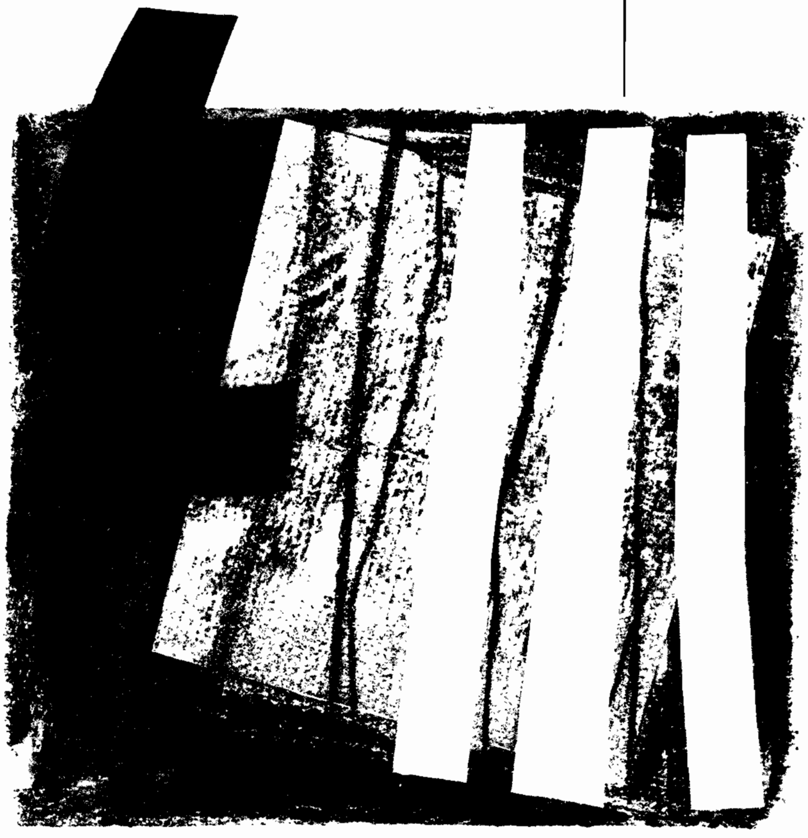
RESAMPLING TIME SERIES BY MISSING VALUES TECHNIQUES

Andrés M. Alonso, Daniel Peña and Juan Romo

00-42



WORKING PAPERS

Universidad Carlos III de Madrid

Working Paper 00-42
Statistics and Econometrics Series 19
July 2000

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-49

RESAMPLING TIME SERIES BY MISSING VALUES TECHNIQUES

Andrés M. Alonso, Daniel Peña and Juan Romo*

Abstract

For strongly dependent data, deleting blocks of observations is expected to produce bias as in the moving block jackknife of Künsch (1989) and Liu and Singh (1992). We propose an alternative technique which considers the blocks of deleted observations in the blockwise jackknife as missing data which are replaced by missing values estimates incorporating the observations dependence structure. Thus, the variance estimator is a weighted sample variance of the statistic evaluated in a "complete" series. We establish consistency for the variance and distribution of the sample mean. Also we extent this missing values approach to the blockwise bootstrap by assuming some missing observations among two consecutive blocks. We present the results of an extensive Monte Carlo study to evaluate the performance of the proposed methods in finite sample sizes in which it is shown that our proposal produces estimates of the variance of several time series statistics with smaller mean squared error than previous procedures.

Keywords: jackknife; bootstrap; missing values; time series.

*Dept. of Statistics and Econometrics, Univ. Carlos III de Madrid, Spain. Alonso, E-mail: alonsoam@est-econ.uc3m.es; Peña, e-mail: dpena@est-econ.uc3m.es; Romo: romo@est-econ.uc3m.es. AMS 1991 Subject Classifications. Primary: 62G09; Secondary: 62M10

1 Introduction

The classical jackknife and bootstrap, as proposed by Quenouille (1949), Tukey (1958) and Efron (1979), are not consistent in the case of dependent observations. During recent years these methods have been modified in order to account for the dependence structure of the data. The main existing procedures could be classified as model based and model free. Model based procedures fit a model to the data and resample the residuals which mimic the i.i.d. errors of the model (see, e.g., Freedman (1984), Efron and Tibshirani (1986), Bose (1990) and Kreiss and Franke (1992)). Model free procedures consider blocks of consecutive observations and resample from these blocks as in the independent case (see, e.g. Carlstein (1986), Künsch (1989) and Liu and Singh(1992)). Sherman (1998) compares these approaches in terms of efficiency and robustness and concludes that for moderate sample size the model based variance estimators provide a small gain under the correct model and, under mild misspecification, have bias similar to model free estimators while being more variable.

In this paper we are interested in the moving blocks jackknife (MBJ) and the moving blocks bootstrap (MBB) introduced in Künsch (1989) and independently in Liu and Singh (1992). These methods allow to estimate the variance of statistics defined by functionals of finite dimensional marginal distributions, which include robust estimators of location and scale, least-squares estimators of the parameters of an AR model and certain versions of the sample correlations.

As it is usual in the jackknife methods, the variance estimator is obtained by a weighted sample variance of the statistic evaluated in a sample where some observations (blocks of consecutive observations, in this case) are deleted or downweighted. Künsch (1989) showed that the MBJ that smoothes transition between observations left out and observations with full weight reduce the bias. Other resampling methods which also reduce the bias are: linear combinations of block bootstrap estimators with different block sizes proposed by Politis and Romano (1995), and the matched-block bootstrap of Carlstein *et al.* (1998) that suggest to use some block joining rule favoring blocks that are more likely to be close.

When the time series has a strong dependence structure, computing autocovariance by deleting blocks of observations is expected to produce bias. An alternative procedure is to assume that the block of observations is missing. For independent data, deleting observations is equivalent to assume that these observations are missing but for autocorrelated data, as shown in Peña (1990), both procedures are very different. Deleting a block of data means to substitute the observations in the block by their marginal expectation. Treating the block as missing values is equivalent to substitute the observations in the block by their conditional expectations given the rest of the data. This is the procedure we propose in this paper. In our case, the observations left out in the MBJ are considered as missing observations and they are substituted by a missing value estimate which takes into account the data dependence structure. Thus, the variance estimator is a weighted sample variance of the statistic evaluated in a "complete" series. This procedure could be interpreted as smooth transition between the two parts with full weight in the blockwise jackknife.

Also, we extend this idea to the blockwise bootstrap, defining a block of missing values between the blocks that form the bootstrap resample. Then, the procedure resemble to a block joining engine. In some sense, the matched-block bootstrap has a common point with the procedure that we propose in this paper, in particular with their autoregressive matching.

In Section 2 we define the MBJ with missing values techniques (M^{2}BJ) and the bootstrap (M^{2}BB). In Section 3 we present the missing values estimation procedures. In Section 4 the results about consistency of both methods as variance and distribution estimators for the sample mean are presented. Finally, the results of a simulation study comparing the MBJ and the M^{2}BJ, and the MBB and the M^{2}BB are presented in Section 5. All the proofs are given in an Appendix.

2 Resampling algorithms

2.1 Moving missing block jackknife

Let $X_1, \dots, X_N$ be observations from a stationary process. Let us suppose that the statistics T_N, whose variance or distribution we want to estimate, is defined by $T_N = T_N(\rho^N)$, where ρ^N is the empirical measure of $X_1, \dots, X_N$. As noted by Künsch (1989), it is impossible to estimate ρ^N without assuming some structure for the stationary processes. Then, in blockwise jackknife (and bootstrap) we suppose that T_N could be written as a functional of empirical m-dimensional distributions, i.e. $T_N = T(\rho_N^m)$, where $\rho_N^m = n^{-1} \sum_{t=1}^n \delta_{Y_t}$ is an empirical m-dimensional marginal measure, $n = N - m + 1$, $Y_t = (X_t, \dots, X_{t+m-1})$ are blocks of m consecutive observations and δ_y denotes the point mass at $y \in \mathbb{R}^m$.

The MBJ deletes or downweights blocks of m-tuples in the calculation of ρ_N^m:

$$\rho_N^{m,(j)} = (n - \|w_n\|_1)^{-1} \sum_{t=1}^{n} (1 - w_n(t-j))\, \delta_{Y_t}, \tag{1}$$

where $\|w_n\|_1 = \sum_{i=1}^{l} w_n(i)$ and $j = 0, 1, \dots, n-l$. The weights satisfy $0 \le w_n(i) \le 1$ for $i \in \mathbb{Z}$, and $w_n(i) > 0$ iff $1 \le i \le l$, and l is the length of the downweighted block. Note that $w_n(i) = 1$ for $1 \leq i \leq l$, corresponds to the deletion of blocks.

The MBJ variance estimator of T_N is defined as follows:

$$\widehat{\sigma}^2_{Jack} = (n - \|w_n\|_1)^2 n^{-1} (n - l + 1)^{-1} \|w_n\|_2^{-2} \sum_{j=0}^{n-l} \left(T_N^{(j)} - T_N^{(\cdot)} \right)^2, \qquad (2)$$

where $T_N^{(j)} = T_N(\rho_N^{m,(j)})$ is the j-th jackknife pseudo-value, $T_N^{(\cdot)} = (n - l + 1)^{-1} \sum_{j=0}^{n-l} T_N^{(j)}$ and $\|w_n\|_2^2 = \sum_{i=1}^{l} w_n(i)^2$.

In our approach we will use the following expression to calculate ρ_N^m:

$$\tilde{\rho}_N^{m,(j)} = n^{-1} \left(\sum_{t=1}^{n} (1 - w_n(t-j))\, \delta_{Y_t} + \sum_{t=1}^{n} w_n(t-j) \delta_{\widehat{Y}_{t,j}} \right), \qquad (3)$$

where $\widehat{Y}_{t,j}$ is an estimate of Y_t supposing that it is a missing value in the j-th sample, and then calculate $\widetilde{T}_N^{(j)} = T_N(\tilde{\rho}_N^{m,(j)})$, for $j = 0, 1, \ldots, n - l$. Note that in (3) instead of eliminating the blocks indexed by $j+1, \ldots, j+l$, we consider those $l + m - 1$ consecutive observations as missing in the time series sequence. The M^{2}BJ and variance estimator is defined by:

$$\widetilde{\sigma}^2_{Jack} = n(n - l + 1)^{-1} \|w_n\|_2^{-2} \sum_{j=0}^{n-l} \left(\widetilde{T}_N^{(j)} - \widetilde{T}_N^{(\cdot)} \right)^2. \qquad (4)$$

Also, we are interested in the distribution of T_N. We define the following jackknife-histograms, as in the subsampling method of Politis and Romano (1994):

$$H_N(x) = (n - l + 1)^{-1} \sum_{j=0}^{n-l} 1 \left\{ \tau_l l^{-1} (n - l)(T_N^{(j)} - T_N) \leq x \right\}, \qquad (5)$$

for the MBJ, and

$$\widetilde{H}_N(x) = (n - l + 1)^{-1} \sum_{j=0}^{n-l} 1 \left\{ \tau_l l^{-1} (n - l)(\widetilde{T}_N^{(j)} - T_N) \leq x \right\}, \qquad (6)$$

for the M^{2}BJ, where τ_l is an appropriate normalizing constant (typically $\tau_l = \sqrt{l}$), and $1\{E\}$ denotes the indicator of the event E.

2.2 Moving missing block bootstrap

In the case of bootstrap, we will use the circular block bootstrap (CBB) of Politis and Romano (1992) and Shao and Yu (1993) which could be described as follows. First, the sample is "extended" with $l-1$ observations:

$$X_{i,n} = \begin{cases} X_i & \text{if } i \in \{1,\ldots,n\} \\ X_{i-n} & \text{if } i \in \{n+1,\ldots,n+l-1\} \end{cases} \tag{7}$$

Second, define blocks of l consecutive observations $Y_{i,n} = (X_{i,n},\ldots,X_{i+l-1,n})$. Then $\{Y_{i,n}\}_{i=1}^n$ is used to obtain resamples $(Y_1^*,\ldots,Y_s^*)$ such that $\Pr^*\{Y_j^* = Y_{i,n}\} = 1/n$, and this implies that $\Pr^*\{X_j^* = X_i\} = 1/n$. Then, the bootstrap estimator is $T_N^* = T_N(\rho_N^*)$, where $\rho_N^* = n^{-1}\sum_{t=1}^n \delta_{Y_t^*}$. The bootstrap variance and distribution of T_N^*,

$$\mathrm{Var}^*\left(T_N^*\right) = \mathrm{E}^*\left[(T_N^* - \mathrm{E}^*\left[T_N^*\right])^2\right] \tag{8}$$

and

$$\Pr^*\left\{\tau_N(T_N^* - \mathrm{E}^*\left[T_N^*\right]) \le x\right\} \tag{9}$$

are used as variance and distribution estimators of T_N. In (9) the normalizing constant τ_N should take into account the number s of blocks in the bootstrap resample; usually s is selected such that $n = sl$.

Other blockwise bootstraps have been proposed, for instance, the moving blocks bootstrap (MBB) of Künsch (1989) and Liu and Singh (1992), the non-overlapping block bootstrap (NBB) based on Carlstein (1986), and the stationary bootstrap (SB) of Politis and Romano (1994). A recent paper of Lahiri (1999) compare these block bootstrap methods and concludes that CBB and MBB have better performance than NBB and SB in term of mean square error as variance and bias estimators.

The method that we propose could be described as follows: given a CBB resample $(Y_1^*,\ldots,Y_s^*)$, the idea of moving missing block bootstrap (M^2BB) is to introduce a block of k "observations" between two consecutive blocks, $\widetilde{Y}_j^*$. Then, the M^2BB resample is $(Y_1^*,\widetilde{Y}_1^*,Y_2^*\ldots,\widetilde{Y}_{s-1}^*,Y_s^*,\widetilde{Y}_s^*)$. For simplicity, we will use a fixed block size k for the blocks included and we will always introduce a final block in order to have ks missing observations. Another way of interpreting the M^2BB resample is to put $l+k$ as the block size in the CBB, and then to consider the last k observations in each block as missing values.

The M^2BB estimator is $\widetilde{T}_N^* = T_N(\widetilde{\rho}_N^*)$, where $\widetilde{\rho}_N^* = n^{-1}\sum_{t=1}^n \delta_{\widehat{Y}_t^*}$, and $\widehat{Y}_t^* = Y_t^*$ if $t \in \{1,\ldots,l,l+k+1,\ldots,2l+k,\ldots,(s-1)(l+k)+1,\ldots,(s-1)(l+k)+l\}$ and $\widehat{Y}_t^*$ is properly an estimate, otherwise. Then the bootstrap variance and distribution of $\widetilde{T}_N^*$,

$$\mathrm{Var}^*\left(\widetilde{T}_N^*\right) = \mathrm{E}^*\left[\left(\widetilde{T}_N^* - \mathrm{E}^*\left[\widetilde{T}_N^*\right]\right)^2\right] \tag{10}$$

and

$$\Pr^*\left\{\tau_N\left(T_N^* - \mathrm{E}^*\left[\widetilde{T}_N^*\right]\right) \le x\right\}, \tag{11}$$

are used as variance and distribution estimators of T_N. In (11), τ_N is a function of $s(l+k)$.

3 Missing values techniques

There are a number of alternatives to obtain $\widehat{Y}_t$ for stationary and invertible linear processes, see e.g. Harvey and Pierse (1984), Ljung (1989), Peña and Maravall (1991), and Beveridge (1992), and for some nonlinear processes as in Abraham and Thavaneswaran (1991). In this paper we will use the generalized least square method presented in Peña and Maravall (1991).

If $\{X_t\}_{t\in\mathbb{Z}}$ is a stationary process that admits an AR(∞) representation: $\Phi(B)(X_t-\mu)=a_t$, where $\Phi(B)=\sum_{j=0}^{\infty}\phi_j B^j$, B is the backshift operator and $\mathrm{E}[X_t]=\mu$, let's denote $z_t=X_t-\mu$, and assume that the finite series z_t has m missing values at times $T_1,T_2,\ldots,T_m$ with $T_i<T_j$. We fill the holes in the series with arbitrary numbers Z_{T_i} and construct an "observed" series Z_t by:

$$Z_t=\begin{cases} z_t+\omega_t & \text{if } t\in\{T_1,T_2,\ldots,T_m\} \\ z_t & \text{otherwise}\end{cases} \tag{12}$$

where ω_t is an unknown parameter. In matrix notation, we can write,

$$Z=z+H\omega, \tag{13}$$

where Z and z are the series expressed as a $N\times 1$ vector, H is $N\times m$ matrix such that $H_{T_i,i}=1$ and $H_{i,j}=0$ otherwise, and ω is a $m\times 1$ vector of unknown parameters. Let Σ be the $N\times N$ autocovariance matrix of the series z_t, then the generalized least squares estimator of ω is:

$$\widehat{\omega}=(H'\Sigma^{-1}H)^{-1}H'\Sigma^{-1}Z, \tag{14}$$

and the missing values estimates are obtained by:

$$\widehat{Z}=Z-H\widehat{\omega}=Z-H(H'\Sigma^{-1}H)^{-1}H'\Sigma^{-1}Z. \tag{15}$$

Note that (15) could be interpreted as a "nonparametric" estimator of the missing values, because does not require any distributional or model assumptions. However, the optimality properties (minimum mean square error and maximum likelihood) are established in the gaussian linear case.

When we apply this method to the j-th jackknife resample, we consider that observations $X_{j+1},\ldots,X_{j+l}$ are missing values, i.e., there are $m=l$ consecutive missing values and the matrix $H=H_j$ takes the form:

$$H_j=\begin{bmatrix} 0_{j\times l} \\ \mathrm{I}_{l\times l} \\ 0_{N-(l+j)\times l}\end{bmatrix}_{N\times l} \tag{16}$$

In the case of the bootstrap, we have $m=k\lceil n/(l+k)\rceil$ missing observations, where l is the length of the block in the bootstrap resample and k is the number of missing observations between two consecutive blocks. The matrix H is fixed and it has the following expression:

$$H = \begin{bmatrix} 0_{l\times l} & 0_{l\times k} & \cdots & 0_{l\times k} & 0_{l\times k} \\ 0_{k\times l} & I_{k\times k} & \cdots & 0_{k\times k} & 0_{l\times k} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0_{l\times l} & 0_{l\times k} & \cdots & 0_{l\times k} & 0_{l\times k} \\ 0_{k\times l} & 0_{k\times k} & \cdots & 0_{k\times k} & I_{k\times k} \end{bmatrix}. \tag{17}$$

Ljung (1989) suggests to use 0 in the missing values positions for Z, but in our case should be more convenient to use z_t (then $\omega_t = 0$); in such a case:

$$z - \widehat{Z} = H\left(H'\Sigma^{-1}H\right)^{-1} H'\Sigma^{-1}z, \tag{18}$$

and since $z = X - \mu$ and calling $\widehat{Z} = \widehat{X} - \mu$, we have

$$X - \widehat{X} = H\left(H'\Sigma^{-1}H\right)^{-1} H'\Sigma^{-1}(X - \mu), \tag{19}$$

which is a more tractable expression. For the bootstrap, the X in (19) is replaced by the X^* forming the bootstrap resample.

4 Consistency results

We now study the consistency of the proposed missing values approaches for jackknife and bootstrap for the sample mean. This case corresponds to $m = 1$, $T(F^1) = \int x\, dF^1(x) = \mathrm{E}[X_t] = \mu$. We will show that both procedures provide consistent estimators of the variance and the distribution of the sample mean. Theorem 1 and Proposition 2 present the fundamental results for the jackknife and Proposition 3 for the bootstrap. Also in Theorem 2 we establish the consistency of the MBJ of Künsch (1989) as distribution estimator of linear statistics.

Starting with the MBJ with missing values replacement we have that according to (3), the statistics evaluated in the j-th completed resample is:

$$\begin{aligned} \widetilde{T}_N^{(j)} &= n^{-1}\left(\textstyle\sum_{t=1}^n (1 - w_n(t-j))\, X_t + \sum_{t=1}^n w_n(t-j)\widehat{X}_{t,j}\right) \\ &= T_N - n^{-1}\textstyle\sum_{t=1}^n w_n(t-j)(X_t - \widehat{X}_{t,j}), \end{aligned} \tag{20}$$

where $T_N = n^{-1}\sum_{t=1}^n X_t$. First, we will consider the expression $\sum_{j=0}^{n-l}(\widetilde{T}_N^{(j)} - T_N)^2$. The use of T_N as a central measure seems more natural than $\widetilde{T}_N^{(\cdot)}$ because $T_N = T(F_N^1)$ (see Liu and Singh (1992)). We have that

$$n(\widetilde{T}_N^{(j)} - T_N) = -\sum_{t=1}^{n} w_n(t-j)(X_t - \widehat{X}_{t,j}) = -w'_{n,j}(X - \widehat{X}_j), \tag{21}$$

where $w_{n,j} = (w_n(1-j), \ldots, w_n(n-j))'$ and $\widehat{X}_j = \left(\widehat{X}_{1,j}, \ldots, \widehat{X}_{n,j}\right)'$, with

$$\widehat{X}_{t,j} = \begin{cases} X_t & \text{if} \quad w_n(t-j) = 0 \\ \widehat{X}_t & \text{if} \quad w_n(t-j) > 0. \end{cases} \tag{22}$$

In order to prove the consistency of jackknife variance estimator we will use the following proposition established in Berk (1974):

Proposition 1 *Suppose that $\{X_t\}_{t\in\mathbb{Z}}$ is a linear process such that $\sum_{i=0}^{\infty} \phi_i x_{t-i} = e_t$, where $\{e_t\}_{t\in\mathbb{Z}}$ are independent and identically distributed r.v.'s with $E[e_t] = 0$ and $E[e_t^2] = \sigma^2$, and $\phi_0 = 1$. Assume also that $\Phi(z) = \sum_{i=0}^{\infty} \phi_i z^i$ is bounded away from zero for $|z| \leq 1$. Then, there are constants F_1 and F_2, $0 < F_1 < F_2$, such that*

$$2\pi F_1 \leq \|\Sigma\|_{spec} \leq 2\pi F_2 \text{ and } (2\pi F_2)^{-1} \leq \|\Sigma^{-1}\|_{spec} \leq (2\pi F_1)^{-1}, \tag{23}$$

where $\|\Sigma\|_{spec} \equiv \max\left\{\sqrt{\lambda} : \lambda \text{ is eigenvalue of } \Sigma'\Sigma\right\}$ *denotes the spectral norm.*

Condition (23) allows us to establish the asymptotical unbiasness of $\widetilde{\sigma}^2_{Jack}$. We substitute in (4) $\widetilde{T}_N^{(\cdot)}$ by T_N and under standard assumptions we prove in Corollary 1 that the effect of this substitution is negligible.

Lemma 1 *If the conditions of Proposition 1, hold, and assuming that $w_n(i) = 1$ iff $1 \leq i \leq l$, $l = l(n) \to \infty$, and $\sum_{k=1}^{\infty} k|\gamma_k| < \infty$, then $E[n\widetilde{\sigma}^2_{Jack}] \to \sigma^2_\infty = \sum_{k=-\infty}^{+\infty} \gamma_k$.*

Now, we must prove that $\mathrm{Var}(n\widetilde{\sigma}^2_{Jack}) \to 0$. We have that

$$\mathrm{Var}(n\widetilde{\sigma}^2_{Jack}) = (n-l+1)^{-2}\|w_n\|_2^{-4} \sum_{j=0}^{n-l}\sum_{i=0}^{n-l} \mathrm{cov}(n^2(\widetilde{T}_N^{(j)} - T_N)^2, n^2(\widetilde{T}_N^{(i)} - T_N)^2.) \tag{24}$$

Note that $n^2(\widetilde{T}_N^{(j)} - T_N)^2 = \widetilde{w}'_{n,j}(X-\mu)(X-\mu)'\widetilde{w}_{n,j}$, where $\widetilde{w}_{n,j} = \Sigma^{-1}H_j (H_j'\Sigma^{-1}H_j)^{-1}H_j'w_{n,j}$, thus the only difference with Theorem 3.3 of Künsch (1989) is replacing $w_{n,j}$ by $\widetilde{w}_{n,j}$. A crucial aspect in his proof is the number of non zero elements ($l = o(n)$) in the vector $w_{n,j}$. The following lemma establishes that $\widetilde{w}_{n,j} = \bar{w}_{n,j} + o(l^{-1/2})$, where $\bar{w}_{n,j}$ has at most $l + 4\lceil l^{1/2}\rceil$ non zero elements.

Lemma 2 *Under the conditions of Lemma 1, and assuming that $\sum_{k=1}^{\infty} k^2|\gamma_k| < \infty$, we have $\widetilde{w}_{n,j} = \bar{w}_{n,j} + o(l^{-1/2})$.*

The next result follows by combining the previous lemmas 1 and 2 and Theorem 3.3 of Künsch (1989).

Theorem 1 *Under the conditions of Lemma 2, and assuming that* $\mathrm{E}[|X_t|^{6+\delta}] < \infty$, *with* $\delta > 0$, $\sum_{k=1}^{\infty} k^2 \alpha_k^{\delta/(6+\delta)} < \infty$ *where* α_k *are the strong mixing coefficients, and* $l = o(n)$, *it follows that* $n\widetilde{\sigma}^2_{Jack} \xrightarrow{P} \sigma^2_{\infty}$.

Corollary 1 *Under the conditions of Lemma 1, and assuming that* $l = o(n^{1/2})$, *we have that*

$$n\widetilde{\sigma}^2_{Jack} = n^2(n-l+1)^{-1}\|w_n\|_2^{-2}\sum_{j=0}^{n-l}\left(\widetilde{T}_N^{(j)} - T_N\right)^2 + o_P(l^{-1}).$$

The previous results assume that the matrix Σ is known; the next lemma shows that the consistency result obtained in Theorem 1 holds if we substitute Σ by an autoregressive estimator $\widehat{\Sigma}$, i.e., the $n \times n$ autocovariance matrix of an AR(p) process, with $p = p(n)$. We will use the matrix column-sum norm $\|A\|_{col} = \max\{\sum_{i=1}^{n}|a_{ij}| : j = 1,\ldots,n\}$, and the vector maximum norm $\|X\|_{\infty} = \max\{x_i : i = 1,\ldots,n\}$.

Lemma 3 *Under the conditions of Theorem 1, and assuming that* $\|\Sigma^{-1}\|_{col} < M < \infty$, $p = o((n/\log n)^{1/6})$, $\|X-\mu\|_{\infty} = O(\log^{1/2} n)$ *a.s., and* $l = o(n^{2/21-\varepsilon})$ *with* $\varepsilon > 0$, *it follows*

$$n\widetilde{\sigma}^2_{Jack} - n\widehat{\widetilde{\sigma}}^2_{Jack} = o(1) \ a.s., \tag{25}$$

where

$$\widehat{\widetilde{\sigma}}^2_{Jack} = n(n-l+1)^{-1}\|w_n\|_2^{-2}\sum_{j=0}^{n-l}\left(\widehat{\widetilde{T}}_N^{(j)} - T_N\right)^2, \tag{26}$$

and $\widehat{\widetilde{T}}_N^{(j)} = w_{n,j}H_j\left(H_j'\widehat{\Sigma}^{-1}H_j\right)^{-1}H_j'\widehat{\Sigma}^{-1}(X-\mu)$.

The condition $\|\Sigma^{-1}\|_{col} < M < \infty$ is satisfied by stationary and invertible ARMA(p,q) process. This is a direct consequence of the representation of Σ^{-1} in Galbraith and Galbraith (1974). For general processes the proof is still valid if $\|\Sigma^{-1}\|_{col} = O(l^{1/4-\alpha})$ for some α such that $0 < \alpha < 1/4$.

Now, we prove that the moving block jackknife (MBJ) of Künsch (1989) could be used as an estimator of the distribution of linear statistics. We will use the analogy between the subsampling of Politis and Romano (1994) and the blockwise jackknife. First, we introduce some notation: $S_{N,t} =$

$T_b(X_t, \ldots, X_{t+b-1})$ is the estimator of $T(\rho)$ based on the block or subsample $(X_t, \ldots, X_{t-b+1})$. Let $J_b(\rho)$ be the sampling distribution of

$$\tau_b(S_{N,1} - T(\rho)), \tag{27}$$

where τ_b is the normalizing constant. Also define the corresponding cumulative distribution function:

$$J_b(x, \rho) = \Pr_\rho \{\tau_b(S_{N,1} - T(\rho)) \leq x\}, \tag{28}$$

and denote $J_N(\rho)$ the sampling distribution of $\tau_n(T_N - T(\rho))$. The approximation of $J_N(\rho)$ proposed by subsampling is

$$L_N(x) = (N - b + 1)^{-1} \sum_{t=1}^{N-b+1} 1\{\tau_b(S_{N,t} - T_N) \leq x\}. \tag{29}$$

The only essential assumption in Politis and Romano's approach is that there exists a limiting law $J(\rho)$ such that $J_N(\rho)$ converges weakly to a limit law $J(\rho)$, as $n \to \infty$.

For simplicity, we only prove the consistency of MBJ for linear statistics, as the sample mean

$$T_N = n^{-1} \sum_{t=1}^{n} f(Y_t), \tag{30}$$

where $Y_t = (X_t, \ldots, X_{t+m-1})$, $n = N - m + 1$ and f is a continuous function on $\mathbb{R}^m$.

In the MBJ we have l deleted blocks $(Y_{j+1}, \ldots, Y_{j+l})$ which corresponds to $b = l + m - 1$ consecutive observations. Using (1), we have

$$\begin{aligned} T_N^{(j)} &= (n-l)^{-1} \textstyle\sum_{t=1}^{n} (1 - w_n(t-j)) f(Y_t) \\ &= (n-l)^{-1} n T_N - (n-l)^{-1} \textstyle\sum_{t=j+1}^{j+l} f(Y_t) \\ &= (n-l)^{-1} n T_N - (n-l)^{-1} S_{N,j+1}. \end{aligned} \tag{31}$$

Assuming without loss of generality that $m = 1$,

$$T_N^{(j)} - T_N = -l(n-l)^{-1}(S_{N,j+1} - T_N), \tag{32}$$

$$-\tau_l l^{-1}(n-l)(T_N^{(j)} - T_N) = \tau_l(S_{N,j+1} - T_N), \tag{33}$$

and

$$L_N(x) = (n - l + 1)^{-1} \sum_{j=0}^{n-l} 1\left\{\tau_l l^{-1}(n-l)(T_N - T_N^{(j)}) \leq x\right\}. \tag{34}$$

The MBJ analogous to $L_N(x)$ is

$$H_N(x) = (n-l+1)^{-1} \sum_{j=0}^{n-l} 1 \left\{ \tau_l l^{-1}(n-l)(T_N^{(j)} - T_N) \leq x \right\}. \quad (35)$$

Roughly speaking $H_N(x) = 1 - L_N(-x)$, in a sense that they could be different at most in a finite set of x, or we could change in (35) the $\leq$ by a $<$ and then the equality hold for all x.

We obtain the consistency under the following assumption:

Assumption 1 *There exists a symmetric limiting law $J(\rho)$ such that $J_n(\rho)$ converges weakly to a limit law $J(\rho)$, as $n \to \infty$.*

The following theorem shows that jackknife-histograms are consistent estimators of the distribution.

Theorem 2 *Assume Assumption 1 and that $\tau_l/\tau_n \to 0$, $l/n \to 0$ and $l \to \infty$ as $n \to \infty$. Also assume that the α-mixing sequence satisfies that $\alpha_X(k) \to 0$ as $k \to \infty$.*

1. *If x is a continuity point of $J(\cdot, \rho)$, then $H_N(x) \to J(x, \rho)$ in probability.*
2. *If $J(\cdot, \rho)$ is continuous, then $\sup_x |H_N(x) - J(x, \rho)| \to 0$ in probability.*

Also, we could use the M^2BJ method as distribution estimator. We establish the consistency for the sample mean. The MBJ and the M^2BJ statistics satisfy

$$T_N^{(j)} - T_N = -(n-l)^{-1} \sum_{t=1}^{n} w_n(t-j)(X_t - T_N), \quad (36)$$

and

$$\widetilde{T}_N^{(j)} - T_N = -n^{-1} \sum_{t=1}^{n} w_n(t-j)(X_t - \widehat{X}_{t,j}). \quad (37)$$

Therefore,

$$l^{-1/2} n(\widetilde{T}_N^{(j)} - T_N) = l^{-1/2}(n-l)(T_N^{(j)} - T_N) + l^{-1/2} \sum_{t=1}^{n} w_n(t-j)(\widehat{X}_{t,j} - T_N). \quad (38)$$

The following proposition establishes that the second term in the right hand side of (38) is $o_P(1)$.

Proposition 2 *Assume that $l/n \to 0$ and $l \to \infty$ as $n \to \infty$. Also assume that $w_n(i) = 1$ iff $1 \le i \le l$, $\sum_{k=1}^{\infty} k|\gamma_k| < \infty$, and $\|\Sigma^{-1}\|_{col} < M < \infty$. Then $l^{-1/2}\sum_{t=1}^{n} w_n(t-j)(\widehat{X}_{t,j} - T_N) = o_P(1)$. Furthermore, if $\|X-\mu\|_\infty = O_{a.s.}(\log^{1/2}(n))$, then it is $o_{a.s.}(1)$ uniformly in j.*

The consistency follows now from Theorem 2.1 in Politis and Romano (1994), Proposition 2 and the asymptotic equivalence lemma (cf. Rao (1973)).

Theorem 3 *Under the conditions in Proposition 2, for all x*

$$\widetilde{H}_N(x) = (n-l+1)^{-1}\sum_{j=0}^{n-l} 1\left\{\tau_l l^{-1}(n-l)(\widetilde{T}_N^{(j)} - T_N) \le x\right\} \to J(x,\rho) \quad (39)$$

in probability.

Remark 1 *In the proof of Lemma 3 we obtained that $n(\widetilde{T}_N^{(j)} - T_N) - n(\widehat{\widetilde{T}}_N^{(j)} - T_N)$ is $o(l^{3/2}(n/\log n)^{-1/3}\log^{1/2} n)$ a.s., thus (39) holds if we substitute Σ by $\widehat{\Sigma}$, and $l = o(n^{1/3})$.*

Now, we prove that the M^{2}BB provide us with consistent estimators of the variance and the distribution of the sample mean. We have the following CBB and M^{2}BB statistics:

$$\bar{X}^*_{n,s} = (sl)^{-1}\sum_{i=1}^{s}\sum_{j=1}^{l} X^*_{(i-1)l+j}, \quad (40)$$

and

$$\widetilde{\bar{X}}^*_{n,s} = (s(l+k))^{-1}\sum_{i=1}^{s}\left(\sum_{j=1}^{l} X^*_{(i-1)(l+k)+j} + \sum_{j=l+1}^{l+k} \widehat{X}^*_{(i-1)(l+k)+j}\right), \quad (41)$$

where $\widehat{X}^*_t$ is an estimate of the "missing observation" X^*_t, that takes into account the dependence structure on the original process $\{X_t\}$.

We could write the M^{2}BB analogous to $(sl)^{1/2}(\bar{X}^*_{n,s} - \bar{X}_n)$ as follows:

$$\begin{aligned}
&(s(l+k))^{1/2}(\widetilde{\bar{X}}^*_{n,s} - \bar{X}_n) \\
&= (s(l+k))^{-1/2}\Big(\textstyle\sum_{i=1}^{s}\sum_{j=1}^{l}(X^*_{(i-1)(l+k)+j} - \bar{X}_n) \\
&\qquad + \textstyle\sum_{i=1}^{s}\sum_{j=l+1}^{l+k}(\widehat{X}^*_{(i-1)(l+k)+j} - \bar{X}_n)\Big) \\
&= (l/(l+k))^{1/2}(sl)^{1/2}(\bar{X}^*_{n,s} - \bar{X}_n) \\
&\quad + (s(l+k))^{-1/2}\textstyle\sum_{i=1}^{s}\sum_{j=l+1}^{l+k}(\widehat{X}^*_{(i-1)(l+k)+j} - \bar{X}_n).
\end{aligned} \quad (42)$$

Notice that if $k/l \to 0$ as $n \to \infty$ we have that $l/(l+k) \to 1$ and then the first term in (42) satisfies the conditions in Theorem 1 in Politis and Romano (1992). The following proposition establishes that the second term in (42) is $o_P(1)$.

Proposition 3 *Assume that $l/n \to 0$, $l \to \infty$, and $k/l \to 0$ as $n \to \infty$. Also assume that $\sum_{k=1}^{\infty} k|\gamma_k| < \infty$, $\sum_{k=1}^{\infty} k^2 \alpha_k^{\delta/(6+\delta)} < \infty$, where α_k are the strong mixing coefficients, and $\|\Sigma^{-1}\|_{col} < M < +\infty$, then $(s(l+k))^{-1/2} \sum_{i=1}^{s} \sum_{j=l+1}^{l+k} (\widehat{X}^*_{(i-1)(l+k)+j} - \bar{X}_n) = o_P(1)$.*

Now, using the statement (1) of Theorem 1 in Politis and Romano (1992), Proposition 3 and the Cauchy-Schwarz inequality we have:

$$\operatorname{Var}^*\left((s(l+k))^{1/2}\left(\tilde{\bar{X}}^*_{n,s} - \bar{X}_n\right)\right) \xrightarrow{P} \sigma^2_\infty, \tag{43}$$

and from statement (2) of Theorem 1 in Politis and Romano (1992), Proposition 3 and the asymptotic equivalence lemma, we obtain the consistency results.

Theorem 4 *Under the conditions in Proposition 3, for all x*

$$\Pr^*\left\{(s(l+k))^{1/2}(\tilde{\bar{X}}^*_{n,s} - \bar{X}_n) \le x | X_1, \ldots, X_n\right\} - \Pr\left\{n^{1/2}(\bar{X}_n - \mu) \le x\right\} \to 0, \tag{44}$$

for almost all sample sequences $X_1, \ldots, X_N$.

5 Simulations

In this section, we compare the performance of the MBJ and MBJ with missing values replacement (M^2BJ), and the MBB and MBB with missing values replacement (M^2BB). We consider the following autoregressive models $X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + \epsilon_t$:

- (M1) AR(1) $\phi_1 = 0.8$, ϵ_t i.i.d. $\mathcal{N}(0,1)$.
- (M2) AR(2) $\phi_1 = 1.372$, $\phi_2 = -0.677$, ϵ_t i.i.d. $\mathcal{N}(0, 0.4982)$.
- (M3) AR(5) $\phi_1 = 0.9$, $\phi_2 = -0.4$, $\phi_3 = 0.3$, $\phi_4 = -0.5$, $\phi_5 = 0.3$, ϵ_t i.i.d. $\mathcal{N}(0,1)$.
- (M4) AR(1) $\phi_1 = -0.8$, ϵ_t i.i.d. $\mathcal{N}(0,1)$.

Models M1-M3 are the same as in Bühlmann (1994) and Bühlmann and Künsch (1994). In all of them the largest root is around 0.8. M4 is included because it presents a considerable amount of repulsion, and Carlstein *et al.* (1998) show that this feature is contrary to matching block bootstrap. The models M2-M4 exhibit a "damped-periodic" autocorrelation function, where the correlations can be negative. In M1 all the autocorrelations are positive. We also consider the following "dual" moving average models:

- (M5) MA(1) $\theta_1 = -0.8$, ϵ_t i.i.d. $\mathcal{N}(0,1)$.
- (M6) MA(2) $\theta_1 = -1.372$, $\theta_2 = 0.677$, ϵ_t i.i.d. $\mathcal{N}(0, 0.4909)$.
- (M7) MA(5) $\theta_1 = -0.9$, $\theta_2 = 0.4$, $\theta_3 = -0.3$, $\theta_4 = 0.5$, $\theta_5 = -0.3$, ϵ_t i.i.d. $\mathcal{N}(0,1)$.
- (M8) MA(1) $\theta_1 = 0.8$, ϵ_t i.i.d. $\mathcal{N}(0,1)$.

For M^2BJ and M^2BB, we use an autoregressive estimator for the autocovariance matrix Σ, choosing the order p of the approximating autoregressive process by minimizing the BIC (cf. Schwarz (1978)) in a range $0 \leq p \leq 10 \log_{10} n$. As in Bühlmann and Künsch (1994) we choose the sample size $N = 480$ and $N = 120$. Our results are based on 1000 simulations, and block size range from $l = 1$ to $l = 95$ for $N = 480$, and from $l = 1$ to $l = 30$ for $N = 120$. The statistics T_N included in the simulation study are the sample mean, median, variance, and autocovariance of order 1 and 5. Notice that in the case of h-th autocovariance, a block size l corresponds to l blocks of size h in MBJ, and $l + h - 1$ missing observations in M^2BJ. We report the estimates for the variance of these statistics and, as recommended in Carlstein *et al.* (1998), we measure the accuracy using the mean square error (MSE) of the logarithm of the variance. The simulations have been done as follows.

First, for each model Mi ($i = 1, \ldots, 8$) $N_T = 1000$ replications have been generated. In each replication the value of the statistic T_N is computed and the "true" value of the variance of this statistic is calculated by

$$\sigma_N^2 = N \frac{\sum_1^{N_T} (T_N^{(i)} - \overline{T}_N)^2}{N_T}$$

where $\overline{T}_N = \sum_1^{N_T} T_N^{(i)} / N_T$. The log of this value, $\log \sigma_N^2$, is reported in all the tables for each model and sample size, N.

Second, in the jackknife simulations (Tables 1 to 5) an estimate of the variance is computed by the following steps: (1) For each model Mi ($i = 1, \ldots, 8$) generate a sample of size N; (2) Select the length l, build the $N - l + 1$ jackknifed series, and compute in each series the value of the statistic T_N; (3) Compute the estimated jackknife variance by (2) and (4) and call them

$\widehat{\sigma}_N^2$ and $\widetilde{\sigma}_N^2$ respectively; and, (4) repeat the steps (1) to (3) 1000 times for each possible value l. The statistics given are E_1, SD_1, the average and standard deviation of the statistic $\log \widehat{\sigma}_N^2$ in the 1000 replications, and $MSE_1 = (\log \sigma_N^2 - E(\log \widehat{\sigma}_N^2))^2 + SD(\log \widehat{\sigma}_N^2)^2$, the mean squared error. The value of l given in L_1 is the block size producing the minimum MSE. The values E_2, SD_2, MSE_2, L_2 have the same interpretation and are computed for the proposed method based on $\widetilde{\sigma}_N^2$. The results with the relative mean square error $RMSE = MSE(\widehat{\sigma}_N^2)/\sigma_N^4$ are similar and therefore are omitted from the tables.

Third, in the bootstrap simulations (Tables 6 to 10) the estimate of the variance of the statistic is computed as follows: (1) For each model Mi ($i = 1, \ldots, 8$) generate a sample of size N; (2) Choose the block length l (l and k in M^{2}BB) and build $B = 250$ bootstrap samples by randomly selecting blocks with replacement among the blocks of observations. Compute in each bootstrap sample the value of the statistic T_N; (3) Compute the estimated bootstrap variance by (8) and (10) and call them $\widehat{\sigma}_N^2$ and $\widetilde{\sigma}_N^2$, respectively; and (4) repeat the steps (1) to (3) 1000 times. The values reported in the tables have the same interpretation as in the jackknife ones. The only difference is that for the method M^{2}BB in column corresponding to L_2, we report also the value of k, the optimal length of the missing value block (k takes values in $\{1, \ldots, 5\}$). Note that the MBB is equivalent to M^{2}BB with $k = 0$.

Due to the large number of simulations, we find a significant difference between the two methods in almost all the cases. However, we are interested in the big differences, e.g. $\text{MSE}(\widehat{\sigma}_N^2)/\text{MSE}(\widetilde{\sigma}_N^2) > 1.25$, i.e. at least a 25 percent of gain. Also, we could use a smaller number of simulations, as in Bühlmann and Künsch (1994) and Bühlmann (1997); in such a case, the results are similar to those of the previous approach.

Our main conclusions for jackknife methods are as follows: (a) In the cases where there exits a substantial difference between the two methods, the missing values replacement generally gives least MSE. In particular, only in the case of the sample mean and models M3 and M8, the MBJ have a better performance; (b) for the median, and models M1 and M5-M8, the M^{2}BJ outperforms the MBJ; (c) the methods are "equivalent" for the variance but, for the first autocovariance the proposed method presents a big difference in three cases; and, (d) for the autocovariance of order 5, which is the statistics that depend on the largest m-dimensional marginal distribution, in all cases and sample sizes the M^{2}BJ have a significant smaller MSE than MBJ. We can conclude that the proposed method works better in general than previous procedures and that the advantages are especially large for autocovariance, especially for lags greater than 2. Other simulation studies (not shown here) have confirmed this advantage of the proposed method in autocovariance for lags larger than 2. Regarding the optimal value of l, it is larger in MBJ than in M^{2}BJ.

In the comparison of bootstrap methods, we observe that: (a) In the cases where there exits a substantial difference between the two methods, the missing values replacement always gives least MSE; (b) for the mean, in almost all models, and for the median, in all models, the M^2BB outperforms the MBB; (c) the methods are "equivalent" in the variance and the autocovariance of order 5 (although the M^2BB outperforms the MBB when the sample size is large, 480) but for the first order autocovariance the M^2BB outperforms the MBB in all the cases and the differences are significantly larger for moving average models.

6 Conclusions

We have presented a generalization of the idea of using blocks for jackknife and bootstrapping estimation in time series. In the jackknife method we propose, instead of deleting observations, to assume that these observations are missing values. Note that for independent data both procedures are equivalent, but they are not for correlated data. It has been shown that with this procedure better results can be obtained in the model free estimation of the variance of the autocovariance of a stationary process. The advantages are especially important for larger lags. The consistency of the estimation of the variance and distribution of the sample mean has been established.

In the block bootstrap case we propose to assume that there are missing observations among two consecutive blocks. In these way the dependency structure among observations is better preserved and it has been shown that this procedure leads to better estimation in general than previous procedures specially for large sample size. The consistency of the estimation of the variance and distribution of the sample mean has been proved.

One additional advantage of this approach is that we are always dealing with complete series and, therefore, the usual routines for computing statistics in a time series can be applied to the jackknife or bootstrap samples generated with the missing value approach. In particular, previous bootstrap procedures can be seen as particular cases in which the length of the missing value block is equal to zero.

Appendix

Proof of Lemma 1: Using (19) and (21), we obtain:

$$n(\widetilde{T}_N^{(j)} - T_N) = -\, w'_{n,j} H_j \left(H_j'\Sigma^{-1}H_j\right)^{-1} H_j'\Sigma^{-1}(X-\mu) \qquad (45)$$

and

$$\mathrm{E}[n^2(\widetilde{T}_N^{(j)} - T_N)^2] = w'_{n,j} H_j \left(H_j'\Sigma^{-1}H_j\right)^{-1} H_j' w_{n,j}. \qquad (46)$$

Let $\alpha_j = \{j+1, j+2, \ldots, j+l\}$. Using the formula for the inverse of a partitioned matrix,

$$\left(H_j'\Sigma^{-1}H_j\right)^{-1} = (\Sigma^{-1}(\alpha_j))^{-1} = \Sigma(\alpha_j) - \Sigma(\alpha_j, \alpha_j')\Sigma(\alpha_j')^{-1}\Sigma(\alpha_j', \alpha_j), \quad (47)$$

where $\Sigma(\alpha_j)$ is the principal submatrix of Σ with the elements indexed by α_j, and $\Sigma(\alpha_j, \alpha_j')$ is the result of taking the rows indicated by α_j and deleting the columns indicated by α_j. $\Sigma(\alpha_j', \alpha_j)$ and $\Sigma(\alpha_j')$ are defined analogously, cf. Horn and Johnson (1990). Note that $\Sigma^{-1}(\alpha_j)$ is a submatrix of Σ^{-1}, while $\Sigma(\alpha_j)^{-1}$ is the inverse of a submatrix of Σ.

Using (4), and (46)-(47), we get:

$$\begin{aligned} \mathrm{E}[n\widetilde{\sigma}^2_{Jack}] &= (n-l+1)^{-1}l^{-1}\textstyle\sum_{j=0}^{n-l}\mathrm{E}[n^2(\widetilde{T}_N^{(j)} - T_N)^2] \\ &= l^{-1}w_n'\Sigma_{ll}w_n - (n-l+1)^{-1}l^{-1}\sum_{j=0}^{n-l} w_n'\Sigma(\alpha_j,\alpha_j')\Sigma(\alpha_j')^{-1}\Sigma(\alpha_j',\alpha_j)w_n, \end{aligned} \quad (48)$$

where $w_n = (w_n(1), \ldots, w_n(l))' = H_j'w_{n,j} = 1_{l\times 1}$, and $\Sigma(\alpha_j) = \Sigma_{ll}$ is the $l \times l$ autocovariance matrix.

Let's prove that $l^{-1}w_n'\Sigma_{ll}w_n \to \sigma^2_\infty$. We have that

$$\begin{aligned} l^{-1}w_n'\Sigma_{ll}w_n &= l^{-1}\left(l\gamma_0 + 2(l-1)\gamma_1 + \cdots + 2\gamma_{l-1}\right) \\ &= \textstyle\sum_{k=-l+1}^{l-1}\gamma_k - 2l^{-1}\sum_{k=1}^{l-1}k\gamma_k, \end{aligned} \quad (49)$$

which has limit σ^2_∞ using that $l(n) \to \infty$ and $\sum_{k=1}^{+\infty} k|\gamma_k| < \infty$.

Now we prove that the second term in (48) is bounded. First, note that

$$\begin{aligned} \|w_n'\Sigma(\alpha_j,\alpha_j')\|_2 &\le \|w_n'\Sigma(\alpha_j,\alpha_j')\|_1 \\ &= \textstyle\sum_{k=1}^{j}\left|\sum_{i=1}^{l}\gamma_{k+i-1}\right| + \sum_{k=1}^{n-l-j}\left|\sum_{i=1}^{l}\gamma_{k+i-1}\right| \\ &\le 2\textstyle\sum_{k=1}^{\max\{j,n-l-j\}}\left|\sum_{i=1}^{l}\gamma_{k+i-1}\right| \\ &\le 2\textstyle\sum_{k=1}^{n-l}\sum_{i=1}^{l}|\gamma_{k+i-1}| \\ &\le 2\textstyle\sum_{k=1}^{n}k|\gamma_k|. \end{aligned} \quad (50)$$

Second, we could write $\Sigma = \begin{bmatrix} A & B & C \\ B' & \Sigma(\alpha_j) & D \\ C' & D' & E \end{bmatrix}$, then $\Sigma(\alpha_j') = \begin{bmatrix} A & C \\ C' & E \end{bmatrix}$.

Let's define $\widetilde{\Sigma} = \begin{bmatrix} \Sigma(\alpha_j) & B' & D \\ B & A & C \\ D' & C' & E \end{bmatrix}$.

Note that $\widetilde{\Sigma}$ is also symmetric and $x'\Sigma x = \widetilde{x}'\widetilde{\Sigma}\widetilde{x}$, where $x' = (x_1, x_2, \ldots, x_n)$ and $\widetilde{x}' = (x_{j+1}, \ldots, x_{j+l}, x_1, \ldots, x_j, x_{j+l+1}, \ldots, x_n)$, then:

$$\lambda_{max}(\Sigma) = \max\left\{\frac{x'\Sigma x}{x'x} : x \neq 0\right\} = \max\left\{\frac{\widetilde{x}'\widetilde{\Sigma}\widetilde{x}}{\widetilde{x}'\widetilde{x}} : \widetilde{x} \neq 0\right\} = \lambda_{max}(\widetilde{\Sigma})$$

and the same is true for $\lambda_{min}(\Sigma)$. Then, we have:

$$\begin{array}{rcccccc} 2\pi F_1 & \leq & \lambda_{min}(\widetilde{\Sigma}) & \leq & \lambda_{max}(\widetilde{\Sigma}) & \leq & 2\pi F_2 \\ (2\pi F_2)^{-1} & \leq & \lambda_{min}(\widetilde{\Sigma}^{-1}) & \leq & \lambda_{max}(\widetilde{\Sigma}^{-1}) & \leq & (2\pi F_1)^{-1}. \end{array}$$

Since $\Sigma(\alpha_j')$ is a principal symmetric submatrix of $\widetilde{\Sigma}$, we have:

$$\begin{array}{rclcrcl} \lambda_{min}(\widetilde{\Sigma}) & \leq & \lambda_{min}(\Sigma(\alpha_j')) & \text{and} & \lambda_{max}(\Sigma(\alpha_j')) & \leq & \lambda_{max}(\widetilde{\Sigma}) \\ \lambda_{max}(\Sigma(\alpha_j')^{-1}) & \leq & \lambda_{max}(\widetilde{\Sigma}^{-1}) & \text{and} & \lambda_{min}(\widetilde{\Sigma}^{-1}) & \leq & \lambda_{min}(\Sigma(\alpha_j')^{-1}). \end{array}$$

Finally,

$$\begin{array}{rcl} w_n'\Sigma(\alpha_j,\alpha_j')\Sigma(\alpha_j')^{-1}\Sigma(\alpha_j',\alpha_j)w_n & \leq & \|\Sigma(\alpha_j')^{-1}\|_{spec}\|w_n'\Sigma(\alpha_j,\alpha_j')\|_2^2 \\ & \leq & (2\pi F_1)^{-1}\left(2\sum_{k=1}^{n} k|\gamma_k|\right)^2, \end{array}$$

and thus the second term in (48) goes to 0 as l goes to infinity.■

Proof of Lemma 2: Let $\Sigma^{-1} = \begin{bmatrix} A_1 & B_1 & C_1 \\ B_1' & \Sigma^{-1}(\alpha_j) & D_1 \\ C_1' & D_1' & E_1 \end{bmatrix}$; then, using (16),

$$Z_j = \Sigma^{-1}H_j\left(H_j'\Sigma^{-1}H_j\right)^{-1}H_j' = \begin{bmatrix} 0_{j\times j} & B_1(\Sigma^{-1}(\alpha_j))^{-1} & 0_{j\times N-l-j} \\ 0_{l\times j} & I_{l\times l} & 0_{l\times N-l-j} \\ 0_{N-l-j\times j} & D_1'(\Sigma^{-1}(\alpha_j))^{-1} & 0_{N-l-j\times N-l-j} \end{bmatrix}, \tag{51}$$

and

$$\widetilde{w}_{n,j} = \begin{bmatrix} B_1(\Sigma^{-1}(\alpha_j))^{-1}1_{l\times 1} \\ 1_{l\times 1} \\ D_1'(\Sigma^{-1}(\alpha_j))^{-1}1_{l\times 1} \end{bmatrix}. \tag{52}$$

The elements in positions $j+1,\ldots,j+l$ are 1's, and the remaining elements depend on the product $\Sigma^{-1}(\alpha_j',\alpha_j)(\Sigma^{-1}(\alpha_j))^{-1}$, because $\begin{bmatrix} B_1 \\ D_1' \end{bmatrix} = \Sigma^{-1}(\alpha_j',\alpha_j)$. Using the expressions for the inverse of a partitioned matrix, we obtain

$$\begin{array}{rcl} \Sigma^{-1}(\alpha_j',\alpha_j) & = & \left(\Sigma(\alpha_j',\alpha_j)\Sigma(\alpha_j)^{-1}\Sigma(\alpha_j,\alpha_j') - \Sigma(\alpha_j')\right)^{-1}\Sigma(\alpha_j',\alpha_j)\Sigma(\alpha_j)^{-1} \\ (\Sigma^{-1}(\alpha_j))^{-1} & = & \Sigma(\alpha_j) - \Sigma(\alpha_j,\alpha_j')\Sigma(\alpha_j')^{-1}\Sigma(\alpha_j',\alpha_j). \end{array}$$

Let's denote $Q_j = \left(\Sigma(\alpha_j',\alpha_j)\Sigma(\alpha_j)^{-1}\Sigma(\alpha_j,\alpha_j') - \Sigma(\alpha_j')\right)^{-1}$; then

$$\begin{array}{rcl} \Sigma^{-1}(\alpha_j',\alpha_j)(\Sigma^{-1}(\alpha_j))^{-1} & = & Q_j\Sigma(\alpha_j',\alpha_j) - (I + Q_j\Sigma(\alpha_j'))\Sigma(\alpha_j')^{-1}\Sigma(\alpha_j',\alpha_j) \\ & = & -\Sigma(\alpha_j')^{-1}\Sigma(\alpha_j',\alpha_j). \end{array} \tag{53}$$

Thus, we could concentrate our attention on $-\Sigma(\alpha_j')^{-1}\Sigma(\alpha_j',\alpha_j)$. We have that

$$\Sigma(\alpha_j',\alpha_j) = \begin{bmatrix} \gamma_j & \gamma_{j+1} & \cdots & \gamma_{j+l-1} \\ \gamma_{j-1} & \gamma_j & \cdots & \gamma_{j+l-2} \\ \vdots & \vdots & & \vdots \\ \gamma_1 & \gamma_2 & \cdots & \gamma_l \\ \gamma_l & \gamma_{l-1} & \cdots & \gamma_1 \\ \vdots & \vdots & & \vdots \\ \gamma_{n-2-j} & \gamma_{n-3-j} & \cdots & \gamma_{n-l-1-j} \\ \gamma_{n-1-j} & \gamma_{n-2-j} & \cdots & \gamma_{n-l-j} \end{bmatrix}_{n-l\times l} . \tag{54}$$

Let $\widetilde{\Sigma}(\alpha_j',\alpha_j)_{n-l\times l}$ be the matrix obtained writing 0 in each position of matrix $\Sigma(\alpha_j',\alpha_j)$ such that the index k of γ_k satisfies $k > \lceil l^{1/2}\rceil$. The difference with $\Sigma(\alpha_j',\alpha_j)$ satisfies

$$\begin{aligned} \|(\Sigma(\alpha_j',\alpha_j) - \widetilde{\Sigma}(\alpha_j',\alpha_j))w_n\|_2 &\leq \|\left(\Sigma(\alpha_j',\alpha_j) - \widetilde{\Sigma}(\alpha_j',\alpha_j)\right)w_n\|_1 \\ &\leq 2\textstyle\sum_{k=\lceil l^{1/2}\rceil+1}^{n-1} k|\gamma_k| = o(l^{-1/2}), \end{aligned} \tag{55}$$

since $\sum_{k=1}^{\infty} k^2|\gamma_k| < \infty$ implies $\sum_{k=\lceil l^{1/2}\rceil+1}^{n-1} k|\gamma_k| = o(l^{-1/2})$.

On the other hand, $\Sigma(\alpha_j') = \begin{bmatrix} \Sigma_{j-1,j-1} & F \\ F' & \Sigma_{n-l-j-1,n-l-j-1} \end{bmatrix}$, where $\Sigma_{h,h}$ is the $h\times h$ autocovariance matrix. Define $\widetilde{\Sigma}(\alpha_j') = \begin{bmatrix} \Sigma_{j-1,j-1} & 0 \\ 0' & \Sigma_{n-l-j-1,n-l-j-1} \end{bmatrix}$; as before, we have that

$$\begin{aligned} \|\Sigma(\alpha_j')^{-1} - \widetilde{\Sigma}(\alpha_j')^{-1}\|_{spec} &= \|\Sigma(\alpha_j')^{-1}(\Sigma(\alpha_j') - \widetilde{\Sigma}(\alpha_j'))\widetilde{\Sigma}(\alpha_j')^{-1}\|_{spec} \\ &\leq \|\Sigma(\alpha_j')^{-1}\|_{spec}\|\Sigma(\alpha_j') - \widetilde{\Sigma}(\alpha_j')\|_{spec}\|\widetilde{\Sigma}(\alpha_j')^{-1}\|_{spec} \\ &\leq (2\pi F_1)^{-2}(\|\Sigma(\alpha_j') - \widetilde{\Sigma}(\alpha_j')\|_{col}\|\Sigma(\alpha_j') - \widetilde{\Sigma}(\alpha_j')\|_{row})^{1/2} \\ &\leq (2\pi F_1)^{-2}\textstyle\sum_{k=l+1}^{n-1}|\gamma_k|, \end{aligned} \tag{56}$$

and $\sum_{k=1}^{\infty} k^2|\gamma_k| < \infty$ implies $\sum_{k=l+1}^{n-1}|\gamma_k| = o(l^{-2})$. Let $\Sigma_{h,h}^a = [\gamma^a]_{h\times h}$ be the $h\times h$ covariance matrix of an AR($\lceil l^{1/2}\rceil$) process such that $\Sigma^a_{\lceil l^{1/2}\rceil,\lceil l^{1/2}\rceil} = \Sigma_{\lceil l^{1/2}\rceil,\lceil l^{1/2}\rceil}$.

We can assume that $\sum_{k=1}^{\infty} k^2|\gamma_k^a| < \infty$, see Bühlmann (1995). Define $\bar{\Sigma}(\alpha_j') = \begin{bmatrix} \Sigma^a_{j-1,j-1} & 0 \\ 0' & \Sigma^a_{n-l-j-1,n-l-j-1} \end{bmatrix}$, then we have the following results:

$$\begin{aligned} \|\widetilde{\Sigma}(\alpha_j')^{-1} - \bar{\Sigma}(\alpha_j')^{-1}\|_{spec} &= \|\widetilde{\Sigma}(\alpha_j')^{-1}\left(\widetilde{\Sigma}(\alpha_j') - \bar{\Sigma}(\alpha_j')\right)\bar{\Sigma}(\alpha_j')^{-1}\|_{spec} \\ &\leq \|\widetilde{\Sigma}(\alpha_j')^{-1}\|_{spec}\|\widetilde{\Sigma}(\alpha_j') - \bar{\Sigma}(\alpha_j')\|_{spec}\|\bar{\Sigma}(\alpha_j')^{-1}\|_{spec} \\ &\leq 2(2\pi F_1)^{-2}\textstyle\sum_{k=\lceil l^{1/2}\rceil+1}^{n-1}(|\gamma_k| + |\gamma_k^a|), \end{aligned} \tag{57}$$

and $\sum_{k=1}^{\infty} k^2|\gamma_k| < \infty$ and $\sum_{k=1}^{\infty} k^2|\gamma_k^a| < \infty$ implies $\sum_{k=\lceil l^{1/2}\rceil+1}^{n-1} |\gamma_k| = o(l^{-1})$ and $\sum_{k=\lceil l^{1/2}\rceil+1}^{n-1} |\gamma_k^a| = o(l^{-1})$.

Note that $\bar{\Sigma}(\alpha_j')^{-1}$ is a $\lceil l^{1/2}\rceil$-diagonal matrix; then $\bar{\Sigma}(\alpha_j')^{-1}\widetilde{\Sigma}(\alpha_j', \alpha_j)w_n$ has at most $4\lceil l^{1/2}\rceil$ non zero elements. Define $\bar{w}_{n,j}$ replacing in $\widetilde{w}_{n,j}$ the matrices $\Sigma(\alpha_j')^{-1}$ and $\Sigma(\alpha_j', \alpha_j)$ with $\bar{\Sigma}(\alpha_j')^{-1}$ and $\widetilde{\Sigma}(\alpha_j', \alpha_j)$, then $\bar{w}_{n,j}$ has at most $l + 4\lceil l^{1/2}\rceil$ non zero elements.

Finally,

$$\begin{aligned}
\|\widetilde{w}_{n,j} - \bar{w}_{n,j}\|_2 &= \|\Sigma(\alpha_j')^{-1}\Sigma(\alpha_j', \alpha_j)w_n - \bar{\Sigma}(\alpha_j')^{-1}\widetilde{\Sigma}(\alpha_j', \alpha_j)w_n\|_2 \\
&\leq \|\Sigma(\alpha_j')^{-1}\Sigma(\alpha_j', \alpha_j)w_n - \bar{\Sigma}(\alpha_j')^{-1}\Sigma(\alpha_j', \alpha_j)w_n\|_2 \\
&\quad + \|\bar{\Sigma}(\alpha_j')^{-1}\Sigma(\alpha_j', \alpha_j)w_n - \bar{\Sigma}(\alpha_j')^{-1}\widetilde{\Sigma}(\alpha_j', \alpha_j)w_n\|_2 \\
&= \|(\Sigma(\alpha_j')^{-1} - \bar{\Sigma}(\alpha_j')^{-1})\Sigma(\alpha_j', \alpha_j)w_n\|_2 \\
&\quad + \|\bar{\Sigma}(\alpha_j')^{-1}(\Sigma(\alpha_j', \alpha_j) - \widetilde{\Sigma}(\alpha_j', \alpha_j))w_n\|_2 \\
&\leq \|\Sigma(\alpha_j')^{-1} - \bar{\Sigma}(\alpha_j')^{-1}\|_{spec}\|\Sigma(\alpha_j', \alpha_j)w_n\|_2 \\
&\quad + \|\bar{\Sigma}(\alpha_j')^{-1}\|_{spec}\|(\Sigma(\alpha_j', \alpha_j) - \widetilde{\Sigma}(\alpha_j', \alpha_j))w_n\|_2,
\end{aligned}$$

and using (55) - (57) we have that $\|\widetilde{w}_{n,j} - \bar{w}_{n,j}\|_2 = o(l^{-1/2})$.■

Proof of Corollary 1: We have

$$n\widetilde{\sigma}_{Jack}^2 = n^2\|w_n\|_2^{-2}\left((n-l+1)^{-1}\sum_{j=0}^{n-l}\left(\widetilde{T}_N^{(j)} - T_N\right)^2 - \left(\widetilde{T}_N^{(\cdot)} - T_N\right)^2\right), \quad (58)$$

and is enough to prove that $\widetilde{S}_N = n\left(\widetilde{T}_N^{(\cdot)} - T_N\right) = (n-l+1)^{-1}\sum_{j=0}^{n-l}\widetilde{w}_{n,j}(X - \mu)$ is $o_p(1)$. It's clear that $\mathrm{E}[\widetilde{S}_N] = 0$, and

$$\mathrm{E}[\widetilde{S}_N^2] = (n-l+1)^{-2}\sum_{j=0}^{n-l}\sum_{i=0}^{n-l}\widetilde{w}_{n,j}'\Sigma\widetilde{w}_{n,i}. \quad (59)$$

As in Lemma 2, we can concentrate our attention on

$$(n-l+1)^{-2}\sum_{j=0}^{n-l}\sum_{i=0}^{n-l}\bar{w}_{n,i}'\Sigma\bar{w}_{n,j}. \quad (60)$$

Put a 0 in each position of matrix Σ with γ_k such that $k > l$. $\widetilde{\Sigma}$ denote the resulting matrix. Since $\|\Sigma - \widetilde{\Sigma}\|_{spec} = o(l^{-1})$, then

$$(n-l+1)^{-2}\sum_{j=0}^{n-l}\sum_{i=0}^{n-l}\bar{w}_{n,j}'\Sigma\bar{w}_{n,i} = (n-l+1)^{-2}\sum_{j=0}^{n-l}\sum_{i=0}^{n-l}\bar{w}_{n,j}'\widetilde{\Sigma}\bar{w}_{n,i} + o(1). \quad (61)$$

On the other hand, note that $\bar{w}'_{n,j}\widetilde{\Sigma}\bar{w}_{n,i}$ is equal to a sum of 1, 2, ..., $l+4\lceil l^{1/2}\rceil$, non zero summands for the different values of i and j. Then

$$\sum_{i=0}^{n-l} |\bar{w}'_{n,j}\Sigma\bar{w}_{n,i}| \leq 2C\left(1+2+\ldots\left(l+4\lceil l^{1/2}\rceil\right)\right) = O(l^2), \tag{62}$$

and

$$(n-l+1)^{-2}\sum_{j=0}^{n-l}\sum_{i=0}^{n-l} \bar{w}'_{n,j}\widetilde{\Sigma}\bar{w}_{n,i} = O\left((n-l+1)^{-1}l^2\right), \tag{63}$$

where $C = \max\{1, (\pi F_1)^{-1}\sum_{k=1}^{\infty} k|\gamma_k|\}\sum_{k=-l}^{l}|\gamma_k|$. Finally, if $l = o(n^{1/2})$, we obtain that $\widetilde{S}_N \xrightarrow{P} 0$.■

Proof of Lemma 3: Under these assumptions, we have that (c.f. Hannan and Kavalieri (1986) and Bühlmann (1995))

$$\max_{0\leq k\leq p} |\widehat{\gamma}_k - \gamma_k| = O((n/\log n)^{-1/2}) \ a.s., \tag{64}$$

and there exists a random variable n_1 such that:

$$\sup_{n\geq n_1}\sum_{k=0}^{\infty} k^2|\gamma_k| < +\infty \ a.s. \tag{65}$$

Thus, we have that:

$$\begin{aligned} \|\Sigma - \widehat{\Sigma}\|_{col} &\leq 2\sum_{k=0}^{\infty}|\widehat{\gamma}_k - \gamma_k| \\ &= 2\left(|\sum_{k=0}^{p}|\widehat{\gamma}_k - \gamma_k| + \sum_{k=p+1}^{\infty}|\widehat{\gamma}_k - \gamma_k|\right) \\ &= O((n/\log n)^{-1/2})p + o(p^{-2}) \ a.s. = o((n/\log n)^{-1/3}) \ a.s., \end{aligned} \tag{66}$$

and

$$\begin{aligned} \|\Sigma^{-1} - \widehat{\Sigma}^{-1}\|_{col} &\leq \|\Sigma^{-1}\|_{col}\|\Sigma - \widehat{\Sigma}\|_{col}\|\widehat{\Sigma}^{-1}\|_{col} \\ &= o((n/\log n)^{-1/3}) \ a.s. \end{aligned} \tag{67}$$

Let's define B_j, A_j, $\widehat{B}_j$ and $\widehat{A}_j$ by:

$$B_j = A_j^2 = n^2(\widetilde{T}_N^{(j)} - T_N)^2 \tag{68}$$

and

$$\widehat{B}_j = \widehat{A}_j^2 = n^2(\widehat{\widetilde{T}}_N^{(j)} - T_N)^2. \tag{69}$$

Note that $|B_j - \widehat{B}_j| = |A_j - \widehat{A}_j||A_j + \widehat{A}_j|$. Next, we find a bound that does not depend on j.

$$|A_j| = |\widetilde{w}'_{n,j}(X-\mu)| \leq \|\widetilde{w}_{n,j}\|_1\|X-\mu\|_\infty, \tag{70}$$

$$\begin{aligned} \|\widetilde{w}_{n,j}\|_1 &\leq \|\Sigma^{-1}H_j\left(H_j'\Sigma^{-1}H_j\right)^{-1}H_j'\|_{col}\|w_{n,j}\|_1 \\ &\leq l\|\Sigma^{-1}\|_{col}\|\left(H_j'\Sigma^{-1}H_j\right)^{-1}\|_{col}, \end{aligned} \tag{71}$$

and

$$\begin{aligned} \|(\Sigma^{-1}(\alpha_j))^{-1}\|_{col} &= \|\Sigma(\alpha_j)+\Sigma(\alpha_j,\alpha_j')\Sigma(\alpha_j')^{-1}\Sigma(\alpha_j',\alpha_j)\|_{col} \\ &\leq \|\Sigma(\alpha_j)\|_{col}+l^{1/2}\|\Sigma(\alpha_j,\alpha_j')\Sigma(\alpha_j')^{-1}\Sigma(\alpha_j',\alpha_j)\|_{spec} \\ &\leq \|\Sigma\|_{col}+l^{1/2}\|\Sigma\|_{col}^2(2\pi F_1)^{-1}. \end{aligned} \tag{72}$$

In (72) we can consider the involved matrices as squared matrices, because defining Δ and $\Delta^{(a)}$ (Δ augmented) by

$$\Delta_{l\times l}=\Sigma(\alpha_j,\alpha_j')\Sigma(\alpha_j')^{-1}\Sigma(\alpha_j',\alpha_j), \tag{73}$$

and

$$\Delta^{(a)}_{n-l\times n-l}=\left[\begin{array}{c}\Sigma(\alpha_j,\alpha_j')\\ 0_{n-2l\times n-l}\end{array}\right]\Sigma(\alpha_j')^{-1}\left[\begin{array}{cc}\Sigma(\alpha_j',\alpha_j) & 0_{n-l\times n-2l}\end{array}\right], \tag{74}$$

and using that $\|A\|_{spec}=\max\{\|Ax\|:\|x\|_2\leq 1\}$, we have that

$$\begin{aligned} \|\Delta\|_{spec} &\leq \|\Delta^{(a)}\|_{spec}\leq \left\|\left[\begin{array}{c}\Sigma(\alpha_j,\alpha_j')\\ 0\end{array}\right]\right\|_{spec}\|\Sigma(\alpha_j')^{-1}\|_{spec}\left\|\left[\begin{array}{cc}\Sigma(\alpha_j',\alpha_j) & 0\end{array}\right]\right\|_{spec} \\ &\leq \|\Sigma(\alpha_j,\alpha_j')\|_{col}\|\Sigma(\alpha_j,\alpha_j')\|_{row}\|\Sigma(\alpha_j')^{-1}\|_{spec}. \end{aligned} \tag{75}$$

From (72) we have $A_j=O(l^{3/2}\log^{1/2}n)$ *a.s.* For $|\widehat{A}_j|$, we can proceed in a similar way. Note that $\|\widehat{\Sigma}\|_{col}$ and $\|\widehat{\Sigma}^{-1}\|_{col}$ are $O(1)$ *a.s.*; also, by symmetry, we have that $\|\widehat{\Sigma}\|_{row}$ and $\|\widehat{\Sigma}^{-1}\|_{row}$ are $O(1)$ *a.s.*, and $\|\widehat{\Sigma}\|_{spec}^2\leq\|\widehat{\Sigma}\|_{col}\|\widehat{\Sigma}\|_{row}$ and $\|\widehat{\Sigma}^{-1}\|_{spec}^2\leq\|\widehat{\Sigma}^{-1}\|_{col}\|\widehat{\Sigma}^{-1}\|_{row}$; then $\widehat{A}_j=O(l^{3/2}\log^{1/2}n)$ *a.s.*

Now,

$$|A_j-\widehat{A}_j|=\left|\left(\widetilde{w}_{n,j}-\widehat{\widetilde{w}}_{n,j}\right)'(X-\mu)\right|\leq\|\widetilde{w}_{n,j}-\widehat{\widetilde{w}}_{n,j}\|_1\|X-\mu\|_\infty, \tag{76}$$

$$\begin{aligned} \|\widetilde{w}_{n,j}-\widehat{\widetilde{w}}_{n,j}\|_1 &= \|\Sigma^{-1}H_j\big(H_j'\Sigma^{-1}H_j\big)^{-1}H_j'w_{n,j}-\widehat{\Sigma}^{-1}H_j\big(H_j'\widehat{\Sigma}^{-1}H_j\big)^{-1}H_j'w_{n,j}\|_1 \\ &\leq l\big(\|\Sigma^{-1}-\widehat{\Sigma}^{-1}\|_{col}\|(\Sigma^{-1}(\alpha_j))^{-1}\|_{col} \\ &\qquad+\|\widehat{\Sigma}^{-1}\|_{col}\|(\Sigma^{-1}(\alpha_j))^{-1}-(\widehat{\Sigma}^{-1}(\alpha_j))^{-1}\|_{col}\big) \end{aligned} \tag{77}$$

and

$$\|\Sigma^{-1}(\alpha_j)^{-1}-\widehat{\Sigma}^{-1}(\alpha_j)^{-1}\|_{col} \le \|\Sigma(\alpha_j)-\widehat{\Sigma}(\alpha_j)\|_{col}+\|\Sigma(\alpha_j,\alpha_j')\Sigma(\alpha_j')^{-1}\Sigma(\alpha_j',\alpha_j) - \widehat{\Sigma}(\alpha_j,\alpha_j')\widehat{\Sigma}(\alpha_j')^{-1}\widehat{\Sigma}(\alpha_j',\alpha_j)\|_{col}$$
$$\le O(l^{1/2})\|\Sigma-\widehat{\Sigma}\|_{col} = o\left(l^{1/2}(n/\log n)^{-1/3}\right) \ a.s. \quad (78)$$

Then,

$$\|\widetilde{w}_{n,j}-\widehat{\widetilde{w}}_{n,j}\|_1 = o\left(l^{3/2}(n/\log n)^{-1/3}\right) \ a.s., \quad (79)$$

$$|A_j-\widehat{A}_j| = o\left(l^{3/2}(n/\log n)^{-1/3}\log^{1/2} n\right) \ a.s., \quad (80)$$

$$|B_j-\widehat{B}_j| = o\left(l^{9/2}(n/\log n)^{-1/3}\log^{3/2} n\right) \ a.s. \quad (81)$$

and

$$|n\widetilde{\sigma}^2_{Jack}-n\widehat{\widetilde{\sigma}}^2_{Jack}| = o\left(l^{7/2}(n/\log n)^{-1/3}\log^{3/2} n\right) \ a.s. \quad (82)$$

To end the proof, only rest to use that $l = o\left(n^{2/21-\varepsilon}\right)$.■

Proof of Theorem 2: We use extensively the relation between H_N and L_N and the symmetry of $J(\rho)$, i.e. $J(x,\rho) = 1 - J(-x,\rho)$ and the following result from Politis and Romano (1994):

Theorem *Assume that there exists a limiting law $J(\rho)$ such that $J_N(\rho)$ converge weakly to a limit law $J(\rho)$, as $n \to \infty$. and that $\tau_b/\tau_n \to 0$, $b/n \to 0$ and $b \to \infty$ as $n \to \infty$. Also assume that α-mixing sequence satisfy that $\alpha_X(k) \to 0$ as $k \to \infty$.*

1. *If x is a continuity point of $J(\cdot,\rho)$, then $L_N(x) \to J(x,\rho)$ in probability.*
2. *If $J(\cdot,\rho)$ is continuous, then $\sup_x |L_N(x) - J(x,\rho)| \to 0$ in probability.*

By symmetry, if x is a continuity point of $J(\cdot,\rho)$, also $-x$ is a continuity point. Then, using statement (1) of the theorem, we have

$$H_N(x) = 1 - L_N(-x) \to 1 - J(-x,\rho) = J(x,\rho) \text{ in probability.} \quad (83)$$

Using the statement (2) of the theorem, we obtain the convergence to 0 in probability, since

$$\begin{aligned}\sup_x |H_N(x) - J(x,\rho)| &= \sup_x |1 - L_N(-x) - (1 - J(-x,\rho))| \\ &= \sup_x |L_N(-x) - J(-x,\rho)|.\end{aligned} \quad (84)$$

Notice that if $\tau_l = \sqrt{l}$, and using that $l/n \to 0$, then the coefficient $\tau_l l^{-1}(n-l)$ is close to $\sqrt{n(n-l)l^{-1}}$ which is the standardizing constant of Wu (1990).■

Proof of Proposition 2: We have that

$$\begin{aligned} l^{-1/2}\sum_{t=1}^{n} w_n(t-j)(\widehat{X}_{t,j}-T_N) &= l^{-1/2}\sum_{t=1}^{n} w_n(t-j)(\widehat{X}_{t,j}-\mu) - l^{1/2}(T_N-\mu) \\ &= l^{-1/2}\sum_{t=1}^{n} w_n(t-j)(\widehat{X}_{t,j}-\mu) + O_P(l^{1/2}n^{-1/2}) \\ &= l^{-1/2}w_{n,j}'(\widehat{X}_j-\mu)+O_P(l^{1/2}n^{-1/2}), \end{aligned} \tag{85}$$

where $w_{n,j}' = (w_n(1-j),\dots,w_n(n-j))$ and $\widehat{X}_j' = (\widehat{X}_{1,j},\dots,\widehat{X}_{n,j})$. Now,

$$\begin{aligned} l^{-1/2}w_{n,j}'(\widehat{X}_j-\mu) &= l^{-1/2}w_{n,j}'\left(I-H_j(H_j'\Sigma^{-1}H_j)^{-1}H_j'\Sigma^{-1}\right)(X-\mu) \\ &= l^{-1/2}(w_{n,j}-\widetilde{w}_{n,j})'(X-\mu). \end{aligned} \tag{86}$$

From the proof of Lemma 2 we have that

$$\widetilde{w}_{n,j} = \Sigma^{-1}H_j(H_j'\Sigma^{-1}H_j)^{-1}H_j'w_{n,j} = \begin{bmatrix} B(\Sigma^{-1}(\alpha_j))^{-1}1_{l\times 1} \\ 1_{l\times 1} \\ D_1'(\Sigma^{-1}(\alpha_j))^{-1}1_{l\times 1} \end{bmatrix} \tag{87}$$

and

$$\begin{aligned} \|\Sigma^{-1}(\alpha_j',\alpha_j)(\Sigma^{-1}(\alpha_j))^{-1}\|_1 &= \|-\Sigma(\alpha_j')^{-1}\Sigma(\alpha_j',\alpha_j)\|_1 \\ &\leq 4M\sum_{k=1}^{\infty} k|\gamma_k|. \end{aligned} \tag{88}$$

Then,

$$\|w_{n,j}-\widetilde{w}_{n,j}\|_1 = \left\|\begin{bmatrix} B(\Sigma^{-1}(\alpha_j))^{-1}1_{l\times 1} \\ D_1'(\Sigma^{-1}(\alpha_j))^{-1}1_{l\times 1} \end{bmatrix}\right\|_1 \leq \left\|\Sigma^{-1}(\alpha_j',\alpha_j)(\Sigma^{-1}(\alpha_j))^{-1}\right\|_1, \tag{89}$$

and for some $0 < \varepsilon < 1/2$,

$$l^{-1/2}(w_{n,j}-\widetilde{w}_{n,j})'(X-\mu) = o_P(l^{-1/2+\varepsilon}). \tag{90}$$

Therefore,

$$l^{-1/2}\sum_{t=1}^{n} w_n(t-j)(\widehat{X}_{t,j}-T_N) = O_P(l^{1/2}n^{-1/2}) + o_P(l^{-1/2+\varepsilon}). \tag{91}$$

On the other hand,

$$\begin{aligned} |l^{-1/2}(w_{n,j}-\widetilde{w}_{n,j})'(X-\mu)| &\leq l^{-1/2}\|w_{n,j}-\widetilde{w}_{n,j}\|_1\|X-\mu\|_\infty \\ &= O_{a.s.}(l^{-1/2}\log^{1/2}(n)). ■ \end{aligned} \tag{92}$$

Proof of Proposition 3: Assuming that $n = s(l+k)$, we have that

$$\begin{aligned} &(s(l+k))^{-1/2}\sum_{i=1}^{s}\sum_{j=l+1}^{l+k}(\widehat{X}^*_{(i-1)(l+k)+j}-\bar{X}_n) \\ &\quad = (s(l+k))^{-1/2}W'\left(I_{s(l+k)}-H(H'\Sigma^{-1}H)^{-1}H'\Sigma^{-1}\right)(X^*-\bar{X}), \end{aligned} \tag{93}$$

where $I_{s(l+k)}$ is the $s(l+k) \times s(l+k)$ identity matrix, $X^* = (X_1^*, \ldots, X_{s(l+k)}^*)'$, $\bar{X} = \bar{X}_n 1_{n\times 1}$, and W is a $s(l+k) \times 1$ vector defined as follows:

$$W' = (\underbrace{0, \ldots, 0}_{l\ times}, \underbrace{1, \ldots, 1}_{k\ times}, \ldots\ldots, \underbrace{0, \ldots, 0}_{l\ times}, \underbrace{1, \ldots, 1}_{k\ times}),$$

i.e., W indicates the missing observations positions.

Analogously to M²BJ, the matrix $H(H'\Sigma^{-1}H)^{-1}H'\Sigma^{-1}$ have submatrices equal to the $k \times k$ identity matrix in the missing observations positions, and the remaining non-zero elements are elements of $-\Sigma(\alpha')^{-1}\Sigma(\alpha', \alpha)$, where $\alpha = (l+1, \ldots, l+k, 2l+k+1, \ldots, 2(l+k), \ldots, \ldots, sl+(s-1)k+1, \ldots, s(l+k))$. Therefore,

$$\begin{aligned} &W'\left(I - H(H'\Sigma^{-1}H)^{-1}H'\Sigma^{-1}\right) \\ &\quad = (a_1, \ldots, a_l, \underbrace{0, \ldots, 0}_{k\ times}, \ldots, a_{(s-1)(l+k)+1}, \ldots\ldots, a_{(s-1)(l+k)+l}, \underbrace{0, \ldots, 0}_{k\ times}), \end{aligned} \tag{94}$$

where the a's are 0 or are the sum of one column of $-\Sigma(\alpha')^{-1}\Sigma(\alpha', \alpha)$, and they satisfy that $\sum_{t=1}^{s(l+k)} |a_t| \leq 4M \sum_{k=1}^{\infty} k|\gamma_k|$. Then,

$$\mathrm{E}^*\left[(s(l+k))^{-1/2} \sum_{t=1}^{s(l+k)} a_t(X_t^* - \bar{X}_n)\right] = 0, \tag{95}$$

and

$$\begin{aligned} &\mathrm{E}^*\left[\left((s(l+k))^{-1/2} \textstyle\sum_{t=1}^{s(l+k)} a_t(X_t^* - \bar{X}_n)\right)^2\right] \\ &\quad = (s(l+k))^{-1} \textstyle\sum_{t=1}^{s(l+k)} \sum_{s=1}^{s(l+k)} a_t a_s \mathrm{E}^*\left[(X_t^* - \bar{X}_n)(X_s^* - \bar{X}_n)\right] \\ &\quad \leq (s(l+k))^{-1} \textstyle\sum_{t=1}^{s(l+k)} \sum_{s=1}^{s(l+k)} |a_t a_s| \, \mathrm{E}^*\left[(X_1^* - \bar{X}_n)^2\right] \\ &\quad = (s(l+k))^{-1} O(1) O_{a.s.}(1) = o_{a.s.}((s(l+k))^{-1+\varepsilon}), \end{aligned} \tag{96}$$

for some $0 < \varepsilon < 1$.

Finally, (95) and (96) imply that $(s(l+k))^{-1/2} \sum_{t=1}^{s(l+k)} a_t(X_t^* - \bar{X}_n) = o_P(1)$ for almost all sample sequences $X_1, \ldots, X_N$.■

References

[1] Abraham, B. y Thavaneswaran, A. (1991) A nonlinear time series model and estimation of missing observations, *Ann. Inst. Statist. Math.*, **43**, 493-504.

[2] Beveridge, S. (1992) Least squares estimation of missing values in time series, *Commun. Statist. Theory Meth.*, **21**, 3479-3496.

[3] Berk, K. N. (1974) Consistent autoregressive spectral estimates, *Ann. Statist.*, **2**, 489-502.

[4] Bose, A. (1990) Bootstrap in moving average models, *Ann. Inst. Statist. Math.*, **42**, 753-768.

[5] Brockwell, P. J. and Davis, R. A. (1991) *Time series: Theory and methods*, Springer, New York.

[6] Bühlmann, P. (1995) Moving-average representation of autoregressive approximations, *Stochastic Processes and their Applications*, **60**, 331-342.

[7] Bühlmann, P. (1996) Locally adaptive lag-window spectral estimation, *J. Time Ser. Anal.*, **17**, 247-270.

[8] Bühlmann, P. (1997) Sieve bootstrap for time series, *Bernoulli*, **8**, 123-148.

[9] Bühlmann, P. and Künsch, H.R. (1994) Block length selection in the bootstrap for time series, *R.R. No. 72 Eidgenössische Technische Hochschule.*

[10] Carlstein, E. (1986) The use of subseries values for estimating the variance of a general statistics from a stationary sequence, *Ann. Statist.*, **14**, 1171-1194.

[11] Carlstein, E., Do, K., Hall, P., Hesterberg, T. and Künsch, H. R. (1996) Matched-block bootstrap for dependent data, *Bernoulli*, **4**, 305-328.

[12] Efron, B. (1979) Bootstrap methods: Another look at the jackknife, *Ann. Statist.*, **7**, 1-26.

[13] Efron, B. and Tibshirani, R. J. (1986) Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy, *Statist. Science*, **1**, 54-77.

[14] Freedman, D. A. (1984) On bootstrapping two-stage least-square estimates in stationary linear models, *Ann. Statist.*, **12**, 827-842.

[15] Galbraith, R. F. and Galbraith, J. I. (1974) On the inverse of some patterned matrices arising in the theory of stationary time series, *J. Appl. Prob.*, **11**, 63-71.

[16] Hannan , E. J. and Kavalieris, L. (1986) Regression, autoregression models, *J. Time Ser. Anal.*, **7**, 27-49.

[17] Harvey, A. C. and Pierse, R. G. (1984) Estimating missing observations in economic time series, *J. Amer. Statist. Assoc.*, **79**, 125-131.

[18] Horn, R. A. and Johnson, C. R. (1985) *Matrix Analysis*, Cambridge University Press, Cambridge.

[19] Kreiss, J. P. and Franke, J. (1992) Bootstrapping stationary autoregressive moving average models, *J. Time Ser. Anal.*, **13**, 297-317.

[20] Künsch, H. R. (1989) The jackknife and the bootstrap for general stationary observations, *Ann. Statist.*, **17**, 1217-1241.

[21] Lahiri, S. N. (1999) Theoretical comparison of block bootstrap methods, *Ann. Statist.*, **27**, 386-404.

[22] Liu, R. Y. and Singh, K. (1992) Moving blocks jackknife and bootstrap capture weak dependence, In: *Exploring the Limits of Bootstrap*, R. Lepage and L. Billard eds., 225-248, Wiley, New York.

[23] Ljung, G. M. (1989) A note on the estimation of missing values in time series, *Commun. Statist. Simula.*, **18**, 459-465.

[24] Peña, D. (1990) Influential observations in time series, *J. Bus. Econ. Statist.*, **8**, 235-241.

[25] Peña, D. and Maravall, A. (1991) Interpolation, outliers, and inverse autocorrelations, *Commun. Statist. Theory Meth.*, **20**, 3175-3186.

[26] Politis, D. N. and Romano, J. F. (1992) A circular block-resampling procedure for stationary data, In: *Exploring the Limits of Bootstrap*, R. Lepage and L. Billard eds., 263-270, Wiley, New York.

[27] Politis, D. N. and Romano, J. F. (1994) Large sample confidence regions based on subsamples under minimal assumptions, *Ann. Statist.*, **22**, 2031-2050.

[28] Politis, D. N. and Romano, J. F. (1994) The stationary bootstrap, *J. Amer. Statist. Assoc.*, **89**, 1303-1313.

[29] Politis, D. N. and Romano, J. F. (1995) Bias-corrected nonparametric spectral estimation, *J. Time Ser. Anal.*, **16**, 67-103.

[30] Quenouille, M. (1949) Approximation test of correlation in time series, *J. R. Statist. Soc. B*, **11**, 18-84.

[31] Romano, J. F. and Thombs, L. A. (1996) Inference for autocorrelations under weak assumptions, *J. Amer. Statist. Assoc.*, **91**, 590-600.

[32] Shao, Q.-M. and Yu, H. (1993) Bootstrapping the sample means for stationary mixing sequences, *Stochastic Processes and their Applications*, **48**, 175-190.

[33] Sherman, M. (1998) Efficiency and robustness in subsampling for dependent data, *J. Statist. Planning Inference*, **75**, 133-146.

[34] Tukey, J. (1958) Bias and confidence in not quite large samples, *Ann. Math. Statist.*, **29**, 614.

Table 1: MBJ and M^2BJ to estimate σ_N^2 in the case of the sample mean. MSE denotes the mean square error of $\log(\widehat{\sigma}_N^2)$. (+) denotes where the MBJ outperforms the M^2BJ. (*) denotes where the M^2BJ outperforms the MBJ.

Model	N	$\log \sigma_N^2$	E_1	SD_1	L_1	MSE_1	E_2	SD_2	L_2	MSE_2
M1	480	3.21	2.95	0.25	35	0.130 (0.011)	2.90	0.24	45	0.152 (0.011)
M1	120	3.18	2.57	0.36	15	0.502 (0.030)	2.51	0.34	20	0.568 (0.032)
M2	480	1.68	1.61	0.09	2	0.012 (0.001)	1.65	0.09	7	0.009 (0.001) *
M2	120	1.70	1.57	0.18	2	0.050 (0.004)	1.55	0.18	7	0.056 (0.006)
M3	480	1.82	1.67	0.14	15	0.045 (0.003)	1.64	0.16	20	0.056 (0.004)
M3	120	1.83	1.52	0.17	4	0.119 (0.006)	1.38	0.23	10	0.251 (0.020) +
M4	480	-1.16	-1.09	0.16	25	0.030 (0.003)	-1.13	0.12	10	0.017 (0.002) *
M4	120	-1.13	-1.09	0.22	15	0.050 (0.005)	-1.07	0.16	4	0.029 (0.003) *
M5	480	-3.14	-3.04	0.22	85	0.057 (0.005)	-3.09	0.22	35	0.048 (0.005)
M5	120	-2.93	-2.57	0.18	30	0.166 (0.008)	-2.89	0.31	15	0.100 (0.012) *
M6	480	-3.05	-2.96	0.21	60	0.054 (0.004)	-3.03	0.10	5	0.011 (0.001) *
M6	120	-2.92	-2.85	0.21	30	0.051 (0.006)	-2.88	0.19	2	0.038 (0.004) *
M7	480	-1.81	-1.72	0.19	40	0.043 (0.004)	-1.71	0.15	15	0.032 (0.003) *
M7	120	-1.76	-1.64	0.24	20	0.073 (0.007)	-1.67	0.25	8	0.073 (0.007)
M8	480	1.18	1.09	0.11	8	0.020 (0.001)	1.05	0.11	10	0.030 (0.002) +
M8	120	1.17	0.99	0.16	4	0.058 (0.005)	0.90	0.22	7	0.118 (0.011) +

Table 2: MBJ and M^2BJ to estimate σ_N^2 in the case of the median. MSE denotes the mean square error of $\log(\widehat{\sigma}_N^2)$. (*) denotes where the M^2BJ outperforms the MBJ.

Model	N	$\log \sigma_N^2$	E_1	SD_1	L_1	MSE_1	E_2	SD_2	L_2	MSE_2
M1	480	3.31	3.04	0.36	35	0.202 (0.022)	3.37	0.31	15	0.101 (0.010) *
M1	120	3.29	2.67	0.52	20	0.654 (0.054)	2.81	0.48	15	0.458 (0.054) *
M2	480	2.04	2.02	0.43	35	0.183 (0.023)	1.93	0.42	20	0.187 (0.024)
M2	120	2.05	1.97	0.56	15	0.315 (0.035)	1.98	0.52	10	0.273 (0.033)
M3	480	2.05	1.91	0.42	20	0.194 (0.024)	1.90	0.39	15	0.177 (0.022)
M3	120	2.03	1.71	0.50	10	0.356 (0.038)	1.77	0.52	6	0.337 (0.036)
M4	480	0.33	0.24	0.56	40	0.329 (0.049)	0.28	0.49	30	0.243 (0.045) *
M4	120	0.34	0.29	0.65	15	0.426 (0.070)	0.18	0.57	15	0.352 (0.050)
M5	480	-0.10	-0.20	0.53	40	0.295 (0.037)	-0.15	0.42	20	0.180 (0.024) *
M5	120	-0.10	-0.22	0.72	15	0.528 (0.056)	-0.05	0.56	10	0.317 (0.035) *
M6	480	-0.28	-0.45	0.57	50	0.349 (0.036)	-0.36	0.53	20	0.288 (0.029)
M6	120	-0.28	-0.35	0.70	15	0.501 (0.060)	-0.12	0.54	10	0.321 (0.034) *
M7	480	0.18	0.06	0.52	55	0.285 (0.042)	0.17	0.45	25	0.201 (0.025) *
M7	120	0.19	-0.01	0.68	20	0.497 (0.071)	0.09	0.58	15	0.342 (0.062) *
M8	480	1.45	1.33	0.44	30	0.208 (0.026)	1.38	0.44	15	0.200 (0.023)
M8	120	1.43	1.25	0.57	10	0.360 (0.039)	1.24	0.47	8	0.263 (0.030) *

Table 3: MBJ and M^2BJ to estimate σ_N^2 in the case of the variance. MSE denotes the mean square error of $\log(\widehat{\sigma}_N^2)$.

Model	N	$\log \sigma_N^2$	E_1	SD_1	L_1	MSE_1	E_2	SD_2	L_2	MSE_2
M1	480	4.22	3.91	0.33	20	0.209 (0.016)	3.94	0.34	30	0.197 (0.017)
M1	120	4.11	3.31	0.55	9	0.941 (0.050)	3.43	0.59	15	0.810 (0.052)
M2	480	3.98	3.73	0.26	15	0.134 (0.010)	3.71	0.28	25	0.151 (0.011)
M2	120	3.98	3.38	0.46	9	0.576 (0.035)	3.35	0.47	15	0.616 (0.040)
M3	480	3.04	2.86	0.19	10	0.069 (0.005)	2.86	0.20	15	0.072 (0.006)
M3	120	3.02	2.58	0.36	7	0.327 (0.021)	2.59	0.37	10	0.327 (0.026)
M4	480	4.25	3.95	0.36	20	0.217 (0.016)	3.94	0.36	30	0.220 (0.018)
M4	120	4.25	3.51	0.58	10	0.883 (0.049)	3.50	0.56	15	0.891 (0.050)
M5	480	2.07	1.94	0.16	6	0.043 (0.003)	1.93	0.16	7	0.047 (0.003)
M5	120	2.08	1.82	0.29	4	0.149 (0.011)	1.79	0.31	5	0.178 (0.013)
M6	480	2.39	2.24	0.20	7	0.059 (0.005)	2.22	0.21	15	0.073 (0.006)
M6	120	2.40	2.10	0.34	5	0.205 (0.018)	2.05	0.37	8	0.255 (0.020)
M7	480	3.47	3.28	0.28	20	0.112 (0.009)	3.26	0.27	20	0.116 (0.009)
M7	120	3.48	3.04	0.47	9	0.423 (0.030)	2.97	0.43	7	0.447 (0.030)
M8	480	2.08	1.95	0.17	7	0.046 (0.004)	1.95	0.18	9	0.049 (0.004)
M8	120	2.06	1.83	0.33	4	0.162 (0.011)	1.82	0.33	5	0.168 (0.012)

Table 4: MBJ and M^2BJ to estimate σ_N^2 in the case of the covariance of order 1. MSE denotes the mean square error of $\log(\widehat{\sigma}_N^2)$. (*) denotes where the M^2BJ outperforms the MBJ.

Model	N	$\log \sigma_N^2$	E_1	SD_1	L_1	MSE_1	E_2	SD_2	L_2	MSE_2
M1	480	4.17	3.85	0.34	20	0.216 (0.018)	3.85	0.32	20	0.205 (0.016)
M1	120	4.04	3.24	0.58	9	0.960 (0.056)	3.31	0.53	10	0.808 (0.050)
M2	480	3.69	3.45	0.26	15	0.124 (0.010)	3.43	0.28	25	0.147 (0.011)
M2	120	3.67	3.08	0.46	8	0.553 (0.034)	3.06	0.47	15	0.579 (0.040)
M3	480	2.66	2.51	0.21	9	0.067 (0.006)	2.49	0.19	10	0.066 (0.005)
M3	120	2.62	2.21	0.38	5	0.317 (0.022)	2.22	0.39	8	0.317 (0.025)
M4	480	4.20	3.91	0.38	25	0.232 (0.018)	3.90	0.37	30	0.229 (0.018)
M4	120	4.19	3.43	0.58	10	0.914 (0.052)	3.38	0.52	10	0.927 (0.049)
M5	480	1.52	1.40	0.19	6	0.054 (0.004)	1.40	0.14	2	0.035 (0.003) *
M5	120	1.51	1.25	0.29	2	0.150 (0.012)	1.29	0.28	2	0.124 (0.011)
M6	480	2.05	1.91	0.21	9	0.064 (0.006)	1.88	0.21	10	0.074 (0.006)
M6	120	2.04	1.69	0.36	3	0.250 (0.020)	1.70	0.41	8	0.290 (0.025)
M7	480	3.34	3.13	0.29	20	0.125 (0.011)	3.12	0.28	20	0.127 (0.010)
M7	120	3.33	2.85	0.52	9	0.501 (0.035)	2.83	0.49	7	0.484 (0.035)
M8	480	1.52	1.39	0.18	4	0.050 (0.004)	1.46	0.17	3	0.032 (0.003) *
M8	120	1.48	1.21	0.32	2	0.173 (0.014)	1.26	0.29	2	0.135 (0.012) *

Table 5: MBJ and M^{2}BJ to estimate σ_N^2 in the case of the autocovariance of order 5. MSE denotes the mean square error of $\log(\widehat{\sigma}_N^2)$. (*) denotes where the M^{2}BJ outperforms the MBJ.

Model	N	$\log \sigma_N^2$	E_1	SD_1	L_1	MSE_1	E_2	SD_2	L_2	MSE_2
M1	480	3.79	3.39	0.40	25	0.319 (0.022)	3.52	0.33	15	0.176 (0.014) *
M1	120	3.55	2.72	0.59	7	1.041 (0.055)	2.97	0.49	9	0.578 (0.042) *
M2	480	3.50	3.22	0.29	20	0.157 (0.012)	3.30	0.24	15	0.097 (0.008) *
M2	120	3.43	2.78	0.48	8	0.651 (0.039)	2.91	0.44	10	0.469 (0.031) *
M3	480	2.39	2.20	0.20	10	0.078 (0.006)	2.42	0.17	5	0.029 (0.003) *
M3	120	2.30	1.92	0.38	6	0.278 (0.022)	2.22	0.37	6	0.144 (0.016) *
M4	480	3.84	3.46	0.41	20	0.311 (0.019)	3.56	0.34	15	0.198 (0.013) *
M4	120	3.79	2.92	0.58	7	1.109 (0.053)	3.18	0.55	10	0.684 (0.042) *
M5	480	1.37	1.25	0.14	4	0.031 (0.003)	1.46	0.12	2	0.023 (0.002) *
M5	120	1.33	1.14	0.29	3	0.121 (0.009)	1.33	0.27	2	0.072 (0.007) *
M6	480	1.69	1.55	0.18	6	0.050 (0.004)	1.77	0.15	4	0.029 (0.003) *
M6	120	1.65	1.40	0.35	5	0.185 (0.016)	1.61	0.32	4	0.102 (0.010) *
M7	480	2.78	2.54	0.27	10	0.133 (0.009)	2.74	0.22	7	0.048 (0.005) *
M7	120	2.73	2.28	0.44	6	0.399 (0.029)	2.53	0.39	6	0.193 (0.021) *
M8	480	1.36	1.26	0.15	5	0.033 (0.003)	1.45	0.13	2	0.025 (0.002) *
M8	120	1.28	1.12	0.31	3	0.122 (0.011)	1.27	0.27	2	0.073 (0.008) *

Table 6: MBB and M^{2}BB to estimate σ_N^2 in the case of the sample mean. MSE denotes the mean square error of $\log(\widehat{\sigma}_N^2)$. (*) denotes where the M^{2}BB outperforms the MBB.

Model	N	$\log \sigma_N^2$	E_1	SD_1	L_1	MSE_1	E_2	SD_2	L_2	MSE_2
M1	480	3.21	2.93	0.24	30	0.135 (0.010)	3.00	0.19	5 15	0.079 (0.006) *
M1	120	3.18	2.56	0.36	15	0.521 (0.031)	2.76	0.31	5 9	0.275 (0.018) *
M2	480	1.68	1.60	0.11	2	0.018 (0.001)	1.70	0.10	1 1	0.011 (0.001) *
M2	120	1.70	1.83	0.20	3	0.055 (0.005)	1.65	0.19	1 1	0.040 (0.004) *
M3	480	1.82	1.66	0.16	15	0.052 (0.004)	1.87	0.11	3 5	0.014 (0.001) *
M3	120	1.83	1.51	0.18	4	0.130 (0.007)	1.86	0.17	2 3	0.028 (0.003) *
M4	480	-1.16	-1.08	0.16	25	0.033 (0.003)	-1.09	0.16	1 15	0.029 (0.003)
M4	120	-1.13	-1.10	0.24	15	0.057 (0.005)	-1.03	0.20	1 8	0.049 (0.004)
M5	480	-3.14	-2.84	0.17	60	0.117 (0.006)	-3.12	0.22	1 55	0.050 (0.005) *
M5	120	-2.93	-2.57	0.19	30	0.167 (0.008)	-2.90	0.24	2 30	0.056 (0.007) *
M6	480	-3.05	-2.96	0.22	60	0.056 (0.005)	-2.96	0.11	1 10	0.020 (0.002) *
M6	120	-2.92	-2.85	0.21	30	0.049 (0.006)	-2.94	0.22	1 15	0.047 (0.004)
M7	480	-1.81	-1.74	0.20	40	0.045 (0.004)	-1.76	0.18	1 25	0.035 (0.004) *
M7	120	-1.76	-1.79	0.27	25	0.072 (0.007)	-1.72	0.26	1 15	0.068 (0.007)
M8	480	1.18	1.09	0.13	9	0.025 (0.002)	1.17	0.08	1 1	0.006 (0.001) *
M8	120	1.17	0.98	0.17	4	0.064 (0.005)	1.15	0.12	1 1	0.015 (0.002) *

Table 7: MBB and M^2BB to estimate σ_N^2 in the case of the median. MSE denotes the mean square error of $\log(\widehat{\sigma}_N^2)$. (*) denotes where the M^2BB outperforms the MBB.

Model	N	$\log \sigma_N^2$	E_1	SD_1	L_1	MSE_1	E_2	SD_2	L_2	MSE_2
M1	480	3.31	3.06	0.29	30	0.147 (0.013)	3.10	0.19	5 9	0.084 (0.006) *
M1	120	3.29	2.76	0.44	15	0.471 (0.035)	2.90	0.36	5 10	0.277 (0.021) *
M2	480	2.04	2.08	0.24	15	0.058 (0.007)	1.98	0.14	1 1	0.023 (0.002) *
M2	120	2.05	2.07	0.33	9	0.111 (0.013)	1.94	0.23	1 1	0.066 (0.006) *
M3	480	2.05	1.89	0.23	15	0.075 (0.008)	2.06	0.14	3 5	0.021 (0.003) *
M3	120	2.03	1.80	0.29	4	0.137 (0.013)	2.06	0.22	2 3	0.051 (0.006) *
M4	480	0.33	0.44	0.29	8	0.097 (0.009)	0.26	0.18	3 5	0.037 (0.004) *
M4	120	0.34	0.45	0.38	4	0.156 (0.016)	0.26	0.27	3 5	0.077 (0.009) *
M5	480	-0.10	-0.04	0.30	15	0.093 (0.009)	-0.22	0.16	3 5	0.038 (0.003) *
M5	120	-0.10	0.05	0.37	7	0.162 (0.016)	-0.12	0.23	3 5	0.054 (0.005) *
M6	480	-0.28	-0.23	0.30	15	0.092 (0.009)	-0.25	0.21	2 5	0.046 (0.004) *
M6	120	-0.28	-0.18	0.40	9	0.168 (0.019)	-0.37	0.27	3 5	0.084 (0.007) *
M7	480	0.18	0.20	0.29	20	0.085 (0.011)	0.14	0.18	3 5	0.033 (0.003) *
M7	120	0.19	0.27	0.40	9	0.167 (0.019)	0.15	0.24	3 5	0.061 (0.007) *
M8	480	1.45	1.36	0.21	6	0.053 (0.005)	1.39	0.12	1 1	0.017 (0.001) *
M8	120	1.43	1.27	0.28	3	0.105 (0.010)	1.39	0.16	1 1	0.028 (0.003) *

Table 8: MBB and M^2BB to estimate σ_N^2 in the case of the variance. MSE denotes the mean square error of $\log(\widehat{\sigma}_N^2)$.

Model	N	$\log \sigma_N^2$	E_1	SD_1	L_1	MSE_1	E_2	SD_2	L_2	MSE_2
M1	480	4.22	3.88	0.33	20	0.219 (0.017)	3.86	0.33	3 20	0.233 (0.018)
M1	120	4.11	3.28	0.53	8	0.970 (0.050)	3.24	0.51	3 7	1.026 (0.054)
M2	480	3.98	3.73	0.27	15	0.139 (0.011)	3.74	0.29	3 20	0.139 (0.011)
M2	120	3.98	3.34	0.47	10	0.622 (0.036)	3.38	0.47	3 10	0.581 (0.037)
M3	480	3.04	2.83	0.20	9	0.082 (0.006)	2.83	0.20	1 10	0.081 (0.006)
M3	120	3.02	2.55	0.37	8	0.353 (0.023)	2.55	0.36	1 7	0.350 (0.022)
M4	480	4.25	3.94	0.35	20	0.222 (0.017)	3.92	0.36	4 25	0.238 (0.019)
M4	120	4.25	3.50	0.56	9	0.875 (0.050)	3.46	0.56	4 7	0.948 (0.053)
M5	480	2.07	1.95	0.17	9	0.044 (0.004)	1.95	0.17	1 8	0.043 (0.004)
M5	120	2.08	1.81	0.28	3	0.152 (0.012)	1.80	0.26	1 2	0.146 (0.011)
M6	480	2.39	2.24	0.20	7	0.062 (0.005)	2.25	0.18	2 4	0.050 (0.005)
M6	120	2.40	2.09	0.36	6	0.225 (0.019)	2.10	0.33	2 4	0.199 (0.017)
M7	480	3.47	3.26	0.27	15	0.118 (0.009)	3.26	0.27	1 15	0.115 (0.009)
M7	120	3.48	3.01	0.46	8	0.438 (0.030)	2.98	0.46	1 7	0.464 (0.029)
M8	480	2.08	1.93	0.18	5	0.052 (0.004)	1.95	0.18	1 5	0.048 (0.004)
M8	120	2.06	1.80	0.31	3	0.168 (0.012)	1.79	0.30	1 4	0.163 (0.012)

Table 9: MBB and M^2BB to estimate σ_N^2 in the case of the autocovariance of order 1. MSE denotes the mean square error of $\log(\widehat{\sigma}_N^2)$. (*) denotes where M^2BB outperforms the MBB.

Model	N	$\log \sigma_N^2$	E_1	SD_1	L_1	MSE_1	E_2	SD_2	L_2	MSE_2
M1	480	4.17	3.79	0.35	25	0.268 (0.019)	3.80	0.32	3 20	0.245 (0.019)
M1	120	4.04	3.13	0.61	15	1.195 (0.065)	3.14	0.53	3 7	1.081 (0.059)
M2	480	3.69	3.41	0.27	20	0.153 (0.011)	3.43	0.24	3 9	0.125 (0.009)
M2	120	3.67	2.99	0.46	10	0.678 (0.036)	3.11	0.45	3 7	0.519 (0.034) *
M3	480	2.66	2.45	0.23	20	0.098 (0.007)	2.48	0.22	2 10	0.080 (0.006)
M3	120	2.62	2.11	0.39	8	0.414 (0.025)	2.18	0.37	2 5	0.336 (0.025)
M4	480	4.20	3.84	0.37	20	0.263 (0.019)	3.86	0.37	4 25	0.247 (0.019)
M4	120	4.19	3.35	0.62	15	1.081 (0.058)	3.37	0.57	4 7	1.002 (0.056)
M5	480	1.52	1.36	0.19	9	0.065 (0.005)	1.39	0.18	1 5	0.049 (0.004) *
M5	120	1.51	1.20	0.32	5	0.194 (0.013)	1.24	0.30	1 3	0.159 (0.012)
M6	480	2.05	1.88	0.24	15	0.086 (0.008)	1.90	0.19	2 4	0.057 (0.005) *
M6	120	2.04	1.65	0.40	8	0.315 (0.025)	1.73	0.35	2 3	0.217 (0.020) *
M7	480	3.34	3.09	0.30	20	0.151 (0.011)	3.10	0.28	2 15	0.133 (0.010)
M7	120	3.33	2.77	0.50	10	0.564 (0.037)	2.81	0.53	2 10	0.544 (0.037)
M8	480	1.52	1.34	0.20	8	0.073 (0.005)	1.41	0.20	1 8	0.051 (0.005) *
M8	120	1.48	1.09	0.29	3	0.231 (0.014)	1.16	0.28	1 1	0.179 (0.013) *

Table 10: MBB and M^2BB to estimate σ_N^2 in the case of the autocovariance of order 5. MSE denotes the mean square error of $\log(\widehat{\sigma}_N^2)$.

Model	N	$\log \sigma_N^2$	E_1	SD_1	L_1	MSE_1	E_2	SD_2	L_2	MSE_2
M1	480	3.79	3.30	0.39	35	0.389 (0.022)	3.31	0.37	3 25	0.366 (0.022)
M1	120	3.55	2.60	0.47	5	1.133 (0.045)	2.59	0.53	2 10	1.204 (0.054)
M2	480	3.50	3.16	0.29	30	0.201 (0.013)	3.18	0.29	4 25	0.189 (0.012)
M2	120	3.43	2.68	0.38	5	0.702 (0.032)	2.76	0.44	3 10	0.643 (0.035)
M3	480	2.39	2.15	0.18	15	0.092 (0.006)	2.19	0.20	2 15	0.081 (0.006)
M3	120	2.30	1.94	0.30	5	0.212 (0.015)	1.92	0.31	1 9	0.241 (0.016)
M4	480	3.84	3.40	0.39	30	0.351 (0.020)	3.40	0.39	2 25	0.337 (0.019)
M4	120	3.79	2.89	0.46	5	1.034 (0.045)	2.92	0.54	3 10	1.053 (0.050)
M5	480	1.37	1.29	0.14	5	0.025 (0.002)	1.29	0.13	1 6	0.022 (0.002)
M5	120	1.33	1.24	0.25	5	0.070 (0.006)	1.22	0.24	1 4	0.069 (0.006)
M6	480	1.69	1.56	0.15	5	0.038 (0.003)	1.60	0.17	2 9	0.034 (0.003)
M6	120	1.65	1.51	0.27	5	0.095 (0.011)	1.50	0.28	2 6	0.100 (0.012)
M7	480	2.78	2.52	0.24	15	0.124 (0.008)	2.54	0.26	2 15	0.124 (0.008)
M7	120	2.73	2.40	0.41	5	0.273 (0.018)	2.34	0.37	1 4	0.292 (0.019)
M8	480	1.36	1.26	0.13	5	0.027 (0.002)	1.28	0.13	1 9	0.024 (0.002)
M8	120	1.28	1.18	0.25	5	0.073 (0.007)	1.17	0.24	1 4	0.069 (0.007)