Econometric Theory, **29**, 2013, 567–589.
doi:10.1017/S0266466612000655

A WARP-SPEED METHOD FOR CONDUCTING MONTE CARLO EXPERIMENTS INVOLVING BOOTSTRAP ESTIMATORS

RAFFAELLA GIACOMINI
University College London/CeMMAP

DIMITRIS N. POLITIS AND HALBERT WHITE
University of California, San Diego

We analyze fast procedures for conducting Monte Carlo experiments involving bootstrap estimators, providing formal results establishing the properties of these methods under general conditions.

1. INTRODUCTION

In spite of ever-increasing computational power, Monte Carlo (MC)-based performance evaluation of bootstrap methods often quickly becomes infeasible due to the multiplicative contribution of every added Monte Carlo iteration to the overall computational cost. In the words of bootstrap pioneer Bradley Efron (2000), "There is some sort of law working here, whereby statistical methodology always expands to strain the current limits of computation." Examples of computationally costly Monte Carlo simulations are, e.g., the comparison of the coverage of alternative confidence intervals, calibration methods for the refinement of bootstrap confidence intervals (see, e.g., Loh, 1988, 1991), selection of optimal block size b for block bootstrap methods (e.g. Politis, Romano, and Wolf, 1999, Ch. 9.3.1), or, in general, any application where the bootstrap is estimating the distribution of a complex statistic. Here we analyze a fast way to perform Monte Carlo experiments whose goal is to assess the performance of bootstrap estimators, test statistics, and confidence intervals. This method was first considered in White (1998), where it was used to investigate the performance of his data snooping reality check. Independently, Davidson and MacKinnon (2000, 2002, 2007)

This paper is dedicated to the memory of Halbert White, inspiring mentor, friend and colleague, master econometrician, and jazz musician. He was a true scholar and an exceptional human being. Raffaella Giacomini gratefully acknowledges financial support from the Economic and Social Research Council through the ESRC Centre for Microdata Methods and Practice grant RES-589-28-0001. Dimitris Politis gratefully acknowledges partial support from NSF Grant DMS-10-07513. Address correspondence to Raffaella Giacomini, University College London, Department of Economics, Gower Street, London WC1E6BT, UK; e-mail: r.giacomini@ucl.ac.uk.

© Cambridge University Press 2013

https://doi.org/10.1017/S0266466612000655 Published online by Cambridge University Press

developed the same key idea to propose a method for dramatically speeding up the double bootstrap (the fast double bootstrap (FDB)).

The goal of a Monte Carlo experiment is typically to assess the performance of a given bootstrap procedure, as summarized by a measure $T(\cdot)$, which could be the accuracy of a bootstrap estimate, the coverage of a bootstrap confidence interval, or the size of a bootstrap hypothesis test. If B is the number of bootstrap resamples and K the number of Monte Carlo replications, the Monte Carlo simulation will involve computing a statistic a total of $K \cdot B$ times. The theoretical justification of the MC approach rests on asymptotic results valid for $B, K \to \infty$, which means that both B and K should be large for the MC experiment to give an accurate estimate of $T(\cdot)$.

When the bootstrap involves computation of complex statistics, as in many important applications, the computational cost can be very high. Examples of studies where the authors admitted having reached the computational boundary and settled for more tractable experiments, are, e.g., Kilian and Chang (2000), Whang (2000), Kim (2001), and Chen and Conley (2001). More recent studies with Monte Carlo simulations that feature small B and K are, e.g., Li and Tkacz (2006) ($B = 100$, $K = 500$), Escanciano and Velasco (2006) ($B = 300$, $K = 1000$), and Canepa (2006) ($B = 400$, $K = 1000$). The key idea underlying the proposals of White (1998) and Davidson and MacKinnon (2000, 2002, 2007) is that taking just *one* bootstrap draw for each simulated sample can suffice to provide a useful approximation to the statistic of interest. Applying this insight to Monte Carlo evaluation of bootstrap-based confidence intervals yields evaluation methods that work with $B = 1$, taking K and n (the underlying estimation sample size) to be large. Because of the resulting dramatic computational savings, we call our method a "Warp-Speed" Monte Carlo method.

Neither White (1998) nor Davidson and MacKinnon (2000, 2002, 2007) provide formal results for their procedures. Thus, a main goal of this paper is to provide explicit formal results for Warp-Speed methods, rigorously establishing their key properties under explicit regularity conditions. Although our Warp-Speed procedure for assessing the coverage of a bootstrap confidence interval is related to Davidson and MacKinnon's (2007) $\widehat{RP}_A$ procedure**,** our regularity conditions differ from those informally relied on in that paper. For example, we do not require our statistics to be asymptotically pivotal, nor do we require differentiability of the root. Davidson and MacKinnon (2007) note that their results hold true "under much less restrictive conditions which are, however, harder to specify precisely." In contrast, our results justify use of our procedures under precise and general conditions.

As should be expected, formal results afford insight into the strengths and weaknesses of the procedure analyzed. In particular, we find that there is a cost to be paid for the significant computational savings of Warp-Speed methods: Compared to standard Monte Carlo-bootstrap methods, Warp-Speed methods are not necessarily able to diagnose situations in which the variance of the bootstrap distribution is performing poorly. Nevertheless, Warp-Speed methods are able to

https://doi.org/10.1017/S0266466612000655 Published online by Cambridge University Press

identify situations in which coverage problems result from nonvanishing bias in the bootstrap distribution, an important aspect of bootstrap behavior.

An appealing feature of our results is that they apply not just to independent and identically distributed (IID) data samples, but also to dependent, possibly heterogeneous data. Further, they apply to bootstrap procedures generally, whether or not the bootstrap works; and they apply to general estimation methods, whether parametric or nonparametric.

The plan of the paper is as follows. In Section 2 we provide a heuristic discussion of standard Monte Carlo-bootstrap methods for evaluating coverage of bootstrap-based confidence intervals, and we introduce the Warp-Speed method. In Section 3 we provide theory for the scalar case with non-IID data. Section 4 extends the results of Section 3 to the case of roots taking values in general spaces. Section 5 provides an illustrative example contrasting the performance of standard methods with our Warp-Speed methods. Section 6 concludes with an application to choosing the block size for the block bootstrap. A mathematical Appendix contains proofs of all our results.

2. HEURISTICS

For concreteness, we consider the case in which the bootstrap is used to construct confidence intervals for a scalar parameter of interest. Given the duality between confidence intervals and hypothesis tests, the results also apply to hypothesis testing, as long as the hypothesis of interest can be expressed as a hypothesis about a parameter.

For clarity in describing the key ideas, for now we consider the IID case originally considered by Efron (1979). Nevertheless, as we establish in the next two sections, our results hold in general contexts, e.g., with stationary data, as in Künsch (1989) and Politis and Romano (1994).

Thus, let $X^n \equiv (X_1, \ldots, X_n)$ be a sample of IID random variables with unknown probability distribution F and let $\theta_F \equiv \boldsymbol{\theta}(F)$ be a scalar parameter of interest. The construction of a confidence interval for θ_F proceeds by first finding a function of both the sample and the parameter, a *root* $R_n(X^n, \theta_F)$, whose distribution is either known or can be consistently estimated. If $\hat{\theta}_n(X^n)$ is an estimator of θ_F based on the sample X^n, familiar choices for the root are $R_n(X^n, \theta_F) = \sqrt{n}(\hat{\theta}_n(X^n) - \theta_F)$ or $R_n(X^n, \theta_F) = \sqrt{n}(\hat{\theta}_n(X^n) - \theta_F)/s_n$, where s_n is an estimate of the (asymptotic) standard deviation of $\hat{\theta}_n$. A confidence interval for θ_F can then be derived by estimating the appropriate quantiles by inverting $J_n(\cdot, F)$, the cumulative distribution function of $R_n(X^n, \theta_F)$.

An approximation to the sampling distribution $J_n(\cdot, F)$ can be obtained using the bootstrap in the following way. Draw B IID bootstrap resamples of size n, $X_1^{*n}, \ldots, X_B^{*n}$, by sampling with replacement from the set $\{X_1, \ldots, X_n\}$, and for each bootstrap sample compute the root $R_n(X_i^{*n}, \hat{\theta}_n(X^n))$. The empirical distribution $J_{n,B}^*(\cdot)$ of the B values $R_n(X_i^{*n}, \hat{\theta}_n(X^n))$ gives the bootstrap approximation to $J_n(\cdot, F)$. Defining the α-quantile of the bootstrap distribution as

https://doi.org/10.1017/S0266466612000655 Published online by Cambridge University Press

$q^*_{n,B}(\alpha) \equiv \inf\{x : J^*_{n,B}(x) \geq \alpha\}$, we obtain a confidence interval (CI) for θ_F of nominal level $1-\alpha$ as

$$CI^*_{n,B}(1-\alpha) = \left\{\theta : q^*_{n,B}(\alpha/2) \leq R_n\left(X^n, \theta\right) \leq q^*_{n,B}(1-\alpha/2)\right\}. \quad \textbf{(1)}$$

Suppose one is interested in analyzing the coverage of the bootstrap confidence interval (1) in samples of size n. Typically, a Monte Carlo experiment is designed to answer this question in the following manner.

- *Standard Monte Carlo experiment.* Draw a set of K IID Monte Carlo samples of size n from the given distribution F. Denote each Monte Carlo sample as $X^n_k \equiv (X_{1,k}, \ldots, X_{n,k})$, $k = 1, \ldots, K$. For each one of these samples, construct a bootstrap confidence interval (1) in the manner described above, which results in a sequence of K confidence intervals

$$CI^*_{n,B,k}(1-\alpha) = \left\{\theta : q^*_{n,B,k}(\alpha/2) \leq R_n\left(X^n_k, \theta\right) \leq q^*_{n,B,k}(1-\alpha/2)\right\}, \quad k = 1, \ldots, K. \quad \textbf{(2)}$$

  The bootstrap quantile $q^*_{n,B,k}(\alpha)$ is obtained from Monte Carlo sample X^n_k by computing the estimator $\hat{\theta}_{n,k} \equiv \hat{\theta}_n(X^n_k)$, drawing B bootstrap resamples $X^{*n}_{k,1}, \ldots, X^{*n}_{k,B}$, computing for each the root $R_n(X^{*n}_{k,i}, \hat{\theta}_{n,k})$, and inverting the empirical distribution $J^*_{n,B,k}$ of the B values $R_n(X^{*n}_{k,i}, \hat{\theta}_{n,k})$. To calculate the empirical coverage of the bootstrap confidence interval, define an indicator variable $A_{n,B,k}$ that equals 1 if θ_F lies inside the kth confidence interval and equals 0 otherwise, and compute the Monte Carlo coverage as $1-\alpha_{n,K,B} = K^{-1}\sum_{k=1}^{K} A_{n,B,k}$.

The theoretical justification for the Monte Carlo procedure rests on the law of large numbers: The estimated Monte Carlo coverage $1-\alpha_{n,K,B}$ will converge to the true finite sample coverage $1-\alpha_n$ of the bootstrap confidence interval for a given sample size n as $\min(K, B) \to \infty$. To guarantee accuracy of the estimated coverage, the number of Monte Carlo replications should therefore be large (typically, one sets $K = 1{,}000$ or $5{,}000$). The Monte Carlo procedure described above involves $B \cdot K$ computations of the root $R_n(X^{*n}_{k,i}, \hat{\theta}_{n,k})$, which can become computationally quite costly for large B and K, particularly when the root involves complex statistics.

An alternative way to estimate the finite sample coverage of the bootstrap confidence interval (1) is to draw just one bootstrap resample, as recommended by White (1998) and Davidson and MacKinnon (2000, 2002, 2007). For assessing confidence interval coverage, the method is as follows.

- *Warp-Speed Monte Carlo experiment.* Draw a set of K IID Monte Carlo samples X^n_k, $k = 1, \ldots, K$, each of size n, from the given F. For each of these samples, draw $B = 1$ bootstrap resample X^{*n}_k and compute the root

https://doi.org/10.1017/S0266466612000655 Published online by Cambridge University Press

$R_n(X_k^{*n}, \hat{\theta}_{n,k})$, $k = 1, \ldots, K$. Let $1\{\cdot\}$ denote the indicator function, and define $\hat{J}_{n,k}(x) \equiv \mathbf{1}\{R_n(X_k^{*n}, \hat{\theta}_n(X_k^n)) \leq x\}$ (i.e., the cdf of a distribution with a point mass at $R_n(X_k^{*n}, \hat{\theta}_n(X_k^n))$), and let

$$\bar{J}_{n,K}(x) \equiv K^{-1} \sum_{k=1}^{K} \hat{J}_{n,k}(x).$$

The empirical distribution $\bar{J}_{n,K}$ can be inverted to calculate the bootstrap quantile $\hat{q}_{n,K}(\alpha)$. A sequence of K confidence intervals is then obtained as

$$\widehat{CI}_{n,K,k}(1-\alpha) = \left\{\theta : \widehat{q}_{n,K}(\alpha/2) \leq R_n\left(X_k^n, \theta\right) \leq \widehat{q}_{n,K}(1-\alpha/2)\right\}, \quad k = 1, \ldots, K. \tag{3}$$

Defining an indicator variable $A_{n,K,k}^{WS}$ such that $A_{n,K,k}^{WS}$ equals 1 if θ_F lies inside the kth confidence interval and equals 0 otherwise, we compute the Monte Carlo coverage as $1 - \alpha_{n,K}^{WS} = K^{-1} \sum_{k=1}^{K} A_{n,K,k}^{WS}$.

In this method, the root $R_n(\cdot)$ is only computed K times. When $K \ll B \cdot K$, this method is much faster than the standard Monte Carlo bootstrap. Accordingly, we call our method a "Warp-Speed" method. Its properties are the subject of the next two sections.

3. SCALAR ROOTS WITH NON-IID DATA

In this section we provide formal results establishing the properties of the Warp-Speed bootstrap for scalar roots without imposing the IID assumption. We let a given probability measure P govern the sequence of random variables $\{X_i\}$. Thus, P also governs each sample of size n, $X^n \equiv (X_1, \ldots, X_n)$. The sample may be IID; it may also be dependent, heterogeneous, or both. Next, let $\theta_P \equiv \boldsymbol{\theta}(P)$ denote the parameter of interest and define the associated real-valued root, $R_n(X^n, \theta_P)$. An example is $R_n(X^n, \theta_P) = \tau_n(\hat{\theta}_n(X^n) - \theta_P)$, where $\hat{\theta}_n(X^n)$ is a consistent estimator of θ_P, and τ_n is the rate of convergence of $\hat{\theta}_n(X^n)$ to θ_P; another example is $R_n(X^n, \theta_P) = (\hat{\theta}_n(X^n) - \theta_P)/\hat{s}_n(X^n)$, where $\hat{s}_n(X^n)$ embodies a consistent estimator of the asymptotic standard deviation of $\hat{\theta}_n(X^n)$.

We impose formal assumptions as follows.

Assumption 3.1 (Data generation)**.** Let $(\Omega, \mathcal{F}, P)$ be a complete probability space on which is defined the (infinite) sequence of random variables $X^\infty = (X_1, X_2, \ldots)$, where each X_i takes values in some space $\mathbf{X}$.

The expectation corresponding to probability measure P is denoted $\mathbb{E}$.

Assumption 3.2 (Parameter of interest)**.** Let $P \to \boldsymbol{\theta}(P)$ be a scalar-valued functional.

Assumption 3.3 (Estimator)**.** For $n = 1, 2, \ldots$, $\hat{\theta}_n : \mathbf{X}^n \to \mathbb{R}$ is a measurable mapping, where $\mathbf{X}^n \equiv \times_{i=1}^n \mathbf{X}$.

https://doi.org/10.1017/S0266466612000655 Published online by Cambridge University Press

Assumption 3.4 (Root)**.** For $n = 1, 2, \ldots$, $R_n : \mathbf{X}^n \times \mathbb{R} \to \mathbb{R}$ is a measurable function.

Suppose a resampling method $\mathcal{B}_n$ generates bootstrap pseudo-samples $X^{*n} \equiv (X_1^*, \ldots, X_n^*)$ from the sample X^n with joint distribution P_n^*, the bootstrap distribution conditional on X^n, and let

$$\hat{J}_n\left(x, P_n^*\right) \equiv P_n^*\left\{R_n\left(X^{*n}, \hat{\theta}_n\left(X^n\right)\right) \leq x\right\}.$$

If $\hat{J}_n(x, P_n^*)$ provides a useful approximation to $J(x, P)$, the limiting distribution of $R_n(X^n, \theta_P)$, e.g., $\hat{J}_n(x, P_n^*) \xrightarrow{P} J(x, P)$ as $n \to \infty$ for all continuity points x of $J(\cdot, P)$, then the bootstrap "works" and can be used to construct confidence intervals, etc.

A standard method of approximating $\hat{J}_n(x, P_n^*)$ is to use P_n^* to generate a large number B of IID bootstrap resamples $X_i^{*n}, i = 1, \ldots, B$, and compute

$$J_{n,B}^*(x) \equiv B^{-1} \sum_{i=1}^{B} \mathbf{1}\left\{R_{n,i}^* \leq x\right\},$$

where $R_{n,i}^* \equiv R_n(X_i^{*n}, \hat{\theta}_n(X^n))$. The law of large numbers then ensures that $J_{n,B}^*(x) \to \hat{J}_n(x, P_n^*)$ a.s. $-\mathcal{P}_n^*$ as $B \to \infty$, where $\mathcal{P}_n^* \equiv \Pi_{i=1}^{\infty} P_n^*$ is the probability measure governing the sequence $\{R_{n,i}^*\}$.

To assess the performance of the bootstrap for a given sample size n, it is common to apply Monte Carlo methods as described in the previous section. Formally, we require

Assumption 3.5 (Monte Carlo)**.** Let $X_1^{\infty}, X_2^{\infty}, \ldots$ be a sequence of IID Monte Carlo random elements, where each of these random elements is an IID version of X^{∞} defined in Assumption 3.1.

The countable sequence $\{X_k^{\infty}, k = 1, 2, \ldots\}$ is thus governed by $\mathbf{P} \equiv \Pi_{k=1}^{\infty} P$, with associated expectation $\mathbf{E}$.

The first n entries of the sequence X_k^{∞} are those of practical use in our experiments. We let X_k^n denote the array consisting of these n entries; thus, $X_k^n, k = 1, 2, \ldots$, are independent copies of X^n defined above. We generate bootstrap resamples from each Monte Carlo sample X_k^n as follows.

Assumption 3.6 (Bootstrap)**.** For any natural numbers n and k, apply $\mathcal{B}_n$ to X_k^n to generate independent bootstrap resamples $X_{k,i}^{*n}$, $i = 1, \ldots, B$, each with corresponding bootstrap distribution $P_{n,k}^*$ conditional on X_k^n.

We let $\mathbb{E}_{n,k}^*$ and $\mathrm{Var}_{n,k}^*$ denote the mean and variance associated with the bootstrap probability measure $P_{n,k}^*$. For each n and k, the countable resample sequence $\{X_{k,i}^{*n}\}$ is governed by the probability measure $\mathcal{P}_{n,k}^* \equiv \Pi_{i=1}^{\infty} P_{n,k}^*$.

We formalize our proposed method as follows.

https://doi.org/10.1017/S0266466612000655 Published online by Cambridge University Press

DEFINITION 3.1 (Warp). *Let n be given, and let K be a finite integer. For each Monte Carlo sample X_k^n, $k = 1, \ldots, K$, draw $B = 1$ bootstrap resample X_k^{*n} and compute the root $R_n(X_k^{*n}, \hat{\theta}_n(X_k^n))$. Let $\hat{J}_{n,k}(x) \equiv \mathbf{1}\{R_n(X_k^{*n}, \hat{\theta}_n(X_k^n)) \leq x\}$, and let*

$$\bar{J}_{n,K}(x) \equiv K^{-1} \sum_{k=1}^{K} \hat{J}_{n,k}(x).$$

Our first result describes the behavior of $\bar{J}_{n,K}(x)$. For this, we make use of the Monte Carlo-bootstrap probability distribution $\mathbf{P}_n^* \equiv \Pi_{k=1}^{\infty} P_{n,k}^*$, and we let $\mathbf{E}_n^*$ and $\mathbf{Var}_n^*$ denote the corresponding expectations and variances. We also define the "global probability" $\mathbf{P}_n \equiv \Pi_{k=1}^{\infty}(P P_{n,k}^*) = \mathbf{P}\mathbf{P}_n^*$, with associated "global expectation" $\mathbf{E}_n = \mathbf{E}\mathbf{E}_n^*$ and "global variance" denoted by $\mathbf{Var}_n$. The fact that the global probability is the product of $\mathbf{P}$ and $\mathbf{P}_n^*$, and the global expectation is the iterated expectation $\mathbf{E}_n = \mathbf{E}\mathbf{E}_n^*$ arises because $\mathbf{P}_n^*$ and $\mathbf{E}_n^*$ represent *conditional* probability and expectation respectively.

THEOREM 1. *Suppose Assumptions 3.1–3.6 hold. Then, for each n and x,*

$$\mathbf{E}_n\left(\left[\bar{J}_{n,K}(x) - \mathbb{E}\hat{J}_n\left(x, P_n^*\right)\right]^2\right) \to 0 \quad \text{as } K \to \infty. \tag{4}$$

Next, we impose a condition describing the behavior of the bootstrap distribution $\hat{J}_n(x, P_n^*)$ as n grows.

Assumption 3.7 (Bootstrap mean convergence)**.** There exists a cumulative distribution function $\breve{J}(\cdot, P)$ such that for all continuity points x of $\breve{J}(\cdot, P)$

$$\mathbb{E}\hat{J}_n\left(x, P_n^*\right) \longrightarrow \breve{J}(x, P) \quad \text{as } n \to \infty. \tag{5}$$

COROLLARY 1. *Suppose Assumptions 3.1–3.7 hold. Then, for each continuity point x of $\breve{J}(\cdot, P)$,*

$$\mathbf{E}_n\left(\left[\bar{J}_{n,K}(x) - \breve{J}(x, P)\right]^2\right) \to 0 \quad \text{as } \min(n, K) \to \infty. \tag{6}$$

Let $J_n(x, P) \equiv P\{R_n(X^n, \theta_P) \leq x\}$ denote the distribution of the root $R_n(X^n, \theta_P)$. The next assumption formally describes the limiting behavior of this distribution.

Assumption 3.8 (Root weak convergence)**.** There exists $J(\cdot, P)$ such that for all continuity points x of $J(\cdot, P)$

$$J_n(x, P) \to J(x, P) \quad \text{as } n \to \infty. \tag{7}$$

Let $q(\alpha, P) \equiv \inf\{x : J(x, P) \geq \alpha\}$, $\breve{q}(\alpha, P) \equiv \inf\{x : \breve{J}(x, P) \geq \alpha\}$, and $\hat{q}_{n,K}(\alpha) \equiv \inf\{x : \bar{J}_{n,K}(x) \geq \alpha\}$. Also, for $k \in \{1, \ldots, K\}$, define the *one-sided* confidence intervals[1]

$$\widehat{CI}_{n,K,k}(\alpha) \equiv \left\{\theta : R_n\left(X_k^n, \theta\right) \leq \hat{q}_{n,K}(\alpha)\right\}.$$

https://doi.org/10.1017/S0266466612000655 Published online by Cambridge University Press

This interval has nominal coverage α. Part (ii) of our next result describes the limiting behavior of its actual coverage. Part (iii) provides the probability limit of the Warp-Speed Bootstrap coverage estimator,

$$\alpha_{n,K}^{WS} \equiv K^{-1} \sum_{k=1}^{K} \mathbf{1}\left\{R_n\left(X_k^n, \theta_P\right) \leq \hat{q}_{n,K}(\alpha)\right\}.$$

COROLLARY 2. *Suppose Assumptions 3.1–3.7 hold and that* $\breve{J}(x, P)$ *is continuous and strictly increasing at* $x = \breve{q}(\alpha, P)$.

(i) Then for any $\delta > 0$,

$$\mathbf{P}_n\left\{|\hat{q}_{n,K}(\alpha) - \breve{q}(\alpha, P)| > \delta\right\} \to 0 \quad \text{as } \min(n, K) \to \infty. \tag{8}$$

(ii) If Assumption 3.8 also holds and $J(x, P)$ *is continuous at* $x = \breve{q}(\alpha, P)$, *then uniformly in k, for* $k = 1, \ldots, K$,

$$\mathbf{P}_n\left\{R_n\left(X_k^n, \theta_P\right) \leq \hat{q}_{n,K}(\alpha)\right\} \longrightarrow J\left(\breve{q}\left(\alpha, P\right), P\right) \quad \text{as } \min(n, K) \to \infty, \tag{9}$$

and thus also

$$K^{-1} \sum_{k=1}^{K} \mathbf{P}_n\left\{R_n\left(X_k^n, \theta_P\right) \leq \hat{q}_{n,K}(\alpha)\right\} \longrightarrow J\left(\breve{q}(\alpha, P), P\right) \quad \text{as } \min(n, K) \to \infty. \tag{10}$$

(iii) Under the assumptions of parts (i) and (ii) above, for any $\epsilon > 0$, *we further have*

$$\mathbf{P}_n\left\{\left|\alpha_{n,K}^{WS} - J\left(\breve{q}(\alpha, P), P\right)\right| > \epsilon\right\} \to 0 \quad \text{as } \min(n, K) \to \infty. \tag{11}$$

So far, we have not assumed that the bootstrap works. Indeed, Corollary 2 justifies use of the Warp-Speed method for assessing bootstrap performance based, for example, on the achieved coverage of confidence intervals.

For the bootstrap to actually work, we need Assumption 3.8, coupled with the following additional assumption.

Assumption 3.9 (Bootstrap consistency)**.** For all continuity points x of $J(\cdot, P)$ we have

$$\hat{J}_n\left(x, P_n^*\right) \xrightarrow{P} J(x, P) \quad \text{as } n \to \infty. \tag{12}$$

Assumption 3.9 is much stronger than Assumption 3.7. To see this, note that since $\hat{J}_n(x, P_n^*)$ is a bounded sequence of random variables, then it is a uniformly integrable sequence, even when raised to any power. Therefore, Assumption 3.9 implies Assumption 3.7, as well as

$$\breve{J}(x, P) = J(x, P) \tag{13}$$

for all x-points of continuity of $J(x, P)$. Assumption 3.9 also implies

https://doi.org/10.1017/S0266466612000655 Published online by Cambridge University Press

$$\mathrm{Var}\left(\hat{J}_n\left(x, P_n^*\right)\right) \to 0 \quad \text{as } n \to \infty. \tag{14}$$

Of course, equation (13) implies $\breve{q}(\cdot, P) = q(\cdot, P)$. We thus have the following immediate corollary.

COROLLARY 3. *Suppose Assumptions 3.1–3.6, 3.8, and 3.9 hold. If $J(x, P)$ is continuous and strictly increasing at $x = q(\alpha, P)$, then for each $k = 1, \ldots, K$,*

$$\mathbf{P}_n\left\{R_n\left(X_k^n, \theta_P\right) \le \hat{q}_{n,K}(\alpha)\right\} \to \alpha \quad \textit{as } \min(n, K) \to \infty. \tag{15}$$

Furthermore, for any $\epsilon > 0$,

$$\mathbf{P}_n\left\{\left|\alpha_{n,K}^{WS} - \alpha\right| > \epsilon\right\} \to 0 \quad \textit{as } \min(n, K) \to \infty. \tag{16}$$

Hence, for any k, the coverage of interval $\widehat{CI}_{n,K,k}(\alpha)$ and the Warp-Speed bootstrap coverage estimator $\alpha_{n,K}^{WS}$ tend to the nominal value α as $\min(n, K) \to \infty$ *when the bootstrap works, but not necessarily otherwise.* In fact, if the bootstrap does not work, the coverage of $\widehat{CI}_{n,K,k}(\alpha)$ and the estimator $\alpha_{n,K}^{WS}$ tend to $J\left(\breve{q}(\alpha, P), P\right)$, where $\breve{q}(\alpha, P)$ is the large-sample α–quantile of the *expectation* of the bootstrap distribution.

Thus the Warp-Speed method is capable of calibrating the finite-sample coverage of bootstrap confidence intervals when the bootstrap is known to work. In addition, the Warp-Speed method is capable of diagnosing bootstrap failure when this is due to a problem in the expected value of the bootstrap distribution; see equation (5).

In contrast, a full-blown Monte Carlo experiment—in which a large number B of bootstrap resamples are drawn for each of the K samples—could also detect bootstrap failure due to "excess variance" of the bootstrap limit distribution, i.e., a situation where equation (14) is violated. One famous such example is the sample mean of heavy-tailed data where the limit of the bootstrap distribution is a random variable (with nonzero variance). These results thus provide direct insight into both the strengths and limitations of the Warp-Speed method.

4. GENERAL ROOTS WITH NON-IID DATA

Now suppose that the parameter of interest $\theta_P \equiv \boldsymbol{\theta}(P)$ takes values in a general space Θ and that the root of interest, which we now denote $\tilde{R}_n(X^n, \theta_P)$, takes values in a general space Γ with norm $||\cdot||$, specifically, a normed linear space. (The spaces Θ and Γ may be the same, but this is not necessary.) For example, let $\tilde{R}_n(X^n, \theta_P) = \hat{C}_n(X^n)^{-1/2}\tau_n(\hat{\theta}_n(X^n) - \theta_P)$, where $\Theta = \Gamma$ is a finitely dimensioned vector space, $\hat{\theta}_n(X^n)$ is a parametric estimator of θ_P, $\hat{C}_n(X^n)$ is a consistent estimator of the asymptotic covariance matrix of $\tau_n(\hat{\theta}_n(X^n) - \theta_P)$, and $||\, r\, || = (r'r)^{1/2}$. Alternatively, with $\tilde{R}_n(X^n, \theta_P) = \tau_n(\hat{\theta}_n(X^n) - \theta_P)$, $\Theta = \Gamma$ could be a space of continuous functions on the unit interval, with $\hat{\theta}_n(X^n)$ a nonparametric estimator of θ_P, and $||\, r\, || = [\int_0^1 r^2(z)dz]^{1/2}$.

https://doi.org/10.1017/S0266466612000655 Published online by Cambridge University Press

To handle this situation, we modify Assumptions 3.2–3.4 as follows.

Assumption 3.2′ (Parameter of interest)**.** Let Θ be a linear space and let $P \to \boldsymbol{\theta}(P)$ be a Θ-valued function.

Assumption 3.3′ (Estimator)**.** For $n = 1, 2, \ldots$, $\hat{\theta}_n : \mathbf{X}^n \to \Theta$ is a measurable mapping, where $\mathbf{X}^n \equiv \times_{i=1}^n \mathbf{X}$.

We apply the theory of the previous section to scalar-valued roots $R_n(X^n, \boldsymbol{\theta}(P))$ constructed according to

Assumption 3.4′ (Root)**.** Let $(\Gamma, ||\cdot||)$ be a normed linear space. For $n = 1, 2, \ldots$, let $\tilde{R}_n : \mathbf{X}^n \times \Theta \to \Gamma$ be a measurable function, and define

$$R_n\left(X^n, \theta_P\right) \equiv g\left(\tilde{R}_n\left(X^n, \theta_P\right)\right),$$

where $g : \Gamma \to \mathbb{R}$ is a nonnegative function, continuous with respect to $||\cdot||$, such that $g(r) = 0$ if and only if $||r|| = 0$.

Let all notation and definitions be as in the previous section with this choice of $R_n(X^n, \theta_P)$. Results analogous to those above now hold immediately.

THEOREM 2. *Suppose Assumptions 3.1, 3.2′, 3.3′, 3.4′, 3.5, and 3.6 hold. Then, for each n and x,*

$$\mathbf{E}_n\left(\left[\bar{J}_{n,K}(x) - \mathbb{E}\hat{J}_n\left(x, P_n^*\right)\right]^2\right) \to 0 \quad \text{as } K \to \infty.$$

We also immediately obtain

COROLLARY 4. *Suppose Assumptions 3.1, 3.2′, 3.4′, 3.5, 3.6, and 3.7 hold. Then, for each continuity point x of $J(\cdot, P)$,*

$$\mathbf{E}_n\left(\left[\bar{J}_{n,K}(x) - \breve{J}(x, P)\right]^2\right) \to 0 \quad \text{as } \min(n, K) \to \infty.$$

COROLLARY 5. *Suppose Assumptions 3.1, 3.2′, 3.3′, 3.4′, 3.5, 3.6, and 3.7 hold and that $\breve{J}(x, P)$ is continuous and strictly increasing at $x = \breve{q}(\alpha, P)$.*

(i) Then, for any $\delta > 0$,

$$\mathbf{P}_n\left\{\left|\hat{q}_{n,K}(\alpha) - \breve{q}(\alpha, P)\right| > \delta\right\} \to 0 \quad \text{as } \min(n, K) \to \infty.$$

(ii) If Assumption 3.8 also holds and $J(x, P)$ is continuous at $x = \breve{q}(\alpha, P)$, then uniformly in k for $k = 1, \ldots, K$,

$$\mathbf{P}_n\left\{R_n\left(X_k^n, \theta_P\right) \le \hat{q}_{n,K}(\alpha)\right\} \longrightarrow J\left(\breve{q}(\alpha, P), P\right) \quad \text{as } \min(n, K) \to \infty.$$

and thus also

$$K^{-1}\sum_{k=1}^{K}\mathbf{P}_n\left\{R_n\left(X_k^n, \theta_P\right) \le \hat{q}_{n,K}(\alpha)\right\} \longrightarrow J\left(\breve{q}(\alpha, P), P\right)$$

as $\min(n, K) \to \infty$.

https://doi.org/10.1017/S0266466612000655 Published online by Cambridge University Press

(iii) Under the assumptions of parts (i) and (ii) above, for any $\epsilon > 0$, we further have

$$\mathbf{P}_n\left\{\left|\alpha_{n,K}^{WS} - J\left(\breve{q}(\alpha, P), P\right)\right| > \epsilon\right\} \to 0 \quad as\ \min(n, K) \to \infty.$$

COROLLARY 6. *Suppose Assumptions 3.1, 3.2′, 3.3′, 3.4′, 3.5, 3.6, 3.8, and 3.9 hold. If $J(x, P)$ is continuous and strictly increasing at $x = q(\alpha, P)$, then for each $k = 1, \ldots, K$,*

$$\mathbf{P}_n\left\{R_n\left(X_k^n, \theta_P\right) \le \hat{q}_{n,K}(\alpha)\right\} \to \alpha \quad as\ \min(n, K) \to \infty.$$

Furthermore, for any $\epsilon > 0$,

$$\mathbf{P}_n\left\{\left|\alpha_{n,K}^{WS} - \alpha\right| > \epsilon\right\} \to 0 \quad as\ \min(n, K) \to \infty.$$

When $\tilde{R}_n(X^n, \theta_P)$ is real-valued such that $\tilde{R}_n(X^n, \theta_P) = \tau_n(\hat{\theta}_n(X^n) - \theta_P)$ or $\tilde{R}_n(X^n, \theta_P) = (\hat{\theta}_n(X^n) - \theta_P)/\hat{s}_n(X^n)$ and $g(r) = \| r \| = | r |$, then $\widehat{CI}_{n,K,k}(\alpha)$ is a *symmetric* confidence interval for θ_P, that is, an interval of the type $\hat{\theta}_n(X^n) \pm A_n$ for some A_n. When $\tilde{R}_n(X^n, \theta_P) = \hat{C}_n(X^n)^{-1/2}\tau_n(\hat{\theta}_n(X^n) - \theta_P)$ and $g(r) = r'r$, where $\hat{C}_n(X^n)$ is a consistent estimator of the asymptotic covariance matrix of $\tau_n(\hat{\theta}_n(X^n) - \theta_P)$, then $\widehat{CI}_{n,K,k}(\alpha)$ is a familiar elliptical confidence interval for θ_P.

5. APPLICATION: CHOOSING THE BLOCK LENGTH FOR THE BLOCK BOOTSTRAP

A useful application of the Warp-Speed method is that of block-length selection for the block bootstrap and for subsampling; see the useful summary in Politis et al. (1999, Ch. 9) as well as Politis and White (2004). In particular, Politis et al. (1999, Ch. 9.3.1) discuss a calibration method for determining the optimal block length when the object of interest is the construction of bootstrap or subsampling confidence intervals. The basic idea is to select a block size b such that the corresponding confidence interval has actual coverage equal to the nominal coverage $1 - \alpha$ in a sample of a given size. This is achieved by estimating the calibration function γ such that $\gamma(b) \approx 1 - \alpha$ through a nested bootstrap procedure similar to the calculation of the empirical coverage considered in this paper, the only difference being that a bootstrap distribution is now used as the pseudo-DGP.

Politis et al. (1999) suggest an algorithm (Alg. 9.3.2) to estimate the calibration function, where the empirical calibration is estimated for a number M of block sizes and the block size yielding the desired coverage is selected. In terms of computational costs and using the notation introduced above, this algorithm involves on the order of $B \cdot K \cdot M$ computations. The algorithm is stated below for easy reference.

https://doi.org/10.1017/S0266466612000655 Published online by Cambridge University Press

Algorithm 1 (Block-length selection via Algorithm 9.3.2 of Politis et al., 1999)**.**

1. Given a sample X^n of size n, calculate an estimate $\hat{\theta}_n$ of the parameter of interest.
2. Select an initial estimate of the block length b_0, as well as a sequence $(b_1, \ldots, b_M)$ of candidate block lengths.
3. Generate K pseudo-samples of size n, $X_1^n, \ldots, X_K^n$, using a block bootstrap resampling scheme[2] with (average) block length b_0 based on the initial sample X^n, and compute $\hat{\theta}_{n,k}$ (the analog of $\hat{\theta}_n$) for each pseudo-sample $X_k^n, k = 1, \ldots, K$.
4. For $m = 1, \ldots, M$, do the following two steps:
   - Apply the block bootstrap with block size b_m to pseudo-sample X_k^n to generate B pseudo-samples, resulting in a $(1-\alpha)100\%$ bootstrap confidence interval $CI^*_{n,B,k}(1-\alpha; b_m)$ for $\hat{\theta}_n$, $k = 1, \ldots, K$.
   - Compute the (pseudo-)empirical coverage corresponding to block size b_m as: $\gamma^*_{n,B,K}(b_m) = \#\{\hat{\theta}_n \in CI^*_{n,B,k}(1-\alpha; b_m)\}/K$.
5. Pick the block length b^* such that $b^* = \arg\min_{b_i} |\gamma^*_{n,B,K}(b_i) - (1-\alpha)|$.

As previously mentioned, the above algorithm involves about $B \cdot K \cdot M$ computations. We propose below a faster version of the calibration algorithm that uses the Warp-Speed approach, thus reducing the number of computations to the order of $K \cdot M$.

Algorithm 2 (Block-length selection via Warp-Speed calibration)**.**

1. Perform steps 1–3 of the above Algorithm 9.3.2 of Politis et al. (1999).
2. For $m = 1, \ldots, M$ and $k = 1, \ldots, K$: apply the block bootstrap with block size b_m to pseudo-sample X_k^n to create a single bootstrap resample $X_k^{*n}(b_m)$, and calculate the root $R_n(X_k^{*n}(b_m), \hat{\theta}_{n,k})$.
3. Invert the empirical distribution $\hat{J}_{n,K}(b_m)$ of the K values $R_n(X_k^{*n}(b_m), \hat{\theta}_{n,k})$, $k = 1, \ldots, K$, to calculate the bootstrap quantile $\hat{q}_{n,K}(\alpha; b_m)$.
4. For a confidence level $1-\alpha$, generate the sequence of K bootstrap confidence intervals

$$\widehat{CI}_{n,K,k}(1-\alpha; b_m) = \left\{\theta : \hat{q}_{n,K}(\alpha/2; b_m) \le R_n\left(X_k^n, \theta\right) \le \hat{q}_{n,K} \times (1-\alpha/2; b_m)\right\}, k = 1, \ldots, K. \quad \textbf{(17)}$$

5. Compute $\hat{\gamma}_{n,K}(b_m) = \#\{\hat{\theta}_n \in \widehat{CI}_{n,K,k}(1-\alpha; b_m)\}/K$.
6. Pick the block length $\hat{b}$ such that $\hat{b} = \arg\min_{b_i} |\hat{\gamma}_{n,K}(b_i) - (1-\alpha)|$.

Remark 1. The choice of the initial block length b_0 has only a second-order effect and can thus be considered relatively unimportant (especially if the stationary bootstrap is employed at that stage).

https://doi.org/10.1017/S0266466612000655 Published online by Cambridge University Press

Remark 2. The confidence interval (17) could be alternatively obtained by intersection of two separately calibrated one-sided intervals. Politis et al. (1999) suggest always adopting the separate calibration approach, to allow for possibly different amounts of calibration needed in the upper and lower tails when the distribution of $\hat{\theta}_n$ is asymmetric.

6. SIMULATION STUDY

6.1. Warp-Speed vs Standard Method

We consider a simple illustrative example where a Monte Carlo experiment is performed to examine the coverage properties of bootstrap confidence intervals for the population mean θ_o of an $N(0,1)$ random variable. Let $X^n \equiv (X_1, \ldots, X_n)$ be an IID sample of $N(0,1)$ variables and $X_i^{*n} \equiv (X_{1,i}^*, \ldots, X_{n,i}^*)$, $i = 1, \ldots, B$, denote IID bootstrap resamples, drawn with replacement from the sample population consisting of the observations $\{X_1, \ldots, X_n\}$. An equal-tailed 95% bootstrap confidence interval for θ_o is constructed as

$$CI_{n,B}^*(.95) = \left[\hat{\theta}_n - q_{n,B}^*(.975), \hat{\theta}_n - q_{n,B}^*(.025)\right], \tag{18}$$

where $q_{n,B}^*(\alpha)$ is the $\alpha-$quantile of the empirical distribution of the root $\hat{\theta}_{n,i}^* - \hat{\theta}_n$, $i = 1, \ldots, B$, and $\hat{\theta}_n$, $\hat{\theta}_{n,i}^*$, respectively, indicate the means of samples X^n and X_i^{*n}.

We compare and contrast the two alternative methods to perform Monte Carlo experiments described in Section 2 in the context of this simple example. In order to assess the empirical coverage of the bootstrap confidence interval (18), we can proceed in two alternative ways.

1. (Standard MC method) Draw K random samples of size n from $N(0,1)$. For each Monte Carlo sample, construct the following sequence of equal-tailed 95% bootstrap confidence intervals

$$CI_{n,B,k}^*(.95) = \left[\hat{\theta}_{n,k} - q_{n,B,k}^*(.975), \hat{\theta}_{n,k} - q_{n,B,k}^*(.025)\right], \qquad k = 1, \ldots, K,$$

where $q_{n,B,k}^*(\alpha)$ is the $\alpha-$quantile of the empirical distribution of the root $\hat{\theta}_{n,k,i}^* - \hat{\theta}_{n,k}$, $i = 1, \ldots, B$, with $\hat{\theta}_{n,k}$ and $\hat{\theta}_{n,k,i}^*$, respectively, indicating the means of the kth sample and of its ith resample. Let the indicator variable $A_{n,B,k}$ equal 1 if the kth confidence interval contains the true mean $\theta_o = 0$ and equal 0 otherwise. Then the empirical coverage is given by $1 - \alpha_{n,K,B} = K^{-1} \sum_{k=1}^{K} A_{n,B,k}$.

2. (Warp-Speed MC method) Draw a random sample of size n from $N(0,1)$, and from this, draw a single bootstrap resample. Repeat the procedure K times, with a different resample for each draw. Compute $\hat{\theta}_{n,k}$ and $\hat{\theta}_{n,k}^*$ as the means of the kth Monte Carlo sample and of its bootstrap resample, respectively. Construct a sequence of equal-tailed 95% confidence intervals as

https://doi.org/10.1017/S0266466612000655 Published online by Cambridge University Press

$$\widehat{CI}_{n,K,k}(.95) = \left[\hat{\theta}_{n,k} - \widehat{q}_{n,K}(.975), \hat{\theta}_{n,k} - \widehat{q}_{n,K}(.025)\right], \qquad k = 1, \ldots, K,$$

where $\widehat{q}_{n,K}(\alpha)$ is the $\alpha-$quantile of the empirical distribution of the roots $\hat{\theta}^*_{n,k} - \hat{\theta}_{n,k}$, $k = 1, \ldots, K$. Defining an indicator variable $A^{WS}_{n,K,k}$ equal to 1 if the kth confidence interval contains the true mean $\theta_o = 0$ and equal to 0 otherwise, we compute the empirical coverage as $1 - \alpha^{WS}_{n,K} = K^{-1} \sum_{k=1}^{K} A^{WS}_{n,K,k}$.

Our goal is to compare the two approaches in assessing the empirical coverage properties of the confidence interval (18). To do so, we replicate the two above Monte Carlo experiments 1,000 times, for sample sizes $n = 200$ (Figure 1) and $n = 1{,}000$ (Figure 2) for several combinations of K and B. For each method, we calculate the average coverage over the 1,000 replications. Figures 1 and 2 show the average empirical coverage over the 1,000 replications for $B = K$ ranging from 150 to 1,000, for the standard Monte Carlo-bootstrap and $B = 1$ for the Warp-Speed designs of the Monte Carlo experiment, respectively. As expected, in both cases the empirical coverage converges to the nominal coverage for increasing K.

The coverages for the standard and the Warp-Speed methods are of comparable sizes, only differing in the third decimal place. It appears that the Warp-Speed MC method yields estimates of the empirical coverage that converge to the nominal coverage slightly faster than the estimates of the coverage obtained from the standard MC method with $B = K$.

This example gives an idea of the computational savings that one can expect when using the Warp-Speed MC method. For example, Figure 1 shows that the estimate of the empirical coverage obtained by the Warp-Speed MC with $K =$ 300 replications is roughly equivalent to the estimate of the coverage yielded by

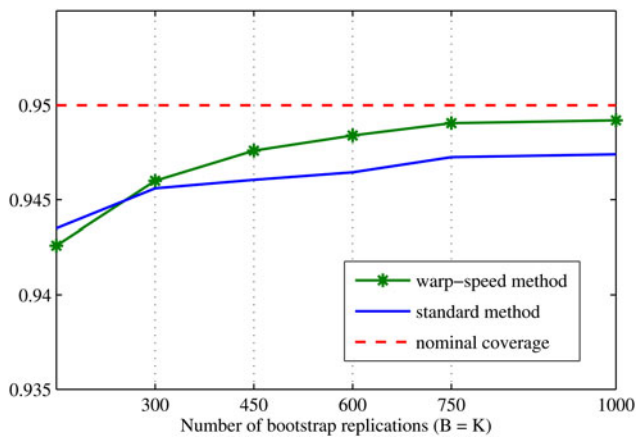


FIGURE 1. Empirical coverage of bootstrap confidence intervals. Sample size $n = 200$.

https://doi.org/10.1017/S0266466612000655 Published online by Cambridge University Press

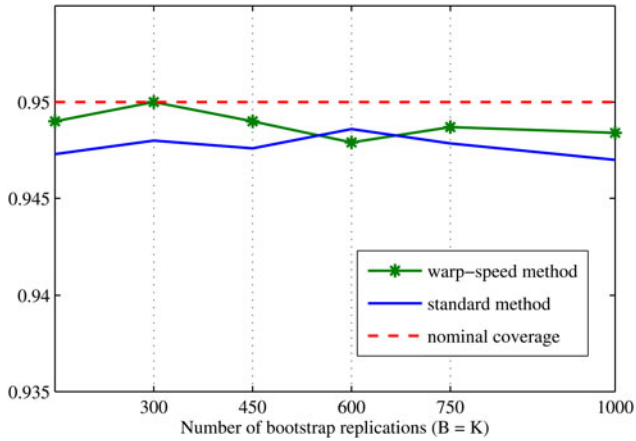


FIGURE 2. Empirical coverage of bootstrap confidence intervals. Sample size $n = 1{,}000$.

the standard MC method with $B = K = 300$. This means that the Warp-Speed method gives comparable results to the standard method, while computing the test statistics only 300, rather than 90, 000 times. Thus, an experiment that might take five hours of computing time for the standard MC approach takes one minute with the Warp-Speed approach.

Intuitively, the faster convergence of the Warp-Speed MC to the nominal coverage could be due to the fact that in the standard MC method the bootstrap confidence intervals are based on a bootstrap distribution that is *conditional* on the particular Monte Carlo draw from the DGP, whereas in the Warp-Speed MC method each bootstrap resample is drawn from a different empirical distribution. The accuracy of the confidence interval in the standard MC method is thus affected by how well the empirical distribution of the MC sample approximates the true distribution, while such reliance on the particular MC draw is dampened in the Warp-Speed MC method. In any case, the difference between the estimated coverage using the two methods should disappear as the sample size n increases, as the comparison of Figure 1 and Figure 2 seems to suggest.

6.2. Performance of Warp-Speed Calibration for Block-Length Selection

We conduct another Monte Carlo experiment to examine the performance of the Warp-Speed calibration algorithm. We construct block-bootstrap confidence intervals for the first autocorrelation of a time series choosing the block length by the Warp-Speed calibration algorithm. We also compare our results to those of Politis, Romano, and Wolf's (2004) calibration algorithm and its application to the block length selection for subsampling confidence intervals.

https://doi.org/10.1017/S0266466612000655 Published online by Cambridge University Press

Let $(\ldots, X_1, X_0, X_1, \ldots)$ be a sequence of zero-mean random variables, where either the series (X_t) itself or the first difference series $(X_t - X_{t-1})$ is assumed to be strictly stationary.[3] Let $\rho \equiv \lim_{t\to\infty} \mathbb{E}X_t X_{t+1}/\mathbb{E}X_t^2$ be the first autocorrelation of (X_t), and let $X^n \equiv (X_1, \ldots, X_n)$ be a time series. Given the ordinary least squares (OLS) estimator $\hat{\rho}_n \equiv \sum_{t=1}^{n-1} X_t X_{t+1} / \sum_{t=1}^{n-1} X_t^2$ and its standard error $\hat{\sigma}_n \equiv [(n-2)^{-1} \sum_{t=1}^{n-1} (X_{t+1} - \hat{\rho}_n X_t)^2 / \sum_{t=1}^{n-1} X_t^2]^{1/2}$, a symmetric $1-\alpha$ confidence interval for ρ is constructed as

$$CI_n^* = \left[\hat{\rho}_n - \hat{\sigma}_n \hat{q}_n(1-\alpha), \hat{\rho}_n + \hat{\sigma}_n \hat{q}_n(1-\alpha)\right], \tag{19}$$

where $\hat{q}_n(1-\alpha)$ is an estimator of the $1-\alpha$-quantile of the root $|\hat{\rho}_n - \rho|/\hat{\sigma}_n$. Politis et al. (2004) use subsampling to estimate this quantile and construct a confidence interval. Alternatively, we consider a block-bootstrap method based on the following algorithm.

Algorithm 3 (Residual block-bootstrap)**.**

1. For each sample X^n, fit an $AR(p)$ model and obtain the estimates $(\hat{\phi}_1, \ldots, \hat{\phi}_p)$ of AR coefficients and residuals $(\hat{\epsilon}_{p+1}, \ldots, \hat{\epsilon}_n)$, where the order p of lags is selected by the *BIC* criterion.
2. Block-bootstrap (or stationary bootstrap) the residuals to obtain $(\epsilon_{p+1}^*, \ldots, \epsilon_n^*)$.
3. Obtain a bootstrap resample $X^{*n} \equiv (X_1^*, \ldots, X_n^*)$ by setting $X_t^* = X_t$ for $t = 1, \ldots, p$ and $X_t^* = \hat{\phi}_1 X_{t-1}^* + \cdots + \hat{\phi}_p X_{t-p}^* + \epsilon_t^*$ for $t = p+1, \ldots, n$.

Let $X_i^{*n} \equiv (X_{1,i}^*, \ldots, X_{n,i}^*), i = 1, \ldots, B$, denote bootstrap resamples generated by the algorithm above. For each resample X_i^{*n}, let $\rho_{n,i}^*$ and $\sigma_{n,i}^*$ be similarly defined. A block-bootstrap confidence interval for ρ can be constructed by substituting the empirical $1-\alpha$-quantile of the root $|\rho_{n,i}^* - \hat{\rho}_n|/\sigma_{n,i}^*$ into (19). For both subsampling and block-bootstrap confidence intervals, the block length can be selected by calibration algorithms. Below, we summarize Politis et al.'s (2004) calibration algorithm and a version of the Warp-Speed calibration algorithm for this example.

1. (Politis et al.'s (2004) subsampling confidence interval with standard calibration algorithm) For each sample X^n, calculate $\hat{\rho}_n$. Set the initial block length b_0 and candidate block lengths $(b_1, \ldots, b_M)$. Generate K pseudo-samples of size n, $X_1^n, \ldots, X_K^n$, applying the stationary bootstrap (Algorithm 3) with average block length b_0. For each X_k^n, let $\hat{\rho}_{n,k}$ and $\hat{\sigma}_{n,k}$ be the OLS estimator and its standard error. For each subsample $(X_{k,l}, X_{k,l+1}, \ldots, X_{k,l+b_m-1})$ of X_k^n, let $\hat{\rho}_{b_m,l}$ and $\hat{\sigma}_{bm,l}$ be defined similarly. For $k = 1, \ldots, K$ and $m = 1, \ldots, M$, apply subsampling to pseudo-sample X_k^n, resulting in a subsampling confidence interval

$$\hat{CI}_{n,B,k}(b_m) = \left[\hat{\rho}_{n,k} - \hat{\sigma}_{n,k}\hat{q}_{n,B,k}(1-\alpha; b_m), \hat{\rho}_{n,k} + \hat{\sigma}_{n,k}\hat{q}_{n,B,k} \times (1-\alpha; b_m)\right],$$

https://doi.org/10.1017/S0266466612000655 Published online by Cambridge University Press

for $\hat{\rho}_n$, $k = 1, \ldots, K$, where $\hat{q}_{n,B,k}(1-\alpha; b_m) \equiv \inf\{x : L_{n,B,k}(x; b_m) \geq 1-\alpha\}$, $L_{n,B,k}(x; b_m) \equiv B^{-1} \sum_l 1\{|\hat{\rho}_{b_m,l} - \hat{\rho}_n| / \hat{\sigma}_{b_m,l} \leq x\}$, and $B \equiv n - b_m + 1$. Compute the (pseudo-)empirical coverage corresponding to block size b_m as $\hat{\gamma}_{n,B,K}(b_m) = \#\{\hat{\rho}_n \in \hat{CI}_{n,B,k}(b_m)\}/K$. Pick the block length $\hat{b}$ such that $\hat{b} = \arg\min_{b_m} |\hat{\gamma}_{n,B,K}(b_m) - (1-\alpha)|$. Use $\hat{b}$ to construct a subsampling confidence interval for ρ.

2. (Block-bootstrap confidence interval with Warp-Speed calibration algorithm) For each sample X^n, calculate $\hat{\rho}_n$. Set the initial block length b_0 and candidate block lengths $(b_1, \ldots, b_M)$. Generate K pseudo-samples of size n, $X_1^n, \ldots, X_K^n$, applying the stationary bootstrap (Algorithm 3) with average block length b_0. For $m = 1, \ldots, M$ and $k = 1, \ldots, K$, apply Algorithm 3 with block size b_m to pseudo-sample X_k^n to create a single bootstrap resample $X_k^{*n}(b_m)$, and calculate the root $R_n(X_k^{*n}(b_m), \hat{\rho}_{n,k}) = |\rho_{n,k}^* - \hat{\rho}_{n,k}| / \sigma_{n,k}^*$. Invert the empirical distribution $J_{n,K}^*(b_m)$ of the K values $R_n(X_k^{*n}(b_m), \hat{\rho}_{n,k})$, $k = 1, \ldots, K$, to calculate the bootstrap quantile $q_{n,K}^*(1-\alpha; b_m)$. Generate the sequence of K bootstrap confidence intervals

$$CI_{n,k}^*(b_m) = \left[\hat{\rho}_{n,k} - \hat{\sigma}_{n,k} q_{n,K}^*(1-\alpha; b_m), \hat{\rho}_{n,k} + \hat{\sigma}_{n,k} q_{n,K}^*(1-\alpha; b_m)\right].$$

Compute $\gamma_{n,K}^*(b_m) = \#\{\hat{\rho}_n \in CI_{n,k}^*(b_m)\}/K$. Pick the block length b^* such that $b^* = \arg\min_{b_m} |\gamma_{n,K}^*(b_m) - (1-\alpha)|$. Use b^* to construct a block bootstrap confidence interval for ρ.

Our goal here is to compare the two approaches in assessing the empirical coverage properties of the confidence interval in (19). To do so, we follow Politis et al. (2004) and generate samples 1,000 times from the following autoregressive moving average (ARMA)(1,1) model and compute the empirical coverage probabilities of the two confidence intervals:

$$X_t = \phi X_{t-1} + \epsilon_t + \theta \epsilon_{t-1}, \qquad t = 1, \ldots, n, \tag{20}$$

where $\epsilon_t = Z_t Z_{t-1}$, and $(Z_1, \ldots, Z_n)$ is an IID sample of $N(0,1)$ variables. The sample size is $n = 128$. The parameter values used in this experiment are $\phi = 1, 0.95$, and 0.8, and $\theta = 0.8, 0$, and -0.8 respectively. For both subsampling and block-bootstrap, we use the initial block length $b_0 = 10$. Politis et al. (2004) use the candidate block lengths $(b_1, \ldots, b_M) \equiv (5, 15, 25, 35)$. For the block-bootstrap, computational savings from the Warp-Speed calibration algorithm allow us to work with a finer grid for the candidate block lengths. Specifically, we use $(b_1, \ldots, b_M) \equiv (3, 4, \ldots, 49, 50)$.

Tables 1 and 2 report the coverage probabilities of the subsampling and the block-bootstrap confidence intervals, respectively.[4] In each table the estimated coverage probabilities of nominal 90% and 95% confidence intervals are reported for several fixed block lengths and for the data-dependent block length. As the tables indicate, the best block length depends on the true parameter value (ϕ, θ). Both of the calibration methods seem to give valid confidence intervals in many cases. Exceptions are the following. For both methods, coverage probabilities

https://doi.org/10.1017/S0266466612000655 Published online by Cambridge University Press

TABLE 1. Estimated coverage probabilities of subsampling confidence intervals

Target	$b = 5$	$b = 15$	$b = 25$	$b = 35$	$\hat{b}$
$\phi = 1, \theta = 0.8$					
0.90	0.99	0.94	0.90	0.86	0.96
0.95	1.00	0.98	0.93	0.91	0.98
$\phi = 1, \theta = 0$					
0.90	0.99	0.89	0.84	0.80	0.93
0.95	1.00	0.95	0.90	0.86	0.96
$\phi = 1, \theta = -0.8$					
0.90	0.19	0.09	0.11	0.11	0.20
0.95	0.48	0.19	0.17	0.15	0.45
$\phi = 0.95, \theta = 0.8$					
0.90	1.00	0.96	0.93	0.89	0.95
0.95	1.00	0.98	0.96	0.94	0.97
$\phi = 0.95, \theta = 0$					
0.90	0.98	0.89	0.85	0.81	0.90
0.95	1.00	0.95	0.90	0.87	0.96
$\phi = 0.95, \theta = -0.8$					
0.90	0.82	0.74	0.72	0.70	0.78
0.95	0.92	0.81	0.78	0.76	0.87
$\phi = 0.8, \theta = 0.8$					
0.90	0.99	0.93	0.89	0.84	0.90
0.95	1.00	0.97	0.94	0.91	0.94
$\phi = 0.8, \theta = 0$					
0.90	0.96	0.85	0.81	0.77	0.88
0.95	0.99	0.92	0.86	0.82	0.93
$\phi = 0.8, \theta = -0.8$					
0.90	0.87	0.80	0.77	0.73	0.82
0.95	0.96	0.87	0.83	0.79	0.90

are below the nominal level when $\theta = -0.8$, which is a well-known problematic case in the literature.[5] Especially when $(\phi, \theta) = (1, -0.8)$, the subsampling confidence intervals perform very poorly even with the data-dependent block length selection. The block-bootstrap outperforms subsampling in this case, although the coverage probabilities are still below the nominal level. On the other hand, the subsampling confidence intervals with the data-dependent block lengths have better coverage properties than their block-bootstrap counterparts when $(\phi, \theta) = (0.8, 0.8)$ and (0.8,0).

Overall, the coverage probabilities of block-bootstrap confidence intervals are more stable across different values of block length than those of subsampling confidence intervals. This might have limited the ability of the calibration algorithm to improve the coverage probability. In sum, the Warp-Speed calibration algorithm seems to perform as well as the standard algorithm. As it provides a way to optimize the block length over a fine grid, it would be especially useful

https://doi.org/10.1017/S0266466612000655 Published online by Cambridge University Press

TABLE 2. Estimated coverage probabilities of block bootstrap confidence intervals

Target	$b = 5$	$b = 15$	$b = 25$	$b = 35$	b^*
$\phi = 1, \theta = 0.8$					
0.90	0.96	0.95	0.94	0.94	0.93
0.95	0.98	0.97	0.97	0.96	0.95
$\phi = 1, \theta = 0$					
0.90	0.94	0.92	0.91	0.90	0.90
0.95	0.97	0.96	0.95	0.93	0.92
$\phi = 1, \theta = -0.8$					
0.90	0.54	0.53	0.54	0.54	0.55
0.95	0.57	0.58	0.56	0.56	0.55
$\phi = 0.95, \theta = 0.8$					
0.90	0.94	0.92	0.90	0.89	0.89
0.95	0.97	0.96	0.95	0.93	0.92
$\phi = 0.95, \theta = 0$					
0.90	0.93	0.91	0.89	0.86	0.86
0.95	0.96	0.96	0.94	0.92	0.92
$\phi = 0.95, \theta = -0.8$					
0.90	0.78	0.77	0.75	0.75	0.76
0.95	0.85	0.84	0.83	0.81	0.83
$\phi = 0.8, \theta = 0.8$					
0.90	0.90	0.89	0.85	0.83	0.83
0.95	0.96	0.93	0.92	0.89	0.90
$\phi = 0.8, \theta = 0$					
0.90	0.86	0.85	0.82	0.81	0.78
0.95	0.92	0.91	0.88	0.86	0.88
$\phi = 0.8, \theta = -0.8$					
0.90	0.82	0.79	0.79	0.77	0.76
0.95	0.89	0.87	0.85	0.83	0.84

when the coverage probability of confidence intervals may vary considerably across the block lengths.

NOTES

1. Equal-tailed two-sided confidence intervals can be constructed by intersections of one-sided intervals like $\widehat{CI}_{n,K,k}(\alpha)$. Analyzing the one-sided interval is thus more general. Symmetric two-sided confidence intervals are addressed specifically below.

2. At this stage, it is advisable to use the stationary bootstrap with expected block length b_0, as this is less sensitive to block size misspecification; see Politis and Romano (1994) and Politis and White (2004).

3. In this section, we use the subscript t instead of i to index observations.

4. In Table 1 we report the coverage probabilities of subsampling confidence intervals replicating Politis et al.'s (2004) Table II. The replicated coverage probabilities are very close to theirs, except for the case $(\phi, \theta) = (1, -0.8)$.

5. See Politis et al. (2004, Sec. 5) for details.

https://doi.org/10.1017/S0266466612000655 Published online by Cambridge University Press

REFERENCES

Canepa, A. (2006) Small sample corrections for linear restrictions on cointegrating vectors: A monte Carlo comparison. *Economics Letters* 91, 330–336.

Chen, X. & T.G. Conley (2001) A new semiparametric spatial model for panel time series. *Journal of Econometrics* 105, 59–83.

Davidson, R. & J. MacKinnon (2000) Improving the Reliability of Bootstrap Tests. Queen's University Working paper no. 995.

Davidson, R. & J. MacKinnon (2002) Fast double bootstrap tests of nonnested linear regression models. *Econometric Reviews* 21, 417–427.

Davidson, R. & J. MacKinnon (2007) Improving the reliability of bootstrap tests with the fast double bootstrap. *Computational Statistics and Data Analysis* 51, 3259–3281.

Efron, B. (1979) Bootstrap methods: Another look at the jackknife. *Annals of Statistics* 7, 1–26.

Efron, B. (2000) The bootstrap and modern statistics. *Journal of the American Statistical Association* 95, 1293–1296.

Escanciano, J.C. & C. Velasco (2006) Generalized spectral tests for the martingale difference hypothesis. *Journal of Econometrics* 134, 151–185.

Kilian, L. & P.L. Chang (2000) How accurate are confidence intervals for impulse responses in large VAR models? *Economics Letters* 69, 299–307.

Kim, J.H. (2001) Bootstrap-after-bootstrap prediction intervals for autoregressive models. *Journal of Business and Economic Statistics* 19, 117–128.

Künsch, H.R. (1989) The jackknife and the bootstrap for general stationary observations. *Annals of Statistics* 17, 1217–1241.

Lehmann, E.L. & J.P. Romano (2005) *Testing Statistical Hypotheses*, 3rd ed., Springer Verlag.

Li, F. & G. Tkacz (2006) A consistent bootstrap test for conditional density functions with time-series data. *Journal of Econometrics* 133, 863–886.

Loh, W.Y. (1988) Discussion of "Theoretical comparison of bootstrap confidence intervals," by P. Hall. *Annals of Statistics* 16, 972–976.

Loh, W.Y. (1991) Bootstrap calibration for confidence interval construction and selection. *Statistica Sinica* 1, 479–495.

Politis, D.N. & J.P. Romano (1994) The stationary bootstrap. *Journal of the American Statistical Association* 89, 1303–1313.

Politis, D.N., J.P. Romano, & M. Wolf (1999) *Subsampling*. Springer Verlag.

Politis, D.N., J.P. Romano, & M. Wolf (2004) Inference for autocorrelations in the possible presence of a unit root. *Journal of Time Series Analysis* 25, 251–263.

Politis, D.N. & H. White (2004) Automatic block-length selection for the dependent bootstrap. *Econometric Reviews* 23, 53–70. (With correction: A. Patton, D.N. Politis, & H. White (2009) Correction to "Automatic block-length selection for the dependent bootstrap" by D. Politis and H. White. *Econometric Reviews* 28, 372–375).

Shao, J. & D. Tu (1995) *The Jackknife and the Bootstrap*. Springer Verlag.

Whang, Y.J. (2000) Consistent bootstrap tests of parametric regression functions. *Journal of Econometrics* 98, 27–46.

White, H. (1998) A reality check for data snooping. QRDA Technical Report 8/98.

APPENDIX: Mathematical Proofs

Proof of Theorem 1. Fix an n and x throughout the proof. Then pick a k, and note that $\hat{J}_{n,k}(x)$ is a bounded random variable whose randomness (conditional on the sample X_k^n) is a result of the "bootstrap randomness" of X_k^{*n} appearing in the root $R_n(X_k^{*n}, \hat{\theta}_n(X_k^n))$. Thus, we can determine its conditional expectation, $\mathbb{E}^*_{n,k}\hat{J}_{n,k}(x)$, via the strong law of

https://doi.org/10.1017/S0266466612000655 Published online by Cambridge University Press

large numbers (SLLN), i.e., by considering the hypothetical experiment of drawing a large number of such roots, say $R^*_{n,k,1}, \ldots, R^*_{n,k,B}$, based on X^n_k.

These B roots are IID, with the same distribution as $R_n(X^{*n}_k, \hat{\theta}_n(X^n_k))$. So, by the SLLN,

$$\lim_{B\to\infty} B^{-1} \sum_{j=1}^{B} \mathbf{1}\left\{R^*_{n,k,j} \le x\right\} = \mathbb{E}^*_{n,k} \hat{J}_{n,k}(x), \tag{A.1}$$

almost surely $(-\mathcal{P}^*_{n,k})$. But the left-hand side (LHS) of equation (A.1) is precisely the bootstrap distribution $\hat{J}_n(x, P^*_{n,k})$. Consequently,

$$\begin{aligned}\mathbf{E}^*_n\left(\bar{J}_{n,K}(x)\right) &= K^{-1} \sum_{k=1}^{K} \mathbf{E}^*_n\left(\hat{J}_{n,k}(x)\right) = K^{-1} \sum_{k=1}^{K} \mathbb{E}^*_{n,k} \hat{J}_{n,k}(x) \\ &= K^{-1} \sum_{k=1}^{K} \hat{J}_n\left(x, P^*_{n,k}\right)\end{aligned} \tag{A.2}$$

for any finite K.

But the variables $\hat{J}_n(x, P^*_{n,k})$ for $k = 1, 2, \ldots$ (whose randomness is governed by $\mathbf{P}$) are also IID under Assumption 3.5, with finite expectation given by

$$\mathbf{E}\left(\hat{J}_n\left(x, P^*_{n,k}\right)\right) = \mathbb{E}\left(\hat{J}_n\left(x, P^*_n\right)\right). \tag{A.3}$$

By construction, the global expectation $\mathbf{E}_n$ is the iterated expectation $\mathbf{E}\mathbf{E}^*_n$, so that

$$\mathbf{E}_n\left(\bar{J}_{n,K}(x)\right) = \mathbf{E}\mathbf{E}^*_n\left(\bar{J}_{n,K}(x)\right) = \mathbb{E}\left(\hat{J}_n\left(x, P^*_n\right)\right). \tag{A.4}$$

Now, $\hat{J}_{n,k}(x)$ is a Bernoulli random variable; hence it has $\mathbf{Var}_n(\hat{J}_{n,k}(x)) \le 1/4$. Also note that $\hat{J}_{n,k}(x)$ is a function of (X^n_k, X^{*n}_k); since (X^n_k, X^{*n}_k) for $k = 1, 2, \ldots$ are IID, the same is true for $\hat{J}_{n,k}(x)$. Thus,

$$\begin{aligned}\mathbf{Var}_n\left(\bar{J}_{n,K}(x)\right) &= K^{-2}\ \mathbf{Var}_n\left(\sum_{k=1}^{K} \hat{J}_{n,k}(x)\right) \\ &= K^{-2} \sum_{k=1}^{K} \mathbf{Var}_n\left(\hat{J}_{n,k}(x)\right) \le 1/(4K).\end{aligned} \tag{A.5}$$

But equations (A.4) and (A.5) imply that

$$\mathbf{E}_n\left(\left[\bar{J}_{n,K}(x) - \mathbb{E}\hat{J}_n\left(x, P^*_n\right)\right]^2\right) \le 1/(4K), \tag{A.6}$$

and the result follows. ■

Proof of Corollary 1. Fix a continuity point x of $\breve{J}(\cdot, P)$. From equation (A.6), we have

$$\begin{aligned}1/(4K) &\ge \mathbf{E}_n\left(\left[\bar{J}_{n,K}(x) - \mathbb{E}\hat{J}_n\left(x, P^*_n\right)\right]^2\right) \\ &= \mathbf{E}_n\left(\left[\bar{J}_{n,K}(x) - \breve{J}(x, P) + \breve{J}(x, P) - \mathbb{E}\hat{J}_n\left(x, P^*_n\right)\right]^2\right)\end{aligned}$$

https://doi.org/10.1017/S0266466612000655 Published online by Cambridge University Press

$$= \mathbf{E}_n\left(\left[\bar{J}_{n,K}(x) - \breve{J}(x,P)\right]^2\right) + 2\left[\breve{J}(x,P) - \mathbb{E}\hat{J}_n\left(x, P_n^*\right)\right]$$
$$\times \mathbf{E}_n\left[\bar{J}_{n,K}(x) - \breve{J}(x,P)\right] + \left[\breve{J}(x,P) - \mathbb{E}\hat{J}_n\left(x, P_n^*\right)\right]^2,$$

using the fact that both $\breve{J}(x,P)$ and $\mathbb{E}(\hat{J}_n(x,P_n^*))$ are nonrandom.

Because of the boundedness of $\bar{J}_{n,K}(x)$ and Assumption 3.7, we have

$$\mathbf{E}_n\left(\left[\bar{J}_{n,K}(x) - \breve{J}(x,P)\right]^2\right) \leq 1/(4K) + C_n,$$

where $C_n \to 0$ as $n \to \infty$, and the result is proven. ■

Proof of Corollary 2.

(i) First note that Corollary 1 implies that, for any $\epsilon > 0$,

$$\mathbf{P}_n\left\{\left|\bar{J}_{n,K}(x) - \breve{J}(x,P)\right| > \epsilon\right\} \to 0 \quad \text{as } \min(n,K) \to \infty.$$

Since $\bar{J}_{n,K}(x)$ tends to $\breve{J}(x,P)$ in probability$-\mathbf{P}_n$, equation (8) now follows from arguments similar to Lemma 1.2.1 of Politis et al. (1999).

(ii) Note that, for any k, Assumption 3.8 implies

$$\mathbf{P}_n\left\{R_n\left(X_k^n, \theta_P\right) \leq \breve{q}(\alpha,P)\right\} \longrightarrow J\left(\breve{q}(\alpha,P),P\right) \quad \text{as } n \to \infty,$$

where the above convergence is uniform in k. Hence, equation (9) follows from equation (8) and Corollary 11.2.3 of Lehmann and Romano (2005). Equation (10) then follows directly.

(iii) Pick $\epsilon > 0$. By the triangle inequality

$$\left|K^{-1}\sum_{k=1}^{K}\mathbf{1}\left\{R_n\left(X_k^n,\theta_P\right) \leq \hat{q}_{n,K}(\alpha)\right\} - J\left(\breve{q}(\alpha,P),P\right)\right| \leq A_1 + A_2,$$

where

$$A_1 = \left|K^{-1}\sum_{k=1}^{K}\mathbf{1}\left\{R_n\left(X_k^n,\theta_P\right) \leq \hat{q}_{n,K}(\alpha)\right\} - \mathbf{E}_n\mathbf{1}\left\{R_n(X_k^n,\theta_P) \leq \hat{q}_{n,K}(\alpha)\right\}\right|$$

and

$$A_2 = \left|K^{-1}\sum_{k=1}^{K}\mathbf{E}_n\mathbf{1}\left\{R_n\left(X_k^n,\theta_P\right) \leq \hat{q}_{n,K}(\alpha)\right\} - J\left(\breve{q}(\alpha,P),P\right)\right|.$$

Now A_2 is deterministic, and $A_2 \to 0$ by equation (10). Therefore, to show equation (11), it will be sufficient to show that

$$\mathbf{P}_n\{A_1 > \epsilon\} \to 0 \quad \text{as } \min(n,K) \to \infty. \tag{A.7}$$

Now, take $\delta > 0$ small enough that $J(x,P)$ is continuous for all $x \in Q_\delta(\alpha)$ where $Q_\delta(\alpha) \equiv [\breve{q}(\alpha,P) - \delta, \breve{q}(\alpha,P) + \delta]$. Let $D_{n,K,\delta}(\alpha)$ denote the event $\{\hat{q}_{n,K}(\alpha) \in Q_\delta(\alpha)\}$; by equation (8), we have that

$$\mathbf{P}_n\left\{D_{n,K,\delta}(\alpha)\right\} \to 1 \quad \text{as } \min(n,K) \to \infty.$$

https://doi.org/10.1017/S0266466612000655 Published online by Cambridge University Press

Define the conditional probability $\tilde{\mathbf{P}}_n\{\cdot\} = \mathbf{P}_n\{\cdot \mid D_{n,K,\delta}(\alpha)\}$. In view of the above convergence, to show equation (11), it will be sufficient to show that

$$\tilde{\mathbf{P}}_n\{A_1 > \epsilon\} \to 0 \quad \text{as } \min(n, K) \to \infty. \tag{A.8}$$

Now, given the event $D_{n,K,\delta}(\alpha)$, we have that $A_1 \leq \sup_{q \in Q_\delta(\alpha)} |B_{n,K,q}|$, where

$$B_{n,K,q} = K^{-1} \sum_{k=1}^{K} \mathbf{1}\left\{R_n\left(X_k^n, \theta_P\right) \leq q\right\} - \mathbf{E}_n \mathbf{1}\left\{R_n\left(X_k^n, \theta_P\right) \leq q\right\}.$$

But, as above, the variance of $B_{n,K,q}$ is bounded above by $1/(4K)$; hence, by Chebychev's inequality, we have that for any $\epsilon^* > 0$,

$$\tilde{\mathbf{P}}_n\left\{\left|B_{n,K,q}\right| > \epsilon\right\} < \epsilon^*,$$

provided $\min(n, K)$ is large enough. Since this ϵ^* does not depend on q, it follows that for $\min(n, K)$ sufficiently large

$$\sup_{q \in Q_\delta(\alpha)} \tilde{\mathbf{P}}_n\left\{\left|B_{n,K,q}\right| > \epsilon\right\} < \epsilon^*. \tag{A.9}$$

Note, however, that

$$\sup_{q \in Q_\delta(\alpha)} \tilde{\mathbf{P}}_n\left\{\left|B_{n,K,q}\right| > \epsilon\right\} = \tilde{\mathbf{P}}_n\left\{\sup_{q \in Q_\delta(\alpha)} \left|B_{n,K,q}\right| > \epsilon\right\}. \tag{A.10}$$

From equations (A.9) and (A.10) it follows that

$$\tilde{\mathbf{P}}_n\left\{\sup_{q \in Q_\delta(\alpha)} \left|B_{n,K,q}\right| > \epsilon\right\} < \epsilon^*.$$

Since $A_1 \leq \sup_{q \in Q_\delta(\alpha)} |B_{n,K,q}|$, equation (A.8) follows, and part (iii) is proven. ■

Proof of Corollary 3. Just note that since $\breve{q}(\alpha, P) = q(\alpha, P)$, we have $J(\breve{q}(\alpha, P), P) = J(q(\alpha, P), P) = \alpha$, and the result follows from Corollary 2. ■

Proof of Theorem 2. Identical to that of Theorem 1, mutatis mutandis. ■

https://doi.org/10.1017/S0266466612000655 Published online by Cambridge University Press