Optimal Bias Correction of the Log-periodogram Estimator of the Fractional Parameter: A Jackknife Approach*

K. Nadarajah[†], Gael M. Martin[‡] & D. S. Poskitt[§]

April 1, 2020

Abstract

We use the jackknife to bias correct the log-periodogram regression (LPR) estimator of the fractional parameter in a stationary fractionally integrated model. The weights for the jackknife estimator are chosen in such a way that bias reduction is achieved without the usual increase in asymptotic variance, with the estimator viewed as 'optimal' in this sense. The theoretical results are valid under both the non-overlapping and moving-block sub-sampling schemes that can be used in the jackknife technique, and do not require the assumption of Gaussianity for the data generating process. A Monte Carlo study explores the finite sample performance of different versions of the jackknife estimator, under a variety of scenarios. The simulation experiments reveal that when the weights are constructed using the parameter values of the true data generating process, a version of the optimal jackknife estimator almost always out-performs alternative semi-parametric bias-corrected estimators. A feasible version of the jackknife estimator, in which the weights are constructed using estimates of the unknown parameters, whilst not dominant overall, is still the least biased estimator in some cases. Even when misspecified short run dynamics are assumed in the construction of the weights, the feasible jackknife still shows significant reduction in bias under certain designs. As is not surprising, parametric maximum likelihood estimation out-performs all semi-parametric methods when the true values of the short memory parameters are known, but is dominated by the semi-parametric methods (in terms of bias) when the short memory parameters need to be estimated, and in particular when the model is misspecified.

Keywords: Long memory; bias adjustment; cumulants; discrete Fourier transform; periodograms; log-periodogram regression.

MSC2010 subject classifications: Primary 62M10, 62M15; Secondary 62G09

JEL classifications: C18, C22, C52

*This research has been supported by Australian Research Council Discovery Grants No. DP150101728 and DP170100729. We would like to thank two anonymous referees for very helpful comments on an earlier version of the paper. We would also like to acknowledge the comments of participants at: the Royal Statistical Society Conference, University of Glasgow, August, 2017; and the 11th International Conference on Computational and Financial Econometrics, University of London, December, 2017. We thank the Monash e-research centre for the use of their computing facilities.

[†]Department of Economics, University of Sheffield, United Kingdom. Corresponding author: k.nadarajah@sheffield.ac.uk.

[‡]Department of Econometrics and Business Statistics, Monash University, Australia.

[§]Department of Econometrics and Business Statistics, Monash University, Australia.

1 Introduction

Data on many climate, hydrological, economic and financial variables exhibit dynamic patterns characterized by a long lasting response to past shocks. Notable examples include, water levels in rivers (Hurst, 1951), rainfall (Gil-Alana, 2012), aggregate output (Diebold and Rudebusch, 1989Hassler and Wolters, 1995), interest rates (Baillie, 1996), exchange rates (Cheung, 2016) and stock market volatility (Bollerslev and Mikkelsen, 1996; Andersen *et al.*, 2003). Such 'long memory processes' are characterized by non-summable autocovariances that decline at a (slow) hyperbolic rate, in contrast to the usual exponential, and summable, decay associated with a short memory process; the fractionally integrated autoregressive moving average (ARFIMA) model of Adenstedt (1974), Granger and Joyeux (1980) and Hosking (1981) being a popular representation. Equivalently, a stationary (potentially) long memory process, $\{Y_t\}$, $t = 0, \pm 1, \pm 2, \ldots$, can be represented by the spectral density,

$$f_{YY}(\lambda) = (2\sin(\lambda/2))^{-2d} f^*_{YY}(\lambda), \text{ for } \lambda \in [-\pi, \pi], \tag{1}$$

where the fractional differencing parameter d satisfies $d \in (-0.5, 0.5)$, and $f^*_{YY}(\cdot)$ is an even function that is continuous on $[-\pi, \pi]$, is bounded above and bounded away from zero, and satisfies $\int_{-\pi}^{\pi} \log f^*_{YY}(\lambda)\, d\lambda = 0$. The process is said to have *long memory* when $d \in (0, 0.5)$, *intermediate memory* when $d \in (-0.5, 0)$ and *short memory* when $d = 0$. The factor $f^*_{YY}(\cdot)$ controls the (remaining) short memory behaviour associated with the process. For detailed expositions of processes described by (1), including applications, see, Beran (1994), Doukhan *et al.* (2003) and Robinson (2004).

In estimating the parameter d, the semi-parametric log-periodogram regression (LPR) estimator of Geweke and Porter-Hudak (1983) and Robinson (1995a,b) has been widely used, due to the simplicity of its construction as an ordinary least squares (OLS) estimator, and its avoidance of potentially incorrect specification of the short memory component. However, consistency of the LPR estimator is achieved only at the cost of both a slower rate of convergence than the usual parametric rate and substantial finite sample bias in the presence of ignored short run dynamics (see, for example, Agiakloglou *et al.*, 1993 and Nielsen and Frederiksen, 2005).

Given this well-documented bias, *bias reduction* of the LPR estimator has been a focus of the literature. Andrews and Guggenberger (2003), for example, include additional frequencies, to degree $2r$ for $r \geq 0$, in the log-periodogram regression that defines the LPR estimator, producing an estimator (denoted hereafter by $\widehat{d}_r^{AG}$) whose bias converges to zero at a faster rate than that of the LPR estimator (recovered by setting $r = 0$), when $r > 1$. Alternative analytical procedures appear in Moulines and Soulier (1999), Hurvich and Brodsky (2001) and Robinson and Henry (2003), whilst a method based on the pre-filtered sieve bootstrap has been introduced by Poskitt *et al.* (2016). Critically, all such bias-correction methods come at a cost: namely, an increase in asymptotic variance. Notably, Guggenberger and Sun (2006) produce a weighted average of LPR estimators over different bandwidths that achieves the same degree of bias reduction as $\widehat{d}_r^{AG}$ for any given r, but with less variance inflation. This estimator, along with that of Poskitt *et al.* (2016), serve as important comparators for the alternative bias-corrected estimator that we develop herein.

The approach to bias adjustment adopted in this paper applies the jackknife principle, with the bias-corrected estimator constructed as a weighted average of LPR estimators computed, in turn, from the full sample and m sub-samples of a given length. The sub-samples may be created by using either the non-overlapping or the moving-block method. Motivated by the jackknife technique proposed by Chen and Yu (2015) in a unit root setting, weights are chosen to remove bias up to a given order and, at the same time, to minimize the increase in asymptotic variance. The weights are 'optimal' in this sense and the associated jackknife estimator referred to as 'optimal' accordingly. In the fractional setting, with the LPR estimator being the method to be adjusted, these optimal weights involve two types of covariance terms: (i) covariances between the full-sample and sub-sample log-periodogram ordinates, and (ii) covariances between distinct sub-sample log-periodogram values. These covariance terms may, in turn, be represented by cumulants of the discrete Fourier transform (DFT) of the time series. Building on results in Brillinger (1981, Chapters 2 and 4), we first derive closed-form expressions for the association between the corresponding DFTs in terms of cumulants. These expressions are used to derive the form of dependence between the periodograms (at a given ordinate or at different ordinates) associated with the full sample and the sub-samples, which allows us to obtain closed-form expressions for the covariances terms, (i) and (ii), and, hence, to evaluate the optimal weights.

We prove the consistency and asymptotic normality of the optimal jackknife estimator. Most notably, we establish that the convergence rate and asymptotic variance are equal to those of the unadjusted LPR estimator. This implies that there is *no* inflation in asymptotic efficiency compared to the *unadjusted* LPR estimator of d, despite the bias reduction that is achieved. This compares with Guggenberger and Sun (2006), in which the goal is to produce an estimator (for a given value of r) with an asymptotic variance that is smaller than that of the corresponding bias-adjusted estimator of Andrews and Guggenberger (2003), as based on the same value of r, $\widehat{d}_r^{AG}$. In particular, in the case where $r = 0$, and no bias adjustment is achieved (with $\widehat{d}_r^{AG}$ equivalent to the raw LPR estimator), the estimator of Guggenberger and Sun is still biased, but with a (possibly) reduced asymptotic variance. In addition, in contrast with Guggenberger and Sun, and the other analytical bias adjustment methods cited above, our theoretical results do not rely on the assumption of Gaussianity. Specifically, expressions for the dominant bias term and variance of the LPR estimator – needed in the construction of the jackknife estimator and as originally derived by Hurvich *et al.* (1998) for fractional *Gaussian* processes - are shown to hold under non-Gaussian assumptions. Hence, all theoretical results for the bias-adjusted estimator hold under similar generality.[1]

Extensive simulation exercises are conducted in order to compare the finite sample performance of the jackknife estimator with that of alternative approaches, including the bias-adjusted estimators of Guggenberger and Sun (2006) and Poskitt *et al.* (2016). Results show that certain versions of the

[1] We refer the reader to Hahn and Newey (2013), Chambers (2013), Chen and Yu (2015) and Robinson and Kaufmann (2015) for other applications of the jackknife in time series settings. To our knowledge the technique has been used only once in a long memory setting *per se*, namely in the numerical work of Ekonomi and Butka (2011), where the method of Chambers (2013) is adopted for the purpose of reducing the bias of the LPR estimator to the first order. However, no rigorous proofs of the properties of the estimator are provided, and no attempt at yielding an optimal estimator in the sense given in the current paper, is made.

optimally bias-corrected jackknife estimator outperform the alternative bias-adjusted estimators of Guggenberger and Sun and Poskitt *et al.*, in terms of bias-reduction and root mean squared error (RMSE), with the RMSE being somewhat close to, or even smaller than, that of the LPR in some cases. In the empirically realistic case where the true values of the parameters - required in order to evaluate the optimal weights in the jackknife estimator - are unknown, we implement the jackknife technique using an iterative procedure. This feasible version of the estimator does not consistently outperform either the bootstrap-based estimator of Poskitt *et al.* or (a feasible version of) the method of Guggenberger and Sun, but is not substantially inferior, in terms of either bias or RMSE, and is sometimes still the least biased estimator of all.

We assess the finite sample performance of all bias-adjusted estimators under scenarios of both correct model specification and misspecification and, for completeness, parametric methods based on maximum likelihood estimation (MLE) and pre-whitening are included in the assessment.[2] As would be anticipated, given the asymptotic efficiency of MLE under correct specification, no semi-parametric method out-performs the optimal parametric approach in terms of RMSE in this case. However, when the short memory dynamics need to be estimated, a semi-parametric method is typically less biased than both parametric methods. When the model is misspecified, the semi-parametric methods are dominant in terms of both bias and RMSE, with the feasible jackknife estimator producing the least bias in some cases, most notably when the true process has a moving average component that is omitted in the model specification.

In summary, the paper makes two important contributions to the literature on semi-parametric estimation in fractional models. First, a new estimator is derived that bias-corrects the popular LPR estimator to a given order, with no associated variance inflation asymptotically. Second, that estimator is shown to perform well in finite samples, under ideal conditions, and to hold its own in empirically relevant scenarios, relative to existing comparators.

The remainder of the paper is organized as follows. In Section 2, we introduce two log-periodogram regression estimators; namely, the LPR estimator originally proposed by Geweke and Porter-Hudak (1983) and the particular bias-reduced estimator of Guggenberger and Sun (2006). In Section 3, we develop the new jackknife estimator that accommodates both bias correction and variance minimization via the appropriate choice of weights. All theoretical results pertaining to the construction of the afore-mentioned covariance terms, and the resultant asymptotic properties of the optimal estimator, are given in Section 4. Section 5 documents the finite sample performance of the estimator by means of a Monte Carlo study.

The proofs of all results are contained in Appendix A, while Appendix B provides various technical results, including the evaluation of the covariances required for the construction of the weights for the optimal jackknife estimator. Appendix C contains Tables 2 to 15, which document the results of the Monte Carlo study, with these results summarized briefly in Table 16 in the same appendix. The following notation is used throughout: "$\to^P$" denotes convergence in probability, "$\to^D$" denotes

[2]We thank a referee for these suggestions.

convergence in distribution, and "$\rightarrow$" is used to indicate the limit as $n \rightarrow \infty$, (unless otherwise stated). The k^{th}-order spectral density function of the time series $\{X_t\}$ is denoted by $f_{X...X}(\lambda_1, \lambda_2, \ldots, , \lambda_{k-1})$, where $\lambda_1, \lambda_2, \ldots, , \lambda_{k-1}$ are fundamental frequencies. For instance, the density function given in (1) is the second-order spectral density of $\{Y_t\}$.

2 Log-periodogram regression estimation methods

In this section we briefly review two log-periodogram regression estimators; namely, the raw (unadjusted) LPR estimator and the bias-reduced weighted-average estimator of Guggenberger and Sun (2006) (GS). These estimators are used as benchmarks for later comparisons, and the raw LPR estimator, of course, underpins the jackknife method developed in Section 3. We summarize the asymptotic properties of these estimators and the assumptions underlying those properties. In contrast to earlier proofs related to the LPR estimator (e.g. Hurvich *et al.*, 1998) we do not assume that the data generating process (DGP) is Gaussian. This extension to non-Gaussian processes means that the properties subsequently derived for the optimal jackknife estimator are also applicable for this general case.

2.1 The log-periodogram regression estimator

Let $\mathbf{y}^\top = (y_1, y_2, ..., y_n)$ be a sample of n observations from a process with a spectral density as given in (1). The LPR estimator, $\widehat{d}_n$, is motivated by the following simple linear regression model that is formed directly from the spectral density given in (1),

$$\log I_Y^{(n)}(\lambda_j) = (\log f_{YY}^*(0) - C) - 2d \log(2\sin(\lambda_j/2)) + \xi_j, \tag{2}$$

where

$$I_Y^{(n)}(\lambda) = |D_Y^{(n)}(\lambda)|^2, \quad D_Y^{(n)}(\lambda) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^{n} y_t \exp(-\imath \lambda t), \tag{3}$$

and $D_Y^{(n)}(\lambda_j)$ is the DFT of the vector of realizations, $\mathbf{y}$, measured at Fourier frequencies, $\lambda_j = 2\pi j/n$, $(j = 1, 2..., N_n)$, $N_n = \lfloor n^\alpha \rfloor$ for $0 < \alpha < 1$, and $\imath = \sqrt{-1}$ is the imaginary unit. Here, the error terms $\xi_j = \log\left(I_Y^{(n)}(\lambda_j)/f_{YY}(\lambda_j)\right) + C + V_j$, $j = 1, 2, ..., N_n$, where

$$V_j = \log\left(f_{YY}^*(\lambda_j)/f_{YY}^*(0)\right), \tag{4}$$

are assumed to be asymptotically independently and identically distributed ($i.i.d.$) and C is the Euler constant. The LPR estimator of d is simply the OLS estimator of the slope parameter in (2) and is given by

$$\widehat{d}_n = \frac{-0.5 \sum_{j=1}^{N_n} (x_j - \overline{x}) z_j}{\sum_{j=1}^{N_n} (x_j - \overline{x})^2}, \tag{5}$$

where $z_j = \log I_Y^{(n)}(\lambda_j)$, $x_j = \log(2\sin(\lambda_j/2))$, and $\overline{x} = \frac{1}{N_n}\sum_{j=1}^{N_n} x_j$. The subscript n is introduced here in order to distinguish this full-sample version of the estimator from that computed subsequently from sub-samples, in the process of applying the jackknife.

Certain statistical properties of the LPR estimator such as its bias, variance, mean-squared-error (MSE) and asymptotic distribution have been derived by Hurvich *et al.* (1998) under given regularity conditions, and with certain approximations invoked. Alternative expressions for the bias and variance of the LPR estimator are provided in Theorem 1 of Andrews and Guggenberger (2003), plus in Theorem 3.1 of Guggenberger and Sun (2006), by setting $r = 0$. Lieberman (2001) also provides a formula for the expectation of the LPR estimator under the same conditions as Hurvich *et al.*; however, his expression is an infinite sum of a quantity that depends on the true values of d and the short memory parameters, which renders a feasible version of the jackknife technique using his expression more cumbersome.

With all results cited above derived under the assumption of Gaussianity, we now extend the results stated in Theorems 1 and 2 of Hurvich *et al.* (1998) to the general (potentially non-Gaussian) case. In particular, the resultant expression for the expectation of the LPR estimator is used in the specification of the optimal jackknife estimator, and in the proof of its properties.

We begin with the following assumptions on the DGP:

$(A.1)$ There exists $G > 0$, such that

$$f_{YY}(\lambda) = G\lambda^{-2d} + O(\lambda^{2-2d}) \text{ as } \lambda \to 0+,$$

where '$\to 0+$' denotes an approach from above.

$(A.2)$ In a neighbourhood $(0, \varepsilon)$ of the origin, $f_{YY}(\lambda)$ is differentiable on $[-\pi, \pi] \setminus \{0\}$ and

$$\left| \frac{d}{d\lambda} \log f_{YY}(\lambda) \right| = O(\lambda^{-1}), \text{ as } \lambda \to 0+ .$$

In addition, for some $0 < \widetilde{B}_2,\ \widetilde{B}_3 < \infty$, $f^{*\prime}_{YY}(0) = 0$, $|f^{*\prime\prime}_{YY}(\lambda)| < \widetilde{B}_2$ and $|f^{*\prime\prime\prime}_{YY}(\lambda)| < \widetilde{B}_3$, where $f^{*\prime}_{YY}(\lambda)$, $f^{*\prime\prime}_{YY}(\lambda)$ and $f^{*\prime\prime\prime}_{YY}(\lambda)$ denote, respectively, the first-, second- and third-order derivatives of f^{*}_{YY} with respect to λ in a neighborhood of zero.

$(A.3)$ $\{Y_t\}$, $t \in \mathbb{Z} := \{0, \pm 1, \pm 2, \cdots\}$, satisfies

$$Y_t - \mu_Y = \sum_{j=0}^{\infty} b_j \varepsilon_{t-j}, \quad \sum_{j=0}^{\infty} b_j^2 < \infty, \quad \left| \frac{d}{d\lambda} b(\lambda) \right| = O(\lambda^{-1}) \quad \text{as } \lambda \to 0+,$$

where $b(\lambda) = \sum_{j=0}^{\infty} b_j \exp(\imath j \lambda)$ and $\{\varepsilon_t\}$ is a strictly stationary process with $E(\varepsilon_t) = 0$ and $E\left(\varepsilon_t^2\right) = 1$.

$(A.4)$ The innovation process $\{\varepsilon_t\}$ satisfies the conditions in $(A.3)$. In addition, $E(\varepsilon_t)^3 < \infty$ and $E(\varepsilon_t)^4 < \infty$.

Assumptions $(A.1) - (A.3)$ are standard in the long memory literature (see, Fox and Taqqu, 1986, Hurvich *et al.*, 1998 Lieberman *et al.*, 2012, among others) and are satisfied by the class of ARFIMA models. The boundedness of the first three derivatives of f^{*}_{YY} in Assumption $(A.2)$ is required to control the fourth-order moment of the sine and cosine components of the standardized DFTs that are used to derive the bias term of the LPR. Assumption $(A.4)$ specifies the third and fourth moments of $\{\varepsilon_t\}$ to be finite, as we do not invoke Gaussianity. The boundedness imposed on the higher-order moments of $\{\varepsilon_t\}$ ensures the asymptotic normality of the DFTs associated with the process $\{Y_t\}$. The asymptotic normality of the DFTs is, in turn, used in proving Theorems 1 – 5.

We now state Theorem 1, which gives the mean, variance and asymptotic distribution of the LPR estimator. We subsequently exploit these results to construct the optimal jackknife estimator, and to prove its properties, in Section 3.

Theorem 1 *Let Assumptions* $(A.1) - (A.3)$ *hold. Given* $N_n \to \infty$, $n \to \infty$, *with* $\frac{N_n \log N_n}{n} \to 0$,

$$E(\widehat{d}_n) = d_0 - \frac{2\pi^2}{9}\frac{f_{YY}^{*\prime\prime}(0)}{f_{YY}^{*}(0)}\frac{N_n^2}{n^2} + o\Big(\frac{N_n^2}{n^2}\Big) + O\Big(\frac{\log^3 N_n}{N_n}\Big), \tag{6}$$

$$Var(\widehat{d}_n) = \frac{\pi^2}{24N_n} + o\Big(\frac{1}{N_n}\Big) \tag{7}$$

and $\widehat{d}_n \to^P d_0$. *Given that* $(A.4)$ *also holds and if* $N_n = o\left(n^{4/5}\right)$ *and* $\log^2 n = o\left(N_n\right)$, *then,*

$$\sqrt{N_n}(\widehat{d}_n - d_0) \to^D N\Big(0, \frac{\pi^2}{24}\Big) \text{ as } n \to \infty. \tag{8}$$

2.2 The weighted-average log-periodogram regression estimator

The motivation for the estimator of Guggenberger and Sun (2006) stems from the work of Andrews and Guggenberger (2003). With (4) being the term that causes the dominant bias in the LPR estimator, Andrews and Guggenberger use a Taylor series expansion around $j = 0$ to approximate (4) as an even polynomial in the frequencies of order r.[3] Including the first $2r$ terms (with $r \geq 1$) in the log-periodogram regression in (2) as additional regressors leads to

$$\ln I_Y^{(n)}(\lambda_j) = (\log f_{YY}^{*}(0) - C) - 2d\log(2\sin(\lambda_j/2)) + \sum_{k=1}^{r}\frac{b_{2k}}{(2k)!}\lambda_j^{2k} + \zeta_j, \tag{9}$$

where $\zeta_j = \xi_j - \sum_{k=1}^{r}\frac{b_{2k}}{(2k)!}\lambda_j^{2k}$. Application of OLS to (9) then yields an estimator of d, $\widehat{d}_r^{AG}$, with reduced bias relative to the raw LPR estimator, $\widehat{d}_n$. The bias-adjusted estimator is shown to be $\sqrt{N_n}$-consistent, with an asymptotic variance equal to $\frac{\pi^2}{24}c_r$, with $c_r > 1$ for $r \geq 1$ and $c_r = 1$ for $r = 0$.

Guggenberger and Sun (2006) proceed to show that an appropriate weighted average of raw LPR estimators, as based on different bandwidths, $N_{n,i} = \lfloor q_i N_n \rfloor$; $i = 1, \ldots, K$, for fixed numbers q_i chosen suitably, has the same asymptotic bias as $\widehat{d}_r^{AG}$ (constructed using N_n), but with a reduced asymptotic variance. That is, bias reduction is achieved at a smaller cost than is the original method of Andrews and Guggenberger (2003). Further, for the case of $r = 0$, the bias of the raw LPR estimator is retained but with reduced asymptotic variance. The authors also demonstrate that the weighted-average estimator, denoted by $\widehat{d}_r^{GS}$ hereafter, can be implemented via a simple two-step procedure. In the first step, a series of K LPR estimates are obtained using the regression model in (2) and for

[3]The odd-order terms of the Taylor's expansion around zero are exactly zero. This leads to the expansion with only even-order terms.

bandwidths, $N_{n,i}$, $i = 1, \ldots, K$. Then, in the second step, the following pseudo-regression is estimated, using the K estimates produced in the first step as observations of the dependent variable in the regression,

$$\widehat{d}_{N_{n,i}} = d + \sum_{j=1}^{r} \beta_{2j} q_i^{2j} + \beta_{2+2r} \left(q_i^{2+2r} - \delta \sum_{p=1}^{K} q_p^{2+2r} \right) + u_i, \; i = 1, \ldots, K, \tag{10}$$

where u_i is the error term, and $\mathbf{u}^\top = (u_1, u_2, ..., u_K)$ has a zero (vector) mean and asymptotic variance-covariance matrix,

$$\boldsymbol{\Omega} = (\Omega_{i,j}) \in \mathbb{R}^{K \times K}, \text{ with } \Omega_{i,j} = \frac{1}{\max(q_i, q_j)}.$$

The tuning parameter δ on the right-hand-side of (10) is a fixed non-zero constant that is used to control the multiplicative constant of the dominant bias term and render that term equivalent to the dominant bias term of $\widehat{d}_r^{AG}$. The estimator, $\widehat{d}_r^{GS}$, is then defined as the first component of the GLS estimator of $\left(d, \beta^\top\right)^\top$, where $\beta^\top = (\beta_2, \beta_4, \ldots, \beta_{2+2r})$, that is,

$$\left(\widehat{d}_r^{GS}, \widehat{\beta}^\top\right)^\top = \left(\mathbf{Z}^\top \boldsymbol{\Omega}^{-1} \mathbf{Z}\right)^{-1} \mathbf{Z}^\top \boldsymbol{\Omega}^{-1} \widehat{\mathbf{d}}, \tag{11}$$

where $\widehat{\mathbf{d}}$ is the $(K \times 1)$ dimensional vector with i^{th} element $\widehat{d}_{N_{n,i}}$, and

$$\mathbf{Z}^\top = \left(\mathbf{z}_1, \ldots, \mathbf{z}_K\right) \in \mathbb{R}^{(2+r) \times K}, \text{ with } \mathbf{z}_i^\top = \left(1, q_i^2, \ldots, q_i^{2r}, \left(q_i^{2+2r} - \delta \sum_{p=1}^{K} q_p^{2+2r}\right)\right).$$

Both the raw LPR estimator, $\widehat{d}_n$, and the weighted-average estimator, $\widehat{d}_r^{GS}$, with $r = 1$, are used as comparators of our proposed jackknife procedure in the Monte Carlo simulation exercises in Section 5.

3 The optimal jackknife log-periodogram regression estimator

3.1 Definition of the jackknife estimator

The idea behind jackknifing is to generate a set of sub-samples, by deleting one or more observations of the original sample, while preserving the structure of dependence within the sub-samples; the aim being to use (weighted) sub-sample estimates to produce a bias-corrected estimator of the parameter of interest. Let $\mathbf{y}_i$ $(i = 1, 2, ..., m)$ denote a set of m sub-samples of $\mathbf{y}$, each of which has equal length, l, such that $n = l \times m$. If sub-samples are chosen using the 'non-overlapping' method, then $\mathbf{y}_i^\top = \left(y_{(i-1)l+1}, \ldots, y_{il}\right)$ for $i = 1, \ldots, m$; alternatively if the sub-sampling scheme is 'moving-block' then $\mathbf{y}_i^\top = (y_i, \ldots, y_{i+l-1})$ for all i. In the current context we use the jackknife technique to bias correct the LPR estimator. Hence, we need to produce the full-sample estimator, $\widehat{d}_n$, and the LPR estimators produced by applying OLS to the model in (2), using the relevant sub-sample. We denote these m sub-sample estimators (based on either the non-overlapping or moving-block method) by $\widehat{d}_i$, $i = 1, 2, ..., m$. We summarize notation corresponding to the full-sample estimation and both forms of sub-sample estimation in Table 1, for ease of subsequent referencing.

Table 1: Quantities related to the full sample and the sub-samples used in the construction of the jackknife estimator

		Full sample	i^{th} sub-sample
(i)	Frequency	$\lambda_j = 2\pi j/n$	$\mu_j = 2\pi j/l = 2\pi jm/n = m\lambda_j$
(ii)	Frequency range	$j = 1, ..., N_n$	$j = 1, ..., N_l$
(iii)	Spectral density	$f_{YY}(\lambda) = (2\sin(\lambda/2))^{-2d} f^*_{YY}(\lambda)$	$f_{Y_iY_i}(\mu) = (2\sin(\mu/2))^{-2d} f^*_{Y_iY_i}(\mu)$
(iv)	DFT	$D^{(n)}_Y(\lambda) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^{n} y_t \exp(-\imath\lambda t)$	$D^{(l)}_{Y_i}(\mu) = \frac{1}{\sqrt{2\pi l}} \sum_{t=1}^{l} y_{t+i'} \exp(-\imath\mu t)$
(v)	Periodogram	$I^{(n)}_Y(\lambda) = \|D^{(n)}_Y(\lambda)\|^2$	$I^{(l)}_{Y_i}(\mu) = \|D^{(l)}_{Y_i}(\mu)\|^2$
(vi)	Error term	$\xi_j = \log\left(I^{(n)}_Y(\lambda_j)/f_{YY}(\lambda_j)\right)$	$\xi^{(i)}_j = \log\left(I^{(l)}_{Y_i}(\mu_j)/f_{Y_iY_i}(\mu_j)\right)$
	Other notation:		
(vii)		$x_j = \log(2\sin(\lambda_j/2))$	$x'_j = \log(2\sin(\mu_j/2))$
$(viii)$		$\overline{x} = \sum_{t=1}^{N_n} x_j \Big/ N_n$	$\overline{x'} = \sum_{t=1}^{N_l} x'_j \Big/ N_l$
(ix)		$a_j = x_j - \overline{x}$	$a'_j = x'_j - \overline{x'}$
(x)		$S_{xx} = \sum_{j=1}^{N_n} a_j^2$	$S'_{xx} = \sum_{j=1}^{N_l} a'^2_j$

Note, regarding the sub-sample notation in point (iv), if the sub-samples are drawn with the non-overlapping scheme then, $i' = (i-1)l$. If the moving-block scheme is used then, $i' = i - 1$.

Define the jackknife estimator, $\widehat{d}_{J,m}$, as

$$\widehat{d}_{J,m} = w_n \widehat{d}_n - \sum_{i=1}^{m} w_i \widehat{d}_i, \tag{12}$$

where w_n and $\{w_i\}_{i=1}^m$ are the weights assigned to the full-sample estimator and the sub-sample estimators, respectively. Re-iterating, $\widehat{d}_n$ is the LPR estimator obtained from the full sample (as defined directly in (5)) and $\widehat{d}_i$ $(i = 1, 2, ..., m)$ denotes the i^{th} sub-sample LPR estimator. Under the conditions of Theorem 1, it is straightforward to show that

$$\begin{aligned} E\big(\widehat{d}_{J,m}\big) &= \big(w_n - \sum_{i=1}^{m} w_i\big) d_0 - \Big(\frac{2\pi^2}{9}\frac{f^{*\prime\prime}_{YY}(0)}{f^*_{YY}(0)}\frac{N_n^2}{n^2} w_n - \frac{2\pi^2}{9}\frac{f^{*\prime\prime}_{Y_iY_i}(0)}{f^*_{Y_iY_i}(0)}\frac{N_l^2}{l^2}\sum_{i=1}^{m} w_i\Big) \\ &\quad + o\Big(\frac{N_n^2}{n^2}\Big) + O\Big(\frac{\log^3 N_n}{N_n}\Big), \end{aligned} \tag{13}$$

and

$$\begin{aligned} Var\big(\widehat{d}_{J,m}\big) &= \frac{\pi^2}{24 N_n} w_n^2 + \frac{\pi^2}{24 N_l}\sum_{i=1}^{m} w_i^2 + 2\sum_{i=1}^{m-1}\sum_{j=i+1}^{m} w_i w_j Cov\big(\widehat{d}_i, \widehat{d}_j\big) \\ &\quad - 2 w_n \sum_{i=1}^{m} w_i Cov\big(\widehat{d}_n, \widehat{d}_i\big) + o\Big(\frac{1}{N_n}\Big). \end{aligned} \tag{14}$$

The covariance between the full-sample LPR estimator and each sub-sample LPR estimator, $Cov\left(\widehat{d}_n, \widehat{d}_i\right)$, and the covariances between the different sub-sample LPR estimators, $Cov\left(\widehat{d}_i, \widehat{d}_j\right)$, for $i \neq j$,

$i, j = 1, 2, ..., m$, are given respectively by,

$$Cov\big(\widehat{d}_n, \widehat{d}_i\big) = \frac{1}{4S_{xx}}\frac{1}{S'_{xx}}\sum_{j=1}^{N_n}\sum_{k=1}^{N_l} a_j a_k^{(i)} Cov\big(\log I_Y^{(n)}(\lambda_j), \log I_{Y_i}^{(l)}(\mu_k)\big) \quad (15)$$

$$Cov\big(\widehat{d}_i, \widehat{d}_{i'}\big) = \frac{1}{4}\frac{1}{(S'_{xx})^2}\sum_{j=1}^{N_l}\sum_{k=1}^{N_l} a'_j a'_k Cov\big(\log I_{Y_i}^{(l)}(\mu_j), \log I_{Y_{i'}}^{(l)}(\mu_k)\big), \quad (16)$$

with all notation as defined in Table 1.

Our aim is to obtain the set of weights, $\{w_n, w_1, \ldots, w_m\}$, such that $\widehat{d}_{J,m}$ has the following properties:

$(P.1)$ $\widehat{d}_{J,m}$ is an asymptotically unbiased estimator of d_0, with bias reduced to an order of $o(N_n^2/n^2)$, and,

$(P.2)$ $\widehat{d}_{J,m}$ achieves minimum variance among all such bias-reduced estimators.

The 'optimal' jackknife estimator so defined is derived via the Lagrangian method in the following section. In Section 4, the asymptotic properties of the covariances in (B.1) and (B.2) that determine the asymptotic behaviour of the estimator are derived, and the asymptotic efficiency of the estimator then proven.

3.2 Derivation of the optimal estimator

The minimization problem is formulated as follows. Produce weights, $\{w_n, w_1, \ldots, w_m\}$, that satisfy:

$$\min_{w_n, \{w_i\}_{i=1}^m} Var\big(\widehat{d}_{J,m}\big), \quad (17)$$

subject to two constraints

$$g^1(w_n, w_1, \ldots, w_m) = w_n - \sum_{i=1}^{m} w_i - 1 = 0, \quad (18)$$

$$g^2(w_n, w_1, ..., w_m) = \frac{N_n^2}{n^2}w_n - m^2\frac{N_l^2}{l^2}\sum_{i=1}^{m} w_i = 0. \quad (19)$$

We refer to the optimal estimator so produced as $\widehat{d}_{J,m}^{Opt}$ hereinafter.

Constraints (18) and (19) ensure that Property $(P.1)$ holds for the resultant estimator. Specifically, (18) ensures that $\widehat{d}_{J,m}^{Opt}$ is asymptotically unbiased for d_0, as can be seen by inspection of (13). The dominant bias term of $\widehat{d}_{J,m}^{Opt}$ will be eliminated if and only if the second component appearing in (13) is set to zero; that is, if and only if

$$\frac{2\pi^2}{9}\frac{f_{YY}^{*\prime\prime}(0)}{f_{YY}^{*}(0)}\frac{N_n^2}{n^2}w_n - \frac{2\pi^2}{9}\frac{f_{Y_iY_i}^{*\prime\prime}(0)}{f_{Y_iY_i}^{*}(0)}\frac{N_l^2}{l^2}\sum_{i=1}^{m} w_i = 0. \quad (20)$$

Using Point (iii) of Table 1, we have that $f_{Y_iY_i}^{*}(0) = f_{YY}^{*}(0)$ and $f_{Y_iY_i}^{*\prime\prime}(0) = m^2 f_{YY}^{*\prime\prime}(0)$. Hence, the condition in (20) collapses to constraint (19). Given (17), Property $(P.2)$ is satisfied by construction.

Henceforth writing, $Cov\big(\widehat{d}_n,\widehat{d}_i\big)=c^*_{n,i}$ and $Cov\big(\widehat{d}_i,\widehat{d}_{i'}\big)=c^{\dagger}_{i,j}$, such that $c^{\dagger}_{i,j}=c^{\dagger}_{j,i}$, the Lagrangian function is given by,

$$\begin{aligned}\tilde{L}\left(w_n,w_1,\ldots,w_m,\delta_1,\delta_2\right) &= \frac{\pi^2}{24N_n}w_n^2+\frac{\pi^2}{24N_l}\sum_{i=1}^{m}w_i^2+2\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_jc^{\dagger}_{i,j}\\ &\quad -2w_n\sum_{i=1}^{m}w_ic^*_{n,i}+\delta_1\big(w_n-\sum_{i=1}^{m}w_i-1\big)\\ &\quad +\delta_2\Big(\frac{N_n^2}{n^2}w_n-m^2\frac{N_l^2}{l^2}\sum_{i=1}^{m}w_i\Big). \end{aligned} \tag{21}$$

The first-order conditions (FOCs) are thus given by,

$$\begin{aligned}\frac{\partial\tilde{L}}{\partial\delta_1} &= 0\Rightarrow w_n-\sum_{i=1}^{m}w_i=1,\\ \frac{\partial\tilde{L}}{\partial\delta_2} &= 0\Rightarrow \frac{N_n^2}{n^2}w_n-m^2\frac{N_l^2}{l^2}\sum_{i=1}^{m}w_i=0,\\ \frac{\partial\tilde{L}}{\partial w_n} &= 0\Rightarrow \frac{2\pi^2}{24N_n}w_n-2\sum_{i=1}^{m}w_ic^*_{n,i}+\delta_1+\frac{N_n^2}{n^2}\delta_2=0,\\ \frac{\partial\tilde{L}}{\partial w_{i,m}} &= 0\Rightarrow -2w_nc^*_{n,i}+\frac{2\pi^2}{24N_l}w_i+2\sum_{j=1,j\neq i}^{m}w_jc^{\dagger}_{i,j}-\delta_1-m^2\frac{N_l^2}{l^2}\delta_2=0;\ i=1,\ldots,m.\end{aligned}$$

Defining

$$\mathbf{A}=\begin{bmatrix} 1 & -1 & \ldots & -1 & 0 & 0\\ \frac{N_n^2}{n^2} & -m^2\frac{N_l^2}{l^2} & \ldots & -m^2\frac{N_l^2}{l^2} & 0 & 0\\ \frac{\pi^2}{12N_n} & -2c^*_{n,1} & \ldots & -2c^*_{n,m} & 1 & \frac{N_n^2}{n^2}\\ -2c^*_{n,1} & \frac{\pi^2}{12N_l} & \ldots & 2c^{\dagger}_{1,m} & -1 & -m^2\frac{N_l^2}{l^2}\\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots\\ -2c^*_{n,m} & 2c^{\dagger}_{1,m} & \ldots & \frac{\pi^2}{12N_l} & -1 & -m^2\frac{N_l^2}{l^2} \end{bmatrix},\ \mathbf{w}=\begin{bmatrix} w_n\\ w_1\\ \vdots\\ w_m\\ \delta_1\\ \delta_2\end{bmatrix}\text{ and }\mathbf{b}=\begin{bmatrix}1\\0\\0\\ \vdots\\0\\0\end{bmatrix}, \tag{22}$$

the optimal solution, $\mathbf{w}^*=\begin{bmatrix} w_n^* & w_1^* & \ldots & w_m^* & \delta_1^* & \delta_2^*\end{bmatrix}^{\top}$, is given by

$$\mathbf{w}^*=\mathbf{A}^{-1}\mathbf{b}. \tag{23}$$

Given the structure of $\mathbf{b}$ this means that the solutions for the weights are given by the elements of the first column of $\mathbf{A}^{-1}$, and the optimal jackknife estimator is accordingly given as:

$$\widehat{d}^{Opt}_{J,m}=w_n^*\widehat{d}_n-\sum_{i=1}^{m}w_i^*\widehat{d}_i, \tag{24}$$

where $w_n^*=\Big[1-\left(N_nl/\left(N_lmn\right)\right)^2\Big]^{-1}$, given immediately by solving the first two FOCs.

To complete the result we need to show that (23) is a local minimizer of $\tilde{L}\left(\cdot\right)$. To do so, we need to show that: (i) the constraint qualification – that the rank of the matrix formed by the first-order derivatives at the solution of the constraints with respect to parameters, except the Lagrangian parameters, is equal to the number of conditions – is met, (ii) the solution of the Lagrangian function satisfies the FOCs, and, (iii) the leading principal minors of the bordered Hessian matrix, $\mathbf{H}^B_{(m+3)\times(m+3)}$, all take the same sign of $(-1)^k$, where k is the number of constraints (see, Chapter 12 of Chiang and Wainwright, 2005, for more details).

In our problem, the number of constraints equals 2 and

$$Rank\begin{bmatrix} \frac{\partial g^1}{\partial w_n} & \frac{\partial g^2}{\partial w_n} \\ \frac{\partial g^1}{\partial w_1} & \frac{\partial g^2}{\partial w_1} \\ \vdots & \vdots \\ \frac{\partial g^1}{\partial w_m} & \frac{\partial g^2}{\partial w_m} \end{bmatrix} = Rank\begin{bmatrix} 1 & 1 \\ \frac{N_n^2}{n^2} & m^2\frac{N_l^2}{l^2} \\ \vdots & \vdots \\ \frac{N_n^2}{n^2} & m^2\frac{N_l^2}{l^2} \end{bmatrix} = 2.$$

Hence, the rank condition is met. The second condition is met by default. The important condition is the third one, where we need to show that the leading principal minors of $\mathbf{H}^B_{(m+3)\times(m+3)}$, exceed zero for every $m = 2, 3, \ldots$. The bordered Hessian matrix for our case is given by

$$\mathbf{H}^B_{(m+3)\times(m+3)} = \begin{bmatrix} 0 & 0 & 1 & -1 & \cdots & -1 \\ 0 & 0 & \frac{N_n^2}{n^2} & -m^2\frac{N_l^2}{l^2} & \ldots & -m^2\frac{N_l^2}{l^2} \\ 1 & \frac{N_n^2}{n^2} & \frac{\pi^2}{12N_n} & -2c^*_{n,1} & \cdots & -2c^*_{n,m} \\ -1 & -m^2\frac{N_l^2}{l^2} & -2c^*_{n,1} & \frac{\pi^2}{12N_l} & \cdots & 2c^\dagger_{1,m} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & -m^2\frac{N_l^2}{l^2} & -2c^*_{n,m} & 2c^\dagger_{1,m} & \cdots & \frac{\pi^2}{12N_l} \end{bmatrix}.$$

The proof of positivity of the principal minors of the above matrix is given in Appendix B. Hence, the solution in (23) is a local minimizer of $\tilde{L}(\cdot)$.

We complete this section with three remarks:

Remark 1 *If we consider only bias reduction to the order N_n^2/n^2, without concurrent variance reduction; that is, we produce an estimator that satisfies only (P.1), and not (P.2), then the formulae for the weights are*

$$w_n^* = \left[1 - \left(\frac{N_n}{N_l}\frac{l}{nm}\right)^2\right]^{-1} \text{ and } w_i^* = \frac{1}{m}\left(w_n^* - 1\right), \text{ for } i = 1, \ldots, m. \tag{25}$$

These weights mimic those of Chambers (2013) in the short memory setting (under a non-overlapping sub-sampling scheme), in which variance minimization was not a consideration.

Remark 2 *When Chambers (2013) considers the moving-block sub-sampling scheme (again, in the short memory setting), he chooses the sub-sample length to be $l = n - m + 1$. In this case, when n is large and m is small, the sub-sample length is $l \approx n$, and the impact of bias correction is reduced as a consequence; something that is in evidence in the Monte Carlo simulation results reported by that author. As a result of this observation, in our investigations we use the common sub-sample length of $l = n/m$, under both the non-overlapping and moving-block schemes.*

Remark 3 *Condition* 3.3 *of Guggenberger and Sun (2006) has a similar purpose to our (19). The difference is that we eliminate the* $O\left(N_n^2 / n^2 \right)$ *term from the bias of the LPR estimator, whereas they eliminate bias up to an order of* N_n^{2r} / n^{2r}*, for some* $r \geq 1$*. The role played by (17) is somewhat different from that played by Condition* 3.4 *of Guggenberger and Sun (2006). The latter condition is imposed mainly to link the bias and variance of* $\widehat{d}_r^{GS}$ *to that of* $\widehat{d}_r^{AG}$*, for any given* r*; this link occurring via the introduction of the tuning parameter,* δ *(see (10) above), on which the finite sample performance of their estimator depends. In our method, (17) is used to control the increase in variance that occurs due to the reduction in bias, with the optimal weights determined by (17)-(19) not depending on any arbitrary quantities.*

4 Asymptotic results

The asymptotic properties of the optimal jackknife estimator depend on the optimal weights which, in turn, are functions of the covariance terms between the log-periodograms associated with the full sample and the sub-samples, as seen in (B.1) and (B.2). Provided that the DGP satisfies assumptions $(A.1) - (A.3)$, Lahiri (2003) has shown that periodogram ordinates are asymptotically independent when the frequencies are at a sufficient distance apart, provided that the set of observations remain the same. However, in our case, we are dealing with periodograms calculated both for the full set of observations, and for subsets of the full set. Thus, two questions that arise here are: (i) Are the periodograms of the full sample and the sub-samples at different frequency ordinates asymptotically independent? and, (ii) When $d \neq 0$, do the periodograms still converge to a chi-square distribution as they do when $d = 0$ (see Theorem 5.2.6 of Brillinger, 1981)? We address both questions in Section 4.1 and provide formulae for calculating the relevant covariance terms algebraically, adopting the procedure used in Brillinger (1981). In Section 4.2 we then use these results to derive the asymptotic properties of the optimal jackknife estimator.

4.1 Stochastic properties of periodograms in the full sample and in sub-samples

We begin by defining $\{X_1, X_2, \ldots, X_h\}$ as an arbitrary set of h stationary time series. We link these series to the full sample and the m sub-samples of observations below. Our use of notation in this section mimics, in large part, that of Brillinger (1981, §. 2.6).

Definition 1 *Suppose* $\{X_1, X_2, \ldots, X_h\}$ *is a set of* h *stationary time series. The* k^{th}*-order cumulant* $\kappa_{X_{a_1},\ldots,X_{a_k}}(u_1, \ldots, u_{k-1})$*, for* $k = 1, 2, \ldots, h,$ *and* $u_j = 0, \pm 1, \pm 2\ldots$ *for* $j = 1, 2, \ldots, k-1$*, is defined as follows,*

$$\kappa_{X_{a_1},\ldots,X_{a_k}}(u_1, \ldots, u_{k-1}) = \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \exp\left(-\imath \sum_{j=1}^{k-1} \lambda_j u_j \right) f_{X_{a_1},\ldots,X_{a_k}}(\lambda_1, \ldots, \lambda_{k-1})\, d\lambda_1 \ldots d\lambda_{k-1}, \quad (26)$$

where $f_{X_{a_1},\ldots,X_{a_k}}(\lambda_1, \ldots, \lambda_{k-1})$ *is the* k^{th}*-order joint spectral density of* $\{X_{a_1}, \ldots, X_{a_k}\}$*, for* $-\pi < \lambda_j < \pi$*,* $j = 1, 2, \ldots, k-1$*, with* $a_1, \ldots, a_k = 1, 2, \ldots, h,$ *and* $k = 1, 2, \ldots$.

For $\sum_{u_1=-\infty}^{\infty}\cdots\sum_{u_{k-1}=-\infty}^{\infty}\left|\kappa_{X_{a_1},\ldots,X_{a_k}}\left(u_1,\ldots,u_{k-1}\right)\right|<\infty$, *then the inverse form of (26) is given by,*

$$f_{X_{a_1},\ldots,X_{a_k}}\left(\lambda_1,\ldots,\lambda_{k-1}\right)=(2\pi)^{-k+1}\sum_{u_1=-\infty}^{\infty}\cdots\sum_{u_{k-1}=-\infty}^{\infty}\kappa_{X_{a_1},\ldots,X_{a_k}}\left(u_1,\ldots,u_{k-1}\right)\exp\Big(-\imath\sum_{j=1}^{k-1}\lambda_j u_j\Big). \tag{27}$$

Now let $X_1=\mathbf{y}$ denote the full sample of n observations on the random variable following the model in (1); whilst $X_{1+i}=\mathbf{y}_i$ denotes the vector of observations for the sub-sample $i=1,2,\ldots,m$, with length l. Set $h=m+1$ in Definition 1. Let $D_{X_1}^{(n)}(.)$ and $D_{X_{1+i}}^{(l)}(.)$ respectively be the DFT of the full sample and i^{th} sub-sample at some frequency. Set

$$L_i=\begin{cases} n & if\ i=1\\ l & otherwise\end{cases}. \tag{28}$$

In Proposition 1 we give the expression for the k^{th}-order joint cumulant of the DFTs of the $h=m+1$ series associated with the full sample and the m sub-samples.

Proposition 1 *Suppose Assumptions* $(A.1)-(A.3)$ *hold. The* k^{th}*-order cumulant of* $\big\{D_{X_{a_1}}^{(L_1)}(\lambda_1)$, $D_{X_{a_2}}^{(L_2)}(\lambda_2)$, ..., $D_{X_{a_k}}^{(L_k)}(\lambda_k)\big\}$, *for* $k=1,2,\ldots,$ *is given by,*

$$\kappa_{D_{X_{a_1}},\ldots,D_{X_{a_k}}}\left(\lambda_1,\ldots,\lambda_{k-1}\right)=L^{-\frac{k}{2}}(2\pi)^{\frac{k}{2}-1}\Delta^{(L)}\Big(\sum_{j=1}^{k}\lambda_j\Big)f_{X_{a_1},\ldots,X_{a_k}}\left(\lambda_1,\ldots,\lambda_{k-1}\right)+o\Big(L^{1-2d-\frac{k}{2}}\Big), \tag{29}$$

where, $L=\min\{L_1,\ldots,L_k\}$.[4]

From Proposition 1 we can derive the relationship between the DFTs corresponding to full sample and the m sub-samples as the sample size increases. The result is given in the following theorem:

Theorem 2 *Suppose Assumptions* $(A.1)-(A.4)$ *hold, and suppose* $\lambda=2\pi r/L_i$ *and* $\omega=2\pi s/L_j$ *for integers* r *and* s*. Then for a fixed value of* L_i *and* L_j, $D_{X_{a_i}}^{(L_i)}(\lambda)$ *and* $D_{X_{a_j}}^{(L_j)}(\mu)$ *are asymptotically independent, whenever* $\max\{L_i\lambda, L_j\mu\}\to\infty$, *for* $i\neq j$.

Theorem 2 immediately implies the asymptotic independence of the periodograms of the full sample and all sub-samples. However, in finite samples, the dependence structure across these periodograms may play an important role in determining the variance of the jackknife estimator in (14), through the form of the covariances in (B.1) and (B.2). Expressions for the covariances between the periodograms corresponding to the full sample and the sub-samples are provided in the following theorem, from which further insights on this point can be gleaned.

[4] The k^{th}-order cumulant associated with the DFTs should, for completeness, be denoted by $\kappa_{D_{X_{a_1}}^{(L_1)},\ldots,D_{X_{a_k}}^{(L_k)}}(.,\ldots,.)$. For notational ease, however, we express the cumulant without making explicit the relevant sample sizes.

Theorem 3 *Let $I^{(L_i)}_{X_{a_i}}(\lambda)$ and $I^{(L_j)}_{X_{a_j}}(\lambda)$ be the periodograms associated with DFTs $D^{(L_i)}_{X_{a_i}}(\lambda)$ and $D^{(L_j)}_{X_{a_j}}(\mu)$ respectively. Suppose Assumptions $(A.1)-(A.3)$ hold. Then,*

$$
\begin{aligned}
Cov\left(I^{(L_i)}_{X_{a_i}}(\lambda), I^{(L_j)}_{X_{a_j}}(\mu)\right) &= \frac{2\pi}{L} f_{X_{a_i},X_{a_i},X_{a_j},X_{a_j}}(\lambda,-\lambda,\mu) + \frac{2\pi}{L}\left[\eta(\lambda-\mu)+\eta(\lambda+\mu)\right]\left\{f_{X_{a_i}X_{a_j}}(\lambda)\right\}^2 \\
&\quad +2\pi\left[\eta(\lambda-\mu)+\eta(\lambda+\mu)\right] f_{X_{a_i}X_{a_j}}(\lambda)\, o\left(L^{-2d}\right) + o\left(L^{-1-2d}\right), \qquad (30)
\end{aligned}
$$

where $\eta(\omega) = \lim_{T\to\infty} \frac{1}{2\pi}\sum_{t=-T}^{T}\exp\{-\imath\omega t\}$, and L is as defined in Proposition 1. When Assumption $(A.4)$ also holds, the periodogram ordinates $I^{(L_i)}_{X_{a_i}}(\mu)$ and $I^{(L_j)}_{X_{a_j}}(\omega)$ with $i \neq j$, are asymptotically $f_{X_1X_1}(\cdot)\,\chi^2_{(2)}\big/ 2$ random variables.

Theorem 3 is a generalization of the result of Theorem 5.2.6 of Brillinger (1981) to the context of jackknifing. Equation (30) provides the first few dominant terms of the covariance between the periodograms associated with the full sample and a particular sub-sample, or between distinct sub-samples, at various frequency ordinates. Further, (30) reflects the fact that, for finite n, the relevant periodograms are positively correlated. This result is to be anticipated given that the sub-samples are subsets of the full sample and, hence, retain the same dependence structure as the full sample. Furthermore, the theorem states that the periodogram ordinates (for either the full sample and a given sub-sample, or between sub-samples) have a limiting joint distribution of the form, $f_{X_1X_1}(\lambda)\,\chi^2_{(2)}\big/ 2$, where $f_{X_1X_1}(.)$ is the spectral density of the time series from which the full sample is generated.

Using the covariance terms and the distribution of the periodograms provided in the above theorem, we can find the joint distribution of the log-periodograms associated with the full sample and any sub-sample (or for two distinct sub-samples). Using the joint distribution of the log-periodograms, we can derive the moment generating function of the joint distribution. This leads to the derivation of the covariance terms for the log-periodogram. This result is provided in Appendix B. The covariances between log-periodograms allow us to obtain the covariances between the full-sample and sub-sample LPR estimators given in (B.1) and (B.2). Exploiting the relationship between the different LPR estimators, we then establish the consistency and asymptotic normality of the optimal jackknife estimator in the following section.

4.2 Asymptotic properties of the optimal jackknife estimator

Using the results established in the previous section, we state the relationship between the full-sample and sub-sample LPR estimators in Theorem 4. The asymptotic properties of the optimal jackknife estimator are then established in Theorem 5.

Theorem 4 *Let $\widehat{d}_n$ and $\widehat{d}_i$ be the LPR estimators for the full sample and the i^{th} sub-sample with sub-sample length, l. Suppose Assumptions $(A.1)-(A.4)$ hold. Then, for a fixed value of m,*

(i) $\widehat{d}_n$ and $\widehat{d}_i$ are asymptotically independent.

(ii) $\widehat{d}_i$ and $\widehat{d}_j$ for $i \neq j$, $i, j = 1, \ldots, m$, are asymptotically independent.

From Theorem 1, the LPR estimator constructed from the full sample is consistent and satisfies (8). Similarly, allowing the number of sub-samples, m, to be fixed (hence l changes as n changes such that $n = m \times l$), as $l \to \infty$, $\widehat{d}_i \to^P d_0$, and $\sqrt{N_l}\big(\widehat{d}_i - d_0\big) \to^D N\Big(0, \frac{\pi^2}{24}\Big)$. This implies the sub-sample LPR estimators have the same limiting distribution as the full-sample estimator. The asymptotic properties of $\widehat{d}_{J,m}^{Opt}$ are given in the following theorem.

Theorem 5 *Under the same assumptions and conditions given in Theorem 1, for a fixed value of* m*,*

$$\widehat{d}_{J,m}^{Opt} \to^P d_0, \text{ and } \sqrt{N_n}\big(\widehat{d}_{J,m}^{Opt} - d_0\big) \to^D N\Big(0, \frac{\pi^2}{24}\Big) \text{ as } n \to \infty$$

where d_0 *is the true value of* d *and* $\widehat{d}_{J,m}^{Opt}$ *is as given in (24).*

Thus, it follows from Theorem 5 that $\widehat{d}_{J,m}^{Opt}$ is consistent for d_0 and achieves a limiting normal distribution with the same variance as the base LPR estimator itself. Further, the rate of convergence of the optimal jackknife estimator, $\sqrt{N_n}$, is also the same as that of the LPR estimator. That is, there is no loss of asymptotic efficiency compared to $\widehat{d}_n$. Importantly, these asymptotic properties of the jackknife estimator do not depend on the number of sub-samples or the sub-sample length, as long as the former is fixed and the latter increases with n such that that $n = m \times l$.

5 Simulation exercise

In this section, Monte Carlo simulation is used to compare the finite sample performance of the proposed jackknife estimator with: (i) the weighted-average estimator of Guggenberger and Sun (2006), $\widehat{d}_r^{GS}$, with $r = 1$, (ii) the bias-corrected pre-filtered sieve bootstrap-based estimator of Poskitt *et al.* (2016), $\widehat{d}^{PFSB}$, (iii) the unadjusted LPR estimator, $\widehat{d}_n$, (iv) the MLE, $\widehat{d}^{MLE}$, and, (v) the pre-whitened (PW) estimator, $\widehat{d}^{PW}$. Performance is assessed in terms of bias and RMSE, and under a variety of true DGPs. In Section 5.1 details of the basic Monte Carlo design are provided. Results under correct and incorrect specification of the model are then documented in Section 5.2 and Section 5.3 respectively, with further computational details that pertain to those specific settings provided therein. In order to assist the reader, we tabulate a ranking of the different estimators, under the variety of settings considered, in Section 5.4. All numerical results are produced using MATLAB 2015b, version 8.6.0.267246, and all tables of results are collected in Appendix C.

5.1 Monte Carlo design

Data are generated from various versions of a Gaussian fractional process, ARFIMA(p_0, d_0, q_0), where p_0 is the lag length of the true autoregressive (AR) component and q_0 the lag length of the true moving average (MA) component. The lag lengths p_0 and q_0 equal either one or zero in all settings. For $p_0 = q_0 = 1$, the process is given by

$$(1 + \phi_0 B)\,(1 - B)^{d_0}\,(Y_t - \mu_0) = (1 + \theta_0 B)\,\varepsilon_t, \tag{31}$$

where B is the backward shift operator, $B^k x_t = x_{t-k}$, for $k = 1, 2, \ldots$, and $\varepsilon_t \sim i.i.d\ N(0,1)$. Here, μ_0 is the mean parameter for Y_t, and without loss of generality we assume that $\mu_0 = 0$. For the parameter of interest, d, we select true values from the set, $d_0 = \{-0.25, 0, 0.25, 0.45\}$. Values from the set $\{-0.9, -0.4, 0.4, 0.9\}$ are adopted for both ϕ_0 and θ_0.[5] Additional details are provided in Sections 5.2 and 5.3.

Sample sizes $n \in \{96, 576\}$ are considered. These values are chosen to reflect the size of samples used in real world examples (see, for example, Diebold *et al.*, 1991, Delgado and Robinson, 1994, Gil-Alana and Robinson, 1997, and Reisen and Lopes, 1999). However, one should note that, in general, the size of data sets from finance, in particular those recorded at high frequency (for example, Granger and Hyung, 2004), or from biology (for example, the tree-ring data set of Contreras-reyes and Palma, 2013), or in certain other of the examples mentioned in the Introduction, are much larger than the sample sizes considered here. On the other hand, these sample sizes are large enough to enable a range of values for the number of sub-samples, m, to be explored, with the chosen range of m being $\{2, 3, 4, 6, 8\}$. We also consider only sub-samples that have equal length, $l = n/m$, under both sub-sampling approaches, with the slightly unorthodox values of $n \in \{96, 576\}$ chosen in order to ensure that l is an integer.

We adopt the following procedure in implementing the *optimal* jackknife bias-adjustment technique:

Step 1: Generate the full sample of size n, $\mathbf{y}$, from the relevant stationary ARFIMA(p_0, d_0, q_0) model.

Step 2: Compute the LPR estimator of d_0, $\widehat{d}_n$ using (5).

Step 3: Draw the sub-samples, $\mathbf{y}_i$ $(i = 1, 2, ..., m)$, from the full sample based on the relevant sub-sampling technique (non-overlapping or moving-block) and compute the LPR estimator of d_0, $\widehat{d}_i$, for each sub-sample.

Step 4: Depending on the sub-sample selection method chosen in Step 3, obtain the optimal weights for the corresponding method based on the parameters of the (true) ARFIMA(p_0, d_0, q_0) model and compute the optimal jackknife estimator, $\widehat{d}^{Opt}_{J,m}$. In the empirically realistic case in which the true model parameters are unknown, a feasible version of the jackknife estimator is implemented, with all details provided in the relevant sections below.

Step 5: Repeat Steps $1-4$ $100,000$ times and compute estimates of the bias and RMSE of the optimal jackknife estimator.

In Steps 2 and 3, the number of frequencies used to calculate the relevant LPR estimator is set to $N_L = \lfloor L^\alpha \rfloor$, with $\alpha = 0.65$, where L is as defined in (28).[6] The optimal jackknife estimators

[5] Additional results based on the assumption that ε_t is distributed as Student t with 5 degrees of freedom are available from the authors on request. This additional design feature did not lead to qualitatively different results.

[6] Certain simulation results based on $\alpha = 0.5$ have also been produced, but are not presented here due to space considerations. These additional numerical results can be provided by the authors on request.

calculated using the non-overlapping (abbreviated to Opt-NO), and moving-block (abbreviated to Opt-MB) schemes, are denoted by $\widehat{d}_{J,m}^{Opt-NO}$ and $\widehat{d}_{J,m}^{Opt-MB}$, respectively.

The weighted-average estimator of Guggenberger and Sun (2006) is computed as described in Section 2.2, with the following additional details. For a given N_n, the set of bandwidths used to calculate the constituent estimators in (10) are $N_{n,i} = \lfloor q_i N_n \rfloor$, where $\mathbf{q}^\top = (q_1, q_2, \ldots, q_K) = (1, 1.05, \ldots, 2)$. We produce the GS estimator (based on $r = 1$) using two different choices of N_n: (i) $N_n = \lfloor n^\alpha \rfloor$, with $\alpha = 0.65$ (denoting this estimator by $\widehat{d}_1^{GS}$), and (ii) the optimal choice of N_n as suggested in Guggenberger and Sun (2006, page 876) (denoting this version by $\widehat{d}_1^{Opt-GS}$). Importantly, bandwidth choice (ii) depends on knowledge of the true values of the short memory parameters, whereas bandwidth choice (i) yields an estimator that is feasible empirically. The parameter δ, required for both versions of the GS estimator, is evaluated using the formula $\delta = \tau_r/(\tau_r^* \sum_{i=1}^K q_k^{2+2r})$, where $\tau_{r-1}^* = -(2\pi)^{2r} r \big/ [(2r)!\,(2r+1)^2]$ and the number τ_r is as defined in Andrews and Guggenberger (2003). Details regarding the construction of the pre-filtered sieve bootstrap estimator ($\widehat{d}^{PFSB}$) can be found in Poskitt *et al.* (2016). In implementing this method, we set the number of bootstrap samples to $B = 1000$.

The MLE, $\widehat{d}^{MLE}$, is obtained by concentrating the Gaussian log-likelihood associated with an assumed ARFIMA(p, d, q) model with respect to μ and σ^2, subtracting from that the resulting constant $(n \log n - n)/2$, and maximizing the profile log-likelihood function

$$L(\eta) = -\frac{n}{2} \log \left[(\mathbf{y} - \widehat{\mu}\mathbf{1})^\top \boldsymbol{\Sigma}_\eta^{-1} (\mathbf{y} - \widehat{\mu}\mathbf{1}) \right] - \frac{1}{2} \log |\boldsymbol{\Sigma}_\eta|, \tag{32}$$

where $\mathbf{1}$ is the vector of ones and $\widehat{\mu} = \mathbf{1}^\top \boldsymbol{\Sigma}_\eta^{-1} \mathbf{y} \big/ \left(\mathbf{1}^\top \boldsymbol{\Sigma}_\eta^{-1} \mathbf{1} \right)$. The parameter $\eta = (d, \phi^\top, \theta^\top)^\top$, with d the fractional differencing parameter, ϕ the p-dimensional vector of AR parameters, θ the q-dimensional vector of MA parameters, and $\sigma^2 \boldsymbol{\Sigma}_\eta := [\gamma_{i-j}(\eta)]$, $i, j = 1, \ldots, n$,with $\gamma_k(\eta)$ being the autocovariance at lag k. When the MLE is implemented under correct model specification the assumed model corresponds to an ARFIMA(p_0, d_0, q_0) with μ_0 set to 0; under misspecification, the ARFIMA(p, d, q) model differs from the true data generating process (DGP) in some way. The estimation procedure under misspecification is detailed in Section 5.3.

The PW estimator of d, $\widehat{d}^{PW}$, is obtained in two steps. In the first step, an autoregressive-moving average (ARMA) model of order (p, q) is fit to the data, and in the second step, d is estimated by minimizing the sum of squares of the resultant residuals. Again, the PW estimator is implemented under correct and incorrect specification of the short memory dynamics.

5.2 Finite sample bias and RMSE: Correct model specification

In this section, we document results under correct specification of the true DGP. Results are presented for the full set of values: $d_0 = \{-0.25, 0, 0.25, 0.45\}$ and $\phi_0/\theta_0 = \{-0.9, -0.4, 0.4, 0.9\}$. The relative performance of the jackknife method is, in turn, assessed under two scenarios: (i) when the values of all parameters in the true DGP are used in the construction of the optimal jackknife weights and, (ii) when all parameters in the true DGP are estimated, but the correct values for lag lengths p_0 and q_0

are still adopted (with the superscript '*Opt*' omitted in this case). An iterative method (described in Section 5.2.2) is used to produce this feasible version of the jackknife estimator. To save on space, results for both $\widehat{d}_{J,m}^{Opt-NO}$ and $\widehat{d}_{J,m}^{NO}$ are recorded for $m = 2, 3, 4, 6, 8$, whilst results for both $\widehat{d}_{J,m}^{Opt-MB}$ and $\widehat{d}_{J,m}^{MB}$ based on only $m = 2$ are documented. We do note that the patterns exhibited (in terms of both bias and RMSE) for $\widehat{d}_{J,m}^{Opt-MB}$ and $\widehat{d}_{J,m}^{MB}$, across m, are similar to those exhibited for $\widehat{d}_{J,m}^{Opt-NO}$ and $\widehat{d}_{J,m}^{NO}$ respectively.

In case (i) we compare the jackknife estimator with the GS estimator obtained with the optimal choice of N_n ($\widehat{d}_1^{Opt-GS}$) - which, of course, relies on the known values of the short memory parameters - and with the sub-optimal estimator, $\widehat{d}_1^{GS}$. In case (ii) results for only $\widehat{d}_1^{GS}$ are included for comparison, as $\widehat{d}_1^{Opt-GS}$ is infeasible when the true values of the short memory parameters are unknown.[7] Note that the finite sample results for the (raw) LPR and PFSB estimators remain the same in both scenarios, (i) and (ii), as the construction of neither estimator relies on knowledge of the true parameters, nor of p_0 and q_0. As concerns the parametric estimators, $\widehat{d}^{MLE}$ and $\widehat{d}^{PW}$, in this correct specification scenario, $p = p_0$, $q = q_0$ and $\mu = \mu_0 = 0$, and d is estimated under the two cases: (i) where the short memory parameters are set at their true values; and (ii) where the short memory parameters are estimated simultaneously with d.

All relevant finite sample results for case (i) and case (ii) are presented and discussed in Section 5.2.1 and Section 5.2.2 respectively.

5.2.1 Case 1: True parameters are known

Tables 2 and 3 record the bias and RMSE of the various optimal jackknife estimators, the two different GS estimators, and the LPR, PFSB, MLE and PW estimators, for the case where the DGP is ARFIMA$(1, d_0, 0)$ and the short memory parameter ϕ_0 is known. The corresponding results for the ARFIMA$(0, d_0, 1)$ DGP are presented in Tables 4 and 5. The lowest biases and RMSEs for each design are marked in boldface. The second lowest values are italicized. Only that number which is smallest at the precision of 8 decimal places is bolded. Values highlighted with a '*' are equally small to 4 decimal places.

- Table 2 here -

- Table 3 here -

- Table 4 here -

- Table 5 here -

With reference to Tables 2 and 3: as would be anticipated in this situation, in which the true model is estimated and the true value of ϕ_0 is imposed, the MLE is the least biased estimator of all

[7]Note that in the case where the short memory parameters are unknown Guggenberger and Sun (2006) suggest that an adaptive procedure for the local Whittle-based estimator that they propose could be extended to the weighted-average estimator based on LPR. Since the adaptive method is not provided in detail in their paper, we do not pursue this option here.

methods considered, and has the smallest RMSE. The parametric PW method has the second least bias in a small number of cases, and also performs relatively well in terms of RMSE.

As is also consistent with expectations, and existing results (see, for example, Agiakloglou *et al.*, 1993, Nielsen and Frederiksen, 2005 and Poskitt *et al.*, 2016), when short memory dynamics are present, the raw, unadjusted, LPR estimator is biased, as the low frequencies are contaminated by the spectral density of the short run dynamics, particularly for *negative* values of ϕ_0 (which corresponds to positive first-order autocorrelation). As is evident from the recorded results, the bias is particularly large when there is a large negative value for ϕ_0 in (31), and it decreases as this value increases. Further, both bias and RMSE decline as the sample size increases, illustrating the consistency of the estimator.

We shall now comment on the performance of all nine bias-corrected semi-parametric estimators under the ARFIMA$(1, d_0, 0)$ process. With reference to Table 2, for the great majority of designs, $\widehat{d}_{J,m}^{Opt-NO}$ with $m = 2$, has the smallest bias of all nine such estimators. For $\phi_0 = -0.9$ and $n = 96$, the bias reduction of $\widehat{d}_{J,m}^{Opt-NO}$ $(m = 2)$, relative to the raw LPR estimator is up to 3.6%, and when $n = 576$, this rises to 5.7%.[8] For the larger values of ϕ_0, when $n = 96$, the bias reduction ranges from 48% to 60%, and from 56% to 97% when $n = 576$. Only occasionally is this particular version of the jackknife estimator inferior to an alternative semi-parametric estimator. Importantly, however, an increase in m leads to an increase in bias for $\widehat{d}_{J,m}^{Opt-NO}$ and, hence, a reduction in its superiority over all alternatives, including the raw LPR method in some cases. The reason is that the increase in m leads to a smaller sub-sample length and, hence, increases the finite sample impact of the dominant bias term on the sub-sample estimators used in the construction of the jackknife estimator.

Now referencing the results in Table 3, we see that, despite the lack of variance inflation in the asymptotic distribution of the optimal jackknife estimator, the reduction in bias does cause some finite sample increase in variance, leading to RMSEs for $\widehat{d}_{J,m}^{Opt-NO}$ that are occasionally slightly larger than the RMSE of the raw LPR estimator. That said, in the vast majority of cases $\widehat{d}_{J,m}^{Opt-NO}$ with $m = 8$, has the smallest RMSE of all semi-parametric estimators (including the raw LPR) and, in many cases, the RMSE of the jackknife estimator with the smallest bias ($\widehat{d}_{J,m}^{Opt-NO}$, $m = 2$) has a RMSE which remains less than that of the raw estimator. In addition, all versions of the jackknife estimator (including the moving-block version) tend to have smaller RMSEs than the three alternative bias-corrected methods ($\widehat{d}_1^{GS}$, $\widehat{d}_1^{Opt-GS}$ and $\widehat{d}_1^{PFSB}$), most notably for the smaller sample size ($n = 96$). As befits the optimality of the estimator, in almost all cases, $\widehat{d}_1^{Opt-GS}$ out-performs $\widehat{d}_1^{GS}$, in terms of both bias and RMSE, although both estimators, as already noted, are virtually always out-performed by a version of the jackknife procedure.

The broad conclusions drawn above obtain under the ARFIMA$(0, d_0, 1)$ DGP, as seen from the results recorded in Tables 4 and 5. The only notable difference is the improved RMSE performance of $\widehat{d}_{J,8}^{Opt-NO}$, with this estimator ranked second overall (after $\widehat{d}^{MLE}$) in terms of this measure.

[8] We remind the reader that when $\phi_0 = -0.9$ *all* estimators remain very biased.

5.2.2 Case 2: True parameters are unknown

Evaluation of the optimal weights in (23), required for the construction of the optimal jackknife estimator, depends on the covariances between both the different sub-sample LPR estimators and between the full-sample and sub-sample estimators, as given in (B.1) and (B.2). These covariances depend, in turn, on covariances between the various log-periodograms and, hence, on the values of the parameters that underpin the true DGP, as is made explicit in (30) and Appendix B. Hence, implementation of the optimal bias-correction procedure via the jackknife is not feasible in practice, without modification. To this end, we propose the following iterative method for obtaining a feasible version of the jackknife-based estimator; one still appropriate, however, for the case where the specified model (i.e. the values of p_0 and q_0) is correct.

An iterative version of the optimal jackknife estimator

1. **Prerequisite:** Estimate the relevant short memory parameter(s) in the ARFIMA(p_0, d_0, q_0) model, using pre-filtered data based on $d^f = \widehat{d}^{GS}$.
2. **Initialization:** Set $k = 1$ and tolerance level $\tau = \tau^{(0)}$.
3. **Recursive step:** For the k^{th} recursion, perform the jackknife bias-correction procedure of Section 3.2, but with the estimates of the short memory parameters from step 1, and $d^f = \widehat{d}^{GS}$, now inserted into the formulae for the covariance terms in (B.1) and (B.2). Denote the resulting estimator by $\widehat{d}^{(k)}_{J,m}$.
4. **Stopping rule:** If $\left|\widehat{d}^{(k+1)}_{J,m} - \widehat{d}^{(k)}_{J,m}\right| > \tau$ set $k = k + 1$ and $\tau = \tau^{(k)}$, and repeat steps 1 and 3 after updating $d^f = \widehat{d}^{(k)}_{J,m}$.

The basic idea behind the algorithm is as follows: estimation of the short memory parameter requires pre-filtering via some preliminary estimate of d_0. An obvious initial (consistent) choice is $d^f = \widehat{d}^{GS}$, as this estimator is already bias-adjusted, and a feasible estimator in the presence of unknown values for the short memory parameter(s). However $\widehat{d}^{GS}$ will still exhibit some bias in finite samples. Hence, iteration of the above algorithm, which involves replacing the initial pre-filtering value with successively less biased values, $d^f = \widehat{d}^{(k)}_{J,m}$, is expected to yield a final feasible version of the jackknife estimator, $\widehat{d}^{(k+1)}_{J,m}$, based on accurate estimates of all unknown parameters. (See also Poskitt *et al.*, 2016 for a related application of this form of iterative procedure). The feasible version of the jackknife statistic at the final iteration is denoted hereafter by $\widehat{d}^{NO}_{J,m}$ if the sub-sampling method is non-overlapping and $\widehat{d}^{MB}_{J,m}$ if the sub-sampling method is moving-block.

- Table 6 here -

- Table 7 here -

- Table 8 here -

- Table 9 here -

Tables 6 and 7 record the bias and RMSE of all versions of the feasible jackknife estimator, the feasible GS estimator, $\widehat{d}_1^{GS}$, and the LPR, PFSB, MLE and PW estimators, for the case where the DGP is ARFIMA$(1, d_0, 0)$ and the short memory parameter ϕ_0 is now estimated. The corresponding results for the ARFIMA$(0, d_0, 1)$ DGP are presented in Tables 8 and 9. The results for the $\widehat{d}_1^{GS}$, LPR and PFSB are the same as in the earlier corresponding tables, as these estimators do not depend on knowledge or estimation of the short memory dynamics. The parametric estimators, MLE and PW, do of course change when ϕ_0 is estimated. Once again, the minimum bias and RMSE are shown in bold font, and the second lowest values are italicized.

Consider the results for the ARFIMA$(1, d_0, 0)$ process (Tables 6 and 7). The (various versions of the) feasible jackknife estimators show similar characteristics to the corresponding optimal estimators, except for exhibiting larger bias and RMSE. This is to be expected given that the optimal weights are now functions of estimates of both d_0 and ϕ_0. The increase in bias (relative to the known parameter case) is particularly marked when $\phi_0 = -0.9$, with the feasible jackknife estimators seen to be more biased overall than the raw LPR estimator itself, in three cases. However, for all other values of ϕ_0, the least biased versions of the feasible jackknife estimators are still almost always less biased than the LPR estimator. For example, when $\phi_0 = -0.4$ and $n = 96$, the bias reduction of $\widehat{d}_{J,m}^{NO}$ with $m = 2$ compared to the raw LPR estimator is up to 35% and when $n = 576$, the bias reduction rises to 69%. Overall, the $\widehat{d}_{J,2}^{NO}$, $\widehat{d}_1^{GS}$, $\widehat{d}_1^{PFSB}$ and $\widehat{d}^{MLE}$ estimators share the title of the least, or second-least biased estimator. The RMSE results in Table 7 indicate the consistency of the feasible jackknife estimators. However, the MLE estimator still exhibits the least RMSE of all estimators considered, even when ϕ_0 is estimated, with the LPR estimator taking second place.

The results in Tables 8 and 9, for the ARFIMA$(0, d_0, 1)$ process, tell a broadly similar story to those for the ARFIMA$(1, d_0, 0)$ case, except for the fact that $\widehat{d}_{J,2}^{NO}$ is now the least biased estimator in more cases than any other competing estimator, and $\widehat{d}_{J,8}^{NO}$ is sometimes ranked second in terms of RMSE.

5.3 Finite sample bias and RMSE: Model misspecification

Misspecification occurs when the true DGP is ARFIMA(p_0, d_0, q_0) and the fitted model is ARFIMA(p, d, q), where p and q are such that $\{p \neq p_0 \cup q \neq q_0\} \setminus \{p_0 \leq p \cap q_0 \leq q\}$. We consider three different forms of misspecification: (*i*) true DGP: ARFIMA$(1, d_0, 0)$; fitted model: ARFIMA$(0, d, 0)$; (*ii*) true DGP: ARFIMA$(0, d_0, 1)$; fitted model: ARFIMA$(0, d, 0)$; and (*iii*) true DGP: ARFIMA$(1, d_0, 1)$; fitted model: ARFIMA$(2, d, 0)$. The first two forms of misspecification mimic a situation in which short memory dynamics are present, but are ignored. In particular, these scenarios allow us to assess the relative performance of the feasible jackknife estimator when no aspect of the short memory specification is used in the calculation of the weights. The third form of misspecification allows for another type of error in the specification of the short memory component. In all cases, we restrict both the DGP and the fitted model to be within the stationary region. In order to reduce the number of results to be tabulated and discussed, in Tables 10 to 13 we present results for the reduced set of values: $d_0 =$

$\{-0.25, 0.25, 0.45\}$ and $\phi_0/\theta_0 = \{-0.9, -0.4, 0.4, 0.9\}$. In Tables 14 and 15, we reduce the settings further by omitting results for $\phi_0/\theta_0 = 0.9$.

Under misspecification, the feasible jackknife estimates are obtained by using the fitted ARFIMA(p, d, q) model in the Prerequisite step in Section 5.2.2, while the remaining steps are unchanged. The MLE and PW estimators are produced as explained in Section 5.1.[9] The estimators $\widehat{d}_n$, $\widehat{d}^{GS}$ and $\widehat{d}^{PFSB}$ are not affected by misspecification of the short memory dynamics, as specification of that component of the model plays no role in their construction. Hence, the results for these estimators in Tables 10 to 13 match the corresponding results in Tables 2 to 5. The results for all estimators in Table 14, under the ARFIMA$(1, d_0, 1)$ DGP, are distinct from results in all other tables.

Tables 10 and 11 display the bias and RMSE results of all estimators under misspecification type (i). The corresponding results for the misspecification types (ii) and (iii) are presented in Tables 12 to 15. As previously, the minimum bias and RMSE are shown in bold font, and the second lowest values are italicized.

- Table 10 here -

- Table 11 here -

- Table 12 here -

- Table 13 here -

- Table 14 here -

- Table 15 here -

From Tables 10 and 11, we observe that under misspecification the (various versions of the) feasible jackknife estimators show similar characteristics to those observed under correct model specification, although with larger bias and RMSE. This is not surprising, given that the weights are now functions of estimates of d only, with information on the true or estimated autoregressive coefficient in the DGP ignored. The increase in bias (relative to the correct specification case) is particularly marked when $\phi_0 = -0.9$. When $\phi_0 > -0.9$, the feasible jackknife estimators still tend to show reduced bias compared to the LPR estimator, almost uniformly for $m = 2$. For example, when $\phi_0 = -0.4$ and $n = 96$, the bias reduction of $\widehat{d}^{NO}_{J,2}$ compared to the raw LPR estimator is up to 12%, and when $n = 576$, the bias reduction rises to 39%. Moreover when $\phi_0 = 0.9$, $\widehat{d}^{NO}_{J,2}$ is either the least, or second-least biased estimator of all estimators considered.

Under this form of misspecification, at least for $\phi_0 > -0.9$, the MLE and PW estimators are more biased than the feasible jackknife estimators, and much more so in some cases. This is expected, as model misspecification has a *direct* impact on the parametric estimators. In contrast, for the jackknife estimators, misspecification impinges more *indirectly,* only via the choice of weights. Overall, the PFSB estimator shows the least bias and the feasible GS estimator shows the second-least bias.

[9] For more details on MLE under mis-specification of the short-memory dynamics see, Martin *et al.* (2020)

The RMSE results in Table 11 demonstrate that neither the feasible jackknife estimators, nor the parametric methods, out-perform the raw LPR estimator, which shows the least RMSE; the second lowest RMSE usually being observed with one of $\widehat{d}_{J,8}^{NO}$, $\widehat{d}_1^{GS}$, $\widehat{d}^{PFSB}$ or $\widehat{d}^{MLE}$.

The bias results for misspecification form (ii) in Table 12 display similar characteristics to those described for misspecification form (i), apart from the much more distinct dominance of $\widehat{d}_{J,2}^{NO}$ in this case. The RMSE results in Table 13 reveal that, once again, $\widehat{d}_n$ has the smallest RMSE values, with $\widehat{d}^{MLE}$ taking the second place.

Under the third form of misspecification, the bias estimates indicate that $\widehat{d}_1^{GS}$ exhibits the least bias, with $\widehat{d}_{J,2}^{NO}$ taking the second place (refer Table 14). In terms of the RMSE results in Table 15, once again the raw LPR has the lowest values most frequently, followed by $\widehat{d}_1^{GS}$, with the *second* smallest RMSE values mostly observed for $\widehat{d}_{J,8}^{NO}$.

5.4 A summary of the simulation results

- Table 16 here -

To assist the reader, in Table 16 we summarize all of the simulation results tabulated in Tables 2 to 15, by ranking the estimators - from first to third - under the different scenarios. Panel A in Table 16 summarizes the results in Tables 2 to 5; Panel B summarizes the results in Tables 6 to 9; and Panel C summarizes the results in Tables 10 to 15, with the three misspecification types – (i) to (iii) – corresponding to those described in the above section. An estimator is ranked first if it has the smallest value (in bold font) for the relevant measure (bias or RMSE) the largest number of times in a given table. The other ranks follow accordingly. If needed to complete the ranking, the method with the largest number of second-smallest values (in italic font) in a given Table is referenced. And so on. The rankings accord with the narrative in the preceding sections.

6 Discussion

With the fractionally integrated autoregressive moving-average model being one of the key model classes for describing long memory processes, much effort has been expended on producing accurate estimates of the fractional differencing parameter, d, in particular. This quest has been hampered by certain problems, for both parametric and semi-parametric approaches. Specifically, the need to fully specify the model for parametric estimation means that any incorrect specification of the short memory dynamics has serious consequences, in terms of both finite sample and asymptotic properties (see, for example, Chen and Deo, 2006 and Martin *et al.*, 2020). On the other hand, the semi-parametric estimators, whilst not requiring explicit modelling of the short memory component, can suffer substantial finite sample bias in the presence of unaccounted for short memory dynamics. It is bias-correction of this latter class of estimator that has been the focus of this paper.

A natural way of producing a bias-corrected version of the commonly used the log-periodogram regression (LPR) estimator is suggested in this article, based on the jackknife technique. Optimality is achieved by allocating weights within the jackknife that are adjusted for the bias to a particular order, and that minimize the increase in variance caused by the reduction in bias. The construction of the optimally bias-corrected estimator requires expressions for the dominant bias term and variance of the unadjusted LPR estimator. We show that the statistical properties of the LPR estimator, as originally established by Hurvich *et al.* (1998), are valid for a more general class of fractional process that is not necessarily Gaussian. Hence, the jackknife estimator that we construct from the optimally weighted average of LPR estimators also has proven optimality under this general form of process. In addition to proving the consistency of the optimal jackknife estimator, we have the important result that the asymptotic variance of the estimator is equivalent to that of the unadjusted LPR estimator. That is, bias adjustment is effected without any associated increase in asymptotic variance.

Our Monte Carlo study shows that, amongst the semi-parametric estimators, the optimal jackknife estimator based on a small number of non-overlapping sub-samples outperforms (in terms of bias reduction) both the pre-filtered sieve bootstrap estimator of Poskitt *et al.* (2016) and the weighted-average estimator of Guggenberger and Sun (2006), albeit in the somewhat artificial case in which the parameters of the DGP are correctly identified and known, for the purpose of computing optimal weights. In the realistic case in which these parameters are not known, we suggest an iterative procedure in which the weights are constructed using consistent estimates. In this case the method is not dominant overall, compared to alternative bias-corrected methods, but is still the least biased in some cases. The relationship between the semi-parametric methods and the two parametric methods is much as anticipated. In particular, the semi-parametric methods dominate in terms of both bias and RMSE when the short memory dynamics are misspecified. Once again, a version of the feasible jackknife method is ranked highly under certain misspecified settings, despite the fact that the misspecification impacts on the construction of the jackknife weights.

Throughout the paper we assume that the number of sub-samples is fixed. One may wish to allow the number of sub-samples to vary and explore the characteristics of the resultant bias-adjusted estimators in this case. Importantly, alternative methods of estimating the weights are to be investigated, including the possible use of a non-parametric estimate of the spectral density (see, Moulines and Soulier, 1999), rather than replacing the true values with their consistent estimates, or the use of an adaptive method in the spirit of that suggested by Guggenberger and Sun (2006).

Finally, although we focus on the LPR estimator, the jackknife procedure can easily be applied to other estimators such as the local Whittle estimator of Künsch (1987), the local polynomial Whittle estimator of Andrews and Sun (2004) or even to the (already analytically) bias-reduced estimators of Andrews and Guggenberger (2003) and Guggenberger and Sun (2006). Another possible extension is to relax the assumption of stationarity of the process using the results Velasco (1999), and to derive the properties the optimal jackknife estimators in the nonstationary setting.

References

Adenstedt, R. K. (1974). On large-sample estimation for the mean of a stationary random sequence. *The Annals of Statistics*, **2**, 1095–1107.

Agiakloglou, C., Newbold, P., and Wohar, M. (1993). Bias in an estimator of the fractional difference parameter. *Journal of Time Series Analysis*, **14**, 235–246.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, **71**, 579–625.

Andrews, D. W. K. and Guggenberger, P. (2003). A bias-reduced log-periodogram regression estimator for the long-memory parameter. *Econometrica*, **71**, 675–712.

Andrews, D. W. K. and Sun, Y. (2004). Adaptive local polynomial Whittle estimation of long-range dependence. *Econometrica*, **72**, 569–614.

Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics*, **73**, 5–59.

Beran, J. (1994). *Statistics for long-memory processes*. Chapman and Hall, New York, 1st edition.

Bollerslev, T. and Mikkelsen, H. O. (1996). Modeling and pricing long memory in stock market volatility. *Journal of Econometrics*, **73**, 151–184.

Brillinger, D. (1981). *Time series: Data Analysis and Theory*. Holden Day, San Francisco, 2nd edition.

Chambers, M. J. (2013). Jackknife estimation of stationary autoregressive models. *Journal of Econometrics*, **172**, 142–157.

Chen, W. W. and Deo, R. S. (2006). Estimation of mis-specified long memory models. *Journal of Econometrics*, **134**, 257–281.

Chen, Y. and Yu, J. (2015). Optimal jackknife for unit root models. *Statistics and Probability Letters*, **99**, 135–142.

Cheung, Y. W. (2016). Long memory in foreign-exchange rates. *Journal of Business & Economic Statistics*, **11**, 93–101.

Chiang, A. C. and Wainwright, K. (2005). *Fundamental methods of mathematical economics*. McGraw-Hill, New York.

Contreras-reyes, J. E. and Palma, W. (2013). Statistical analysis of autoregressive fractionally integrated moving average models in R. *Computational Statistics*, **28**, 2309–2331.

Delgado, M. A. and Robinson, P. M. (1994). New methods for the analysis of long memory time series: Application to Spanish inflation. *Journal of Forecasting*, **13**, 97–107.

Diebold, F. X. and Rudebusch, G. D. (1989). Long memory and persistence in aggregate output. *Journal of Monetary Economics*, **24**, 189–209.

Diebold, F. X., Husted, S., and Rush, M. (1991). Real exchange rates under the gold standard. *Journal of Political Economy*, **99**, 1252–1271.

Doukhan, P., Oppenheim, G., and Taqqu, S. M. (2003). *Theory and applications of long-range dependence*. Springer Science & Business Media.

Ekonomi, L. and Butka, A. (2011). Jackknife and bootstrap with cycling blocks for the estimation of fractional parameter in ARFIMA model. *Turkish Journal of Mathematics*, **35**, 151–158.

Fox, R. and Taqqu, S. M. (1986). Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *The Annals of Statistics*, **14**, 517–532.

Geweke, J. and Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis*, **4**, 221–238.

Gil-Alana, L. A. (2012). U.K. rainfall data: A long-term persistence approach. *Journal of Applied Meteorology and Climatology*, **51**, 1904–1913.

Gil-Alana, L. A. and Robinson, P. M. (1997). Testing of unit root and other nonstationary hypotheses in macroeconomic time series. *Journal of Econometrics*, **80**, 241–268.

Granger, C. W. J. and Hyung, N. (2004). Occasional structural breaks and long memory with an application to the S&P500 absolute stock returns. *Journal of Empirical Finance*, **11**, 399–421.

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

Guggenberger, P. and Sun, Y. (2006). Bias-reduced log-periodogram and Whittle estimation of the long-memory Parameter without variance inflation. *Econometric Theory*, **22**, 863–912.

Hahn, J. and Newey, W. K. (2013). Jackknife and analytical bias reduction for nonlinear panel models. *Econometrica*, **72**, 1295–1319.

Hassler, U. and Wolters, J. (1995). Long memory in inflation rates: International evidence. *Journal of Business & Economic Statistics*, **13**, 37–45.

Hosking, J. R. M. (1981). Fractional differencing. *Biometrika*, **68**, 165–175.

Hurst, H. E. (1951). Methods of using long-term storage in reservoirs. *Transactions of American Civil Engineers*, **33**, 619–626.

Hurvich, C. M. and Brodsky, J. (2001). Broadband semiparametric estimation of the memory parameter of a long-memory time series using fractional exponential models. *Journal of Time Series Analysis*, **22**, 221–249.

Hurvich, C. M., Deo, R. S., and Brodsky, J. (1998). The mean squared error of Geweke and Porter-Hudak's estimator of the memory parameter of a long-memory time series. *Journal of Time Series Analysis*, **19**, 19–46.

Krishnaiah, P. R., Haigs, P. J., and Steinberg, L. (1963). A note on the bivariate chi distribution. *SIAM Review*, **5**, 140–144.

Künsch, H. R. (1987). Statistical aspects of self-similar processes. In *Proceedings of the First World Congress of the Bernoulli Society*, pages 67–74, Utrecht. VNU Science Press.

Lahiri, S. N. (2003). A necessary and sufficient condition for asymptotic independence of discrete Fourier transforms under short- and long-range dependence. *Annals of Statistics*, **31**, 613–641.

Lieberman, O. (2001). The exact bias of the log-periodogram regression estimator. *Econometric Reviews*, **20**, 369–383.

Lieberman, O., Rosemarin, R., and Rousseau, J. (2012). Asymptotic theory for maximum likelihood estimation of the memory parameter in stationary Gaussian processes. *Econometric Theory*, **28**, 457–470.

Martin, G. M., Nadarajah, K., and Poskitt, D. S. (2020). Issues in the estimation of mis-specified models of fractionally integrated processes. *Journal of Econometrics*, **215**, 559–573.

Moulines, E. and Soulier, P. (1999). Broadband log-periodogram regression of time series with long-range dependence. *Annals of Statistics*, **27**, 1415–1439.

Nielsen, M. O. and Frederiksen, P. H. (2005). Finite sample comparison of parametric, semiparametric, and wavelet estimators of fractional integration. *Econometric Reviews*, **24**, 405–443.

Poskitt, D. S., Martin, G. M., and Grose, S. D. (2016). Bias correction of semiparametric long memory parameter estimators via the prefiltered sieve bootstrap. *Econometric Theory*, **1**, 1–32.

Reisen, V. A. and Lopes, S. (1999). Some simulations and applications of forecasting long-memory time-series models. *Journal of Statistical Planning and Inference*, **80**, 269–287.

Robinson, P. M. (1995a). Gaussian semiparametric estimation of the long range dependence. *The Annals of Statistics*, **23**, 1630–1661.

Robinson, P. M. (1995b). Log-periodogram regression of time series with long range dependence. *Annals of Statistics*, **23**, 1048–1072.

Robinson, P. M. (2004). *Time series with long memory.* Oxford University Press, New York.

Robinson, P. M. and Henry, M. (2003). Higher-order kernel semiparametric M-estimation of long memory. *Journal of Econometrics*, **114**, 1–27.

Robinson, Y. K. and Kaufmann, H. (2015). Bias-corrected estimation in mildly explosive autoregressions. In *In Annual Conference 2015 (Muenster): Economic Development-Theory and Policy (No. 112897)*. Verein für Socialpolitik/German Economic Association.

Velasco, C. (1999). Non-stationary log-periodogram regression. *Journal of Econometrics*, **91**, 325–371.

Yajima, Y. (1989). A central limit theorem of Fourier transforms of strongly dependent stationary processes. *Journal of Time Series Analysis*, **10**, 375–383.

Appendix A: Proofs of Theorems and Lemmas

Proof of Theorem 1. Under Assumptions $(A.1)-(A.4)$, the proof of the theorem follows immediately after applying the results of Corollary A.1 of Martin *et al.* (2020) to Lemmas, 2, 5, 6 and 7 of Hurvich *et al.* (1998). Hence we omit the proof. ■

Prior to providing the proofs of the other theorems and lemmas, we will introduce the following definition, and its properties, to be used hereinafter.

Define $\Delta^{(T)}(\lambda) = \sum_{t=1}^{T} \exp(-\imath\lambda t)$. Then,

$$\Delta^{(T)}(\lambda) = \exp\left(-\imath\frac{\lambda}{2}(T+1)\right)\frac{\sin\left(\frac{\lambda T}{2}\right)}{\sin\left(\frac{\lambda}{2}\right)} = \begin{cases} 0 & \textit{if} \quad \lambda \not\equiv 0(mod\ \pi) \\ T & \textit{if} \quad \lambda \equiv 0(mod\ 2\pi) \\ 0 \quad \text{or} \quad T & \textit{if} \quad \lambda = \pm\pi, \pm 3\pi, \ldots \end{cases} \tag{A.1}$$

where, $a \equiv b\,(mod\alpha)$ means that the difference $(a-b)$ is an integral multiple of α for $\alpha, x, y \in \mathbb{R}$.

Consider

$$\begin{aligned}\sum_{t=-T}^{T} \exp\{-\imath\lambda t\} &= 1 + \sum_{t=1}^{T}\exp\{-\imath\lambda t\} + \sum_{t=1}^{T}\exp\{-\imath(-\lambda)t\} \\ &= 1 + 2\Delta^{(T)}(\lambda), \text{ using (A.1)}.\end{aligned}$$

This immediately gives that

$$\lim_{T\to\infty}\frac{1}{2\pi}\sum_{t=-T}^{T}\exp\{-\imath\lambda t\} = \eta(\lambda). \tag{A.2}$$

We will derive the following two properties of $\Delta^{(T)}(\lambda)$.

1. Sum:

$$\begin{aligned}\lim_{T\to\infty}\left[\Delta^{(T)}(\lambda) + \Delta^{(T)}(-\lambda)\right] &= \lim_{T\to\infty}\left(\sum_{t=-T}^{T}\exp\{\imath\lambda t\} - 1\right) \\ &= 2\pi\eta(\lambda) - 1, \text{ by (A.2)}.\end{aligned} \tag{A.3}$$

2. Product:

$$T^{-2}\Delta^{(T)}(-\lambda)\,\Delta^{(T)}(\lambda) = T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T}\exp\{-\imath\lambda(t-s)\}$$
$$= T^{-2}\sum_{t=-(T-1)}^{T-1}(T-|t|)\exp\{-\imath\lambda t\}$$
$$= T^{-1}\sum_{t=-(T-1)}^{T-1}\exp\{-\imath\lambda t\} - \sum_{t=-(T-1)}^{T-1}\frac{|t|}{T^2}\exp\{-\imath\lambda t\}. \tag{A.4}$$

Consider the second term in the above expression,

$$\Big|\sum_{t=-(T-1)}^{T-1}\frac{|t|}{T^2}\exp\{-\imath\lambda t\}\Big| \le \Big|\sum_{t=-(T-1)}^{T-1}\frac{|t|}{T^2}\Big| \to 0 \text{ as } T\to\infty.$$

Hence the expression in (A.4) is given by,

$$T^{-2}\Delta^{(T)}(-\lambda)\,\Delta^{(T)}(\lambda) = T^{-1}2\pi\eta(\lambda) + o(1). \tag{A.5}$$

Lemma 1 *Let $\mathbf{W}_t$ be a stationary h vector-valued time series with n observations satisfying the spectral density given in (1). Suppose that Assumptions (A.1) − (A.3) hold. The k^{th}-order cumulant of the multivariate series, $\kappa\left\{D^{(n)}_{W_{a_1}}(\lambda_1), ..., D^{(n)}_{W_{a_k}}(\lambda_k)\right\}$ is*

$$n^{-\frac{k}{2}}(2\pi)^{\frac{k}{2}-1}\Delta^{(n)}\left(\sum_{j=1}^{k}\lambda_j\right) f_{W_{a_1}...W_{a_k}}(\lambda_1,...,\lambda_{k-1}) + o\big(n^{1-2d-\frac{k}{2}}\big). \tag{A.6}$$

where $f_{W_{a_1}...W_{a_k}}(\lambda_1,...,\lambda_{k-1})$ is the k^{th}-order spectrum of the series $\mathbf{W}_t$, with $a_1,\dots,a_k = 1,2,\dots,h$, and $k = 1,2,\dots$.

Proof. By Lemma P4.2 of Brillinger (1981), the cumulant, $\kappa\{D^{(n)}_{W_{a_1}}(\lambda_1), ..., D^{(n)}_{W_{a_k}}(\lambda_k)\}$ has the form

$$\sum_{t_1=-\infty}^{\infty}...\sum_{t_k=-\infty}^{\infty}\exp\left(-\imath\sum_{j=1}^{k}\lambda_j t_j\right)\kappa_{W_{a_1}...W_{a_k}}(t_1-t_k,...,t_{k-1}-t_k)$$

Substituting, $u_j = t_j - t$ where $t = t_k$, and $-S \le u_j \le S$, for $j = 1,\dots,k-1$ with $S = 2(n-1)$ we have that

$$\begin{aligned}
&\kappa\{D^{(n)}_{W_{a_1}}(\lambda_1), D^{(n)}_{W_{a_2}}(\lambda_2), ..., D^{(n)}_{W_{a_k}}(\lambda_k)\}\\
=\ & (2\pi n)^{-\frac{k}{2}}\sum_{t=-\infty}^{\infty}\sum_{u_1=-S}^{S}\cdots\sum_{u_k=-S}^{S}\exp\left(-\imath\sum_{j=1}^{k}\lambda_j(u_j+t)\right)\kappa_{W_{a_1}...W_{a_k}}(u_1,...,u_{k-1})\\
=\ & (2\pi n)^{-\frac{k}{2}}\sum_{u_1=-S}^{S}\cdots\sum_{u_k=-S}^{S}\exp\left(-\imath\sum_{j=1}^{k-1}\lambda_j u_j\right)\kappa_{W_{a_1}...W_{a_k}}(u_1,...,u_{k-1})\sum_{t=-\infty}^{\infty}\exp\left(-\imath\sum_{j=1}^{k}\lambda_j t\right)\\
=\ & (2\pi)^{-\frac{k}{2}+1}n^{-\frac{k}{2}}\Delta^{(n)}\left(\textstyle\sum_{j=1}^{k}\lambda_j\right)\sum_{u_1=-S}^{S}\cdots\sum_{u_k=-S}^{S}\exp\left(-\imath\sum_{j=1}^{k-1}\lambda_j u_j\right)\kappa_{W_{a_1}...W_{a_k}}(u_1,...,u_{k-1}).
\end{aligned}$$

The rapidity of the convergence of $\sum_{u_1=-S}^{S}\cdots\sum_{u_k=-S}^{S}\exp(-\imath\sum_{j=1}^{k-1}\lambda_j u_j)\kappa_{W_{a_1}\ldots W_{a_k}}(u_1,...,u_{k-1})$ to $f_{W_{a_1}\ldots W_{a_k}}(\lambda_1,\ldots,\lambda_{k-1})$ as $n\to\infty$ is measured as follows.

$$\begin{aligned}
&\left|\sum_{u_1=-S}^{S}\cdots\sum_{u_k=-S}^{S}\exp\left(-\imath\sum_{j=1}^{k-1}\lambda_j u_j\right)\kappa_{W_{a_1}\ldots W_{a_k}}(u_1,...,u_{k-1})-f_{W_{a_1}\ldots W_{a_k}}(\lambda_1,\ldots,\lambda_{k-1})\right| \\
=&\left|\sum_{|u_1|>S}\cdots\sum_{|u_k|>S}\exp\left(-\imath\sum_{j=1}^{k-1}\lambda_j u_j\right)\kappa_{W_{a_1}\ldots W_{a_k}}(u_1,...,u_{k-1})\right| \\
\leq&\sum_{|u_1|>S}\cdots\sum_{|u_k|>S}\left|\kappa_{W_{a_1}\ldots W_{a_k}}(u_1,...,u_{k-1})\right| \\
\leq& n^{-1+2d}\sum_{|u_1|>S}\cdots\sum_{|u_k|>S}\left(\left|\frac{u_1}{n}\right|^{1-2d}+\cdots+\left|\frac{u_{k-1}}{n}\right|^{1-2d}\right)\left|\kappa_{W_{a_1}\ldots W_{a_k}}(u_1,...,u_{k-1})\right|.
\end{aligned}$$

Hence the proof is completed since Assumption $(A.1)$ holds and $n^{-1+2d}(|u_1|+\cdots+|u_{k-1}|)\to 0$ as $n\to\infty$. ■

The above Lemma shows that when the DFTs correspond to multivariate time series with the same number of observations in their sample, the k^{th}-order cumulant of the multivariate series can be approximated with the expression given in (A.6). The only difference between this Lemma and Proposition 1 is that the proposition deals with different sample sizes for the time series in the multivariate set-up.

Proof of Proposition 1. The proof of the proposition can be established in a similar fashion to the above proof. Hence, we omit the proof here. ■

Proof of Theorem 2. The expectation of the DFT of the full sample or the sub-sample is

$$\begin{aligned}
E\left(D_{X_{a_i}}^{(L_i)}(\lambda)\right) &= \tfrac{1}{\sqrt{2\pi n}}\sum\nolimits_{t=1}^{n}\exp(-\imath\lambda t)\,E(y_t) \\
&= \tfrac{\mu_Y}{\sqrt{2\pi L_i}}\Delta^{(L_i)}(\lambda) \\
&= \begin{cases} 0 & if \quad \lambda\not\equiv 0\,(mod\ \pi) \\ \sqrt{\frac{L_i}{2\pi}}\mu_Y & if \quad \lambda\equiv\pi\,(mod\ 2\pi) \\ 0\ or\ \sqrt{\frac{L_i}{2\pi}}\mu_Y & if \quad \lambda=\pm\pi,\pm3\pi,\ldots \end{cases},
\end{aligned}$$

where $E(y_t)=\mu_Y$. Therefore, $D_{X_{a_i}}^{(L_i)}(\lambda)$ behaves in the manner required by the theorem as the first-order cumulant provides the mean of the random variable of interest.

The covariance between $D_{X_{a_i}}^{(L_i)}(\lambda)$ and $D_{X_{a_j}}^{(L_j)}(\mu)$ is measured by the second-order cumulant and Proposition 1 gives that

$$Cov\left(D_{X_{a_i}}^{(L_i)}(\lambda),D_{X_{a_j}}^{(L_j)}(\mu)\right)=\tfrac{1}{L}\Delta^{(L)}(\lambda+\mu)\,f_{X_{a_i},X_{a_j}}(\lambda)+o\left(L^{-2d}\right),$$

where $L=\min(L_i,L_j)$. Thus, the covariance between the DFTs of the full sample and the sub-sample tends to 0 as $n\to\infty$. ■

Proof of Theorem 3. The covariance between $I^{(L_i)}_{X_{a_i}}(\lambda)$ and $I^{(L_j)}_{X_{a_j}}(\mu)$ is given by,

$$\begin{aligned} Cov\left(I^{(L_i)}_{X_{a_i}}(\lambda), I^{(L_j)}_{X_{a_j}}(\mu)\right) &= E\left(I^{(L_i)}_{X_{a_i}}(\lambda) I^{(L_j)}_{X_{a_j}}(\mu)\right) - E\left(I^{(L_i)}_{X_{a_i}}(\lambda)\right) E\left(I^{(L_j)}_{X_{a_j}}(\mu)\right) \\ &= E\left(D^{(L_i)}_{X_{a_i}}(\lambda) D^{(L_i)}_{X_{a_i}}(-\lambda) D^{(L_j)}_{X_{a_j}}(\mu) D^{(L_j)}_{X_{a_j}}(-\mu)\right) \\ &\quad - E\left(D^{(L_i)}_{X_{a_i}}(\lambda) D^{(L_i)}_{X_{a_i}}(-\lambda)\right) E\left(D^{(L_j)}_{X_{a_j}}(\mu) D^{(L_j)}_{X_{a_j}}(-\mu)\right). \end{aligned}$$

Since the expectations can be expressed in terms of cumulants (see Appendix B for more details), we may express the covariance term as follows,

$$\begin{aligned} Cov\left(I^{(L_i)}_{X_{a_i}}(\lambda), I^{(L_j)}_{X_{a_j}}(\mu)\right) &= \kappa\left(D^{(L_i)}_{X_{a_i}}(\lambda), D^{(L_i)}_{X_{a_i}}(-\lambda), D^{(L_j)}_{X_{a_j}}(\mu), D^{(L_j)}_{X_{a_j}}(-\mu)\right) \\ &\quad + \kappa\left(D^{(L_i)}_{X_{a_i}}(-\lambda), D^{(L_j)}_{X_{a_j}}(\mu)\right) \kappa\left(D^{(L_i)}_{X_{a_i}}(\lambda), D^{(L_j)}_{X_{a_j}}(-\mu)\right) \\ &\quad + \kappa\left(D^{(L_i)}_{X_{a_i}}(\lambda), D^{(L_j)}_{X_{a_j}}(\mu)\right) \kappa\left(D^{(L_i)}_{X_{a_i}}(-\lambda), D^{(L_j)}_{X_{a_j}}(-\mu)\right). \end{aligned}$$

Then Proposition 1 gives us that,

$$\begin{aligned} Cov\left(I^{(L_i)}_{X_{a_i}}(\lambda), I^{(L_j)}_{X_{a_j}}(\mu)\right) &= L^{-2}(2\pi)\Delta^{(L)}(\lambda+\mu-\lambda-\mu) f_{X_{a_i}X_{a_i}X_{a_j}X_{a_j}}(\lambda,-\lambda,\mu) + o\left(L^{-1-2d}\right) \\ &\quad + \left(L^{-1}\Delta^{(L)}(-\lambda+\mu) f_{X_{a_i}X_{a_j}}(-\lambda) + o\left(L^{-2d}\right)\right) \\ &\quad \times \left(L^{-1}\Delta^{(L)}(\lambda-\mu) f_{X_{a_i}X_{a_j}}(\lambda) + o\left(L^{-2d}\right)\right) \\ &\quad + \left(L^{-1}\Delta^{(L)}(\lambda+\mu) f_{X_{a_i}X_{a_j}}(\lambda) + o\left(L^{-2d}\right)\right) \\ &\quad \times \left(L^{-1}\Delta^{(L)}(-\lambda-\mu) f_{X_{a_i}X_{a_j}}(-\lambda) + o\left(L^{-2d}\right)\right) \\ &= L^{-2}(2\pi)\Delta^{(L)}(0) f_{X_{a_i}X_{a_i}X_{a_j}X_{a_j}}(\lambda,-\lambda,\mu) + o\left(L^{-1-2d}\right) \\ &\quad + L^{-2}\Delta^{(L)}(-\lambda+\mu)\Delta^{(L)}(\lambda-\mu)\left(f_{X_{a_i}X_{a_j}}(\lambda)\right)^2 \\ &\quad + L^{-1}\left(\Delta^{(L)}(-\lambda+\mu) + \Delta^{(L)}(\lambda-\mu)\right) f_{X_{a_i}X_{a_j}}(\lambda)\, o\left(L^{-2d}\right) \\ &\quad + L^{-2}\Delta^{(L)}(\lambda+\mu)\Delta^{(L)}(-\lambda-\mu)\left(f_{X_{a_i}X_{a_j}}(\lambda)\right)^2 \\ &\quad + L^{-1}\Delta^{(L)}(\lambda+\mu) f_{X_{a_i}X_{a_j}}(-\lambda) + \Delta^{(L)}(-\lambda-\mu) f_{X_{a_i}X_{a_j}}(-\lambda)\, o\left(L^{-2d}\right) \\ &= L^{-1}(2\pi) f_{X_{a_i}X_{a_i}X_{a_j}X_{a_j}}(\lambda,-\lambda,\mu) + L^{-2}\Big[\Delta^{(L)}(-\lambda+\mu)\Delta^{(L)}(\lambda-\mu) \\ &\quad + \Delta^{(L)}(\lambda+\mu)\Delta^{(L)}(-\lambda-\mu)\Big]\left(f_{X_{a_i}X_{a_j}}(\lambda)\right)^2 + \Big[\Delta^{(L)}(-\lambda+\mu) \\ &\quad + \Delta^{(L)}(\lambda-\mu) + \Delta^{(L)}(\lambda+\mu) + \Delta^{(L)}(-\lambda-\mu)\Big] f_{X_{a_i}X_{a_j}}(\lambda)\, o\left(L^{-2d}\right) \\ &\quad + o\left(L^{-1-2d}\right) + o\left(L^{-4d}\right). \end{aligned} \tag{A.7}$$

Using the two properties in (A.3) and (A.5), the covariance in (A.7) is simplified further as follows,

$$\begin{aligned} Cov\left(I^{(L_i)}_{X_{a_i}}(\lambda), I^{(L_j)}_{X_{a_j}}(\mu)\right) &= \frac{2\pi}{L}\left[\eta(\lambda-\mu) + \eta(\lambda+\mu)\right]\left\{f_{X_{a_i}X_{a_j}}(\lambda)\right\}^2 + \frac{2\pi}{l^\dagger} f_{X_{a_i}X_{a_i}X_{a_j}X_{a_j}}(\lambda,-\lambda,\mu) \\ &\quad + 2\pi\left[\eta(\lambda-\mu) + \eta(\lambda+\mu)\right] f_{X_{a_i}X_{a_j}}(\lambda)\, o\left({l^\dagger}^{-2d}\right) + o\left(L^{-1-2d}\right). \end{aligned}$$

Now let us consider the asymptotic distribution of $I_{X_{a_i}}^{(L_i)}(\lambda)$. We may re-write the periodogram as follows,

$$I_{X_{a_i}}^{(L_i)}(\lambda) = \left[ReD_{X_{a_i}}^{(L_i)}(\lambda)\right]^2 + \left[ImD_{X_{a_i}}^{(L_i)}(\lambda)\right]^2,$$

where

$$ReD_{X_{a_i}}^{(L_i)}(\lambda) = \frac{1}{\sqrt{2\pi L_i}} \sum_{t=1}^{L_i} y_t \cos(\lambda t), \text{ and, } ImD_{X_{a_i}}^{(L_i)}(\lambda) = \frac{1}{\sqrt{2\pi L_i}} \sum_{t=1}^{L_i} y_t \sin(\lambda t).$$

Following Theorem 2.1 of Lahiri (2003), we have that

$$\begin{bmatrix} \dfrac{ReD_{X_{a_i}}^{(L_i)}(\lambda) - E\left(ReD_{X_{a_i}}^{(L_i)}(\lambda)\right)}{\sqrt{L_i f_{X_{a_i}X_{a_i}}(\lambda)}} \\ \dfrac{ImD_{X_{a_i}}^{(L_i)}(\lambda) - E\left(ImD_{X_{a_i}}^{(L_i)}(\lambda)\right)}{\sqrt{L_i f_{X_{a_i}X_{a_i}}(\lambda)}} \end{bmatrix} \to^D N(\mathbf{0}, \mathbf{I}_2).$$

Hence the result. ■

Proof of Theorem 4. Recall that $x_j = \ln(2\sin(\lambda_j/2))$, $a_j = x_j - \overline{x}$ and $S_{xx} = \sum_{j=1}^{N_n} \left(X_j - \overline{X}\right)^2$. From Hurvich *et al.* (1998) we have that $S_{xx} = N_n(1 + o(1))$ and $a_j = \log j - \log N_n + 1 + o(1) + o\left(\frac{N_n^2}{n^2}\right)$, $j = 1, \ldots, N_n$. Thus,

$$\sup_j |a_j| = 1 + o(1) + O\left(\frac{N_n^2}{n^2}\right).$$

Using Appendix B we have that

$$\begin{aligned} Cov\left(\log I_{X_{a_i}}^{(L_i)}(\lambda_j), \log I_{X_{a_j}}^{(L_j)}(\mu_k)\right) &= (1-\rho^2)^{\frac{1}{2}} \sum_{k=1}^{\infty} \left(\Psi\left(\frac{1}{2}+k\right) + \Psi\left(\frac{1}{2}\right)\right)^2 \frac{\Gamma\left(\frac{1}{2}+k\right)}{\Gamma\left(\frac{1}{2}\right)} \frac{(\rho^2)^k}{k!} \\ &\quad - (1-\rho^2)\left(\sum_{k=1}^{\infty} \left(\Psi\left(\frac{1}{2}+k\right) + \Psi\left(\frac{1}{2}\right)\right) \frac{\Gamma\left(\frac{1}{2}+k\right)}{\Gamma\left(\frac{1}{2}\right)} \frac{(\rho^2)^k}{k!}\right)^2 \\ &\le (1-\rho^2)^{\frac{1}{2}} \sum_{k=1}^{\infty} \left(\Psi\left(\frac{1}{2}+k\right) + \Psi\left(\frac{1}{2}\right)\right)^2 \frac{\Gamma\left(\frac{1}{2}+k\right)}{\Gamma\left(\frac{1}{2}\right)} \frac{(\rho^2)^k}{k!}, \end{aligned}$$

where $\rho = Corr\left(I_{X_{a_i}}^{(L_i)}(\lambda_j), I_{X_{a_j}}^{(L_j)}(\mu_k)\right) = o\left(n^{-1}\right)$ by Theorem 3. Thus,

$$Cov\Big(\log I_{X_{a_i}}^{(L_i)}(\lambda_j), \log I_{X_{a_j}}^{(L_j)}(\mu_k)\Big) = o\left(n^{-1}\right).$$

This leads to

$$\begin{aligned}
Cov\left(\widehat{d}_n, \widehat{d}_i\right) &= \frac{1}{4S_{xx}}\frac{1}{S'_{xx}}\sum_{j=1}^{N_n}\sum_{k=1}^{N_l} a_j a_k^{(i)} Cov\left(\log I_{X_{a_i}}^{(L_i)}(\lambda_j), \log I_{X_{a_j}}^{(L_j)}(\mu_k)\right) \\
&\leq \sup_{j,k}\frac{1}{4S_{xx}}\frac{1}{S'_{xx}} N_n N_l \left|a_j a_k^{(i)} Cov\left(\log I_{X_{a_i}}^{(L_i)}(\lambda_j), \log I_{X_{a_j}}^{(L_j)}(\mu_k)\right)\right| \\
&= \frac{(1+o(1))^{-2}}{4}\sup_{j,k}|a_j|\left|a_k^{(i)}\right|\left|Cov\left(\log I_{X_{a_i}}^{(L_i)}(\lambda_j), \log I_{X_{a_j}}^{(L_j)}(\mu_k)\right)\right| \\
&= \frac{(1+o(1))^{-2}}{4}\left(1+o(1)+O\left(\frac{N_n^2}{n^2}\right)\right)^2 \sup_{j,k}\left|Cov\left(\log I_{X_{a_i}}^{(L_i)}(\lambda_j), \log I_{X_{a_j}}^{(L_j)}(\mu_k)\right)\right| \\
&= o\left(n^{-1}\right).
\end{aligned}$$

Similarly, we can prove that $Cov\left(\widehat{d}_i, \widehat{d}_j\right) = o\left(n^{-1}\right)$. Hence the result. ■

Proof of Theorem 5. Consider,

$$\left(\widehat{d}_{J,m}^{Opt} - d_0\right) = w_n^*\left(\widehat{d}_n - d_0\right) - \sum_{i=1}^{m} w_i^*\left(\widehat{d}_{i,m} - d_0\right). \tag{A.8}$$

Recall that $w_n^* = \left[1 - \left(\frac{1}{m}\frac{N_n}{n}\frac{l}{N_l}\right)^2\right]^{-1}$ and $\sum_{i=1}^m w_i^* = w_n^* - 1$; for $i = 1, \ldots, m$. Let us firstly consider w_n^*. For fixed m and for the choice of N_n such that $N_n \log N_n/n \to 0$,

$$w_n^* = \frac{1}{1 - \left(n^{-1} l n^{-1+\alpha} l^{1-\alpha}\right)^2} = 1 + o(1), \tag{A.9}$$

and hence

$$\sum_{i=1}^{m} w_i^* = o(1), \tag{A.10}$$

with $w_i^* \to 0$ as $n \to \infty$ (see the proof of Theorem 4).

By virtue of the consistency of $\widehat{d}_n$, we have that the first term in (A.8) such that $w_n^*\left(\widehat{d}_n - d\right) = o_p(1)$, using (A.9).

Now, we show that the second term in (A.8) is $o_p(1)$.

$$\begin{aligned}
\lim_{n\to\infty}\Pr\left[\left|\sum_{i=1}^{m} w_i^*\left(\widehat{d}_i - d_0\right)\right| \geq \varepsilon\right] &\leq \lim_{n\to\infty}\frac{E\left(\sum_{i=1}^m w_i^*\left(\widehat{d}_i - d_0\right)\right)^2}{\varepsilon^2} \\
&= \lim_{n\to\infty}\frac{Var\left(\widehat{d}_i\right)}{\varepsilon^2}\sum_{i=1}^{m}\left(w_i^*\right)^2 \\
&\quad + \frac{2}{\varepsilon^2}\lim_{n\to\infty}\sum_{i=1}^{m}\sum_{j=i+1}^{m} w_i^* w_j^* Cov\left(\widehat{d}_i, \widehat{d}_j\right) \\
&= 0,
\end{aligned}$$

since $\lim_{n\to\infty} Var\left(\widehat{d}_i\right) = 0$ from Theorem 1, $\lim_{n\to\infty} Cov\left(\widehat{d}_i, \widehat{d}_j\right) = 0$ directly from Theorem 2 and the limit of $\sum_{i=1}^m w_i^*$ given in (A.10). This completes the proof of consistency.

The proof of asymptotic normality of the optimal jackknife estimator depends on the joint convergence of $\widehat{d}_n$ and $\widehat{d}_{i,m}$. Firstly, let us consider the following standardized optimal jackknife estimator,

$$\sqrt{N_n}\left(\widehat{d}_{J,m}^{Opt}-d_0\right)=w_n^*\sqrt{N_n}\left(\widehat{d}_n-d_0\right)-\sum_{i=1}^{m}w_i^*\sqrt{N_n}\left(\widehat{d}_i-d_0\right). \tag{A.11}$$

Using Theorem 1 we have that $\sqrt{N_n}\left(\widehat{d}_n-d_0\right)\rightarrow^D N\left(0,\frac{\pi^2}{24}\right)$. Therefore, regarding the first component in (A.11), it immediately follows that

$$w_n^*\sqrt{N_n}\left(\widehat{d}_n-d_0\right)\rightarrow^d N\left(0,\frac{\pi^2}{24}\right),\text{ using (A.9).}$$

Now, let us consider the second term in (A.11):

$$\begin{aligned}\lim_{n\to\infty}\Pr\left[\left|\sum_{i=1}^{m}w_i^*\sqrt{N_n}\left(\widehat{d}_i-d_0\right)\right|\geq\varepsilon\right]&\leq\lim_{n\to\infty}\frac{E\left(\sum_{i=1}^{m}w_i^*\left(\widehat{d}_i-d_0\right)\right)^2}{\varepsilon^2}N_n\\&=\lim_{n\to\infty}\frac{Var\left(\widehat{d}_i\right)}{\varepsilon^2}N_n\sum_{i=1}^{m}(w_i^*)^2\\&\quad+\lim_{n\to\infty}\frac{2N_n}{\varepsilon^2}\sum_{i=1}^{m}\sum_{j=i+1}^{m}w_i^*w_j^*Cov\left(\widehat{d}_i,\widehat{d}_j\right).\end{aligned} \tag{A.12}$$

By considering the first term in (A.12), for fixed m we have that

$$\lim_{n\to\infty}\frac{Var\left(\widehat{d}_i\right)}{\varepsilon^2}N_n\sum_{i=1}^{m}(w_i^*)^2=\lim_{n\to\infty}\frac{\sum_{i=1}^{m}(w_i^*)^2}{\varepsilon^2}\left[\frac{\pi^2}{24}+o(1)\right]=0,$$

using Theorem 1 and (A.9). The second term in (A.12) would give us that,

$$\lim_{n\to\infty}\frac{2N_n}{\varepsilon^2}\sum_{i=1}^{m}\sum_{j=i+1}^{m}w_i^*w_j^*Cov\left(\widehat{d}_i,\widehat{d}_j\right)=0,$$

immediately from (A.9). Therefore, $\Pr\left[\left|\sum_{i=1}^{m}w_i^*\sqrt{N_l}\left(\widehat{d}_i-d_0\right)\right|\geq\varepsilon\right]\rightarrow 0$ as $n\rightarrow\infty$. Hence the proof completes. ■

Appendix B: Additional technical results

Recall that the covariance between the full-sample LPR estimator and each sub-sample LPR estimator, $Cov\left(\widehat{d}_n,\widehat{d}_i\right)$, and the covariances between the different sub-sample LPR estimators, $Cov\left(\widehat{d}_i,\widehat{d}_j\right)$, for $i\neq j$, $i,j=1,2,...,m$, are given respectively by,

$$Cov\big(\widehat{d}_n,\widehat{d}_i\big)=\frac{1}{4S_{xx}}\frac{1}{S_{xx}'}\sum_{j=1}^{N_n}\sum_{k=1}^{N_l}a_ja_k^{(i)}Cov\big(\log I_Y^{(n)}(\lambda_j),\log I_{Y_i}^{(l)}(\mu_k)\big) \tag{B.1}$$

$$Cov\big(\widehat{d}_i,\widehat{d}_{i'}\big)=\frac{1}{4}\frac{1}{(S_{xx}')^2}\sum_{j=1}^{N_l}\sum_{k=1}^{N_l}a_j'a_k'Cov\big(\log I_{Y_i}^{(l)}(\mu_j),\log I_{Y_{i'}}^{(l)}(\mu_k)\big), \tag{B.2}$$

with all notation as defined in Table 1.

Evaluation of the covariance terms in (B.1) and (B.2)

The main purpose of this exercise is to calculate the covariances between the full-sample and sub-sample LPR estimators (refer to (B.1)) and the covariance between two distinct sub-sample LPR estimators (refer to (B.2)). These covariance terms depend on the covariance between the log-periodograms associated with either the full sample and a given sub-sample or two different sub-samples.

To obtain the covariance between the log-periodograms associated with the full sample and a given sub-sample, or between sub-samples, we follow the method stated below.

Step 1: Write down the joint distribution of the periodograms $(I_{X_{a_i}}^{(L_i)}(\lambda), I_{X_{a_j}}^{(L_j)}(\mu))$.

Step 2: Write down the joint distribution of the log transformed periodograms $(\log I_{X_{a_i}}^{(L_i)}(\lambda), \log I_{X_{a_j}}^{(L_j)}(\mu))$ using the expression of the covariance between the two different periodograms.

Step 3: Find the expression for the covariance between the above mentioned log-periodograms, $Cov(\log I_{X_{a_i}}^{(L_i)}(\lambda), \log I_{X_{a_j}}^{(L_j)}(\mu))$, using the moment generating function.

In relation to Step 1: Using the results of Theorem 3, we can say that the periodograms associated with the full sample and the sub-sample have a limiting distribution of the form $f_{X_1X_1}(\lambda)\, \chi^2_{(2)} \big/ 2$. For notational convenience, let us denote by (U,V) the bivariate χ^2_k random variables, $(I_{X_{a_i}}^{(L_i)}(\lambda), I_{X_{a_j}}^{(L_j)}(\mu))$. Although $k=2$, we use the generic notation for the degrees of freedom, k. Note that we ignore the constant term $f_{X_1X_1}(\lambda)/2$ for convenience, as these terms will disappear in the calculation of the covariance between two different LPR estimators (either the full- and sub-sample LPR estimators or two distinct sub-sample LPR estimators).

The joint probability density function (pdf), $f_{U,V}(u,v)$, is defined by (see, Krishnaiah *et al.*, 1963)

$$f_{U,V}(u,v) = \left(1-\rho^2\right)^{\frac{k-1}{2}} \sum_{i=0}^{\infty} \frac{\Gamma\left(\frac{k-1}{2}+i\right)\rho^{2i}\,(uv)^{\frac{k-3+2i}{2}} \exp\left[-\frac{u+v}{2(1-\rho^2)}\right]}{\Gamma\left(\frac{k-1}{2}\right) i! \left[2^{\frac{k-1}{2}+i}\Gamma\left(\frac{k-1}{2}+i\right)\left(1-\rho^2\right)^{\frac{k-1}{2}+i}\right]^2},$$

where $\rho = \dfrac{\sigma_{uv}}{\sigma_u \sigma_v}$. Here, $\sigma_{uv} = cov(U,V)$. Then, the marginal densities of U and V, $f_U(u)$ and $f_V(v)$, are respectively given by,

$$f_U(u) = \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} u^{\frac{k}{2}} \exp\left\{-\frac{u}{2}\right\}, \text{ and, } f_V(v) = \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} v^{\frac{k}{2}} \exp\left\{-\frac{v}{2}\right\}.$$

In relation to Step 2: Let $W = \log U = \log I_{X_{a_i}}^{(L_i)}(\lambda)$ and $Z = \log V = \log I_{X_{a_j}}^{(L_j)}(\mu)$. Then, the joint pdf of W and Z is given by,

$$\begin{aligned}
f_{W,Z}(w,z) &= f_{U,V}(\exp w, \exp z) \begin{vmatrix} \frac{\partial \exp w}{\partial w} & \frac{\partial \exp w}{\partial z} \\ \frac{\partial \exp z}{\partial w} & \frac{\partial \exp z}{\partial z} \end{vmatrix} \\
&= \left(1-\rho^2\right)^{\frac{k-1}{2}} \sum_{i=0}^{\infty} \frac{\Gamma\left(\frac{k-1}{2}+i\right)\rho^{2i}\left(\exp w \exp z\right)^{\frac{k-3+2i}{2}} \exp\left[-\frac{\exp w + \exp z}{2(1-\rho^2)}\right]}{\Gamma\left(\frac{k-1}{2}\right) i! \left[2^{\frac{k-1}{2}+i}\Gamma\left(\frac{k-1}{2}+i\right)\left(1-\rho^2\right)^{\frac{k-1}{2}+i}\right]^2} \exp w \exp z \\
&= \left(1-\rho^2\right)^{\frac{k-1}{2}} \sum_{i=0}^{\infty} \frac{\Gamma\left(\frac{k-1}{2}+i\right)\rho^{2i} \exp\left(\frac{k-1}{2}+i\right)(w+z) \exp\left[-\frac{\exp w + \exp z}{2(1-\rho^2)}\right]}{\Gamma\left(\frac{k-1}{2}\right) i! \left[2^{\frac{k-1}{2}+i}\Gamma\left(\frac{k-1}{2}+i\right)\left(1-\rho^2\right)^{\frac{k-1}{2}+i}\right]^2}.
\end{aligned}$$

In relation to Step 3: The moment generating function (MGF) of (W, Z) is given by,

$$\begin{aligned}
M_{W,Z}(t_1,t_2) = E\left(\exp\left(t_1 W + t_2 Z\right)\right) &= \int_0^\infty \int_0^\infty \exp\left(t_1 w + t_2 z\right) f_{W,Z}(w,z)\, dw dz \\
&= \left(1-\rho^2\right)^{\frac{k-1}{2}} \sum_{i=0}^{\infty} \frac{\Gamma\left(\frac{k-1}{2}+i\right)\rho^{2i}}{\Gamma\left(\frac{k-1}{2}\right) i! \left[2^{\frac{k-1}{2}+i}\Gamma\left(\frac{k-1}{2}+i\right)\left(1-\rho^2\right)^{\frac{k-1}{2}+i}\right]^2} \\
&\quad \times \int_0^\infty \int_0^\infty \exp\left(t_1 w + t_2 z\right) \exp\left(\frac{k-1}{2}+i\right)(w+z) \exp\left[-\frac{\exp w + \exp z}{2(1-\rho^2)}\right] dw dz \\
&= \left(1-\rho^2\right)^{\frac{k-1}{2}} \sum_{i=0}^{\infty} \frac{\Gamma\left(\frac{k-1}{2}+i\right)\rho^{2i}}{\Gamma\left(\frac{k-1}{2}\right) i! \left[2^{\frac{k-1}{2}+i}\Gamma\left(\frac{k-1}{2}+i\right)\left(1-\rho^2\right)^{\frac{k-1}{2}+i}\right]^2} \\
&\quad \times \int_0^\infty \exp\left(\frac{k-1}{2}+t_1+i\right) w \exp\left[-\frac{\exp w}{2(1-\rho^2)}\right] dw \\
&\quad \times \int_0^\infty \exp\left(\frac{k-1}{2}+t_2+i\right) z \exp\left[-\frac{\exp z}{2(1-\rho^2)}\right] dz. \qquad \text{(B.3)}
\end{aligned}$$

Now let us consider the form of the last expression in (B.3). Let $\alpha_1 = \frac{k-1}{2} + t_2 + i$ and $\alpha_2 = \frac{1}{2(1-\rho^2)}$. Then, substituting $x = \exp z$ would give us that

$$\int_0^\infty \exp \alpha_1 z \exp\left[-\alpha_2 \exp z\right] dz = \int_0^\infty x^{\alpha_1 - 1} \exp\left[-\alpha_2 x\right] dx = \frac{\Gamma(\alpha_1)}{\alpha_2^{\alpha_1}}. \qquad \text{(B.4)}$$

Therefore, using (B.4), the MGF given in (B.3) may be re-arranged as follows,

$$\begin{aligned}
M_{W,Z}(t_1,t_2) &= \left[2\left(1-\rho^2\right)\right]^{t_1+t_2} \left(1-\rho^2\right)^{\frac{k-1}{2}} \sum_{i=0}^{\infty} \frac{\Gamma\left(\frac{k-1}{2}+i\right)\rho^{2i}\Gamma\left(\frac{k-1}{2}+t_2+i\right)\Gamma\left(\frac{k-1}{2}+t_1+i\right)}{i!\Gamma\left(\frac{k-1}{2}\right)\left[\Gamma\left(\frac{k-1}{2}+i\right)\right]^2} \\
&= \left[2\left(1-\rho^2\right)\right]^{t_1+t_2} \left(1-\rho^2\right)^{\frac{k-1}{2}} \frac{\Gamma\left(\frac{k-1}{2}+t_1\right)\Gamma\left(\frac{k-1}{2}+t_2\right)}{\left[\Gamma\left(\frac{k-1}{2}\right)\right]^2} \\
&\quad \times \sum_{i=0}^{\infty} \frac{\Gamma\left(\frac{k-1}{2}+t_1+i\right)\Gamma\left(\frac{k-1}{2}+t_2+i\right)\Gamma\left(\frac{k-1}{2}\right)}{\Gamma\left(\frac{k-1}{2}+t_1\right)\Gamma\left(\frac{k-1}{2}+t_2\right)\Gamma\left(\frac{k-1}{2}+i\right)} \frac{\left(\rho^2\right)^i}{i!} \\
&= \left[2\left(1-\rho^2\right)\right]^{t_1+t_2} \left(1-\rho^2\right)^{\frac{k-1}{2}} \frac{\Gamma\left(\frac{k-1}{2}+t_1\right)\Gamma\left(\frac{k-1}{2}+t_2\right)}{\left[\Gamma\left(\frac{k-1}{2}\right)\right]^2} \\
&\quad \times {}_2F_1\left(\frac{k-1}{2}+t_1, \frac{k-1}{2}+t_2; \frac{k-1}{2}; \rho^2\right).
\end{aligned}$$

Setting $k = 2$ gives,

$$M_{W,Z}(t_1,t_2) = \left[2\left(1-\rho^2\right)\right]^{t_1+t_2}\left(1-\rho^2\right)^{\frac{1}{2}}\frac{\Gamma\left(\frac{1}{2}+t_1\right)\Gamma\left(\frac{1}{2}+t_2\right)}{\left[\Gamma\left(\frac{1}{2}\right)\right]^2}\ {}_2F_1\left(\tfrac{1}{2}+t_1,\tfrac{1}{2}+t_2;\tfrac{1}{2};\rho^2\right).$$

Therefore the cumulant generating function is given by $K(t_1,t_2) = \log M_{W,Z}(t_1,t_2)$ and

$$\begin{aligned} K(t_1,t_2) &= (t_1+t_2)\log\left[2\left(1-\rho^2\right)\right] + \tfrac{1}{2}\log\left(1-\rho^2\right) + \log\Gamma\left(\tfrac{1}{2}+t_1\right) \\ &\quad + \log\Gamma\left(\tfrac{1}{2}+t_2\right) - 2\log\left[\Gamma\left(\tfrac{1}{2}\right)\right] + \log{}_2F_1\left(\tfrac{1}{2}+t_1,\tfrac{1}{2}+t_2;\tfrac{1}{2};\rho^2\right). \end{aligned}$$

The covariance between W and Z when $k = 2$, is given by, $cov(W,Z) = \left.\frac{\partial^2 K(t_1,t_2)}{\partial t_1 \partial t_2}\right|_{t_1=0,t_2=0}$. Therefore, let us firstly evaluate $\partial K(t_1,t_2)/\partial t_1$, as

$$\begin{aligned} \frac{\partial K(t_1,t_2)}{\partial t_1} &= \log\left[2\left(1-\rho^2\right)\right] + \Psi\left(\tfrac{1}{2}+t_1\right) \\ &\quad + \left({}_2F_1\left(\tfrac{1}{2}+t_1,\tfrac{1}{2}+t_2;\tfrac{1}{2};\rho^2\right)\right)^{-1}\frac{\partial_2F_1\left(\frac{1}{2}+t_1,\frac{1}{2}+t_2;\frac{1}{2};\rho^2\right)}{\partial t_1}, \end{aligned} \tag{B.5}$$

where $\Psi(.)$ is the digamma function and $\partial_2F_1\left(\frac{1}{2}+t_1,\frac{1}{2}+t_2;\frac{1}{2};\rho^2\right)/\partial t_1$ is given by,

$$\begin{aligned} &\sum_{i=1}^{\infty}\frac{\partial\,\Gamma\left(\frac{1}{2}+t_1+i\right)/\Gamma\left(\frac{1}{2}+t_1\right)}{\partial t_1}\frac{\Gamma\left(\frac{1}{2}+t_2+i\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}+t_2\right)\Gamma\left(\frac{1}{2}+i\right)}\frac{\left(\rho^2\right)^i}{i!} \\ &= \sum_{i=1}^{\infty}\left(\frac{\Gamma\left(\frac{1}{2}+t_1\right)\Gamma\left(\frac{1}{2}+t_1+i\right)\Psi\left(\frac{1}{2}+t_1+i\right)}{\left(\Gamma\left(\frac{1}{2}+t_1\right)\right)^2} + \frac{\Gamma\left(\frac{1}{2}+t_1+i\right)\Psi\left(\frac{1}{2}+t_1\right)\Gamma\left(\frac{1}{2}+t_1\right)}{\left(\Gamma\left(\frac{1}{2}+t_1\right)\right)^2}\right) \\ &\quad\times\frac{\Gamma\left(\frac{1}{2}+t_2+i\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}+t_2\right)\Gamma\left(\frac{1}{2}+i\right)}\frac{\left(\rho^2\right)^i}{i!} \\ &= \sum_{i=1}^{\infty}\left(\frac{\Gamma\left(\frac{1}{2}+t_1+i\right)\Psi\left(\frac{1}{2}+t_1+i\right)+\Gamma\left(\frac{1}{2}+t_1+i\right)\Psi\left(\frac{1}{2}+t_1\right)}{\Gamma\left(\frac{1}{2}+t_1\right)}\right)\frac{\Gamma\left(\frac{1}{2}+t_2+i\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}+t_2\right)\Gamma\left(\frac{1}{2}+i\right)}\frac{\left(\rho^2\right)^i}{i!}. \end{aligned} \tag{B.6}$$

This leads to,

$$\left.\frac{\partial_2F_1\left(\frac{1}{2}+t_1,\frac{1}{2}+t_2;\frac{1}{2};\rho^2\right)}{\partial t_1}\right|_{t_1=0,t_2=0} = \sum_{i=1}^{\infty}\left(\Psi\left(\tfrac{1}{2}+i\right)+\Psi\left(\tfrac{1}{2}\right)\right)\frac{\Gamma\left(\frac{1}{2}+i\right)}{\Gamma\left(\frac{1}{2}\right)}\frac{\left(\rho^2\right)^i}{i!}.$$

The first derivative of ${}_2F_1\left(\frac{1}{2}+t_1,\frac{1}{2}+t_2;\frac{1}{2};\rho^2\right)$ with respect to t_2 is also given by (B.6).

Now let us evaluate the second order derivative of $K(t_1,t_2)$,

$$\begin{aligned} \frac{\partial^2 K(t_1,t_2)}{\partial t_1\partial t_2} &= \frac{\partial\left({}_2F_1\left(\frac{1}{2}+t_1,\frac{1}{2}+t_2;\frac{1}{2};\rho^2\right)\right)^{-1}\frac{\partial_2F_1\left(\frac{1}{2}+t_1,\frac{1}{2}+t_2;\frac{1}{2};\rho^2\right)}{\partial t_1}}{\partial t_2} \\ &= \left({}_2F_1\left(\tfrac{1}{2}+t_1,\tfrac{1}{2}+t_2;\tfrac{1}{2};\rho^2\right)\right)^{-1}\frac{\partial^2\ {}_2F_1\left(\frac{1}{2}+t_1,\frac{1}{2}+t_2;\frac{1}{2};\rho^2\right)}{\partial t_1\partial t_2} \\ &\quad - \left({}_2F_1\left(\tfrac{1}{2}+t_1,\tfrac{1}{2}+t_2;\tfrac{1}{2};\rho^2\right)\right)^{-2}\frac{\partial_2F_1\left(\frac{1}{2}+t_1,\frac{1}{2}+t_2;\frac{1}{2};\rho^2\right)}{\partial t_2} \\ &\quad \times\frac{\partial_2F_1\left(\frac{1}{2}+t_1,\frac{1}{2}+t_2;\frac{1}{2};\rho^2\right)}{\partial t_1}, \end{aligned}$$

where $\partial^2 \,{}_2F_1\left(\frac{1}{2}+t_1,\frac{1}{2}+t_2;\frac{1}{2};\rho^2\right)/\,\partial t_1\partial t_2$ is given by,

$$\sum_{i=1}^{\infty}\left(\frac{\Gamma\left(\frac{1}{2}+t_1+i\right)\Psi\left(\frac{1}{2}+t_1+i\right)}{\Gamma\left(\frac{1}{2}+t_1\right)}+\frac{\Gamma\left(\frac{1}{2}+t_1+i\right)\Psi\left(\frac{1}{2}+t_1\right)}{\Gamma\left(\frac{1}{2}+t_1\right)}\right)\frac{\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}+i\right)}\frac{\left(\rho^2\right)^i}{i!}$$
$$\times\left(\frac{\Gamma\left(\frac{1}{2}+t_2+i\right)\Psi\left(\frac{1}{2}+t_2+i\right)}{\Gamma\left(\frac{1}{2}+t_2\right)}+\frac{\Gamma\left(\frac{1}{2}+t_2+i\right)\Psi\left(\frac{1}{2}+t_2\right)}{\Gamma\left(\frac{1}{2}+t_2\right)}\right),$$

with

$$\left.\frac{\partial^2 \,{}_2F_1\left(\frac{1}{2}+t_1,\frac{1}{2}+t_2;\frac{1}{2};\rho^2\right)}{\partial t_1\partial t_2}\right|_{t_1=0,t_2=0}=\sum_{i=1}^{\infty}\left(\Psi\left(\tfrac{1}{2}+i\right)+\Psi\left(\tfrac{1}{2}\right)\right)^2\frac{\Gamma\left(\frac{1}{2}+i\right)}{\Gamma\left(\frac{1}{2}\right)}\frac{\left(\rho^2\right)^i}{i!}.$$

Hence $cov\left(W,Z\right)$ is given by,

$$\left(1-\rho^2\right)^{\frac{1}{2}}\left.\frac{\partial^2 \,{}_2F_1\left(\frac{1}{2}+t_1,\frac{1}{2}+t_2;\frac{1}{2};\rho^2\right)}{\partial t_1\partial t_2}\right|_{t_1=0,t_2=0}$$
$$-\left(1-\rho^2\right)\left.\frac{\partial_2F_1\left(\frac{1}{2}+t_1,\frac{1}{2}+t_2;\frac{1}{2};\rho^2\right)}{\partial t_1}\frac{\partial_2F_1\left(\frac{1}{2}+t_1,\frac{1}{2}+t_2;\frac{1}{2};\rho^2\right)}{\partial t_2}\right|_{t_1=0,t_2=0}$$
$$=\left(1-\rho^2\right)^{\frac{1}{2}}\sum_{i=1}^{\infty}\left(\Psi\left(\tfrac{1}{2}+i\right)+\Psi\left(\tfrac{1}{2}\right)\right)^2\frac{\Gamma\left(\frac{1}{2}+i\right)}{\Gamma\left(\frac{1}{2}\right)}\frac{\left(\rho^2\right)^i}{i!}$$
$$-\left(1-\rho^2\right)\left(\sum_{i=1}^{\infty}\left(\Psi\left(\tfrac{1}{2}+i\right)+\Psi\left(\tfrac{1}{2}\right)\right)\frac{\Gamma\left(\frac{1}{2}+i\right)}{\Gamma\left(\frac{1}{2}\right)}\frac{\left(\rho^2\right)^i}{i!}\right)^2, \qquad \text{(B.7)}$$

using the fact ${}_1F_0\left(a;;z\right)=\left(1-z\right)^{-a}$.

Let us now provide the expression for ρ in (B.7). For example, consider calculating the correlation between the full- and sub-sample periodograms. Using the similar arguments, the correlation between two sub-samples periodograms, $\rho=corr\left(I_Y^{(n)}\left(\lambda\right),I_{Y_i}^{(l)}\left(\mu\right)\right)$ can be derived using

$$Cov\left(I_Y^{(n)}\left(\lambda\right),I_{Y_i}^{(l)}\left(\mu\right)\right)\approx\frac{2\pi}{l}f_{YYY_iY_i}\left(\lambda,-\lambda,\mu\right)+l^{-2}\left[\Delta^{(l)}\left(-\lambda+\mu\right)\Delta^{(l)}\left(\lambda-\mu\right)\right.$$
$$\left.+\,\Delta^{(l)}\left(\lambda+\mu\right)\Delta^{(l)}\left(-\lambda-\mu\right)\right]\left|f_{YY_i}\left(\lambda\right)\right|^2, \qquad \text{(B.8)}$$

and $Var\left(I_Y^{(n)}\left(\lambda\right)\right)$ and $Var\left(I_{Y_i}^{(l)}\left(\mu\right)\right)$ can be calculated from the above given covariance formula. The covariance and variance terms rely upon certain joint spectral densities. Those spectral densities can be expressed in closed form as follows. Let us firstly consider the cross spectrum corresponding to the full sample and j^{th} sub-sample, $f_{YY_j}\left(\lambda\right)$. Suppose we consider the jackknife approach using non-overlapping sub-samples. Then, the general definition of spectral density gives that

$$\begin{aligned} f_{YY_j}(\lambda) &= \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\exp(-ik\lambda)\,\kappa\left(Y_{t+k},Y_{t+(j-1)l}\right) \\ &= \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\exp(-ik\lambda)\,\gamma(k-(j-1)\,l) \\ &= \frac{\exp(-i(j-1)\,l\lambda)}{2\pi}\sum_{k=-\infty}^{\infty}\exp(-i(k-(j-1)\,l)\,\lambda)\,\gamma(k-(j-1)\,l) \\ &= \exp(-i(j-1)\,l\lambda)\,f_{YY}(\lambda)\,. \end{aligned}$$

Similarly, for moving-block sub-samples we have the relationship $f_{YY_j}(\lambda) = \exp(-i(j+l-1)\,\lambda)\,f_{YY}(\lambda)$ and $f_{Y_jY_k}(\lambda) = \exp(-i(j-k)\,l\lambda)\,f_{YY}(\lambda)\,.$

Lemma 2 of Yajima (1989) immediately gives that,

$$f_{YYYY}(\lambda,-\lambda,\mu) = \frac{1}{(2\pi)^3}b(\lambda)\,b(-\lambda)\,b(\mu)\,b(-\mu)\,f_{\varepsilon\varepsilon\varepsilon\varepsilon}(\lambda,-\lambda,\mu)\,,$$

where $b(\lambda) = \sum_{j=0}^{\infty} b_j \exp(\imath j\omega)$ with $b_j = \sum_{r=0}^{j}\dfrac{k(j-r)\,\Gamma(r+d)}{\Gamma(r+1)\,\Gamma(d)}$, and $k(z)$ is the transfer function of a stable and invertible autoregressive moving average (ARMA) process such that $\sum_{j=0}^{\infty}|k(j)| < \infty$. Here,

$$f_{\varepsilon\varepsilon\varepsilon\varepsilon}(\lambda,-\lambda,\mu) = \sum_{u_1=-\infty}^{\infty}\sum_{u_2=-\infty}^{\infty}\sum_{u_3=-\infty}^{\infty}\exp(-i(\lambda u_1-\lambda u_2+\mu u_3))\,\kappa_{\varepsilon\varepsilon\varepsilon\varepsilon}(u_1,u_2,u_3)\,,$$

where

$$\begin{aligned} \kappa_{\varepsilon\varepsilon\varepsilon\varepsilon}(u_1,u_2,u_3) &= \kappa(\varepsilon_{t+u_1},\varepsilon_{t+u_2},\varepsilon_{t+u_3},\varepsilon_t) \\ &= E(\varepsilon_{t+u_1}\varepsilon_{t+u_2}\varepsilon_{t+u_3}\varepsilon_t) - E(\varepsilon_{t+u_1}\varepsilon_{t+u_2})\,E(\varepsilon_{t+u_3}\varepsilon_t) \\ &\quad - E(\varepsilon_{t+u_2}\varepsilon_{t+u_3})\,E(\varepsilon_{t+u_1}\varepsilon_t) - E(\varepsilon_{t+u_1}\varepsilon_{t+u_3})\,E(\varepsilon_{t+u_2}\varepsilon_t)\,. \end{aligned}$$

Suppose the errors are $i.i.d$ normal random variables with zero mean and a constant variance σ^2,

$$\begin{aligned} \kappa_{\varepsilon\varepsilon\varepsilon\varepsilon}(u_1,u_2,u_3) &= \begin{cases} E\left(\varepsilon_t^4\right) - 3\left(E\left(\varepsilon_t^2\right)\right)^2, & if\ u_1=u_2=u_3=0 \\ 0, & otherwise \end{cases} \\ &= \begin{cases} 3\sigma^4, & if\ u_1=u_2=u_3=0 \\ 0, & otherwise \end{cases}. \end{aligned}$$

Then $f_{YYYY}(\lambda,-\lambda,\mu)$ is simplified as follows using the fact that $f_{YY}(\lambda) = \dfrac{\sigma^2}{2\pi}b(\lambda)\,b(-\lambda)$.

$$f_{YYYY}(\lambda,-\lambda,\mu) = \frac{3\sigma^4}{(2\pi)^3}b(-\lambda)\,b(\lambda)\,b(\mu)\,b(-\mu) = \frac{3}{2\pi}f_{YY}(\lambda)\,f_{YY}(\mu)\,.$$

Now let us consider $f_{YYY_jY_j}(\lambda,-\lambda,\mu)$.

$$\begin{aligned} f_{YYY_jY_j}(\lambda,-\lambda,\mu) &= \frac{1}{(2\pi)^3}\sum_{u_1=-\infty}^{\infty}\sum_{u_2=-\infty}^{\infty}\sum_{u_3=-\infty}^{\infty}\exp\left(-\imath\left(\lambda u_1-\lambda u_2+\mu u_3\right)\right)\\ &\quad\times\kappa\left(Y_{t+u_1},Y_{t+u_2},Y_{t+(j-1)l+u_3},Y_{t+(j-1)l}\right)\\ &= \frac{1}{(2\pi)^3}\sum_{u_1=-\infty}^{\infty}\sum_{u_2=-\infty}^{\infty}\sum_{u_3=-\infty}^{\infty}\exp\left(-\imath\left(\lambda\left(u_1-(j-1)l\right)-\lambda\left(u_2-(j-1)l\right)+\mu u_3\right)\right)\\ &\quad\times\kappa\left(Y_{t-(j-1)l+u_1},Y_{t-(j-1)l+u_2},Y_{t+u_3},Y_t\right)\\ &= f_{YYYY}(\lambda,-\lambda,\mu)\,. \end{aligned}$$

The covariance and variance terms in (B.8) can thus be simplified as follows.

$$\begin{aligned} Cov\left(I_Y^{(n)}(\lambda),I_{Y_i}^{(l)}(\mu)\right) &\approx \frac{3}{l}f_{YY}(\lambda)f_{YY}(\mu)+\frac{1}{l^2}\left[\Delta^{(l)}(-\lambda+\mu)\Delta^{(l)}(\lambda-\mu)\right.\\ &\quad\left.+\Delta^{(l)}(\lambda+\mu)\Delta^{(l)}(-\lambda-\mu)\right]\left(f_{YY}(\lambda)\right)^2,\\ Var\left(I_Y^{(n)}(\lambda)\right) &\approx \left[1+\frac{3}{l}+\frac{1}{l^2}\Delta^{(l)}(2\lambda)\Delta^{(l)}(-2\lambda)\right]\left(f_{YY}(\lambda)\right)^2. \end{aligned}$$

Hence, the correlation is given by,

$$\rho\approx\frac{\frac{3}{l}+\frac{1}{l^2}\left[\Delta^{(l)}(-\lambda+\mu)\Delta^{(l)}(\lambda-\mu)+\Delta^{(l)}(\lambda+\mu)\Delta^{(l)}(-\lambda-\mu)\right]\frac{f_{YY}(\lambda)}{f_{YY}(\mu)}}{\sqrt{\left(1+\frac{3}{l}+\frac{1}{l^2}\Delta^{(l)}(2\lambda)\Delta^{(l)}(-2\lambda)\right)}\sqrt{\left(1+\frac{3}{l}+\frac{1}{l^2}\Delta^{(l)}(2\mu)\Delta^{(l)}(-2\mu)\right)}}.$$

Positiveness of the principle minors of the bordered Hessian matrix

Here we show that for every $m\in\mathbb{N}$, $\left|\mathbf{H}^B_{(m+3)\times(m+3)}\right|>0$ using mathematical induction. For our convenience, we assume that

$$\varphi_{\min}\left(\mathbf{H}^B_{(m+3)\times(m+3)}\right)>(m+3)^2\,\frac{12N_l}{\pi^2},$$

where $\varphi_{\min}(\mathbf{A})$ is the minimum eigenvalue corresponding to the matrix $\mathbf{A}$.

Let us start with $m=1$. The first minor of the bordered Hessian matrix, $\mathbf{H}^B_{4\times4}$, is,

$$\begin{aligned} \left|\mathbf{H}^B_{4\times4}\right| &= \begin{vmatrix} 0 & 0 & -m^2\frac{N_l^2}{l^2}\\ 1 & \frac{N_n^2}{n^2} & -2c^*_{n,1}\\ -1 & -m^2\frac{N_l^2}{l^2} & \frac{\pi^2}{12N_l} \end{vmatrix} + \begin{vmatrix} 0 & 0 & \frac{N_n^2}{n^2}\\ 1 & \frac{N_n^2}{n^2} & \frac{\pi^2}{12N_n}\\ -1 & -m^2\frac{N_l^2}{l^2} & -2c^*_{n,1} \end{vmatrix}\\ &= -m^2\frac{N_l^2}{l^2}\left(-m^2\frac{N_l^2}{l^2}+\frac{N_n^2}{n^2}\right)+\frac{N_n^2}{n^2}\left(-m^2\frac{N_l^2}{l^2}+\frac{N_n^2}{n^2}\right)=\left(\frac{N_n^2}{n^2}-m^2\frac{N_l^2}{l^2}\right)^2>0. \end{aligned}$$

That is, $\left|\mathbf{H}^B_{(m+3)\times(m+3)}\right|>0$ for $m=1$.

Suppose that $\left|\mathbf{H}^B_{(m+3)\times(m+3)}\right|>0$ is true for $m=k$, then we need to show that it is true for

$m = k+1$. To do so, we consider the partition of $\mathbf{H}^B_{(k+4)\times(k+4)}$ is as follows:

$$\mathbf{H}^B_{(k+4)\times(k+4)} = \begin{pmatrix} \mathbf{H}^B_{(k+3)\times(k+3)} & \mathbf{U} \\ \mathbf{U}^T & \frac{\pi^2}{12N_l} \end{pmatrix},$$

where $\mathbf{U}^\top = \begin{bmatrix} -1 & -(k+1)^2\frac{N_l^2}{l^2} & -2c^*_{n,k+1} & 2c^\dagger_{1,k+1} & \dots & 2c^\dagger_{k,k+1} \end{bmatrix}$. Then,

$$\left|\mathbf{H}^B_{(k+4)\times(k+4)}\right| = \left|\mathbf{H}^B_{(k+3)\times(k+3)}\right| \left(\frac{\pi^2}{12N_l} - \mathbf{U}^\top \left(\mathbf{H}^B_{(k+3)\times(k+3)}\right)^{-1} \mathbf{U}\right).$$

Since $\left|\mathbf{H}^B_{(k+3)\times(k+3)}\right| > 0$,

$$0 < \mathbf{U}^\top \left(\mathbf{H}^B_{(k+3)\times(k+3)}\right)^{-1} \mathbf{U} \leq \frac{1}{\varphi_{\min}\left(\mathbf{H}^B_{(k+3)\times(k+3)}\right)} \max_{\mathbf{U}\in\mathbb{R}^{k+3}\setminus\{\mathbf{0}\}} \mathbf{U}^\top\mathbf{U} < \frac{\pi^2}{12N_l}, \text{ as } \max_{\mathbf{U}\in\mathbb{R}^{k+3}\setminus\{\mathbf{0}\}} \mathbf{U}^\top\mathbf{U} = 1.$$

Hence this completes the proof.

Appendix C: Monte Carlo results: Tables 2 to 16

Table 2: Bias estimates of the unadjusted LPR estimator, the optimal jackknife estimator based on 2,3,4,6,8 non-overlapping (NO) sub-samples, the optimal jackknife estimator based on 2 moving block (MB) sub-samples, both versions of the GS estimator, the pre-filtered sieve bootstrap estimator, the maximum likelihood estimator (MLE) and the pre-whitened (PW) estimator, for the DGP: ARFIMA$(1, d_0, 0)$ with Gaussian innovations. The optimal jackknife estimates are evaluated as described in Section 5.1. The estimates are obtained by setting α = 0.65 and assuming the model is correctly specified. The lowest values are **bold-faced** and the second lowest values are *italicized.*

ϕ_0	d_0	n	$\widehat{d}_n$	$\widehat{d}_{J,2}^{Opt\text{-}NO}$	$\widehat{d}_{J,3}^{Opt\text{-}NO}$	$\widehat{d}_{J,4}^{Opt\text{-}NO}$	$\widehat{d}_{J,6}^{Opt\text{-}NO}$	$\widehat{d}_{J,8}^{Opt\text{-}NO}$	$\widehat{d}_{J,2}^{Opt\text{-}MB}$	$\widehat{d}_1^{GS}$	$\widehat{d}_1^{Opt\text{-}GS}$	$\widehat{d}^{PFSB}$	$\widehat{d}^{MLE}$	$\widehat{d}^{PW}$
-0.9	-0.25	96	0.8145	0.7852	0.7903	0.7995	0.8072	0.8120	0.8156	0.8002	0.7902	0.7908	**0.6408**	*0.7047*
		576	0.5945	0.5614	0.5682	0.5726	0.5804	0.5946	0.5841	0.5724	0.5657	0.5898	**0.5051**	*0.5520*
	0	96	0.8053	0.7865	0.7945	0.7988	0.8042	0.8169	0.7927	0.8015	0.7957	0.7955	**0.7026**	*0.7373*
		576	0.5912	0.5581	0.5627	0.5699	0.5773	0.5843	0.5608	0.5761	0.5630	0.5888	**0.4905**	*0.5264*
	0.25	96	0.7752	*0.7477*	0.7515	0.7694	0.7747	0.7804	0.7799	0.7673	0.7517	0.7685	**0.7182**	0.7589
		576	0.5883	0.5553	0.5622	0.5687	0.5731	0.5816	0.5673	0.5716	0.5628	0.5638	**0.4943**	*0.5381*
	0.45	96	0.7006	0.6783	0.6842	0.6905	0.7046	0.7172	0.6945	0.6946	0.6846	*0.6705*	**0.6182**	0.6858
		576	0.5748	0.5423	0.5487	0.5535	0.5586	0.5629	0.5567	0.5659	0.5580	0.5451	**0.4941**	*0.5225*
-0.4	-0.25	96	0.1756	*0.1223*	0.1344	0.1459	0.1563	0.1660	0.1560	0.1367	0.1286	0.1435	**0.1108**	0.1444
		576	0.0607	**0.0043**	0.0429	0.0534	0.0585	0.0599	0.0599	0.0304	*0.0245*	0.0286	0.0362	0.0488
	0	96	0.1653	*0.1203*	0.1216	0.1395	0.1596	0.1674	0.1674	0.1304	0.1276	0.1353	**0.0904**	0.1206
		576	0.0560	**0.0127**	0.0253	0.0307	0.0479	0.0569	0.0369	0.0264	*0.0152*	0.0249	0.0216	0.0371
	0.25	96	0.1629	*0.1190*	0.1274	0.1314	0.1508	0.1665	0.0731	0.1329	0.1276	0.1294	**0.1084**	0.1243
		576	0.0571	*0.0179*	0.0243	0.0341	0.0431	0.0599	0.0599	0.0289	0.0181	0.0251	**0.0178**	0.0239
	0.45	96	0.1653	*0.1154*	0.1226	0.1353	0.1560	0.1702	0.1702	0.1400	0.1245	0.1277	**0.1042**	0.1215
		576	0.0625	*0.0203*	0.0325	0.0495	0.0518	0.0667	0.0667	0.0359	0.0217	0.0261	**0.0197**	0.0258
0.4	-0.25	96	-0.0363	-0.0194	-0.0136	-0.0259	-0.0323	-0.0493	-0.0393	**-0.0047**	*-0.0068*	-0.0147	-0.0068*	-0.0256
		576	-0.0056	*-0.0004**	-0.0037	-0.0046	-0.0057	-0.0076	-0.0076	0.0056	-0.0027	**-0.0004**	-0.0026	-0.0122
	0	96	-0.0534	-0.0114	-0.0145	-0.0298	-0.0360	-0.0449	-0.0549	*-0.0089*	-0.0092	-0.0175	**-0.0065**	-0.0178
		576	-0.0125	*-0.0007*	-0.0049	-0.0038	-0.0031	-0.0028	-0.0128	-0.0008	-0.0007*	-0.0040	**-0.0006**	-0.0064
	0.25	96	-0.0559	-0.0121	-0.0188	-0.0281	-0.0350	-0.0458	-0.0558	*-0.0068*	**-0.0050**	-0.0153	-0.0072	-0.0196
		576	-0.0115	**-0.0003**	-0.0014	-0.0024	-0.0079	-0.0100	-0.0100	0.0017	*-0.0008*	-0.0027	-0.0016	-0.0063
	0.45	96	-0.0501	-0.0091	-0.0092	-0.0302	-0.0460	-0.0486	-0.0486	**0.0032**	0.0090	-0.0111	*-0.0085*	-0.0129
		576	-0.0058	**-0.0003**	-0.0037	-0.0054	-0.0062	-0.0078	-0.0028	0.0089	-0.0061	*0.0004*	-0.0007	-0.0082
0.9	-0.25	96	-0.0291	*-0.0150*	-0.0167	-0.0213	-0.0276	-0.0312	-0.0245	-0.0175	-0.0153	-0.0162	**-0.0039**	-0.0166
		576	-0.0058	*-0.0003*	-0.0020	-0.0035	-0.0059	-0.0080	-0.0040	-0.0034	-0.0011	-0.0023	**-0.0002**	-0.0054
	0	96	-0.0170	*-0.0076*	-0.0131	-0.0149	-0.0184	-0.0222	-0.0117	-0.0140	-0.0115	-0.0082	**-0.0024**	-0.0101
		576	-0.0029	*-0.0001**	-0.0009	-0.0018	-0.0023	-0.0044	-0.0011	-0.0009	-0.0004	-0.0005	**-0.0001**	-0.0028
	0.25	96	-0.0249	*-0.0112*	-0.0156	-0.0184	-0.0207	-0.0269	-0.0156	-0.0162	-0.0129	-0.0117	**-0.0017**	-0.0177
		576	-0.0044	*-0.0019*	-0.0038	-0.0041	-0.0068	-0.0081	-0.0055	-0.0032	-0.0020	-0.0020	**-0.0002**	-0.0033
	0.45	96	-0.0241	*-0.0095*	-0.0157	-0.0198	-0.0226	-0.0275	-0.0218	-0.0175	-0.0112	-0.0126	**-0.0011**	-0.0185
		576	-0.0077	*-0.0017*	-0.0029	-0.0038	-0.0042	-0.0065	-0.0021	-0.0038	-0.0026	-0.0018	**-0.0003**	-0.0031

Table 3: RMSE estimates of the unadjusted LPR estimator, the optimal jackknife estimator based on 2,3,4,6,8 non-overlapping (NO) sub-samples, the optimal jackknife estimator based on 2 moving block (MB) sub-samples, both versions of the GS estimator, the pre-filtered sieve bootstrap estimator, the maximum likelihood estimator (MLE) and the pre-whitened (PW) estimator, for the DGP: ARFIMA$(1, d_0, 0)$ with Gaussian innovations. The optimal jackknife estimates are evaluated as described in Section 5.1. The estimates are obtained by setting α = 0.65 and assuming the model is correctly specified. The lowest values are **bold-faced** and the second lowest values are *italicized.*

ϕ_0	d_0	n	$\widehat{d}_n$	$\widehat{d}^{Opt\text{-}NO}_{J,2}$	$\widehat{d}^{Opt\text{-}NO}_{J,3}$	$\widehat{d}^{Opt\text{-}NO}_{J,4}$	$\widehat{d}^{Opt\text{-}NO}_{J,6}$	$\widehat{d}^{Opt\text{-}NO}_{J,8}$	$\widehat{d}^{Opt\text{-}MB}_{J,2}$	$\widehat{d}^{GS}_1$	$\widehat{d}^{Opt\text{-}GS}_1$	$\widehat{d}^{PFSB}$	$\widehat{d}^{MLE}$	$\widehat{d}^{PW}$
-0.9	-0.25	96	1.0359	1.0627	1.0532	1.0596	1.0358	1.0286	1.1837	1.3386	1.1864	1.2885	**0.7257**	*0.9158*
		576	0.7398	0.7490	0.7403	0.7372	0.7325	0.7299	0.7382	0.7371	0.7200	0.7359	**0.6353**	*0.6994*
	0	96	1.1148	1.1398	1.1275	1.1158	1.1080	1.0966	1.1576	1.1819	1.1120	1.2167	**0.7380**	*0.9181*
		576	0.8288	0.8370	0.8311	0.8294	0.8216	0.8157	0.8215	0.8173	0.8173	0.8053	**0.5261**	*0.5429*
	0.25	96	1.1618	1.1857	1.1066	1.0971	1.0944	1.0913	1.1162	1.1484	1.1285	1.2299	**0.7492**	*0.9726*
		576	0.9175	0.9250	0.9203	0.9186	0.9128	0.9076	0.9115	1.1171	1.0172	1.1130	**0.5258**	*0.5530*
	0.45	96	1.1286	1.1552	1.1325	1.1294	1.1200	1.1168	1.1132	1.4331	1.3331	1.5385	**0.6482**	*0.9438*
		576	0.9708	0.9781	0.9732	0.9650	0.9558	0.9546	0.9687	1.1124	1.0524	1.1647	**0.5263**	*0.5492*
-0.4	-0.25	96	0.2568	0.2292	0.2568	0.2422	0.2384	0.2376	0.2576	0.2594	0.2441	0.3028	**0.1308**	*0.1953*
		576	0.1098	0.0978	0.0974	0.0884	*0.0873*	0.0896	0.1096	0.1118	0.0995	0.1272	**0.0662**	0.0948
	0	96	0.2498	0.2395	0.2284	0.2146	0.2138	0.2117	0.2517	0.2560	0.2416	0.2930	**0.1309**	*0.1999*
		576	0.1069	0.0837	0.0879	0.0819	0.0787	*0.0778*	0.1078	0.1104	0.0967	0.1247	**0.0530**	0.1065
	0.25	96	0.2490	0.2678	0.2574	0.2435	0.2354	0.2254	0.3254	0.2580	0.2404	0.2879	**0.1382**	*0.1896*
		576	0.1079	0.1036	0.0965	0.0901	0.0819	*0.0797*	0.1097	0.1115	0.1029	0.1239	**0.0528**	0.1047
	0.45	96	0.2506	0.2615	0.2563	0.2434	0.2390	0.2243	0.2544	0.2616	0.2511	0.2506	**0.1371**	*0.1966*
		576	0.1115	0.0963	0.0878	0.0808	0.0777	*0.0742*	0.1142	0.1143	0.1005	0.1230	**0.0593**	0.1028
0.4	-0.25	96	0.1917	0.1721	0.1654	0.1629	0.1544	0.1529	0.1929	0.2212	0.2157	0.2717	**0.0904**	*0.1445*
		576	0.0919	0.0762	0.0747	0.0665	0.0632	*0.0624*	0.0924	0.1081	0.0695	0.1198	**0.0335**	0.0764
	0	96	0.1946	0.1726	0.1717	0.1631	0.1569	0.1557	0.1957	0.2203	0.2162	0.2546	**0.0872**	*0.1439*
		576	0.0920	0.0890	0.0793	0.0751	0.0730	0.0724	0.0924	0.1073	*0.0684*	0.1166	**0.0434**	0.0753
	0.25	96	0.1960	0.2107	0.2063	0.2008	0.1913	0.1966	0.1966	0.2209	0.2091	0.2482	**0.0912**	*0.1535*
		576	0.0922	0.0705	0.0696	0.0644	0.0627	*0.0624*	0.0924	0.1076	0.0688	0.1158	**0.0381**	0.0736
	0.45	96	0.1955	0.2178	0.2140	0.2085	0.2061	0.2058	0.1958	0.2218	0.2143	0.2453	**0.0944**	*0.1538*
		576	0.0926	0.0710	0.0684	0.0667	0.0634	*0.0569*	0.0929	0.1089	0.0701	0.1149	**0.0499**	0.0752
0.9	-0.25	96	0.1115	0.1039	0.1006	0.0994	0.0913	0.0886	0.0932	0.1365	0.1132	0.1266	**0.0482**	*0.0872*
		576	0.0624	0.0522	0.0513	0.0482	0.0440	0.0402	0.0399	0.0708	0.0659	0.0600	**0.0127**	*0.0331*
	0	96	0.1010	0.1012	0.0954	0.0911	0.0827	0.0813	0.0955	0.1121	0.0992	0.1093	**0.0438**	*0.0838*
		576	0.0602	0.0504	0.0486	0.0455	0.0422	0.0391	0.0400	0.0698	0.0632	0.0705	**0.0121**	*0.0323*
	0.25	96	0.1114	0.1053	0.1011	0.0942	0.0930	0.0913	0.1106	0.1328	0.1179	0.1282	**0.0463**	*0.0880*
		576	0.0518	0.0500	0.0482	0.0438	0.0419	0.0374	0.0491	0.0626	0.0573	0.0581	**0.0139**	*0.0341*
	0.45	96	0.1053	0.0992	0.0914	0.0824	0.0862	*0.0801*	0.0937	0.1253	0.1188	0.1215	**0.0418**	0.0868
		576	0.0526	0.0518	0.0583	0.0503	0.0455	0.0412	0.0527	0.0769	0.0600	0.0684	**0.0122**	*0.0351*

Table 4: Bias estimates of the unadjusted LPR estimator, the optimal jackknife estimator based on 2,3,4,6,8 non-overlapping (NO) sub-samples, the optimal jackknife estimator based on 2 moving block (MB) sub-samples, both versions of the GS estimator, the pre-filtered sieve bootstrap estimator, the maximum likelihood estimator (MLE) and the pre-whitened (PW) estimator, for the DGP: ARFIMA$(0, d_0, 1)$ with Gaussian innovations. The optimal jackknife estimates are evaluated as described in Section 5.1. The estimates are obtained by setting α = 0.65 and assuming the model is correctly specified. The lowest values are **bold-faced** and the second lowest values are *italicized*.

θ_0	d_0	n	$\widehat{d}_n$	$\widehat{d}^{Opt\text{-}NO}_{J,2}$	$\widehat{d}^{Opt\text{-}NO}_{J,3}$	$\widehat{d}^{Opt\text{-}NO}_{J,4}$	$\widehat{d}^{Opt\text{-}NO}_{J,6}$	$\widehat{d}^{Opt\text{-}NO}_{J,8}$	$\widehat{d}^{Opt\text{-}MB}_{J,2}$	$\widehat{d}^{GS}_1$	$\widehat{d}^{Opt\text{-}GS}_1$	$\widehat{d}^{PFSB}$	$\widehat{d}^{MLE}$	$\widehat{d}^{PW}$
-0.9	-0.25	96	-0.5671	-0.5276	-0.5348	-0.5429	-0.5574	-0.5653	-0.5536	-0.5450	-0.5329	-0.5466	**-0.3341**	*-0.4090*
		576	-0.4527	-0.4149	-0.4266	-0.4357	-0.4404	-0.4595	-0.4375	-0.4385	-0.4248	-0.4285	**-0.1068**	*-0.1539*
	0	96	-0.7042	-0.6416	-0.6502	-0.6642	-0.6743	-0.6869	-0.6724	-0.6575	-0.6476	-0.6664	**-0.3050**	*-0.4042*
		576	-0.5594	-0.5112	-0.5259	-0.5384	-0.5469	-0.5572	-0.5346	-0.5256	-0.5156	-0.5375	**-0.0972**	*-0.1577*
	0.25	96	-0.7763	-0.7299	-0.7345	-0.7466	-0.7547	-0.7681	-0.7367	-0.7524	-0.7425	-0.7661	**-0.3175**	*-0.4132*
		576	-0.5880	-0.5299	-0.5374	-0.5450	-0.5581	-0.5623	-0.5348	-0.5473	-0.5373	-0.5621	**-0.0962**	*-0.1521*
	0.45	96	-0.8004	-0.7414	-0.7588	-0.7615	-0.7741	-0.7878	-0.7649	-0.7600	-0.7501	-0.7854	**-0.3135**	*-0.4142*
		576	-0.5880	-0.5061	-0.5127	-0.5349	-0.5457	-0.5537	-0.5224	-0.5351	-0.5151	-0.5527	**-0.0984**	*-0.1564*
-0.4	-0.25	96	-0.1437	*-0.1013*	-0.1152	-0.1105	-0.1211	-0.1371	-0.1271	-0.1120	-0.1057	-0.1240	**-0.0512**	-0.1038
		576	-0.0476	-0.0342	-0.0234	-0.0139	-0.0234	-0.0303	-0.0303	-0.0187	**-0.0123**	-0.0271	*-0.0135*	-0.0392
	0	96	-0.1653	-0.1199	-0.1213	-0.1293	-0.1394	-0.1472	-0.1472	-0.1305	-0.1209	-0.1248	**-0.0569**	*-0.1023*
		576	-0.0560	*-0.0226*	-0.0353	-0.0407	-0.0579	-0.0570	-0.0370	-0.0265	-0.0274	-0.0307	**-0.0118**	-0.0374
	0.25	96	-0.1692	-0.1136	-0.1273	-0.1292	-0.1398	-0.1496	-0.1496	-0.1297	-0.1170	-0.1200	**-0.0591**	*-0.1061*
		576	-0.0552	**-0.0122**	-0.0366	-0.0475	-0.0529	-0.0543	-0.0443	-0.0243	*-0.0160*	-0.0287	-0.0163	-0.0339
	0.45	96	-0.1630	*-0.0712*	-0.1374	-0.1510	-0.1605	-0.1620	-0.1420	-0.1190	-0.1036	-0.1118	**-0.0546**	-0.1033
		576	-0.0493	*-0.0155*	-0.0177	-0.0314	-0.0436	-0.0436	-0.0268	-0.0169	**-0.0126**	-0.0244	-0.0182	-0.0312
0.4	-0.25	96	0.0637	**0.0036**	0.0475	0.0563	0.0628	0.0637	0.0437	0.0154	*0.0092*	0.0651	0.0119	0.0339
		576	0.0175	**0.0037**	0.0092	0.0068	0.0141	0.0161	0.0061	0.0049	*0.0040*	0.0132	0.0062	0.0185
	0	96	0.0525	0.0202	0.0234	0.0288	0.0351	0.0340	0.0340	0.0081	*0.0077*	0.0603	**0.0067**	0.0342
		576	0.0125	0.0088	0.0148	0.0137	0.0130	0.0128	0.0088	*0.0006*	0.0007	0.0100	**0.0005**	0.0169
	0.25	96	0.0504	0.0164	0.0397	0.0511	0.0566	0.0535	0.0335	0.0110	*0.0095*	0.0574	**0.0085**	0.0218
		576	0.0136	*0.0028*	0.0048	0.0072	0.0083	0.0157	0.0057	0.0031	0.0030	0.0108	**0.0018**	0.0102
	0.45	96	0.0549	0.0192	0.0375	0.0474	0.0641	0.0592	0.0393	0.0204	*0.0112*	0.0570	**0.0077**	0.0225
		576	0.0192	*0.0049*	0.0072	0.0069	0.0073	0.0129	0.0119	0.0103	0.0050	0.0132	**0.0013**	0.0116
0.9	-0.25	96	0.0359	0.0082	0.0106	0.0166	0.0203	0.0245	0.0246	0.0109	*0.0076*	0.0085	**0.0051**	0.0199
		576	0.0065	*0.0009*	0.0011	0.0025	0.0033	0.0053	0.0031	0.0020	0.0009*	0.0014	**0.0002**	0.0068
	0	96	0.0347	*0.0073*	0.0086	0.0098	0.0106	0.0132	0.0101	0.0091	0.0081	0.0076	**0.0038**	0.0184
		576	0.0052	*0.0007*	0.0010	0.0016	0.0021	0.0039	0.0037	0.0015	0.0009	0.0010	**0.0005**	0.0061
	0.25	96	0.0293	*0.0065*	0.0072	0.0089	0.0115	0.0146	0.0102	0.0130	0.0070	0.0073	**0.0032**	0.0121
		576	0.0083	*0.0012*	0.0018	0.0021	0.0034	0.0047	0.0014	0.0057	0.0012*	0.0019	**0.0010**	0.0055
	0.45	96	0.0235	*0.0068*	0.0079	0.0095	0.0129	0.0168	0.0119	0.0132	0.0086	0.0075	**0.0043**	0.0132
		576	0.0195	*0.0035*	0.0058	0.0069	0.0105	0.0121	0.0086	0.0071	0.0037	0.0042	**0.0006**	0.0063

Table 5: RMSE Bias estimates of the unadjusted LPR estimator, the optimal jackknife estimator based on 2,3,4,6,8 non-overlapping (NO) sub-samples, the optimal jackknife estimator based on 2 moving block (MB) sub-samples, both versions of the GS estimator, the pre-filtered sieve bootstrap estimator, the maximum likelihood estimator (MLE) and the pre-whitened (PW) estimator, for the DGP: ARFIMA$(0, d_0, 1)$ with Gaussian innovations. The optimal jackknife estimates are evaluated as described in Section 5.1. The estimates are obtained by setting $\alpha = 0.65$ and assuming the model is correctly specified. The lowest values are **bold-faced** and the second lowest values are *italicized.*

θ_0	d_0	n	$\widehat{d}_n$	$\widehat{d}_{J,2}^{Opt\text{-}NO}$	$\widehat{d}_{J,3}^{Opt\text{-}NO}$	$\widehat{d}_{J,4}^{Opt\text{-}NO}$	$\widehat{d}_{J,6}^{Opt\text{-}NO}$	$\widehat{d}_{J,8}^{Opt\text{-}NO}$	$\widehat{d}_{J,2}^{Opt\text{-}MB}$	$\widehat{d}_1^{GS}$	$\widehat{d}_1^{Opt\text{-}GS}$	$\widehat{d}^{PFSB}$	$\widehat{d}^{MLE}$	$\widehat{d}^{PW}$
-0.9	-0.25	96	0.6233	0.6345	0.6275	0.6177	0.6112	0.6020	0.6284	0.6385	0.6086	0.8247	**0.3671**	*0.3729*
		576	0.4794	0.4812	0.4723	0.4662	0.4553	0.4492	0.4671	0.4885	0.4686	0.4977	**0.1352**	*0.1945*
	0	96	0.7361	0.8081	0.7972	0.7875	0.7726	0.7642	0.7815	0.8413	0.7214	0.8510	**0.6705**	*0.6938*
		576	0.5687	0.5919	0.5822	0.5719	0.5641	*0.5527*	0.5637	0.5838	0.5639	0.5942	**0.5426**	0.5941
	0.25	96	0.7996	0.8096	0.7918	0.7872	0.7716	*0.7615*	0.7715	0.8268	0.7869	0.8430	**0.7592**	0.8081
		576	0.5951	0.6193	0.6022	0.5976	0.5843	*0.5693*	0.5826	0.6219	0.6019	0.6590	**0.5513**	0.5993
	0.45	96	0.8219	0.8410	0.8325	0.8224	0.8135	*0.8064*	0.8231	0.8590	0.8190	0.8327	**0.7883**	0.8184
		576	0.5950	0.6066	0.5953	0.5871	0.5763	*0.5642*	0.5783	0.6298	0.6198	0.6487	**0.5609**	0.6063
-0.4	-0.25	96	0.2376	0.2253	0.2218	0.2198	0.2133	0.2102	0.2401	0.2488	0.2255	0.3103	**0.1682**	*0.2094*
		576	0.1037	0.0923	0.0895	0.0745	0.0672	*0.0652*	0.1052	0.1098	0.1004	0.1254	**0.0526**	0.1027
	0	96	0.2497	0.2385	0.2278	0.2142	0.2136	*0.2015*	0.2514	0.2559	0.2512	0.2883	**0.1644**	0.2043
		576	0.1070	0.0936	0.0979	0.0819	0.0887	*0.0778*	0.1078	0.1105	0.0845	0.1215	**0.0511**	0.1011
	0.25	96	0.2527	0.2451	0.2425	0.2379	0.2343	0.2335	0.2535	0.2560	0.2495	0.2782	**0.1679**	*0.2087*
		576	0.1068	0.0987	0.1052	0.1057	0.0964	*0.0867*	0.1067	0.1103	0.0934	0.1199	**0.0518**	0.1128
	0.45	96	0.2496	0.2524	0.2459	0.2476	0.2493	0.2495	0.2495	0.2518	0.2441	0.2725	**0.1682**	*0.2093*
		576	0.1047	0.0928	0.0900	0.0855	0.0830	*0.0740*	0.1040	0.1098	0.0991	0.1188	**0.0566**	0.1066
0.4	-0.25	96	0.1982	0.1894	0.1875	0.1825	0.1793	0.1687	0.1987	0.2212	0.2153	0.2809	**0.1083**	*0.1422*
		576	0.0932	0.0858	0.0988	0.0947	0.0935	0.0933	0.0933	0.1078	0.0812	0.1268	**0.0594**	*0.0739*
	0	96	0.1944	0.1826	0.1815	0.1729	0.1666	0.1654	0.1955	0.2203	0.2146	0.2701	**0.1042**	*0.1492*
		576	0.0919	0.0890	0.0893	0.0850	0.0829	0.0824	0.0924	0.1072	0.0930	0.1243	**0.0518**	*0.0725*
	0.25	96	0.1947	0.1945	0.1918	0.1878	0.1780	*0.1762*	0.1962	0.2213	0.2048	0.2663	**0.1015**	0.1786
		576	0.0925	0.0942	0.1079	0.0983	0.0942	0.0832	0.0932	0.1077	0.0924	0.1238	**0.0539**	*0.0731*
	0.45	96	0.1964	0.1769	0.1649	0.1544	*0.1407*	0.1483	0.1984	0.2223	0.2175	0.2643	**0.1028**	0.1818
		576	0.0943	0.0902	0.0831	0.0846	0.0772	*0.0756*	0.0955	0.1090	0.0939	0.1229	**0.0541**	0.0857
0.9	-0.25	96	0.0886	0.0983	0.0944	0.0907	0.0883	*0.0864*	0.0938	0.1105	0.1073	0.1253	**0.0543**	0.0912
		576	*0.0344*	0.0518	0.0504	0.0493	0.0426	0.0376	0.0467	0.0561	0.0538	0.0589	**0.0215**	0.0421
	0	96	0.0863	0.1086	0.1011	0.0977	0.0945	0.0912	0.0975	0.1209	0.1158	0.1288	**0.0607**	*0.0832*
		576	*0.0312*	0.0572	0.0542	0.0519	0.0482	0.0460	0.0493	0.0674	0.0625	0.0729	**0.0202**	0.0404
	0.25	96	*0.0865*	0.1113	0.1086	0.1012	0.0974	0.0928	0.0972	0.1287	0.1176	0.1286	**0.0653**	0.0926
		576	*0.0304*	0.0574	0.0548	0.0519	0.0496	0.0433	0.0487	0.0692	0.0614	0.0706	**0.0195**	0.0451
	0.45	96	*0.0885*	0.1268	0.1206	0.1158	0.1107	0.1073	0.1069	0.1181	0.1093	0.1197	**0.0654**	0.0933
		576	*0.0378*	0.0592	0.0541	0.0517	0.0482	0.0459	0.0528	0.0647	0.0580	0.0695	**0.0122**	0.0424

Table 6: Bias estimates of the unadjusted LPR estimator, the feasible jackknife estimator based on 2,3,4,6,8 non-overlapping (NO) sub-samples, the feasible jackknife estimator based on 2 moving block (MB) sub-samples, both versions of the GS estimator, the pre-filtered sieve bootstrap estimator, the maximum likelihood estimator (MLE) and the pre-whitened (PW) estimator, for the DGP: ARFIMA$(1, d_0, 0)$ with Gaussian innovations. The feasible jackknife estimates are evaluated using the iterative procedure described in Section 5.2.2. The estimates are obtained by setting $\alpha = 0.65$ and assuming the model is correctly specified. The lowest values are **bold-faced** and the second lowest values are *italicized*.

ϕ_0	d_0	n	$\widehat{d}_n$	$\widehat{d}_{J,2}^{NO}$	$\widehat{d}_{J,3}^{NO}$	$\widehat{d}_{J,4}^{NO}$	$\widehat{d}_{J,6}^{NO}$	$\widehat{d}_{J,8}^{NO}$	$\widehat{d}_{J,2}^{MB}$	$\widehat{d}_1^{GS}$	$\widehat{d}^{PFSB}$	$\widehat{d}^{MLE}$	$\widehat{d}^{PW}$
-0.9	-0.25	96	0.8145	0.8274	0.8301	0.8362	0.8427	0.8488	0.8322	0.8002	*0.7908*	**0.7331**	0.8022
		576	0.5945	*0.5743*	0.5792	0.5815	0.5880	0.5903	0.5853	**0.5724**	0.5898	0.5918	0.6120
	0	96	0.8053	0.8158	0.8226	0.8279	0.8104	0.8272	0.8184	0.8015	*0.7955*	**0.7215**	0.7965
		576	0.5912	0.6263	0.6342	0.6393	0.6428	0.6528	0.6335	**0.5761**	*0.5888*	0.5953	0.6097
	0.25	96	0.7752	**0.7664**	0.7738	0.7702	0.7816	0.7953	0.7729	*0.7673*	0.7685	0.7722	0.8158
		576	0.5883	0.5737	0.5820	0.5896	0.5942	0.6054	0.5728	0.5716	**0.5638**	*0.5669*	0.6136
	0.45	96	0.7006	*0.6930*	0.7056	0.7174	0.7227	0.7287	0.7036	0.6946	**0.6705**	0.7378	0.8079
		576	0.5748	**0.5428**	0.5549	0.5591	0.5608	0.5691	0.5584	0.5659	*0.5451*	0.5463	0.6121
-0.4	-0.25	96	0.1756	0.1418	0.1465	0.1508	0.1577	0.1626	0.1427	**0.1367**	0.1435	*0.1375*	0.1754
		576	0.0607	**0.0189**	*0.0254*	0.0338	0.0590	0.0616	0.0392	0.0304	0.0286	0.0412	0.0512
	0	96	0.1653	**0.1143**	*0.1186*	0.1201	0.1295	0.1328	0.1276	0.1304	0.1353	0.1325	0.1706
		576	0.0560	0.0288	0.0313	0.0370	0.0422	0.0465	0.0321	*0.0264*	**0.0249**	0.0328	0.0471
	0.25	96	0.1629	**0.1116**	*0.1174*	0.1223	0.1320	0.1387	0.1122	0.1329	0.1294	0.1382	0.1743
		576	0.0571	0.0284	0.0291	0.0326	0.0379	0.0429	0.0358	0.0289	**0.0251**	*0.0253*	0.0431
	0.45	96	0.1653	**0.1073**	*0.1132*	0.1245	0.1302	0.1444	0.1520	0.1400	0.1277	0.1352	0.1615
		576	0.0625	0.0289	0.0315	0.0384	0.0438	0.0557	0.0529	0.0359	**0.0261**	*0.0269*	0.0499
0.4	-0.25	96	-0.0363	-0.0174	-0.0236	-0.0288	-0.0317	-0.0392	-0.0284	**-0.0047**	-0.0147	*-0.0115*	-0.0402
		576	-0.0056	-0.0086	-0.0113	-0.0157	-0.0187	-0.0216	-0.0122	0.0056	**-0.0004**	*-0.0026*	-0.0153
	0	96	-0.0534	-0.0178	-0.0210	-0.0256	-0.0303	-0.0382	-0.0234	**-0.0089**	-0.0175	*-0.0098*	-0.0340
		576	-0.0125	-0.0085	-0.0096	-0.0114	-0.0155	-0.0184	-0.0032	**-0.0008**	-0.0040	*-0.0010*	-0.0086
	0.25	96	-0.0559	-0.0116	-0.0150	-0.0182	-0.0238	-0.0296	-0.0221	**-0.0068**	-0.0153	*-0.0112*	-0.0357
		576	-0.0115	-0.0082	-0.0066	-0.0045	-0.0064	-0.0094	-0.0056	**0.0017**	-0.0027	*-0.0026*	-0.0088
	0.45	96	-0.0501	*-0.0095*	-0.0110	-0.0182	-0.0249	-0.0319	-0.0197	**0.0032**	-0.0111	-0.0152	-0.0396
		576	-0.0058	-0.0073	-0.0025	-0.0059	-0.0043	-0.0088	-0.0174	0.0089	**0.0004**	*-0.0018*	-0.0092
0.9	-0.25	96	-0.0291	*-0.0100*	-0.0127	-0.0152	-0.0185	-0.0217	-0.0121	-0.0175	-0.0162	**-0.0098**	-0.0206
		576	-0.0058	*-0.0017*	-0.0020	-0.0036	-0.0054	-0.0074	-0.0035	-0.0034	-0.0023	**-0.0005**	-0.0062
	0	96	-0.0170	-0.0096	-0.0119	-0.0148	-0.0173	-0.0195	-0.0128	-0.0140	**-0.0082**	*-0.0093*	-0.0194
		576	-0.0029	-0.0010	-0.0011	-0.0026	-0.0044	-0.0036	-0.0012	-0.0009	**-0.0005**	*-0.0007*	-0.0051
	0.25	96	-0.0249	**-0.0095**	-0.0135	-0.0182	-0.0215	-0.0248	-0.0126	-0.0162	-0.0117	*-0.0098*	-0.0267
		576	-0.0044	*-0.0018*	-0.0032	-0.0041	-0.0059	-0.0073	-0.0045	-0.0032	-0.0020	**-0.0004**	-0.0044
	0.45	96	-0.0241	**-0.0089**	-0.0114	-0.0176	-0.0199	-0.0249	-0.0179	-0.0175	-0.0126	*-0.0090*	-0.0192
		576	-0.0077	-0.0019	-0.0034	-0.0039	-0.0046	-0.0063	-0.0017	-0.0038	*-0.0018*	**-0.0004**	-0.0040

Table 7: RMSE estimates of the unadjusted LPR estimator, the feasible jackknife estimator based on 2,3,4,6,8 non-overlapping (NO) sub-samples, the feasible jackknife estimator based on 2 moving block (MB) sub-samples, both versions of the GS estimator, the pre-filtered sieve bootstrap estimator, the maximum likelihood estimator (MLE) and the pre-whitened (PW) estimator, for the DGP: ARFIMA$(1, d_0, 0)$ with Gaussian innovations. The feasible jackknife estimates are evaluated using the iterative procedure described in Section 5.2.2. The estimates are obtained by setting $\alpha = 0.65$ and assuming the model is correctly specified. The lowest values are **bold-faced** and the second lowest values are *italicized*.

ϕ_0	d_0	n	$\widehat{d}_n$	$\widehat{d}_{J,2}^{NO}$	$\widehat{d}_{J,3}^{NO}$	$\widehat{d}_{J,4}^{NO}$	$\widehat{d}_{J,6}^{NO}$	$\widehat{d}_{J,8}^{NO}$	$\widehat{d}_{J,2}^{MB}$	$\widehat{d}_1^{GS}$	$\widehat{d}^{PFSB}$	$\widehat{d}^{MLE}$	$\widehat{d}^{PW}$
-0.9	-0.25	96	*1.0359*	1.2162	1.2053	1.1907	1.1853	1.1814	1.2193	1.3386	1.2885	**0.9538**	1.0562
		576	0.7398	0.7688	0.7621	0.7586	0.7549	0.7451	0.7618	0.7371	*0.7359*	**0.6315**	0.7365
	0	96	1.1148	1.1491	1.1400	1.1365	1.1334	1.1240	1.1391	1.1819	1.2167	**0.9243**	*0.9847*
		576	0.8288	0.8418	0.8397	0.8322	0.8282	0.8239	0.8306	0.8173	0.8053	**0.6221**	*0.6648*
	0.25	96	1.1618	1.1835	1.1773	1.1666	1.1537	1.1428	1.1588	1.1484	1.2299	**0.9428**	*1.0275*
		576	0.9175	0.9409	0.9348	0.9275	0.9212	0.9187	0.9334	1.1171	1.1130	**0.6385**	*0.6611*
	0.45	96	1.1286	1.2150	1.2061	1.1982	1.1933	1.1869	1.2186	1.4331	1.5385	**0.9382**	*1.0335*
		576	0.9708	0.9842	0.9711	0.9672	0.9627	0.9574	0.9775	1.1124	1.1647	**0.6415**	*0.7069*
-0.4	-0.25	96	*0.2568*	0.2841	0.2726	0.2699	0.2606	0.2515	0.2635	0.2594	0.3028	**0.1863**	0.2671
		576	*0.1098*	0.1249	0.1168	0.1134	0.1121	0.1276	0.1149	0.1118	0.1272	**0.0946**	0.1339
	0	96	0.2498	0.2772	0.2724	0.2685	0.2576	*0.2418*	0.2643	0.2560	0.2930	**0.1792**	0.2496
		576	0.1069	0.1278	0.1218	0.1106	0.1073	*0.1005*	0.1055	0.1104	0.1247	**0.0867**	0.1348
	0.25	96	0.2490	0.2835	0.2782	0.2737	0.2688	0.2630	0.3108	0.2580	0.2879	**0.1814**	*0.2473*
		576	0.1079	0.1374	0.1326	0.1248	0.1160	*0.1053*	0.1164	0.1115	0.1239	**0.0992**	0.1340
	0.45	96	0.2506	0.2833	0.2761	0.2619	0.2598	0.2541	0.2759	0.2616	0.2506	**0.1836**	*0.2497*
		576	*0.1115*	0.1428	0.1411	0.1337	0.1276	0.1128	0.1221	0.1143	0.1230	**0.0934**	0.1306
0.4	-0.25	96	*0.1917*	0.2350	0.2335	0.2278	0.2210	0.2172	0.2266	0.2212	0.2717	**0.1244**	0.1984
		576	0.0919	0.1229	0.1189	0.1144	0.1075	0.1035	0.1020	0.1081	0.1198	**0.0531**	*0.0909*
	0	96	0.1946	0.2295	0.2251	0.2177	0.2114	0.2001	0.2163	0.2203	0.2546	**0.1194**	*0.1939*
		576	*0.0920*	0.1246	0.1208	0.1145	0.1185	0.1099	0.1176	0.1073	0.1166	**0.0617**	0.0946
	0.25	96	*0.1960*	0.2281	0.2219	0.2163	0.2267	0.2296	0.2225	0.2209	0.2482	**0.1273**	0.2088
		576	*0.0922*	0.1168	0.1113	0.1087	0.1055	0.1019	0.1150	0.1076	0.1158	**0.0566**	0.0999
	0.45	96	*0.1955*	0.2379	0.2318	0.2206	0.2284	0.2178	0.2174	0.2218	0.2453	**0.1282**	0.2094
		576	0.0926	0.1241	0.1241	0.1179	0.1055	0.1013	0.1084	0.1089	0.1149	**0.0476**	*0.0913*
0.9	-0.25	96	*0.1115*	0.1385	0.1306	0.1282	0.1243	0.1210	0.1284	0.1365	0.1266	**0.0712**	0.1160
		576	0.0624	0.0687	0.0660	0.0616	0.0579	*0.0548*	0.0599	0.0708	0.0600	**0.0369**	0.0649
	0	96	*0.1010*	0.1162	0.1123	0.1105	0.1088	0.1023	0.1187	0.1121	0.1093	**0.0681**	0.1097
		576	0.0602	0.0629	0.0609	0.0549	*0.0533*	0.0521	0.0577	0.0698	0.0705	**0.0344**	0.0539
	0.25	96	*0.1114*	0.1324	0.1318	0.1268	0.1222	0.1198	0.1229	0.1328	0.1282	**0.0771**	0.1142
		576	*0.0518*	0.0695	0.0641	0.0616	0.0589	0.0552	0.0572	0.0626	0.0581	**0.0322**	0.0530
	0.45	96	0.1053	0.1284	0.1307	0.1284	0.1229	0.1216	0.1179	0.1253	0.1215	**0.0725**	*0.1031*
		576	*0.0526*	0.0681	0.0635	0.0610	0.0595	0.0549	0.0581	0.0769	0.0684	**0.0349**	0.0528

Table 8: Bias estimates of the unadjusted LPR estimator, the feasible jackknife estimator based on 2,3,4,6,8 non-overlapping (NO) sub-samples, the feasible jackknife estimator based on 2 moving block (MB) sub-samples, both versions of the GS estimator, the pre-filtered sieve bootstrap estimator, the maximum likelihood estimator (MLE) and the pre-whitened (PW) estimator, for the DGP: ARFIMA$(0, d_0, 1)$ with Gaussian innovations. The feasible jackknife estimates are evaluated using the iterative procedure described in Section 5.2.2. The estimates are obtained by setting $\alpha = 0.65$ and assuming the model is correctly specified. The lowest values are **bold-faced** and the second lowest values are *italicized*.

θ_0	d_0	n	$\widehat{d}_n$	$\widehat{d}_{J,2}^{NO}$	$\widehat{d}_{J,3}^{NO}$	$\widehat{d}_{J,4}^{NO}$	$\widehat{d}_{J,6}^{NO}$	$\widehat{d}_{J,8}^{NO}$	$\widehat{d}_{J,2}^{MB}$	$\widehat{d}_1^{GS}$	$\widehat{d}^{PFSB}$	$\widehat{d}^{MLE}$	$\widehat{d}^{PW}$
-0.9	-0.25	96	-0.5671	-0.5487	-0.5519	-0.5586	-0.5627	-0.5738	-0.5862	*-0.5450*	-0.5466	**-0.5372**	-0.6278
		576	-0.4527	**-0.4259**	-0.4364	-0.4429	-0.4581	-0.4694	-0.4575	-0.4385	*-0.4285*	-0.4294	-0.4663
	0	96	-0.7042	**-0.6510**	-0.6627	-0.6681	-0.6729	-0.6786	-0.6692	*-0.6575*	-0.6664	-0.6586	-0.6638
		576	-0.5594	*-0.5237*	-0.5344	-0.5492	-0.5542	-0.5649	-0.5437	-0.5256	-0.5375	**-0.5131**	-0.5445
	0.25	96	-0.7763	*-0.7482*	-0.7535	-0.7686	-0.7715	-0.7899	-0.7559	-0.7524	-0.7661	**-0.7375**	-0.7784
		576	-0.5880	*-0.5462*	-0.5561	-0.5548	-0.5726	-0.5772	-0.5628	-0.5473	-0.5621	**-0.5367**	-0.5647
	0.45	96	-0.8004	**-0.7515**	*-0.7587*	-0.7653	-0.7749	-0.7841	-0.7736	-0.7600	-0.7854	-0.7615	-0.7935
		576	-0.5880	*-0.5349*	-0.5395	-0.5438	-0.5485	-0.5509	-0.5394	-0.5351	-0.5527	**-0.5345**	-0.5732
-0.4	-0.25	96	-0.1437	**-0.1116**	-0.1186	-0.1274	-0.1348	-0.1486	-0.1357	*-0.1120*	-0.1240	-0.1153	-0.1465
		576	-0.0476	-0.0382	-0.0405	-0.0458	-0.0495	-0.0517	-0.0393	**-0.0187**	-0.0271	*-0.0243*	-0.0485
	0	96	-0.1653	**-0.1234**	-0.1282	-0.1340	-0.1494	-0.1538	-0.1488	-0.1305	*-0.1248*	-0.1286	-0.1436
		576	-0.0560	-0.0290	-0.0379	-0.0454	-0.0510	-0.0686	-0.0395	*-0.0265*	-0.0307	**-0.0239**	-0.0588
	0.25	96	-0.1692	*-0.1275*	-0.1304	-0.1399	-0.1438	-0.1544	-0.1317	-0.1297	**-0.1200**	-0.1362	-0.1527
		576	-0.0552	**-0.0184**	-0.0397	-0.0472	-0.0565	-0.0605	-0.0492	*-0.0243*	-0.0287	-0.0275	-0.0573
	0.45	96	-0.1630	**-0.0837**	-0.1399	-0.1623	-0.1708	-0.1769	-0.1433	-0.1190	*-0.1118*	-0.1385	-0.1586
		576	-0.0493	**-0.0160**	-0.0184	-0.0398	-0.0416	-0.0483	-0.0305	*-0.0169*	-0.0244	-0.0283	-0.0526
0.4	-0.25	96	0.0637	**0.0084**	0.0490	0.0615	0.0684	0.0709	0.0534	*0.0154*	0.0651	0.0384	0.0494
		576	0.0175	*0.0051*	0.0117	0.0180	0.0233	0.0168	0.0095	**0.0049**	0.0132	0.0083	0.0212
	0	96	0.0525	*0.0313*	0.0396	0.0424	0.0356	0.0397	0.0388	**0.0081**	0.0603	0.0317	0.0375
		576	0.0125	0.0135	0.0271	0.0315	0.0340	0.0284	0.0094	**0.0006**	0.0100	*0.0076*	0.0199
	0.25	96	0.0504	*0.0239*	0.0418	0.0568	0.0599	0.0633	0.0428	**0.0110**	0.0574	0.0364	0.0475
		576	0.0136	*0.0065*	0.0083	0.0105	0.0173	0.0209	0.0073	**0.0031**	0.0108	0.0088	0.0154
	0.45	96	0.0549	*0.0245*	0.0386	0.0495	0.0626	0.0600	0.0455	**0.0204**	0.0570	0.0347	0.0426
		576	0.0192	**0.0063**	0.0095	*0.0074*	0.0099	0.0129	0.0103	0.0103	0.0132	0.0091	0.0182
0.9	-0.25	96	0.0359	*0.0097*	0.0119	0.0176	0.0237	0.0262	0.0254	0.0109	**0.0085**	0.0153	0.0246
		576	0.0065	**0.0012**	0.0018	0.0028	0.0035	0.0056	0.0042	0.0020	*0.0014*	0.0027	0.0181
	0	96	0.0347	*0.0084*	0.0092	0.0095	0.0101	0.0140	0.0119	0.0091	**0.0076**	0.0089	0.0279
		576	0.0052	**0.0010**	0.0011	0.0018	0.0025	0.0045	0.0041	0.0015	*0.0010*	0.0034	0.0126
	0.25	96	0.0293	*0.0074*	0.0077	0.0092	0.0120	0.0136	0.0117	0.0130	**0.0073**	0.0117	0.0281
		576	0.0083	**0.0013**	0.0028	0.0029	0.0033	0.0051	0.0026	0.0057	*0.0019*	0.0059	0.0122
	0.45	96	0.0235	**0.0075**	0.0080	0.0095	0.0110	0.0173	0.0117	0.0132	*0.0075*	0.0121	0.0260
		576	0.0195	**0.0039**	0.0066	0.0077	0.0119	0.0118	0.0091	0.0071	*0.0042*	0.0060	0.0130

Table 9: RMSE estimates of the unadjusted LPR estimator, the feasible jackknife estimator based on 2,3,4,6,8 non-overlapping (NO) sub-samples, the feasible jackknife estimator based on 2 moving block (MB) sub-samples, both versions of the GS estimator, the pre-filtered sieve bootstrap estimator, the maximum likelihood estimator (MLE) and the pre-whitened (PW) estimator, for the DGP: ARFIMA$(0, d_0, 1)$ with Gaussian innovations. The feasible jackknife estimates are evaluated using the iterative procedure described in Section 5.2.2. The estimates are obtained by setting $\alpha = 0.65$ and assuming the model is correctly specified. The lowest values are **bold-faced** and the second lowest values are *italicized*.

θ_0	d_0	n	$\widehat{d}_n$	$\widehat{d}^{NO}_{J,2}$	$\widehat{d}^{NO}_{J,3}$	$\widehat{d}^{NO}_{J,4}$	$\widehat{d}^{NO}_{J,6}$	$\widehat{d}^{NO}_{J,8}$	$\widehat{d}^{MB}_{J,2}$	$\widehat{d}^{GS}_{1}$	$\widehat{d}^{PFSB}$	$\widehat{d}^{MLE}$	$\widehat{d}^{PW}$
-0.9	-0.25	96	*0.6233*	0.6561	0.6463	0.6405	0.6348	0.6319	0.6653	0.6385	0.8247	**0.5982**	0.6663
		576	0.4794	0.4980	0.4927	0.4854	0.4771	*0.4727*	0.4785	0.4888	0.4977	**0.4406**	0.4858
	0	96	*0.7361*	0.8327	0.8291	0.8247	0.8189	0.8006	0.8114	0.8413	0.8510	**0.7234**	0.7604
		576	*0.5687*	0.6371	0.6034	0.6152	0.6038	0.5972	0.6241	0.5838	0.5942	**0.5621**	0.6329
	0.25	96	0.7996	0.8238	0.8186	0.8013	0.7926	*0.7884*	0.8108	0.8268	0.8430	**0.7429**	0.8215
		576	0.5951	0.6339	0.6257	0.6108	0.6075	*0.5922*	0.6249	0.6219	0.6590	**0.5513**	0.6359
	0.45	96	0.8219	0.8562	0.8414	0.8393	0.8242	*0.8107*	0.8233	0.8590	0.8327	**0.8107**	0.8353
		576	*0.5950*	0.6384	0.6279	0.6211	0.6184	0.6124	0.6589	0.6298	0.6487	**0.5918**	0.6337
-0.4	-0.25	96	0.2376	0.2517	0.2441	0.2384	0.2300	*0.2283*	0.2566	0.2488	0.3103	**0.2173**	0.2923
		576	*0.1037*	0.1352	0.1239	0.1192	0.1116	0.1085	0.1232	0.1098	0.1254	**0.0954**	0.1326
	0	96	0.2497	0.2743	0.2662	0.2545	0.2449	*0.2406*	0.2457	0.2559	0.2883	**0.2284**	0.2873
		576	0.1070	0.1366	0.1245	0.1184	0.1105	*0.1044*	0.1193	0.1105	0.1215	**0.1020**	0.1478
	0.25	96	0.2527	0.2719	0.2636	0.2591	0.2530	*0.2418*	0.2495	0.2648	0.2782	**0.2305**	0.2995
		576	*0.1068*	0.1384	0.1315	0.1251	0.1144	0.1084	0.1267	0.1103	0.1199	**0.1032**	0.1526
	0.45	96	*0.2496*	0.2737	0.2608	0.2549	0.2521	0.2512	0.2528	0.2518	0.2725	**0.2311**	0.2880
		576	*0.1047*	0.1392	0.1384	0.1300	0.1243	0.1154	0.1276	0.1098	0.1188	**0.1044**	0.1539
0.4	-0.25	96	*0.1982*	0.2348	0.2217	0.2106	0.2058	0.2007	0.2245	0.2212	0.2809	**0.1729**	0.2035
		576	*0.0932*	0.1079	0.1155	0.1249	0.1163	0.1096	0.0972	0.1078	0.1268	**0.0883**	0.1185
	0	96	*0.1944*	0.2315	0.2242	0.2215	0.2194	0.2072	0.2159	0.2203	0.2701	**0.1637**	0.1927
		576	*0.0919*	0.1232	0.1119	0.1076	0.1026	0.0944	0.1036	0.1072	0.1243	**0.0906**	0.1053
	0.25	96	0.1947	0.2224	0.2153	0.2018	0.1982	*0.1902*	0.2247	0.2213	0.2663	**0.1625**	0.1901
		576	*0.0925*	0.1105	0.1151	0.1172	0.1069	0.1010	0.1108	0.1077	0.1238	**0.0853**	0.1136
	0.45	96	0.1964	0.2247	0.2172	0.2033	0.1946	*0.1916*	0.2265	0.2223	0.2643	**0.1639**	0.2084
		576	*0.0943*	0.1221	0.1134	0.1076	0.1016	0.0992	0.1157	0.1090	0.1229	**0.0927**	0.1120
0.9	-0.25	96	*0.0886*	0.1213	0.1189	0.1083	0.1035	0.0953	0.1166	0.1105	0.1253	**0.0834**	0.1117
		576	*0.0344*	0.0695	0.0637	0.0559	0.0521	0.0486	0.0578	0.0561	0.0589	**0.0315**	0.0538
	0	96	*0.0863*	0.1136	0.1120	0.1084	0.1043	0.0997	0.1045	0.1209	0.1288	**0.0807**	0.1013
		576	*0.0312*	0.0613	0.0595	0.0546	0.0504	0.0488	0.0491	0.0674	0.0729	**0.0299**	0.0526
	0.25	96	*0.0865*	0.1224	0.1210	0.1171	0.1123	0.1161	0.1140	0.1287	0.1286	**0.0818**	0.1157
		576	*0.0304*	0.0628	0.0654	0.0613	0.0592	0.0568	0.0606	0.0692	0.0706	**0.0253**	0.0530
	0.45	96	*0.0885*	0.1262	0.1241	0.1109	0.1185	0.1136	0.1175	0.1181	0.1197	**0.0824**	0.1128
		576	*0.0378*	0.0630	0.0581	0.0542	0.0473	0.0495	0.0536	0.0647	0.0695	**0.0271**	0.0563

Table 10: Bias estimates of the unadjusted LPR estimator, the feasible GS estimator and the pre-filtered sieve bootstrap estimator, for the DGP: ARFIMA$(1, d_0, 0)$ with Gaussian innovations. The bias estimates of the feasible jackknife estimator based on 2,3,4,6,8 non-overlapping (NO) sub-samples, the feasible jackknife estimator based on 2 moving block (MB) sub-samples, the maximum likelihood estimator (MLE) and the pre-whitened (PW) estimator are obtained under the misspecified model: ARFIMA$(0, d, 0)$ using the approach described in Section 5.3. The estimates are obtained by setting $\alpha = 0.65$. The lowest values are **bold-faced** and the second lowest values are *italicized*.

ϕ_0	d_0	n	$\widehat{d}_n$	$\widehat{d}_{J,2}^{NO}$	$\widehat{d}_{J,3}^{NO}$	$\widehat{d}_{J,4}^{NO}$	$\widehat{d}_{J,6}^{NO}$	$\widehat{d}_{J,8}^{NO}$	$\widehat{d}_{J,2}^{MB}$	$\widehat{d}_1^{GS}$	$\widehat{d}^{PFSB}$	$\widehat{d}^{MLE}$	$\widehat{d}^{PW}$
-0.9	-0.25	96	0.8145	0.8566	0.8595	0.8643	0.8681	0.8734	0.8540	*0.8002*	**0.7908**	0.8195	0.8318
		576	0.5945	0.6389	0.6422	0.6490	0.6553	0.6608	0.6321	**0.5724**	*0.5898*	0.6104	0.6444
	0.25	96	0.7752	0.8075	0.8134	0.8217	0.8246	0.8392	0.8273	**0.7673**	*0.7685*	0.7945	0.8258
		576	0.5883	0.6464	0.6429	0.6326	0.6255	0.6233	0.6356	*0.5716*	**0.5638**	0.5943	0.6305
	0.45	96	0.7006	0.7426	0.7482	0.7538	0.7599	0.7646	0.7520	*0.6946*	**0.6705**	0.7895	0.8377
		576	0.5748	0.6105	0.6154	0.6203	0.6286	0.6397	0.6118	*0.5659*	**0.5451**	0.5986	0.6390
-0.4	-0.25	96	0.1756	0.1543	0.1626	0.1691	0.1753	0.1804	0.1629	**0.1367**	*0.1435*	0.1947	0.2021
		576	0.0607	0.0419	0.0484	0.0507	0.0541	0.0585	0.0426	*0.0304*	**0.0286**	0.0743	0.1035
	0.25	96	0.1629	0.1536	0.1588	0.1614	0.1669	0.1760	0.1512	*0.1329*	**0.1294**	0.1840	0.2064
		576	0.0571	0.0348	0.0376	0.0491	0.0527	0.0598	0.0506	*0.0289*	**0.0251**	0.0587	0.0996
	0.45	96	0.1653	0.1546	0.1592	0.1648	0.1688	0.1735	0.1648	*0.1400*	**0.1277**	0.1886	0.2145
		576	0.0625	0.0445	0.0498	0.0572	0.0645	0.0686	0.0749	*0.0359*	**0.0261**	0.0534	0.1031
0.4	-0.25	96	-0.0363	-0.0286	-0.0359	-0.0397	-0.0462	-0.0481	-0.0388	**-0.0047**	*-0.0147*	-0.0385	-0.0490
		576	-0.0056*	-0.0105	-0.0128	-0.0156	-0.0184	-0.0229	-0.0154	*-0.0056*	**-0.0004**	-0.0163	-0.0258
	0.25	96	-0.0559	-0.0267	-0.0293	-0.0319	-0.0350	-0.0424	-0.0372	**-0.0068**	*-0.0153*	-0.0372	-0.0593
		576	-0.0115	-0.0104	-0.0131	-0.0177	-0.0176	-0.0239	-0.0168	**0.0017**	*-0.0027*	-0.0095	-0.0269
	0.45	96	-0.0501	-0.0279	-0.0249	-0.0325	-0.0381	-0.0458	-0.0294	**0.0032**	*-0.0111*	-0.0314	-0.0627
		576	*-0.0058*	-0.0115	-0.0157	-0.0186	-0.0195	-0.0210	-0.0153	0.0089	**0.0004**	-0.0088	-0.0251
0.9	-0.25	96	-0.0291	**-0.0123**	*-0.0129*	-0.0145	-0.0193	-0.0224	-0.0148	-0.0175	-0.0162	-0.0156	-0.0339
		576	-0.0058	**-0.0020**	-0.0028	-0.0041	-0.0066	-0.0075	-0.0049	-0.0034	*-0.0023*	-0.0037	-0.0142
	0.25	96	-0.0249	**-0.0106**	-0.0132	-0.0220	-0.0241	-0.0269	-0.0153	-0.0162	*-0.0117*	-0.0163	-0.0326
		576	-0.0044	*-0.0021*	-0.0047	-0.0055	-0.0073	-0.0084	-0.0062	-0.0032	**-0.0020**	-0.0034	-0.0174
	0.45	96	-0.0241	**-0.0095**	-0.0126*	-0.0217	-0.0229	-0.0241	-0.0175	-0.0175	*-0.0126*	-0.0145	-0.0366
		576	-0.0077	*-0.0026*	-0.0035	-0.0042	-0.0048	-0.0065	-0.0029	-0.0038	**-0.0018**	-0.0048	-0.0134

Table 11: RMSE estimates of the unadjusted LPR estimator, the feasible GS estimator and the pre-filtered sieve bootstrap estimator, for the DGP: ARFIMA$(1, d_0, 0)$ with Gaussian innovations. The bias estimates of the feasible jackknife estimator based on 2,3,4,6,8 non-overlapping (NO) sub-samples, the feasible jackknife estimator based on 2 moving block (MB) sub-samples, the maximum likelihood estimator (MLE) and the pre-whitened (PW) estimator are obtained under the misspecified model: ARFIMA$(0, d, 0)$ using the approach described in Section 5.3. The estimates are obtained by setting $\alpha = 0.65$. The lowest values are **bold-faced** and the second lowest values are *italicized*.

ϕ_0	d_0	n	$\widehat{d}_n$	$\widehat{d}_{J,2}^{NO}$	$\widehat{d}_{J,3}^{NO}$	$\widehat{d}_{J,4}^{NO}$	$\widehat{d}_{J,6}^{NO}$	$\widehat{d}_{J,8}^{NO}$	$\widehat{d}_{J,2}^{MB}$	$\widehat{d}_1^{GS}$	$\widehat{d}^{PFSB}$	$\widehat{d}^{MLE}$	$\widehat{d}^{PW}$
-0.9	-0.25	96	**1.0359**	1.2624	1.2506	1.2406	1.2416	1.2384	1.2395	1.3386	1.2885	*1.2247*	1.2372
		576	0.7398	0.7851	0.7827	0.7795	0.7628	0.7695	0.7726	*0.7371*	**0.7359**	0.8030	0.8428
	0.25	96	*1.1618*	1.2137	1.2154	1.2042	1.2069	1.1958	1.2073	**1.1484**	1.2299	1.2149	1.2226
		576	**0.9175**	0.9769	0.9732	0.9618	0.9594	*0.9537*	0.9678	1.1171	1.1130	0.9556	0.9638
	0.45	96	**1.1286**	1.2446	1.2384	1.2329	1.2280	*1.2186*	1.2349	1.4331	1.5385	1.2285	1.2385
		576	**0.9708**	1.0975	1.0299	1.0183	0.9716	*0.9824*	1.0198	1.1124	1.1647	0.9929	1.0064
-0.4	-0.25	96	**0.2568**	0.3063	0.3015	0.2940	0.2874	0.2711	0.2726	*0.2594*	0.3028	0.2736	0.2915
		576	**0.1098**	0.1408	0.1377	0.1305	0.1249	0.1153	0.1281	*0.1118*	0.1272	0.1242	0.1463
	0.25	96	**0.2490**	0.3034	0.2956	0.2874	0.2713	0.2624	0.3097	*0.2580*	0.2879	0.2794	0.2905
		576	**0.1079**	0.1410	0.1393	0.1318	0.1275	0.1248	0.1290	0.1375	*0.1239*	0.1329	0.1539
	0.45	96	**0.2506**	0.3087	0.3016	0.2971	0.2840	0.2737	0.2972	0.2616	*0.2506*	0.2840	0.3017
		576	**0.1115**	0.1672	0.1508	0.1438	0.1378	0.1225	0.1482	*0.1143*	0.1230	0.1385	0.1228
0.4	-0.25	96	**0.1917**	0.2441	0.2318	0.2473	0.2319	0.2301	0.2333	*0.2212*	0.2717	0.2425	0.2573
		576	**0.0919**	0.1323	0.1295	0.1206	0.1134	0.1199	0.1210	0.1253	*0.1198*	0.1347	0.1836
	0.25	96	**0.1960**	0.2359	0.2306	0.2239	0.2406	0.2378	0.2381	*0.2209*	0.2482	0.2385	0.2529
		576	**0.0922**	0.1330	0.1289	0.1211	0.1281	0.1166	0.1199	0.1366	0.1158	*0.1093*	0.1545
	0.45	96	**0.1955**	0.2416	0.2345	0.2394	0.2296	0.2244	0.2267	*0.2218*	0.2453	0.2236	0.2530
		576	**0.0926**	0.1305	0.1287	0.1226	0.1148	0.1106	0.1292	0.1089	0.1149	*0.1058*	0.1551
0.9	-0.25	96	**0.1115**	0.1419	0.1381	0.1305	0.1267	*0.1212*	0.1284	0.1365	0.1266	0.1315	0.1589
		576	*0.0624*	0.0795	0.0754	0.0713	0.0691	0.0627	0.0553	0.0708	**0.0600**	0.0703	0.0846
	0.25	96	**0.1114**	0.1482	0.1469	0.1337	0.1304	0.1289	0.1376	0.1328	*0.1282*	0.1293	0.1428
		576	**0.0518**	0.0800	0.0786	0.0711	0.0678	0.0624	0.0545	0.0626	0.0581	*0.0540*	0.0722
	0.45	96	**0.1053**	0.1495	0.1461	0.1398	0.1376	0.1268	0.1272	0.1253	0.1215	*0.1214*	0.1473
		576	**0.0526**	0.0714	0.0790	0.0774	0.0722	*0.0659*	0.0698	0.0769	0.0684	0.0648	0.0739

Table 12: Bias estimates of the unadjusted LPR estimator, the feasiblel GS estimator and the pre-filtered sieve bootstrap estimator, for the DGP: ARFIMA$(0, d_0, 1)$ with Gaussian innovations. The bias estimates of the feasible jackknife estimator based on 2,3,4,6,8 non-overlapping (NO) sub-samples, the feasible jackknife estimator based on 2 moving block (MB) sub-samples, the maximum likelihood estimator (MLE) and the pre-whitened (PW) estimator are obtained under the misspecified model: ARFIMA$(0, d, 0)$ using the approach described in Section 5.3. The estimates are obtained by setting $\alpha = 0.65$. The lowest values are **bold-faced** and the second lowest values are *italicized*.

θ_0	d_0	n	$\widehat{d}_n$	$\widehat{d}_{J,2}^{NO}$	$\widehat{d}_{J,3}^{NO}$	$\widehat{d}_{J,4}^{NO}$	$\widehat{d}_{J,6}^{NO}$	$\widehat{d}_{J,8}^{NO}$	$\widehat{d}_{J,2}^{MB}$	$\widehat{d}_1^{GS}$	$\widehat{d}^{PFSB}$	$\widehat{d}^{MLE}$	$\widehat{d}^{PW}$
-0.9	-0.25	96	-0.5671	**-0.5355**	*-0.5398*	-0.5437	-0.5541	-0.5626	-0.5584	-0.5450	-0.5466	-0.5692	-0.6361
		576	-0.4527	**-0.4216**	*-0.4273*	-0.4349	-0.4479	-0.4581	-0.4469	-0.4385	-0.4285	-0.4375	-0.4858
	0.25	96	-0.7763	**-0.7394**	*-0.7401*	-0.7475	-0.7523	-0.7670	-0.7416	-0.7524	-0.7661	-0.7718	-0.8149
		576	-0.5880	-0.5482	-0.5529	-0.5653	-0.5776	-0.5798	-0.5629	*-0.5473*	-0.5621	**-0.5426**	-0.5846
	0.45	96	-0.8004	**-0.7469**	*-0.7504*	-0.7638	-0.7769	-0.7716	-0.7553	-0.7600	-0.7854	-0.7713	-0.7942
		576	-0.5880	**-0.5033**	*-0.5059*	-0.5249	-0.5384	-0.5529	-0.5247	-0.5351	-0.5527	-0.5484	-0.5927
-0.4	-0.25	96	-0.1437	-0.1468	-0.1288	-0.1201	*-0.1175*	-0.1435	-0.1586	**-0.1120**	-0.1240	-0.1379	-0.1728
		576	-0.0476	-0.0296	-0.0379	-0.0334	-0.0562	-0.0560	-0.0392	**-0.0187**	*-0.0271*	-0.0584	-0.0957
	0.25	96	-0.1692	**-0.1172**	-0.1243	-0.1289	-0.1350	-0.1462	-0.1471	-0.1297	*-0.1200*	-0.1381	-0.1783
		576	-0.0552	**-0.0222**	-0.0318	-0.0425	-0.0531	-0.0573	-0.0469	*-0.0243*	-0.0287	-0.0558	-0.0990
	0.45	96	-0.1630	**-0.0716**	*-0.1076*	-0.1277	-0.1392	-0.1436	-0.1318	-0.1190	-0.1118	-0.1254	-0.1739
		576	-0.0493	**-0.0152**	-0.0183	-0.0309	-0.0473	-0.0414	-0.0205	*-0.0169*	-0.0244	-0.0509	-0.0948
0.4	-0.25	96	0.0637	*0.0139*	0.0226	0.0274	0.0395	0.0433	0.0381	**0.0154**	0.0651	0.0312	0.0661
		576	0.0175	*0.0066*	0.0091	0.0095	0.0122	0.0146	0.0070	**0.0049**	0.0132	0.0198	0.0283
	0.25	96	0.0504	*0.0141*	0.0364	0.0490	0.0526	0.0548	0.0301	**0.0110**	0.0574	0.0248	0.0529
		576	0.0136	0.0082	0.0087	0.0096	0.0121	0.0137	*0.0050*	**0.0031**	0.0108	0.0153	0.0238
	0.45	96	0.0549	**0.0189**	0.0321	0.0458	0.0529	0.0585	0.0362	*0.0204*	0.0570	0.0274	0.0668
		576	0.0192	**0.0053**	*0.0075*	0.0083	0.0104	0.0129	0.0083	0.0103	0.0132	0.0135	0.0269
0.9	-0.25	96	0.0359	**0.0075**	*0.0083*	0.0092	0.0116	0.0142	0.0117	0.0109	0.0085	0.0131	0.0354
		576	0.0065	0.0024	0.0039	0.0069	0.0086	0.0109	0.0078	*0.0020*	**0.0014**	0.0045	0.0151
	0.25	96	0.0293	*0.0096*	0.0138	0.0157	0.0199	0.0237	0.0195	0.0130	**0.0073**	0.0099	0.0322
		576	0.0083	*0.0022*	0.0036	0.0043	0.0055	0.0063	0.0029	0.0057	**0.0019**	0.0052	0.0137
	0.45	96	0.0235	*0.0109*	0.0128	0.0146	0.0165	0.0198	0.0149	0.0132	**0.0075**	0.0084	0.0353
		576	0.0195	**0.0030**	0.0075	0.0081	0.0076	0.0116	0.0084	0.0071	*0.0042*	0.0044	0.0120

Table 13: RMSE estimates of the unadjusted LPR estimator, the feasible GS estimator and the pre-filtered sieve bootstrap estimator, for the DGP: ARFIMA$(0, d_0, 1)$ with Gaussian innovations. The bias estimates of the feasible jackknife estimator based on 2,3,4,6,8 non-overlapping (NO) sub-samples, the feasible jackknife estimator based on 2 moving block (MB) sub-samples, the maximum likelihood estimator (MLE) and the pre-whitened (PW) estimator are obtained under the misspecified model: ARFIMA$(0, d, 0)$ using the approach described in Section 5.3. The estimates are obtained by setting $\alpha = 0.65$. The lowest values are **bold-faced** and the second lowest values are *italicized*.

θ_0	d_0	n	$\widehat{d}_n$	$\widehat{d}^{NO}_{J,2}$	$\widehat{d}^{NO}_{J,3}$	$\widehat{d}^{NO}_{J,4}$	$\widehat{d}^{NO}_{J,6}$	$\widehat{d}^{NO}_{J,8}$	$\widehat{d}^{MB}_{J,2}$	$\widehat{d}^{GS}_1$	$\widehat{d}^{PFSB}$	$\widehat{d}^{MLE}$	$\widehat{d}^{PW}$
-0.9	-0.25	96	**0.6233**	0.6882	0.6854	0.6825	0.6797	0.6774	0.6828	*0.6385*	0.8247	0.6619	0.7538
		576	**0.4794**	0.5335	0.5391	0.5447	0.5482	0.5505	0.5549	*0.4885*	0.4977	0.5043	0.5578
	0.25	96	**0.7996**	0.8695	0.8662	0.8533	0.8504	0.8442	0.8451	0.8268	0.8430	*0.8072*	0.8459
		576	**0.5951**	0.6749	0.6685	0.6612	0.6553	0.6490	0.6514	0.6219	0.6590	*0.6049*	0.6318
	0.45	96	**0.8219**	0.8806	0.8829	0.8781	0.8700	0.8651	0.8588	0.8590	*0.8327*	0.8381	0.8637
		576	**0.5950**	0.6474	0.6433	0.6419	0.6349	0.6355	0.6671	0.6298	0.6487	*0.6015*	0.6372
-0.4	-0.25	96	**0.2376**	0.2996	0.2963	0.2927	0.2903	0.2847	0.2834	*0.2488*	0.3103	0.2688	0.3142
		576	**0.1037**	0.1534	0.1588	0.1556	0.1485	0.1414	0.1549	*0.1098*	0.1254	0.1236	0.1540
	0.25	96	**0.2527**	0.2958	0.2964	0.2942	0.2846	0.2812	0.2833	*0.2560*	0.2782	0.2645	0.3162
		576	**0.1068**	0.1576	0.1529	0.1486	0.1455	0.1438	0.1482	*0.1103*	0.1199	0.1182	0.1573
	0.45	96	**0.2496**	0.2999	0.2927	0.2900	0.2867	0.2815	0.2795	*0.2518*	0.2725	0.2691	0.3029
		576	**0.1047**	0.1553	0.1518	0.1489	0.1421	0.1358	0.1365	*0.1098*	0.1188	0.1105	0.1661
0.4	-0.25	96	**0.1982**	0.2589	0.2556	0.2528	0.2438	0.2482	0.2473	0.2212	0.2809	*0.2140*	0.2344
		576	**0.0932**	0.1442	0.1418	0.1397	0.1362	0.1315	0.1428	0.1078	0.1268	*0.1026*	0.1473
	0.25	96	**0.1947**	0.2546	0.2439	0.2418	0.2317	0.2338	0.2496	0.2213	0.2663	*0.2187*	0.2365
		576	**0.0925**	0.1483	0.1426	0.1474	0.1322	0.1272	0.1349	0.1077	0.1238	*0.1043*	0.1558
	0.45	96	**0.1964**	0.2565	0.2534	0.2429	0.2412	0.2324	0.2448	0.2223	0.2643	*0.2125*	0.2424
		576	**0.0943**	0.1338	0.1288	0.1265	0.1169	0.1142	0.1396	0.1090	0.1229	*0.0975*	0.1466
0.9	-0.25	96	**0.0886**	0.1384	0.1357	0.1329	0.1310	0.1294	0.1367	*0.1105*	0.1253	0.1254	0.1427
		576	**0.0344**	0.0712	0.0696	0.0653	0.0628	0.0611	0.0636	*0.0561*	0.0589	0.0685	0.0749
	0.25	96	**0.0865**	0.1393	0.1329	0.1314	0.1299	0.1245	0.1343	0.1287	0.1286	*0.0952*	0.1426
		576	**0.0304**	0.0794	0.0768	0.0752	0.0746	0.0699	0.0712	0.0692	0.0706	*0.0487*	0.0758
	0.45	96	**0.0885**	0.1375	0.1348	0.1311	0.1297	0.1249	0.1324	0.1181	0.1197	*0.0991*	0.1424
		576	**0.0378**	0.0728	0.0715	0.0676	0.0619	0.0588	0.0662	0.0647	0.0695	*0.0454*	0.0743

Table 14: Bias estimates of the unadjusted LPR estimator, the feasible GS estimator and the pre-filtered sieve bootstrap estimator, for the DGP: ARFIMA$(1, d_0, 1)$ with Gaussian innovations. The bias estimates of the feasible jackknife estimator based on 2,3,4,6,8 non-overlapping (NO) sub-samples, the feasible jackknife estimator based on 2 moving block (MB) sub-samples, the maximum likelihood estimator (MLE) and the pre-whitened (PW) estimator are obtained under the misspecified model: ARFIMA$(2, d, 0)$ using the approach described in Section 5.3. The estimates are obtained by setting $\alpha = 0.65$. The lowest values are **bold-faced** and the second lowest values are *italicized*.

ϕ_0	θ_0	d_0	n	$\widehat{d}_n$	$\widehat{d}_{J,2}^{NO}$	$\widehat{d}_{J,3}^{NO}$	$\widehat{d}_{J,4}^{NO}$	$\widehat{d}_{J,6}^{NO}$	$\widehat{d}_{J,8}^{NO}$	$\widehat{d}_{J,2}^{MB}$	$\widehat{d}_1^{GS}$	$\widehat{d}^{PFSB}$	$\widehat{d}^{MLE}$	$\widehat{d}^{PW}$
-0.9	-0.9	-0.25	96	-0.5908	-0.5874	-0.5986	-0.6020	-0.6135	-0.6159	-0.6954	**-0.5065**	*-0.5557*	-0.6790	-0.6939
			576	-0.4027	*-0.3853*	-0.3899	-0.3936	-0.4007	-0.4048	-0.3941	**-0.3164**	-0.3860	-0.3946	-0.4226
		0.25	96	-0.6384	-0.5932	-0.5975	-0.6061	-0.6126	-0.6194	-0.5946	**-0.5353**	*-0.5780*	-0.6295	-0.6354
			576	-0.5379	*-0.4940*	-0.5008	-0.5049	-0.5174	-0.5236	-0.5073	**-0.4801**	-0.5128	-0.5396	-0.5585
		0.45	96	-0.7769	-0.7739	-0.7824	-0.7884	-0.7935	-0.7997	-0.7834	**-0.6807**	*-0.7504*	-0.7648	-0.7920
			576	-0.5886	-0.5677	-0.5785	-0.5842	-0.5878	-0.5911	-0.5970	**-0.5307**	*-0.5622*	-0.5764	-0.5985
	-0.4	-0.25	96	-0.0697	*-0.0523*	-0.0572	-0.0621	-0.0663	-0.0694	-0.0575	**-0.0491**	-0.0541	-0.0639	-0.0828
			576	*-0.0092*	**-0.0086**	-0.0094	-0.0108	-0.0137	-0.0179	-0.0095	-0.0369	-0.0144	-0.0188	-0.0395
		0.25	96	-0.2476	-0.2038	-0.2080	-0.2131	-0.2185	-0.2260	-0.2048	**-0.1294**	*-0.1543*	-0.2563	-0.2884
			576	-0.0721	-0.0554	-0.0613	-0.0675	-0.0726	-0.0798	-0.0657	**-0.0250**	*-0.0343*	-0.0885	-0.0967
		0.45	96	-0.2392	-0.2046	-0.2095	-0.2147	-0.2189	-0.2206	-0.2068	**-0.1106**	*-0.1530*	-0.2476	-0.2859
			576	-0.0635	-0.0498	-0.0548	-0.0624	-0.0665	-0.0701	-0.0639	**-0.0148**	*-0.0281*	-0.0545	-0.0984
	0.4	-0.25	96	**-0.0044**	-0.0056*	-0.0069	-0.0097	-0.0126	-0.0148	-0.0092	0.0197	0.0071	*-0.0056*	-0.0376
			576	0.0086	**0.0064**	*0.0085*	0.0103	0.0130	0.0196	0.0105	0.0099	0.0159	0.0102	0.0105
		0.25	96	0.0221	*0.0105*	0.0157	0.0194	0.0228	0.0259	0.0176	**0.0079**	0.0207	0.0354	0.0339
			576	**0.0013**	*0.0041*	0.0060	0.0095	0.0106	0.0187	0.0085	0.0052	0.0060	0.0193	0.0140
		0.45	96	0.0414	0.0319	0.0369	0.0427	0.0482	0.0517	*0.0189*	0.0399	**0.0180**	0.0368	0.0353
			576	**0.0035**	*0.0036*	0.0063	0.0088	0.0105	0.0120	0.0055	0.0094	0.0063	0.0147	0.0142
-0.4	-0.9	-0.25	96	-0.4968	-0.4469	-0.4527	-0.4636	-0.4690	-0.4742	-0.4506	**-0.3998**	*-0.4012*	-0.5026	-0.5293
			576	-0.3861	*-0.3357*	-0.3463	-0.3554	-0.3628	-0.3680	-0.3539	**-0.3141**	-0.3498	-0.3729	-0.4098
		0.25	96	-0.8033	*-0.7530*	-0.7629	-0.7683	-0.7753	-0.7822	-0.7749	**-0.7328**	-0.7932	-0.8013	-0.8347
			576	-0.5966	-0.5639	-0.5685	-0.5700	-0.5775	-0.5893	-0.5730	**-0.5413**	*-0.5550*	-0.5739	-0.5941
		0.45	96	-0.8501	*-0.8046*	-0.8153	-0.8203	-0.8279	-0.8328	-0.8276	**-0.7819**	-0.8092	-0.8362	-0.8504
			576	-0.5943	*-0.5537*	-0.5584	-0.5629	-0.5669	-0.5725	-0.5648	**-0.5396**	-0.5560	-0.5835	-0.6095
	-0.4	-0.25	96	-0.1826	-0.1455	-0.1538	-0.1597	-0.1638	-0.1687	-0.1594	**-0.0885**	*-0.1352*	-0.1944	-0.2263
			576	-0.0505	-0.0226	-0.0275	-0.0312	-0.0379	-0.0413	-0.0395	**-0.0085**	*-0.0155*	-0.0493	-0.1028
		0.25	96	-0.2172	-0.1648	-0.1728	-0.1749	-0.1884	-0.1926	-0.1754	**-0.1169**	*-0.1351*	-0.2274	-0.2343
			576	-0.0650	-0.0356	-0.0372	-0.0418	-0.0479	-0.0535	-0.0473	**-0.0202**	*-0.0328*	-0.0426	-0.1124
		0.45	96	-0.2261	-0.1748	-0.1812	-0.1856	-0.1940	-0.1972	-0.1749	**-0.1227**	*-0.1297*	-0.2375	-0.2559
			576	-0.0598	-0.0324	-0.0376	-0.0418	-0.0463	-0.0495	-0.0437	**-0.0137**	*-0.0269*	-0.0468	-0.1076
	0.4	-0.25	96	0.0147	**0.0115**	0.0179	0.0206	0.0238	0.0269	0.0148	*0.0121*	0.0249	0.0253	0.0563
			576	0.0089	0.0080	0.0088	0.0093	0.0104	0.0134	0.0093	*0.0075*	**0.0063**	0.0079	0.0176
		0.25	96	0.0182	**0.0096**	*0.0114*	0.0145	0.0186	0.0209	0.0189	0.0168	0.0115	0.0182	0.0412
			576	*0.0025*	0.0029	0.0045	0.0068	0.0083	0.0095	0.0057	0.0034	**0.0015**	0.0045	0.0134
		0.45	96	0.0167	0.0132	0.0176	0.0199	0.0205	0.0258	0.0155	**0.0067**	*0.0090*	0.0281	0.0458
			576	0.0106	*0.0093*	0.0104	0.0146	0.0188	0.0211	0.0178	0.0134	**0.0022**	0.0102	0.0195
0.4	-0.9	-0.25	96	0.2462	0.1905	0.1969	0.2023	0.2069	0.2135	0.1958	**0.1307**	*0.1461*	0.2646	0.2958
			576	0.0806	0.0548	0.0613	0.0649	0.0695	0.0710	0.0637	**0.0333**	*0.0450*	0.0817	0.1139
		0.25	96	0.2225	0.1842	0.1936	0.1970	0.2044	0.2153	0.1973	**0.1141**	*0.1372*	0.2309	0.2555
			576	0.0774	0.0515	0.0546	0.0563	0.0628	0.0694	0.0603	**0.0331**	*0.0475*	0.0832	0.1149
		0.45	96	0.2308	0.1948	0.2063	0.2159	0.2190	0.2243	0.2175	**0.1356**	*0.1470*	0.2294	0.2656
			576	0.0811	*0.0453*	0.0496	0.0528	0.0579	0.0639	0.0582	**0.0374**	0.0712	0.0957	0.1127
	-0.4	-0.25	96	0.0144	**0.0087**	*0.0095*	0.0126	0.0174	0.0183	0.0131	0.0131	0.0249	0.0255	0.0578
			576	0.0110	**0.0062**	0.0086	0.0098	0.0132	0.0148	0.0114	0.0092	*0.0068*	0.0108	0.0186
		0.25	96	0.0097	**0.0089**	*0.0096*	0.0113	0.0158	0.0169	0.0147	0.0186	0.0115	0.0156	0.0490
			576	*0.0010*	0.0025	0.0036	0.0051	0.0073	0.0089	0.0064	**0.0008**	0.0015	0.0084	0.0147
		0.45	96	0.0103	**0.0057**	*0.0066*	0.0081	0.0098	0.0109	0.0079	0.0084	0.0090	0.0275	0.0389
			576	0.0037	*0.0022*	0.0038	0.0049	0.0037	0.0046	0.0033	0.0043	**0.0022**	0.0089	0.0165
	0.4	-0.25	96	0.2349	*0.1828*	0.1950	0.1986	0.2058	0.2158	0.2074	**0.1412**	0.2509	0.2548	0.2873
			576	0.0740	*0.0453*	0.0515	0.0548	0.0594	0.0613	0.0553	**0.0293**	0.0580	0.0577	0.1098
		0.25	96	0.2172	*0.1775*	0.1824	0.1869	0.1936	0.2004	0.1783	**0.1303**	0.2243	0.2263	0.2464
			576	0.0707	*0.0379*	0.0438	0.0518	0.0549	0.0593	0.0476	**0.0279**	0.0525	0.0954	0.1153
		0.45	96	0.2193	*0.1846*	0.1895	0.1938	0.1996	0.2026	0.1857	**0.1426**	0.2139	0.2295	0.2367
			576	0.0738	*0.0438*	0.0489	0.0527	0.0579	0.0637	0.0554	**0.0332**	0.0497	0.0840	0.1081

Table 15: RMSE estimates of the unadjusted LPR estimator, the feasible GS estimator and the pre-filtered sieve bootstrap estimator, for the DGP: ARFIMA$(1, d_0, 1)$ with Gaussian innovations. The bias estimates of the feasible jackknife estimator based on 2,3,4,6,8 non-overlapping (NO) sub-samples, the feasible jackknife estimator based on 2 moving block (MB) sub-samples, the maximum likelihood estimator (MLE) and the pre-whitened (PW) estimator are obtained under the misspecified model: ARFIMA$(2, d, 0)$ using the approach described in Section 5.3. The estimates are obtained by setting $\alpha = 0.65$. The lowest values are **bold-faced** and the second lowest values are *italicized*.

ϕ_0	θ_0	d_0	n	$\widehat{d}_n$	$\widehat{d}^{NO}_{J,2}$	$\widehat{d}^{NO}_{J,3}$	$\widehat{d}^{NO}_{J,4}$	$\widehat{d}^{NO}_{J,6}$	$\widehat{d}^{NO}_{J,8}$	$\widehat{d}^{MB}_{J,2}$	$\widehat{d}^{GS}_{1}$	$\widehat{d}^{PFSB}$	$\widehat{d}^{MLE}$	$\widehat{d}^{PW}$
-0.9	-0.9	-0.25	96	**0.6822**	0.7230	0.7187	0.7146	0.7048	0.6932	0.7044	*0.6912*	0.6943	0.7867	0.9373
			576	**0.4804**	0.5062	0.4935	0.4895	0.4893	*0.4811*	0.4928	0.4830	0.4916	0.5538	0.6328
		0.25	96	**0.6832**	0.7048	0.6964	0.6929	0.6924	*0.6843*	0.6923	0.6849	0.7020	0.7571	0.8299
			576	**0.5506**	0.5885	0.5837	0.5735	0.5692	*0.5606*	0.5682	0.5675	0.5608	0.6059	0.6738
		0.45	96	**0.8021**	0.8420	0.8356	0.8310	0.8278	*0.8109*	0.8339	0.8191	0.8486	0.8546	0.9461
			576	**0.5963**	0.6142	0.6099	0.6034	0.5963	0.6025	0.6062	*0.6024*	0.5975	0.6259	0.7026
	-0.4	-0.25	96	**0.2310**	0.2539	0.2485	0.2426	0.2377	*0.2318*	0.2447	0.2549	0.2798	0.2810	0.3138
			576	**0.1089**	0.1354	0.1247	0.1183	0.1162	*0.1094*	0.1153	0.1274	0.1792	0.1955	0.2200
		0.25	96	**0.3075**	0.3418	0.3356	0.3295	0.3249	0.3241	0.3166	*0.3113*	0.3180	0.3257	0.3547
			576	*0.1153*	0.1367	0.1281	0.1214	0.1185	**0.1073**	0.1258	0.1192	0.1165	0.1328	0.1505
		0.45	96	**0.3057**	0.3369	0.3298	0.3248	0.3176	0.3119	0.3165	*0.3104*	0.3034	0.3273	0.3442
			576	**0.1109**	0.1387	0.1350	0.1319	0.1282	0.1221	0.1379	*0.1159*	0.1172	0.1351	0.1518
	0.4	-0.25	96	**0.1927**	0.2319	0.2284	0.2166	0.2078	*0.1946*	0.2053	0.2362	0.2425	0.2550	0.2863
			576	**0.0873**	0.1057	0.1011	0.0978	0.0943	*0.0910*	0.0967	0.1039	0.1247	0.1349	0.1478
		0.25	96	**0.1979**	0.2229	0.2195	0.2169	0.2143	*0.2016*	0.2038	0.2399	0.2414	0.2666	0.2949
			576	*0.0886*	0.1006	0.0978	0.0922	0.0906	**0.0883**	0.0946	0.1062	0.1159	0.1328	0.1455
		0.45	96	**0.1806**	0.2074	0.1940	0.1863	*0.1810*	0.1839	0.1947	0.2256	0.2694	0.2862	0.3050
			576	**0.0965**	0.1285	0.1239	0.1157	0.1075	*0.1028*	0.1068	0.1147	0.1172	0.1374	0.1528
-0.4	-0.9	-0.25	96	**0.5789**	0.6079	0.6030	0.5982	0.5893	*0.5830*	0.5871	0.5831	0.5940	0.6159	0.6819
			576	**0.4357**	0.4661	0.4568	0.4477	0.4446	*0.4394*	0.4465	0.4644	0.4696	0.4760	0.5229
		0.25	96	0.8289	0.8375	0.8236	0.8164	0.8092	*0.7973*	0.8084	**0.7727**	0.8217	0.8421	0.8931
			576	**0.6037**	0.6369	0.6295	0.6231	0.6203	0.6169	0.6138	*0.6124*	0.6131	0.6302	0.6529
		0.45	96	**0.8724**	0.9146	0.9073	0.9074	0.8940	0.8920	0.8856	*0.8777*	0.8759	0.8913	0.9204
			576	**0.6008**	0.6373	0.6319	0.6277	0.6186	0.6159	0.6270	*0.6096*	0.6108	0.6278	0.6559
	-0.4	-0.25	96	**0.2724**	0.2968	0.2926	0.2883	0.2815	*0.2796*	0.2843	0.2819	0.3181	0.3319	0.3463
			576	**0.1081**	0.1253	0.1218	0.1168	0.1130	*0.1098*	0.1189	0.1121	0.1228	0.1430	0.1738
		0.25	96	**0.2870**	0.3216	0.3155	0.3079	0.3026	0.2954	0.3165	*0.2943*	0.3015	0.3276	0.3424
			576	**0.1136**	0.1388	0.1342	0.1265	0.1248	*0.1219*	0.1278	0.1222	0.1294	0.1448	0.1575
		0.45	96	*0.2922*	0.3353	0.3297	0.3169	0.3082	**0.2905**	0.3057	0.2958	0.2987	0.3132	0.3406
			576	**0.1084**	0.1279	0.1220	0.1195	0.1126	0.1104	0.1138	*0.1092*	0.1134	0.1473	0.1738
	0.4	-0.25	96	**0.1840**	0.2156	0.2109	0.2086	0.1958	*0.1876*	0.1945	0.2260	0.2600	0.3119	0.3362
			576	**0.0923**	0.1160	0.1148	0.1099	0.1031	*0.0984*	0.1179	0.1094	0.1112	0.1380	0.1594
		0.25	96	**0.1808**	0.2188	0.2119	0.2055	0.1974	*0.1945*	0.2051	0.2218	0.2546	0.2699	0.2883
			576	**0.0910**	0.1279	0.1236	0.1178	0.1127	*0.1028*	0.1149	0.1082	0.1057	0.1168	0.1378
		0.45	96	**0.1809**	0.2051	0.2012	0.1963	0.1912	*0.1888*	0.1973	0.2190	0.2482	0.2605	0.2756
			576	**0.0943**	0.1126	0.1098	0.0942	0.0919	*0.0958*	0.1028	0.1120	0.1072	0.1285	0.1564
0.4	-0.9	-0.25	96	**0.3064**	0.3385	0.3365	0.3286	0.3202	*0.3160*	0.3194	0.3172	0.3118	0.3326	0.3658
			576	*0.1210*	0.1432	0.1373	0.1338	0.1293	0.1249	0.1376	**0.1130**	0.1275	0.1475	0.1692
		0.25	96	*0.2916*	0.3164	0.3127	0.3089	0.3044	0.2962	0.3083	**0.2565**	0.3104	0.3326	0.3566
			576	0.1196	0.1347	0.1289	0.1235	0.1210	*0.1178*	0.1275	**0.1125**	0.1326	0.1549	0.1741
		0.45	96	*0.2933*	0.3276	0.3190	0.3139	0.3077	0.3023	0.3198	**0.2583**	0.3015	0.3285	0.3478
			576	*0.1224*	0.1436	0.1382	0.1327	0.1295	0.1241	0.1326	**0.1148**	0.1250	0.1463	0.1653
	-0.4	-0.25	96	**0.1896**	0.2435	0.2359	0.2268	0.2144	*0.2037*	0.2254	0.2332	0.2600	0.2844	0.2727
			576	**0.0936**	0.1237	0.1210	0.1174	0.1133	*0.1085*	0.1153	0.1105	0.1112	0.1375	0.1540
		0.25	96	**0.1842**	0.2352	0.2276	0.2215	0.2146	*0.2072*	0.2378	0.2255	0.2546	0.2768	0.2803
			576	**0.0923**	0.1328	0.1268	0.1174	0.1057	0.1175	0.1245	*0.1095*	0.1057	0.1239	0.1456
		0.45	96	**0.1878**	0.2367	0.2257	0.2193	0.2142	*0.2018*	0.2229	0.2288	0.2482	0.2680	0.2807
			576	**0.0939**	0.1295	0.1238	0.1176	0.1116	*0.1063*	0.1165	0.1113	0.1072	0.1254	0.1482
	0.4	-0.25	96	**0.2992**	0.3273	0.3190	0.3128	0.3067	*0.3018*	0.3176	0.3412	0.3540	0.3707	0.3964
			576	*0.1164*	0.1368	0.1341	0.1289	0.1226	0.1200	0.1348	**0.1108**	0.1375	0.1442	0.1508
		0.25	96	0.2839	0.3156	0.3073	0.2958	0.2895	0.2759	*0.2836*	**0.2564**	0.3382	0.3538	0.3511
			576	*0.1171*	0.1327	0.1289	0.1230	0.1187	0.1142	0.1284	**0.1131**	0.1309	0.1518	0.1695
		0.45	96	0.2905	0.3219	0.3176	0.3064	0.2954	*0.2882*	0.2972	**0.2721**	0.3351	0.3564	0.3774
			576	0.1187	0.1388	0.1326	0.1259	0.1202	*0.1147*	0.1274	**0.1144**	0.1337	0.1532	0.1660

Table 16: A ranking of the estimation methods.

Panel A: Correct Specification; known parameters (Tables 2-5)

True DGP	Bias First	Bias Second	Bias Third	RMSE First	RMSE Second	RMSE Third
ARFIMA(1,d_0, 0)	$\widehat{d}^{MLE}$	$\widehat{d}_{J,2}^{Opt-NO}$	$\widehat{d}_1^{GS}$	$\widehat{d}^{MLE}$	$\widehat{d}^{PW}$	$\widehat{d}_{J,8}^{Opt-NO}$
ARFIMA(0,d_0, 1)	$\widehat{d}^{MLE}$	$\widehat{d}_{J,2}^{Opt-NO}$	$\widehat{d}_1^{Opt-GS}$	$\widehat{d}^{MLE}$	$\widehat{d}_{J,8}^{Opt-NO}$	$\widehat{d}^{PW}$

Panel B: Correct Specification; unknown parameters (Tables 6-9)

True DGP	Bias First	Bias Second	Bias Third	RMSE First	RMSE Second	RMSE Third
ARFIMA(1,d_0, 0)	$\widehat{d}_1^{GS}/\widehat{d}^{PFSB}$	$\widehat{d}_{J,2}^{NO}$	$\widehat{d}^{MLE}$	$\widehat{d}^{MLE}$	$\widehat{d}_n$	$\widehat{d}^{PW}$
ARFIMA(0,d_0, 1)	$\widehat{d}_{J,2}^{NO}$	$\widehat{d}_1^{GS}$	$\widehat{d}^{MLE}$	$\widehat{d}^{MLE}$	$\widehat{d}_n$	$\widehat{d}_{J,8}^{NO}$

Panel C: Misspecification (Tables 10-15)

Form of misspec.	Bias First	Bias Second	Bias Third	RMSE First	RMSE Second	RMSE Third
(i)	$\widehat{d}^{PFSB}$	$\widehat{d}_1^{GS}$	$\widehat{d}_{J,2}^{NO}$	$\widehat{d}_n$	$\widehat{d}^{PFSB}$	$\widehat{d}_1^{GS}$
(ii)	$\widehat{d}_{J,2}^{NO}$	$\widehat{d}_1^{GS}$	$\widehat{d}^{PFSB}$	$\widehat{d}_n$	$\widehat{d}^{MLE}$	$\widehat{d}_1^{GS}$
(iii)	$\widehat{d}_1^{GS}$	$\widehat{d}_{J,2}^{NO}$	$\widehat{d}^{PFSB}$	$\widehat{d}_n$	$\widehat{d}_1^{GS}$	$\widehat{d}_{J,8}^{NO}$