International Journal of Statistics and Probability; Vol. 12, No. 4; July 2023
ISSN 1927-7032 E-ISSN 1927-7040
Published by Canadian Center of Science and Education

Bootstrap Probability Errors of the Whittle MLE for Linear Regression Processes with Strongly Dependent Disturbances

Mosisa Aga

Correspondence: Department of Mathematics, Auburn University at Montgomery, Montgomery, AL, USA. E-mail: maga@aum.edu

Received: April 12, 2023 Accepted: July 11, 2023 Online Published: July 27, 2023

doi:10.5539/ijsp.v12n4p26 URL: https://doi.org/10.5539/ijsp.v12n4p26

Abstract

This paper determines bounds on the asymptotic orders of the coverage probability errors of parametric bootstrap confidence intervals (CIs) and tests for the covariance parameters of a time series generated by a regression model with Gaussian, stationary, and strongly dependent errors. The CIs and tests are based on the plug-in Whittle maximum likelihood (PWML) estimators. It is shown that, under some sets of conditions on the regression coefficients, the spectral density function, and the parameter values, the bounds on the coverage probability errors of symmetric two-sided and one-sided parametric bootstrap confidence intervals on the plug-in Whittle log-likelihood function are shown to be $O(n^{-3/2} \ln n)$ and $O(n^{-1} \ln n)$, respectively. Apart from the $\ln n$ term, the magnitudes of the coverage probability errors of the one-sided bootstrap confidence intervals for our model is shown to be essentially the same as that of the independent and identically distributed (iid) data. The error for the two-sided confidence intervals is not as small as the error $O(n^{-2})$ that has been established for many confidence intervals in the literature, see Hall (1992), pp 102-108.

Keywords: Edgeworth Expansion, parametric bootstrap, t-statistic, linear regression, long memory

2000 Mathematics Subject Classification: 62M10

1. Introduction

Introduced by Efron (1979), the bootstrap approximation technique has a history of less than half a century. Most research work in the literature during the last few decades of the introduction of the bootstrap was geared towards its application in the iid case or short memory time series in the nonparametric context. During subsequent years Efron and others have expanded the applicability of the bootstrap for confidence intervals, hypothesis testing, regression models, and more complex problems. Such ideas were explored in Efron (1982), Efron and Gong (1983), Diaconis and Efron (1983), and Efron and Tibshirani (1986). The explosion of Bootstrap papers grew at an exponential rate during the 1980s and 1990s when more key results appeared in the works of Singh (1981), Bickel and Freedman (1981, 1984), Beran (1982), Martin (1990), Hall (1986, 1988), Hall and Martin (1988), and Navidi (1989) among others. Today bootstrapping has become a popular technique for approximating the distribution of many statistics of time series mainly in situations of short term persistence. While its use with long memory time series has increased significantly over the last two decades, there exist few theoretical justifications of its validity in the context of long memory processes.

Time series exhibiting long range dependence have applications in a variety of fields including astronomy, hydrology, economics, and finance where correlations may decrease particularly slowly between observations over time (see Hurst (1951), Maldelbrot and Van Ness (1968), and Beran (1994) among others.) While the bulk of material in the literature on linear regression models is focused on data whose error components are short memory, regression models whose error terms exhibit long-range disturbances have also been studied extensively; see for example, Kunsch (1986), Yajima (1988, 1991), Dalhaus (1995), Robinson and Hidalgo (1997), Sibbertsen (2001), Koul, Baillie, and Surgailis (2004), Choy and Taniguchi (2001), Ivanov and Leonenko (2008) among others. Yajima (1988) investigated estimation of the regression parameters by the least square estimator (LSE) and those describing the correlation structure of the error terms by using the residuals obtained from the LSE. Yajima (1991) established many asymptotic results for least squares error estimators and best linear unbiased estimators. Dahlhaus (1995) studied the estimation of the coefficients of a regression model in which the error terms exhibit long-range dependence. Robinson and Hidalgo (1997) investigated stochastic regressors where the components of the desgin matrix follow stationary process in the presence of long-range dependence in both errors and stochastic regressors.

Among the works towards the justification of the application of bootstrap on long memory time series are Lahiri (1993), Hidalgo (2003), Franco et al. (2004), Andrews et al. (2006), Aga et al. (2007), Hidalgo (2021) and some other more recent publications. Lahiri (1993) has shown that the moving block bootstrap provides valid approximation to the distribution of the normalized sample mean for a class of long-range dependent observations if and only if the underlying statistic is asymptotically normal. He has immensely contributed to the development of the use of the block bootstrap technique in his subsequent extensive publications. Hidalgo (2003) proposed an alternative approach to the moving block bootstrap resampling for the estimator of the parameters for time-series regression models. Franco et al. (2004) comparatively investigated the various available bootstrap techniques in semiparametric estimation methods for Auto-Regressive Fractionally Integrated Moving Average (ARFIMA) models through Monte Carlo simulation. Andrews et al. (2006) provided the coverage probability errors of both delta method and parametric bootstrap confidence intervals for both the plug-in maximum likelihood (PML) and plug-in Whittle maximum likelihood (PWML) estimators for the covariance parameters of stationary long-memory Gaussian time series. Aga (2007) extended the work of Andrews et al. (2006) to a linear regression model with Gaussian, stationary, and long-memory errors and provided coverage probability errors of the parametric bootstrap for the PML estimators of the covariance parameters of the model. This article determines the coverage probability errors when the Whittle maximum likelihhod estimators are used to approximate the plug-in maximum likelihood estimators of the parameters of the model.

Consider a linear regression model $\{Y_t = Z_t'\beta + \varepsilon_t, t \geq 1\}$, where $\beta = (\beta_1, \beta_2, ..., \beta_p)'$ is a p vector of deterministic but unknown real numbers, $\{Z_t = (z_{t1}, z_{t2}, ..., z_{tp})' \in \mathbb{R}^p, t \geq 1, p \geq 1\}$ are non-stochastic regressors, and the error terms $\{\varepsilon_t, t \geq 1\}$ are stationary, Gaussian, and strongly dependent discrete time series. The process $\{\varepsilon_t, t \geq 1\}$ is assumed to have mean zero and spectral density $f_\theta(\lambda)$ for $\lambda \in (-\pi, \pi)$, where $\theta = (\theta_1, \theta_2, \ldots, \theta_r)' \in \mathbb{R}^r$ and

$$f_\theta(\lambda) = O(|\lambda|^{-2d-\delta})$$

as $|\lambda| \downarrow 0, \forall\delta > 0, d \in (0, 1/2)$, and $\theta_1 = d$, referred to as the "long-memory parameter" of the process (see Andrews et al. (2006)).

Let $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)'$ be an observed sample of size n and $\mathcal{E} = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)'$ be the corresponding error terms, where for each $i = 1, 2, ..., n$, $Y_i = Z_i'\beta + \varepsilon_i$. We note that the covariance matrices of $\mathbf{Y}$ and $\mathcal{E}$ are the same.

If $\hat{\beta} = \left(\hat{\beta_1}, \hat{\beta_2}, \ldots, \hat{\beta_p}\right)'$ denote the least square estimate (LSE) of β, then $\hat{\beta} = V^{-1}\sum_{t=1}^n Y_t Z_t$, where $V = \sum_{t=1}^n (Z_t Z_t')$ is a $p \times p$ matrix. Let $\mu = (\mu_1, \mu_2, \ldots, \mu_n)$ be the true mean of $\mathbf{Y}$. Then, an estimator of μ is $\hat{\mu} = (\hat{\mu_1}, \ldots, \hat{\mu_n})$, where $\hat{\mu_t} = Z_t'\hat{\beta}$, $t = 1, 2, ..., n$. Let $\mathbf{Z}$ denote the design matrix given by $\mathbf{Z} = (z_{ij})$ for $i = 1, ..., n$ and $j = 1, ..., p$, where the rank of $\mathbf{Z}$ is p. We note that the matrix V is symmetric and positive definite.

The $n \times n$ (Toeplitz) covariance matrix corresponding to $f_\theta(\lambda)$ is denoted by $\mathcal{T}_n(f_\theta)$ and has (j, k) element defined by:

$$\mathcal{T}_n(f_\theta)_{j,k} = \int_{-\pi}^{\pi} \exp(i(j-k)\lambda) f_\theta(\lambda) d\lambda. \tag{1.1}$$

The log-likelihood function is

$$\mathcal{L}(\theta, \mu) = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\ln(\det(\mathcal{T}_n(f_\theta))) - \frac{1}{2}(Y-\mu)'\mathcal{T}_n^{-1}(f_\theta)(Y-\mu). \tag{1.2}$$

Based on the fact that

$$\frac{1}{n}\ln(\det(\mathcal{T}_n(f_\theta))) \to \frac{1}{2\pi}\int_{-\pi}^{\pi} \ln(f_\theta(\lambda)) d\lambda \tag{1.3}$$

as $n \to \infty$ and

$$\mathcal{T}_n((2\pi)^{-2} f_\theta^{-1}) \to \mathcal{T}_n^{-1}(f_\theta) \tag{1.4}$$

as $n \to \infty$ (Beran, 1994), the log-likelihood function (1.2) can now be approximated by

$$\mathcal{L}_W(\theta, \mu) = -\frac{n}{2}\ln(2\pi) - \frac{n}{4\pi}\int_{-\pi}^{\pi} \ln(f_\theta(\lambda)) d\lambda - \frac{1}{2}(Y-\mu)'\mathcal{T}_n((2\pi)^{-2} f_\theta^{-1})(Y-\mu). \tag{1.5}$$

We refer to $\mathcal{L}_W(\theta, \hat{\mu})$, where $\hat{\mu}$ is replaced for μ in (1.5) above, as the *plug-in Whittle log-likelihood* (PWLL) function. Let $R_n = \mathbf{Z}V^{-1}\mathbf{Z}'$ and let $\mathcal{M}_n = I_n - R_n$, where I_n is the $n \times n$ identity matrix. One can easily verify that the matrices $\mathcal{M}_n$ and R_n have the following properties: (a) Both $\mathcal{M}_n$ and R_n are symmetric. (b) $\mathbf{Y}'\mathcal{M}_n = (\mathbf{Y} - \hat{\mu})'$. (c) If $U = \mathbf{Y} - \mu$, then $\mathcal{M}_n\mathbf{Y} = \mathcal{M}_n U$. (d) There exists an $n \times p$ matrix $\mathbf{B}$ such that

$$R_n = \mathbf{B}\mathbf{B}'. \tag{1.6}$$

Using the properties (a) through (d), the PWLL function can now be written as

$$\mathcal{L}_W(\theta, \hat{\mu}) = -\frac{n}{2}\ln(2\pi) - \frac{n}{4\pi}\int_{-\pi}^{\pi} \ln((f_\theta)(\lambda)) - \frac{1}{2}Y'\mathcal{M}_n\mathcal{T}_n((2\pi)^{-2} f_\theta^{-1})\mathcal{M}_n Y \tag{1.7}$$

On the other hand, the last term in (1.5) can be approximated as

$$\begin{aligned}\frac{1}{2}(Y-\mu)'\mathcal{T}_n((2\pi)^{-2}f_\theta^{-1})(Y-\mu) &\approx \sum_{j,l=1}^{n}(Y_j-\mu_j)\frac{1}{8\pi^2}\int_{-\pi}^{\pi}f_\theta^{-1}(\lambda)e^{i(l-j)\lambda}d\lambda(Y_l-\mu_l)\\ &= \frac{1}{8\pi^2}\int_{-\pi}^{\pi}f_\theta^{-1}(\lambda)\sum_{j,l=1}^{n}(Y_j-\mu_j)(Y_l-\mu_l)e^{i(l-j)\lambda}d\lambda\\ &= \frac{1}{8\pi^2}\int_{-\pi}^{\pi}f_\theta^{-1}(\lambda)\left|\sum_{j=1}^{n}(Y_j-\mu_j)e^{i\lambda j}\right|^2\\ &= \frac{n}{4\pi}\int_{-\pi}^{\pi}f_\theta^{-1}(\lambda)I(\lambda)d\lambda,\end{aligned} \tag{1.8}$$

and therefore, the PWLL function becomes

$$\mathcal{L}_W(\theta,\hat{\mu}) = -\frac{n}{2}\ln(2\pi) - \frac{n}{4\pi}\int_{-\pi}^{\pi}(\ln(f_\theta(\lambda)+f_\theta^{-1}(\lambda)I_n(\lambda))d\lambda, \tag{1.9}$$

where $I_n(\lambda) = |\frac{1}{2n\pi}\sum_{j=1}^{n} e^{ij\lambda}(Y_j-\hat{\mu}_j)|^2$ is the periodogram.
By definition, the Whittle Maximum Likelihood Estimator (WMLE), $\hat{\theta}_n$, solves the equation:

$$\int_{\pi}^{\pi}\frac{\partial}{\partial\theta_r}(\ln f_\theta(\lambda)+f_\theta^{-1}(\lambda)I_n(\lambda))d\lambda = 0 \tag{1.10}$$

for $r = 1,...,d_\theta$, where $d_\theta = dim(\theta)$.

Andrews et al. (2006) have established the coverage probability errors of the symmetric two-sided and one-sided parametric bootstrap confidence intervals based on the Plug-in maximum likelihood (PML) and plug-in Whittle maximum likelihood (PWML) estimators of the parameter of the error component $\{\varepsilon_t, t\geq 1\}$ given above. In this paper we extend the work of Andrews et al. (2006) and establish the asymptotic order of magnitude of the coverage probability errors of the parametric bootstrap based on the PWML estimator of the linear regression processes described above by imposing an additional condition on the regression coefficients and a mild additional condition on the spectral density function.

The remainder of the paper proceeds as follows. Section 2 provides some preliminaries and background assumptions. In this section we present some preliminaries on cumulants and Edgeworth expansions, introduce the general set-up of parametric bootstrap confidence intervals and tests, describe the Whittle log-likelihood derivatives, and state the background assumptions. Section 3 presents and proves several lemmas and two main theorems on the coverage probability errors of the parametric bootstrap of the PWML estimator of our linear regression model.

2. Preliminaries and Assumptions

2.1 Cumulants and Edgeworth Expansions

For a random variable Y with a characteristic function $\chi(t) = E(e^{itY})$, the jth cumulant, κ_j, of Y is defined to be the coefficient of $\frac{1}{j!}(it)^j$ in a power series expansion of $\log\chi(t) = \sum_{j\geq 1}\frac{1}{j!}\kappa_j(it)^j$. Moreover, the *jth* moment μ_j of Y is defined by $\mu_j = E(Y^j)$. Using these notations we have

$$\chi(t) = E(e^{itY}) = 1+\mu_1 it+\frac{1}{2!}\mu_2(it)^2+...+\frac{1}{j!}\mu_j(it)^j+..., \tag{2.1}$$

and

$$\begin{aligned}\sum_{j\geq 1}\frac{1}{j!}\kappa_j(it)^j &= \log\left[1+\sum_{j\geq 1}\frac{1}{j!}\mu_j(it)^j\right]\\ &= \sum_{j\geq 1}(-1)^{k+1}\frac{1}{k}\left(\sum_{j\geq 1}\frac{1}{j!}\mu_j(it)^j\right)^k.\end{aligned} \tag{2.2}$$

By equating the coefficients of $\frac{1}{j!}(it)^j$ we obtain $\kappa_1 = \mu_1, \kappa_2 = \mu_2-\mu_1^2, \kappa_3 = \mu_3-3\mu_1\mu_2+2\mu_1^3, \kappa_4 = \mu_4-4\mu_1\mu_3-3\mu_2^2+12\mu_1^2\mu_2-6\mu_1^4$, and so on.

Let $X_1, ..., X_n$ be independent and identically distributed with mean θ_0 and finite variance σ^2 and let $\hat{\theta}$ be the sample mean. Then by the Central Limit Theorem $S_n = \sqrt{n}(\hat{\theta} - \theta_0)/\sigma$ is asymptotically normally distributed with zero mean and unit variance and the characteristic function $\chi_n(t)$ of S_n can be written as

$$\begin{aligned} \chi_n(t) &= E(\exp itS_n) \\ &= \exp\{-\frac{1}{2}t^2 + n^{-1/2}\frac{1}{3!}\kappa_3(it)^3 + ... + n^{-(j-2)/2}\frac{1}{j!}\kappa_j(it)^j + ...\} \\ &= e^{-t^2/2} + n^{-1/2}r_1(it)e^{-t^2/2} + ... + n^{-j/2}r_j(it)e^{-t^2/2} + ..., \end{aligned} \tag{2.2}$$

where r_j is a polynomial with real coefficients depending on the cumulants $\kappa_3, ..., \kappa_{j+2}$. Using the inverse Fourier transform of (2.2) we obtain

$$P(S_n \le x) = \Phi(x) + n^{-1/2}\pi_1(x)\phi(x) + n^{-1}\pi_2(x)\phi(x)... + n^{-j/2}\pi_j(x)\phi(x) + ...$$

where the polynomial π_j is in terms of cumulants, Φ and ϕ denote the standard normal cumulative distribution function and probability distribution function, respectively. The right hand side (rhs) of the last equation above is an Edgeworth expansion of the distribution function $P(S_n \le x)$. (For more general Edgeworth expansion theory see Hall (1992), pp 39-81.)

2.2 Confidence Intervals and Tests of the Parametric Bootstrap

To help us introduce the bootstrap coverage probability errors, we first define the parametric bootstrap sample and formulate the general set up of bootstrap confidence intervals and tests.

Bootstrapping is one of the different re-sampling techniques in which a series of random samples are drawn a large number of times with replacement from an original sample $\mathcal{X}$ obtained from the population of interest. The statistic of interest is then calculated from each of the bootstrap samples and an approximate of the sampling distribution of the statistic is obtained from the calculated values. This is what is known as the non-parametric bootstrap. When the distribution $\mathcal{F}_{\theta_0}$ of the population is assumed to be completely known up to a vector θ_0 of unknown parameters, then a parametric bootstrap may be used. If $\hat{\theta}$ is an estimate of θ_0 computed from the sample $\mathcal{X}$, then a parametric bootstrap sample $\mathcal{X}^*$ is obtained by drawing from the distribution $\mathcal{F}_{\hat{\theta}}$. Then, a large number of parametric bootstrap samples are generated in this way and the statistic of interest to be computed from these samples (see Hall (1992) for details).

Now, let $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)'$ be a sample from our linear regression model with strongly dependent errors as described in section 1 above. Then, the parametric bootstrap sample $\mathbf{Y}^*= (Y_1^*, ..., Y_n^*)$ is the same as the distribution of the original sample except that the true parameters are $(\hat{\theta}_n, \hat{\mu})$ instead of (θ_0, μ). In other words, $\mathbf{Y}^*$ consists of random variables from a linear regression process with stationary, Gaussian, and strongly dependent errors having mean $\hat{\mu}$ and spectral density $f_{\hat{\theta}_n}(\lambda)$ conditional on the original sample $\mathbf{Y}$.

In order to establish the coverage probability errors of the bootstrap confidence intervals, we will need to define the boostrap analogues of sample mean, the PWLL function, the bootstrap estimator of the true parameter θ_0, the bootstrap t-statistic, and the one-sided and two-sided bootstrap confidence intervals.
(a) The bootstrap sample mean $\hat{\mu}^*$ is then given by $\hat{\mu}^* = (\hat{\mu}_1^*, ..., \hat{\mu}_n^*)$, where, for $t = 1, ..., n$, $\hat{\mu}_t^* = Z_t'\hat{\beta}^*$, $\hat{\beta}^* = V^{-1}\sum_{t=1}^n Y_t^* Z_t$, where $V = \sum_{t=1}^n (Z_t Z_t')$ as defined in section 1.

(b) The bootstrap PWLL function $\mathcal{L}_W(\theta, \hat{\mu}^*)$ is defined in the same way as the PWLL function $\mathcal{L}_W(\theta, \hat{\mu})$ (see (1.7) above) but with $\mathbf{Y}^*$ and $\hat{\mu}^*$ replacing $\mathbf{Y}$ and $\hat{\mu}$, respectively.

(c) Let Θ^* denote the set of solutions in the parameter space Θ to the first order conditions for the bootstrap PWLL function. The bootstrap estimator $\hat{\theta}_n^*$ can now be defined as that value of θ that maximizes the bootstrap PWLL function $\mathcal{L}_W(\theta, \hat{\mu}^*)$. Observe that the true parameter of the bootstrap sample is $\hat{\theta}_n$, and hence $\hat{\theta}_n^*$ is a PWML estimator of $\hat{\theta}_n$.

Let θ_h denote some element of Θ, the parameter space. Let $\theta_{0,r}$, $\theta_{h,r}$, and $\hat{\theta}_{n,r}$ denote the r-th elements of θ_0, θ_h, and $\hat{\theta}_n$, respectively. The asymptotic covariance matrix of a consistent PWML estimator $\hat{\theta}_n$ is $\Sigma(\theta_0)$, where

$$\Sigma(\theta) = \left(\frac{1}{4\pi}\int_{-\pi}^{\pi}\frac{\partial}{\partial\theta}\ln(f_\theta(\lambda))\frac{\partial}{\partial\theta'}\ln(f_\theta(\lambda))d\lambda\right)^{-1}.$$

A consistent estimator of $\Sigma(\theta_0)$ is $\Sigma(\hat{\theta}_n)$, provided that $f_\theta(\lambda)$ is smooth with respect of θ. Let $\Sigma_{r,r}(\hat{\theta}_n)$ denote the (r,r)-th element of $\Sigma(\hat{\theta}_n)$. Let z_α denote the $1-\alpha$ quantile of the standard normal distribution.

(d) We define the t-statistic by

$$\tau_n(\theta_{0,r}) = \frac{n^{1/2}(\hat{\theta}_{n,r} - \theta_{0,r})}{\sum_{r,r}^{1/2}(\hat{\theta}_n)} \tag{2.3}$$

and the bootstrap t-statistic by

$$\tau_n^*(\hat{\theta}_{n,r}) = \frac{\sqrt{n}(\hat{\theta}_{n,r}^* - \hat{\theta}_{n,r})}{\Sigma_{r,r}^{1/2}(\hat{\theta}_n^*)} \quad (2.4).$$

where $\hat{\theta}_{n,r}^*$ denotes the r-th element of $\hat{\theta}_n^*$.

(e) Let $z_{|\tau|,\alpha}^*$ and $z_{\tau,\alpha}^*$ denote the $1-\alpha$ quantiles of $|\tau_n^*(\hat{\theta}_{n,r})|$ and $\tau_n^*(\hat{\theta}_{n,r})$, respectively. To be precise, $z_{|\tau|,\alpha}^*$ is defined to be a value that minimizes $|P^*(|\tau_n^*(\hat{\theta}_{n,r})| \leq z) - (1-\alpha)|$ over $z \in \mathbb{R}$. (The precise definition of $z_{\tau,\alpha}^*$ is analogous.)

(f) The *symmetric two-sided bootstrap* CI for $\theta_{0,r}$ with approximate confidence level $100(1-\alpha)\%$ based on the PWML estimator $\hat{\theta}_n$ is

$$\mathcal{I}_2(\hat{\theta}_n) = \left[\hat{\theta}_{n,r} - \frac{z_{|\tau|,\alpha}^* \Sigma_{r,r}^{1/2}(\hat{\theta}_n)}{\sqrt{n}}, \hat{\theta}_{n,r} + \frac{z_{|\tau|,\alpha}^* \Sigma_{r,r}^{1/2}(\hat{\theta}_n)}{\sqrt{n}}\right].$$

(g) The *upper one-sided bootstrap* $100(1-\alpha)\%$ CI for $\theta_{0,r}$ is

$$\mathcal{I}_1(\hat{\theta}_n) = \left[\hat{\theta}_{n,r} - \frac{z_{\tau,\alpha}^* \Sigma_{r,r}^{1/2}(\hat{\theta}_n)}{\sqrt{n}}, \infty\right).$$

(h) The *symmetric two-sided bootstrap t test* of $H_0 : \theta_{0,r} = \theta_{h,r}$ versus $H_1 : \theta_{0,r} \neq \theta_{h,r}$ with significance level α rejects H_0 if $|\tau_n(\theta_{h,r})| > z_{|\tau|,\alpha}^*$.

(i) The *one sided bootstrap t test* of $H_0 : \theta_{0,r} \leq \theta_{h,r}$ versus $H_1 : \theta_{0,r} > \theta_{h,r}$ with significance level α rejects H_0 if $\tau_n(\theta_{h,r}) > z_{\tau,\alpha}^*$.

2.3 The Whittle Log-likelihood Derivatives

Let $v = (r_1, r_2, ..., r_q)'$ denote a q-vector of positive integers each less than or equal to $dim(\theta)$. Let $\mathcal{L}_W(\theta, \hat{\mu})$ be as defined in (1.9) and let $\mathcal{M}_n$ be as in the paragraph preceding (1.6). We write the real valued q-th order partial derivative of the PWLL function indexed by v as

$$\mathcal{L}_{W,v} = D_v \mathcal{L}_W(\theta, \hat{\mu}) = \frac{\partial^q}{\partial\theta_{r_1} \ldots \partial\theta_{r_q}} \mathcal{L}_W(\theta, \hat{\mu}) = \mathcal{F}_{n,v}(\theta) + \mathbf{Y}'\mathcal{M}_n \mathcal{G}_{n,v}(\theta)\mathcal{M}_n \mathbf{Y} \quad (2.5)$$

where

$$\mathcal{F}_{n,v}(\theta) = -\frac{n}{4\pi}\int_{-\pi}^{\pi} D_v \ln(f_\theta(\lambda))d\lambda \quad (2.6)$$

and

$$\mathcal{G}_{n,v}(\theta) = -\frac{1}{2} D_v \mathcal{T}_n((2\pi)^{-2} f_\theta^{-1}). \quad (2.7)$$

Equations (2.5)-(2.7) are modified versions of equations (A.3) and (A.4) of Andrews *et al.* [2006] in which we used the matrix $\mathcal{M}_n$ and the estimator $\hat{\mu}$ of our sample $\mathbf{Y}$ as defined in section 1 above. We shall introduce some more notations. Let

$$\mathcal{U}_n(\theta) = (\mathcal{L}_{W,v(1)}(\theta), \ldots, \mathcal{L}_{W,v(r)}(\theta)), \quad (2.8)$$

where each vector $v(j)$ is of the same form as v defined in (2.5)-(2.7) above for $r = dim(\mathcal{U}_n(\theta))$ and $j = 1, 2, ..., r$. Let

$$\mathcal{W}_n(\theta) = n^{-1/2}(\mathcal{U}_n(\theta) - E_\theta \mathcal{U}_n(\theta)). \quad (2.9)$$

Without loss of generality we may assume that $E_\theta \mathcal{U}_n(\theta) = 0$. Let

$$\mathcal{D}_n(\theta) = E[\mathcal{W}_n(\theta)\mathcal{W}_n(\theta)'] \quad (2.10)$$

and let $\mathcal{D}(\theta) = \lim_{n\to\infty} \mathcal{D}_n(\theta)$.
Because $\mathcal{W}_n(\theta)$ is a vector of centeral quadratic forms in Gaussian variables plus a vector of nonrandom quantities we have

$$\mathcal{D}_n(\theta)i, j = tr(\mathcal{G}_{n,v_i}\mathcal{T}_n(f_\theta)\mathcal{G}_{n,v_j}\mathcal{T}_n(f_\theta)) \quad (2.11).$$

(See Anderson, 1984 for details.)

2.4 Assumptions

In this subsection we present the assumptions for which the results of this paper hold. Assumptions A1-A3 impose conditions on the parameter space. In particular the PWML estimators for which we establish bootstrap coverage probability errors are required to be consistent by A2 and the matrices $\mathcal{D}_n(\theta)$ and $\mathcal{D}(\theta)$ in (2.10) are required to be positive definite by A3. Assumptions A4-A9 are needed to control the behavior of the spectral density function, its inverse, and their derivatives in the neighborhood of the origin. These assumptions depend on a positive integer $s \geq 3$ that indexes the order of the Whittle Log-likelihood Derivatives (WLLDs) that are used in the construction of the bootstrap confidence intervals. Assumption A10 gives some restriction on the design matrix. It is mainly due to this assumption that we extend the results of Andrews et al. (2006) on establishing the magnitude of errors of the bootstrap confidence intervals and tests to our current model. In particular, Theorem 3.3 of Aga (2021) uses this assumption to establish that the rth cumulants $\kappa_r(\theta)$ of the WLLDs in the Edgeworth expansion are bounded by $O(n)$ which in turn are used to prove lemmas presented in section 3 of this paper.

A1. The parameter space Θ is a subset of $\mathbb{R}^r$ where $r = dim(\theta_0)$ with non-empty interior, where θ_0 is the true parameter.

A2. For all $\varepsilon > 0$ and all compact subsets Θ_c of Θ, the sequence of PWML estimators $\{\bar{\theta}_n : n \geq 1\}$ for which the results of this paper hold satisfy

$$\sup_{\theta_0 \in \Theta_c} P_{\theta_0}(\|\bar{\theta}_n - \theta_0\| > n^{-1/2}\ln(n)\varepsilon) = o(n^{1-s/2}) \;\; as \;\; n \to \infty$$

for some integer $s \geq 3$.

A3. The matrices $\mathcal{D}_n(\theta)$ and $\mathcal{D}(\theta)$ in (2.10) are positive definite.

A4. For some integer $s \geq 3$, $g(\theta) = \int_{-\pi}^{\pi} \ln f_\theta(\lambda) d\lambda$ and $h(\theta) = \int_{-\pi}^{\pi} f_\theta^{-1}(\lambda) I_n(\lambda) d\lambda$ can be differentiated $s+1$ times under the integral sign.

A5. $f_\theta(\lambda)$ is continuous at all (λ, θ) for which $\lambda \neq 0$, $f_\theta^{-1}(\lambda)$ is continuous at all (λ, θ), and $\forall \delta > 0$, $\exists c_1(\theta, \delta) < \infty$ such that

$$|f_\theta(\lambda)| \leq c_1(\theta, \delta)|\lambda|^{-2d-\delta}$$

for all λ in the neighborhood N_δ of the origin, where $\theta = (d, \theta_2, ..., \theta_r)$ and $d \in (0, 1/2)$.

A6. For all $(j_1, ..., j_k)$ with $k \leq s+1$ and $j_i \in \{1, ..., r\}$, $(\partial^k/(\partial\theta_{j1}...\partial\theta_{jk}))f_\theta^{-1}(\lambda)$ is continuous at all (λ, θ) and $\forall \delta > 0$, $\exists c_2(\theta, \delta) < \infty$ such that

$$\left| \frac{\partial^k f_\theta^{-1}(\lambda)}{\partial\theta_{j1}...\partial\theta_{jk}} \right| \leq c_2(\theta, \delta)|\lambda|^{2d-\delta}, \forall \lambda \in N_\delta.$$

A7. $(\partial/\partial\lambda)f_\theta(\lambda)$ is continuous at all (λ, θ) for which $\lambda \neq 0$ and $\forall \delta > 0$, $\exists c_4(\theta, \delta) < \infty$ such that

$$\left| \frac{\partial f_\theta(\lambda)}{\partial\lambda} \right| \leq c_4(\theta, \delta)|\lambda|^{2d-1-\delta}, \forall \lambda \in N_\delta.$$

A8. For all $(j_1, ..., j_k)$ with $k \leq s+1$ and $j_i \in \{1, ..., r\}$, $(\partial^{k+1}/(\partial\lambda\partial\theta_{j1}...\partial\theta_{jk}))f_\theta^{-1}(\lambda)$ is continuous at all (λ, θ) for which $\lambda \neq 0$ and $\forall \delta > 0$, $\exists c_5(\theta, \delta) < \infty$ such that

$$\left| \frac{\partial^{k+1} f_\theta^{-1}(\lambda)}{\partial\lambda\partial\theta_{j1}...\partial\theta_{jk}} \right| \leq c_4(\theta, \delta)|\lambda|^{2d-1-\delta}, \forall \lambda \in N_\delta.$$

A9. For any compact subset Θ_c of the parameter space there exists a constant $C(\Theta_c, \delta) < \infty$ such that the constants $c_i(\theta, \delta)$ for $i = 1, ..., 4$ given above are bounded by $C(\Theta_c, \delta)$, $\forall \theta \in \Theta_c$ and $\forall \delta > 0$.

A10. The design matrix $\mathbf{Z}$ is chosen in such a way that for the matrix

$$B = (e_{ij}), i = 1, ..., n, j = 1, ..., p, \tag{2.12}$$

defined by (1.6) above, there exists a constant $M < \infty$ such that $|e_{ij}| \leq \frac{M}{\sqrt{n}}$ for $i = 1, ..., n$, $j = 1, ..., p$.

Most of the assumptions stated above are standard assumptions in asymptotic theory and have appeared in numerous papers in the literature under different contexts including Dahlhaus (1989), Lieberman et al. (2003), Andrews et al. (2006), and Aga et al. (2007) among others. Assumptions A1 and A4-A9 are essentially Assumptions W1-W7 of Andrews et al. (2006) restated here for convenience and Assumption A3 is Assumption VIII of Lieberman et al. (2003). Assumption A2 is Condition C_s of Andrews et al. (2006) suitably adjusted.

One drawback of the results of this paper is that the sequence of PWML estimators for which the bootstrap coverage probability errors are established are required to satisfy Assumption A2. The same drawback exists in Dahlhaus (1989), Lieberman et al (2003) and Aga et al. (2007). While Lemma 1 of Andrews et al. (2006) shows the existence of PWML estimators that satisfy this assumption, it is generally unknown whether or not all sequences of PWML estimators satisfy this assumption.

Another drawback of the result of this paper (and that of Aga et al. (2007) and Aga (2021)) is the restriction imposed on the design matrix by A10 which requires the elements of the design matrix to be bounded by $\frac{M}{\sqrt{n}}$ for $M < \infty$. While Lemma 3.1 of Aga et al. (2007) provides a useful example of design matrix that satisfies this assumption, we do not generally show that the results of this paper and others about the current regression model will go through without this condition.

3. Bootstrap Coverage Probability Errors of the PWML Estimators

In this section we present the main result of establishing the bootstrap coverage prpbability errors of the Whittle maximum likelihood estimators of our linear regression model. The next lemma is one of the key ingredients to achieve this goal. First we introduce some additional notations.

Let $\Phi(\cdot)$ denote the distribution function of the standard normal distribution. Define $D_{\omega,\eta} = \frac{\partial^q}{\partial\omega_{\eta_1}\ldots\partial\omega_{\eta_q}}$, for $\eta = (\eta_1, ..., \eta_q)$. Let $\varphi_n(\omega, \theta) = E_\theta \exp(i\omega'\mathcal{U}_n(\theta))$ denote the characteristic function of $\mathcal{U}_n(\theta)$ where $\omega \in \mathbb{R}^d$ and let $\kappa_n(\theta)_\eta$ denote the η cumulants of $\mathcal{U}_n(\theta)$ (see equation (2.8) above). By definition, $\kappa_n(\theta)_\eta = i^{-q} D_{\omega,\eta} \ln(\varphi_n(\omega, \theta))|_{\omega=0}$, where $i = \sqrt{-1}$. The vector $\kappa_n(\theta)$ is composed of elements $\kappa_n(\theta)_\eta$ for vectors η of dimension $q \le s$, where s is as given in Assumption A2. Let $\bar{\kappa}_n(\theta) = \frac{\kappa_n(\theta)}{n}$. By Theorem 3.3 of Aga (2021), the elements of $\bar{\kappa}_n(\theta)$ are $O(1)$.

Let $P_j(\Delta, \bar{\kappa}_n(\theta))$ be a polynomial in $\Delta = \partial/\partial z$ whose coefficients are polynomials in the elements of $\bar{\kappa}_n(\theta)$ and for which $P_j(\Delta, \bar{\kappa}_n(\theta))\Phi(x)$ is an even function of x when j is odd and an odd function of x when j is even for $j = 1, 2, ..., s-2$. (see for example Hall (1992), pp. 41-45).

Lemma 3.1. Let $\{\xi_n(\theta_0) \in \mathbb{R}^d : n \ge 1\}$ and $\{\mathcal{Z}_n(\theta_0) \in \mathbb{R}^d : n \ge 1\}$ be a sequence of random vectors such that $\sup_{\theta_0\in\Theta_c} P_{\theta_0}(\|\xi_n(\theta_0)\| > \epsilon_n) = \epsilon_n$ and $\{\mathcal{Z}_n(\theta_0) \in \mathbb{R}^d : n \ge 1\}$ have Edgeworth expansions for each $\theta_0 \in \Theta_c$ with coefficients of order $O(1)$ and remainders of order $\epsilon_n = o(n^{-(s-2)/2})$ both uniformly over $\theta_0 \in \Theta_c$. Then,

$$\sup_{\theta_0\in\Theta_c} \sup_{C\in C_d} |P_{\theta_0}(\mathcal{Z}_n(\theta_0) + \xi_n(\theta_0) \in C) - P_{\theta_0}(\mathcal{Z}_n(\theta_0) \in C)| = o(n^{-(s-2)/2}), \tag{3.1}$$

where C_d denotes the class of all convex sets in $\mathbb{R}^d$.

Proof.

Let $\tilde{P}(z, \theta_0) = 1 + \sum_{i=1}^{s-2} n^{-i/2}\pi_{ni}(z, \theta_0)\phi(z)$ be an Edgeworth expansion to order $s-2$ of $\{\mathcal{Z}_n(\theta_0) : n \ge 1\}$ where $\{\pi_{ni}(z, \theta_0) : i = 1, ..., s-2, n \ge 1\}$ are polynomials in z whose coefficients are $O(1)$ uniformly over $\theta_0 \in \Theta_c$ such that

$$\sup_{\theta_0\in\Theta_c} \sup_{C\in C_d} |P_{\theta_0}(\mathcal{Z}_n(\theta_0) \in C) - \int_C (\tilde{P}(z, \theta_0))dz| = o(n^{-(s-2)/2}), \tag{3.2}$$

where $\phi(z)$ is the density function of a $N(0, \Sigma_n(\theta_0))$ random vector, $\Sigma_n(\theta_0)$ has eigenvalues that are bounded away from zero and infinity as $n \to \infty$ uniformly over $\theta_0 \in \Theta_c$. Now, for any convex set $C \in \mathbb{R}^d$ and $r > 0$, let $C_r^+ = \{x \in \mathbb{R}^d : \|x - y\| \le r\}$ for some $y \in C$. Let $A_n(\theta) = \mathcal{Z}_n(\theta_0) + \xi_n(\theta_0)$ and let $\Omega_n = P_{\theta_0}(A_n(\theta) \in C) - P_{\theta_0}(\mathcal{Z}_n(\theta_0) \in C)$. Then

$$\begin{aligned}
\sup_{\theta_0\in\Theta_c} \sup_{C\in C_d} \Omega_n = & \sup_{\theta_0\in\Theta_c} \sup_{C\in C_d} (P_{\theta_0}(A_n(\theta) \in C, \|\xi_n(\theta_0)\| \le \epsilon_n) - P_{\theta_0}(\mathcal{Z}_n(\theta_0) \in C)) \\
& + P_{\theta_0}(A_n(\theta) \in C, \|\xi_n(\theta_0)\| > \epsilon_n) \\
\le & \sup_{\theta_0\in\Theta_c} \sup_{C\in C_d} (P_{\theta_0}(\mathcal{Z}_n(\theta_0) \in C_{\epsilon_n}^+) - P_{\theta_0}(\mathcal{Z}_n(\theta_0) \in C)) \\
& + \sup_{\theta_0\in\Theta_C} P_{\theta_0}(\|\xi_n(\theta_0)\| > \epsilon_n).
\end{aligned} \tag{3.3}$$

By assumption, $\sup_{\theta_0\in\Theta_C} P_{\theta_0}(\|\xi_n(\theta_0)\| > \epsilon_n) = o(n^{-(s-2)/2})$. Because $\mathcal{Z}_n(\theta_0)$ has an Edgeworth expansion with remainder ϵ_n, the expression $\sup_{\theta_0\in\Theta_c} \sup_{C\in C_d} (P_{\theta_0}(\mathcal{Z}_n(\theta_0) \in C_{\epsilon_n}^+) - P_{\theta_0}(\mathcal{Z}_n(\theta_0) \in C))$ in (3.3) is less than or equal to

$$\sup_{\theta_0\in\Theta_c} \sup_{C\in C_d} \left(\int_{C_{\epsilon_n}^+} (\tilde{P}(z, \theta_0))dz - \int_C (\tilde{P}(z, \theta_0))dz \right) + \epsilon_n. \tag{3.4}$$

We observe that $\phi(z)$ and its derivatives of all orders are bounded over $z \in R^d$ and that by assumption the polynomials $\{\pi_{ni}(z, \theta_0) : i = 1, ..., s-2\}$ have coefficient that are $O(1)$ uniformly for $\theta_0 \in \Theta_c$. Therefore, the expression in (3.4) is $O(\omega_n) = o(n^{-(s-2)/2})$ and consequently, $\sup_{\theta_0\in\Theta_c} \sup_{C\in C_d} (P_{\theta_0}(\mathcal{Z}_n(\theta_0) + \xi_n(\theta_0) \in C) - P_{\theta_0}(\mathcal{Z}_n(\theta_0) \in C)) \le \epsilon_n$ holds.

The t-statistic is given by $\tau_n(\theta_{0,r})$ as in equation (2.3). Let $\frac{1}{n}\mathcal{U}_n^+(\theta_0)$ denote the vector $\frac{1}{n}\mathcal{U}_n(\theta_0)$ of normalized Whittle log-likelihood derivatives augmented to include the vector of expected values of all partial derivatives with respect to θ of order s of $\frac{1}{n}\mathcal{L}_W(\theta_0,\hat{\mu})$ (see (1.9)). The next lemma shows that $\tau_n(\hat{\theta}_{n,r})$ can be approximated by a smooth function of $\frac{1}{n}\mathcal{U}_n(\theta_0)$.

Lemma 3.2. Suppose Assumptions A1-A10 hold. Then, there is an infinitely differentiable function $\tilde{\Gamma}$ that satisfies $\tilde{\Gamma}(n^{-1}E_{\theta_0}\mathcal{U}_n^+(\theta)) = 0$ for n suffciently large and $\forall\theta_0 \in \Theta_c$ where Θ_c is some compact subset of the parameter space Θ, and

$$\sup_{\theta_0\in\Theta_c}\sup_{C\in\mathcal{C}_d}|P_{\theta_0}(\tau_n(\hat{\theta}_{0,r}) \in C) - P_{\theta_0}(\sqrt{n}\tilde{\Gamma}(n^{-1}\mathcal{U}_n^+(\theta_0)) \in C)| = o(n^{1-s/2}). \tag{3.5}$$

Proof.
Let $\boldsymbol{h}_n(\theta) = \frac{1}{n}\mathcal{L}_W(\theta,\hat{\mu})$ and $\widetilde{\boldsymbol{h}}_n(\hat{\theta}) = \frac{\partial}{\partial\theta}\boldsymbol{h}_n(\hat{\theta}_n)$. By Assumption A2, we have $\inf_{\theta_0\in\Theta_c} P_{\theta_0}(\hat{\theta}_n \in \Theta^0) = 1 - o(n^{1-s/2})$, where Θ^0 denotes the interior of Θ and

$$\inf_{\theta_0\in\Theta_c} P_{\theta_0}(\frac{\partial}{\partial\theta}\boldsymbol{h}_n(\hat{\theta}_n) = 0) = 1 - o(n^{1-s/2}). \tag{3.6}$$

By Taylor's expansion of $\widetilde{\boldsymbol{h}}_n(\hat{\theta})$ about θ_0 of order $s-1$, there exists $\bar{\theta}_n$ that lies between $\hat{\theta}_n$ and θ_0 such that

$$\widetilde{\boldsymbol{h}}_n(\hat{\theta}) = \widetilde{\boldsymbol{h}}_n(\theta_0) + \sum_{j=1}^{s-2}\frac{D^j\widetilde{\boldsymbol{h}}_n(\theta_0)}{j!}(\hat{\theta}_n-\theta_0,...,\hat{\theta}_n-\theta_0) + \frac{D^{s-1}\widetilde{\boldsymbol{h}}_n(\bar{\theta}_n)}{(s-1)!}(\hat{\theta}_n-\theta_0,...,\hat{\theta}_n-\theta_0), \tag{3.7}$$

where $D^j\widetilde{\boldsymbol{h}}_n(\theta_0)(\hat{\theta}_n-\theta_0,...,\hat{\theta}_n-\theta_0)$ denotes $D^j\widetilde{\boldsymbol{h}}_n(\theta_0)$ as a j-linear map, whose coefficients are partial derivatives of $\widetilde{\boldsymbol{h}}_n(\theta_0)$ of order j, applied to the j-tuple $(\hat{\theta}_n-\theta_0,...,\hat{\theta}_n-\theta_0)$.
Let,

$$\varepsilon_1(\theta_0) = \frac{1}{(s-1)!}(D^{s-1}\widetilde{\boldsymbol{h}}_n(\bar{\theta}_n) - D^{s-1}\widetilde{\boldsymbol{h}}_n(\theta_0))(\hat{\theta}_n-\theta_0,...,\hat{\theta}_n-\theta_0), \textit{ and}$$

$$\varepsilon_2(\theta_0) = \frac{1}{(s-1)!}(D^{s-1}\widetilde{\boldsymbol{h}}_n(\theta_0) - ED^{s-1}\widetilde{\boldsymbol{h}}_n(\theta_0))(\hat{\theta}_n-\theta_0,...,\hat{\theta}_n-\theta_0).$$

Then, (3.7) above can be written as

$$\begin{aligned}\widetilde{\boldsymbol{h}}_n(\hat{\theta}) &= \widetilde{\boldsymbol{h}}_n(\theta_0) + \textstyle\sum_{j=1}^{s-2}\frac{D^j\widetilde{\boldsymbol{h}}_n(\theta_0)}{j!}(\hat{\theta}_n-\theta_0,...,\hat{\theta}_n-\theta_0)\\ &+ \tfrac{1}{(s-1)!}ED^{s-1}\widetilde{\boldsymbol{h}}_n(\bar{\theta}_n)(\hat{\theta}_n-\theta_0,...,\hat{\theta}_n-\theta_0) + \varepsilon_1(\theta_0) + \varepsilon_2(\theta_0) = 0.\end{aligned} \tag{3.8}$$

Note that, by definition, $\frac{1}{n}\mathcal{U}_n^+(\theta_0)$ is the column vector whose elements are the non-redundant components of $\widetilde{\boldsymbol{h}}_n(\theta_0)$, $D^1\widetilde{\boldsymbol{h}}_n(\theta_0)$, ..., $D^{s-2}\widetilde{\boldsymbol{h}}_n(\theta_0)$ plus the components of $ED^{s-1}\widetilde{\boldsymbol{h}}_n(\theta_0)$. Let $\zeta_n(\theta_0) = ((\varepsilon_1(\theta_0)+\varepsilon_2(\theta_0))', 0, .., 0)'$ be conformable with $\mathcal{U}_n^+(\theta_0)$. Then, (3.8) can be written as

$$\boldsymbol{g}\left(\frac{1}{n}\mathcal{U}_n^+(\theta_0) + \zeta_n(\theta_0), \hat{\theta}_n - \theta_0\right) = 0, \tag{3.9}$$

where $\boldsymbol{g}(.,.)$ is an infinitely differentiable function that satisfies

$$\boldsymbol{g}\left(\frac{1}{n}E_{\theta_0}\mathcal{U}_n^+(\theta_0), 0\right) = 0, \tag{3.10}$$

for all $n \geq 1$. From (3.8), we can see that the function $\boldsymbol{g}$ also satisfies

$$\frac{\partial}{\partial x}\boldsymbol{g}\left(\frac{1}{n}E_{\theta_0}\mathcal{U}_n^+(\theta_0), x\right)|_{x=0} = \frac{1}{n}E_{\theta_0}\frac{\partial}{\partial\theta\partial\theta'}\boldsymbol{h}_n(\theta_0) \tag{3.11}$$

where $x = \hat{\theta}_n - \theta_0$. Using the information matrix equality, the right hand side of (3.11) converges to $-\Sigma^{-1}(\theta_0)$ as $n \to \infty$, and, hence, is negative definite for n large because the later is negative definite by Assumption A3, where $\Sigma(\theta)$ is the asymptotic covariance matrix of the PWML estimator. Thus, since $\boldsymbol{g}(n^{-1}E_{\theta_0}\mathcal{U}_n^+(\theta_0), 0) = 0$, the Implicit Function Theorem can be applied to the function $\boldsymbol{g}(.,.)$ at the point $(n^{-1}E_{\theta_0}\mathcal{U}_n^+(\theta_0), 0)$. That is, there is an infinitely differentiable function Ψ, defined near $n^{-1}E_{\theta_0}\mathcal{U}_n^+(\theta_0)$, and that does not depend on n or θ_0 such that

$$\boldsymbol{g}(n^{-1}Z_n^+(\theta_0) + \zeta_n(\theta_0), \Psi(n^{-1}Z_n^+(\theta_0) + \zeta_n(\theta_0))) = 0, \tag{3.12}$$

where Ψ satisfies $\Psi(n^{-1}E_{\theta_0}\mathcal{U}_n^+(\theta_0)) = 0$. Combining (3.6), (3.9), and (3.12) we obtain

$$\inf_{\theta_0\in\Theta_c} P_{\theta_0}(\hat{\theta}_n - \theta_0 = \Psi(n^{-1}\mathcal{U}_n^+(\theta_0) + \zeta_n(\theta_0))) = 1 - o(n^{1-s/2}). \tag{3.13}$$

We now apply Lemma 3.1 with $\mathcal{Z}_n(\theta_0) = \sqrt{n}\Psi(n^{-1}\mathcal{U}_n^+(\theta_0))$ and $\xi_n(\theta_0) = \sqrt{n}(\Psi(n^{-1}\mathcal{U}_n^+(\theta_0) + \zeta_n(\theta_0)) - \Psi(n^{-1}\mathcal{U}_n^+(\theta_0)))$ to obtain

$$|P_{\theta_0}(\sqrt{n}\Psi(n^{-1}\mathcal{U}_n^+(\theta_0) + \zeta_n(\theta_0)) \in C) - P_{\theta_0}(\sqrt{n}\Psi(n^{-1}\mathcal{U}_n^+(\theta_0)) \in C)| = o(n^{-(s-2)/2}), \tag{3.14}$$

uniformly over $\theta_0 \in \Theta_c$ and $C \in C_d$.

We are now ready to approximate the t-statistic $\tau_n(\theta_{0,r})$ by a smooth function of $\mathcal{U}_n^+(\theta_0)/n$. Because the $\tau_n(\theta_{0,r})$ is a function of $\hat{\theta}_n$ (see (2.3)), taking a Taylor expansion of $\tau_n(\theta_{0,r})/\sqrt{n}$ about $\hat{\theta}_n = \theta_0$ to order $s-1$, where the highest-order term involves the expectation of the partial derivatives, we obtain

$$\tau_n(\theta_{0,r}) = \sqrt{n}(\widetilde{\mathcal{A}}(n^{-1}\mathcal{U}_n^+(\theta_0), \hat{\theta}_n - \theta_0) + \widetilde{\rho}_n(\theta_0)), \tag{3.15}$$

where $\widetilde{\mathcal{A}}$ is an infinitely differentiable function that does not depend on θ_0, $\widetilde{\mathcal{A}}(n^{-1}\mathcal{U}_n^+(\theta_0), 0) = 0$ for large n, $\widetilde{\rho}_n(\theta_0)$ is the remainder term in the Taylor expansion, and $\|\widetilde{\rho}_n(\theta_0)\| = O(\|\hat{\theta}_n - \theta_0\|^s)$. Combining (3.11) with (3.15) gives

$$\tau_n(\theta_{0,r}) = \sqrt{n}(\widetilde{\mathcal{A}}(n^{-1}\mathcal{U}_n^+(\theta_0), \Psi(n^{-1}\mathcal{U}_n^+(\theta_0) + \epsilon_n(\theta_0)) + \widetilde{\rho}_n(\theta_0)). \tag{3.16}$$

Again we apply Lemma 3.1 with $\mathcal{Z}_n(\theta_0) = \sqrt{n}\widetilde{\mathcal{A}}(n^{-1}\mathcal{U}_n^+(\theta_0), \Psi(n^{-1}\mathcal{U}_n^+(\theta_0)))$ to obtain

$$\sup_{\theta_0\in\Theta_c, C\in C_d} |P_{\theta_0}(\mathcal{Z}_n(\theta_0) + \widetilde{\rho}_n(\theta_0) \in C) - P_{\theta_0}(\mathcal{Z}_n(\theta_0) \in C)| = o(n^{1-s/2}). \tag{3.17}$$

Define a function Ψ' by $\Psi'(x) = (\widetilde{\mathcal{A}}(x), \Psi(x))$. Then $\Psi'(.)$ is infinitely differentiable and satisfies

$$\Psi'(\tfrac{1}{n}E_{\theta_0}\mathcal{U}_n^+(\theta_0)) = \widetilde{\mathcal{A}}(\tfrac{1}{n}E_{\theta_0}\mathcal{U}_n^+(\theta_0), \Psi(\tfrac{1}{n}E_{\theta_0}\mathcal{U}_n^+(\theta_0))) = \widetilde{\mathcal{A}}(\tfrac{1}{n}E_{\theta_0}\mathcal{U}_n^+(\theta_0), 0) = 0, \tag{3.18}$$

for all n large. Combining (3.17) and (3.18) gives the result of Lemma 3.2. □

Let $\delta > 0$ and let $dist(\theta, \Theta_c) = \inf\{\|\theta - \theta_c\| : \theta_c \in \Theta_c\}$. For Θ_c a compact subset of the parameter space let $\Theta_c^+ = \{\theta \in R^r : dist(\theta, \Theta_c) \leq \delta\}$ be a compact subset of the parameter space Θ that is larger than Θ_c by a radius of δ. Let $B(\theta, \epsilon)$ denote an open ball of radius $\epsilon > 0$ centered at θ. To obtain the coverage probability errors of bootstrap confidence intervals we need to establish asymptotic expansion of the bootstrap t-statistic of our model that holds uniformly for the true parameter lying in the larger set Θ_c^+. The next lemma is essentially Lemma 11 of Andrews et al. (2006) and is used to establish such expansion of the bootstrap t-statistic provided in Lemma 3.4 (b) below which in turn is used in the proof of the main results of this paper given in Theorems 3.6 and 3.7.

Lemma 3.3. Suppose $\sup_{\theta_0\in\Theta_c} P_{\theta_0}(\hat{\theta}_n \notin B(\theta_0, \delta)) = o(n^{-(s-2)/2})$, where Θ_c is a compact subset of Θ and δ is as in the definition of Θ_c^+, and $\{\lambda_n(\theta) : n \geq 1\}$ is a sequence of non-random real functions on Θ_c^+ that satisfies $\sup_{\theta\in\Theta_c^+} |\lambda_n(\theta)| = o(n^{-(s-2)/2})$. Then, for all $\varepsilon > 0$,

$$\sup_{\theta_0\in\Theta_c} P_{\theta_0}(|\lambda_n(\hat{\theta}_n)| > n^{-(s-2)/2}\varepsilon) = o(n^{-(s-2)/2)}).$$

We now determine the error bounds between the parametric bootstrap t-statistic and its formal asymptotic expansion. We first introduce some additional notations. Let $G(x) = P_{\theta_0}(\tau_n(\theta_{0,r}) \leq x)$, $\widetilde{G}(x) = \Phi(x) + \sum_{j=1}^{s-2} n^{-j/2}P_j(\Delta, \bar{\kappa}_n(\theta_0))\Phi(x)$, $G^*(x) = P^*_{\hat{\theta}_n}(\tau_n^*(\hat{\theta}_{n,r}) \leq x)$, and $\widetilde{G}^*(x) = \Phi(x) + \sum_{j=1}^{s-2} n^{-j/2}P_j(\Delta, \bar{\kappa}_n(\hat{\theta}_n))\Phi(x)$, where $\widetilde{G}(x)$ and $\widetilde{G}^*(x)$ are formal Edgeworth expansions of $G(x)$ and $G^*(x)$, respectively.

Lemma 3.4. Suppose Assumptions A1-A10 hold, and let $s \geq 3$ be an integer which satisfies assumptions A2 and A4. Then, for all $\varepsilon > 0$,
(a) $\sup_{\theta_0\in\Theta_c} \sup_{x\in\mathbb{R}} |G(x) - \widetilde{G}(x)| = o(n^{1-s/2})$.
(b) $P_{\theta_0}\left(\sup_{x\in\mathbb{R}} |G^*(x) - \widetilde{G}^*(x)| > n^{1-s/2}\varepsilon\right) = o(n^{1-s/2})$
uniformly over $\theta_0 \in \Theta_c$.
Proof.
(a) We observe that

$$\begin{aligned} |G(x) - \widetilde{G}(x)| \leq{}& |G(x) - P_{\theta_0}\sqrt{n}\Psi(n^{-1}\mathcal{U}_n^+(\theta_0)) \leq x)| \\ &+ |P_{\theta_0}(\sqrt{n}\Psi(n^{-1}\mathcal{U}_n^+(\theta_0)) \leq z) - \widetilde{G}(x)|. \end{aligned} \tag{3.19}$$

The first term of the right hand side of (3.19) is $o(n^{1-s/2})$ because it is just the statement of (3.5) proved in Lemma 3.2, and the second term is also $o(n^{1-s/2})$ because it is shown in the proof of Lemma 3.2 that $n^{1/2}\Psi(n^{-1}\mathcal{U}_n^+(\theta_0))$ possesses an Edgeworth expansion with an error $o(n^{1-s/2})$, and thus (3.19) follows.

Part (b) of the lemma follows from Lemma 3.3 above with

$$\lambda_n(\hat{\theta}_n) \quad = \quad \sup_{x\in\mathbb{R}} |G^*(x) - \widetilde{G}^*(x)|. \tag{3.20}$$

upon checking the conditions of the lemma. The first condition of Lemma 3.3 holds by Assumption A2 and the second condition holds by Lemma 3.2.

Lemma 3.5. Suppose Assumptions A1-A10 hold, and let Θ_c and $s \geq 3$ be as given in Assumptions A2 and A4. Then, for all $\varepsilon > 0$,

$$\sup_{\theta_0\in\Theta_c} P_{\theta_0}(\sqrt{n}\|\bar{\kappa}_n(\hat{\theta}_n) - \bar{\kappa}_n(\theta_0\| > ln(n)\varepsilon) = o(n^{-(s-2)/2}),$$

where $\bar{\kappa}_n(\theta)$ denotes the vector of cumulants of the PWLL function.

Proof
Let $\bar{\kappa}_n(\theta)_\eta$ denote an element of $\bar{\kappa}_n(\theta)$. By a mean value expansion, for all $\theta_0 \in \Theta_c$ and all $\theta \in \Theta_c^+$ such that $\|\theta - \theta_0\| < \delta$ (where δ and Θ_c^+ are as defined in the paragraph preceding Lemma 3.3), $|\bar{\kappa}_n(\theta)_\eta - \bar{\kappa}_n(\theta_0)_\eta| \leq K_n\|\theta - \theta_0\|$, where

$$K_n = \sup_{\theta\in\Theta_c^+, i=1,\ldots,dim(\theta)} |\frac{\partial}{\partial\theta_i}\bar{\kappa}_n(\theta)_\eta|. \tag{3.21}$$

We first show that K_n is a constant that satisfies $\limsup_{n\to\infty} K_n < \infty$. But this holds provided that

$$\sup_{\theta\in\Theta_c} |\frac{\partial}{\partial\theta_i}\kappa_n(\theta)_\eta| = O(n) \tag{3.22}$$

for all $i \leq dim(\theta)$.

We assume for the moment that (3.22) holds and establish the lemma. Let $\gamma > 0$ satisfy

$$\gamma < \frac{\varepsilon}{\sqrt{dim(\bar{\kappa})}\limsup_{n\to\infty} K_n} < \infty,$$

where, $dim(\bar{\kappa})$ denotes the dimension of $\bar{\kappa}_n(\theta)$. We have

$$\begin{aligned}
&\sup_{\theta_0\in\Theta_c} P_{\theta_0}(\sqrt{n}\|\bar{\kappa}_n(\hat{\theta}_n) - \bar{\kappa}_n(\theta_0\| > ln(n)\varepsilon) \\
&\leq \sup_{\theta_0\in\Theta_c} P_{\theta_0}(\sqrt{n}\|\bar{\kappa}_n(\hat{\theta}_n) - \bar{\kappa}_n(\theta_0\| > ln(n)\varepsilon,\ \sqrt{n}\|\theta - \theta_0\| \leq \ln(n)\gamma) \\
&+ \sup_{\theta_0\in\Theta_c} \sqrt{n}\|\theta - \theta_0\| \leq \ln(n)\gamma) \\
&\leq \sup_{\theta_0\in\Theta_c} P_{\theta_0}(\sqrt{dim(\bar{\kappa})}\limsup_{n\to\infty} K_n \sqrt{n}\|\hat{\theta}_n - \theta_0\| > ln(n)\varepsilon,\ \sqrt{n}\|\theta - \theta_0\| \leq \ln(n)\gamma) \\
&+ o(n^{-(s-2)/2}) \\
&= o(n^{-(s-2)/2})
\end{aligned} \tag{3.23}$$

where, the second inequality above uses (3.21) and Assumption A2 and the last equality holds because

$$P_{\theta_0}(\sqrt{dim\bar{\kappa}}\limsup_{n\to\infty} K_n n^{1/2}\|\tilde{\theta}_n - \theta_0\| > ln(n)\varepsilon, n^{1/2}\|\theta - \theta_0\| \leq \ln(n)\gamma) = 0,$$

since $\gamma < \frac{\varepsilon}{\sqrt{dim\bar{\kappa}}\limsup_{n\to\infty} K_n}$. To complete the proof it remains to prove that (3.22) holds. Suppose $\kappa_n(\theta)_\eta$ is a cumulant of order two or greater. By Lemma 6 (c) of Andrew et al. (2006) and the chain rule, $\frac{\partial}{\partial\theta_i}\kappa_n(\theta)_\eta$ is a finite sum of terms of the form

$$C_q\left(\prod_{r=1}^{q}(\mathcal{M}_n\bar{\mathcal{G}}_r\mathcal{M}_n\bar{\mathcal{T}}_r)\right),$$

where $\mathcal{M}_n$, $\mathcal{G}$, and $\mathcal{T}_n$ are as given in equations (2.5-2.7), $\bar{\mathcal{G}}_r$ equals either $\mathcal{G}_{n,\nu(\eta_r)}(\theta)$ or $\frac{\partial}{\partial\theta_i}\mathcal{G}_{n,\nu(\eta_r)}(\theta)$ and $\bar{\mathcal{T}}_r$ equals either $\mathcal{T}_n((2\pi)^{-2}f_\theta^{-1})$ or $\frac{\partial}{\partial\theta_i}\mathcal{T}_n((2\pi)^{-2}f_\theta^{-1})$.

We observe that $\mathcal{G}_{n,\nu(\eta_r)}(\theta)$ and $\frac{\partial}{\partial\theta_i}\mathcal{G}_{n,\nu(\eta_r)}(\theta)$ have the same form because they are both partial derivatives of $-\frac{1}{2}\mathcal{T}_n((2\pi)^{-2}f_\theta^{-1})$, (see (2.7)). It follows that $\frac{\partial}{\partial\theta_i}\kappa_n(\theta)_\eta$ has the same form as $\kappa_n(\theta)_\eta$ itself. (3.22) now follows from Lemma 6 (c) of Andrews et al. (2006). □

We now state the main bootstrap coverage probability results of this paper.

Theorem 3.6. Suppose Assumptions A1-A10 hold, and let $s \geq 3$ be an integer which satisfies assumptions A2 and A4. Let Θ_c be a compact subset of Θ as stated in Assumptions A2 and A9. Then, for $s = 5$ we have

$$\sup_{\theta_0 \in \Theta_c} |P_{\theta_0}(\theta_0 \in \mathcal{I}_2(\hat{\theta}_n)) - (1-\alpha)| = o(n^{-3/2}\ln(n)). \tag{3.24}$$

Proof.

We first note that $P_{\theta_0}(\theta_{0,r} \in \mathcal{I}_2(\hat{\theta}_n)) = P_{\theta_0}(|t_n(\theta_{0,r})| \leq z^*_{t,\alpha})$. To prove (3.24) it suffices to show that the later equals $1-\alpha+o(n^{-3/2}\ln(n))$ uniformly over $\theta_0 \in \Theta_c$.

Let $H_5(\hat{\theta}_n) = 1 + \sum_{i=1}^{3} n^{-i/2} P_j(\Delta, \bar{\kappa}_n(\hat{\theta}_n))$ where $P_j(\Delta, \bar{\kappa}_n(\theta))$ is as defined in the paragraph preceding Lemma 3.1. Using Lemma 3.4 (b) with $s = 5$ and the fact that $P_{\hat{\theta}_n}(|t^*_n(\hat{\theta}_{n,r})| \leq z) = P_{\hat{\theta}_n}(t^*_n(\hat{\theta}_{n,r}) \leq z) - P_{\hat{\theta}_n}(t^*_n(\hat{\theta}_{n,r}) \leq -z)$, we obtain for all $\varepsilon > 0$:

$$\sup_{\theta_0 \in \Theta} P_{\theta_0}\Big(\sup_{z \in \mathbb{R}} |P^*_{\hat{\theta}_n}(|t^*_n(\hat{\theta}_{n,r})| \leq z) - [H_5(\hat{\theta})](\Phi(z) - \Phi(-z))| > n^{-3/2}\varepsilon\Big) = o(n^{-3/2}). \tag{3.25}$$

We observe that because $P_j(\Delta, \bar{\kappa}_n(\hat{\theta}_n))(\Phi(z))$ is an even function for odd j (see the paragraph preceding Lemma 3.1), it follows that $P_1(\Delta, \bar{\kappa}_n(\hat{\theta}_n))(\Phi(z) - \Phi(-z)) = P_3(\Delta, \bar{\kappa}_n(\hat{\theta}_n))(\Phi(z) - \Phi(-z)) = 0$. Therefore, (3.25) above is equivalent to:

$$\sup_{\theta_0 \in \Theta} P_{\theta_0}\left(\sup_{z \in \mathbb{R}} |P^*_{\hat{\theta}_n}(|t^*_n(\hat{\theta}_{n,r})| \leq z) - [1 + n^{-1}P_2(\Delta, \bar{\kappa}_n(\hat{\theta}_n))](\Phi(z) - \Phi(-z))| > n^{-3/2}\varepsilon\right)$$
$$= o(n^{-3/2}). \tag{3.26}$$

Likewise, using Lemma 3.4 (a) and Lemma 3.5 we obtain, respectively, (3.27) and (3.28) below

$$\sup_{\theta_0 \in \Theta_c} \sup_{z \in \mathbb{R}} |P_{\theta_0}(|t_n(\theta_{0,r})| \leq z) - [1 + n^{-1}P_2(\Delta, \bar{\kappa}_n(\theta_0)))](\Phi(z) - \Phi(-z))| = o(n^{-3/2}), \tag{3.27}$$

$$\sup_{\theta_0 \in \Theta_c} P_{\theta_0}\left(\sup_{z \in \mathbb{R}} |[P_2(\Delta, \bar{\kappa}_n(\hat{\theta}_n)) - P_2(\Delta, \bar{\kappa}_n(\theta_0))](\Phi(z) - \Phi(-z))| > n^{-1/2}\ln(n)\varepsilon\right)$$
$$= o(n^{-3/2}). \tag{3.28}$$

Combining (3.26), (3.27), and (3.28) above we obtain:

$$\sup_{\theta_0 \in \Theta} P_{\theta_0}\left(\sup_{z \in \mathbb{R}} |P^*_{\hat{\theta}_n}(|t^*_n(\hat{\theta}_{n,r})| \leq z) - P_{\theta_0}(|t_n(\theta_{0,r})| \leq z)| > n^{-3/2}\ln(n)\varepsilon\right) = o(n^{-3/2}). \tag{3.29}$$

Now, let $\mathcal{S}(z) = H_5(\hat{\theta}_n)\Phi(z)$. $\mathcal{S}$ is essentailly an Edgeworth expansion of the bootstrap t-statistic $t^*_n(\hat{\theta}_n)$ given in Lemma 3.4 (b). Since $\mathcal{S}$ is continuous in z, there exists $\tilde{z}_{\tau,\alpha}$ such that $\mathcal{S}(\tilde{z}_{\tau,\alpha}) = 1-\alpha$. Using this and the definition of $z^*_{\tau,\alpha}$ (see the last paragraph of section (2.2)), we have:

$$\begin{aligned} |P^*_{\hat{\theta}_n}(t^*_n(\hat{\theta}_{n,r}) \leq \tilde{z}_{\tau,\alpha}) - \mathcal{S}(\tilde{z})| &= |P^*_{\hat{\theta}_n}(t^*_n(\hat{\theta}_{n,r}) \leq \tilde{z}_{\tau,\alpha}) - (1-\alpha)| \\ &\geq |P^*_{\hat{\theta}_n}(t^*_n(\hat{\theta}_{n,r}) \leq z^*_{\tau,\alpha}) - (1-\alpha)|. \end{aligned} \tag{3.30}$$

Moreover, by Lemma 3.4 (b) we obtain:

$$\sup_{\theta_0 \in \Theta_c} P_{\theta_0}\left(\sup_{z \in \mathbb{R}} |P^*_{\hat{\theta}_n}(t^*_n(\hat{\theta}_{n,r}) \leq z^*_{\tau,\alpha}) - (1-\alpha)| > n^{-3/2}\varepsilon\right) = o(n^{-3/2}). \tag{3.31}$$

Taking $z = z^*_{\tau,\alpha}$ in (3.29) and combining it with (3.31) yields:

$$\sup_{\theta_0 \in \Theta_c} P_{\theta_0}(|1 - \alpha - P_{\theta_0}(t_n(\theta_{0,r}) \leq z^*_{\tau,\alpha})| > n^{-3/2}\ln(n)\varepsilon) = o(n^{-3/2}). \tag{3.32}$$

Hence for large n, (3.32) reduces to:

$|1-\alpha-P_{\theta_0}(t_n(\theta_{0,r})\le z^*_{\tau,\alpha})|<n^{-3/2}\ln(n)\varepsilon$, which establishes (3.24). □

Theorem 3.7. Suppose Assumptions A1-A10 hold, and let $s\ge 3$ be an integer which satisfies assumptions A2 and A4. Let Θ_c be a compact subset of Θ as stated in Assumptions A2 and A9. Then, for $s=4$ we have

$$\sup_{\theta_0\in\Theta_c}|P_{\theta_0}(\theta_0\in\mathcal{I}_1(\hat\theta_n))-(1-\alpha)|=o(n^{-1}\ln(n)). \tag{3.33}$$

Proof. The proof is analogous to that of Theorem 3.6 above and therefore some details are omitted.
Let $H_4(\hat\theta_n)=1+\sum_{i=1}^2 n^{-i/2}P_j(\Delta,\bar\kappa_n(\hat\theta_n))$. By Lemma 3.4 (b) with $s=4$ we obtain for all $\varepsilon>0$:

$$\sup_{\theta_0\in\Theta}P_{\theta_0}\Big(\sup_{z\in\mathbb{R}}|P^*_{\hat\theta_n}(|t^*_n(\hat\theta_{n,r})|\le z)-[H_4(\hat\theta)](\Phi(z)-\Phi(-z))|>n^{-1}\varepsilon\Big)=o(n^{-1}). \tag{3.34}$$

Moreover, by Lemma 3.4 (a) and Lemma 3.5 with $s=4$, respectively, and using the evenness of $P_j(\Delta,\bar\kappa_n(\theta))(\Phi(z)-\Phi(-z))$ for $j=1$ we obtain:

$$\sup_{\theta_0\in\Theta_c}\sup_{z\in\mathbb{R}}|P_{\theta_0}(|t_n(\theta_{0,r})|\le z)-H_4(\theta_0)(\Phi(z)-\Phi(-z))|=o(n^{-1}), \tag{3.35}$$

and

$$\sup_{\theta_0\in\Theta_c}P_{\theta_0}\left(\sup_{z\in\mathbb{R}}|[P_2(\Delta,\bar\kappa_n(\hat\theta_n))-P_2(\Delta,\bar\kappa_n(\theta_0))](\Phi(z)-\Phi(-z))|>n^{-1/2}\ln(n)\varepsilon\right)$$
$$=o(n^{-1}). \tag{3.36}$$

Combining (3.34), (3.35), and (3.36) above we obtain:

$$\sup_{\theta_0\in\Theta}P_{\theta_0}\left(\sup_{z\in\mathbb{R}}|P^*_{\hat\theta_n}(|t^*_n(\hat\theta_{n,r})|\le z)-P_{\theta_0}(|t_n(\theta_{0,r})|\le z)|>n^{-1}\ln(n)\varepsilon\right)=o(n^{-1}). \tag{3.37}$$

Observing the similarities between equations (3.37) above and (3.29) in the proof of Theorem 3.6 we can see that the remainder of the proof follows analogously to establish (3.33). □

Comments: 1. For $s=3$ it can be analogously shown that the error in (3.24) and (3.33) is $o(n^{-1/2})$.

2. Theorems 3.6 and 3.7 provide coverage probability errors for parametric bootstrap confidence intervals as well as boostrap tests based on the t-statistic $t_n(\theta_{h,r})$. For parametric values θ_0 for which $\theta_{0,r}=\theta_{h,r}$, we have $P_{\theta_0}(\theta_{0,r}\in\mathcal{I}_2(\hat\theta_n))=P_{\theta_0}(|t_n(\theta_{h,r})|\le z^*_{\tau,\alpha})$. Similar results hold for upper level bootstrap confidnce intervals and tests. Therefore, Theorems 3.6 and 3.7 also establish bounds, respectively, on the rejection error rates of symmetric two-sided and upper one-sided bootstrap tests.

References

Aga, M., & Sun, T. (2007). Parametric Bootstrap Confidence Intervals for Linear Regression Processes with long-memory Errors. *Sankhya: The Indian Journal of Statistics, 69*(4), 615-634. https://www.jstor.org/stable/25664582

Aga, M. (2021). Edgeworth Expansion for the Whittle Maximum Likelihood Estimator of Linear Regression Processes with Long Memory Residuals. *International Journal of Statistics and Probability, 10*(4), 2021. https://doi.org/10.5539/ijsp.v10n4p119

Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis* (2nd ed.). New York: Wiley.

Andrews, D., Lieberman, O., & Marmer, V. (2006). Higher-order Improvements of the Parametric Bootstrap for Long-memory Gaussian Processes. *Journal of Econometrics, 133*(2), 673-702. https://doi.org/10.1016/j.jeconom.2005.06.010

Beran, R. J. (1982). Estimating Sampling Distributions: The Bootstrap and Competitors. *Annals of Statistics, 10*, 212-225.

Beran, R. J. (1994). *Statistics for Long Memory Process.* Chapman and Hall. New York.

Bickel, P. J., & Freedman, D. A. (1981). Some Asymptotic Theory for the Bootstrap. *Annals of Statistics, 9*, 1196-1217. https://doi.org/10.1214/aos/1176345637

Bickel, P. J., & Freedman, D. A. (1984). Asymptotic Normality and the Bootstrap in Stratified Sampling. *Annals of Statistics, 12*, 470-482. https://doi.org/10.1214/aos/1176346500

Choy, K., & Taniguchi, M. (2001). Stochastic regression model with dependent disturbances. *Journal of Time Series Analysis, 22*, 175-196. https://doi.org/10.1111/1467-9892.00218

Dahlhaus, R. (1989). Efficient Parameter Estimation for Self-similar Processes. *Annals of Statistics, 17*(4), 1749-1766. https://doi.org/10.1214/aos/1176347393

Dahlhaus, R. (1995). Efficient location and regression estimation for long-range dependent regression models. *Annals of Statistics, 23*, 1029-1047. https://doi.org/10.1214/aos/1176324635

Diaconis, P., & Efron, B. (1983). Computer-intensive Methods in Statistics. *Scientific American, 248*, 116-130. https://doi.org/10.1038/scientificamerican0583-116

Efron, B. (1979). Bootstrap Methods: another look at the jackknife. *Annals of Statistics, 7*(1), 1-26. https://doi.org/10.1214/aos/1176344552

Efron, B. (1982). The Jackknife, the bootstrap and other resampling plans. *Volume 38 of CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM*. https://doi.org/10.1137/1.9781611970319

Efron, B., & Gong, G. (1983). A Leisurely look at the Bootstrap, the Jackknife and Cross-validation. *American Statistician, 37*, 36-48. https://doi.org/10.2307/2685844

Efron, B., & Tibshirani, R. (1986). Bootstrap Measures for Standard Errors, Confidence Intervals and Other Measures of Statistical Accuracy. *Statistical Science, 1*, 54-77. https://doi.org/10.1214/ss/1177013815

Franco, G., & Reisen V. (2004). Bootstrap techniques in semiparametric estimation methods for ARFIMA models: A comparison study. *Computational Statistics, 19*, 243-259.

Hall, P. (1986). On the Bootstrap and Confidence Intervals. *Annals of Statistics, 14*, 1431-1452.

Hall, P. (1988). Theoretical Comparison of Bootstrap Confidence Intervals. *Annals of Statistics, 16*, 927-985, 1988. https://doi.org/10.1214/aos/1176350168

Hall, P., & Martin, M. A. (1988). On Bootstrap Resampling and Iteration. *Biometrika, 75*, 661-671. https://doi.org/10.1093/biomet/75.4.661

Hall, P. (1992). The Bootstrap and Edgeworth Expansion. *Springer Series in Statistics*, 39-81. https://doi.org/10.1007/978-1-4612-4384-7_2

Hidalgo, J. (2003). An alternative bootstrap to moving blocks for time series regression models. *Journal of Econometrics, 117*, 369-399. https://doi.org/10.1016/S0304-4076(03)00154-4

Hidalgo, J. (2021). Bootstrap Long Memory Process in the Frequency Domain. *Annals of Statistics, 15*(3), 1407-1435. https://doi.org/10.1214/20-AOS2006

Hurst, H. E. (1951). The Long-Term Storage Capacity of Reservoir. *Transactions of the American Society of Civil Engineers, 116*, 770-808. https://doi.org/10.1017/S0266466604203036

Koul, H. L., Baillie, R. T., & Surgailis, D. (2004). Regression Model Fitting with a Long-memory covariate Process. *Econometric Theory, 20*, 485-512. https://www.jstor.org/stable/3533529

Ivanov, A., & Leonenko, N. (2008). Robust Estimators in Non-linear Regression Models with Long-Range Dependence. *Springer Optimization and Its Applications book series, 28*, 193-221.

Kunsch, H. (1986). Discrimination between monotonic trends and long-range dependence. *Journal of Applied Probability, 23*, 1025-1030.

Lahiri, S. N. (1993). On the moving block bootstrap under long range dependence. *Statistics and Probability Letters, 18*, 405-413. https://doi.org/10.1016/0167-7152(93)90035-H

Lieberman, O., Rousseau, J., & Zucker, M. (2003). Valid Edgeworth Expansions For the Maximum Likelihood Estimator of the Parameter of a Stationary, Gaussian, Strongly Dependent Processes. *Annals of Statistics, 31*(2), 586-612.

Maldelbrot, B. B., & Van Ness, J. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Review, 10*, 422-437. https://doi.org/10.1137/1010093

Martin, M. A. (1990). On the Bootstrap Iteration for Coverage Correction in Confidence Intervals. *Journal of American Statistical Association, 85*, 1105-1118. https://doi.org/10.1080/01621459.1990.10474982

Navidi, W. (1989). Edgeworth Expansions for Bootstrapping in Regression Models. *Annals of Statistics, 17*, 1472-1478. https://doi.org/10.1214/aos/1176347375

Robinson, P. M., & Hidalgo, F. J.(1997). Time series regression with long-range dependence. *Annals of Statistics, 25*, 77-104. https://doi.org/10.1214/aos/1034276622

Sibbertsen P. (2001). S-Estimation in the linear regression model with long-memory error terms under trend. *Journal of Time Series Analysis, 22*, 353-363. https://doi.org/10.1111/1467-9892.00228

Singh, K. (1981). On the Asymptotic Accuracy of Efron's Bootstrap. *Annals of Statistics, 9*, 1187-1195. https://doi.org/10.1214/aos/1176345636

Yajima, Y. (1988). On estimation of a regression model with long memory errors. *Annals of Statistics, 16*, 791-807. https://doi.org/10.1214/aos/1176350837

Yajima, Y. (1991). Asymptotic properties of the LSE in a regression model with long-memory stationary errors. *Annals of Statistics, 19*, 158-177. https://doi.org/10.1214/aos/1176347975

Copyrights