Some Improvements of the Bootstrap over the Delta Method Probability Errors for Whittle Estimators

Mosisa Aga

Correspondence: Department of Mathematics, Auburn University at Montgomery, Montgomery, AL, USA

Received: July 11, 2024 Accepted: August 27, 2024 Online Published: October 23, 2024

doi:10.5539/jmr.v16n5p10 URL: https://doi.org/10.5539/jmr.v16n5p10

Abstract

The purpose of this paper is to compare the coverage probability errors of the parametric bootstrap with that of the delta method for the covariance parameters of a regression model with auto-regressive fractionally integrated moving average (ARFIMA) errors. We consider the coverage probability errors of both confidence intervals (CIs) and tests based on the the plug-in Whittle maximum likelihood (PWML) estimators. We first show 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 the two-sided delta method and parametric bootstrap confidence intervals on the plug-in Whittle likelihood estimator of the covariance parameter are shown to be $O(n^{-1})$ and $o(n^{-3/2} \ln n)$, respectively, where n is the sample size. Next, we show that those of the one-sided confidence intervals are shown to be $O(n^{-1/2})$ and $o(n^{-1} \ln n)$, respectively. These results show that for both one-sided and two-sided confidence intervals and tests, the bootstrap provides a significant improvement over that of the delta method.

Keywords: parametric bootstrap, t-statistic, linear regression, Whittle estimator, long memory

2020 Mathematics Subject Classification: 62M10,62F12, 60E05, 60G05, 60G10

1. Introduction

Since the introduction of the bootstrap by Efron (1979), most research has focused on its application to independent and identically distributed (iid) data or short-memory time series in nonparametric settings. Over the years, Efron and other researchers have broadened the bootstrap's applications to include confidence intervals, hypothesis testing, regression models, and more complex problems. Key contributions in this expansion are found in works such as Efron (1982), Efron et al. (1982), Diaconis et al. (1983), and Efron et al. (1986).

The volume of bootstrap-related research surged exponentially during the 1980s and 1990s, with significant findings from Singh (1981), Bickel et al. (1981, 1984), Beran (1982), Martin (1990), Hall (1986, 1988), Hall et al. (1988), and Navidi (1989), among others. Today, bootstrapping is widely used to approximate the distributions of various statistics in time series analysis, particularly in cases of short-term persistence. Despite its growing use with long-memory time series over the past two decades, there remains a limited amount of theoretical justification for its validity in the context of long-memory processes.

Time series exhibiting long-range dependence are prevalent in various fields such as astronomy, hydrology, economics, and finance, where correlations between observations diminish very slowly over time (see Hurst (1951); Mandelbrot et al. (1968); Beran (1994)). While most literature on linear regression models focuses on data with short-memory error components, there has also been extensive research on regression models with long-range dependent errors. Key studies in this area include Kunsch (1986), Yajima (1988, 1991), Dahlhaus (1995), Robinson et al. (1997), Sibbertsen (2001), Koul et al. (2004), Choy et al. (2001), and Ivanov et al. (2008).

Yajima (1988) examined the estimation of regression parameters using the least squares estimator (LSE) and investigated the correlation structure of error terms through residuals obtained from the LSE. In a subsequent study, Yajima (1991) derived several asymptotic results for least squares error estimators and best linear unbiased estimators. Dahlhaus (1995) focused on estimating coefficients in regression models with long-range dependent errors. Robinson et al. (1997) explored models with stochastic regressors, where components of the design matrix follow stationary processes and both errors and stochastic regressors exhibit long-range dependence.

Research aimed at justifying the use of bootstrap methods for long memory time series includes works by Lahiri (1993), Hidalgo (2003), Franco et al. (2004), Andrews et al. (2006), Aga et al. (2007), Hidalgo (2021), and several more recent studies. Lahiri (1993) demonstrated that the moving block bootstrap can effectively approximate the distribution of the normalized sample mean for certain long-range dependent data, provided that the underlying statistic is asymptotically normal. His subsequent extensive work has significantly advanced the application of the block bootstrap technique.

Hidalgo (2003) introduced an alternative to the moving block bootstrap method for estimating parameters in time-series regression models. Franco et al. (2004) conducted a comparative study of various bootstrap techniques used in semiparametric estimation for Auto-Regressive Fractionally Integrated Moving Average (ARFIMA) models through Monte Carlo simulations. Andrews et al. (2006) analyzed the coverage probability errors for delta method and parametric bootstrap confidence intervals for both plug-in maximum likelihood (PML) and plug-in Whittle maximum likelihood (PWML) estimators of covariance parameters in stationary, long-memory, and Gaussian time series. Aga et al. (2007) extended this research to linear regression models with Gaussian, stationary, and long-memory errors, examining the coverage probability errors of the parametric bootstrap for PML estimators of the model's covariance parameters. The current article explores the delta method and parametric boostrap coverage probability errors when using the Whittle maximum likelihood estimators to approximate PML estimators of the model parameters.

In the 1980s, following the advent of the bootstrap, there was extensive research into applying bootstrapping methods to regression models. Pioneering contributions by Bickel et al. (1981), Freedman (1981), Shorack (1982), Freedman et al. (1984), Weber (1984), Wu (1986), Shao (1988), Efron (1991), and others rigorously expanded the bootstrap's applicability to various facets of regression analysis. More recently, Eck (2018) extended Freedmans (1981) work to multivariate linear regression. These studies predominantly addressed cases where the error terms are independent and focused mainly on regression coefficient parameters.

Research into bootstrapping regression models with dependent errors includes notable works by Stute (1995), McKnight et al. (2000), Aga (2015, 2022, 2024a) among others. Stute (1995) and McKnight et al. (2000) examined bootstrapping techniques for regression coefficients in models with short-memory autoregressive (AR) errors. Aga (2015) and Aga (2022) focused on Edgeworth expansions of the parametric bootstrap t-statistic and the t-statistic, respectively, for PML estimators of these models. Aga (2024a) provided an Edgeworth expansion for the bootstrap t-statistic of the plug-in Whittle maximum likelihood (PWML) estimator in linear regression models with long-memory errors.

The current work builds upon Aga (2024a) by further developing the Edgeworth expansions of both the delta method and the parametric bootstrap t-statistic of the Whittle estimator. It also expands upon Aga et al. (2007) and explores the bootstrap and delta method for two-sided and upper one-sided confidence interval probability errors of the PWML estimator in the same context of regression models with long-memory errors. The primary objective of this study is to mathematically justify that the bootstrap provides improved probability error estimates over the traditional delta method confidence intervals and tests when Whittle estimators are used in regression models with long-memory errors.

Consider a linear regression model

$$Y_t = X_t'\beta + \mathcal{E}_t, t \geq 1, \tag{1.1}$$

where $\beta = (\beta_1, \beta_2, ..., \beta_k)'$ is a k vector of deterministic but unknown real numbers, $\{X_t = (x_{t1}, x_{t2}, ..., x_{tk})' \in \mathbb{R}^k, t \geq 1, k \geq 1\}$ are non-stochastic regressors. In this paper, the error terms $\{\mathcal{E}_t, t \geq 1\}$ are specifically assumed to be the ARFIMA (p,d,q) long memory processes introduced by Hosking (1980) and Grangr et al. (1981), defined by

$$\Phi(B)\mathcal{E}_t = \Psi(B)(1 - B)^{-d}\epsilon_t, \tag{1.2}$$

where B is the back shift operator, $\Phi(B) = 1 + \phi_1 B + ... + \phi_p B^p$ and $\Psi(B) = 1 + \psi_1 B + ... + \psi_q B^q$ are autoregressive and moving-average operators, $\Phi(B)$ and $\Psi(B)$ have no common roots, $d \in (0, \frac{1}{2})$, $(1 - B)^{-d}$ is defined by the binomial formula $(1 - B)^{-d} = \sum_{j=0}^{\infty} \eta_j B^j$, where

$$\eta_j = \frac{\Gamma(j + d)}{\Gamma(j + 1)\Gamma(d)}, \tag{1.3}$$

and Γ is the gamma function, and ϵ_t is a white noise sequence with finite variance σ^2.
Let $Y = (Y_1, Y_2, \ldots, Y_n)'$ be an observed sample of size n and $\mathcal{E} = (\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_n)'$ be the corresponding error terms, where for each $t = 1, 2, ..., n$, $Y_t = X_t'\beta + \mathcal{E}_t$ is as in (1.1). Observe that because the $X_t'\beta$ component in (1.1) is deterministic, it follows that the covariance matrix of Y is the same as that of $\mathcal{E}$.
The ordinary least squares estimate (OLSE) $\hat{\beta}$ is given by

$$\hat{\beta} = \left(\hat{\beta}_1, \ldots, \hat{\beta}_k\right)' = (X'X)^{-1}X'Y = P^{-1}X'Y \tag{1.4}$$

where $P = X'X$ is a $k \times k$ matrix. One can show that P is positive definite and symmetric. The $n \times k$ design matrix X is assumed to have rank k and in practice $k < n$. Let $\mu = (\mu_1, \ldots, \mu_n)$ be the true mean of Y. Then, an estimator of μ is $\hat{\mu} = (\hat{\mu_1}, \ldots, \hat{\mu_n}) = (X_1'\hat{\beta}, ..., X_n'\hat{\beta})$. The spectral density function f_θ of the ARFIMA (p,d,q) process (1.2) is given by

$$f_\theta(\lambda) = \frac{\sigma^2}{2\pi} \frac{|\Psi(e^{i\lambda})|^2}{|\Phi(e^{i\lambda})|^2} |1 - e^{i\lambda}|^{-2d}, \tag{1.5}$$

where $\theta = (\sigma^2; d; \phi_1, ..., \phi_p; \psi_1, ..., \psi_q)$. [Brockwell et al. (1991), Equation (13.2.18).] The $n \times n$ covariance matrix corresponding to $f_\theta(\lambda)$, denoted by $\Sigma_n(f_\theta)$ and has the (j, k) element defined by

$$\Sigma_n(f_\theta)_{j,k} = \int_{-\pi}^{\pi} e^{i(j-k)\lambda} f_\theta(\lambda) d\lambda = \int_{-\pi}^{\pi} e^{i(j-k)\lambda} \frac{\sigma^2}{2\pi} \frac{|\Psi(e^{i\lambda})|^2}{|\Phi(e^{i\lambda})|^2} |1 - e^{i\lambda}|^{-2d} d\lambda. \tag{1.6}$$

The log-likelihood function is

$$\Lambda(\theta, \mu) = -\frac{n}{2} \ln(2\pi) - \frac{1}{2} \ln(\det(\Sigma_n(f_\theta))) - \frac{1}{2}(Y - \mu)'\Sigma_n^{-1}(f_\theta)(Y - \mu). \tag{1.7}$$

To find the exact maximum likelihood estimator (MLE) in (1.7), one needs to compute the inverse of the $n \times n$ covariance matrix $\Sigma_n(f_\theta)$. For large values of n, this is numerically a cumbersome task to handle. Especially for values of d near $\frac{1}{2}$, storing the whole covariance matrix requires excessive computer memory. Moreover, evaluation of the inverse of $\Sigma_n(f_\theta)$ may be unstable. For instance, Beran (1994), pp. 108 shows that the determinant of the correlation matrix of a fractionally Gaussian noise with $d = 0.1, 0.2, 0.3$, and 0.4 is equal to $0.06, 2.2(10^{-6}), 2.7(10^{-16})$, and $5.0(10^{-39})$, respectively.

To alleviate this problem, an approximate likelihood estimator was introduced by Whittle (1953, 1957) and has been widely used since then. In our current set up, the two terms in (1.7) that depend on θ and pose computational difficulty are the logarithm of the determinant of the covariance matrix $\ln\det(\Sigma_n(f_\theta))$ and the quadratic form $(Y - \mu)'(\Sigma_n^{-1}(f_\theta))(Y - \mu)$.

Using the fact that

$$\frac{1}{n} \ln(\det(\Sigma_n(f_\theta))) \to \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln(f_\theta(\lambda)) d\lambda \tag{1.8}$$

and

$$\Sigma_n((2\pi)^{-2} f_\theta^{-1}) \to \Sigma_n^{-1}(f_\theta) \tag{1.9}$$

as $n \to \infty$ (Beran (1994)), and replacing $\frac{1}{n} \ln\det(\Sigma_n(f_\theta))$ by $\frac{1}{2\pi} \int_{-\pi}^{\pi} \ln(f_\theta(\lambda)) d\lambda$ and $\Sigma_n^{-1}(f_\theta)$ by $\Sigma_n((2\pi)^{-2} f_\theta^{-1})$, the log-likelihood function (1.7) can now be approximated by

$$\Lambda_W(\theta, \mu) = -\frac{n}{2} \ln(2\pi) - \frac{n}{4\pi} \int_{-\pi}^{\pi} \ln(f_\theta(\lambda)) d\lambda - \frac{1}{2}(Y - \mu)'\Sigma_n((2\pi)^{-2} f_\theta^{-1})(Y - \mu). \tag{1.10}$$

We refer to $\Lambda_W(\theta, \hat{\mu})$, where $\hat{\mu}$ is replaced for μ in (1.10) above, as the *plug-in Whittle log-likelihood* (PWLL) function. Let $Q_n = XP^{-1}X'$ and let $\mathcal{M}_n = I_n - Q_n$, where I_n is the $n \times n$ identity matrix. One can easily verify that the matrices $\mathcal{M}_n$ and Q_n have the following properties: (a) Both $\mathcal{M}_n$ and Q_n are symmetric. (b) $Y'\mathcal{M}_n = (Y - \hat{\mu})'$. (c) If $U = Y - \mu$, then $\mathcal{M}_n Y = \mathcal{M}_n U$. (d) There exists an $n \times p$ matrix $\mathcal{B}$ such that

$$Q_n = \mathcal{B}\mathcal{B}'. \tag{1.11}$$

In section 2.1, we impose a condition on the design matrix through the matrix $\mathcal{B}$ given in (1.11).
Using the properties (a) through (d), the PWLL function can now be written as

$$\Lambda_W(\theta, \hat{\mu}) = -\frac{n}{2} \ln(2\pi) - \frac{n}{4\pi} \int_{-\pi}^{\pi} \ln((f_\theta)(\lambda)) - \frac{1}{2} Y'\mathcal{M}_n \Sigma_n((2\pi)^{-2} f_\theta^{-1})\mathcal{M}_n Y \tag{1.12}$$

On the other hand, the last term in (1.10) can be approximated as

$$\frac{1}{2}(Y - \mu)'\Sigma_n((2\pi)^{-2} f_\theta^{-1})(Y - \mu) \approx \frac{n}{4\pi} \int_{-\pi}^{\pi} f_\theta^{-1}(\lambda) I(\lambda) d\lambda, \tag{1.13}$$

and therefore, the PWLL function becomes

$$\Lambda_W(\theta, \hat{\mu}) = -\frac{n}{2} \ln(2\pi) - \frac{n}{4\pi} \int_{-\pi}^{\pi} (\ln(f_\theta(\lambda) + f_\theta^{-1}(\lambda) I_n(\lambda)) d\lambda, \tag{1.14}$$

where $I_n(\lambda) = |\frac{1}{2n\pi} \sum_{j=1}^{n} e^{ij\lambda}(Y_j - \hat{\mu}_j)|^2$ is the periodogram (see Aga (2024a), equations (1.9) and (1.10)).

By definition, the Plug-in Whittle Maximum Likelihood (PWML) estimator, $\hat{\theta}_n$, solves the equation

$$\int_{-\pi}^{\pi} D_{\theta_r}(\ln f_\theta(\lambda) + f_\theta^{-1}(\lambda) I_n(\lambda)) d\lambda = 0 \tag{1.15}$$

for $r = 1, ..., p+q+1$, where $D_{\theta_r} = \frac{\partial}{\partial \theta_r}$ and $p+q+1 = dim(\theta)$ as in (1.5) above. Andrews et al. (2006) have analyzed the symmetric two-sided and one-sided confidence intervals of both delta method and parametric bootstrap based on the PML and the PWML estimators of the variance parameter for a more general long memory process than the error component $\{\mathcal{E}_t, t \geq 1\}$ given above, without the linear regression component. This paper builds on the work of Andrews et al. (2006) and analyzes the coverage probability errors both the delta method and the parametric bootstrap of the PWML estimator of our current model by imposing an additional condition on the regression coefficients.

The remainder of the paper proceeds as follows. Section 2 presents the background assumptions and gives a brief description of delta method and parametric bootstrap confidence intervals and tests. Section 3 presents and proves two lemmas and two main theorems on the one sided and two sided confidence interval probability errors of the delta method and parametric bootstrap of the PWML estimator of our model. Section 4 presents a brief conclusion.

2. Assumptions and Background Preliminaries

2.1 Assumptions

In this subsection we describe the assumptions for which the results of this paper hold. There are a number of standard assumptions that are routinely stated in asymptotic theory for long memory process which are also needed in our current analysis. For example, Assumptions A4-A9 of Aga (2024a) are also stated in Andrews et al. (2006), Aga et al. (2007), Dahlhaus (1989), and Lieberman et al. (2003) among others and are needed to control the behavior of the spectral density function, its inverse, and their derivatives in the neighborhood of the origin. Roughly speaking, some of these assumptions state that both f_θ and f_θ^{-1} are continuous with respect to (λ, θ) and bounded around the origin, and others state that the partial derivatives with respect to λ and θ of both f_θ and f_θ^{-1} are continuous and bounded near the origin. These assumptions depend on a positive integer $s \geq 1$ that indexes the order of the Whittle Log-likelihood Derivatives (WLLDs) presented in section (2.4) of this paper.

For brevity, we omit the above standard assumptions from the list of assumptions stated below. In this subsection we present only those assumptions that are referenced in the proofs of the lemmas and theorems of this paper. It is under the assumption that all other standard assumptions mentioned above hold that the lemmas and theorems of this paper are presented without actually stating them.

Assumptions A1-A3 below 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 (2.14) below are required to be positive definite by A3. Assumption A4 states that both f_θ and f_θ^{-1} can be differentiated $s+1$ times inside the integral sign for some integer $s \geq 3$. Assumption A5 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 (and to that of Aga (2007) and Aga (2024a)). In particular, Lemma 3.2 and 3.3 of Aga (2024a) use 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 was used to prove Theorem 3.4 of Aga (2024a) (or Lemma 3.1 of our current paper).

A1. The parameter space Θ is a subset of $\mathbb{R}^r$ where $r = dim(\theta_0) = p+q+1$ 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 (2.14) below 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. The design matrix X is chosen in such a way that for the matrix

$$\mathcal{B} = (e_{ij}), i = 1, ..., n; j = 1, ..., p \tag{2.1}$$

defined by (1.11) above, there exists a constant $M < \infty$ such that $|e_{ij}| \leq \frac{M}{\sqrt{n}}$ for $1 \leq i \leq n; 1 \leq j \leq p$.

2.2 A Brief Description of Delta Method CIs and Tests

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, where θ_0 is the true parameter, and $\hat{\theta}_n$ is the PWML estimator. The asymptotic covariance matrix of a consistent PWML estimator $\hat{\theta}_n$ is $\mathcal{V}(\theta_0)$, where

$$\mathcal{V}(\theta) = \left(\frac{1}{4\pi} \int_{-\pi}^{\pi} D_\theta \ln(f_\theta(\lambda)) D_{\theta'} \ln(f_\theta(\lambda)) d\lambda \right)^{-1} \tag{2.2}$$

where $D_\theta \ln(f_\theta(\lambda)) = \frac{\partial}{\partial\theta} \ln(f_\theta(\lambda))$ and θ' is the transpose of θ. A consistent estimator of $\mathcal{V}(\theta_0)$ is $\mathcal{V}(\hat{\theta}_n)$, provided that $f_\theta(\lambda)$ is smooth with respect of θ. Let z_α denote the $1-\alpha$ quantile of the standard normal distribution and let $\mathcal{V}_{r,r}(\hat{\theta}_n)$ denote the (r,r)-th element of $\mathcal{V}(\hat{\theta}_n)$. Then, the precise definitions of the t statistic and the two sided and one sided confidence intervals and tests of the PWML estimators are given in terms of $\mathcal{V}_{r,r}(\hat{\theta}_n)$ as follows.

1. The *t* statistic for testing the null hypothesis $H_0 : \theta_{0,r} = \theta_{h,r}$ is

$$\tau_n(\theta_{h,r}) = \frac{\sqrt{n}(\hat{\theta}_{n,r} - \theta_{h,r})}{\mathcal{V}_{r,r}^{1/2}(\hat{\theta}_n)}. \tag{2.3}$$

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

$$\mathcal{I}_{sym}(\hat{\theta}_n) = \left[\hat{\theta}_{n,r} - z_{\alpha/2} \frac{\mathcal{V}_{r,r}^{1/2}(\hat{\theta}_n)}{\sqrt{n}}, \hat{\theta}_{n,r} + z_{\alpha/2} \frac{\Sigma_{r,r}^{1/2}(\hat{\theta}_n)}{\sqrt{n}} \right]. \tag{2.4}$$

3. The *upper one-sided delta method* $100(1-\alpha)\%$ CI for $\theta_{0,r}$ is

$$\mathcal{I}_{up}(\hat{\theta}_n) = \left[\hat{\theta}_{n,r} - z_\alpha \frac{\mathcal{V}_{r,r}^{1/2}(\hat{\theta}_n)}{\sqrt{n}}, \infty \right). \tag{2.5}$$

4. The *two-sided delta method 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 $|\tau_n(\theta_{h,r})| > z_{\alpha/2}$.

5. The *one sided 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 $\tau_n(\theta_{h,r}) > z_\alpha$.

2.3 Confidence Intervals and Tests of the Parametric Bootstrap

Let $Y = (Y_1, Y_2, \ldots, Y_n)'$ be a sample from our linear regression model with strongly dependent errors as described in section 1 above. Then, the parametric bootstrap sample $Y^* = (Y_1^*, ..., Y_n^*)$ 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, Y^* consists of random variables from a linear regression process with, ARFIMA (p,d,q) errors having mean $\hat{\mu}$ and spectral density $f_{\hat{\theta}_n}(\lambda)$ conditional on the original sample 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 and tests.

1. The bootstrap sample mean $\hat{\mu}^*$ is then given by $\hat{\mu}^* = (\hat{\mu}_1^*, ..., \hat{\mu}_n^*) = (X_1'\hat{\beta}^*, ..., X_n'\hat{\beta}^*)$, where $\hat{\beta}^* = P^{-1}X'Y^*$ and $P = X'X$ as defined in section 1.

2. The bootstrap PWLL function $\Lambda_W(\theta, \hat{\mu}^*)$ is defined in the same way as the PWLL function $\Lambda_W(\theta, \hat{\mu})$ (see (1.10) above) but with Y^* and $\hat{\mu}^*$ replacing Y and $\hat{\mu}$, respectively.

3. 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 $\Lambda_W(\theta, \hat{\mu}^*)$. 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 $\theta_{0,r}$, $\theta_{h,r}$, and $\hat{\theta}_{n,r}$ be as defined in the paragraph preceding (2.2) and let $z^*_{|\tau|,\alpha}$ denote the value that minimizes $|P^*(|\tau_n^*(\hat{\theta}_{n,r})| \leq z) - (1-\alpha)|$ and let $z^*_{\tau,\alpha}$ denote the value that minimizes $|P^*(\tau_n^*(\hat{\theta}_{n,r}) \leq z) - (1-\alpha)|$ over $z \in \mathbb{R}$.

4. We define the bootstrap t-statistic by

$$\tau_n^*(\hat{\theta}_{n,r}) = \frac{\sqrt{n}(\hat{\theta}^*_{n,r} - \hat{\theta}_{n,r})}{\mathcal{V}_{r,r}^{1/2}(\hat{\theta}_n^*)} \quad (2.6).$$

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

5. 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

$$\mathcal{I}^*_{sym}(\hat{\theta}_n) = \left[\hat{\theta}_{n,r} - \frac{z^*_{|\tau|,\alpha}\mathcal{V}_{r,r}^{1/2}(\hat{\theta}_n)}{\sqrt{n}}, \hat{\theta}_{n,r} + \frac{z^*_{|\tau|,\alpha}\mathcal{V}_{r,r}^{1/2}(\hat{\theta}_n)}{\sqrt{n}}\right]. \quad (2.7)$$

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

$$\mathcal{I}^*_{up}(\hat{\theta}_n) = \left[\hat{\theta}_{n,r} - \frac{z^*_{\tau,\alpha}\mathcal{V}_{r,r}^{1/2}(\hat{\theta}_n)}{\sqrt{n}}, \infty\right). \quad (2.8)$$

7. 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 $|\tau_n(\theta_{h,r})| > z^*_{|\tau|,\alpha}$.

8. 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 $\tau_n(\theta_{h,r}) > z^*_{\tau,\alpha}$.

2.4 The Whittle Log-likelihood Derivatives

We now present the set up of the Whittle likelihood derivatives as it is applied in this paper. Let $\nu = (r_1, r_2, ..., r_q)'$ denote a q-vector of positive integers each less than or equal to $dim(\theta)$. Let $\Lambda_W(\theta, \hat{\mu})$ be as defined in (1.12) and let $\mathcal{M}_n$ be as in the paragraph preceding (1.11). We write the real valued q-th order partial derivative of the PWLL function indexed by ν as

$$\Lambda_{W,\nu} = D_\nu \Lambda_W(\theta, \hat{\mu}) = \frac{\partial^q}{\partial\theta_{r_1}\ldots\partial\theta_{r_q}}\Lambda_W(\theta, \hat{\mu}) = \mathcal{A}_{n,\nu}(\theta) + Y'\mathcal{M}_n\mathcal{B}_{n,\nu}(\theta)\mathcal{M}_nY \quad (2.9)$$

where

$$\mathcal{A}_{n,\nu}(\theta) = -\frac{n}{4\pi}\int_{-\pi}^{\pi} D_\nu \ln(f_\theta(\lambda))d\lambda \quad (2.10)$$

and

$$\mathcal{B}_{n,\nu}(\theta) = -\frac{1}{2}D_\nu\Sigma_n((2\pi)^{-2}f_\theta^{-1}). \quad (2.11)$$

Equations (2.9)-(2.11) 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 Y as defined in section 1 above. We shall introduce some more notations. Let

$$\mathcal{Z}_n(\theta) = (\Lambda_{W,\nu(1)}(\theta), \ldots, \Lambda_{W,\nu(\eta)}(\theta)), \quad (2.12)$$

where each vector $\nu(j)$ is of the same form as ν defined in (2.9)-(2.11) above for $\eta = dim(\mathcal{Z}_n(\theta))$ and $j = 1, 2, ..., \eta$. Let

$$\mathcal{W}_n(\theta) = n^{-1/2}(\mathcal{Z}_n(\theta) - E_\theta\mathcal{Z}_n(\theta)). \quad (2.13)$$

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

$$\mathcal{D}_n(\theta) = E[\mathcal{W}_n(\theta)\mathcal{W}_n(\theta)'] \quad (2.14)$$

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

$$\mathcal{D}_n(\theta)i, j = tr(\mathcal{B}_{n,\nu_i}\Sigma_n(f_\theta)\mathcal{B}_{n,\nu_j}\Sigma_n(f_\theta)) \quad (2.15).$$

(See Anderson, 1984 for details.)

3. Delta Method and Bootstrap Coverage Probability Errors of the PWML Estimators

In this section we present two main results of this paper. In Theorem 3.3, we establish both the delta method and bootstrap coverage probability errors of the two sided confidence intervals for the PWML estimators. In Theorem 3.4, we establish both the delta method and bootstrap coverage probability errors of the one sided confidence intervals for the same estimators. In both theorems a comparison of the performances of the bootstrap and the traditional delta method shows that the bootstrap provides a significant improvement of the probability errors over the delta method. The next two lemmas are key ingredients in the proofs of the two theorems. First we introduce some additional notations.
Let $\Phi(\cdot)$ denote the distribution function of the standard normal distribution. Define

$$D_{z,\upsilon} = \frac{\partial^q}{\partial z_{\upsilon_1} \dots \partial z_{\upsilon_q}}, \quad for \;\; \upsilon = (\upsilon_1, ..., \upsilon_q). \tag{3.1}$$

Let $\varphi_n(z, \theta) = E_\theta \exp(iz'\mathcal{Z}_n(\theta))$ denote the characteristic function of $\mathcal{Z}_n(\theta)$ where $z \in \mathbb{R}^d$ and let $\kappa_n(\theta)_\upsilon$ denote the υ cumulants of $\mathcal{Z}_n(\theta)$ (see equation (2.12) above). By definition,

$$\kappa_n(\theta)_\upsilon = i^{-q} D_{z,\upsilon} \ln(\varphi_n(z, \theta))|_{z=0}, \;\; where \;\; i = \sqrt{-1}. \qquad (3.2$$

The vector $\kappa_n(\theta)$ is composed of elements $\kappa_n(\theta)_\upsilon$ for vectors υ of dimension $q \le s$, where s is as given in Assumption A2. Let $\bar{\kappa}_n(\theta) = \frac{\kappa_n(\theta)}{n}$. By Lemmas 3.2 and 3.3 of Aga (2024a), the elements of $\bar{\kappa}_n(\theta)$ are $O(1)$.

Let $\pi_j(\delta, \bar{\kappa}_n(\theta))$ be a polynomial in $\delta = \partial/\partial z$ whose coefficients are polynomials in the elements of $\bar{\kappa}_n(\theta)$ and for which $\pi_j(\delta, \bar{\kappa}_n(\theta))\Phi(x)$ 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 below provides an Edgeworth expansion of the bootstrap t-statistic (2.6) of the PWML estimators and is proved in Aga (2024a). It is used in the proofs of Theorems 3.3 and 3.4 below. We restate it here for convenience.

Lemma 3.1. Suppose assumptions A1-A5 hold, and let $s \ge 3$ be an integer. Then, for all $\varepsilon > 0$

$$P_{\theta_0}\left(\sup_{x\in\mathbb{R}} |P^*_{\hat{\theta}_n}(\tau^*_n(\hat{\theta}_{n,r}) \le x) - \Phi(x)\left(1 + \sum_{j=1}^{s-2} n^{-j/2}\pi_j(\delta, \bar{\kappa}_n(\hat{\theta}_n))\right)| > n^{1-s/2}\varepsilon\right) = o(n^{1-s/2}) \tag{3.3}$$

uniformly over $\theta_0 \in \Theta_c$.

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 $\Theta_c^+ = \{\theta \in R^r : dist(\theta, \Theta_c) \le \delta\}$ 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. The next lemma is used in the proofs of the main results of this paper.

Lemma 3.2. Suppose Assumptions A1-A5 hold, and let Θ_c and $s \ge 3$ be as given in Assumptions A2 and A4. Then, for all $\varepsilon > 0$,

$$\sup_{\theta_0\in\Theta_c} P_{\theta_0}(\sqrt{n}||\bar{\kappa}_n(\hat{\theta}_n) - \bar{\kappa}_n(\theta_0|| > ln(n)\varepsilon) = o(n^{-(s-2)/2}), \tag{3.4}$$

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 $||\theta - \theta_0|| < \delta$, $|\bar{\kappa}_n(\theta)_\eta - \bar{\kappa}_n(\theta_0)_\eta| \le C_n||\theta - \theta_0||$, where

$$C_n = \sup_{\theta\in\Theta_c^+} |D_{\theta_i}\bar{\kappa}_n(\theta)_\eta|. \tag{3.5}$$

for all $i = 1, ..., p+q+1 = dim(\theta)$, and $D_{\theta_i} = \frac{\partial}{\partial\theta_i}$, We first show that $\limsup_{n\to\infty} C_n < \infty$, but to this end it suffices to show that

$$\sup_{\theta\in\Theta_c} |D_{\theta_i}\kappa_n(\theta)_\eta| = O(n) \tag{3.6}$$

for all $i \le p+q+1$.
To prove that (3.6) 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, $D_{\theta_i}\kappa_n(\theta)_\eta$ is a finite sum of terms of the form

$$C_q\left(\prod_{r=1}^{q}(\mathcal{M}_n\bar{\mathcal{B}}_r\mathcal{M}_n\bar{\Sigma}_r)\right), \tag{3.10}$$

where $\mathcal{M}_n$, $\mathcal{B}$, and Σ_n are as given in equations (2.9-2.11), $\bar{\mathcal{B}}_r$ equals either $\mathcal{B}_{n,\nu(\eta_r)}(\theta)$ or $D_{\theta_i}\mathcal{B}_{n,\nu(\eta_r)}(\theta)$ and $\bar{\Sigma}_r$ equals either $\Sigma_n((2\pi)^{-2}f_\theta^{-1})$ or $D_{\theta_i}\Sigma_n((2\pi)^{-2}f_\theta^{-1})$.
We observe that $\mathcal{B}_{n,\nu(\eta_r)}(\theta)$ and $D_{\theta_i}\mathcal{B}_{n,\nu(\eta_r)}(\theta)$ have the same form because they are both partial derivatives of $-\frac{1}{2}\Sigma_n((2\pi)^{-2}f_\theta^{-1})$, (see (2.11)). It follows that $D_{\theta_i}\kappa_n(\theta)_\eta$ has the same form as $\kappa_n(\theta)_\eta$ itself. (3.6) now follows from Lemma 6 (c) of Andrews et al. (2006).
To complete the proof of the lemma, let $\vartheta > 0$ satisfy

$$\vartheta < \varepsilon\left(\rho \limsup_{n\to\infty} C_n\right)^{-1} < \infty, \tag{3.7}$$

where, $\rho = \sqrt{dim(\bar{\kappa}_n(\theta))}$. We have

$$\begin{aligned}
&\sup_{\theta_0\in\Theta_c} P_{\theta_0}(\sqrt{n}\|\bar{\kappa}_n(\hat{\theta}_n) - \bar{\kappa}_n(\theta_0\| > ln(n)\varepsilon)\\
&\le \sup_{\theta_0\in\Theta_c} P_{\theta_0}(\sqrt{n}\|\bar{\kappa}_n(\hat{\theta}_n) - \bar{\kappa}_n(\theta_0\| > ln(n)\varepsilon,\ \sqrt{n}\|\theta-\theta_0\| \le \ln(n)\vartheta)\\
&+ \sup_{\theta_0\in\Theta_c} \sqrt{n}\|\theta-\theta_0\| \le \ln(n)\vartheta)\\
&\le \sup_{\theta_0\in\Theta_c} P_{\theta_0}(\rho \limsup_{n\to\infty} C_n \sqrt{n}\|\hat{\theta}_n - \theta_0\| > ln(n)\varepsilon,\ \sqrt{n}\|\theta-\theta_0\| \le \ln(n)\vartheta) + o(n^{-(s-2)/2})\\
&= o(n^{-(s-2)/2})
\end{aligned} \tag{3.8}$$

where, the second inequality above uses (3.5) and Assumption A2 and the last equality holds because

$$P_{\theta_0}(\rho \limsup_{n\to\infty} C_n n^{1/2}\|\tilde{\theta}_n - \theta_0\| > ln(n)\varepsilon, n^{1/2}\|\theta-\theta_0\| \le \ln(n)\vartheta) = 0, \tag{3.9}$$

since $\vartheta < \varepsilon\,(\rho \limsup_{n\to\infty} C_n)^{-1}$ by (3.7) and $\limsup_{n\to\infty} C_n < \infty$ by (3.6). This completes the proof of the lemma. □

We now state one of the main bootstrap coverage probability results of this paper. Theorem 3.3 below establishes that the bootstrap probability errors for a two sided probability errors show an improvement over the delta method counter parts.

Theorem 3.3. Suppose assumptions A1-A5 hold and let $s \ge 3$ as in Assumption A4. Consider $\{\hat{\theta}_n : n \ge 1\}$ as given in assumption A2 and let Θ_c be any compact subset of Θ, the parameter space. Then, for $s = 5$ we have
(a) $\sup_{\theta_0\in\Theta_c} P_{\theta_0}(\theta_0 \in \mathcal{I}_{sym}(\hat{\theta}_n)) = (1-\alpha) + O(n^{-1})$.
(b) $\sup_{\theta_0\in\Theta_c} |P_{\theta_0}(\theta_0 \in \mathcal{I}^*_{sym}(\hat{\theta}_n)) - (1-\alpha)| = o(n^{-3/2}\ln(n))$
Proof.
(a) Let $n^{-1}\mathcal{Z}^+_n(\theta_0)$ denote the vector $n^{-1}\mathcal{Z}_n(\theta_0)$ of normalized LLDs augmented to include the vector of expected values of all partial derivatives with respect to θ of order s of $n^{-1}\Lambda_W(\theta_0)$. By Theorem 3(b) of Bhattacharya and Ghosh (1978) the normalized PWML estimator and the t statistic $\tau_n(\theta_{0,r})$ can be approximated by smooth functions of $n^{-1}\mathcal{Z}^+_n(\theta_0)$. Specifically, there is an infinitely differentiable function $\zeta(.)$ that does not depend on θ_0 that satisfies $\zeta(n^{-1}E_{\theta_0}\mathcal{Z}^+_n(\theta)) = 0$ for all n large and all $\theta_0 \in \Theta_c$ and

$$\sup_{\theta_0\in\Theta_c}\sup_{B\in\mathcal{B}_r} |P_{\theta_0}(\tau_n(\theta_0) \in B) - P_{\theta_0}(n^{1/2}\zeta(n^{-1}Z^+_n(\theta_0)) \in B)| = o(n^{-(s-2)/2}) \tag{3.11}$$

where $r = dim(\theta)$, $\mathcal{B}_r$ denotes the set of all convex sets in R^r.
For simplicity of notation, let

$$\Omega_s(\theta_0) = 1 + \sum_{i=1}^{s-2} n^{-i/2}\pi_i(\delta, \bar{\kappa}_n(\theta_0)) \tag{3.12}$$

where $s \ge 3$ as in assumptions A2 and A4. Then,

$$\begin{aligned}
|P_{\theta_0}(\tau_n(\theta_{0,r}) \le z) - \Omega_s(\theta_0)\Phi(z)| &\le |P_{\theta_0}(\tau_n(\theta_{0,r}) \le z) - P_{\theta_0}(n^{1/2}\zeta(n^{-1}\mathcal{Z}^+_n(\theta_0)) \le z)|\\
&+ |P_{\theta_0}(n^{1/2}\zeta(n^{-1}\mathcal{Z}^+_n(\theta_0)) \le z) - \Omega_s(\theta_0)\Phi(z)|.
\end{aligned} \tag{3.13}$$

The first term of the right hand side of (3.13) is equal to $o(n^{-(s-2)/2})$ by Lemma 10 of Andrews *et al.* [2006]. Thus, we only need to show that

$$|P_{\theta_0}(n^{1/2}\zeta(n^{-1}\mathcal{Z}^+_n(\theta_0)) \le z) - \Omega_s(\theta_0)\Phi(z)| = o(n^{-(s-2)/2}) \tag{3.14}$$

uniformly over $\theta_0 \in \Theta_c$. We note that an asymptotic expansion of $\mathcal{W}_n(\theta_0) = n^{-1/2}(\mathcal{Z}_n(\theta_0) - E_\theta\mathcal{Z}_n(\theta_0))$ is established by Theorem 3.4(b) of Aga (2021) for each $\theta_0 \in \Theta_c$. Now, an expansion for $n^{1/2}\zeta(n^{-1}\mathcal{Z}^+_n(\theta_0))$ is obtained from that of $n^{-1/2}(\mathcal{Z}_n(\theta_0) - E_\theta\mathcal{Z}_n(\theta_0))$ by the argument of Battacharya and Ghosh (1978), Theorem 2, p. 436 using the smoothness of $\zeta(.)$ satisfying $\zeta(n^{-1}E_{\theta_0}\mathcal{Z}^+_n(\theta_0)) = 0$ for all $n \ge 1$ and all $\theta_0 \in \Theta_c$, and assumption A3.
Combining (3.13) and (3.14) we obtain

$$|P_{\theta_0}(\tau_n(\theta_{0,r}) \le z) - [1 + \sum_{i=1}^{s-2} n^{-i/2}\pi_i(\delta, \bar{\kappa}_n(\theta_0))]\Phi(z)| = o(n^{1-s/2}) \tag{3.15}$$

for all $\theta_0 \in \Theta_c$ and all $z \in \mathbb{R}$.

Now, from the definition of $\mathcal{I}_{sym}(\hat{\theta}_n)$ given in (2.4) we observe that

$$\begin{aligned} P_{\theta_0}(\theta_0 \in \mathcal{I}_{sym}(\hat{\theta}_n)) &= P_{\theta_0}(\hat{\theta}_{n,r} - z_{\alpha/2}\frac{\sqrt{\mathcal{V}_{r,r}}}{\sqrt{n}} \leq \theta_{0,r} \leq \hat{\theta}_{n,r} - z_{\alpha/2}\frac{\sqrt{\mathcal{V}_{r,r}}}{\sqrt{n}}) \\ &= P_{\theta_0}(-z_{\alpha/2} \leq \frac{\sqrt{n}(\theta_{0,r}-\hat{\theta}_{n,r})}{\sqrt{\mathcal{V}_{r,r}(\hat{\theta}_n)}} \leq z_{\alpha/2}) \\ &= P_{\theta_0}(|\tau_n(\theta_{0,r})| \leq z_{\alpha/2}) \end{aligned} \tag{3.16}$$

and that

$$P_{\theta_0}(|\tau_n(\theta_{0,r})| \leq z_{\alpha/2}) = P_{\theta_0}(\tau_n(\theta_{0,r}) \leq z_{\alpha/2}) - P_{\theta_0}(\tau_n(\theta_{0,r}) \leq -z_{\alpha/2}). \tag{3.17}$$

Putting $s = 5$ in (3.12) to get

$$\Omega_5(\theta_0) = 1 + n^{-1/2}\pi_1(\delta, \bar{\kappa}_{n,5}(\theta_0)) + n^{-1}\pi_2(\delta, \bar{\kappa}_{n,5}(\theta_0)) + n^{-1}\pi_3(\delta, \bar{\kappa}_{n,5}(\theta_0)) \tag{3.18}$$

and substituting z by $z_{\alpha/2}$ and $s = 5$ in (3.15) and using (3.17) and (3.18) we obtain

$$|P_{\theta_0}(|\tau_n(\theta_{0,r})| \leq z_{\alpha/2}) - \Omega_5(\theta_0)(\Phi(z_{\alpha/2}) - \Phi(-z_{\alpha/2}))| = o(n^{-3/2}). \tag{3.19}$$

We observe that $\pi_1(\delta, \bar{\kappa}_{n,4}(\theta_0))\Phi(z)$ and $\pi_3(\delta, \bar{\kappa}_{n,4}(\theta_0))\Phi(z)$ are even functions of z which leads to $\pi_1(\delta, \bar{\kappa}_{n,5}(\theta_0))(\Phi(z_{\alpha/2}) - \Phi(-z_{\alpha/2})) = 0$ and $\pi_3(\delta, \bar{\kappa}_{n,5}(\theta_0))(\Phi(z_{\alpha/2}) - \Phi(-z_{\alpha/2})) = 0$. It follows that

$$\sup_{\theta_0 \in \Theta_c} |P_{\theta_0}(|\tau_n(\theta_{0,r})| \leq z_{\alpha/2}) - [1 + n^{-1}\pi_2(\delta, \bar{\kappa}_{n,5}(\theta_0))(\Phi(z_{\alpha/2}) - \Phi(-z_{\alpha/2}))]| = o(n^{-3/2}). \tag{3.20}$$

Now, because $\Phi(z_{\alpha/2}) - \Phi(-z_{\alpha/2}) = 1 - \alpha$ and $n^{-1}\pi_2(\delta, \bar{\kappa}_{n,5}(\theta_0))(\Phi(z_{\alpha/2}) - \Phi(-z_{\alpha/2})) = O(n^{-1})$, we have

$$P_{\theta_0}(\theta_0 \in \mathcal{I}_{sym}(\hat{\theta}_n)) = P_{\theta_0}(|\tau_n(\theta_{0,r})| \leq z_{\alpha/2}) = O(n^{-1}) \tag{3.21}$$

for all $\theta_0 \in \Theta_c$. This establishes part (a).

(b) We need to show that for $s = 5$,

$$\sup_{\theta_0 \in \Theta_c} |P_{\theta_0}(\theta_0 \in \mathcal{I}^*_{sym}(\hat{\theta}_n)) - (1 - \alpha)| = o(n^{-3/2}\ln(n)) \tag{3.22}$$

for all $\theta_0 \in \Theta_c$. To this end we first observe that, arguing as in (3.16), $P_{\theta_0}(\theta_{0,r} \in \mathcal{I}^*_{sym}(\hat{\theta}_n)) = P_{\theta_0}(|\tau_n(\theta_{0,r})| \leq z^*_{\tau,\alpha})$. To establish (3.22) it is sufficient to show that

$$P_{\theta_0}(|\tau_n(\theta_{0,r})| \leq z^*_{\tau,\alpha}) = 1 - \alpha + o(n^{-3/2}\ln(n)) \tag{3.23}$$

uniformly over $\theta_0 \in \Theta_c$.

As in (3.12), noting that in a bootstrap scheme the PMLE $\hat{\theta}_n$ is the true parameter, let

$$\Omega_5(\hat{\theta}_n) = 1 + \sum_{i=1}^{3} n^{-i/2}\pi_j(\delta, \bar{\kappa}_n(\hat{\theta}_n)). \tag{3.24}$$

Using Lemma 3.1 with $s = 5$ and the fact that

$$P_{\hat{\theta}_n}(|\tau^*_n(\hat{\theta}_{n,r})| \leq z) = P_{\hat{\theta}_n}(\tau^*_n(\hat{\theta}_{n,r}) \leq z) - P_{\hat{\theta}_n}(\tau^*_n(\hat{\theta}_{n,r}) \leq -z), \tag{3.25}$$

we obtain for all $\varepsilon > 0$

$$P_{\theta_0}\Big(\sup_{z \in \mathbb{R}} |P^*_{\hat{\theta}_n}(|\tau^*_n(\hat{\theta}_{n,r})| \leq z) - \Omega_5(\hat{\theta})(\Phi(z) - \Phi(-z))| > n^{-3/2}\varepsilon\Big) = o(n^{-3/2}) \tag{3.26}$$

for all $\theta_0 \in \Theta_c$. Now, since $\pi_j(\delta, \bar{\kappa}_n(\hat{\theta}_n))(\Phi(z))$ is an even function for odd j, it follows that

$$\pi_1(\delta, \bar{\kappa}_n(\hat{\theta}_n))(\Phi(z) - \Phi(-z)) = \pi_3(\delta, \bar{\kappa}_n(\hat{\theta}_n))(\Phi(z) - \Phi(-z)) = 0. \tag{3.27}$$

Therefore, (3.26) above becomes

$$P_{\theta_0}\left(\sup_{z \in \mathbb{R}} |P^*_{\hat{\theta}_n}(|\tau^*_n(\hat{\theta}_{n,r})| \leq z) - [1 + n^{-1}\pi_2(\delta, \bar{\kappa}_n(\hat{\theta}_n))](\Phi(z) - \Phi(-z))| > n^{-3/2}\varepsilon\right) = o(n^{-3/2}) \tag{3.28}$$

for all $\theta_0 \in \Theta_c$. Similarly, substituting $s = 5$ in (3.15) and arguing as in (3.26)-(3.28) above we obtain

$$\sup_{z\in\mathbb{R}} |P_{\theta_0}(|\tau_n(\theta_{0,r})| \leq z) - [1 + n^{-1}\pi_2(\delta, \bar{\kappa}_n(\theta_0)))](\Phi(z) - \Phi(-z))| = o(n^{-3/2}), \tag{3.29}$$

for all $\theta_0 \in \Theta_c$ and combining (3.28) and (3.29) we obtain

$$P_{\theta_0}\left(\sup_{z\in\mathbb{R}} |[\pi_2(\delta, \bar{\kappa}_n(\hat{\theta}_n)) - \pi_2(\delta, \bar{\kappa}_n(\theta_0))](\Phi(z) - \Phi(-z))| > n^{-1/2}\ln(n)\varepsilon\right) = o(n^{-3/2}) \tag{3.30}$$

for all $\theta_0 \in \Theta_c$. Combining (3.28), (3.29), and (3.30) yields

$$P_{\theta_0}\left(\sup_{z\in\mathbb{R}} |P^*_{\hat{\theta}_n}(|\tau^*_n(\hat{\theta}_{n,r})| \leq z) - P_{\theta_0}(|\tau_n(\theta_{0,r})| \leq z)| > n^{-3/2}\ln(n)\varepsilon\right) = o(n^{-3/2}) \tag{3.31}$$

for all $\theta_0 \in \Theta_c$. Now, let $\Upsilon(z) = \Omega_5(\hat{\theta}_n)\Phi(z)$. Υ is essentially an Edgeworth expansion of the bootstrap t-statistic $\tau^*_n(\hat{\theta}_n)$ given in Lemma 3.1 above. Since Υ is continuous in z, there exists $\tilde{z}_{\tau,\alpha}$ such that $\Upsilon(\tilde{z}_{\tau,\alpha}) = 1 - \alpha$. Using this and the definition of $z^*_{\tau,\alpha}$ (see the last line of section (2.3)), we have:

$$\begin{aligned} |P^*_{\hat{\theta}_n}(\tau^*_n(\hat{\theta}_{n,r}) \leq \tilde{z}_{\tau,\alpha}) - \Upsilon(\tilde{z}))| &= |P^*_{\hat{\theta}_n}(\tau^*_n(\hat{\theta}_{n,r}) \leq \tilde{z}_{\tau,\alpha}) - (1-\alpha)| \\ &\geq |P^*_{\hat{\theta}_n}(\tau^*_n(\hat{\theta}_{n,r}) \leq z^*_{\tau,\alpha}) - (1-\alpha)|. \end{aligned} \tag{3.32}$$

Moreover, by Lemma 3.1 we have

$$P_{\theta_0}\left(\sup_{z\in\mathbb{R}} |P^*_{\hat{\theta}_n}(t^*_n(\hat{\theta}_{n,r}) \leq z^*_{\tau,\alpha}) - (1-\alpha)| > n^{-3/2}\varepsilon\right) = o(n^{-3/2}). \tag{3.33}$$

for all $\theta_0 \in \Theta_c$. Taking $z = z^*_{\tau,\alpha}$ in (3.31) and combining it with (3.33) yields:

$$\sup_{\theta_0\in\Theta_c} P_{\theta_0}(|1 - \alpha - P_{\theta_0}(\tau_n(\theta_{0,r}) \leq z^*_{\tau,\alpha})| > n^{-3/2}\ln(n)\varepsilon) = o(n^{-3/2}). \tag{3.34}$$

Hence (3.34) reduces to

$$|1 - \alpha - P_{\theta_0}(\tau_n(\theta_{0,r}) \leq z^*_{\tau,\alpha})| < n^{-3/2}\ln(n)\varepsilon, \tag{3.35}$$

for all $\theta_0 \in \Theta_c$ which proves (3.22). □

Theorem 3.4. Suppose Assumptions A1-A5 hold, and let $s \geq 3$ be an integer and let Θ_c be a compact subset of Θ. Then, for $s = 4$ we have

(a) $\sup_{\theta_0\in\Theta_c} P_{\theta_0}(\theta_0 \in \mathcal{I}_{up}(\hat{\theta}_n)) = (1-\alpha) + O(n^{-1/2})$

(b) $\sup_{\theta_0\in\Theta_c} |P_{\theta_0}(\theta_0 \in \mathcal{I}^*_{up}(\hat{\theta}_n)) - (1-\alpha)| = o(n^{-1}\ln(n))$ **Proof.**

(a) We have $P_{\theta_0}(\theta_0 \in \mathcal{I}_{up}(\hat{\theta}_n)) = P_{\theta_0}(\tau_n(\theta_{0,r}) \leq z_\alpha)$. By (3.15) with $s = 4$, and using the fact that $\Phi(z_\alpha) = 1 - \alpha$ we have

$$\begin{aligned} \sup_{\theta_0\in\Theta_c} |P_{\theta_0}(\theta_0 \in \mathcal{I}_{up}(\hat{\theta}_n)) - (1-\alpha)| &= \sup_{\theta_0\in\Theta_c} |P_{\theta_0}(\tau_n(\theta_{0,r}) \leq z_\alpha) - (1-\alpha)| \\ \leq \sup_{\theta_0\in\Theta_c} (|P_{\theta_0}(\tau_n(\theta_{0,r}) \leq z_\alpha) &- [1 + n^{-1/2}p_1(\delta, \bar{\kappa}_n(\theta_0)) + n^{-1}\pi_2(\delta, \bar{\kappa}_n(\theta_0))]\Phi(z_\alpha)| \\ &+ |[1 + n^{-1/2}\pi_1(\delta, \bar{\kappa}_n(\theta_0)) + n^{-1}\pi_2(\delta, \bar{\kappa}_n(\theta_0))]\Phi(z_\alpha) - (1-\alpha)|) \\ \leq \sup_{\theta_0\in\Theta_c} |(1-\alpha)n^{-1/2}\pi_1(\delta, \bar{\kappa}_n(\theta_0)) &+ n^{-1}\pi_2(\delta, \bar{\kappa}_n(\theta_0))| + o(n^{-1/2}) = O(n^{-1/2}). \end{aligned} \tag{3.36}$$

This proves part (a) of the theorem.

(b) The proof is analogous to that of Theorem 3.3 (b) above and therefore some details are omitted. As in (3.12), let $\Omega_4(\hat{\theta}_n) = 1 + \sum_{i=1}^{2} n^{-i/2}\pi_j(\delta, \bar{\kappa}_n(\hat{\theta}_n))$. By Lemma 3.1 above with $s = 4$ we obtain for all $\varepsilon > 0$

$$P_{\theta_0}\Big(\sup_{z\in\mathbb{R}} |P^*_{\hat{\theta}_n}(|\tau^*_n(\hat{\theta}_{n,r})| \leq z) - \Omega_4(\hat{\theta})(\Phi(z) - \Phi(-z))| > n^{-1}\varepsilon\Big) = o(n^{-1}) \tag{3.37}$$

for all $\theta_0 \in \Theta_c$. Moreover, by Lemma 3.1 and Lemma 3.2 with $s = 4$, respectively, and using the evenness of $\pi_j(\delta, \bar{\kappa}_n(\theta))(\Phi(z) - \Phi(-z))$ for $j = 1$ we obtain:

$$\sup_{z\in\mathbb{R}} |P_{\theta_0}(|\tau_n(\theta_{0,r})| \leq z) - \Omega_4(\theta_0)(\Phi(z) - \Phi(-z))| = o(n^{-1}), \tag{3.38}$$

for all $\theta_0 \in \Theta_c$ and

$$P_{\theta_0}\left(\sup_{z\in\mathbb{R}} |[\pi_2(\delta, \bar{\kappa}_n(\hat{\theta}_n)) - \pi_2(\delta, \bar{\kappa}_n(\theta_0))](\Phi(z) - \Phi(-z))| > n^{-1/2}\ln(n)\varepsilon\right) = o(n^{-1}) \tag{3.39}$$

for all $\theta_0 \in \Theta_c$. Combining (3.37), (3.38), and (3.39) above we obtain:

$$P_{\theta_0}\left(\sup_{z\in\mathbb{R}} |P^*_{\hat{\theta}_n}(|\tau^*_n(\hat{\theta}_{n,r})| \leq z) - P_{\theta_0}(|\tau_n(\theta_{0,r})| \leq z)| > n^{-1}\ln(n)\varepsilon\right) = o(n^{-1}) \tag{3.40}$$

for all $\theta_0 \in \Theta_c$. One can see the resemblance between (3.31) and (3.40) and therefore the remainder of the proof proceeds as in part (b) of Theorem 3.3 above to obtain

$$|P_{\theta_0}(\theta_0 \in \mathcal{I}^*_{up}(\hat{\theta}_n)) - (1 - \alpha)| = o(n^{-1}\ln(n)) \tag{3.41}$$

for all $\theta_0 \in \Theta_c$. □

4. Conclusion

In this paper, we analyze and compare the coverage probability errors of the delta method and the parametric bootstrap of both one-sided and two-sided confidence intervals for the Whittle estimator in a linear regression model with ARFIMA long memory time series errors. In both one-sided and two-sided cases, it has been proved that the parametric bootstrap shows a significant improvement over the traditional delta method in terms of reducing the magnitude of the probability errors. The coverage probability errors of the symmetric two-sided delta method and parametric bootstrap confidence intervals for the covariance parameter θ_0 are shown to be $O(n^{-1})$ and $o(n^{-3/2}\ln n)$ (Theorem 3.3), respectively. Likewise, the coverage probability errors of delta method and parametric bootstrap one-sided confidence intervals are shown to be $O(n^{-1/2})$ and $o(n^{-1}\ln n)$ (Theorem 3.4), respectively. Both of these results show that the bootstrap exhibits a significant improvement of confidence interval probability errors over the delta method counterparts.

The method employed to establish these results is briefly outlined as follows.

1. We show that the t statistic based on the PWML estimators can be approximated arbitrarily closely by a smooth function of a vector of PWML derivatives of sufficiently high order. Such argument follows that of Bhattacharya and Gosh (1978), Theorem 3(b).
2. This result along with the Edgeworth expansion for the normalized and centered PWML estimators (see (2.12)) obtained in Aga (2021), are utilized to give Edgeworth expansions for the distributions of the t statistics that hold uniformly over parameter values in a compact set (see the proof of Theorem 3.3). The Edgeworth expansions for the t statistics are used to determine the coverage probability errors of the delta method confidence intervals in both Theorems 3.3 and 3.4.
3. It has been proved in Aga (2024a) Lemmas 3.2 and 3.3 that the elements of the cumulants $\bar{\kappa}(\theta)$ are $O(1)$. Moreover, Theorem 3.4(a) of Aga (2024a) provides an Edgeworth expansion of the bootstrap t statistic (2.6) of the PWML estimators (stated here as Lemma 3.1). We utilize these results along with Lemma 3.2 stated and proved in this paper to establish the coverage probability errors of the parametric bootstrap in Theorems 3.3 and 3.4.

It is worthwhile noting that the current paper builds on the results obtained by Andrews et al. (2006) in which only the long memory error component of our current model is considered. The results are also similar in structure with those of Aga et al. (2007) in which the PML estimators of the same model are considered instead of the PWML estimators.

One limitation of the findings presented in this paper is that the sequence of PWML estimators, for which the Edgeworth expansion is derived, must satisfy Assumption A2, which requires that the PWML estimators for which the results of the paper hold be consistent. This same limitation is found in Dahlhaus (1989), Lieberman et al. (2003), and Aga (2024a) and others. While Lemma 1 of Andrews et al. (2006) shows the existence of PWML estimators that satisfy this assumption, it is generally unknown to date whether or not the results of this paper and others still hold without the assumption.

Another drawback of the results of this paper (and those of Aga et al. (2007), Aga (2021, 2024a), and others) is the restriction imposed on the design matrix by A5, requiring the elements of the design matrix to be bounded by $\frac{M}{\sqrt{n}}$ for $M < \infty$. A natural question would be whether one can find a design matrix of practical value that satisfies the condition imposed in this assumption. The n x p matrix:

$$\mathcal{Z} = \begin{pmatrix} 1 & 1 & 1 & . & . & 1 \\ 1 & 2 & 2^2 & . & . & 2^{p-1} \\ . & . & . & . & . & \\ 1 & n & n^2 & . & . & n^{p-1} \end{pmatrix}. \tag{4.1}$$

is a special case of the so called Vandermonde Matrix, which arises in many applications such as polynomial least squares fitting, Lagrange interpolating polynomials, and the reconstruction of a statistical distribution from the distributions moments. The fact that the matrix in (4.1) satisfies Assumption A5 is proved in Lemma 3.1 of Aga et al. (2007).

One advantage of the parametric bootstrap over the delta method for long memory time series is its relative robustness to issues such as estimator bias, deviations from normality, and inaccuracies in the underlying model. Additionally, it handles the complexities of long memory processes more effectively. However, while the parametric bootstrap can offer valuable insights into the distribution of estimators and aid in constructing confidence intervals, it does not inherently correct for any bias present in the estimators.

In particular, Whittle estimators can exhibit bias in finite samples, especially when the sample size is small or if the model assumptions are not fully met. This bias can stem from various sources, such as the approximation of the log-likelihood and the estimation of the periodogram (see (1.14)). In cases of long memory, the bias might be more pronounced due to the intricate dependence structure. The effectiveness of the parametric bootstrap method in mitigating this inherent bias depends on how accurately the parametric model represents the data and on the sample size. To overcome such limitations of the parametric bootstrap, non-parametric models such as the non-parametric bootstrap can be employed. Unlike parametric methods, non-parametric bootstrap does not assume a specific distribution and can be more robust when less is known about the true data distribution. Often, addressing the bias of the original estimator requires additional techniques beyond bootstrapping.

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. (2015). Edgeworth expansion for the parametric bootstrap t-statistic for linear regression processes with strongly dependent errors. *Journal of Statistical Theory and Applications, 14*(1), 52-59. https://www.atlantis-press.com/journals/jsta/18187

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), 119. https://doi.org/10.5539/ijsp.v10n4p119

Aga, M. (2022). *Edgeworth expansion of the t-statistic of the Whittle MLE for linear regression processes with long-memory disturbances.* Communications in Statistics - Theory and Methods. Advance online publication. https://doi.org/10.1080/03610926.2022.2111525

Aga, M. (2024a). *Valid Edgeworth expansion of the bootstrap t-statistic of the Whittle MLE for linear regression models with long-memory residuals.* Sankhya A: The Indian Journal of Statistics. Advance online publication. https://doi.org/10.1007/s13171-024-00361-x

Anderson, T. W. (1984). *An introduction to multivariate statistical analysis* (2nd ed.). 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.2005.06.003

Bhattacharya, R. N., & Ghosh, J. K. (1978). On the validity of the formal Edgeworth expansion. *Annals of Statistics, 6*(2), 434-451. https://doi.org/10.1214/aos/1176344134

Beran, R. J. (1982). Estimating sampling distributions: The bootstrap and competitors. *Annals of Statistics, 10,* 212-225. https://doi.org/10.1214/aos/1176345704

Beran, R. J. (1994). *Statistics for long-memory processes.* Chapman and Hall.

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

Brockwell, P. J., & Davis, R. A. (1991). *Time series: Theory and methods* (2nd ed.). Springer.

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://www.jstor.org/stable/2241663

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

Eck, D. J. (2018). Bootstrapping multivariate linear regression models. *Statistics and Probability Letters, 134,* 141-149. https://doi.org/10.1016/j.spl.2017.11.001

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. SIAM. (CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. 38)

Efron, B., & Gong, G. (1982). A leisurely look at the bootstrap, the jackknife, and cross-validation. *The American Statistician, 37*, 36-48. https://doi.org/10.2307/2685844

Efron, B. (1991). Regression percentiles using asymmetric squared error loss. *Statistica Sinica, 1*(1), 93-125. https://www.jstor.org/stable/24303995

Efron, B., & Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science, 1*, 54-75. http://www.jstor.org/stable/2245500

Franco, G., & Reisen, V. (2004). Bootstrap techniques in semiparametric estimation methods for ARFIMA models: A comparison study. *Computational Statistics, 19*, 243-259. https://doi.org/10.1007/BF02892059

Freedman, D. A. (1981). Bootstrapping regression models. *Annals of Statistics, 9*(6), 1218-1228. https://doi.org/10.1214/aos/1176345638

Freedman, D. A., & Peters, S. C. (1984). Bootstrapping regression equations: Some empirical results. *Journal of the American Statistical Association, 79*(385), 97-106. https://doi.org/10.2307/2288341

Granger, C. W. J., & Joyeux, R. (1980). An introduction to long-range time series and fractional differencing. *Journal of Time Series Analysis, 1*(1), 15-29.

Hall, P. (1986). On the bootstrap and confidence intervals. *Annals of Statistics, 14*, 1431-1452. https://doi.org/10.1214/aos/1176350168

Hall, P. (1988). Theoretical comparison of bootstrap confidence intervals. *Annals of Statistics, 16*, 927-985. https://doi.org/10.1214/aos/1176350933

Hall, P. (1992). The bootstrap and Edgeworth expansion. Springer Series in Statistics, pp. 39-81.

Hall, P., & Martin, M. A. (1988). On bootstrap resampling and iteration. *Biometrika, 75,* 661-671. https://doi.org/10.2307/2336307

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

Hosking, J. R. M. (1981). Fractional differencing. *Biometrika, 68*, 165-176. https://doi.org/10.1093/biomet/68.1.165

Hurst, H. E. (1951). Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers, 116,* 770-808.

Ivanov, A., & Leonenko, N. (2008). Robust estimators in non-linear regression models with long-range dependence. In Springer Optimization and Its Applications (Vol. 28, pp. 193-221). Springer. https://doi.org/10.1007/978-0-387-79936-0_9

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

Knsch, H. (1986). Discrimination between monotonic trends and long-range dependence. *Journal of Applied Probability, 23*, 1025-1030. https://doi.org/10.2307/3214476

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 process. *Annals of Statistics, 31*(2), 586-612. https://doi.org/10.1214/aos/1051027882

Mandelbrot, B. B., & Van Ness, J. W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Review, 10*(4), 422-437. https://doi.org/10.1137/1010093

Martin, M. A. (1990). On the bootstrap iteration for coverage correction in confidence intervals. *Journal of the American Statistical Association, 85*, 1105-1118. https://doi.org/10.2307/2289608

McKnight, S. D., McKean, J. W., & Huitema, B. E. (2000). A double bootstrap method to analyze linear models with autoregressive error terms. *Psychological Methods, 5*(1), 87-101. https://doi.org/10.1037/1082-989X.5.1.87

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

Shao, J. (1988). On resampling methods for variance and bias estimation in linear models. *Annals of Statistics, 16*(3), 986-1008. https://www.jstor.org/stable/2241616

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

Shorack, G. R. (1982). Bootstrapping robust regression. *Communications in Statistics - Theory and Methods, 11*(9), 961-972. https://doi.org/10.1080/03610928208828286

Stute, W. (1995). Bootstrap of a linear model with AR-error structure. *Metrika, 42*, 395-410. https://doi.org/10.1007/BF01894336

Weber, N. C. (1984). On resampling techniques for regression models. *Statistics and Probability Letters, 2*(5), 275-278. https://doi.org/10.1016/0167-7152(84)90064-6

Whittle, P. (1953). The analysis of multiple stationary time series. *Journal of the Royal Statistical Society, Series B (Methodological), 15*, 125-139. https://doi.org/10.1111/j.2517-6161.1953.tb00131.x

Whittle, P. (1957). Curve and periodogram smoothing. *Journal of the Royal Statistical Society, Series B (Methodological), 19*, 38-63. https://doi.org/10.1111/j.2517-6161.1957.tb00242.x

Wu, C. J. F. (1986). Jackknife, bootstrap and other resampling methods in regression analysis (with discussion). *Annals of Statistics, 14*(4), 1261-1295. https://www.jstor.org/stable/2241454

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