NBS Working paper
7/2020

On Using Triples to Assess Symmetry Under Weak Dependence

Zacharias Psaradakis, Marián Vávra

NATIONAL BANK OF SLOVAKIA
EUROSYSTEM

www.nbs.sk

research@nbs.sk

This publication is available on the NBS website
www.nbs.sk/en/publications-issued-by-the-nbs/research-publications

ISSN 2585-9269 (online)

On Using Triples to Assess Symmetry Under Weak Dependence

Zacharias Psaradakis* Marián Vávra†‡

December 2, 2020

Abstract

The problem of assessing symmetry about an unspecified center of the one-dimensional marginal distribution of strictly stationary random processes is considered. A well-known U-statistic based on data triples is used to detect deviations from symmetry, allowing the underying process to satisfy suitable mixing or near-epoch dependence conditions. We suggest using subsampling for inference on the target parameter, establish the asymptotic validity of the method in our setting, and discuss data-driven rules for selecting the size of subsamples. The small-sample properties of the proposed procedures are examined by means of Monte Carlo simulations and an application to real output growth rates is also presented.

Keywords: Mixing; Near-epoch dependence; Subsampling; Symmetry; U-statistics; Weak dependence.

JEL-Codes: C12, C22.

1. Introduction

Assessing whether a probability distribution is symmetric about a specified or unspecified center is a problem that has attracted considerable attention. This is not surprising in view of the fact that symmetry plays a fundamental role in many statistical inference and model identification procedures. A variety of nonparametric and robust inference procedures rely heavily on the (often untested) assumption of symmetry and tend to

*Department of Economics, Mathematics and Statistics, Birkbeck, University of London

†Research Department, National Bank of Slovakia, E-mail: marian.vavra@gmail.com

‡Institute of Economics, Comenius University in Bratislava

perform rather poorly under deviations from this assumption. Symmetry is also important in terms of the definition and estimation of location since the center of symmetry of a distribution is its only natural location parameter – and is a location parameter that can be estimated robustly or even adaptively (see, e.g., Beran (1978)). Some well-known problems, such as, for instance, detecting time-reversibility of a random process (Chen, Chou, and Kuan (2000), Psaradakis (2008)) or evaluating the lack of a treatment effect through paired comparisons (e.g., Lehmann and Romano (2005, Sec. 6.8)), may be reduced to that of assessing distributional symmetry of appropriately transformed data. In the context of statistical model building, detecting possible deviations from symmetry of the one-dimensional marginal distribution of the data is a useful model checking tool since asymmetry implies that certain families of parametric models are not valid candidate models. For example, finite-parameter models such as autoregressive moving average (ARMA) models or nonlinear Markovian models with skew-symmetric autoregressive functions (Pemberton and Tong (1981)), whose independent and identically distributed (i.i.d.) driving noise has a symmetric distribution, are inappropriate when the marginal distribution of the underlying random process is asymmetric. Assessing symmetry may also be useful as a way of evaluating the empirical validity of different hypotheses and theoretical models to the extent that the latter rely on or imply distributional symmetry, as is the case, for example, with many option and asset pricing models, rational expectations models, and dynamic stochastic general equilibrium models found in the economics and finance literature. Psaradakis (2016) provides examples from this literature in which the question of whether or not the marginal distribution of economic and financial time series is symmetric is of much theoretical interest.

The present paper focuses on assessing symmetry of the one-dimensional marginal distribution of dependent data. Specifically, we consider using a U-statistic involving triples of observations to detect deviations from symmetry, without specifying or estimating the center of symmetry. Such statistics, which may be thought of as estimators of a measure of skewness of the underlying distribution, have been previously used by Davis and Quade (1978) and Randles, Fligner, Policello, and Wolfe (1980) to develop tests for symmetry under the assumption that the data are realizations of i.i.d. random variables. Our objective in this paper is to extend triples-based procedures to the case of strictly stationary sequences of weakly dependent random variables, thus expanding considerably the range of data sets with which such procedures may be validly used.

Alternative approaches to detecting asymmetry of the marginal distribution of dependent data (in the case of an unspecified location) include, among others, approaches based on moment conditions (Bai and Ng (2005), Psaradakis (2016)), distribution distance measures (Psaradakis (2003), Maasoumi and Racine (2009)), the characteristic function (Leucht (2012)), and order statistics (Psaradakis and Vávra (2015)). In a recent study, Psaradakis and Vávra (2019) investigated the properties of tests for symmetry based on some of these approaches, as well as of tests which have been designed for i.i.d. data. As a way of robustifying tests to deviations from the assumption of independence and/or controlling their levels for a fixed sample size, they explored the possibility of using resampling procedures appropriate for dependent data to construct critical regions for the tests. In a comparison of twenty well-known tests for symmetry, the majority of them developed under i.i.d. conditions, a bootstrap-assisted version of a test based on a U-statistic involving data triples was found to be a serious competitor to all other tests in the presence of serial correlation in the data, providing the best overall performance in terms of finite-sample level accuracy and power. The analysis of Eubank, Iariccia, and Rosenstein (1992), under i.i.d. assumptions, also suggests that the triples test is the test of choice against unimodal asymmetric alternatives. The focus in the present paper on triples-based inference procedures is motivated in part by these findings.

Under suitable regularity conditions, the triples U-statistic is shown to have a Gaussian asymptotic distribution for a large class of strictly stationary random processes that includes absolutely regular processes, strongly mixing processes, and near-epoch dependent functionals of absolutely regular processes. However, unless the infinite-dimensional distributions of such processes are fully specified, the asymptotic variance of the triples statistic is unknown. Rather than relying on the Gaussian asymptotic approximation for inference purposes, we suggest to use the model-free subsampling methodology of Politis and Romano (1994a) to approximate the distribution of the triples U-statistic, estimate its asymptotic variance, and construct confidence intervals and/or hypothesis tests for the target parameter. The basic idea of subsampling is to treat overlapping blocks of adjacent observations as replicates of the original data structure, compute the statistic of interest (in our case the triples U-statistic) over such 'subsamples', and use the subsample replicates of the statistic to approximate its distribution or estimate its variance nonparametrically. As is clear from the thorough review of subsampling by Politis, Romano, and Wolf (1999), the method has wide applicability, is easy to implement in practice, and its asymptotic validity often requires little more than the statistic of interest having a nondegenerate asymptotic distribution (when suitably normalized).

The remainder of the paper is organized as follows. Section 2 introduces the U-statistic based on triples and obtains its asymptotic distribution for large classes of weakly dependent random processes. Section 3 details how subsampling may be used to construct confidence intervals and/or hypotheses tests for the parameter of interest, establishes the asymptotic validity of the method, and discusses data-dependent rules for selecting the subsample size. Section 4 examines the finite-sample properties of the proposed inference procedures by means of Monte Carlo experiments. Section 5 presents a real-data example. Section 6 summarizes and concludes. Proofs are collected in Appendix A.

2. Triples Statistic and its Asymptotic Distribution

Let $\mathbf{X}_n := \{X_1, X_2, \ldots, X_n\}$, $n \in \mathbb{N}$, be an observable segment of a real-valued, strictly stationary random process $\mathbf{X} := \{X_t, t \in \mathbb{Z}\}$ with continuous one-dimensional distribution function $F(x) := \mathsf{P}(X_0 \leqslant x)$, $x \in \mathbb{R}$. The objective is to assess whether F is symmetric about some unspecified centre $\mu \in \mathbb{R}$, that is,

$$F(\mu - x) + F(\mu + x) = 1, \qquad x \in \mathbb{R}, \tag{1}$$

or, equivalently, that $X_0 - \mu$ and $\mu - X_0$ are identically distributed. (As usual, $\mathbb{R}$, $\mathbb{Z}$, $\mathbb{N}_0$, and $\mathbb{N}$ are used throughout to denote the sets of real numbers, integers, nonnegative integers, and positive integers, respectively).

Similarly to Randles, Fligner, Policello, and Wolfe (1980), we consider identifying departures from (1) by means of the U-statistic

$$T_n := \frac{6}{n(n-1)(n-2)} \sum\sum\sum_{1 \leqslant t_1 < t_2 < t_3 \leqslant n} \psi(X_{t_1}, X_{t_2}, X_{t_3}), \qquad n \geqslant 3,$$

with kernel $\psi : \mathbb{R}^3 \to \mathbb{R}$ given by

$$\psi(x_1, x_2, x_3) := \tfrac{1}{3} \left\{ \operatorname{sgn}(x_1 + x_2 - 2x_3) + \operatorname{sgn}(x_1 + x_3 - 2x_2) + \operatorname{sgn}(x_2 + x_3 - 2x_1) \right\},$$

where $\operatorname{sgn}(x) := x^{-1}|x|$ for $x \neq 0$ and $\operatorname{sgn}(0) := 0$. An equivalent formulation was considered by Davis and Quade (1978). If $\mathbf{X}$ is an i.i.d. sequence, then $\mathsf{E}(T_n) = \mathsf{E}[\psi(X_1, X_2, X_3)] = 0$ whenever F satisfies (1).

In the sequel, we relax the independence assumption maintained in Davis and Quade (1978) and Randles, Fligner, Policello, and Wolfe (1980), and allow $\mathbf{X}$ to be a weakly dependent process satisfying suitable mixing conditions. As measures of the degree of dependence, we use the Rozanov–Volkonskii coefficients of absolute regularity

$$\beta(k) := \mathsf{E}\left(\sup_{A\in\mathcal{F}_k^\infty} \left|\mathsf{P}(A|\mathcal{F}_{-\infty}^0) - \mathsf{P}(A)\right|\right), \quad k \in \mathbb{N},$$

and Rosenblatt's strong-mixing coefficients

$$\alpha(k) := \sup_{(A',A)\in\mathcal{F}_{-\infty}^0\times\mathcal{F}_k^\infty} \left|\mathsf{P}(A'\cap A) - \mathsf{P}(A')\mathsf{P}(A)\right|, \quad k \in \mathbb{N},$$

where $\mathcal{F}_{-\infty}^0$ and $\mathcal{F}_k^\infty$ denote the σ-fields generated by $\{X_t, t \leqslant 0\}$ and $\{X_t, t \geqslant k\}$, respectively. The (strictly stationary) process $\mathbf{X}$ is said to be absolutely regular if $\beta(k) \to 0$ as $k \to \infty$, and strongly mixing if $\alpha(k) \to 0$ as $k \to \infty$. Under suitable conditions, the strictly stationary, causal solutions of many commonly used time-series models are known to be absolutely regular and/or strongly mixing (often with geometrically decaying mixing coefficients); examples include ARMA models, nonlinear models with an ergodic Markovian structure, linear state space models, autoregressive conditionally heteroskedastic models, and stochastic volatility models (see, e.g., Doukhan (1994, Sec. 2.4)). Because $\beta(k) \geqslant 2\alpha(k)$ for all $k \in \mathbb{N}$, if $\mathbf{X}$ is absolutely regular, then it is also strongly mixing (and, hence, ergodic). The case where $\mathbf{X}$ is a q-dependent process, for some $q \in \mathbb{N}_0$, is a special case in which $\beta(k) = \alpha(k) = 0$ for all $k > q$.

In addition to absolute regularity and strong mixing, we also consider the case where $\mathbf{X}$ is a near-epoch dependent (two-sided) functional of a mixing sequence. More specifically, for a real-valued, strictly stationary random process $\mathbf{V} := \{V_t, t \in \mathbb{Z}\}$ and a measurable function $f : \mathbb{R}^\mathbb{Z} \to \mathbb{R}$, let $\mathbf{X}$ be such that $X_t = f(\{V_{t+j}, j \in \mathbb{Z}\})$ for each $t \in \mathbb{Z}$. If X_0 is integrable and there exists a sequence of nonnegative constants $\{\xi(m), m \in \mathbb{N}_0\}$ such that $\xi(m) \to 0$ as $m \to \infty$ and

$$\mathsf{E}(|X_0 - \mathsf{E}(X_0|\mathcal{G}_{-m}^m)|) \leqslant \xi(m), \quad m \in \mathbb{N}_0,$$

where $\mathcal{G}_{-m}^m$ denotes the σ-field generated by $\{V_t, -m \leqslant t \leqslant m\}$, then $\mathbf{X}$ is said to be near-epoch dependent on $\mathbf{V}$ (in $\mathbb{L}^1$-norm), or an 1-approximating functional of $\mathbf{V}$, with approximation constants $\{\xi(m)\}$. Restricting f in this fashion so that it can be sufficiently well approximated by a finite-variate function is an idea that goes back to Ibragimov (1962). Under suitable regularity conditions, the strictly (and/or second-order) stationary, causal solutions of many time-series models are near-epoch dependent, including ARMA models, autoregressive conditionally heteroskedastic models, nonlinear autoregressive models, and nonlinear models that admit a Volterra series expansion, and so are observables that arise in many dynamical systems (see Borovkova, Burton, and Dehling (2001) and Davidson (2002), inter alia). Near-epoch dependence has the advantage of holding in cases where absolute regularity or strong mixing may not. For example, a causal linear process with absolutely summable coefficients and zero-mean i.i.d. noise is near-epoch dependent on the noise sequence; in comparison, absolute regularity additionally requires the process to be invertible and the one-dimensional marginal distribution of the noise to admit a sufficiently smooth Lebesgue density (cf. Doukhan (1994, Theorem 2, p. 79)). In what follows, $\{\tilde{\beta}(k), k \in \mathbb{N}\}$ denote the coefficients of absolute regularity of the underlying process $\mathbf{V}$ (defined analogously to those of $\mathbf{X}$), and it is assumed that $\tilde{\beta}(k) \to 0$ as $k \to \infty$ at an appropriate rate. Hence, the near-epoch dependent process $\mathbf{X}$ is ergodic and strictly stationary but need not be absolutely regular or strongly mixing.

In view of the boundedness of the kernel ψ, the strong law of large numbers for U-statistics due to Aaronson, Burton, Dehling, Gilat, Hill, and Weiss (1996, Theorem U) ensures that, under absolute regularity of $\mathbf{X}$, T_n is a strongly consistent estimator for the parameter

$$\theta := \mathsf{E}[\psi(Y_1, Y_2, Y_3)] = \mathsf{P}(Y_1 + Y_2 - 2Y_3 > 0) - \mathsf{P}(Y_1 + Y_2 - 2Y_3 < 0),$$

where Y_1, Y_2, Y_3 are independent random variables, independent of $\mathbf{X}$, with common distribution function F. This is also true if $\mathbf{X}$ is strongly mixing or near-epoch dependent on an absolutely regular process, provided F is such that the points of discontinuity of ψ form a negligible set with respect to the distribution of the vector (Y_1, Y_2, Y_3). Note that T_n is not necessarily unbiased for θ in the presence of dependence; for example, $\mathsf{E}(T_n) = \theta + \mathcal{O}((9n)^{-1/2})$ if $\beta(k) = \mathcal{O}(k^{-\varrho})$ for some $\varrho \geqslant 1$ (cf. Han (2018, Theorem 3.2)). The expectation of $\psi(Y_1, Y_2, Y_3)$ may be thought of as a measure of skewness for F. Although $\theta = 0$ does not necessarily imply symmetry of F, $\theta = 0$ for any F satisfying (1).

In order to consider the asymptotic distribution of T_n, it is useful to define a function $\psi_1 : \mathbb{R} \to \mathbb{R}$ by

$$\psi_1(x) := \mathsf{E}[\psi(x, Y_2, Y_3)] - \theta = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi(x, y_2, y_3) \mathrm{d}F(y_2) \mathrm{d}F(y_3) - \theta,$$

and put

$$\tau := \sum_{h=-\infty}^{\infty} \mathrm{Cov}[\psi_1(X_0), \psi_1(X_h)].$$

The two-sided series above is convergent to some $\tau \geqslant 0$ under appropriate conditions on F and on the dependence structure of $\mathbf{X}$. Such conditions can be found in Theorem 1 below, which gives the limiting distribution (as $n \to \infty$) of the centered and normed transform $S_n := \sqrt{n}(T_n - \theta)$. Summability of the coefficients of absolute regularity of $\mathbf{X}$ is sufficient for S_n to be asymptotically normal. Under the weaker strong-mixing condition, a suitable polynomial rate of decay of the mixing coefficients, finiteness of some related absolute moment of F, and some smoothness of ψ with respect to F are required. More specifically, it will be assumed that there exist positive constants M, M', κ_0 and κ_0' such that, for every $\kappa \in (0, \kappa_0)$ and $\kappa' \in (0, \kappa_0')$, and any triple of integers (t_1, t_2, t_3) such that $-\infty < t_1 < t_2 < t_3 < \infty$,

$$\mathsf{E}\left(\sup_{\|(x_1,x_2,x_3)-(Y_1,Y_2,Y_3)\|\leqslant\kappa} |\psi(x_1, x_2, x_3) - \psi(Y_1, Y_2, Y_3)|\right) \leqslant M\kappa, \tag{2}$$

$$\mathsf{E}\left(\sup_{|x_{t_1}-X'_{t_1}|\leqslant\kappa'} \left|\psi(x_{t_1}, X_{t_2}, X'_{t_3}) - \psi(X'_{t_1}, X_{t_2}, X'_{t_3})\right|\right) \leqslant M'\kappa', \tag{3}$$

$$\mathsf{E}\left(\sup_{|x_{t_1}-X'_{t_1}|\leqslant\kappa'} \left|\psi(x_{t_1}, X_{t_2}, X_{t_3}) - \psi(X'_{t_1}, X_{t_2}, X_{t_3})\right|\right) \leqslant M'\kappa', \tag{4}$$

where $\|\cdot\|$ denotes the Euclidean vector norm and $\{X'_t, t \in \mathbb{Z}\}$ are i.i.d. random variables that are independent of $\mathbf{X}$ and have distribution function F. The regularity conditions (2)–(4) are analogous to the variation conditions of Denker and Keller (1986) and may be understood as a form of Lipschitz continuity of ψ with respect to the distribution of X_0 (cf. Fischer, Fried, and Wendler (2016)). These variation conditions are also required under near-epoch dependence, along with suitable polynomial rates of decay for the approximation constants and for the coefficients of absolute regularity of the underlying process.

Theorem 1 *Suppose one of the following sets of conditions is satisfied:*

(i) $\mathbf{X}$ *is absolutely regular with* $\sum_{k=1}^{\infty} \beta(k) < \infty$*;*

(ii) $\mathbf{X}$ *is strongly mixing,* $\mathsf{E}(|X_0|^\gamma) < \infty$ *for some* $\gamma > 0$*,* $\alpha(k) = \mathcal{O}(k^{-\eta})$ *for some* $\eta > (2\gamma + 1)/\gamma$*, and* (2)–(4) *hold;*

(iii) $\mathbf{X}$ *is near-epoch dependent on an absolutely regular process* $\mathbf{V}$*,* $\tilde{\beta}(k) = \mathcal{O}(k^{-\nu})$ *and* $\xi(m) = \mathcal{O}(m^{-\nu-2})$ *for some* $\nu > 1$*, and* (2)–(4) *hold.*

Then, if $\tau > 0$*,* $S_n/\sigma \to \mathcal{N}(0, 1)$ *in distribution as* $n \to \infty$*, where* $\sigma^2 := 9\tau$*.*

Remark 1 *As an absolutely regular process is trivially near-epoch dependent on itself, with* $\xi(m) = 0$ *for all* $m \in \mathbb{N}_0$*, part (iii) of Theorem* 1 *contains a version of part (i). The reason for considering the absolutely regular case separately is that the central limit theorem for*

T_n can be obtained under weaker conditions than it is possible under the more general assumption of near-epoch dependence.

If $\tau = 0$ under the conditions of Theorem 1, then it is easily verified that $S_n \to 0$ in probability as $n \to \infty$. In the nondegenerate case where $\tau > 0$, although the distribution of S_n is asymptotically normal, inference about the parameter θ based on hypotheses tests or confidence sets is complicated by the fact that the asymptotic variance σ^2 is unknown and depends on the correlation structure of the underlying process $\mathbf{X}$. We discuss next how these difficulties may be overcome by using suitable nonparametric estimators based on subsamples.

3. Subsampling-Based Inference

In this Section, we consider the use of subsampling to estimate the distribution function and asymptotic variance of S_n and to construct confidence intervals (and hypotheses tests) for θ. We establish the asymptotic validity of subsampling in our setting and discuss data-driven procedures for selecting the subsample size.

3.1. Subsampling Estimators and Asymptotic Validity

For a fixed sample size n and an integer $\ell := \ell(n)$ satisfying $n > \ell \geqslant 3$, let

$$T_{\ell,i} := \frac{6}{\ell(\ell-1)(\ell-2)} \sum\sum\sum_{i \leqslant t_1 < t_2 < t_3 \leqslant i+\ell-1} \psi(X_{t_1}, X_{t_2}, X_{t_3}), \qquad i \in \{1, 2, \ldots, n-\ell+1\},$$

so that, for each i, $T_{\ell,i}$ is a replicate of T_n based on the subsample $\{X_i, X_{i+1}, \ldots, X_{i+\ell-1}\}$. The subsampling estimator of the distribution function of S_n is given by the empirical distribution function associated with $\sqrt{\ell}(T_{\ell,1} - T_n), \ldots, \sqrt{\ell}(T_{\ell,n-\ell+1} - T_n)$, i.e., by

$$H_{n,\ell}(x) := \frac{1}{n-\ell+1} \sum_{i=1}^{n-\ell+1} \mathbf{1}\left\{\sqrt{\ell}(T_{\ell,i} - T_n) \leqslant x\right\}, \qquad x \in \mathbb{R},$$

where $\mathbf{1}\{A\}$ denotes the indicator of A. The asymptotic variance of S_n may be estimated by

$$\begin{aligned}\hat{\sigma}^2_{n,\ell} &:= \int_{-\infty}^{\infty} x^2 \mathrm{d}H_{n,\ell}(x) - \left(\int_{-\infty}^{\infty} x \mathrm{d}H_{n,\ell}(x)\right)^2 \\ &= \frac{\ell}{n-\ell+1}\sum_{i=1}^{n-\ell+1} T^2_{\ell,i} - \ell\left(\frac{1}{n-\ell+1}\sum_{i=1}^{n-\ell+1} T_{\ell,i}\right)^2.\end{aligned}$$

These estimators are consistent under suitable dependence conditions, provided the subsample size ℓ diverges to infinity with n but does so more slowly than n. The following is true when $\mathbf{X}$ is absolutely regular or strongly mixing.

Theorem 2 *Suppose conditions* (i) *or* (ii) *of Theorem* 1 *are satisfied,* $n^{-1}\ell(n) + \ell(n)^{-1} \to 0$ *as* $n \to \infty$*, and* $\tau > 0$*. Then:* (a) $\sup_{x\in\mathbb{R}} |H_{n,\ell}(x) - \mathsf{P}(S_n \leqslant x)| \to 0$ *in probability as* $n \to \infty$*;* (b) $\hat{\sigma}^2_{n,\ell} \to \sigma^2$ *in probability as* $n \to \infty$*.*

Remark 2 *Without invoking the asymptotic normality of S_n in Theorem* 1*, it can be shown that* $\hat{\sigma}^2_{n,\ell} \to \sigma^2$ *in quadratic mean as* $n \to \infty$*, provided* $\alpha(k) \to 0$ *as* $k \to \infty$*,* $\mathsf{E}(S_n^2) \to \sigma^2 > 0$ *as* $n \to \infty$*,* $\{S_n^4, n \geqslant 3\}$ *is uniformly integrable, and* $n^{-1}\ell(n) + \ell(n)^{-1} \to 0$ *as* $n \to \infty$ *(cf. Fukuchi (1999, Theorem 1(a))).*

The subsampling estimators $H_{n,\ell}$ and $\hat{\sigma}^2_{n,\ell}$ are also consistent when $\mathbf{X}$ is near-epoch dependent on an absolutely regular process, as long as, for any fixed $x \in \mathbb{R}$, the indicator random variables $U_{\ell,i}(x) := \mathbf{1}\{S_{\ell,i} \leqslant x\}$, $i \in \{1, 2, \ldots, n-\ell+1\}$, are such that

$$\frac{1}{n-\ell+1}\sum_{h=0}^{n-\ell} |\mathrm{Cov}[U_{\ell,1}(x), U_{\ell,1+h}(x)]| \to 0 \quad \text{as } n \to \infty, \tag{5}$$

where $S_{\ell,i} := \sqrt{\ell}(T_{\ell,i}-\theta)$ is a replicate of S_n based on the subsample $\{X_i, X_{i+1}, \ldots, X_{i+\ell-1}\}$.

Theorem 3 *Suppose conditions* (iii) *of Theorem* 1 *are satisfied,* $n^{-1}\ell(n) + \ell(n)^{-1} \to 0$ *as* $n \to \infty$*,* (5) *holds, and* $\tau > 0$*. Then, the conclusions of Theorem 2 hold true.*

Remark 3 *Condition (5) requires the autocovariances of the subsample quantities $\{U_{\ell,i}(x)\}$, viewed as a process indexed by i, to be strongly Cesàro convergent to zero. Since near-epoch dependence is not necessarily preserved by measurable transformations, obtaining sufficient conditions for (5) in terms of more primitive conditions on the near-epoch dependence characteristics of $\mathbf{X}$ is not as straightforward as it is under strong-mixing or absolute regularity. However, if the subsample replicates $\{T_{\ell,i}, 1 \leqslant i \leqslant n-\ell+1\}$ of T_n, viewed* *as a process indexed by i, retain the near-epoch dependence on $\mathbf{V}$, then, for a fixed ℓ and each fixed $x \in \mathbb{R}$, $\{U_{\ell,i}(x)\}$ is itself near-epoch dependent on $\mathbf{V}$, provided the indicator functions of intervals $(-\infty, x]$ in $\mathbb{R}$ satisfy some suitable continuity condition with respect to the distribution of $S_{\ell,1}$. It suffices, for example, to assume that the indicator function of $(-\infty, x]$ satisfies, uniformly in $x \in \mathbb{R}$, an 1-continuity condition (cf. Borovkova, Burton, and Dehling (2001, Proposition 2.11)) or a variation condition (cf. Wendler (2011, Lemma 3.5)) with respect to the distribution of $S_{\ell,1}$; these conditions hold under continuity and Lipschitz continuity, respectively, of the distribution function of $S_{\ell,1}$. The covariance inequality in Lemma 2.18(i) of Borovkova, Burton, and Dehling (2001) then ensures that $|\mathrm{Cov}[U_{\ell,1}(x), U_{\ell,1+h}(x)]| \to 0$ as $h \to \infty$, from which (5) follows by the convergence lemma of Cesàro sums.*

Theorems 2 and 3 justify the use of quantiles of $H_{n,\ell}$ to construct asymptotically correct subsampling confidence intervals for θ. More specifically, for any given $\delta \in (0,1)$, an (approximate) level-$(1-\delta)$ equal-tailed, two-sided confidence interval for θ is given by

$$\mathcal{C}^{(1)}_{n,\ell}(\delta) := \left[T_n - n^{-1/2} H^{-1}_{n,\ell}(1-\delta/2),\ T_n - n^{-1/2} H^{-1}_{n,\ell}(\delta/2)\right], \tag{6}$$

where $\varphi^{-1}(u) := \inf\{x : \varphi(x) \geqslant u\}$ for an arbitrary nondecreasing function $\varphi : \mathbb{R} \to [0,1]$. Alternatively, an (approximate) level-$(1-\delta)$ symmetric, two-sided confidence interval for θ can be obtained as

$$\mathcal{C}^{(2)}_{n,\ell}(\delta) := \left[T_n - n^{-1/2} \bar{H}^{-1}_{n,\ell}(1-\delta),\ T_n + n^{-1/2} \bar{H}^{-1}_{n,\ell}(1-\delta)\right], \tag{7}$$

where $\bar{H}_{n,\ell}$ is the subsampling estimator of the distribution function of $|S_n|$ given by

$$\bar{H}_{n,\ell}(x) := \frac{1}{n-\ell+1} \sum_{i=1}^{n-\ell+1} \mathbf{1}\left\{\sqrt{\ell}\,|T_{\ell,i} - T_n| \leqslant x\right\}, \qquad x \in [0,\infty).$$

Symmetric confidence intervals are known to enjoy improved coverage in many circumstances and can be shorter than equal-tailed confidence intervals (see Politis, Romano, and Wolf (1999, Ch. 10)).

The following result shows that the subsampling confidence intervals defined in (6) and (7) have asymptotically correct coverage.

Corollary 1 *Suppose the assumptions of Theorem 2 or Theorem 3 are satisfied. Then, for any $\delta \in (0,1)$ and $s \in \{1,2\}$, $\mathsf{P}(\mathcal{C}^{(s)}_{n,\ell}(\delta) \ni \theta) \to 1-\delta$ as $n \to \infty$.*

Another possibility for constructing a confidence interval for θ is to rely on the subsampling variance estimator $\hat{\sigma}^2_{n,\ell}$ and the Gaussian asymptotic approximation to the distribution of S_n, exploiting the fact that, under the conditions of Theorem 2 or Theorem 3, $S_n/\hat{\sigma}_{n,\ell} \to \mathcal{N}(0,1)$ in distribution as $n \to \infty$. A two-sided confidence interval for θ, with asymptotic coverage $1-\delta$, may thus be obtained as

$$\mathcal{C}_{n,\ell}^{(3)}(\delta) := \left[T_n + n^{-1/2}\hat{\sigma}_{n,\ell}\Phi^{-1}(\delta/2),\ T_n - n^{-1/2}\hat{\sigma}_{n,\ell}\Phi^{-1}(\delta/2)\right], \tag{8}$$

where Φ denotes the distribution function of a $\mathcal{N}(0,1)$ random variable.

Remark 4 *An alternative estimator of σ^2 that may be used in place of $\hat{\sigma}_{n,\ell}^2$ to construct a 'Gaussian' confidence interval like (8) is*

$$\tilde{\sigma}_{n,\omega}^2 := 9 \sum_{h=1-n}^{n-1} K(\omega^{-1}|h|) \left(n^{-1} \sum_{t=1}^{n-|h|} \tilde{\psi}_1(X_t)\tilde{\psi}_1(X_{t+|h|}) \right),$$

where $K : [0,\infty) \to \mathbb{R}$ is a bounded weight function with $K(0)=1$, $\omega := \omega(n) > 0$ is a bandwidth parameter such that $n^{-1/2}\omega(n) + \omega(n)^{-1} \to 0$ as $n \to \infty$, and $\tilde{\psi}_1$ is the empirical analogue of ψ_1 given by

$$\tilde{\psi}_1(x) := n^{-2}\sum_{t_2=1}^{n}\sum_{t_3=1}^{n}\psi(x, X_{t_2}, X_{t_3}) - n^{-3}\sum_{t_1=1}^{n}\sum_{t_2=1}^{n}\sum_{t_3=1}^{n}\psi(X_{t_1}, X_{t_2}, X_{t_3}), \quad x \in \mathbb{R}.$$

Estimators of this type were shown by Dehling, Vogel, Wendler, and Wied (2017) and Fischer (2017) to be consistent (under near-epoch dependence conditions stronger than those in Theorem 1). Their practical use requires an appropriate choice of bandwidth ω for a fixed sample size n, a problem not too dissimilar to choosing the subsample size ℓ for the estimator $\hat{\sigma}_{n,\ell}^2$. Since we are mostly interested in subampling-based inference on θ, we will not consider weighted-autocovariances estimators like $\tilde{\sigma}_{n,\omega}^2$ here.

Remark 5 *(i) Although our discussion focuses primarily on confidence intervals for θ (because, unlike tests, they are informative about the degree of uncertainty associated with a point estimate of θ), tests of hypotheses about θ can be easily constructed using (6), (7) and (8). By the familiar duality between hypothesis tests and confidence sets, an asymptotically level-δ equal-tailed test for testing the null hypothesis $\theta = 0$ versus the alternative $\theta \neq 0$ rejects if, and only if, $\mathcal{C}_{n,\ell}^{(1)}(\delta)$ does not contain zero, i.e., if $\sqrt{n}T_n < H_{n,\ell}^{-1}(\delta/2)$ or $\sqrt{n}T_n > H_{n,\ell}^{-1}(1-\delta/2)$. Similarly, an asymptotically level-δ symmetric test rejects if, and only if, zero is not a member of $\mathcal{C}_{n,\ell}^{(2)}(\delta)$, i.e., if $\sqrt{n}\,|T_n| > \bar{H}_{n,\ell}^{-1}(1-\delta)$. The asymptotically level-δ test corresponding to $\mathcal{C}_{n,\ell}^{(3)}(\delta)$ rejects when $\sqrt{n}\,|T_n/\hat{\sigma}_{n,\ell}| > \Phi^{-1}(1-\delta/2)$.*

(ii) A test of $\theta = 0$ versus $\theta \neq 0$ is viewed in Davis and Quade (1978) and Randles, Fligner, Policello, and Wolfe (1980) as a test of the symmetry hypothesis (1) against a general

alternative corresponding to $F(\mu - x_0) + F(\mu + x_0) \neq 1$ for some $x_0 \in \mathbb{R}$. However, since $\theta = 0$ is necessary but not sufficient for (1) to hold, a test which rejects for large values of $|T_n|$ has non-trivial power only against asymmetric alternatives for which $\theta \neq 0$. The difficulty clearly remains when confidence intervals for θ are used to assess deviations from (1). Randles, Fligner, Policello, and Wolfe (1980, p. 169) argue that the class of asymmetric distributions for which $\theta = 0$ is small.

3.2. Choice of Subsample Size

An important issue that arises in the use of subsampling techniques in practice is the selection of a reasonable subsample size $\ell := \ell(n)$ for a given sample size n, a problem akin to that of selecting the block length for blockwise bootstrap methods (see, e.g., Lahiri (2003, Ch. 7)). The choice of ℓ matters because the size of subsamples can affect significantly the performance of subsampling estimators in finite samples. Unfortunately, the asymptotic results in Theorems 2 and 3 give no guidance for the selection of an appropriate subsample size beyond the non-restrictive requirement that it grows at a slower rate than n. To circumvent this difficulty, we consider here two data-driven methods for choosing a subsample size $\ell^* := \ell^*(n)$ from a collection of candidate subsample sizes $\Lambda_n := \{\ell \in \mathbb{N} : 2 < \ell_1(n) \leqslant \ell \leqslant \ell_2(n) < n\}$, based on the discussion in Politis, Romano, and Wolf (1999, Sec. 9.3), namely a 'calibration' method and a 'minimum volatility' method.

The basic idea behind the calibration method is to adjust the subsample size so that a subsampling confidence interval of a fixed nominal level has coverage probability close to the nominal level in a sample of a given size. The procedure is described formally in Algorithm 1.

Algorithm 1 (Calibration)

1.1 For a large $B \in \mathbb{N}$ and some $1 \leqslant l < n$, generate pseudo-samples $\mathbf{X}_n^{*b} := \{X_{n,1}^{*b}, \ldots, X_{n,n}^{*b}\}$, $b = 1, \ldots, B$, of size n by means of a block-resampling scheme based on $\mathbf{X}_n$, with (expected) block length l.

1.2 For a given $\delta \in (0, 1)$, each $b \in \{1, \ldots, B\}$ and each $\ell \in \Lambda_n$, construct a level-$(1-\delta)$ subsampling confidence interval $[I_{n,\ell,1}^{*b}, I_{n,\ell,2}^{*b}]$ for θ using $\mathbf{X}_n^{*b}$ in place of $\mathbf{X}_n$.

1.3 For each $\ell \in \Lambda_n$, compute $\hat{\pi}_n(\ell) := B^{-1} \sum_{b=1}^{B} \mathbf{1}\{I_{n,\ell,1}^{*b} \leqslant T_n \leqslant I_{n,\ell,2}^{*b}\}$.

1.4 Set $\ell^* = \arg\min_{\ell \in \Lambda_n} |\hat{\pi}_n(\ell) - (1 - \delta)|$.

There are several block-resampling schemes that may be used to construct pseudo-samples $\mathbf{X}_n^{*b}$ from a model-free approximation to the distribution of $\mathbf{X}_n$ (see, e.g., Lahiri (2003, pp. 25–36)). These are required in order to obtain an estimate $\hat{\pi}_n$ of a calibration function $\ell \mapsto \pi_n(\ell)$, where $\pi_n(\ell)$ is the coverage probability of a confidence interval for θ with nominal level $1-\delta$ based on subsamples of size ℓ. In Sections 4 and 5, we rely on the resampling scheme associated with the stationary bootstrap of Politis and Romano (1994b). This amounts to constructing $\mathbf{X}_n^{*b}$ from overlapping blocks of adjacent observations from the periodically extended sequence $\{X_{t(\mathrm{mod}\ n)}, t \in \mathbb{N}\}$, with $X_0 = X_n$, the random length of each block being geometrically distributed with mean l. Unlike other block-resampling schemes, the stationary bootstrap produces pseudo-observations $\mathbf{X}_n^{*b}$ that are stationary (conditionally on $\mathbf{X}_n$) and is less sensitive to misspecification of the (expected) block length. The asymptotic validity of the stationary bootstrap for U-statistics (of degree 2) was established by Hwang and Shin (2015) under strong-mixing conditions. For implementing the procedure, we set $l^{-1} = \min\{|2\hat{\rho}_n/(1-\hat{\rho}_n^2)|^{-2/3}n^{-1/3}, 0.99\}$, where $\hat{\rho}_n$ is the lag-1 sample autocorrelation of $\mathbf{X}$ (cf. Carlstein (1986, p. 1178)). Since the choice of the expected block length is of second-order importance in the context of calibration, this approach provides a simple data-dependent choice for l. In order to keep the cost of computations at a manageable level, we set $B = 100$ in the simulations in Section 4, while $B = 1,000$ is used for the real-data application in Section 5.

The minimum-volatility approach to choosing ℓ^* amounts to constructing subsampling confidence intervals of a fixed nominal level for different subsample sizes and then identifying a region where the intervals do not exhibit substantial variability. A formal description of the procedure in our setting is given in Algorithm 2.

Algorithm 2 (Minimum Volatility)

2.1 For a given $\delta \in (0,1)$, a small $d \in \mathbb{N}$, and for each integer ℓ such that $2 < \ell_1(n) - d \leqslant \ell \leqslant \ell_2(n) + d < n$, construct a level-$(1-\delta)$ subsampling confidence interval $[I_{n,\ell,1},\ I_{n,\ell,2}]$ for θ.

2.2 For each $\ell \in \Lambda_n$, compute the volatility index

$$D_n(\ell) := \sum_{s=1}^{2} \left\{ \frac{1}{2d} \sum_{j=-d}^{d} \left(I_{n,\ell+j,s} - \bar{I}_{n,\ell,s} \right)^2 \right\}^{1/2},$$

where $\bar{I}_{n,\ell,s} := (1+2d)^{-1} \sum_{j=-d}^{d} I_{n,\ell+j,s}$.

2.3 Set $\ell^* = \arg\min_{\ell \in \Lambda_n} D_n(\ell)$.

Minimizing the volatility of the endpoints of subsampling confidence intervals, as in

Algorithm 2, is arguably more attractive computationally than the calibration approach in Algorithm 1, especially in the context of Monte Carlo simulations, because it does not require the use of a bootstrap procedure to estimate confidence interval coverage. Since the algorithm is relatively insensitive to the choice of d, we set $d = 2$ in Sections 4 and 5, following the recommendation in Politis, Romano, and Wolf (1999, pp. 199–200) and Romano and Wolf (2001, p. 1297).

Remark 6 *In Algorithms 1 and 2, and in their implementations in Sections 4 and 5, we consider all integers in the interval $[\ell_1(n), \ell_2(n)]$ as candidate subsample sizes. An appropriate subset of Λ_n may alternatively be used in order to reduce the computational burden.*

We conclude this subsection by noting that, under appropiate conditions, the assertions of Theorem $2(\mathrm{a})$ and Corollary 1 remain true when a random (data-dependent) subsample size such as ℓ^* is used instead of a fixed subsample size. From the proof of Theorem 4.1 of Politis, Romano, and Wolf (2001), it can be seen that, for the consistency results to go through in this case, it is sufficient that, in addition to the assumptions already needed to guarantee asymptotic normality of S_n (with $\sigma > 0$): (i) Λ_n is such that, as $n \to \infty$, $\ell_1(n) \to \infty$ and $n^{-1}\ell_2(n) \to 0$; (ii) for each fixed $x \in \mathbb{R}$ and every $\epsilon > 0$, the limit, as $n \to \infty$, of

$$(\ell_2(n) - \ell_1(n) + 1) \sup_{\ell \in \Lambda_n} \mathsf{P}\left(\left|\frac{1}{n-\ell+1}\sum_{i=1}^{n-\ell+1} \tilde{U}_{\ell,i}(x)\right| > \epsilon\right) \tag{9}$$

is zero, where $\tilde{U}_{\ell,i}(x) := \mathbf{1}\{S_{\ell,i} \leqslant x\} - \mathsf{P}(S_{\ell,1} \leqslant x)$.

Under absolute regularity or strong mixing of $\mathbf{X}$, the concentration inequality in Bosq (1996, Theorem 1.3, pp. 25–26) provides an upper bound for (9) that tends to zero as $n \to \infty$, as long as $\alpha(k) = \mathcal{O}(k^{-\varrho})$ for some $\varrho > 1$ (cf. Politis, Romano, and Wolf (2001, Theorem 4.1)). An analogous bound may be obtained, via a Hoeffding-type exponential inequality, if the random variables $\{\tilde{U}_{\ell,i}(x), 1 \leqslant i \leqslant n-\ell+1\}$ form a near-epoch dependent sequence on $\mathbf{V}$ with suitable polynomial rates of decay for the approximation constants and for the coefficients of absolute regularity of $\mathbf{V}$. To see how such an inequality may be established, note that, if $\{\tilde{U}_{\ell,i}(x)\}$ is near-epoch dependent on $\mathbf{V}$, when viewed as a process indexed by $i \in \mathbb{N}$ at any fixed ℓ and x, then $\{\tilde{U}_{\ell,i}(x), \mathcal{G}_{-\infty}^{i}\}$ is a strictly stationary mixingale (in $\mathbb{L}^1$-norm), $\mathcal{G}_{-\infty}^{i}$ being the σ-field generated by $\{V_t, t \leqslant i\}$. Moreover, since $\tilde{U}_{\ell,i}(x)$ is integrable to any order, the rate of convergence to zero of the mixingale coefficients of $\{\tilde{U}_{\ell,i}(x), \mathcal{G}_{-\infty}^{i}\}$ is the same as the slower of the rates at which the approximation constants of $\{\tilde{U}_{\ell,i}(x)\}$ and the coefficients of absolute regularity of $\mathbf{V}$ approach zero (cf. Davidson (1994, Theorem 17.5(i),

p. 264)). Therefore, provided the mixingale decay rate is fast enough for the mixingale coefficients to be summable, $\tilde{U}_{\ell,i}(x)$ admits the decomposition $\tilde{U}_{\ell,i}(x) = W_i + Z_i - Z_{i+1}$, where $\{W_i, \mathcal{G}_{-\infty}^i\}$ is a strictly stationary martingale-difference sequence and $\{Z_i\}$ is a strictly stationary sequence of integrable random variables (Davidson (1994, Theorem 16.6, p. 250)); furthermore, $\{W_i\}$ and $\{Z_i\}$ are uniformly bounded by virtue of the uniform boundedness of $\{\tilde{U}_{\ell,i}(x)\}$ (cf. Vaněček (2006, p. 704)). Hence, arguing as in the proof of Lemma 8 of Vaněček (2006) and using the Azuma–Hoeffding inequality for martingale differences (e.g., Davidson (1994, Theorem 15.20, p. 245)), it can be deduced that there exist constants $c_1, c_2 \in (0, \infty)$ such that

$$\mathsf{P}\left(\left|\sum_{i=1}^{n-\ell+1} \tilde{U}_{\ell,i}(x)\right| > (n-\ell+1)\epsilon\right) \leqslant c_1 \exp(-c_2(n-\ell+1)\epsilon^2),$$

for any $\epsilon > 0$. Consequently, (9) is bounded above by $c_1' \ell_2(n) \exp(-c_2'[n - \ell_2(n) + 1]\epsilon^2)$, for some $c_1' > 0$ and $c_2' > 0$, which tends to zero as $n \to \infty$ and $n^{-1}\ell_2(n) \to 0$.

4. Monte Carlo Simulations

In this section, we report and discuss the results of a simulation study of the finite-sample properties of confidence intervals for the skewness parameter θ.

4.1. Experimental Design

The experimental design is similar to that in Psaradakis and Vávra (2019), and includes both linear and nonlinear data-generating mechanisms. Specifically, we consider artificial data generated according to the following models ($t \in \mathbb{Z}$):

M1: $X_t = 0.8X_{t-1} + \varepsilon_t$,

M2: $X_t = 0.6X_{t-1} - 0.5X_{t-2} + \varepsilon_t$,

M3: $X_t = 0.6X_{t-1} + 0.3\varepsilon_{t-1} + \varepsilon_t$.

M4: $X_t = 0.9X_{t-1}\mathbf{1}\{|X_{t-1}| \leqslant 1\} - 0.3X_{t-1}\mathbf{1}\{|X_{t-1}| > 1\} + \varepsilon_t$,

M5: $X_t = \omega_t\varepsilon_t, \quad \omega_t^2 = 0.05 + (0.1\varepsilon_{t-1}^2 + 0.85)\omega_{t-1}^2$,

M6: $X_t = 0.7X_{t-2}\varepsilon_{t-1} + \varepsilon_t$.

In each case, $\{\varepsilon_t\}$ are i.i.d. random variables whose distribution is either $\mathcal{N}(0,1)$ (labelled N in the various tables) or generalized lambda, with quantile function $u \mapsto \lambda_1 + \lambda_2^{-1}\{u^{\lambda_3} - (1-u)^{\lambda_4}\}$, $u \in (0,1)$, recentered at zero and rescaled to have variance 1 (see Ramberg and Schmeiser (1974)). The values of $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ used in the experiments, taken from Randles, Fligner, Policello, and Wolfe (1980), can be found in Table 1, along with associated measures of skewness and kurtosis based on standardized third and fourth cumulants; distributions S1–S3 are symmetric, whereas A1–A4 are asymmetric. Models M1–M3 represent ARMA processes, the one-dimensional distribution of which is symmetric if ε_t is symmetrically distributed. Models M4, M5 and M6 represent a self-exciting threshold autoregressive process, a generalized autoregressive conditionally heteroskedastic process, and a bilinear process, respectively; in all three cases, the third cumulant of X_t is zero if ε_t is symmetric about zero (Pemberton and Tong (1981), Martins (1999)).

Table 1: Noise Distributions

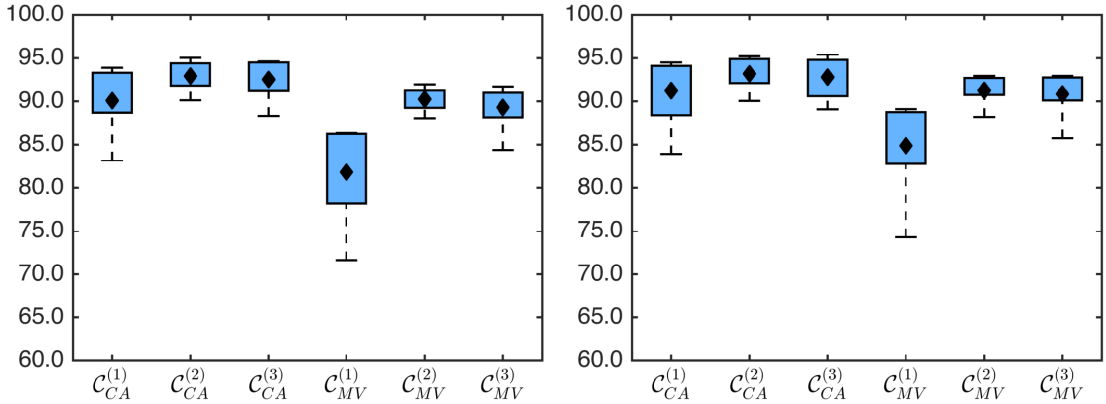
	λ_1	λ_2	λ_3	λ_4	Skewness	Kurtosis
N	–	–	–	–	0.0	3.0
S1	0.000000	-1.000000	-0.080000	-0.080000	0.0	6.0
S2	0.000000	-0.397912	-0.160000	-0.160000	0.0	11.6
S3	0.000000	-1.000000	-0.240000	-0.240000	0.0	126.0
A1	0.000000	-1.000000	-0.007500	-0.030000	1.5	7.5
A2	0.000000	-1.000000	-0.100900	-0.180200	2.0	21.2
A3	0.000000	-1.000000	-0.001000	-0.130000	3.2	23.8
A4	0.000000	-1.000000	-0.000100	-0.170000	3.9	40.7

For each design point, 1,000 independent realizations of $\{X_t\}$ of length $100 + n$, with $n \in \{100, 200\}$, are generated. The first 100 data points of each realization are discarded to minimize initialization effects and the remaining n data points are used to compute the confidence intervals for θ defined in (6), (7) and (8). The subsample size is selected by means of the bootstrap-based calibration algorithm and the minimum-volatility algorithm described in Section 3.2, with $\ell_1(n) = \lfloor (1/2)\sqrt{n} \rfloor$ and $\ell_2(n) = \lfloor (5/2)\sqrt{n} \rfloor$, where $\lfloor a \rfloor$ denotes the integer part of $a \geqslant 0$; these values are in line with the recommendation of Romano and Wolf (2001, p. 1297).

4.2. Simulation Results

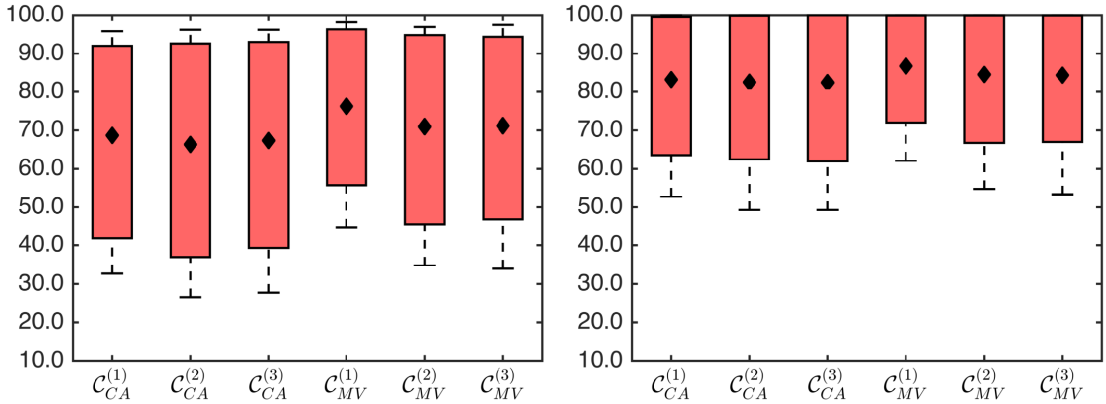
The simulation results over all 24 design points under which the distribution of X_t is symmetric are summarized graphically in Figure 1. This shows boxplots of the estimated coverage probabilities (in percentage) of various confidence intervals for θ, of nominal level $1 - \delta = 0.95$, computed as the percentage of Monte Carlo replications in which each confidence interval correctly includes $\theta = 0$. The top and bottom of each coloured box represent the 25th and 75th percentiles, respectively, of the estimated coverage probabilities, the black diamond inside the box indicates the mean value, and the whiskers indicate the 10th and 90th percentiles. The confidence intervals considered are: the equal-tailed and symmetric subsampling intervals $\mathcal{C}^{(1)}_{n,\ell}(\delta)$ and $\mathcal{C}^{(2)}_{n,\ell}(\delta)$, and the Gaussian approximation interval $\mathcal{C}^{(3)}_{n,\ell}(\delta)$, with subsample size determined using the calibration method (labelled $\mathcal{C}^{(1)}_{CA}$, $\mathcal{C}^{(2)}_{CA}$ and $\mathcal{C}^{(3)}_{CA}$, respectively, in the figures and tables); the corresponding intervals with subsample size determined using the minimum volatility method (labelled $\mathcal{C}^{(1)}_{MV}$, $\mathcal{C}^{(2)}_{MV}$ and $\mathcal{C}^{(3)}_{MV}$, respectively). Detailed results for individual design points can be found in Table 3 in Appendix B.

Figure 1: Monte Carlo results under symmetry; estimated probabilities (in percentage) that 95% confidence intervals contain $\theta = 0$



(a) $n = 100$ (b) $n = 200$

Figure 2: Monte Carlo results under asymmetry; estimated probabilities (in percentage) that 95% confidence intervals do not contain $\theta = 0$



(a) $n = 100$ (b) $n = 200$

It is clear that symmetric subsampling confidence intervals outperform all other competitors when the subsample size is selected by means of the calibration method, having coverage probabilities which are close to the nominal 0.95 level for the vast majority of design points. Selecting the subsample size for such intervals by minimizing the volatility of their endpoints generally leads to somewhat lower coverage, but without the magnitude of the coverage errors making the intervals unattractive for applications.

The confidence interval based on the Gaussian large-sample approximation, used in conjunction with the subsampling variance estimator $\hat{\sigma}^2_{n,\ell}$ and the calibration method, is also a good competitor and often outperforms equal-tailed subsampling confidence intervals. The latter tend to undercover somewhat, the problem being more pronounced when the subsample size is determined via the minimum volatility method.

The simulation results over all 24 design points under which the distribution of X_t is asymmetric are summarized graphically in Figure 2. This show boxplots of the estimated probabilities (in percentage) of $\theta = 0$ being excluded from confidence intervals for θ (of nominal level 0.95), computed as the percentage of Monte Carlo replications in which $\theta = 0$ falls outside each of the confidence intervals. Detailed results for individual design points can be found in Table 4 in Appendix B. Notwithstanding the fact that $\theta = 0$ is not necessarily precluded by asymmetry, the simulation results show that the ability of the various confidence intervals to exclude the value of θ which is typically consistent with symmetry is generally high, with no particular confidence interval dominating. As expected, improved performance is observed with increasing skewness and leptokurtosis in the noise distribution, as well as with an increasing sample size.

5. Real-Data Example

As an illustrative example, we investigate the distributional symmetry of real gross domestic product (GDP), a prominent economic variable analyzed in many studies of the asymmetric behaviour of business cycles (see, inter alia, DeLong and Summers (1986), Verbrugge (1997), Razzak (2001), Narayan and Popp (2009), and Psaradakis (2016)). Our data set consists of time series on real GDP from 15 OECD countries: Australia (AUS), Belgium (BEL), Canada (CAN), Finland (FIN), France (FRA), Italy (ITA), Japan (JAP), Korea (KOR), Netherlands (NLD), Norway (NOR), Portugal (PRT), Spain (ESP), Sweden (SWE), United Kingdom (UK), and United States (US); these represent approximately 35% of the world real GDP (measured in constant 2011 US Dollars). All time series are quarterly, seasonally adjusted, and span the period 1961:1 to 2018:4 (232 observations in total). The data can be downloaded from the OECD website (https://stats.oecd.org/).

DeLong and Summers (1986) characterized asymmetry of the business cycle by the asymmetry of the one-dimensional marginal distribution of the growth rate of a measure of economic output. This type of asymmetry is typically referred to as 'growth-rate' or 'steepness' asymmetry (contractions are steeper than expansions, or vice versa), and is an example of what Ramsey and Rothman (1996) classified as 'longitudinal' asymmetry (asymmetry in the direction of movement of the business cycle). Our analysis is based, therefore, on the growth rates $X_t = 100(Y_t/Y_{t-1} - 1)$, where Y_t is real GDP observed at time t.

For each time series, we compute the subsampling p-value for an equal-tailed test of the null hypothesis $\theta = 0$ versus the alternative $\theta \neq 0$, defined as $P^{(1)}_{n,\ell} := \min\{2P^+_{n,\ell}, 2(1 - P^+_{n,\ell})\}$, where

$$P^+_{n,\ell} := \frac{1}{n-\ell+1} \sum_{i=1}^{n-\ell+1} \mathbf{1}\left\{\sqrt{\ell}(T_{\ell,i} - T_n) \geqslant \sqrt{n} T_n\right\},$$

as well as the subsampling p-value for the corresponding symmetric test, defined as

$$P^{(2)}_{n,\ell} := \frac{1}{n-\ell+1} \sum_{i=1}^{n-\ell+1} \mathbf{1}\left\{\sqrt{\ell}\,|T_{\ell,i} - T_n| \geqslant \sqrt{n}\,|T_n|\right\}.$$

As in the construction of confidence intervals, the subsampling p-values are based on subsample statistics centered at T_n, as recommended by Berg, McMurry, and Politis (2010). In each case, the subsample size ℓ is determined by calibrating the coverage probability of the corresponding subsampling confidence interval for θ (of nominal level 0.95) or by minimizing the volatility of the endpoints of such an interval (see Section 3.2). The resulting p-values are labelled $P^{(1)}_{CA}$, $P^{(2)}_{CA}$, $P^{(1)}_{MV}$ and $P^{(2)}_{MV}$ in Table 2.

Table 2: Empirical Results

	$P^{(1)}_{CA}$	$P^{(2)}_{CA}$	$P^{(1)}_{MV}$	$P^{(2)}_{MV}$
AUS	0.000	0.019	0.000	0.005
BEL	0.378	0.443	0.167	0.500
CAN	0.462	0.527	0.151	0.599
FIN	0.391	0.315	0.304	0.253
FRA	0.124	0.230	0.010	0.286
ITA	0.000	0.117	0.028	0.059
JAP	0.027	0.095	0.000	0.014
KOR	0.000	0.014	0.000	0.000
NLD	0.500	0.648	0.352	0.783
NOR	0.201	0.137	0.071	0.096
PRT	0.391	0.489	0.859	0.485
ESP	0.391	0.600	0.201	0.576
SWE	0.161	0.302	0.215	0.306
UK	0.951	0.805	0.995	0.922
US	0.871	0.762	0.709	0.767

On the basis of symmetric subsampling p-values $P^{(2)}_{CA}$, with the subsample size selected via the calibration method (the best performing combination in our simulations), evidence in favour of asymmetry in real GDP growth rates, at the conventional 0.05 significance level, is found only for Australia and Korea; the test rejects for Japan too if p-values $P^{(2)}_{CA}$ obtained via the minimum volatility method are used instead. The subsampling p-values $P^{(1)}_{CA}$ for equal-tailed tests additionally reject for Italy, and also for France if $P^{(1)}_{MV}$ are used. We conclude, therefore, that steepness does not appear to be a universal characteristic of international business cycles. We finally note that Verbrugge (1997) and Razzak (2001) also used a triples U-statistic to test for business cycle asymmetry; the former relied on Monte Carlo critical values to implement the test, assuming that the data are generated according to an ARMA model with symmetric i.i.d. noise, while the latter treated the data as independent.

6. Conclusion

This paper has considered using a U-statistic based on data triples to assess symmetry of the one-dimensional marginal distribution of strictly stationary random processes satisfying suitable weak dependence conditions. The results given here allow for absolutely regular processes, strongly mixing process, and near-epoch dependent processes with an absolutely regular base. We have discussed how subsampling may be used to draw asymptotically valid inferences about the target skewness parameter. A simulation study has demonstrated that symmetric subsampling confidence intervals based on a data-dependent subsample size determined via calibration have good finite-sample properties and generally outperform equal-tailed subsampling intervals and confidence intervals based on a Gaussian asymptotic approximation. An empirical illustration using time series of output growth rates has also been discussed. The related problem of assessing conditional symmetry of a random process around a parametric or nonparametric function using a triples-based U-statistic is certainly worthy of consideration. Such an extension is not trivial, not least because the kernel of the relevant triples statistic will typically depend on unknown parameters that have to be estimated, and we leave it for future research.

References

Aaronson, J., R. Burton, H. Dehling, D. Gilat, T. Hill, and B. Weiss (1996): "Strong laws for L- and U-statistics," *Transactions of the American Mathematical Society*, 348, 2845–2866.

Arcones, M. A. (1995): "On the central limit theorem for U-statistics under absolute regularity," *Statistics and Probability Letters*, 24, 245–249.

Bai, J., and S. Ng (2005): "Tests for skewness, kurtosis, and normality for time series data," *Journal of Business and Economic Statistics*, 23, 49–60.

Beran, R. (1978): "An efficient and robust adaptive estimator of location," *Annals of Statistics*, 6, 292–313.

Berg, A., T. L. McMurry, and D. N. Politis (2010): "Subsampling p-values," *Statistics and Probability Letters*, 80, 1358–1364.

Borovkova, S., R. Burton, and H. Dehling (2001): "Limit theorems for functionals of mixing processes with applications to U-statistics and dimension estimation," *Transactions of the American Mathematical Society*, 353, 4261–4318.

Bosq, D. (1996): *Nonparametric Statistics for Stochastic Processes*. Springer-Verlag, New York.

Carlstein, E. (1986): "The use of subseries methods for estimating the variance of a general statistic from a stationary time series," *Annals of Statistics*, 14, 1171–1179.

Chen, Y.-T., R. Y. Chou, and C.-M. Kuan (2000): "Testing time reversibility without moment restrictions," *Journal of Econometrics*, 95, 199–218.

Davidson, J. (1994): *Stochastic Limit Theory*. Oxford University Press, Oxford.

——— (2002): "Establishing conditions for the functional central limit theorem in nonlinear and semiparametric time series processes," *Journal of Econometrics*, 106, 243–269.

Davis, C. E., and D. Quade (1978): "U-statistics for skewness or symmetry," *Communications in Statistics – Theory and Methods*, 7, 413–418.

Dehling, H., D. Vogel, M. Wendler, and D. Wied (2017): "Testing for changes in Kendall's tau," *Econometric Theory*, 33, 1352–1386.

DeLong, J. B., and L. H. Summers (1986): "Are Business Cycles Symmetrical?," in *The American Business Cycle: Continuity and Change*, ed. by R. J. Gordon, pp. 166–179. The University of Chicago Press, Chicago.

DENKER, M., AND G. KELLER (1986): "Rigorous statistical procedures for data from dynamical systems," *Journal of Statistical Physics*, 44, 67–93.

DOUKHAN, P. (1994): *Mixing: Properties and Examples*. Springer-Verlag, New York.

EUBANK, R. L., V. N. IARICCIA, AND R. B. ROSENSTEIN (1992): "Testing symmetry about an unknown median, via linear rank procedures," *Journal of Nonparametric Statistics*, 1, 301–311.

FISCHER, A. (2017): "Generalized linear statistics for near epoch dependent processes with application to EGARCH-processes," Institute of Hydrology, Ruhr-Universitochum (arXiv: 1705.04545 [math.ST]).

FISCHER, S., R. FRIED, AND M. WENDLER (2016): "Multivariate generalized linear-statistics of short range dependent data," *Electronic Journal of Statistics*, 10, 646–682.

FUKUCHI, J.-I. (1999): "Subsampling and model selection in time series," *Biometrika*, 86, 591–604.

HAN, F. (2018): "An exponential inequality for U-statistics under mixing conditions," *Journal of Theoretical Probability*, 31, 556–578.

HWANG, E., AND D. W. SHIN (2015): "Stationary bootstrap for U-statistics under strong mixing," *Communications for Statistical Applications and Methods*, 22, 81–93.

IBRAGIMOV, I. A. (1962): "Some limit theorems for stationary processes," *Theory of Probability and Its Applications*, 7, 349–382.

IBRAGIMOV, I. A., AND Y. V. LINNIK (1971): *Independent and Stationary Sequences of Random Variables*. Wolters-Noordhoff, Groningen.

LAHIRI, S. N. (2003): *Resampling Methods for Dependent Data*. Springer-Verlag, New York.

LEHMANN, E. L., AND J. P. ROMANO (2005): *Testing Statistical Hypotheses*. 3rd Edition, Springer, New York.

LEUCHT, A. (2012): "Characteristic function-based hypothesis tests under weak dependence," *Journal of Multivariate Analysis*, 108, 67–89.

MAASOUMI, E., AND J. S. RACINE (2009): "A robust entropy-based test of asymmetry for discrete and continuous processes," *Econometric Reviews*, 28, 246–261.

MARTINS, C. M. (1999): "A note on the third-order moment structure of a bilinear model with non-independent shocks," *Portugaliae Mathematica*, 56, 115–125.

NARAYAN, P. K., AND S. POPP (2009): "Investigating business cycle asymmetry for the G7 countries: Evidence from over a century of data," *International Review of Economics & Finance*, 18, 583–591.

PEMBERTON, J., AND H. TONG (1981): "A note on the distributions of non-linear autoregressive stochastic models," *Journal of Time Series Analysis*, 2, 49–52.

POLITIS, D. N., AND J. P. ROMANO (1994a): "Large sample confidence regions based on subsamples under minimal assumptions," *Annals of Statistics*, 22, 2031–2050.

——— (1994b): "The stationary bootstrap," *Journal of the American Statistical Association*, 89, 1303–1313.

POLITIS, D. N., J. P. ROMANO, AND M. WOLF (1999): *Subsampling*. Springer-Verlag, New York.

——— (2001): "On the asymptotic theory of subsampling," *Statistica Sinica*, 11, 1105–1124.

PSARADAKIS, Z. (2003): "A bootstrap test for symmetry of dependent data based on a Kolmogorov–Smirnov type statistic," *Communications in Statistics – Simulation and Computation*, 32, 113–126.

——— (2008): "Assessing time-reversibility under minimal assumptions," *Journal of Time Series Analysis*, 29, 881–905.

——— (2016): "Using the bootstrap to test for symmetry under unknown dependence," *Journal of Business and Economic Statistics*, 34, 406–415.

PSARADAKIS, Z., AND M. VÁVRA (2015): "A Quantile-based test for symmetry of weakly dependent processes," *Journal of Time Series Analysis*, 36, 587–598.

——— (2019): "Bootstrap-assisted tests of symmetry for dependent data," *Journal of Statistical Computation and Simulation*, 89, 1203–1226.

RAMBERG, J. S., AND B. W. SCHMEISER (1974): "An approximate method for generating asymmetric random variables," *Communications of the ACM*, 17, 78–82.

RAMSEY, J. B., AND P. ROTHMAN (1996): "Time irreversibility and business cycle asymmetry," *Journal of Money, Credit and Banking*, 28, 1–21.

RANDLES, R. H., M. A. FLIGNER, G. E. POLICELLO, AND D. A. WOLFE (1980): "An asymptotically distribution-free test for symmetry versus asymmetry," *Journal of the American Statistical Association*, 75, 168–172.

RAZZAK, W. A. (2001): “Business cycle asymmetries: International evidence,” *Review of Economic Dynamics*, 4, 230–243.

ROMANO, J. P., AND M. WOLF (2001): “Subsampling intervals in autoregressive models with linear time trends,” *Econometrica*, 69, 1283–1314.

SERFLING, R. J. (1980): *Approximation Theorems of Mathematical Statistics*. Wiley, New York.

TEWES, J., D. N. POLITIS, AND D. J. NORDMAN (2019): “Convolved subsampling estimation with applications to block bootstrap,” *Annals of Statistics*, 47, 468–496.

VANĚČEK, P. (2006): “Rate of convergence for a class of RCA estimators,” *Kybernetika*, 42, 699–709.

VERBRUGGE, R. (1997): “Investigating cyclical asymmetries,” *Studies in Nonlinear Dynamics and Econometrics*, 2, 15–22.

WENDLER, M. (2011): “Bahadur representation for U-quantiles of dependent data,” *Journal of Multivariate Analysis*, 102, 1064–1079.

7. Appendix A: Proofs

Proof of Theorem 1: By Hoeffding's decomposition of a U-statistic with kernel ψ (e.g., Serfling (1980, pp. 177–178)),

$$\begin{aligned} S_n &= \frac{3}{\sqrt{n}}\sum_{t=1}^{n}\psi_1(X_t) + \frac{6}{\sqrt{n}(n-1)}\sum\sum_{1\leqslant t_1<t_2\leqslant n}\psi_2(X_{t_1},X_{t_2}) \\ &\quad + \frac{6}{\sqrt{n}(n-1)(n-2)}\sum\sum\sum_{1\leqslant t_1<t_2<t_3\leqslant n}\psi_3(X_{t_1},X_{t_2},X_{t_3}) \\ &=: R_{1,n}+R_{2,n}+R_{3,n}, \end{aligned} \tag{10}$$

where

$$\psi_2(x_1,x_2) := \int_{-\infty}^{\infty}\psi(x_1,x_2,y)\mathrm{d}F(y) - \psi_1(x_1) - \psi_1(x_2) - \theta, \quad x_1,x_2\in\mathbb{R},$$

$$\psi_3(x_1,x_2,x_3) := \psi(x_1,x_2,x_3) - \sum_{i=1}^{3}\psi_1(x_i) - \sum\sum_{1\leqslant i<j\leqslant 3}\psi_2(x_i,x_j) - \theta, \quad x_1,x_2,x_3\in\mathbb{R}.$$

Under (i), and since ψ_1 is measurable, $\{\psi_1(X_t)\}$ is a strictly stationary and uniformly bounded sequence of zero-mean random variables whose coefficients of absolute regularity are bounded above by those of $\mathbf{X}$. Consequently, $\tau<\infty$ and $R_{1,n}\to\mathcal{N}(0,9\tau)$ in distribution as $n\to\infty$, on account of Theorem 18.5.4 in Ibragimov and Linnik (1971, p. 347). Furthermore, noting that $\sup_{-\infty<t_1<t_2<t_3<\infty}|\psi(X_{t_1},X_{t_2},X_{t_3})|<\infty$ almost surely, we have

$$\begin{aligned} \mathsf{E}(R_{2,n}^2) &= \frac{36}{n(n-1)^2}\sum\sum_{1\leqslant t_1<t_2\leqslant n}\sum\sum_{1\leqslant t_3<t_4\leqslant n}\mathsf{E}[\psi_2(X_{t_1},X_{t_2})\psi_2(X_{t_3},X_{t_4})] \\ &\leqslant \frac{36}{n^3}\sum\sum_{1\leqslant t_1<t_2\leqslant n}\sum\sum_{1\leqslant t_3<t_4\leqslant n}|\mathsf{E}[\psi_2(X_{t_1},X_{t_2})\psi_2(X_{t_3},X_{t_4})]|\to 0, \end{aligned} \tag{11}$$

as $n\to\infty$, by the bound given in Lemma 3 of Arcones (1995) and an argument similar to that used in the proof of his Theorem 1. An analogous argument allows us to deduce that

$$\mathsf{E}(R_{3,n}^2)\to 0 \quad \text{as } n\to\infty. \tag{12}$$

Therefore, by Chebyshev's inequality, both $R_{2,n}$ and $R_{3,n}$ converge in probability to zero as $n\to\infty$, and the statement of the theorem follows by Slutsky's lemma.

The stated results under (ii) and (iii) follow as special cases of Theorem 2.3 of Fischer, Fried, and Wendler (2016) and Theorem 2.1 of Fischer (2017), respectively. ■

Proof of Theorem 2: Recalling the decomposition of S_n in (10) and noting that $\tau < \infty$ under the conditions of the theorem, we have

$$\mathsf{E}(R_{1,n}^2) = 9 \sum_{h=1-n}^{n-1} \left(1 - n^{-1}\,|h|\right) \mathsf{E}[\psi_1(X_0)\psi_1(X_h)] \to \sigma^2 \quad \text{as } n \to \infty,$$

by Kronecker's lemma. Moreover, in view of Lemma 3 of Arcones (1995), Lemma 4.3 of Fischer, Fried, and Wendler (2016), and the fact that $\sum_{k=1}^{n} k\{\alpha(k)\}^{\gamma/(2\gamma+1)} \leqslant \sum_{k=1}^{n} k^{1-\eta\gamma/(2\gamma+1)} = \mathcal{O}(n^{\varrho})$ for some $0 < \varrho < 1$, the convergence results in (11) and (12) hold under conditions (i) and (ii) of Theorem 1. Therefore, as $n \to \infty$, $\mathsf{E}[(R_{2,n} + R_{3,n})^2] \to 0$ and $\mathsf{E}[R_{1,n}(R_{2,n} + R_{3,n})] \to 0$, via the C_r-inequality and the Cauchy–Bunyakovskii–Schwarz inequality, respectively, and, hence, $\mathsf{E}(S_n^2) \to \sigma^2$ as $n \to \infty$. Upon noting that the latter result, together with Theorem 1, ensures that the sequence $\{S_n^2, n \geqslant 3\}$ is uniformly integrable (e.g., Serfling (1980, Lemma B, p. 15)), the stated convergence of $H_{n,\ell}$ and $\hat{\sigma}_{n,\ell}^2$ follows from Corollary 2 of Tewes, Politis, and Nordman (2019). ■

Proof of Theorem 3: Since, by Theorem 1 and the continuity of the standard normal distribution function Φ, $\sup_{x\in\mathbb{R}} |\mathsf{P}(S_n \leqslant x) - \Phi(x/\sigma)| \to 0$ as $n \to \infty$, to establish consistency of $H_{n,\ell}$ for the distribution function of S_n it is enough to show that $\sup_{x\in\mathbb{R}} |H_{n,\ell}(x) - \Phi(x/\sigma)| \to 0$ in probability as $n \to \infty$. Hence, by the same argument as in the proof of Theorem 3.2.1 in Politis, Romano, and Wolf (1999, pp. 70–72), it suffices to verify that, for each fixed $x \in \mathbb{R}$, $\bar{U}_{n,\ell}(x) := (n - \ell + 1)^{-1} \sum_{i=1}^{n-\ell+1} U_{\ell,i}(x)$ converges in probability to $\Phi(x/\sigma)$ as $n \to \infty$. Because $\mathsf{E}[\bar{U}_{n,\ell}(x)] = \mathsf{P}(S_{\ell,1} \leqslant x)$ converges to $\Phi(x/\sigma)$ as $n \to \infty$, on account of Theorem 1 and the assumption on ℓ, it remains to show that $\mathrm{Var}[\bar{U}_{n,\ell}(x)] \to 0$ as $n \to \infty$ for each $x \in \mathbb{R}$. To this end, observe that, by the strict stationarity of $\{U_{\ell,i}(x)\}$, viewed as a process indexed by i,

$$\begin{aligned}
\mathrm{Var}[\bar{U}_{n,\ell}(x)] &= \frac{1}{(n-\ell+1)^2} \sum_{h=\ell-n}^{n-\ell} (n - \ell + 1 - |h|)\mathrm{Cov}[U_{\ell,1}(x), U_{\ell,1+|h|}(x)] \\
&\leqslant \frac{1}{(n-\ell+1)^2} \sum_{h=\ell-n}^{n-\ell} (n - \ell + 1 - |h|) \left|\mathrm{Cov}[U_{\ell,1}(x), U_{\ell,1+|h|}(x)]\right| \\
&\leqslant \frac{2}{n-\ell+1} \sum_{h=0}^{n-\ell} \left|\mathrm{Cov}[U_{\ell,1}(x), U_{\ell,1+h}(x)]\right|,
\end{aligned}$$

where, under the assumption in (5), the majorant side converges to zero as $n \to \infty$. Thus, $|H_{n,\ell}(x) - \Phi(x/\sigma)| \to 0$ in probability as $n \to \infty$, for each $x \in \mathbb{R}$, from which uniform convergence in probability follows by the continuity of Φ and a standard subsequence argument.

To prove consistency of $\hat{\sigma}^2_{n,\ell}$ for σ^2, note that

$$\sum_{k=0}^{n} k\tilde{\beta}(k) + \sum_{k=0}^{n} k\left(2\sum_{m=k}^{\infty}\xi(m)\right)^{1/2} \leqslant C\sum_{k=1}^{n} k\left(k^{-\nu} + n^{-(\nu+1)/2}\right)$$
$$= \mathcal{O}(n^{\varrho}) + \mathcal{O}(n^{2-(\nu+1)/2}) = \mathcal{O}(n^{\varrho}),$$

for some $C > 0$ and $\varrho \geqslant 0$. Hence, using Lemma A.2 of Fischer (2017), it is easy to verify that the convergence results in (11) and (12) hold under the conditions of the theorem. By the same argument as in the proof of Theorem 2, it then follows that the sequence $\{S_n^2, n \geqslant 3\}$ is uniformly integrable. This, together with the fact that $|H_{n,\ell}(x) - \Phi(x/\sigma)| \to 0$ in probability as $n \to \infty$, uniformly in $x \in \mathbb{R}$, ensures the stated convergence of $\hat{\sigma}^2_{n,\ell}$ via Theorem 3(iii) of Tewes, Politis, and Nordman (2019). ■

Proof of Corollary 1: By virtue of Theorem 1, Theorem $2(\mathrm{a})$, the continuity of Φ, and the continuous mapping theorem, we have that, as $n \to \infty$, $\sup_{x\in\mathbb{R}} |H_{n,\ell}(x) - \Phi(x/\sigma)| \to 0$ in probability and $\sup_{x\geqslant 0} \left|\bar{H}_{n,\ell}(x) - (2\Phi(x/\sigma) - 1)\right| \to 0$ in probability. Hence, as $n \to \infty$, $H^{-1}_{n,\ell}(\delta) \to \sigma\Phi^{-1}(\delta)$ in probability and $\bar{H}^{-1}_{n,\ell}(1-\delta) \to \sigma\Phi^{-1}(1-\delta/2)$ in probability, for any $\delta \in (0,1)$, from which the stated asymptotic coverage of $\mathcal{C}^{(1)}_{n,\ell}(\delta)$ and $\mathcal{C}^{(2)}_{n,\ell}(\delta)$ follows readily using Slutsky's lemma. ■

8. Appendix B: Tables

Table 3: Monte Carlo Results Under Symmtery

		$\mathcal{C}_{CA}^{(1)}$				$\mathcal{C}_{CA}^{(2)}$				$\mathcal{C}_{CA}^{(3)}$				$\mathcal{C}_{MV}^{(1)}$				$\mathcal{C}_{MV}^{(2)}$				$\mathcal{C}_{MV}^{(3)}$			
n	DGP	N	S1	S2	S3	N	S1	S2	S3	N	S1	S2	S3	N	S1	S2	S3	N	S1	S2	S3	N	S1	S2	S3
100	M1	89.6	84.4	80.6	81.4	95.4	91.6	89.4	88.6	93.2	89.4	86.8	87.6	72.3	72.1	67.2	67.0	89.1	88.1	88.4	86.4	85.3	84.6	82.3	82.4
	M2	94.6	92.8	92.2	90.8	94.4	93.0	93.8	92.4	93.8	92.6	94.0	92.8	88.6	86.3	85.0	86.3	93.1	90.9	89.4	90.4	92.5	91.4	89.8	91.7
	M3	90.2	88.2	86.2	83.4	95.0	92.0	91.0	90.2	94.6	91.0	90.6	88.4	79.4	80.2	76.9	74.8	90.7	91.2	88.4	87.6	89.4	89.8	86.9	86.6
	M4	93.4	93.4	94.6	93.4	93.2	94.6	94.4	93.8	94.6	94.8	94.4	94.2	86.3	86.0	86.7	86.1	90.7	90.3	91.3	89.9	90.8	90.9	91.7	90.6
	M5	92.8	91.8	91.6	93.2	95.0	93.2	93.4	95.6	94.6	94.6	93.4	95.4	86.3	84.1	85.0	85.0	91.3	91.0	91.3	90.8	90.9	90.8	91.6	90.6
	M6	93.8	90.6	90.0	89.2	94.2	92.2	92.4	91.4	93.8	92.4	92.2	91.4	86.2	85.3	85.0	85.4	92.4	91.2	90.2	91.9	91.0	90.4	90.4	91.0
200	M1	87.0	84.0	83.2	82.2	92.0	92.0	88.8	87.0	90.2	89.8	88.0	84.4	76.9	74.4	73.2	71.9	90.4	88.4	86.4	85.0	87.1	85.8	85.3	83.6
	M2	94.0	93.8	94.4	94.4	94.6	93.4	94.8	95.0	94.6	94.0	95.6	95.0	87.5	89.0	89.1	89.1	92.6	92.1	92.8	93.7	92.8	92.7	92.8	92.9
	M3	89.4	89.8	87.4	85.0	92.6	92.2	90.2	90.6	91.0	92.2	89.2	90.2	83.8	82.4	83.2	79.0	92.7	91.0	91.6	88.7	92.2	90.5	91.4	88.5
	M4	93.2	93.6	92.6	94.0	94.8	94.0	93.2	95.2	94.6	93.8	92.6	95.4	88.7	90.8	87.4	88.8	92.0	92.9	91.8	91.7	92.8	93.3	92.5	91.8
	M5	94.2	95.8	93.2	94.4	94.4	95.8	95.0	95.2	94.4	95.8	95.0	94.8	89.1	86.5	88.6	87.6	93.3	91.1	92.7	92.0	93.6	89.7	92.7	91.7
	M6	92.2	95.4	93.6	93.0	92.4	95.4	94.4	93.6	92.8	94.8	94.2	94.6	87.5	87.9	87.5	87.5	92.7	92.2	92.0	90.5	92.3	91.9	91.6	90.7

Note: Entries are estimated probabilities (in percentage) that 95% confidence intervals for θ contain $\theta = 0$.

Table 4: Monte Carlo Results Under Asymmtery

		$\mathcal{C}_{CA}^{(1)}$				$\mathcal{C}_{CA}^{(2)}$				$\mathcal{C}_{CA}^{(3)}$				$\mathcal{C}_{MV}^{(1)}$				$\mathcal{C}_{MV}^{(2)}$				$\mathcal{C}_{MV}^{(3)}$			
n	DGP	A1	A2	A3	A4	A1	A2	A3	A4	A1	A2	A3	A4	A1	A2	A3	A4	A1	A2	A3	A4	A1	A2	A3	A4
100	M1	35.2	31.8	67.2	73.2	31.2	23.2	66.6	73.8	33.6	27.2	68.0	73.6	50.8	45.5	76.8	82.2	38.0	28.2	72.1	78.3	39.3	31.8	73.1	79.0
	M2	63.0	33.6	95.2	95.4	55.6	29.8	95.8	95.0	57.2	30.8	95.8	95.2	70.8	43.0	96.6	98.0	62.8	34.8	94.8	96.6	63.1	34.0	94.4	97.2
	M3	54.0	35.4	88.2	92.2	48.2	27.8	88.6	93.4	49.6	31.0	89.2	94.2	63.2	46.9	93.9	96.0	53.0	35.4	90.7	94.8	54.3	36.8	90.5	94.3
	M4	69.6	32.8	94.4	91.8	59.8	26.2	90.6	89.8	61.6	26.0	91.8	90.0	79.3	44.9	96.7	96.7	70.7	35.4	93.2	93.7	69.4	33.6	93.6	93.5
	M5	89.0	48.2	99.8	99.6	89.8	42.6	99.8	100.0	90.6	45.0	99.8	100.0	94.7	60.6	99.6	99.7	92.7	55.4	99.8	99.8	93.2	55.1	100.0	99.9
	M6	50.0	26.6	90.0	91.8	50.4	26.6	91.8	94.8	50.8	27.8	91.6	95.2	61.6	41.4	94.9	95.3	57.9	34.1	94.8	95.9	57.7	34.2	94.6	95.0
200	M1	57.6	47.2	93.8	92.4	54.8	40.6	93.6	93.4	53.6	40.0	93.0	92.8	63.8	52.4	92.2	95.0	55.4	42.4	93.1	95.7	57.0	42.9	92.4	95.4
	M2	87.4	53.4	100.0	100.0	89.8	51.4	100.0	100.0	89.0	52.4	100.0	100.0	92.2	65.4	100.0	100.0	90.6	57.5	100.0	100.0	89.6	56.7	100.0	100.0
	M3	76.4	56.6	99.0	99.4	74.6	50.6	99.0	100.0	75.8	49.6	98.8	100.0	78.5	63.5	99.8	100.0	75.9	55.0	99.7	100.0	76.9	54.2	99.8	100.0
	M4	94.0	53.6	99.8	99.6	91.8	49.6	100.0	99.6	91.8	49.8	100.0	99.6	95.2	62.8	100.0	99.9	93.4	55.5	100.0	99.9	93.6	53.6	100.0	99.7
	M5	99.0	69.4	99.8	100.0	99.2	70.0	99.8	100.0	99.2	70.4	100.0	100.0	99.9	81.6	100.0	100.0	99.9	79.7	100.0	100.0	99.9	79.9	100.0	100.0
	M6	74.4	45.4	99.0	99.6	76.2	46.6	99.4	100.0	75.8	46.6	99.4	100.0	84.4	56.0	99.7	99.7	82.7	52.5	99.8	99.9	81.9	51.2	99.9	99.7

Note: Entries are the estimated probabilities (in percentage) of exclusion of $\theta = 0$ from 95% confidence intervals.