_{\lambda'}}\left(1+2^{\lambda'}(2^{\lambda'-1}\vee 1)(2+((k-2))^{\lambda'\vee 1}C_2)\right).$$

Clearly for every k this term goes to zero 0 as $n \to \infty$, because $\|F_n - F\|_{\phi_{\lambda'}} \leq \|F_n - F\|_{\phi_\lambda} \to 0$ as $n \to \infty$ by assumption. This together with the fact that Equation (20) goes to zero 0 as $n \to \infty$ shows that Equation (19) goes to zero in $\|\cdot\|_{\phi_{\lambda'}}$ as $n \to \infty$. Therefore, the second summand in Equation (18) goes to zero as $n \to \infty$.

It remains to consider the first term in Equation (18). We find

$$\begin{aligned}
&2^{\lambda'}\|\widetilde{v} - v\|_{\phi_\lambda}\left(\|\mathbb{1}_{[0,\infty)} - (F_n+\varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell}\|_{\phi_{\lambda'}} + 1\right) \\
&\leq 2^{\lambda'}\|\widetilde{v} - v\|_{\phi_{\lambda'}}\left(\|\mathbb{1}_{[0,\infty)} - (F_n+\varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell}\|_{\phi_{\lambda'}} + 1\right) \\
&\leq 2^{\lambda'}\|\widetilde{v} - v\|_{\phi_{\lambda'}}\left(\|\mathbb{1}_{[0,\infty)} - F^{*(k-1)} + F^{*(k-1)} - (F_n+\varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell}\|_{\phi_{\lambda'}} + 1\right) \\
&\leq 2^{\lambda'}\|\widetilde{v} - v\|_{\phi_{\lambda'}}\left(\|\mathbb{1}_{[0,\infty)} - F^{*(k-1)}\|_{\phi_{\lambda'}} + \|F^{*(k-1)} - (F_n+\varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell}\|_{\phi_{\lambda'}} + 1\right) \\
&\leq 2^{\lambda'}\|\widetilde{v} - v\|_{\phi_{\lambda'}}(2^{\lambda'-1}\vee 1)\left(1+k^{\lambda\vee 1}\int |x|^{\lambda'}\,dF(x)\right) \\
&\quad + 2^{\lambda'}\|\widetilde{v} - v\|_{\phi_{\lambda'}}\left(\left\|F^{*(k-1)} - (F_n+\varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell}\right\|_{\phi_{\lambda'}} + 1\right),
\end{aligned} \tag{22}$$

where for the last inequality we used Formula (2.4) of Pitts (1994). In the following, Equation (19) we showed that $\|F^{*(k-1)} - (F_n + \varepsilon_n v_n)^{*(k-1-\ell)} * F_n^{*\ell}\|_{\phi_{\lambda'}}$ goes to zero as $n \to \infty$ for every k and

$\ell \in \{0, \dots, k-1\}$. Hence, for every such k and ℓ, it is uniformly bounded in $n \in \mathbb{N}$. Therefore, we can make Equation (22) arbitrarily small by making $\|\widetilde{v} - v\|_{\phi_{\lambda'}}$ small which, as mentioned above, is possible according to Lemma 4.2 of Pitts (1994). This finishes the proof. □

3.3. Composition of Average Value at Risk Functional and Compound Distribution Functional

Here, we consider the composition of the Average Value at Risk functional $\mathcal{R}_\alpha$ defined in Equation (1) and the compound distribution functional $\mathcal{C}_p$ introduced in Section 2.1. As a consequence of Propositions 1 and 2, we obtain the following Corollary 3. Note that, for any $\lambda > 1$, Lemma 2.2 in Pitts (1994) yields $\mathcal{C}_p(\mathbf{F}_{\phi_\lambda}) \subseteq \mathbf{F}_1$ so that the composition $\mathcal{R}_\alpha \circ \mathcal{C}_p$ is well defined on $\mathbf{F}_{\phi_\lambda}$.

Corollary 3. *Let $\lambda > 1$ and assume $\sum_{k=1}^\infty p_k\, k^{1+\lambda} < \infty$. Let $F \in \mathbf{F}_{\phi_\lambda}$, and assume that $\mathcal{C}_p(F)$ takes the value α only once. Then, the map $T_{\alpha,p} := \mathcal{R}_\alpha \circ \mathcal{C}_p : \mathbf{F}_{\phi_\lambda}(\subseteq \mathbf{D}) \to \mathbb{R}$ is uniformly quasi-Hadamard differentiable at F tangentially to $\mathbf{D}_{\phi_\lambda}\langle \mathbf{D}_{\phi_\lambda}\rangle$, and the uniform quasi-Hadamard derivative $\dot{T}_{\alpha,p;F} : \mathbf{D}_{\phi_\lambda} \to \mathbb{R}$ is given by $\dot{T}_{\alpha,p;F} = \dot{\mathcal{R}}_{\alpha;\mathcal{C}_p(F)} \circ \dot{\mathcal{C}}_{p;F}$, i.e.,*

$$\dot{T}_{\alpha,p;F}(v) = \int g'_\alpha(\mathcal{C}_p(F)(x))(v * H_{p,F})(x)\, dx \qquad \textit{for all } v \in \mathbf{D}_{\phi_\lambda}$$

with g'_α and $v * H_{p,F}$ as in Proposition 1 and 2, respectively.

Proof. We intend to apply Lemma A1 to $H = \mathcal{C}_p : \mathbf{F}_{\phi_\lambda} \to \mathbf{F}_1$ and $\widetilde{H} = \mathcal{R}_\alpha : \mathbf{F}_1 \to \mathbb{R}$. To verify that the assumptions of the lemma are fulfilled, we first recall from the comment directly before Corollary 3 that $\mathcal{C}_p(\mathbf{F}_{\phi_\lambda}) \subseteq \mathbf{F}_1$. It remains to show that the Assumptions (a)–(c) of Lemma A1 are fulfilled. According to Proposition 2 we have that for every $\lambda' \in (1, \lambda)$ the functional $\mathcal{C}_p$ is uniformly quasi-Hadamard differentiable at F tangentially to $\mathbf{D}_{\phi_\lambda}\langle \mathbf{D}_{\phi_\lambda}\rangle$ with trace $\mathbf{D}_{\phi_{\lambda'}}$, which is the first part of Assumption (b). The second part of Assumption (b) means $\dot{\mathcal{C}}_{p,F}(\mathbf{D}_{\phi_\lambda}) \subseteq \mathbf{D}_{\phi_{\lambda'}}$ and follows from

$$\begin{aligned}
\|\dot{\mathcal{C}}_{p;F}(v)\|_{\phi_{\lambda'}} &= \Big\| v * \sum_{k=1}^\infty p_k\, k F^{*(k-1)} \Big\|_{\phi_{\lambda'}} \\
&\le 2^{\lambda'} \|v\|_{\phi_{\lambda'}} \sum_{k=1}^\infty p_k\, k \Big(1 + (2^{\lambda'-1} \vee 1)\Big(1 + k^{\lambda' \vee 1} \int |x|^{\lambda'}\, dF(x)\Big)\Big)
\end{aligned}$$

(for which we applied Lemma 2.3 and Inequality (2.4) in Pitts (1994)), the convergence of the latter series (which holds by assumption), and $\|v\|_{\phi_{\lambda'}} \le \|v\|_{\phi_\lambda} < \infty$. Further, it follows from Proposition 1 that the map $\mathcal{R}_\alpha$ is uniformly quasi-Hadamard differentiable tangentially to $\mathbf{D}_{\phi_{\lambda'}}\langle \mathbf{D}_{\phi_{\lambda'}}\rangle$ at every distribution function of $\mathbf{F}_{\phi_{\lambda'}}$ that takes the value $1 - \alpha$ only once. This is Assumption (c) of Lemma A1.

It remains to show that Assumption (a) of Lemma A1 also holds true. In the present setting, Assumption (a) means that for every sequence $(F_n) \subseteq \mathbf{F}_{\phi_\lambda}$ with $\|F_n - F\|_{\phi_\lambda} \to 0$ we have $\mathcal{C}_p(F_n) \to \mathcal{C}_p(F)$ pointwise. We show that we even have $\|\mathcal{C}_p(F_n) - \mathcal{C}_p(F)\|_{\phi_{\lambda'}} \to 0$. Thus, let $(F_n) \subseteq \mathbf{F}_{\phi_\lambda}$. Then,

$$\begin{aligned}
\|\mathcal{C}_p(F_n) - \mathcal{C}_p(F)\|_{\phi_{\lambda'}} &= \Big\| \sum_{k=1}^\infty p_k (F_n^{*k} - F^{*k}) \Big\|_{\phi_{\lambda'}} \\
&= \Big\| (F_n - F) * \sum_{k=1}^\infty p_k H_k(F_n, F) \Big\|_{\phi_{\lambda'}} \\
&\le 2^{\lambda'} \|F_n - F\|_{\phi_{\lambda'}} \sum_{k=1}^\infty p_k\, k \big(1 + 2^{\lambda'}(2^{\lambda'-1} \vee 1)\big(2 + (k-1)^{\lambda' \vee 1} C_2\big)\big),
\end{aligned}$$

where we used Equation (16) for the second "=" and applied Part (ii) of Lemma 1 to the summands of H_k to obtain the latter inequality. Since the series converges, we obtain $\|\mathcal{C}_p(F_n) - \mathcal{C}_p(F)\|_{\phi_{\lambda'}} \to 0$ when assuming $\|F_n - F\|_{\phi_\lambda} \to 0$. □

As an immediate consequence of Corollary A4, Examples A1 and A2, and Corollary 3, we obtain the following corollary.

Corollary 4. *Let F, $\widehat{F}_n$, $\widehat{F}_n^*$, $\widehat{C}_n$, and B_F be as in Example A1 (S1. or S2.) or as in Example A2, respectively, and assume that the assumptions discussed in Example A1 or in Example A2 respectively are fulfilled for $\phi = \phi_\lambda$ for some $\lambda > 1$ (in particular $F \in \mathbf{F}_1$). Moreover, assume $\sum_{k=1}^{\infty} p_k k^{1+\lambda} < \infty$ and that $\mathcal{C}_p(F)$ takes the value α only once. Then,*

$$\sqrt{n}\big(T_{\alpha,p}(\widehat{F}_n) - T_{\alpha,p}(F)\big) \rightsquigarrow \dot{T}_{\alpha,p;F}(B_F) \qquad \text{in } (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

and

$$\sqrt{n}\big(T_{\alpha,p}(\widehat{F}_n^*(\omega,\cdot)) - T_{\alpha,p}(\widehat{C}_n(\omega))\big) \rightsquigarrow \dot{T}_{\alpha,p;F}(B_F) \qquad \text{in } (\mathbb{R}, \mathcal{B}(\mathbb{R})), \qquad \mathbb{P}\text{-a.e. } \omega.$$

4. Conclusions

In this paper, we consider the sub-additive risk measure Average Value at Risk and presented in Sections 2.1 and 2.2 results on *almost sure* bootstrap consistency for the corresponding empirical plug-in estimator based on i.i.d. or strictly stationary, geometrically β-mixing observations. Our results supplement those by Beutner and Zähle (2016) on bootstrap consistency *in probability* and those by Sun and Cheng (2018) on bootstrap consistency in probability for the *Tail Conditional Expectation* (which is not sub-additive). In Section 2.1, we also look at the case where one is interested in Average Value of Risk in the collective risk model. Note that one might interpret the collective risk model as a pooling of independent risks. In the context of Solvency II, pooling of risks has received increased attention (see, for example, Bølviken and Guillen 2017). However, one should keep in mind that our results of Section 2.1 can typically not be applied in the Solvency II context. In Solvency II applications risks are usually dependent, whereas in the collective risk model the different risks (claims) are assumed to be independent.

Appendix A. Convergence in Distribution$^\circ$

Let $(\mathbf{E}, d)$ be a metric space and $\mathcal{B}^\circ$ be the σ-algebra on $\mathbf{E}$ generated by the open balls $B_r(x) := \{y \in \mathbf{E} : d(x,y) < r\}$, $x \in \mathbf{E}$, $r > 0$. We refer to $\mathcal{B}^\circ$ as *open-ball σ-algebra*. If $(\mathbf{E}, d)$ is separable, then $\mathcal{B}^\circ$ coincides with the Borel σ-algebra $\mathcal{B}$. If $(\mathbf{E}, d)$ is not separable, then $\mathcal{B}^\circ$ might be strictly smaller than $\mathcal{B}$ and thus a continuous real-valued function on $\mathbf{E}$ is not necessarily $(\mathcal{B}^\circ, \mathcal{B}(\mathbb{R}))$-measurable. Let C_b° be the set of all bounded, continuous and $(\mathcal{B}^\circ, \mathcal{B}(\mathbb{R}))$-measurable real-valued functions on $\mathbf{E}$, and $\mathcal{M}_1^\circ$ be the set of all probability measures on $(\mathbf{E}, \mathcal{B}^\circ)$.

Let X_n be an $(\mathbf{E}, \mathcal{B}^\circ)$-valued random variable on some probability space $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$ for every $n \in \mathbb{N}_0$. Then, referring to Billingsley (1999, sct. 1.6), the sequence $(X_n) = (X_n)_{n\in\mathbb{N}}$ is said to *converge in distribution$^\circ$* to X_0 if

$$\int f \, d\mathbb{P} \circ X_n^{-1} \longrightarrow \int f \, d\mathbb{P}_0 \circ X_0^{-1} \qquad \text{for all } f \in C_b^\circ.$$

In this case, we write $X_n \rightsquigarrow^\circ X_0$. This is the same as saying that the sequence $(\mathbb{P}_n \circ X_n^{-1})$ converges to $\mathbb{P}_0 \circ X_0^{-1}$ in the weak$^\circ$ topology on $\mathcal{M}_1^\circ$; for details see Appendix A of Beutner and Zähle (2016). It is worth mentioning that two probability measures $\mu, \nu \in \mathcal{M}_1^\circ$ coincide if $\mu[\mathbf{E}_0] = \nu[\mathbf{E}_0] = 1$ for some separable $\mathbf{E}_0 \in \mathcal{B}^\circ$ and $\int f \, d\mu = \int f \, d\nu$ for all uniformly continuous $f \in C_b^\circ$ (see, for instance, (Billingsley 1999, Theorem 6.2)).

In Appendices A–C in Beutner and Zähle (2016), several properties of convergence in distribution$^\circ$ (and weak$^\circ$ convergence) have been discussed. The following two subsections complement this discussion.

Appendix A.1. Slutsky-Type Results for the Open-Ball σ-Algebra

For a sequence (X_n) of $(\mathbf{E}, \mathcal{B}^\circ)$-valued random variables that are all defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, the sequence (X_n) is said to *converge in probability*$^\circ$ to X_0 if the mappings $\omega \mapsto d(X_n(\omega), X_0(\omega))$, $n \in \mathbb{N}$, are $(\mathcal{F}, \mathcal{B}(\mathbb{R}_+))$-measurable and satisfy

$$\lim_{n\to\infty} \mathbb{P}[d(X_n, X_0) \geq \varepsilon] = 0 \quad \text{for all } \varepsilon > 0. \tag{A1}$$

In this case, we write $X_n \to^{\mathsf{p},\circ} X_0$. The superscript $^\circ$ points to the fact that measurability of the mapping $\omega \mapsto d(X_n(\omega), X_0(\omega))$ is a requirement of the definition (and not automatically valid). Note, however, that in the specific situation where $X_0 \equiv x_0$ for some $x_0 \in \mathbf{E}$, measurability of the mapping $\omega \mapsto d(X_n(\omega), X_0(\omega))$ does hold (see Lemma B.3 in Beutner and Zähle (2016)). In addition, note that the measurability always hold when $(\mathbf{E}, d)$ is separable; in this case, we also write $\to^{\mathsf{p}}$ instead of $\to^{\mathsf{p},\circ}$.

Theorem A1. *Let (X_n) and (Y_n) be two sequences of $(\mathbf{E}, \mathcal{B}^\circ)$-valued random variables on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and assume that the mapping $\omega \mapsto d(X_n(\omega), Y_n(\omega))$ is $(\mathcal{F}, \mathcal{B}(\mathbb{R}_+))$-measurable for every $n \in \mathbb{N}$. Let X_0 be an $(\mathbf{E}, \mathcal{B}^\circ)$-valued random variable on some probability space $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$ with $\mathbb{P}_0[X_0 \in \mathbf{E}_0] = 1$ for some separable $\mathbf{E}_0 \in \mathcal{B}^\circ$. Then, $X_n \rightsquigarrow^\circ X_0$ and $d(X_n, Y_n) \to^{\mathsf{p}} 0$ together imply $Y_n \rightsquigarrow^\circ X_0$.*

Proof. In view of $X_n \rightsquigarrow^\circ X$, we obtain for every fixed $f \in \mathrm{BL}_1^\circ$

$$\begin{aligned}
&\limsup_{n\to\infty} \left| \int f \, d\mathbb{P}_{Y_n} - \int f \, d\mathbb{P}_{X_0} \right| \\
&\leq \limsup_{n\to\infty} \left| \int f \, d\mathbb{P}_{Y_n} - \int f \, d\mathbb{P}_{X_n} \right| + \limsup_{n\to\infty} \left| \int f \, d\mathbb{P}_{X_n} - \int f \, d\mathbb{P}_{X_0} \right| \\
&\leq \limsup_{n\to\infty} \int |f(Y_n) - f(X_n)| \, d\mathbb{P}.
\end{aligned}$$

Since f lies in BL_1° and we assumed $d(X_n, Y_n) \to^{\mathsf{p}} 0$, we also have

$$\begin{aligned}
\limsup_{n\to\infty} \int |f(Y_n) - f(X_n)| \, d\mathbb{P} &\leq \limsup_{n\to\infty} \int |f(Y_n) - f(X_n)| \mathbb{1}_{\{d(X_n,Y_n)\geq\varepsilon\}} \, d\mathbb{P} + 2\varepsilon \\
&\leq 2 \limsup_{n\to\infty} \mathbb{P}[d(X_n, Y_n) \geq \varepsilon] + 2\varepsilon
\end{aligned}$$

for every $\varepsilon > 0$. Thus, $\limsup_{n\to\infty} \int |f(Y_n) - f(X_n)| \, d\mathbb{P} = 0$ which together with the Portmanteau theorem (in the form of (Beutner and Zähle 2016, Theorem A.4)) implies the claim. □

Set $\overline{\mathbf{E}} := \mathbf{E} \times \mathbf{E}$ and let $\overline{\mathcal{B}}^\circ$ be the σ-algebra on $\overline{\mathbf{E}}$ generated by the open balls with respect to the metric

$$\overline{d}((x_1, x_2), (y_1, y_2)) := \max\{d(x_1, y_1); d(x_2, y_2)\}.$$

Recall that $\overline{\mathcal{B}}^\circ \subseteq \mathcal{B}^\circ \otimes \mathcal{B}^\circ$, where the inclusion may be strict.

Corollary A1. *Let (X_n) and (Y_n) be two sequences of $(\mathbf{E}, \mathcal{B}^\circ)$-valued random variables on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let X_0 be an $(\mathbf{E}, \mathcal{B}^\circ)$-valued random variable on some probability space $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$ with $\mathbb{P}_0[X_0 \in \mathbf{E}_0] = 1$ for some separable $\mathbf{E}_0 \in \mathcal{B}^\circ$. Let $y_0 \in \mathbf{E}_0$. Let $(\widetilde{\mathbf{E}}, \widetilde{d})$ be a metric space equipped with the corresponding open-ball σ-algebra $\widetilde{\mathcal{B}}^\circ$. Then, $X_n \rightsquigarrow^\circ X_0$ and $Y_n \to^{\mathsf{p},\circ} y_0$ together imply:*

(i) $(X_n, Y_n) \rightsquigarrow^\circ (X_0, y_0)$.
(ii) $h(X_n, Y_n) \rightsquigarrow^\circ h(X_0, y_0)$ *for every continuous and* $(\overline{\mathcal{B}}^\circ, \widetilde{\mathcal{B}}^\circ)$*-measurable* $h : \overline{\mathbf{E}} \to \widetilde{\mathbf{E}}$.

Proof. Assertion (ii) is an immediate consequence of Assertion (i) and the Continuous Mapping theorem in the form of (Billingsley 1999, Theorem 6.4); take into account that (X_0, y_0) takes values only

in $\overline{\mathbf{E}}_0 := \mathbf{E}_0 \times \mathbf{E}_0$ and that $\mathbf{E}_0 \times \mathbf{E}_0$ is separable with respect to $\overline{d}$. Thus, it suffices to show Assertion (i). First note that we have

$$(X_n, y_0) \rightsquigarrow^\circ (X_0, y_0). \tag{A2}$$

Indeed, for every $f \in \overline{C}_b^\circ$ (with $\overline{C}_b^\circ$ the set of all bounded, continuous and $(\overline{\mathcal{B}}^\circ, \mathcal{B}(\mathbb{R}))$-measurable real-valued functions on $\overline{\mathbf{E}}$) we have $\lim_{n\to\infty} \int f(X_n, y_0)\, d\mathbb{P} = \int f(X_0, y_0)\, d\mathbb{P}_0$ by the assumption $X_n \rightsquigarrow^\circ X_0$ and the fact that the mapping $x \mapsto f(x, y_0)$ lies in C_b° (the latter was shown in the proof of Theorem 3.1 in Beutner and Zähle (2016)).

Second, the distance $\overline{d}((X_n, Y_n), (X_n, y_0)) = d(Y_n, y_0)$ is $(\mathcal{F}, \mathcal{B}(\mathbb{R}_+))$-measurable for every $n \in \mathbb{N}$, because Y_n is $(\mathcal{F}, \mathcal{B}^\circ)$-measurable and $x \mapsto d(x, y_0)$ is $(\mathcal{B}^\circ, \mathcal{B}(\mathbb{R}))$-measurable (due to Lemma B.3 in Beutner and Zähle (2016)). Along with $Y_n \to^{\mathsf{p},\circ} y_0$, we obtain in particular that $\overline{d}((X_n, Y_n), (X_n, y_0)) \to^{\mathsf{p}} 0$. Together with Equation (A2) and Theorem A1 (applied to $X_n' := (X_n, y_0)$, $X_0' := (X_0, y_0)$, $Y_n' := (X_n, Y_n)$), this implies $(X_n, Y_n) \rightsquigarrow^\circ (X_0, y_0)$; take into account again that (X_0, y_0) takes values only in $\overline{\mathbf{E}}_0 := \mathbf{E}_0 \times \mathbf{E}_0$ and that $\mathbf{E}_0 \times \mathbf{E}_0$ is separable with respect to $\overline{d}$. □

Corollary A2. *Let $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$ be a normed vector space and d be the induced metric defined by $d(x_1, x_2) := \|x_1 - x_2\|_{\mathbf{E}}$. Let (X_n) and (Y_n) be two sequences of $(\mathbf{E}, \mathcal{B}^\circ)$-valued random variables on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let X_0 be an $(\mathbf{E}, \mathcal{B}^\circ)$-valued random variable on some probability space $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$ with $\mathbb{P}_0[X_0 \in \mathbf{E}_0] = 1$ for some separable $\mathbf{E}_0 \in \mathcal{B}^\circ$. Let $y_0 \in \mathbf{E}_0$. Assume that the map $h : \overline{\mathbf{E}} \to \mathbf{E}$ defined by $h(x_1, x_2) := x_1 + x_2$ is $(\overline{\mathcal{B}}^\circ, \mathcal{B}^\circ)$-measurable. Then, $X_n \rightsquigarrow^\circ X_0$ and $Y_n \to^{\mathsf{p},\circ} y_0$ together imply $X_n + Y_n \rightsquigarrow^\circ X_0 + y_0$.*

Proof. The assertion is an immediate consequence of Corollary A1 and the fact that h is clearly continuous. □

Appendix A.2. Delta-Method and Chain Rule for Uniformly Quasi-Hadamard Differentiable Maps

Now, assume that $\mathbf{E}$ is a subspace of a vector space $\mathbf{V}$. Let $\|\cdot\|_{\mathbf{E}}$ be a norm on $\mathbf{E}$ and assume that the metric d is induced by $\|\cdot\|_{\mathbf{E}}$. Let $\widetilde{\mathbf{V}}$ be another vector space and $\widetilde{\mathbf{E}} \subseteq \widetilde{\mathbf{V}}$ be any subspace. Let $\|\cdot\|_{\widetilde{\mathbf{E}}}$ be a norm on $\widetilde{\mathbf{E}}$ and $\widetilde{\mathcal{B}}^\circ$ be the corresponding open-ball σ-algebra on $\widetilde{\mathbf{E}}$. Let $0_{\widetilde{\mathbf{E}}}$ denote the null in $\widetilde{\mathbf{E}}$. Moreover, let $\overline{\widetilde{\mathbf{E}}} := \widetilde{\mathbf{E}} \times \widetilde{\mathbf{E}}$ and $\overline{\widetilde{\mathcal{B}}^\circ}$ be the σ-algebra on $\overline{\widetilde{\mathbf{E}}}$ generated by the open balls with respect to the metric $\overline{\widetilde{d}}((\widetilde{x}_1, \widetilde{x}_2), (\widetilde{y}_1, \widetilde{y}_2)) := \max\{\|\widetilde{x}_1 - \widetilde{y}_1\|_{\widetilde{\mathbf{E}}}; \|\widetilde{x}_2 - \widetilde{y}_2\|_{\widetilde{\mathbf{E}}}\}$.

Let $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$ be a probability space and $\widehat{T}_n : \Omega_n \to \mathbf{V}$ be any map for every $n \in \mathbb{N}$. Recall that $\rightsquigarrow^\circ$ and $\to^{\mathsf{p},\circ}$ refer to convergence in distribution$^\circ$ and convergence in probability$^\circ$, respectively. Moreover, recall Definition A1 of quasi-Hadamard differentiability.

Theorem A2. *Let $H : \mathbf{V}_H \to \widetilde{\mathbf{E}}$ be a map defined on some $\mathbf{V}_H \subseteq \mathbf{V}$. Let $\mathbf{E}_0 \in \mathcal{B}^\circ$ be some $\|\cdot\|_{\mathbf{E}}$-separable subset of $\mathbf{E}$. Let $(\theta_n) \subseteq \mathbf{V}_H$ and define the singleton set $\mathcal{S} := \{(\theta_n)\}$. Let (a_n) be a sequence of positive real numbers tending to ∞, and consider the following conditions:*

(a) *$\widehat{T}_n$ takes values only in $\mathbf{V}_H$.*
(b) *$a_n(\widehat{T}_n - \theta_n)$ takes values only in $\mathbf{E}$, is $(\mathcal{F}_n, \mathcal{B}^\circ)$-measurable and satisfies*

$$a_n(\widehat{T}_n - \theta_n) \rightsquigarrow^\circ \xi \qquad \text{in } (\mathbf{E}, \mathcal{B}^\circ, \|\cdot\|_{\mathbf{E}}) \tag{A3}$$

for some $(\mathbf{E}, \mathcal{B}^\circ)$-valued random variable ξ on some probability space $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$ with $\xi(\Omega_0) \subseteq \mathbf{E}_0$.
(c) *$a_n(H(\widehat{T}_n) - H(\theta_n))$ takes values only in $\widetilde{\mathbf{E}}$ and is $(\mathcal{F}_n, \widetilde{\mathcal{B}}^\circ)$-measurable.*
(d) *The map H is uniformly quasi-Hadamard differentiable with respect to $\mathcal{S}$ tangentially to $\mathbf{E}_0\langle\mathbf{E}\rangle$ with trace $\widetilde{\mathbf{E}}$ and uniform quasi-Hadamard derivative $\dot{H}_{\mathcal{S}} : \mathbf{E}_0 \to \widetilde{\mathbf{E}}$.*
(e) *$(\Omega_n, \mathcal{F}_n, \mathbb{P}_n) = (\Omega, \mathcal{F}, \mathbb{P})$ for all $n \in \mathbb{N}$.*
(f) *The uniform quasi-Hadamard derivative $\dot{H}_{\mathcal{S}}$ can be extended to $\mathbf{E}$ such that the extension $\dot{H}_{\mathcal{S}} : \mathbf{E} \to \widetilde{\mathbf{E}}$ is continuous at every point of $\mathbf{E}_0$ and $(\mathcal{B}^\circ, \widetilde{\mathcal{B}}^\circ)$-measurable.*
(g) *The map $h : \overline{\widetilde{\mathbf{E}}} \to \widetilde{\mathbf{E}}$ defined by $h(\widetilde{x}_1, \widetilde{x}_2) := \widetilde{x}_1 - \widetilde{x}_2$ is $(\overline{\widetilde{\mathcal{B}}^\circ}, \widetilde{\mathcal{B}}^\circ)$-measurable.*

Then, the following two assertions hold:

(i) *If Conditions (a)–(d) hold true, then* $\dot{H}_{\mathcal{S}}(\xi)$ *is* $(\mathcal{F}_0, \widetilde{\mathcal{B}}^\circ)$*-measurable and*

$$a_n\big(H(\widehat{T}_n) - H(\theta_n)\big) \rightsquigarrow^\circ \dot{H}_{\mathcal{S}}(\xi) \qquad \text{in } (\widetilde{\mathbf{E}}, \widetilde{\mathcal{B}}^\circ, \|\cdot\|_{\widetilde{\mathbf{E}}}).$$

(ii) *If Conditions (a)–(g) hold true, then*

$$a_n\big(H(\widehat{T}_n) - H(\theta_n)\big) - \dot{H}_{\mathcal{S}}\big(a_n(\widehat{T}_n - \theta_n)\big) \to^{\mathsf{p},\circ} 0_{\widetilde{\mathbf{E}}} \qquad \text{in } (\widetilde{\mathbf{E}}, \|\cdot\|_{\widetilde{\mathbf{E}}}). \tag{A4}$$

Proof. The proof is very similar to the proof of Theorem C.4 in Beutner and Zähle (2016).

(i): For every $n \in \mathbb{N}$, let $\mathbf{E}_n := \{x_n \in \mathbf{E} : \theta_n + a_n^{-1} x_n \in \mathbf{V}_H\}$ and define the map $h_n : \mathbf{E}_n \to \widetilde{\mathbf{E}}$ by

$$h_n(x_n) := \frac{H(\theta_n + a_n^{-1} x_n) - H(\theta_n)}{a_n^{-1}}.$$

Moreover, define the map $h_0 : \mathbf{E}_0 \to \widetilde{\mathbf{E}}$ by

$$h_0(x) := \dot{H}_{\mathcal{S}}(x).$$

Now, the claim would follow by the extended Continuous Mapping theorem in the form of Theorem C.1 in Beutner and Zähle (2016) applied to the functions h_n, $n \in \mathbb{N}_0$, and the random variables $\xi_n := a_n(\widehat{T}_n - \theta_n)$, $n \in \mathbb{N}$, and $\xi_0 := \xi$ if we can show that the assumptions of Theorem C.1 in Beutner and Zähle (2016) are satisfied. First, by Assumption (a) and the last part of Assumption (b), we have $\xi_n(\Omega_n) \subseteq \mathbf{E}_n$ and $\xi_0(\Omega_0) \subseteq \mathbf{E}_0$. Second, by Assumption (c), we have that $h_n(\xi_n) = a_n(H(\widehat{T}_n) - H(\theta_n))$ is $(\mathcal{F}_n, \widetilde{\mathcal{B}}^\circ)$-measurable. Third, the map h_0 is continuous by the definition of the quasi-Hadamard derivative. Thus, h_0 is $(\mathcal{B}_0^\circ, \widetilde{\mathcal{B}}^\circ)$-measurable, because the trace σ-algebra $\mathcal{B}_0^\circ := \mathcal{B}^\circ \cap \mathbf{E}_0$ coincides with the Borel σ-algebra on $\mathbf{E}_0$ (recall that $\mathbf{E}_0$ is separable). In particular, $\dot{H}_{\mathcal{S}}(\xi)$ is $(\mathcal{F}_0, \widetilde{\mathcal{B}}^\circ)$-measurable. Fourth, Condition (a) of Theorem C.1 in Beutner and Zähle (2016) holds by Assumption (b). Fifth, Condition (b) of Theorem C.1 in Beutner and Zähle (2016) is ensured by Assumption (d).

(ii): For every $n \in \mathbb{N}$, let $\mathbf{E}_n$ and h_n be as above and define the map $\overline{h}_n : \mathbf{E}_n \to \overline{\widetilde{\mathbf{E}}}$ by

$$\overline{h}_n(x_n) := (h_n(x_n), \dot{H}_{\mathcal{S}}(x_n)).$$

Moreover, define the map $\overline{h}_0 : \mathbf{E}_0 \to \overline{\widetilde{\mathbf{E}}}$ by

$$\overline{h}_0(x) := (h_0(x), \dot{H}_{\mathcal{S}}(x)) = (\dot{H}_{\mathcal{S}}(x), \dot{H}_{\mathcal{S}}(x)).$$

We first show that

$$\overline{h}_n(a_n(T_n - \theta_n)) \rightsquigarrow^\circ \overline{h}_0(\xi) \qquad \text{in } (\overline{\widetilde{\mathbf{E}}}, \overline{\widetilde{\mathcal{B}}^\circ}, \overline{\widetilde{d}}). \tag{A5}$$

For Equation (A5), it suffices to show that the assumption of the extended Continuous Mapping theorem in the form of Theorem C.1 in Beutner and Zähle (2016) applied to the functions $\overline{h}_n$ and ξ_n (as defined above) are satisfied. The claim then follows by Theorem C.1 in Beutner and Zähle (2016). First, we have already observed that $\xi_n(\Omega_n) \subseteq \mathbf{E}_n$ and $\xi_0(\Omega_0) \subseteq \mathbf{E}_0$. Second, we have seen in the proof of Part (i) that $h_n(\xi_n)$ is $(\mathcal{F}_n, \widetilde{\mathcal{B}}^\circ)$-measurable, $n \in \mathbb{N}$. By Assumption (f), the extended map $\dot{H}_{\mathcal{S}} : \mathbf{E} \to \widetilde{\mathbf{E}}$ is $(\mathcal{B}^\circ, \widetilde{\mathcal{B}}^\circ)$-measurable, which implies that $\dot{H}_{\mathcal{S}}(\xi_n)$ is $(\mathcal{F}_n, \widetilde{\mathcal{B}}^\circ)$-measurable. Thus, $\overline{h}_n(\xi_n) = (h_n(\xi_n), \dot{H}_{\mathcal{S}}(\xi_n))$ is $(\mathcal{F}_n, \widetilde{\mathcal{B}}^\circ \otimes \widetilde{\mathcal{B}}^\circ)$-measurable (to see this note that, in view of $\widetilde{\mathcal{B}}^\circ \otimes \widetilde{\mathcal{B}}^\circ = \sigma(\pi_1, \pi_2)$ for the coordinate projections π_1, π_2 on $\overline{\widetilde{E}} = \widetilde{E} \times \widetilde{E}$, Theorem 7.4 of Bauer (2001) shows that the map $(h_n(\xi_n), \dot{H}_{\mathcal{S}}(\xi_n))$ is $(\mathcal{F}_n, \widetilde{\mathcal{B}}^\circ \otimes \widetilde{\mathcal{B}}^\circ)$-measurable if and only if the maps $h_n(\xi_n) = \pi_1 \circ (h_n(\xi_n), \dot{H}_{\mathcal{S}}(\xi_n))$ and $\dot{H}_{\mathcal{S}}(\xi_n) = \pi_2 \circ (h_n(\xi_n), \dot{H}_{\mathcal{S}}(\xi_n))$ are $(\mathcal{F}_n, \widetilde{\mathcal{B}}^\circ)$-measurable). In particular, the map $\overline{h}_n(\xi_n) = (h_n(\xi_n), \dot{H}_{\mathcal{S}}(\xi_n))$ is $(\mathcal{F}_n, \overline{\widetilde{\mathcal{B}}^\circ})$-measurable, $n \in \mathbb{N}$. Third, we have seen in the proof of Part (i) that the map $h_0 = \dot{H}_{\mathcal{S}}$ is $(\mathcal{B}_0^\circ, \widetilde{\mathcal{B}}^\circ)$-measurable. Thus, the map $\overline{h}_0$ is $(\mathcal{B}_0^\circ, \widetilde{\mathcal{B}}^\circ \otimes \widetilde{\mathcal{B}}^\circ)$-measurable (one can argue

as above) and in particular $(\mathcal{B}_0^\circ, \overline{\widetilde{\mathcal{B}}^\circ})$-measurable. Fourth, Condition (a) of Theorem C.1 in Beutner and Zähle (2016) holds by Assumption (b). Fifth, Condition (b) of Theorem C.1 in Beutner and Zähle (2016) is ensured by Assumption (d) and the continuity of the extended map $\dot{H}_{\mathcal{S}}$ at every point of $\mathbf{E}_0$ (recall Assumption (f)). Hence, Equation (A5) holds.

By Assumption (g) and the ordinary Continuous Mapping theorem (see (Billingsley 1999, Theorem 6.4)) applied to Equation (A5) and the map $h : \widetilde{\mathbf{E}} \to \widetilde{\mathbf{E}}$, $(\widetilde{x}_1, \widetilde{x}_2) \mapsto \widetilde{x}_1 - \widetilde{x}_2$, we now have

$$h_n(a_n(\widehat{T}_n - \theta_n)) - \dot{H}_{\mathcal{S}}(a_n(\widehat{T}_n - \theta_n)) \rightsquigarrow^\circ \dot{H}_{\mathcal{S}}(\xi) - \dot{H}_{\mathcal{S}}(\xi),$$

i.e.,

$$a_n\big(H(\widehat{T}_n) - H(\theta_n)\big) - \dot{H}_{\mathcal{S}}\big(a_n(\widehat{T}_n - \theta_n)\big) \rightsquigarrow^\circ 0_{\widetilde{\mathbf{E}}}.$$

By Proposition B.4 in Beutner and Zähle (2016), we can conclude Equation (A4). □

The following lemma provides a chain rule for uniformly quasi-Hadamard differentiable maps (a similar chain rule with different $\mathcal{S}$ was found in Varron (2015)). To formulate the chain rule, let $\widetilde{\widetilde{\mathbf{V}}}$ be a further vector space and $\widetilde{\widetilde{\mathbf{E}}} \subseteq \widetilde{\widetilde{\mathbf{V}}}$ be a subspace equipped with a norm $\|\cdot\|_{\widetilde{\widetilde{\mathbf{E}}}}$.

Lemma A1. *Let $H : \mathbf{V}_H \to \widetilde{\mathbf{V}}_{\widetilde{H}}$ and $\widetilde{H} : \widetilde{\mathbf{V}}_{\widetilde{H}} \to \widetilde{\widetilde{\mathbf{V}}}$ be maps defined on subsets $\mathbf{V}_H \subseteq \mathbf{V}$ and $\widetilde{\mathbf{V}}_{\widetilde{H}} \subseteq \widetilde{\mathbf{V}}$ such that $H(\mathbf{V}_H) \subseteq \widetilde{\mathbf{V}}_{\widetilde{H}}$. Let $\mathbf{E}_0$ and $\widetilde{\mathbf{E}}_0$ be subsets of $\mathbf{E}$ and $\widetilde{\mathbf{E}}$, respectively. Let $\mathcal{S}$ and $\widetilde{\mathcal{S}}$ be sets of sequences in $\mathbf{V}_H$ and $\widetilde{\mathbf{V}}_{\widetilde{H}}$, respectively, and assume that the following three assertions hold.*

(a) For every $(\theta_n) \in \mathcal{S}$, we have $(H(\theta_n)) \in \widetilde{\mathcal{S}}$.
(b) H is uniformly quasi-Hadamard differentiable with respect to $\mathcal{S}$ tangentially to $\mathbf{E}_0\langle\mathbf{E}\rangle$ with trace $\widetilde{\mathbf{E}}$ and uniform quasi-Hadamard derivative $\dot{H}_{\mathcal{S}} : \mathbf{E}_0 \to \widetilde{\mathbf{E}}$, and we have $\dot{H}_{\mathcal{S}}(\mathbf{E}_0) \subseteq \widetilde{\mathbf{E}}_0$.
(c) $\widetilde{H}$ is uniformly quasi-Hadamard differentiable with respect to $\widetilde{\mathcal{S}}$ tangentially to $\widetilde{\mathbf{E}}_0\langle\widetilde{\mathbf{E}}\rangle$ with trace $\widetilde{\widetilde{\mathbf{E}}}$ and uniform quasi-Hadamard derivative $\dot{\widetilde{H}}_{\widetilde{\mathcal{S}}} : \widetilde{\mathbf{E}}_0 \to \widetilde{\widetilde{\mathbf{E}}}$.

Then, the map $T := \widetilde{H} \circ H : \mathbf{V}_H \to \widetilde{\widetilde{\mathbf{V}}}$ is uniformly quasi-Hadamard differentiable with respect to $\mathcal{S}$ tangentially to $\mathbf{E}_0\langle\mathbf{E}\rangle$ with trace $\widetilde{\widetilde{\mathbf{E}}}$, and the uniform quasi-Hadamard derivative $\dot{T}_{\mathcal{S}}$ is given by $\dot{T}_{\mathcal{S}} := \dot{\widetilde{H}}_{\widetilde{\mathcal{S}}} \circ \dot{H}_{\mathcal{S}}$.

Proof. Obviously, since $H(\mathbf{V}_H) \subseteq \widetilde{\mathbf{V}}_{\widetilde{H}}$ and $\widetilde{H}$ is associated with trace $\widetilde{\widetilde{\mathbf{E}}}$, the map $\widetilde{H} \circ H$ can also be associated with trace $\widetilde{\widetilde{\mathbf{E}}}$.

Now, let $((\theta_n), x, (x_n), (\varepsilon_n))$ be a quadruple with $(\theta_n) \in \mathcal{S}$, $x \in \mathbf{E}_0$, $(x_n) \subseteq \mathbf{E}$ satisfying $\|x_n - x\|_{\mathbf{E}} \to 0$ as well as $(\theta_n + \varepsilon_n x_n) \subseteq \mathbf{V}_H$, and $(\varepsilon_n) \subseteq (0, \infty)$ satisfying $\varepsilon_n \to 0$. Then,

$$\begin{aligned}
&\left\| \dot{\widetilde{H}}_{\widetilde{\mathcal{S}}}(\dot{H}_{\mathcal{S}}(x)) - \frac{\widetilde{H}(H(\theta_n + \varepsilon_n x_n)) - \widetilde{H}(H(\theta_n))}{\varepsilon_n} \right\|_{\widetilde{\widetilde{\mathbf{E}}}} \\
&= \left\| \dot{\widetilde{H}}_{\widetilde{\mathcal{S}}}(\dot{H}_{\mathcal{S}}(x)) - \frac{\widetilde{H}\big(H(\theta_n) + \varepsilon_n \frac{H(\theta_n + \varepsilon_n x_n) - H(\theta_n)}{\varepsilon_n}\big) - \widetilde{H}(H(\theta_n))}{\varepsilon_n} \right\|_{\widetilde{\widetilde{\mathbf{E}}}}.
\end{aligned}$$

Note that by assumption, $H(\theta_n) \in \widetilde{\mathbf{V}}_{\widetilde{H}}$ and in particular $(H(\theta_n)) \in \widetilde{\mathcal{S}}$. By the uniform quasi-Hadamard differentiability of H with respect to $\mathcal{S}$ tangentially to $\mathbf{E}_0\langle\mathbf{E}\rangle$ with trace $\widetilde{\mathbf{E}}$,

$$\lim_{n\to\infty} \left\| \frac{H(\theta_n + \varepsilon_n x_n) - H(\theta_n)}{\varepsilon_n} - \dot{H}_{\mathcal{S}}(x) \right\|_{\widetilde{E}} = 0.$$

Moreover, $(H(\theta_n + \varepsilon_n x_n) - H(\theta_n))/\varepsilon_n \in \widetilde{\mathbf{E}}$ and $\dot{H}_{\mathcal{S}}(x) \in \widetilde{\mathbf{E}}_0$, because H is associated with trace $\widetilde{\mathbf{E}}$ and $\dot{H}_{\mathcal{S}}(\mathbf{E}_0) \subseteq \widetilde{\mathbf{E}}_0$. Hence, by the uniform quasi-Hadamard differentiability of $\widetilde{H}$ with respect to $\widetilde{\mathcal{S}}$ tangentially to $\widetilde{\mathbf{E}}_0\langle\widetilde{\mathbf{E}}\rangle$, we obtain

$$\lim_{n\to\infty} \left\| \tilde{\dot{H}}_{\widetilde{\mathcal{S}}}(\dot{H}_{\mathcal{S}}(x)) - \frac{\widetilde{H}\big(H(\theta_n) + \varepsilon_n \frac{H(\theta_n+\varepsilon_n x_n)-H(\theta_n)}{\varepsilon_n}\big) - \widetilde{H}(H(\theta_n))}{\varepsilon_n} \right\|_{\widetilde{\widetilde{\mathbf{E}}}} = 0.$$

This completes the proof. □

Appendix B. Delta-Method for the Bootstrap

The functional delta-method is a widely used technique to derive bootstrap consistency for a sequence of plug-in estimators with respect to a map H from bootstrap consistency of the underlying sequence of estimators. An essential limitation of the classical functional delta-method for proving bootstrap consistency in probability (or outer probability) is the condition of Hadamard differentiability on H (see Theorem 3.9.11 of van der Vaart Wellner (1996)). It is commonly acknowledged that Hadamard differentiability fails for many relevant maps H. Recently, it was demonstrated in Beutner and Zähle (2016) that a functional delta-method for the bootstrap *in probability* can also be proved for *quasi*-Hadamard differentiable maps H. Quasi-Hadamard differentiability is a weaker notion of "differentiability" than Hadamard differentiability and can be obtained for many relevant statistical functionals H (see, e.g., Beutner et al. 2012; Beutner and Zähle 2010, 2012; Krätschmer et al. 2013; Krätschmer and Zähle 2017). Using the classical functional delta-method to prove almost sure (or outer almost sure) bootstrap consistency for a sequence of plug-in estimators with respect to a map H from almost sure (or outer almost sure) bootstrap consistency of the underlying sequence of estimators requires *uniform* Hadamard differentiability on H (see Theorem 3.9.11 of van der Vaart Wellner (1996)). In this section, we introduce the notion of *uniform quasi*-Hadamard differentiability and demonstrate that one can even obtain a functional delta-method for the *almost sure* bootstrap and *uniformly quasi*-Hadamard differentiable maps H.

To explain the background and the contribution of this section more precisely, assume that we are given an estimator $\widehat{T}_n$ for a parameter θ in a vector space, with n denoting the sample size, and that we are actually interested in the aspect $H(\theta)$ of θ. Here, H is any map taking values in a vector space. Then, $H(\widehat{T}_n)$ is often a reasonable estimator for $H(\theta)$. One of the main objects in statistical inference is the distribution of the error $H(\widehat{T}_n) - H(\theta)$, because the error distribution can theoretically be used to derive confidence regions for $H(\theta)$. However, in applications, the exact specification of the error distribution is often hardly possible or even impossible. A widely used way out is to derive the *asymptotic* error distribution, i.e., the weak limit μ of $\mathrm{law}\{a_n(H(\widehat{T}_n) - H(\theta))\}$ for suitable normalizing constants a_n tending to infinity, and to use μ as an approximation for $\mu_n := \mathrm{law}\{a_n(H(\widehat{T}_n) - H(\theta))\}$ for large n. Since μ usually still depends on the unknown parameter θ, one should use the notation μ_θ instead of μ. In particular, one actually uses $\mu_{\widehat{T}_n} := \mu_\theta|_{\theta=\widehat{T}_n}$ as an approximation for μ_n for large n.

Not least because of the estimation of the parameter θ of μ_θ, the approximation of μ_n by $\mu_{\widehat{T}_n}$ is typically only moderate. An often more efficient alternative technique to approximate μ_n is the bootstrap. The bootstrap has been introduced by Efron (1979) and many variants of his method have been introduced since then. One may refer to Davison and Hinkley (1997); Efron (1994); Lahiri (2003); Shao and Tu (1995) for general accounts on this topic. The basic idea of the bootstrap is the following. Re-sampling the original sample according to a certain re-sampling mechanism (depending on the particular bootstrap method) one can sometimes construct a so-called bootstrap version $\widehat{T}_n^*$ of $\widehat{T}_n$ for which the conditional law of $a_n(H(\widehat{T}_n^*) - H(\widehat{T}_n))$ "given the sample" has the same weak limit μ_θ as the law of $a_n(H(\widehat{T}_n) - H(\theta))$ has. The latter is referred to as bootstrap consistency. Since $\widehat{T}_n^*$ depends only on the sample and the re-sampling mechanism, one can at least numerically determine the conditional law of $a_n(H(\widehat{T}_n^*) - H(\widehat{T}_n))$ "given the sample" by means of a Monte Carlo simulation based on $L \gg n$ repetitions. The resulting law μ_L^* can then be used as an approximation of μ_n, at least for large n.

In applications, the roles of θ and $\widehat{T}_n$ are often played by a distribution function F and the empirical distribution function $\widehat{F}_n$ of n random variables that are identically distributed according to F, respectively. Not least for this particular setting several results on bootstrap consistency for $\widehat{T}_n$ are known (see also Appendix B.2). The functional delta-method then ensures that bootstrap consistency also holds for $H(\widehat{T}_n)$ when H is suitably differentiable at θ. Technically speaking, as indicated above, one has to distinguish between two types of bootstrap consistency. First bootstrap consistency *in probability* for $H(\widehat{T}_n)$ can be associated with

$$\lim_{n\to\infty} \mathbb{P}^{\mathsf{out}}\big[\{\omega \in \Omega : d^{\circ}_{\mathrm{BL}}(P_n(\omega,\cdot), \mu_\theta) \geq \delta\}\big] = 0 \quad \text{for all } \delta > 0, \tag{A6}$$

where ω represents the sample, $P_n(\omega,\cdot)$ denotes the conditional law of $a_n(H(\widehat{T}^*_n) - H(\widehat{T}_n))$ given the sample ω, d°_{BL} is the bounded Lipschitz distance, and the superscript ^{out} refers to outer probability. At this point, it is worth pointing out that we consider weak convergence (respectively, convergence in distribution) with respect to the open-ball σ-algebra, in symbols $\Rightarrow^{\circ}$ (respectively, $\rightsquigarrow^{\circ}$), as defined in (Billingsley 1999, sct. 6) (see also Dudley 1966, 1967; Pollard 1984; Shorack and Wellner 1986) and that by the Portmanteau theorem A.3 in Beutner and Zähle (2016) weak convergence $\mu_n \Rightarrow^{\circ} \mu$ holds if and only if $d^{\circ}_{\mathrm{BL}}(\mu_n, \mu) \to 0$. Second bootstrap consistency *almost surely* for $H(\widehat{T}_n)$ means that

$$\mathrm{law}\big\{a_n\big(H(\widehat{T}^*_n(\omega,\cdot)) - H(\widehat{T}_n(\omega))\big)\big\} \Rightarrow^{\circ} \mu_\theta \qquad \mathbb{P}\text{-a.e. } \omega. \tag{A7}$$

In Beutner and Zähle (2016), it has been shown that Equation (A6) follows from the respective analogue for $\widehat{T}_n$ when H is suitably quasi-Hadamard differentiable at θ. This extends Theorem 3.9.11 of van der Vaart Wellner (1996) which covers only Hadamard differentiable maps. In this section, we show that Equation (A7) follows from the respective analogue for $\widehat{T}_n$ when H is suitably *uniformly* quasi-Hadamard differentiable at θ; the notion of uniform quasi-Hadamard differentiable is introduced in Definition A1 below. This extends Theorem 3.9.13 of van der Vaart Wellner (1996) which covers only Hadamard differentiable maps.

Appendix B.1. Abstract Delta-Method for the Bootstrap

Theorem A4 provides an abstract delta-method for the almost sure bootstrap. It is based on the notion of uniform quasi-Hadamard differentiability which we introduce first. This sort of differentiability extends the notion of quasi-Hadamard differentiability as introduced in Beutner and Zähle (2010, 2016). The latter corresponds to the differentiability concept in (i) of Definition A1 ahead with $\mathcal{S}$ and $\widetilde{\mathbf{E}}$ as in (iii) and (v) of this definition. Let $\mathbf{V}$ and $\widetilde{\mathbf{V}}$ be vector spaces. Let $\mathbf{E} \subseteq \mathbf{V}$ and $\widetilde{\mathbf{E}} \subseteq \widetilde{\mathbf{V}}$ be subspaces equipped with norms $\|\cdot\|_{\mathbf{E}}$ and $\|\cdot\|_{\widetilde{\mathbf{E}}}$, respectively. Let

$$H : \mathbf{V}_H \longrightarrow \widetilde{\mathbf{V}}$$

be any map defined on some subset $\mathbf{V}_H \subseteq \mathbf{V}$.

Definition A1. *Let $\mathbf{E}_0$ be a subset of $\mathbf{E}$, and $\mathcal{S}$ be a set of sequences in $\mathbf{V}_H$.*

(i) The map H is said to be uniformly quasi-Hadamard differentiable with respect to $\mathcal{S}$ tangentially to $\mathbf{E}_0\langle\mathbf{E}\rangle$ with trace $\widetilde{\mathbf{E}}$ if $H(y_1) - H(y_2) \in \widetilde{\mathbf{E}}$ for all $y_1, y_2 \in \mathbf{V}_H$ and there is some continuous map $\dot{H}_{\mathcal{S}} : \mathbf{E}_0 \to \widetilde{\mathbf{E}}$ such that

$$\lim_{n\to\infty} \left\| \dot{H}_{\mathcal{S}}(x) - \frac{H(\theta_n + \varepsilon_n x_n) - H(\theta_n)}{\varepsilon_n} \right\|_{\widetilde{\mathbf{E}}} = 0 \tag{A8}$$

holds for each quadruple $((\theta_n), x, (x_n), (\varepsilon_n))$, with $(\theta_n) \in \mathcal{S}$, $x \in \mathbf{E}_0$, $(x_n) \subseteq \mathbf{E}$ satisfying $\|x_n - x\|_{\mathbf{E}} \to 0$ as well as $(\theta_n + \varepsilon_n x_n) \subseteq \mathbf{V}_H$, and $(\varepsilon_n) \subseteq (0,\infty)$ satisfying $\varepsilon_n \to 0$. In this case, the map $\dot{H}_{\mathcal{S}}$ is called uniform quasi-Hadamard derivative of H with respect to $\mathcal{S}$ tangentially to $\mathbf{E}_0\langle\mathbf{E}\rangle$.

(ii) If $\mathcal{S}$ consists of all sequences $(\theta_n) \subseteq \mathbf{V}_H$ with $\theta_n - \theta \in \mathbf{E}$, $n \in \mathbb{N}$, and $\|\theta_n - \theta\|_{\mathbf{E}} \to 0$ for some fixed $\theta \in \mathbf{V}_H$, then we replace the phrase " with respect to $\mathcal{S}$" by "at θ" and "$\dot{H}_{\mathcal{S}}$" by "$\dot{H}_\theta$".

(iii) If $\mathcal{S}$ consists only of the constant sequence $\theta_n = \theta$, $n \in \mathbb{N}$, then we skip the phrase "uniformly" and replace the phrase " with respect to $\mathcal{S}$" by "at θ" and "$\dot{H}_{\mathcal{S}}$" by "$\dot{H}_\theta$". In this case, we may also replace "$H(y_1) - H(y_2) \in \widetilde{\mathbf{E}}$ for all $y_1, y_2 \in \mathbf{V}_H$" by "$H(y) - H(\theta) \in \widetilde{\mathbf{E}}$ for all $y \in \mathbf{V}_H$".

(iv) If $\mathbf{E} = \mathbf{V}$, then we skip the phrase "quasi-".

(v) If $\widetilde{\mathbf{E}} = \widetilde{\mathbf{V}}$, then we skip the phrase "with trace $\widetilde{\mathbf{E}}$".

The conventional notion of uniform Hadamard differentiability as used in Theorem 3.9.11 of van der Vaart Wellner (1996) corresponds to the differentiability concept in (i) with $\mathcal{S}$ as in (ii), $\mathbf{E}$ as in (iv), and $\widetilde{\mathbf{E}}$ as in (v). Proposition 1 shows that it is beneficial to refrain from insisting on $\mathbf{E} = \mathbf{V}$ as in (iv). It was recently discussed in Belloni et al. (2017) that it can be also beneficial to refrain from insisting on the assumption of (ii). For $\mathbf{E} = \mathbf{V}$ ("non-quasi" case), uniform Hadamard differentiability in the sense of Definition B.1 in Belloni et al. (2017) corresponds to uniform Hadamard differentiability in the sense of our Definition A1 (Parts (i) and (iv)) when $\mathcal{S}$ is chosen as the set of all sequences (θ_n) in a compact metric space $(\mathbf{K}_\theta, d_{\mathbf{K}})$ with $\theta \in \mathbf{K}_\theta \subseteq \mathbf{V}_H$ for which $d_{\mathbf{K}}(\theta_n, \theta) \to 0$. In Comment B.3 of Belloni et al. (2017), it is illustrated by means of the quantile functional that this notion of differentiability (subject to a suitable choice of $(\mathbf{K}_\theta, d_{\mathbf{K}})$) is strictly weaker than the notion of uniform Hadamard differentiability that was used in the classical delta-method for the almost sure bootstrap, Theorem 3.9.11 in van der Vaart Wellner (1996). Although this shows that the flexibility with respect to $\mathcal{S}$ in our Definition A1 can be beneficial, it is somehow even more important that we allow for the "quasi" case.

Of course, the smaller the family $\mathcal{S}$ the weaker the condition of uniform quasi-Hadamard differentiability with respect to $\mathcal{S}$. On the other hand, if the set $\mathcal{S}$ is too small, then Condition (e) in Theorem A4 ahead may fail. That is, for an application of the functional delta-method in the form of Theorem A4 the set $\mathcal{S}$ should be large enough for Condition (e) to be fulfilled and small enough for being able to establish uniform quasi-Hadamard differentiability with respect to $\mathcal{S}$ of the map H.

We now turn to the abstract delta-method. As mentioned in Section 1, convergence in distribution will always be considered for the open-ball σ-algebra. We use the terminology *convergence in distribution*$^\circ$ (symbolically $\rightsquigarrow^\circ$) for this sort of convergence; for details see Appendix A and Appendices A–C of Beutner and Zähle (2016). In a separable metric space the notion of convergence in distribution$^\circ$ boils down to the conventional notion of convergence in distribution for the Borel σ-algebra. In this case, we use the symbol $\rightsquigarrow$ instead of $\rightsquigarrow^\circ$.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and $(\widehat{T}_n)$ be a sequence of maps

$$\widehat{T}_n : \Omega \longrightarrow \mathbf{V}.$$

Regard $\omega \in \Omega$ as a sample drawn from $\mathbb{P}$, and $\widehat{T}_n(\omega)$ as a statistic derived from ω. Somewhat unconventionally, we do not (need to) require at this point that $\widehat{T}_n$ is measurable with respect to any σ-algebra on $\mathbf{V}$. Let $(\Omega', \mathcal{F}', \mathbb{P}')$ be another probability space and set

$$(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}}) := (\Omega \times \Omega', \mathcal{F} \otimes \mathcal{F}', \mathbb{P} \otimes \mathbb{P}').$$

The probability measure $\mathbb{P}'$ represents a random experiment that is run independently of the random sample mechanism $\mathbb{P}$. In the sequel, $\widehat{T}_n$ will frequently be regarded as a map defined on the extension $\overline{\Omega}$ of Ω. Let

$$\widehat{T}_n^* : \overline{\Omega} \longrightarrow \mathbf{V}$$

be any map. Since $\widehat{T}_n^*(\omega, \omega')$ depends on both the original sample ω and the outcome ω' of the additional independent random experiment, we may regard $\widehat{T}_n^*$ as a bootstrapped version of $\widehat{T}_n$. Moreover, let

$$\widehat{C}_n : \Omega \longrightarrow \mathbf{V}$$

be any map. As with $\widehat{T}_n$, we often regard $\widehat{C}_n$ as a map defined on the extension $\overline{\Omega}$ of Ω. We use $\widehat{C}_n$ together with a scaling sequence to get weak convergence results for $\widehat{T}_n^*$. The role of $\widehat{C}_n$ is often played by $\widehat{T}_n$ itself (see Example A1), but sometimes also by a different map (see Example A2). Assume that $\widehat{T}_n$, $\widehat{T}_n^*$, and $\widehat{C}_n$ take values only in $\mathbf{V}_H$.

Let $\mathcal{B}^\circ$ and $\widetilde{\mathcal{B}}^\circ$ be the open-ball σ-algebras on $\mathbf{E}$ and $\widetilde{\mathbf{E}}$ with respect to the norms $\|\cdot\|_{\mathbf{E}}$ and $\|\cdot\|_{\widetilde{\mathbf{E}}}$, respectively. Note that $\mathcal{B}^\circ$ coincides with the Borel σ-algebra on $\mathbf{E}$ when $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$ is separable. The same is true for $\widetilde{\mathcal{B}}^\circ$. Set $\overline{\widetilde{\mathbf{E}}} := \widetilde{\mathbf{E}} \times \widetilde{\mathbf{E}}$ and let $\overline{\widetilde{\mathcal{B}}^\circ}$ be the σ-algebra on $\overline{\widetilde{\mathbf{E}}}$ generated by the open balls with respect to the metric $\overline{\widetilde{d}}((\widetilde{x}_1, \widetilde{x}_2), (\widetilde{y}_1, \widetilde{y}_2)) := \max\{\|\widetilde{x}_1 - \widetilde{y}_1\|_{\widetilde{\mathbf{E}}}; \|\widetilde{x}_2 - \widetilde{y}_2\|_{\widetilde{\mathbf{E}}}\}$. Recall that $\overline{\widetilde{\mathcal{B}}^\circ} \subseteq \widetilde{\mathcal{B}}^\circ \otimes \widetilde{\mathcal{B}}^\circ$, because any $\overline{\widetilde{d}}$-open ball in $\overline{\widetilde{\mathbf{E}}}$ is the product of two $\|\cdot\|_{\widetilde{\mathbf{E}}}$-open balls in $\widetilde{\mathbf{E}}$.

Theorem A3 is a consequence of Theorem A2 in Appendix A.2 as we assume that $\widehat{T}_n$ takes values only in $\mathbf{V}_H$. The proof of the measurability statement of Theorem A3 is given in the proof of Theorem A4. Theorem A3 is stated here because, together with Theorem A4, it implies almost sure bootstrap consistency whenever the limit ξ is the same in Theorem A3 and Theorem A4.

Theorem A3. *Let (θ_n) be a sequence in $\mathbf{V}_H$ and $\mathcal{S} := \{(\theta_n)\}$. Let $\mathbf{E}_0 \subseteq \mathbf{E}$ be a separable subspace and assume that $\mathbf{E}_0 \in \mathcal{B}^\circ$. Let (a_n) be a sequence of positive real numbers with $a_n \to \infty$, and assume that the following assertions hold:*

(a) *$a_n(\widehat{T}_n - \theta_n)$ takes values only in $\mathbf{E}$, is $(\mathcal{F}, \mathcal{B}^\circ)$-measurable, and satisfies*

$$a_n(\widehat{T}_n - \theta_n) \rightsquigarrow^\circ \xi \qquad \text{in } (\mathbf{E}, \mathcal{B}^\circ, \|\cdot\|_{\mathbf{E}}) \tag{A9}$$

for some $(\mathbf{E}, \mathcal{B}^\circ)$-valued random variable ξ on some probability space $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$ with $\xi(\Omega_0) \subseteq \mathbf{E}_0$.

(b) *$a_n(H(\widehat{T}_n) - H(\theta_n))$ takes values only in $\widetilde{\mathbf{E}}$ and is $(\mathcal{F}, \widetilde{\mathcal{B}}^\circ)$-measurable.*

(c) *H is uniformly quasi-Hadamard differentiable with respect to $\mathcal{S}$ tangentially to $\mathbf{E}_0\langle\mathbf{E}\rangle$ with trace $\widetilde{\mathbf{E}}$ and uniform quasi-Hadamard derivative $\dot{H}_{\mathcal{S}}$.*

Then, $\dot{H}_{\mathcal{S}}(\xi)$ is $(\mathcal{F}_0, \widetilde{\mathcal{B}}^\circ)$-measurable and

$$a_n\big(H(\widehat{T}_n) - H(\theta_n)\big) \rightsquigarrow^\circ \dot{H}_{\mathcal{S}}(\xi) \qquad \text{in } (\widetilde{\mathbf{E}}, \widetilde{\mathcal{B}}^\circ, \|\cdot\|_{\widetilde{\mathbf{E}}}).$$

Theorem A4. *Let $\mathcal{S}$ be any set of sequences in $\mathbf{V}_H$. Let $\mathbf{E}_0 \subseteq \mathbf{E}$ be a separable subspace and assume that $\mathbf{E}_0 \in \mathcal{B}^\circ$. Let (a_n) be a sequence of positive real numbers with $a_n \to \infty$, and assume that the following assertions hold:*

(a) *$a_n(\widehat{T}_n^* - \widehat{C}_n)$ takes values only in $\mathbf{E}$, is $(\overline{\mathcal{F}}, \mathcal{B}^\circ)$-measurable, and satisfies*

$$a_n\big(\widehat{T}_n^*(\omega, \cdot\,) - \widehat{C}_n(\omega)\big) \rightsquigarrow^\circ \xi \qquad \text{in } (\mathbf{E}, \mathcal{B}^\circ, \|\cdot\|_{\mathbf{E}}), \qquad \mathbb{P}\text{-a.e. } \omega \tag{A10}$$

for some $(\mathbf{E}, \mathcal{B}^\circ)$-valued random variable ξ on some probability space $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$ with $\xi(\Omega_0) \subseteq \mathbf{E}_0$.

(b) *$a_n(H(\widehat{T}_n^*) - H(\widehat{C}_n))$ takes values only in $\widetilde{\mathbf{E}}$ and is $(\overline{\mathcal{F}}, \widetilde{\mathcal{B}}^\circ)$-measurable.*

(c) *H is uniformly quasi-Hadamard differentiable with respect to $\mathcal{S}$ tangentially to $\mathbf{E}_0\langle\mathbf{E}\rangle$ with trace $\widetilde{\mathbf{E}}$ and uniform quasi-Hadamard derivative $\dot{H}_{\mathcal{S}}$.*

(d) *The uniform quasi-Hadamard derivative $\dot{H}_{\mathcal{S}}$ can be extended from $\mathbf{E}_0$ to $\mathbf{E}$ such that the extension $\dot{H}_{\mathcal{S}} : \mathbf{E} \to \widetilde{\mathbf{E}}$ is $(\mathcal{B}^\circ, \widetilde{\mathcal{B}}^\circ)$-measurable and continuous at every point of $\mathbf{E}_0$.*

(e) *$(\widehat{C}_n(\omega)) \in \mathcal{S}$ for $\mathbb{P}$-a.e. ω.*

(f) *The map $h : \overline{\widetilde{\mathbf{E}}} \to \widetilde{\mathbf{E}}$ defined by $h(\widetilde{x}_1, \widetilde{x}_2) := \widetilde{x}_1 - \widetilde{x}_2$ is $(\overline{\widetilde{\mathcal{B}}^\circ}, \widetilde{\mathcal{B}}^\circ)$-measurable.*

Then, $\dot{H}_{\mathcal{S}}(\xi)$ is $(\mathcal{F}_0, \widetilde{\mathcal{B}}^\circ)$-measurable and

$$a_n\big(H(\widehat{T}_n^*(\omega,\,\cdot\,)) - H(\widehat{C}_n(\omega))\big) \rightsquigarrow^\circ \dot{H}_{\mathcal{S}}(\xi) \qquad \text{in } (\widetilde{\mathbf{E}}, \widetilde{\mathcal{B}}^\circ, \|\cdot\|_{\widetilde{\mathbf{E}}}), \qquad \mathbb{P}\text{-a.e. } \omega. \tag{A11}$$

Remark A1. *In Condition (a) of Theorem A4, it is assumed that $a_n(\widehat{T}_n^* - \widehat{C}_n)$ is $(\overline{\mathcal{F}}, \mathcal{B}^\circ)$-measurable for $\overline{\mathcal{F}} := \mathcal{F} \otimes \mathcal{F}'$. Thus, the mapping $\omega' \mapsto a_n(\widehat{T}_n^*(\omega,\omega') - \widehat{C}_n(\omega))$ is $(\mathcal{F}', \mathcal{B}^\circ)$-measurable for every fixed $\omega \in \Omega$. That is, $a_n(\widehat{T}_n^*(\omega,\cdot) - \widehat{C}_n(\omega))$ can be seen as an $(\mathbf{E}, \mathcal{B}^\circ)$-valued random variable on $(\Omega', \mathcal{F}', \mathbb{P}')$ for every fixed $\omega \in \Omega$, so that assertion (A10) makes sense. By the same line of reasoning one can regard $a_n(H(\widehat{T}_n^*(\omega,\cdot)) - H(\widehat{C}_n(\omega)))$ as an $(\widetilde{\mathbf{E}}, \widetilde{\mathcal{B}}^\circ)$-valued random variable on $(\Omega', \mathcal{F}', \mathbb{P}')$ for every fixed $\omega \in \Omega$, so that also assertion (A11) makes sense.*

Remark A2. *Condition (c) in Theorem A3 (respectively, Theorem A4) assumes that the trace is given by $\widetilde{\mathbf{E}}$, which implies that the first part of Condition (b) in Theorem A3 (respectively, Theorem A4) is automatically satisfied.*

Remark A3. *Condition (f) of Theorem A4 is automatically fulfilled when $(\widetilde{\mathbf{E}}, \|\cdot\|_{\widetilde{\mathbf{E}}})$ is separable. Indeed, in this case we have $\overline{\widetilde{\mathcal{B}}^\circ} = \widetilde{\mathcal{B}}^\circ \otimes \widetilde{\mathcal{B}}^\circ$ so that every continuous map $h : \overline{\widetilde{\mathbf{E}}} \to \widetilde{\mathbf{E}}$ (such as $h(\widetilde{x}_1, \widetilde{x}_2) := \widetilde{x}_1 - \widetilde{x}_2$) is $(\overline{\widetilde{\mathcal{B}}^\circ}, \widetilde{\mathcal{B}}^\circ)$-measurable.*

Proof. Proof of Theorem A4. First note that by the assumption imposed on ξ (see Assumption (a)) and Assumption (c) the map $\dot{H}_{\mathcal{S}}(\xi)$ is $(\mathcal{F}_0, \widetilde{\mathcal{B}}^\circ)$-measurable. Next, note that

$$\begin{aligned} & a_n\big(H(\widehat{T}_n^*(\omega,\omega')) - H(\widehat{C}_n(\omega))\big) \\ = \;& \big\{a_n\big(H(\widehat{T}_n^*(\omega,\omega')) - H(\widehat{C}_n(\omega))\big) - \dot{H}_{\mathcal{S}}\big(a_n(\widehat{T}_n^*(\omega,\omega') - \widehat{C}_n(\omega))\big)\big\} \\ & + \dot{H}_{\mathcal{S}}\big(a_n(\widehat{T}_n^*(\omega,\omega') - \widehat{C}_n(\omega))\big) \\ =: \;& S_1(n,\omega,\omega') + S_2(n,\omega,\omega'). \end{aligned}$$

By Equation (A10) in Assumption (a) and the Continuous Mapping theorem in the form of (Billingsley 1999, Theorem 6.4) (along with $\mathbb{P}_0 \circ \xi^{-1}[\mathbf{E}_0] = 1$ and the continuity of $\dot{H}_{\mathcal{S}}$), we have that $S_2(n,\omega,\cdot) \rightsquigarrow^\circ \dot{H}_{\mathcal{S}}(\xi)$ for $\mathbb{P}$-a.e. ω. Moreover, for every fixed ω we have that $\omega' \mapsto S_1(n,\omega,\omega')$ is $(\mathcal{F}', \widetilde{\mathcal{B}}^\circ)$-measurable by Assumption (f), and for $\mathbb{P}$-a.e. ω we have

$$a_n\big(H_n(\widehat{T}_n^*(\omega,\,\cdot\,)) - H_n(\widehat{C}_n(\omega))\big) - \dot{H}_{\mathcal{S}}\big(a_n(\widehat{T}_n^*(\omega,\,\cdot\,) - \widehat{C}_n(\omega))\big) \to^{\mathsf{P},\circ} 0_{\widetilde{\mathbf{E}}}$$

by Part (ii) of Theorem A2 (recall that $\widehat{T}_n^*$ was assumed to take values only in $\mathbf{V}_H$), where $\to^{\mathsf{P},\circ}$ refers to convergence in probability$^\circ$ (see Appendix A.1) and $\widehat{T}_n^*(\omega,\cdot)$, $\widehat{C}_n(\omega)$, $\{(\widehat{C}_n(\omega))\}$ play the roles of $\widehat{T}_n(\cdot)$, θ_n, $\mathcal{S}$, respectively. Hence, from Corollary A2, we get that Equation (A11) holds. □

Appendix B.2. Application to Plug-In Estimators of Statistical Functionals

Let $\mathbf{D}$, $\mathbf{D}_\phi$, $\mathcal{B}_\phi^\circ$ be as introduced at the beginning of Section 3. Let $\mathbf{C}_\phi \subseteq \mathbf{D}_\phi$ be a $\|\cdot\|_\phi$-separable subspace and assume $\mathbf{C}_\phi \in \mathcal{B}_\phi^\circ$. Moreover, let $H : \mathbf{D}(H) \to \widetilde{\mathbf{V}}$ be a map defined on a set $\mathbf{D}(H)$ of distribution functions of finite (not necessarily probability) Borel measures on $\mathbb{R}$, where $\widetilde{\mathbf{V}}$ is any vector space. In particular, $\mathbf{D}(H) \subseteq \mathbf{D}$. In the following, $\mathbf{D}$, $(\mathbf{D}_\phi, \mathcal{B}_\phi^\circ, \|\cdot\|_\phi)$, $\mathbf{C}_\phi$, and $\mathbf{D}(H)$ play the roles of $\mathbf{V}$, $(\mathbf{E}, \mathcal{B}^\circ, \|\cdot\|_{\mathbf{E}})$, $\mathbf{E}_0$, and $\mathbf{V}_H$, respectively. As before, we let $(\widetilde{\mathbf{E}}, \|\cdot\|_{\widetilde{\mathbf{E}}})$ be a normed subspace of $\widetilde{\mathbf{V}}$ equipped with the corresponding open-ball σ-algebra $\widetilde{\mathcal{B}}^\circ$.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $(F_n) \subseteq \mathbf{D}(H)$ be any sequence and (X_i) be a sequence of real-valued random variables on $(\Omega, \mathcal{F}, \mathbb{P})$. Moreover, let $\widehat{F}_n : \Omega \to \mathbf{D}$ be the empirical distribution function of $X_1, \ldots, X_n$, which will play the role of $\widehat{T}_n$. It is defined by

$$\widehat{F}_n := \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{[X_i,\infty)}. \tag{A12}$$

Assume that $\widehat{F}_n$ takes values only in $\mathbf{D}(H)$. Let $(\Omega', \mathcal{F}', \mathbb{P}')$ be another probability space and set $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}}) := (\Omega \times \Omega', \mathcal{F} \otimes \mathcal{F}', \mathbb{P} \otimes \mathbb{P}')$. Moreover, let $\widehat{F}_n^* : \overline{\Omega} \to \mathbf{D}$ be any map. Assume that $\widehat{F}_n^*$ take values only in $\mathbf{D}(H)$. Furthermore, let $\widehat{C}_n : \overline{\Omega} \to \mathbf{D}$ be any map that takes values only in $\mathbf{D}(H)$. In the present setting Theorems A3 and A4 can be reformulated as follows, where we recall from Remark A3 that Condition (f) of Theorem A4 is automatically fulfilled when $(\widetilde{\mathbf{E}}, \|\cdot\|_{\widetilde{\mathbf{E}}})$ is separable.

Corollary A3. *Let (F_n) be a sequence in $\mathbf{D}(H)$ and $\mathcal{S} := \{(F_n)\}$. Let (a_n) be a sequence of positive real numbers with $a_n \to \infty$, and assume that the following assertions hold:*

(a) *$a_n(\widehat{F}_n - F_n)$ takes values only in $\mathbf{D}_\phi$ and satisfies*

$$a_n(\widehat{F}_n - F_n) \rightsquigarrow^\circ B \qquad \text{in } (\mathbf{D}_\phi, \mathcal{B}_\phi^\circ, \|\cdot\|_\phi) \tag{A13}$$

for some $(\mathbf{D}_\phi, \mathcal{B}_\phi^\circ)$-valued random variable B on some probability space $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$ with $B(\Omega_0) \subseteq \mathbf{C}_\phi$.

(b) *$a_n(H(\widehat{F}_n) - H(F_n))$ takes values only in $\widetilde{\mathbf{E}}$ and is $(\mathcal{F}, \widetilde{\mathcal{B}}^\circ)$-measurable.*

(c) *H is uniformly quasi-Hadamard differentiable with respect to $\mathcal{S}$ tangentially to $\mathbf{C}_\phi\langle\mathbf{D}_\phi\rangle$ with trace $\widetilde{\mathbf{E}}$ and uniform quasi-Hadamard derivative $\dot{H}_\mathcal{S}$.*

Then, $\dot{H}_\mathcal{S}(B)$ is $(\mathcal{F}_0, \widetilde{\mathcal{B}}^\circ)$-measurable and

$$a_n\big(H(\widehat{F}_n) - H(F_n)\big) \rightsquigarrow^\circ \dot{H}_\mathcal{S}(B) \qquad \text{in } (\widetilde{\mathbf{E}}, \widetilde{\mathcal{B}}^\circ, \|\cdot\|_{\widetilde{\mathbf{E}}}).$$

Note that the measurability assumption in Condition (a) of Theorem A3 is automatically satisfied in the present setting (and is therefore omitted in Condition (a) of Corollary A3). Indeed, $a_n(\widehat{F}_n - F)$ is easily seen to be $(\mathcal{F}, \mathcal{B}_\phi^\circ)$-measurable, because $\mathcal{B}_\phi^\circ$ coincides with the trace σ-algebra of $\mathcal{D}$.

Corollary A4. *Let $\mathcal{S}$ be any set of sequences in $\mathbf{D}(H)$. Let (a_n) be a sequence of positive real numbers with $a_n \to \infty$, and assume that the following assertions hold:*

(a) *$a_n(\widehat{F}_n^* - \widehat{C}_n)$ takes values only in $\mathbf{D}_\phi$, is $(\overline{\mathcal{F}}, \mathcal{B}_\phi^\circ)$-measurable, and*

$$a_n\big(\widehat{F}_n^*(\omega, \cdot) - \widehat{C}_n(\omega)\big) \rightsquigarrow^\circ B \qquad \text{in } (\mathbf{D}_\phi, \mathcal{B}_\phi^\circ, \|\cdot\|_\phi), \qquad \mathbb{P}\text{-a.e. } \omega \tag{A14}$$

for some $(\mathbf{D}_\phi, \mathcal{B}_\phi^\circ)$-valued random variable B on some probability space $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$ with $B(\Omega_0) \subseteq \mathbf{C}_\phi$.

(b) *$a_n(H(\widehat{F}_n^*) - H(\widehat{C}_n))$ takes values only in $\widetilde{\mathbf{E}}$ and is $(\overline{\mathcal{F}}, \widetilde{\mathcal{B}}^\circ)$-measurable.*

(c) *H is uniformly quasi-Hadamard differentiable with respect to $\mathcal{S}$ tangentially to $\mathbf{C}_\phi\langle\mathbf{D}_\phi\rangle$ with trace $\widetilde{\mathbf{E}}$ and uniform quasi-Hadamard derivative $\dot{H}_\mathcal{S}$.*

(d) *The uniform quasi-Hadamard derivative $\dot{H}_\mathcal{S}$ can be extended from $\mathbf{C}_\phi$ to $\mathbf{D}_\phi$ such that the extension $\dot{H}_\mathcal{S} : \mathbf{D}_\phi \to \widetilde{\mathbf{E}}$ is $(\mathcal{B}_\phi^\circ, \widetilde{\mathcal{B}}^\circ)$-measurable, and continuous at every point of $\mathbf{C}_\phi$.*

(e) *$(\widehat{C}_n(\omega)) \in \mathcal{S}$ for $\mathbb{P}$-a.e. ω.*

(f) *The map $h : \overline{\widetilde{\mathbf{E}}} \to \widetilde{\mathbf{E}}$ defined by $h(\widetilde{x}_1, \widetilde{x}_2) := \widetilde{x}_1 - \widetilde{x}_2$ is $(\overline{\widetilde{\mathcal{B}}^\circ}, \widetilde{\mathcal{B}}^\circ)$-measurable.*

Then, $\dot{H}_\mathcal{S}(B)$ is $(\mathcal{F}_0, \widetilde{\mathcal{B}}^\circ)$-measurable and

$$a_n\big(H(\widehat{F}_n^*(\omega, \cdot)) - H(\widehat{C}_n(\omega))\big) \rightsquigarrow^\circ \dot{H}_\mathcal{S}(B) \qquad \text{in } (\widetilde{\mathbf{E}}, \widetilde{\mathcal{B}}^\circ, \|\cdot\|_{\widetilde{\mathbf{E}}}), \qquad \mathbb{P}\text{-a.e. } \omega.$$

The following examples illustrate $\widehat{F}_n^*$ and $\widehat{C}_n$. In Example A1, we have $\widehat{C}_n = \widehat{F}_n$, and in Example A2 $\widehat{C}_n$ may differ from $\widehat{F}_n$. Examples for uniformly quasi-Hadamard differentiable functionals H can be found in Section 3. In the examples in Sections 3.1 and 3.3 we have $\widetilde{\mathbf{V}} = \widetilde{\mathbf{E}} = \mathbb{R}$, and in the Example in Section 3.2 we have $\widetilde{\mathbf{V}} = \mathbf{D}$ and $\widetilde{\mathbf{E}} = \mathbf{D}_\phi$ for some ϕ.

Example A1. *Let (X_i) be a sequence of i.i.d. real-valued random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ with distribution function F satisfying $\int \phi^2 \, dF < \infty$, and $\widehat{F}_n$ be given by Equation (A12). Let (W_{ni}) be a triangular array of*

nonnegative real-valued random variables on $(\Omega', \mathcal{F}', \mathbb{P}')$ *such that Setting S1. or Setting S2. of Section 2.1 is met. Define the map* $\widehat{F}_n^* : \overline{\Omega} \to \mathbf{D}$ *by* $\widehat{F}_n^*(\omega, \omega') := \frac{1}{n}\sum_{i=1}^n W_{ni}(\omega')\, \mathbb{1}_{[X_i(\omega),\infty)}$. *Recall that Setting S1. is nothing but* Efron's boostrap *(Efron (1979)), and that Setting S2. is in line with the* Bayesian bootstrap *of Rubin (1981) if* Y_1 *is exponentially distribution with parameter* 1.

In Section 5.1 in Beutner and Zähle (2016), it was proved with the help of results of Shorack and Wellner (1986) and van der Vaart Wellner (1996) that respectively Condition (a) of Corollary A3 (with $F_n := F$*) and Condition (a) of Corollary A4 (with* $\widehat{C}_n := \widehat{F}_n$*) hold for* $a_n := \sqrt{n}$ *and* $B := B_F$*, where* B_F *is an F-Brownian bridge. Here,* $\mathbf{C}_\phi$ *can be chosen to be the set* $\mathbf{C}_{\phi,F}$ *of all* $v \in \mathbf{D}_\phi$ *whose discontinuities are also discontinuities of F. In addition, note that, in view of* $\widehat{C}_n = \widehat{F}_n$*, Condition (e) holds if* $\mathcal{S}$ *is (any subset of) the set of all sequences* (G_n) *of distribution functions on* $\mathbb{R}$ *satisfying* $G_n - F \in \mathbf{D}_\phi$*,* $n \in \mathbb{N}$*, and* $\|G_n - F\|_\phi \to 0$ *(see, for instance, Theorem 2.1 in Zähle (2014)).*

Example A2. *Let* (X_i) *be a strictly stationary sequence of β-mixing random variables on* $(\Omega, \mathcal{F}, \mathbb{P})$ *with distribution function F, and* $\widehat{F}_n$ *be given by Equation (A12). Let* (ℓ_n) *be a sequence of integers such that* $\ell_n \nearrow \infty$ *as* $n \to \infty$*, and* $\ell_n < n$ *for all* $n \in \mathbb{N}$*. Set* $k_n := \lceil n/\ell_n \rceil$ *for all* $n \in \mathbb{N}$*. Let* $(I_{nj})_{n\in\mathbb{N}, 1\le j\le k_n}$ *be a triangular array of random variables on* $(\Omega', \mathcal{F}', \mathbb{P}')$ *such that* $I_{n1}, \ldots, I_{nk_n}$ *are i.i.d. according to the uniform distribution on* $\{1, \ldots, n - \ell_n + 1\}$ *for every* $n \in \mathbb{N}$*. Define the map* $\widehat{F}_n^* : \overline{\Omega} \to \mathbf{D}$ *by* $\widehat{F}_n^*(\omega, \omega') := \frac{1}{n}\sum_{i=1}^n W_{ni}(\omega')\, \mathbb{1}_{[X_i(\omega),\infty)}$ *with* W_{ni} *given by Equation (8), and recall from Section 2.2 that this is the* blockwise bootstrap. *Similar as in Lemma 5.3 in Beutner and Zähle (2016) it follows that* $a_n(\widehat{F}_n^* - \widehat{C}_n)$*, with* $\widehat{C}_n := \mathbb{E}'[\widehat{F}_n^*]$*, takes values only in* $\mathbf{D}_\phi$ *and is* $(\overline{\mathcal{F}}, \mathcal{B}_\phi^\circ)$*-measurable. That is, the first part of Condition (a) of Corollary A4 holds true for* $\widehat{C}_n := \mathbb{E}'[\widehat{F}_n^*]$*. Now, assume that Assumptions A1.–A3. of Section 2.2 hold true. Then, as discussed in Example 4.4 and Section 5.2 of Beutner and Zähle (2016), it can be derived from a result in Arcones and Yu (1994) that under Assumptions A1. and A2. we have that Condition (a) of Corollary A3 holds for* $a_n := \sqrt{n}$*,* $B := B_F$*, and* $F_n := F$*, where* B_F *is a centered Gaussian process with covariance function* $\Gamma(t_0, t_1) = F(t_0 \wedge t_1)(1 - F(t_0 \vee t_1)) + \sum_{i=0}^1 \sum_{k=2}^\infty \mathbb{C}\mathrm{ov}(\mathbb{1}_{\{X_1 \le t_i\}}, \mathbb{1}_{\{X_k \le t_{1-i}\}})$*. Here,* $\mathbf{C}_\phi$ *can be chosen to be the set* $\mathbf{C}_{\phi,F}$ *of all* $v \in \mathbf{D}_\phi$ *whose discontinuities are also discontinuities of F. Moreover, Theorem A5 below shows that under the assumptions A1.–A3. the second part of Condition (a) (i.e., Equation (A14)) and Condition (e) of Corollary A4 hold for* $\widehat{C}_n := \mathbb{E}'[\widehat{F}_n^*] = \frac{1}{n}\sum_{i=1}^n w_{ni}\mathbb{1}_{[X_i,\infty)}$ *with* $w_{ni} := \mathbb{E}'[W_{ni}]$ *(see also Equation (9)) and the same choice of* a_n*, B, and* F_n*, when* $\mathcal{S}$ *is the set of all sequences* $(G_n) \subseteq \mathbf{D}(H)$ *with* $G_n - F \in \mathbf{D}_\phi$*,* $n \in \mathbb{N}$*, and* $\|G_n - F\|_\phi \to 0$*.*

Further examples for Condition (a) in Corollary A4 for dependent observations can, for example, be found in Bühlmann (1994); Naik-Nimbalkar and Rajarshi (1994); Peligrad (1998).

Theorem A5. *In the setting of Example A2 assume that assertions A1.–A3. of Section 2.2 hold, and let* $\mathcal{S}$ *be the set of all sequences* $(G_n) \subseteq \mathbf{D}(H)$ *with* $G_n - F \in \mathbf{D}_\phi$*,* $n \in \mathbb{N}$*, and* $\|G_n - F\|_\phi \to 0$*. Then, the second part of assertion (a) (i.e., Equation (A14)) and assertion (e) in Corollary A4 hold.*

Proof. *Proof of second part of (a):* It is enough to show that under assumptions A1.–A3. the Assumptions (A1)–(A4) of Theorem 1 in Bühlmann (1995) hold when the class of functions is $\mathbf{F}_\phi := \mathbf{F}_\phi^- \cup \mathbf{F}_\phi^+$. Here, $\mathbf{F}_\phi^- := \{f_x : x \le 0\}$ and $\mathbf{F}_\phi^+ := \{f_x : x > 0\}$ with $f_x(\cdot) := \phi(x)\mathbb{1}_{(-\infty,x]}(\cdot)$ for $x \le 0$ and $f_x(\cdot) := -\phi(x)\mathbb{1}_{(x,\infty)}(\cdot)$ for $x > 0$. Due to A2. and A3. we only have to verify Assumptions (A3) and (A4) of Theorem 1 in Bühlmann (1995). That is, we show that the following two assertions hold.

(1) There exist constants $b, c > 0$ such that $N_{[\,]}(\varepsilon, \mathbf{F}_\phi, \|\cdot\|_p) \le c\,\varepsilon^{-b}$ for all $\varepsilon > 0$.
(2) $\int \overline{f}^p dF < \infty$ for the envelope function $\overline{f}(z) := \sup_{x\in\mathbb{R}} |f_x(z)|$.

Here, the bracketing number $N_{[\,]}(\varepsilon, \mathbf{F}_\phi, \|\cdot\|_p)$ is the minimal number of ε-brackets with respect to $\|\cdot\|_p$ (L^p-norm with respect to dF) to cover $\mathbf{F}_\phi$, where an ε-bracket with respect to $\|\cdot\|_p$ is the set, $[\ell, u]$, of all functions f with $\ell \le f \le u$ for some Borel measurable functions $\ell, u : \mathbb{R} \to \mathbb{R}_+$ with $\ell \le u$ pointwise and $\|u - \ell\|_p \le \varepsilon$.

(1) We only show that (1) with $\mathbf{F}_\phi$ replaced by $\mathbf{F}_\phi^-$ holds true. Analogously, one can show that the same holds true for $\mathbf{F}_\phi^+$ (and therefore for $\mathbf{F}_\phi$). On the one hand, since $I_p^- := \int_{(-\infty,0]} \phi^p \, dF < \infty$ by Assumption (a), we can find for every $\varepsilon > 0$ a finite partition $-\infty = y_0^\varepsilon < y_1^\varepsilon < \cdots < y_{k_\varepsilon}^\varepsilon = 0$ such that

$$\max_{i=1,\ldots,k_\varepsilon} \int_{(y_{i-1}^\varepsilon, y_i^\varepsilon]} \phi^p \, dF \le (\varepsilon/2)^p \tag{A15}$$

and $k_\varepsilon \le \lceil I_p^- / (\varepsilon/2)^p \rceil$. On the other hand, using integration by parts we obtain

$$\int_{(-\infty,0]} F \, d(-\phi^p) \;=\; \phi(0)^p F(0) - \int_{(-\infty,0]} (-\phi^p) \, dF \;=\; \phi(0)^p F(0) + I_p^-,$$

so that we can find a finite partition $-\infty = z_0^\varepsilon < z_1^\varepsilon < \cdots < z_{m_\varepsilon}^\varepsilon = 0$ such that

$$\max_{i=1,\ldots,m_\varepsilon} \int_{(z_{i-1}^\varepsilon, z_i^\varepsilon]} F \, d(-\phi^p) \le (\varepsilon/2)^p \tag{A16}$$

and $m_\varepsilon \le \lceil (\phi(0)^p F(0) + I_p^-)/(\varepsilon/2)^p \rceil$.

Now, let $-\infty = x_0^\varepsilon < x_1^\varepsilon < \cdots < x_{k_\varepsilon + m_\varepsilon}^\varepsilon = 0$ be the partition consisting of all points y_i^ε and z_i^ε, and set

$$\begin{aligned} \ell_i^\varepsilon(\cdot) &:= \phi(x_i^\varepsilon) \mathbb{1}_{(-\infty, x_{i-1}^\varepsilon]}(\cdot), \\ u_i^\varepsilon(\cdot) &:= \phi(x_{i-1}^\varepsilon) \mathbb{1}_{(-\infty, x_{i-1}^\varepsilon]}(\cdot) + \phi(\cdot) \mathbb{1}_{(x_{i-1}^\varepsilon, x_i^\varepsilon]}(\cdot). \end{aligned} \tag{A17}$$

Then, $\ell_i^\varepsilon \le u_i^\varepsilon$. Moreover,

$$\begin{aligned} \|u_i^\varepsilon - \ell_i^\varepsilon\|_p &= \Big(\int (u_i^\varepsilon - \ell_i^\varepsilon)^p \, dF \Big)^{1/p} \\ &\le \Big(\int_{(-\infty, x_{i-1}^\varepsilon]} (\phi(x_{i-1}^\varepsilon) - \phi(x_i^\varepsilon))^p \, dF \Big)^{1/p} + \Big(\int_{(x_{i-1}^\varepsilon, x_i^\varepsilon]} \phi^p \, dF \Big)^{1/p} \\ &\le \Big(\int_{(-\infty, x_{i-1}^\varepsilon]} (\phi(x_{i-1}^\varepsilon)^p - \phi(x_i^\varepsilon)^p) \, dF \Big)^{1/p} + \varepsilon/2 \\ &\le \Big((\phi(x_{i-1}^\varepsilon)^p - \phi(x_i^\varepsilon)^p) F(x_{i-1}^\varepsilon) \Big)^{1/p} + \varepsilon/2 \end{aligned}$$

where we used Minkovski's inequality and Equation (A15), and that ϕ is non-increasing on $(-\infty, 0]$ and $x_{i-1}^\varepsilon \le x_i^\varepsilon$. Since F is at least $F(x_{i-1}^\varepsilon)$ on $(x_{i-1}^\varepsilon, x_i^\varepsilon]$, we have

$$(\phi(x_{i-1}^\varepsilon)^p - \phi(x_i^\varepsilon)^p) F(x_{i-1}^\varepsilon) \le \int_{(x_{i-1}^\varepsilon, x_i^\varepsilon]} F \, d(-\phi^p) \le (\varepsilon/2)^p$$

due to Equation (A16). Thus, $\|u_i^\varepsilon - \ell_i^\varepsilon\|_p \le \varepsilon$, so that $[\ell_i^\varepsilon, u_i^\varepsilon]$ provides an ε-bracket with respect to $\|\cdot\|_p$. It is moreover obvious that the ε-brackets $[\ell_i^\varepsilon, u_i^\varepsilon]$, $i = 1, \ldots, k_\varepsilon + m_\varepsilon$, cover $\mathbf{F}_\phi^-$. Thus, $N_{[\,]}(\varepsilon, \mathbf{F}_\phi^-, \|\cdot\|_p) \le c\,\varepsilon^{-p}$ for a suitable constant $c > 0$ and all $\varepsilon > 0$.

(2) The envelope function $\overline{f}$ is given by $\overline{f}(y) = \phi(y)$ for $y \le 0$ and by $\overline{f}(y) = \phi(y-) = \phi(y)$ (recall that ϕ is continuous) for $y > 0$. Then, under Assumption (a) the integrability condition 2) holds.

Proof of (e): We have to show that $\|\widehat{C}_n - F\|_\phi = \sup_{x \in \mathbb{R}} |\widehat{C}_n(x) - F(x)|\phi(x) \to 0$ $\mathbb{P}$-a.s. We only show that

$$\sup_{x \in (-\infty, 0]} |\widehat{C}_n(x) - F(x)|\phi(x) \longrightarrow 0 \qquad \mathbb{P}\text{-a.s.}, \tag{A18}$$

because the analogue for the positive real line can be shown in the same way. Let ℓ_i^ε and u_i^ε be as defined in Equation (A17). By assumption A1. we have $\int \phi \, dF < \infty$, so that similar as above we can find a finite partition $-\infty = x_0^\varepsilon < x_1^\varepsilon < \cdots < x_{k_\varepsilon + m_\varepsilon}^\varepsilon = 0$ such that $[\ell_i^\varepsilon, u_i^\varepsilon]$, $i = 1, \ldots, k_\varepsilon + m_\varepsilon$, are

ε-brackets with respect to $\|\cdot\|_1$ (L^1-norm with respect to F) covering the class $\mathbf{F}_\phi := \{f_x : x \in \mathbb{R}\}$ introduced above. We proceed in two steps.

Step 1. First we show that

$$\sup_{x\le 0}\left|\widehat{C}_n(x) - F(x)\right|\phi(x) \le \max_{i=1,\dots,k_\varepsilon+m_\varepsilon} \max\left\{\int u_i^\varepsilon\, d(\widehat{C}_n - F)\,;\, \int \ell_i^\varepsilon\, d(F-\widehat{C}_n)\right\} + \varepsilon \tag{A19}$$

holds true for every $\varepsilon > 0$. Since $(\widehat{C}_n(x) - F(x))\phi(x) = \int f_x\, d\widehat{C}_n - \int f_x\, dF$, for Equation (A19) it suffices to show

$$\begin{aligned}&\sup_{x\le 0}\left|\int f_x\, d\widehat{C}_n - \int f_x\, dF\right| \\ &\le \max_{i=1,\dots,k_\varepsilon+m_\varepsilon} \max\left\{\int u_i^\varepsilon\, d(\widehat{C}_n - F)\,;\, \int \ell_i^\varepsilon\, d(F-\widehat{C}_n)\right\} + \varepsilon.\end{aligned} \tag{A20}$$

To prove Equation (A20), we note that for every $x \in (-\infty, y]$ there is some $i_x \in \{1,\dots,k_\varepsilon+m_\varepsilon\}$ such that $f_x \in [\ell_{i_x}^\varepsilon, u_{i_x}^\varepsilon]$ (see Step 1). Therefore, since $[\ell_{i_x}^\varepsilon, u_{i_x}^\varepsilon]$ is an ε-bracket with respect to $\|\cdot\|_1$,

$$\begin{aligned}\int f_x\, d\widehat{C}_n - \int f_x\, dF &\le \int u_{i_x}^\varepsilon\, d\widehat{C}_n - \int f_x\, dF \\ &= \int u_{i_x}^\varepsilon\, d(\widehat{C}_n - F) + \int (u_{i_x}^\varepsilon - f_x)\, dF \\ &\le \int u_{i_x}^\varepsilon\, d(\widehat{C}_n - F) + \int (u_{i_x}^\varepsilon - \ell_{i_x}^\varepsilon)\, dF \\ &\le \max_{i=1,\dots,k_\varepsilon+m_\varepsilon} \int u_i^\varepsilon\, d(\widehat{C}_n - F) + \varepsilon.\end{aligned}$$

Analogously, we obtain

$$\int f_x\, d\widehat{C}_n - \int f_x\, dF \ge -\Big(\max_{i=1,\dots,k_\varepsilon+m_\varepsilon} \int \ell_i^\varepsilon\, d(F-\widehat{C}_n) + \varepsilon\Big).$$

That is, Equation (A19) holds true.

Step 2. Because of Equation (A19), for Equation (A18) to be true, it suffices to show that

$$\int \ell_i^\varepsilon\, d(F-\widehat{C}_n) \longrightarrow 0 \quad \text{and} \quad \int u_i^\varepsilon\, d(\widehat{C}_n - F) \longrightarrow 0 \quad \mathbb{P}\text{-a.s.} \tag{A21}$$

for every $i = 1,\dots,k_\varepsilon+m_\varepsilon$. We only show the second convergence in Equation (A21), the first convergence can be shown even easier. We have

$$\begin{aligned}\int u_i^\varepsilon\, d(\widehat{C}_n - F) &= \frac{1}{n}\sum_{j=1}^n \Big(w_{ni}\,\phi(x_{i-1}^\varepsilon)\mathbb{1}_{(-\infty,x_{i-1}^\varepsilon]}(X_j) - \mathbb{E}\big[\phi(x_{i-1}^\varepsilon)\mathbb{1}_{(-\infty,x_{i-1}^\varepsilon]}(X_1)\big]\Big) \\ &\quad + \frac{1}{n}\sum_{j=1}^n \Big(w_{ni}\,\phi(X_j)\mathbb{1}_{(x_{i-1}^\varepsilon,x_i^\varepsilon]}(X_j) - \mathbb{E}\big[\phi(X_1)\mathbb{1}_{(x_{i-1}^\varepsilon,x_i^\varepsilon]}(X_1)\big]\Big) \\ &=: S_1(n) + S_2(n).\end{aligned}$$

The first summand on the right-hand side of

$$\begin{aligned}S_2(n) &= \frac{1}{n}\sum_{j=1}^n \Big(\phi(X_j)\mathbb{1}_{(x_{i-1}^\varepsilon,x_i^\varepsilon]}(X_j) - \mathbb{E}\big[\phi(X_1)\mathbb{1}_{(x_{i-1}^\varepsilon,x_i^\varepsilon]}(X_1)\big]\Big) \\ &\quad + \frac{1}{n}\sum_{j=1}^n (w_{ni}-1)\phi(X_j)\mathbb{1}_{(x_{i-1}^\varepsilon,x_i^\varepsilon]}(X_j)\end{aligned}$$

converges $\mathbb{P}$-a.s. to 0 by Theorem 1 (ii) (and Application 5, p. 924) in Rio (1995) and our assumption A1. The second summand converges $\mathbb{P}$-a.s. to 0 too, which can be seen as follows. From Equation (9), we obtain for n sufficiently large

$$|w_{ni} - 1| \leq \begin{cases} 2 & , \quad i = 1, \ldots, \ell_n \\ \frac{\ell_n - 1}{n - \ell_n + 1} & , \quad i = \ell_n + 1, \ldots, n - \ell_n \\ 2 & , \quad i = n - \ell_n + 1, \ldots, n \end{cases},$$

so that for n sufficiently large

$$\begin{aligned} &\left| \frac{1}{n} \sum_{j=1}^{n} (w_{ni} - 1) \phi(X_j) \mathbb{1}_{(x_{i-1}^{\varepsilon}, x_i^{\varepsilon}]}(X_j) \right| \\ &\leq \quad \frac{\ell_n - 1}{n - \ell_n + 1} \frac{1}{n} \sum_{j=\ell_n+1}^{n-\ell_n} \phi(X_j) \mathbb{1}_{(x_{i-1}^{\varepsilon}, x_i^{\varepsilon}]}(X_j) \\ &\quad + 2 \frac{2\ell_n}{n} \frac{1}{2\ell_n} \Big(\sum_{j=1}^{\ell_n} \phi(X_j) \mathbb{1}_{(x_{i-1}^{\varepsilon}, x_i^{\varepsilon}]}(X_j) + \sum_{j=n-\ell_n+1}^{n} \phi(X_j) \mathbb{1}_{(x_{i-1}^{\varepsilon}, x_i^{\varepsilon}]}(X_j) \Big) \\ &=: \quad S_{2,1}(n) + S_{2,2}(n). \end{aligned}$$

We have seen above that $\frac{1}{n} \sum_{j=1}^{n} \phi(X_j) \mathbb{1}_{(x_{i-1}^{\varepsilon}, x_i^{\varepsilon}]}(X_j)$ converges $\mathbb{P}$-a.s. to the constant $\mathbb{E}[\phi(X_1)\, \mathbb{1}_{(x_{i-1}^{\varepsilon}, x_i^{\varepsilon}]}(X_1)]$. Since ℓ_n converges to ∞ at a slower rate than n (by assumption A3.), it follows that $S_{2,1}(n)$ converges $\mathbb{P}$-a.s. to 0. Using the same arguments we obtain that $S_{2,2}(n)$ converges $\mathbb{P}$-a.s. to 0. Hence, $S_2(n)$ converges $\mathbb{P}$-a.s. to 0. Analogously, one can show that $S_1(n)$ converges $\mathbb{P}$-a.s. to 0. □

Author Contributions: Both authors contributed equllay to all sections of the article.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

Acerbi, Carlo. 2002. Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking & Finance* 26: 1505–18.

Acerbi, Carlo, and Balazs Szekely. 2014. *Backtesting Expected Shortfall*. New York: Morgan Stanley Capital International.

Acerbi, Carlo, and Dirk Tasche. 2002a. On the coherence of expected shortfall. *Journal of Banking & Finance* 26: 1487–503.

Acerbi, Carlo, and Dirk Tasche. 2002b. Expected Shortfall: A natural coherent alternative to Value at Risk. *Economic Notes* 31: 379–88. [CrossRef]

Arcones, Miguel Angel, and Bin Yu. 1994. Central limit theorems for empirical and U-processes of stationary mixing sequences. *Journal of Theoretical Probability* 7: 47–71. [CrossRef]

Bauer, Heinz. 2001. *Measure and Integration Theory*. Berlin: De Gruyter.

Belloni, Alexandre, Victor Chernozhukov, Ivan Frenández-Val, and Christian B. Hansen. 2017. Program evaluation and causal inference with high-dimensional data. *Econometrica* 85: 233–98. [CrossRef]

Beutner, Eric, Wei Biao Wu, and Henryk Zähle. 2012. Asymptotics for statistical functionals of long-memory sequences. *Stochastic Processes and their Applications* 122: 910–29. [CrossRef]

Beutner, Eric, and Henryk Zähle. 2010. A modified functional delta method and its application to the estimation of risk functionals. *Journal of Multivariate Analysis* 101: 2452–63. [CrossRef]

Beutner, Eric, and Henryk Zähle. 2012. Deriving the asymptotic distribution of U- and V-statistics of dependent data using weighted empirical processes. *Bernoulli* 18: 803–22. [CrossRef]

Beutner, Eric, and Henryk Zähle. 2016. Functional delta-method for the bootstrap of quasi-Hadamard differentiable functionals. *Electronic Journal of Statistics* 10: 1181–222. [CrossRef]

Billingsley, Patrick. 1999. *Convergence of Probability Measures*. New York: Wiley.

Bølviken, Eric, and Montserrat Guillen. 2017. Risk aggregation in Solvency II through recursive log-normals. *Insurance: Mathematics and Economics* 73: 20–26. [CrossRef]

Bühlmann, Peter. 1994. Blockwise bootstrapped empirical process for stationary sequences. *Annals of Statistics* 22: 995–1012. [CrossRef]

Bühlmann, Peter. 1995. The blockwise bootstrap for general empirical processes of stationary sequences. *Stochastic Processes and their Applications* 58: 247–65. [CrossRef]

Davison, Anthony C., and David Victor Hinkley. 1997. *Bootstrap Methods and Their Application*. Cambridge: Cambridge University Press.

Dudley, Richard Mansfield. 1966. Weak convergence of probabilities on nonseparable metric spaces and empirical measures on Euclidean spaces. *Illinois Journal of Mathematics* 10: 109–26.

Dudley, Richard Mansfield. 1967. Measures on non-separable metric spaces. *Illinois Journal of Mathematics* 11: 449–53.

Efron, Bradley. 1979. Bootstrap methods: Another look at the jackknife. *Annals of Statistics* 7: 1–26. [CrossRef]

Efron, Bradley, and Robert Tibshirani. 1994. *An introduction to the Bootstrap*. New York: Chapman & Hall.

Emmer, Susanne, Marie Kratz, and Dirk Tasche. 2015. What is the best risk measure in practice? A comparison of standard measures. *Journal of Risk* 18: 31–60. [CrossRef]

Gilat, David, and Roelof Helmers. 1997. On strong laws for generalized L-statistics with dependent data. *Commentationes Mathtematicae Universitatis Carolinae* 38: 187–92.

Gribkova, Nadezhda. 2002. Bootstrap approximation of distributions of the L-statistics. *Journal of Mathematical Sciences* 109: 2088–102. [CrossRef]

Gribkova, Nadezhda. 2016. Department of Theory of Probability and Mathematical Statistics, Saint Petersburg State University, Saint Petersburg, Russia. Personal communication.

Helmers, Roelof, Paul Janssen, and Robert Serfling. 1990. Berry-Esséen and bootstrap results for generalized L-statistics. *Scandinavian Journal of Statistics* 17: 65–77.

Jones, Bruce L., and Ričardas Zitikis. 2003. Empirical estimation of risk measures and related quantities. *North American Actuarial Journal* 7: 44–54. [CrossRef]

Krätschmer, Volker, Alexander Schied, and Henryk Zähle. 2013. Quasi-Hadamard differentiability of general risk functionals and its application. *Statistics and Risk Modeling* 32: 25–47. [CrossRef]

Krätschmer, Volker, and Henryk Zähle. 2017. Statistical inference for expectile-based risk measures. *Scandinavian Journal of Statistics* 44: 425–54.

Lahiri, Soumendra Nath. 2003. *Resampling Methods for Dependent Data*. New York: Springer.

Lauer, Alexandra, and Henryk Zähle. 2015. Nonparametric estimation of risk measures of collective risks. *Statistics and Risk Modeling* 32: 89–102. [CrossRef]

Lauer, Alexandra, and Henryk Zähle. 2017. Bootstrap consistency and bias correction in the nonparametric estimation of risk measures of collective risks. *Insurance: Mathematics and Economics* 74: 99–108. [CrossRef]

Mehra, K. L., and Sudhakara Rao. 1975. On functions of order statistics for mixing processes. *Annals of Statistics* 3: 874–83. [CrossRef]

Naik-Nimbalkar, Uttara V., and M.B. Rajarshi. 1994. Validity of blockwise bootstrap for empirical processes with stationary observations. *Annals of Statistics* 22: 980–94. [CrossRef]

Peligrad, Magda. 1998. On the blockwise bootstrap for empirical processes for stationary sequences. *Annals of Probability* 26: 877–901. [CrossRef]

Pitts, Susan M. 1994. Nonparametric estimation of compound distributions with applications in insurance. *Annals of the Institute of Mathematical Statistics* 46: 537–55.

Pollard, David. 1984. *Convergence of Stochastic Processes*. New York: Springer.

Rio, Emmanuel. 1995. A maximal inequality and dependent Marcinkiewicz-Zygmund strong laws. *Annals of Probability* 23: 918–37. [CrossRef]

Rubin, Donald. 1981. The Bayesian bootstrap. *Annals of Statistics* 9: 130–34. [CrossRef]

Shao, Jun, and Dongsheng Tu. 1995. *The Jackknife and Bootstrap*. New York: Springer.

Shorack, Galen R. 1972. Linear functions of order statistics. *Annals of Mathematical Statistics* 43: 412–27. [CrossRef]

Shorack, Galen R., and Jon A. Wellner. 1986. *Empirical Processes with Applications to Statistics*. New York: Wiley.

Stigler, Stephen M. 1974. Linear functions of order statistics with smooth weight functions. *Annals of Statistics* 2: 676–93. [CrossRef]

Sun, Shuxia, and Fuxia Cheng. 2018. Bootstrapping the Expected Shortfall. *Theoretical Economics Letters* 8: 685–98. [CrossRef]

Tsukahara, Hideatsu. 2013. Estimation of distortion risk measures. *Journal of Financial Econometrics* 12: 213–35. [CrossRef]

Van der Vaart, Aad W., and Jon A. Wellner. 1996. *Weak Convergence and Empirical Processes*. New York: Springer.

Van Zwet, Willem R. 1980. A strong law for linear functionals of order statistics. *Annals of Probability* 8: 986–90. [CrossRef]

Varron, Davit. 2015. Laboratoire de Mathématiques de Besançon, University of Franche-Comté, Besançon, France. Personal communication.

Zähle, Henryk. 2014. Marcinkiewicz–Zygmund and ordinary strong laws for empirical distribution functions and plug-in estimators. *Statistics* 48: 951–64. [CrossRef]

© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).