Inference for likelihood-based estimators of generalized long-memory processes[☆]

Paul M Beaumont[a], Aaron D Smallwood[b,*]

[a]*Department of Economics, Florida State University, Tallahassee, FL 32306, USA*
[b]*Department of Economics, University of Texas Arlington, 701 S. West Street. (mailbox: 19479), Arlington, TX 76019, USA*

Abstract

Despite a recent proliferation of research on cyclical long memory models, very little is known regarding the asymptotic properties of likelihood-based estimation methods. Several estimators have been studied for the Gegenbauer autoregressive moving average (GARMA) process. However, only Chung (1996a,b) present asymptotic results for all parameters and even those, which lack an initial consistency proof, are tenuous. In this paper, we review the GARMA process and study the properties of frequency and time domain likelihood-based estimators using Monte Carlo analysis. Our results show both estimators perform well, and there is some support for the theory of Chung concerning the parameter governing the cycle length. However, we find that asymptotic confidence bands are unreliable in small samples with weak long memory, and the distribution theory under the null of an infinitely long cycle is unreliable. As an alternative testing procedure, we propose a parametric bootstrap approach that we demonstrate with an application to US unemployment rate data.

Keywords: long memory, GARMA, CSS estimator, Whittle estimator
JEL Classification Codes: C22, C40, C58, G12

1. Introduction

Few contributions to time series analysis have generated as much interest as the introduction of long memory by Granger and Joyeux (1980) and Hosking (1981). These methods allow for slowly decaying autocorrelation functions and spectral density functions with one or more singularity. In economics, long memory has provided a major breakthrough in allowing researchers to bridge the gap between unit roots and transitory I(0) dynamics.

As emphasized by Dissanayake et al. (2018), attention has recently focused on methods that can accommodate cyclical long memory, including seasonal long memory and Gegenbauer autoregressive moving average (GARMA) models. The GARMA model, which has received considerable attention, is defined as follows,

$$(1 - 2\eta L + L^2)^\lambda \phi(L)(x_t - \mu) = \theta(L)\varepsilon_t, \tag{1}$$

where $\phi(L)$ and $\theta(L)$ are p and q order polynomials in the lag operator L, and ε_t is a disturbance sequence with $E(\varepsilon_t^2) = \sigma^2$ and no serial correlation. Letting $\nu = \cos^{-1}(\eta)$, the spectral density function is given by

$$f(\omega) = \frac{\sigma^2}{2\pi} \left| \frac{\theta(e^{-i\omega})}{\phi(e^{-i\omega})} \right| 2|\cos(\omega) - \cos(\nu)|^{-2\lambda}. \tag{2}$$

When $p = q = 0$, the process has an autocorrelation function at lag j that is proportional to $\cos(j\nu)j^{2\lambda-1}$. When $\nu = 0$, the result is an ARFIMA$(p, 2\lambda, q)$ process, as studied by Granger and Joyeux (1980) and

[☆]The authors acknowledge the Texas Advanced Computing Center (TACC) for providing HPC resources that were used to generate simulation results in the paper. URL: http://www.tacc.utexas.edu.

[*]Corresponding author: Aaron Smallwood, Department of Economics, University of Texas Arlington, 701 S. West Street (mailbox: 19479), Arlington, TX 76019, USA

Email addresses: `beaumont@fsu.edu` (Paul M Beaumont), `smallwood@uta.edu` (Aaron D Smallwood)

Hosking (1981). The process above is covariance stationary provided $\lambda < 1/4$ when $\nu \in \{0, \pi\}$, or for $\lambda < 1/2$ when $\nu \in (0, 1)$ (Gray et al. 1989).

This model, and its extension, the k-factor GARMA model, have been applied across virtually every discipline that uses time series methods. Examples include atmospheric C02 (Woodward et al. 1998), sunspots (Chung 1996b; Artiach and Arteche 2012), dust pollution (Reisen et al. 2014), river flow (Diongue and Ndongo 2016), electricity demand (Leschinski and Sibbertsen 2019), and traffic patterns (Ferrara and Guégan 2001). In economics and finance, the GARMA and k-factor GARMA models have been applied to study information related to equities (Lu and Guegan 2011; Caporale and Gil-Alana 2014; Beaumont and Smallwood 2019), interest rates (Ramachandran and Beaumont 2001; Gil-Alanaa 2007; Asai et al. 2018), inflation (Arteche and Robinson 2000; Caporale and Gil-Alana 2011; Peiris and Asai 2016), and unemployment (Gil-Alanaa 2007).

In spite of this increased interest in the GARMA model, there is not a unifying estimation approach for all parameters, most notably η, or η_i in the multiple frequency case. Several estimators exist, including estimators in both the time domain (Chung 1996a,b; Dissanayake et al. 2018) and frequency domain (Giraitis et al. 2001) and semi-parametric estimators (Hidalgo and Soulier 2004; Hidalgo 2005) extending the log-periodogrom regressions of Geweke and Porter-Hudak (1983) and Robinson (1995). However, a full set of accepted asymptotics does not exist. For the single frequency case, Yajima (1996) showed that maximization of the periodogram could be used to consistently estimate the position of spectral poles, which would include ν. Extending these results, for semi-parametric estimators, Hidalgo and Soulier (2004) estimate ν using the periodogram and provide an asymptotic result for λ based on (1), showing that the distribution for ν known and unknown is identical. Based on this result, many existing studies fix the position of spectral singularities before estimating other parameters (Caporale and Gil-Alana 2014; Diongue and Ndongo 2016). This approach is sensible for those only interested in point estimates, but it does not allow one to obtain a measure related to the uncertainty of their estimates of η and ν. Obtaining valid inference for η is important for researchers interested in generating confidence bands for estimated cycles and is imperative for those interested in tests for specific values, such as $\eta = 1$.

Some incomplete solutions have been been proposed. Hidalgo (2005) extends the results of Hidalgo and Soulier (2004) and provides an estimator for ν that is asymptotically normal converging at rate T^{β}, with $\beta < 1$, and T denoting the sample size, and whose distribution depends on whether $\nu = \{0, \pi\}$ or $\nu \in (0, \pi)$. Giraitis et al. (2001) provide the distribution for λ for the Whittle estimator, and establish rate T convergence for $\hat{\nu}$. Unfortunately, Giraitis et al. (2001) are unable to provide a full set of results for $\hat{\nu}$. Possibly the most promising results were proposed by Chung (1996a,b), who argued that the constrained sum of squares (CSS) estimator for η converges at either rate T, if $|\eta| < 1$, or T^2, if $|\eta| = 1$, to ratios of functionals of Brownian motion processes. Estimates for the remaining parameters are $O_p(T^{-0.5})$ and achieve asymptotic normality. The CSS based results have recently been extended to consider multiple long memory cycles (Beaumont and Smallwood 2019) and heteroskedastic diisturbances (Peiris and Asai 2016). Regrettably, given a closed parameter space and a potential discontinuity in the distribution for η, an initial consistency proof for the CSS estimators has proven quite elusive. Specifically, Chung and related extensions rely on the observation that the score evaluated at the true parameters is zero, which, as pointed out by Giraitis et al. (2001), may not be sufficient to establish consistency. Given the concerns regarding the theoretical results for the CSS estimator, it remains an open question as to whether or not the theory is useful for inference in practical applications.

In this paper, we review the properties of two commonly applied likelihood-based estimators for the GARMA parameters, the CSS and Whittle methods, and consider the consequences of using the distribution theory of Chung (1996a,b). Our analyses show that both estimators of η yield desirable results in terms of mean bias and root mean squared error. Using an algorithm that does not constrain the parameter space, however, there are some concerns regarding Chung's theory. For $|\eta| < 1$, the estimates of η have fatter tails and a more peaked density than theory implies. Although the problem may become negligible for large samples, differences between theoretical and empirical densities can become severe for small samples and with values of λ close to zero, such that asymptotic confidence bands can be quite unreliable. Further, the results show that empirical test-sizes can at times be inaccurate, which can be problematic when testing $\eta = 1$. Here, the use of distribution theory can specifically lead to severe size distortion, potentially complicating issues for researchers attempting to test if a finite cycle exists. We find that the the use of more conservative test sizes and confidence bands could be used, and even more preferably, computational methods can likely be employed to overcome any deficiencies in the existing theory. A parametric bootstrapped likelihood ratio test applied to US unemployment supports this assertion.

The rest of the paper is organized as follows. In Section 2, we motivate the problem using the US unemployment rate and present the two estimators of the GARMA process due to Chung (1996a,b) and Giraitis et al. (2001). In Section 3, we present the Monte Carlo results, concentrating on the calculation of confidence bands. Section 4 offers evidence specific to testing $|\eta| = 1$ and explores an alternative test. A final section concludes.

2. Parametric Estimators of the GARMA Process

The two methods studied here use time and frequency domain approximations for the Gaussian log-likelihood function based on the GARMA process in (1). The first estimator we consider is a Whittle type method. For a sample size T, let $\tilde{T} = [T/2]$, where $[\cdot]$ denotes the integer part. Let Δ denote the set of all possible parameter values for $\delta = \{\phi', \theta', \lambda, \mu\}$, and let Q_T denote the set of Fourier frequencies, $\omega_j = 2\pi j/T$. Based on the spectrum $f(\omega_j)$ defined in (2), Giraitis et al. (2001) propose the following estimator (GHR),

$$\begin{pmatrix}\hat{\delta}\\ \hat{\nu}\end{pmatrix} = \text{argmin}_{\Delta\times Q_T} S(\delta,\nu), \qquad S(\delta,\nu) = \left[\frac{1}{\tilde{T}}\sum_{j=0}^{\tilde{T}} \frac{I(\omega_j)}{f(\omega_j)}\right], \tag{3}$$

where $I(\omega_j) = \frac{1}{2\pi T}\left|\sum_{t=1}^{T} x_t e^{it\omega_j}\right|^2$. Importantly, note that $\hat{\nu}$ is obtained with respect to the discrete set Q_T, and that the true value of ν need not be in this set. Under suitable regularity conditions, Giraitis et al. (2001) establish asymptotic normality for $\hat{\delta}$ and prove that $\hat{\nu}$ is consistent. However, a limiting distribution for $\hat{\nu}$ is not available because the function in (3) is not minimized for all values in the interval $[0, \pi]$.

Concentrating out the residual variance, and under a Gaussian assumption for ε_t, our second estimator is based on the constrained sum of squares (CSS) function of the parameters ϕ', θ', λ, η, μ, and is given by

$$L(\phi',\theta',\lambda,\eta,\mu) = -\frac{T}{2}\left[\ln(2\pi) + \ln\left(\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t^2\right) + 1\right]. \tag{4}$$

Under an initialization assumption, maximization of the CSS function produces a set of estimates that is asymptotically equivalent to the maximum likelihood values.

The proposed theory of Chung (1996b) establishes asymptotic normality for $\hat{\delta} = (\hat{\lambda}, \hat{\phi}', \hat{\theta}')'$, where, for example, the element from the information matrix for λ is given by $I_\lambda = 2(\pi^2/3 - \pi\nu + \nu^2)$. For $\hat{\eta}$, Chung has the following result in his Theorem 2:

$$T(\hat{\eta} - \eta) \underset{d}{\to} \frac{\sin(\nu)}{\lambda} Y_0 = \frac{\sin(\nu)}{\lambda} \frac{\int_0^1 W_1 dW_2 - \int_0^1 W_2 dW_1}{\int_0^1 W_1^2(r)dr + \int_0^1 W_2^2(r)dr}, \quad \text{for } |\eta| < 1, \text{ and} \tag{5}$$

$$T^2(\hat{\eta} \pm 1) \underset{d}{\to} \mp\frac{1}{2\lambda} Y_1 = \mp\frac{1}{2\lambda}\frac{\int_0^1 \left[\int_0^r W(s)\,ds\right] dW(r)}{\int_0^1 \left[\int_0^r W(s)\,ds\right]^2 dr}, \quad \text{for } \eta = \pm 1 \tag{6}$$

where W, W_1, and W_2 are independent Brownian motions. Percentiles of Y_0 and Y_1 can be simulated to yield confidence bands and test statistics for η.

To illustrate a concern regarding these asymptotic results, the following GARMA (2,1) model was obtained using the CSS estimator for the US unemployment rate, u_t,[1]

$$(1 - 2 * 0.9986L + L^2)^{0.326}(1 - 0.934L - 0.026L^2)(u_t - 6.16) = (1 - 0.567L)\varepsilon_t. \tag{7}$$

With $\hat{\eta} = 0.9986 < 1$ and $\hat{\lambda} = 0.326 < 0.5$, the estimated model represents a stationary process with highly persistent cycles. In contrast, if $\eta = 1$, the model is a non-stationary ARFIMA process with difference

[1] Seasonally adjusted monthly data are from January 1980 through July 2019 and are collected from the St. Louis Federal Reserve Bank. The GARMA(2,1) model had the lowest AIC with p and q less than 3.

parameter $d = 2\lambda = 0.652$, and, therefore it lacks a finite cumulative impulse response function. Using equation 6, 99% bands about 1 do not contain $\hat{\eta}$, implying we reject $\eta = 1$, while the corresponding bands about $\hat{\eta}$ based on rate T convergence are $[0.9972, 1.0001]$, implying we fail to reject $\eta = 1$.[2] Although more formal testing procedures advocated below generally yield evidence in support of $\eta < 1$, the conflicting test results above could prove confusing in practice. In the following sections, we provide Monte Carlo analysis to shed light on the applicability of theoretical results both in forming confidence bands and for testing $\eta = 1$.

3. The Monte Carlo Results

In this section, we present Monte Carlo results to assess how the estimators perform in both small and larger samples. From a computational perspective, an advantage of the Whittle based method of Giraitis et al. (2001) is its relative simplicity. For each Fourier frequency, ω_j, we minimize the function $S(\delta, \omega_j)$ in (3) with respect to δ, and track the value of the objective function for $j = 0, 1, \ldots, \tilde{T}$. The estimate of ν, $\hat{\nu}$, is the frequency associated with the minimum value of the objective function amongst the $\tilde{T} + 1$ alternatives. Then, the estimate of δ is the value that minimizes the objective function with the frequency fixed at $\hat{\nu}$. An estimate of η can be obtained through the functional relationship, $\eta = \cos(\nu)$.

For the CSS estimator, note that the polynomial $(1 - 2\eta L + L^2)^{-\lambda}$ is related to the Gegenbauer polynomials, c_j, as follows (Gray et al. 1989):

$$(1 - 2\eta L + L^2)^{-\lambda} = \sum_{j=0}^{\infty} c_j L^j, \text{ where, } c_j = \sum_{k=0}^{[j/2]} \frac{(-1)^k \Gamma(\lambda + j - k)(2\eta)^{j-2k}}{\Gamma(\lambda)\Gamma(k+1)\Gamma(j-2k+1)}, \tag{8}$$

where $\Gamma(\cdot)$ is the gamma function. Starting from $c_0 = 1$ and $c_1 = 2\eta\lambda$, the Gegenbauer coefficients can be computed recursively as

$$c_j = 2\eta\left(\frac{\lambda - 1}{j} + 1\right) c_{j-1} - \left(2\frac{\lambda - 1}{j} + 1\right) c_{j-2}. \tag{9}$$

With $z_t = x_t - \mu$, the disturbance sequence in (1) is calculated as,

$$\varepsilon_t = z_t - \sum_{i=1}^{p} \phi_i z_{t-i} - \sum_{j=1}^{t-1} c_j \varepsilon_{t-j} - \sum_{i=1}^{q} \left[\theta_i \sum_{j=0}^{t-1-i} c_j \varepsilon_{t-j-i}\right]. \tag{10}$$

To calculate the model parameters, Chung (1996a,b) and Gray et al. (1989), advocate a line-search for η.[3] This implies that the parameter space being searched over is a discrete set. This is unfortunate as the theory is developed under the assumption of a continuous parameter space, and an important advantage of the CSS estimator relative to the GHR counterpart is lost. If the value of η is unknown, a discretization implies that potentially large biases may result or that a very fine grid would need to be employed. Further, if a large number of grid points is selected, it almost certainly becomes necessary to impose boundary constraints on η for computational purposes.

Here, we advocate the use of a double gradient-based procedure as in Ramachandran and Beaumont (2001). A set of starting values for η is selected, which is typically a grid from -1 to 1, using a small step size to avoid potential local minima. Conditional on each value of η in the grid, an estimate of δ is obtained using a gradient based method. Then, conditional on this $\hat{\delta}$, we estimate η, again using a search algorithm. The procedure continues as we update the value of η along the grid. Once a neighborhood for the maximum value of the CSS function is obtained, we iterate over δ and η using a double-gradient procedure until the norm of the estimated parameters between steps is sufficiently small. This procedure allows one to search

[2] These quantities are calculated using (5) and (6) based on Chung's simulated values for Y_0 and Y_1. For example, the 0.5 and 99.5 percentiles for Y_0 are -4.238 and 4.238. With $\nu = 0.0525$ and a sample size of $T = 475$, confidence bands for η can be calculated as $[-4.238\sin(\nu)/(475 * 0.326), 4.238\sin(\nu)/(475 * 0.326)]$.

[3] Chung (1996a,b) proposes estimation of δ using a gradient based method for each value of η along the grid, while Gray et al. (1989) also estimate λ using a line search.

over all possible values of η and does not impose boundary constraints that could artificially improve the fit of the CSS estimator in a simulation environment.

For the Monte Carlo experiments, we considered a total of eight cases, including six GARMA(0,0) (Table 1) and two GARMA(1,0) (Table 2) models. The true values of η are $\{-1, -0.9995, -0.50, 0.50, 0.9995, 1\}$ for the GARMA(0,0) cases. For $|\eta| = 1$, we fix $\lambda = 0.20$, where $\lambda = 0.40$ otherwise. For the GARMA(1,0) cases, we fixed $\lambda = 0.40$ and $\phi = 0.80$, allowing the true values of η to be 0.50 and 0.9995. Note that the last model has short memory dynamics and is parametrically close to the non-stationary border. We thus anticipate that this model may produce poor results. For each process, we performed 2500 simulations and considered sample sizes of 100, 300, 500, 1000, and 2000 observations. To generate a data series, x_t, we calculated the autocovariances of the long memory processes and obtained the Cholesky factorization of the Toeplitz matrix, which is then multiplied by a sequence of normal random variates of the desired length.[4] Data are generated through recursion for each GARMA(1,0) case, where μ is set to 0 throughout.

Tables 1 and 2 report the results of the mean bias and RMSE for each model from both the CSS and GHR estimators. To help interpret the results, the method that yields the smallest absolute bias and RMSE for a given sample size is shown in bold type. For both estimators, the absolute value of the mean bias associated with η is quite small, with a value that decreases rapidly as the sample size increases. The CSS estimator outperforms GHR in terms of the mean bias for η.[5] There are instances where the improvement in mean and RMSE of CSS relative to GHR can be large, especially when $|\eta| = 0.50$. This likely results from the fact that the true value of η is not typically in the discrete parameter space for the GHR estimator except when ν is 0 or π. When $\nu \neq 0$ ($\eta \neq 1$), the GHR estimator tends to dominate in mean bias for λ. In terms of RMSE, for $|\eta| \neq 1$, the CSS method tends to dominate for η, λ, and in the cases of the GARMA(1,0) model, for ϕ as well. It should be noted that the RMSE for both estimators of λ and ϕ are generally quite similar, and compare favorably with the computed asymptotic standard deviations of these parameters from Theorem 3 in Chung (1996b), with one exception. The performance of the estimators for the GARMA(1,0) model with $\eta = 0.9995$, $\lambda = 0.40$, and $\phi = 0.80$, tends to be quite poor. With $T < 2000$, the CSS and GHR procedures can result in a mean bias for λ of -0.2023 and -0.3017, respectively. A similar picture emerges for ϕ, where the mean bias of $\hat{\phi}$ can be as large as 0.0992 for the CSS estimator, while the GHR method tends to underestimate ϕ with a mean bias that is typically quite large in absolute value. The results for the GARMA(1,0) cases show that the mean bias in $\hat{\lambda}$ tends to be inversely related to the mean bias in $\hat{\phi}$, especially for the CSS estimator. As is well known to researchers using parametric estimators in the ARFIMA context, it can be difficult to distinguish between high frequency and low frequency pieces (Nielsen and Frederiksen 2005).

For researchers interested in obtaining point estimates for GARMA parameters, and η specifically, the GHR and CSS estimators appear to provide highly robust options. However, the question remains as to whether or not proposed distribution theory can be used for inference and the construction of confidence bands. To this end, Table 3 displays the estimated and theoretical percentiles from two potentially problematic cases, one with η near 1 and another with an autoregressive component.[6] Below the reported percentiles, we present the distribution values calculated from Chung (1996a), using his equation 19 and Table 1, followed by the empirical distribution of the same quantity resulting from both the CSS and GHR estimators for each sample size.

We are primarily interested in the CSS estimator, noting two things about the GHR method. First, the empirical distribution of $T(\hat{\eta} - \eta)$ confirms the estimator is $O_p(T^{-1})$ as shown by Giraitis et al. (2001). Second, we note that the underlying parameter space is discrete. Consider, for example, the empirical distribution of $T(\hat{\eta} - \eta)$ when the true values of η/λ are 0.9995/0.40 for $T = 300$. Of the 2500 values of $\hat{\eta}$, 1162 are exactly equal to 0.99912, the closest possible value in the grid to 0.9995. While GHR unquestionably performs well, this discretization can naturally result in small biases, which again helps to explain the findings in Tables 1 and 2, where the CSS method tends to dominate. This also highlights our concern in using a CSS-based algorithm that establishes a line search for η using a discrete set as in Chung

[4]See McElroy and Holan (2012) for details concerning the autocovariances of Gegenbauer processes.

[5]In important contributions with fixed ν, ? and Diongue and Ndongo (2016) also demonstrate the relative efficiency of CSS methods for the memory parameter relative to Whittle estimators under heteroskedastic and infinite variance disturbances.

[6]For brevity, we do not include all results from Tables 1 and 2, which are available upon request. Briefly summarizing, they provide conclusions that are qualitatively identical to those reported in Table 3, except when $\phi = 0.80$ and $\eta = 0.9995$. Here, not surprisingly, we find the distribution of $T(\hat{\eta} - \eta)$ is skewed left.

(1996a) and Gray et al. (1989). In what follows, we concentrate on the properties of the CSS estimator and the algorithm we proposed above.

For the models in Table 3, we generally see that the CSS estimator of η has an empirical distribution that is fairly well captured by the asymptotic distribution provided by Chung (1996a,b). The values of the empirical percentiles are typically close to the reported percentiles of Chung, especially for the 2.5% and 5.0% levels, which are commonly used for testing. Based on $T = 500$, for example, with $\eta = 0.50$ and $\phi = 0.8$, the empirical 5^{th} percentile for $T(\hat{\eta} - \eta)$ is -4.417, which is close to the proposed theoretical quantity of -4.70.

In spite of providing some support for the proposed distributional results, Table 3 hints that the finite sample distribution may have a more peaked density and fatter tails than theory implies. Consider, for example, Figure 1, which provides kernel density plots of $2000(\hat{\eta} - \eta)$ and the corresponding theoretical quantity using the theory for the CSS estimator based on the GARMA(0,0) model with $\eta = 0.50$ and $\lambda = 0.40$. The figure shows that there are several places where the associated kernel density plots cross, suggesting that in finite samples, the empirical distribution may have larger kurtosis and a more peaked density than implied by theory. We now turn to the question of whether these issues impact the practical usefulness of the distributional results in constructing confidence bands.

A large number of simulations generally reveal that theoretical confidence bands provide accurate coverage relative to empirical counterparts when $|\eta|$ is in the neighborhood of unity and/or λ is near the non-stationary boundary. However, there can be very serious concerns with the use of theory, especially when T is small and λ is near 0. Tables 4 and 5 highlight these issues, where we present the associated biases that would result from the use of theory based on equation 5 in constructing confidence bands. More specifically, the tables report the difference between the theoretical values of η at the upper and lower 68%, 90%, 95%, and 99% confidence bands and the associated sorted empirical quantities based on 5000 simulated values for GARMA(0,0) models with $\eta = 0.50$ and 0.98. For each value of η, we allow λ to take on the values of $\{0.1, 0.2, 0.3, 0.4\}$, and as above, we consider several sample sizes ranging from 100 to 2000. As a reference, the theoretical bands for sample sizes of 500 observations are presented in bold font.

From Table 4, we see that the amplified empirical kurtosis is especially problematic for small T and λ. Generally speaking, the 99% bands appear to be uninformative when $\lambda = 0.10$, even for moderately large sample sizes. For $T = 500$, for example, theoretically, only 0.5% of estimated values of η should be less than 0.4267. In contrast, the empirical coverage areas differ substantially, as 0.5% of all values of $\hat{\eta}$ are less than 0.2644, yielding a difference between the theoretical and empirical lower confidence band equal to 0.1623. Although somewhat reliable results can be obtained for 68% bands and sample sizes of at least 1000, the results with $\lambda = 0.10$ show that the theory may need to be exercised with caution. We do note that the time series with $\lambda = 0.10$ might be viewed as extreme in light of the fact that they display characteristics that are difficult to distinguish from short memory. The theoretical first order autocorrelation coefficient, for example, is 0.1015 and, after the first lag, there is no ACF value in excess of 0.05 in absolute value.

Remaining results in Table 4 generally support the use of CSS theory for construction of confidence bands, especially when $T > 100$ and for narrower bands. Overall, biases decrease with both T and λ. For example, consider the model shown in the bottom panel with $\lambda = 0.40$ with $T = 2000$. Using the simulated values of Y_0 based on equation 5, the theoretical 95% bands for $\hat{\eta}$ are 0.4970 to 0.5030, implying a range for the Gegenbauer frequencies between 1.0437 and 1.0507. Assuming monthly data, these narrow bands, which result from a larger sample, would yield a periodicity between $2\pi/1.0507 = 5.98$ and $2\pi/1.0437 = 6.02$ months. For this process, 2.5% of the values, $\hat{\eta}$, were smaller than 0.4962 and and 2.5% were larger than 0.5040, thus yielding the associated biases in Table 4 equal to 0.0008 (0.4970-0.4962) and -0.0010 (0.5030-0.5040). The empirical coverage area would suggest periodicity between 5.975 and 6.027 months, such that the use of theoretical bands would provide appropriate estimates for researchers interested in providing statements regarding the uncertainty of estimated cycle lengths The results in Table 5 indicate generally small biases in calculating confidence intervals with asymptotic quantities decline further as $\eta \to 1$. For all cases, except when $\lambda = 0.10$ and T is smaller than 500, the empirical coverage areas are well captured by asymptotic quantities. Especially when λ is large, the differences between empirical percentiles and the theoretical quantities become negligible. Displayed in the final panel in Table 5, we present results with $\lambda = 0.40$ where $\eta = 0.9995$, a parameterization very close to a unit root. Except in the case of the 99% confidence bands with $T = 100$, the associated biases are never greater than 0.0013 and are essentially zero for $T > 1000$.

4. Hypothesis Testing for η

Overall, the Monte Carlo results indicate that the proposed distribution theory for maximum likelihood-based estimators works well in constructing confidence bands as $|\eta| \to 1$. However, the question remains as to how useful these results are for testing purposes. It would be straightforward to use the confidence bands about a hypothesized value η_0 to conduct inference for any given test size, α. Using the methods outlined in the previous section, for example, if $\hat{\eta}$ lies outside the $1-\alpha\%$ confidence bands about η_0, then we reject the hypothesis $\eta = \eta_0$. It is perhaps not surprising, given the results in Tables 4 and 5, that the tests for a specific value of η can be incorrectly sized, particularly when when λ is close to 0. For example, for a sample of 300 observations, with $\eta = 0.50$ and $\lambda = 0.20$, we found that 19.1% of the estimated values, $\hat{\eta}$, are outside the 95% confidence bands about the true $\eta = 0.50$. The remaining results, which are available upon request, do improve as λ and the sample size increase, but they highlight potential concerns with the use of Chung's theory.

Although the calculation of confidence bands is arguably the most important application of theory for researchers working with data strongly suspected to have finite cycles, there will be specific interest in many disciplines for determining if $|\eta| = 1$. For example, as recently emphasized by Dissanayake et al. (2018), there is specific interest in the hypothesis $H_0 : \eta = 1$ versus $H_A : \eta < 1$. It is precisely at this point where we anticipate potential inferential problems due to the discontinuity that exists in the proof of the proposed theory of Chung (1996a) at that point. These concerns are validated in Table 6, which provide simulation results for the CSS estimator when the true value of η is 1.

Turning to the specific percentile values in 6a, the empirical distribution of $T^2(\hat{\eta} - 1)$ does not match the asymptotic distribution of Chung (1996a). For example, with $\lambda = 0.20$, the value of the empirical 1^{st} percentile for $T^2(\hat{\eta} - 1)$ can be more than 11 times larger than the value implied by Chung. In other words, the empirical distribution is dramatically more skewed left than the theoretical results would imply. Further, the empirical distribution takes on fewer positive entries than the proposed asymptotic distribution.

To analyze how empirical and theoretical distributional disparities impact inference, we consider the proposed tests of Chung (1996a,b) in Table 6b for $|\eta| = 1$, when the true value of η is 1 or -1. The hypothesis can be tested by calculating the test statistic $T^2(|\eta| - 1)$. Rejection occurs when this quantity is less than the associated percentile for Y_1 standardized by $2\hat{\lambda}$ as in equation 6. The left-hand side of Table 6b reports the empirical size based on the 95% and 99% critical values when the true value of η is 1, while the right hand side of Table 6b presents the same results when $\eta = -1$. The table shows that the implementation of the proposed distribution theory under the null will result in very large size distortions, with only mild relief as the sample size increases. Consider the case where the generated data are ARFIMA processes with $\eta = 1$. Even with 2000 observations and a 5% test size, we reject H_0 14.8% of the time. In some cases, the rejection rates of the true null $\eta = 1$ can be larger than 19%. Considering a 1% test size, we still see that the rejection rates exceed 9.6%. Throughout, the results are slightly worse when $\eta = -1$. Even for large samples, these results suggest that the proposed distributional theory is unlikely to be useful to researchers interested in determining if the true data generating process is an ARFIMA or GARMA process. This can be especially problematic for those interested in testing for stationarity, where non-stationarity occurs for all values of $\lambda \geq 0.25$ when $|\eta| = 1$, but only occurs when $\lambda \geq 0.50$, otherwise. Clearly, more suitable testing procedures are needed.

In part, the empirical test-sizes for $\eta = 1$ could be inflated due to the factor T^2. A more conservative approach would be to assume rate T convergence and use the confidence bands about $\hat{\eta}$ and testing procedures under the alternative, $\eta < 1$ from equation (5). If, for example, the upper limit of these more conservative confidence bands is less than 1, we can feel more comfortable that $\eta = 1$ is rejected. Our simulation results (available upon request) show that, although more reliable rejection rates can be obtained, even this more conservative approach can sometimes provide inaccurate empirical test sizes. This approach is also somewhat unsatisfactory as the distribution theory developed in (5) explicitly excludes the point $\eta = 1$. It seems additional procedures are needed.

A more formal test has recently been advocated by Dissanayake et al. (2018), who discuss the use of quasi-likelihood ratio test statistics based on a state space representation of the GARMA model. Following a similar approach, one could test the null $\eta = 1$ by computing the value of the likelihood function in (4), with and without $\eta = 1$ imposed, and form the test statistic as follows,

$$LR = 2 \left[\max_{\phi', \theta', \lambda, \eta, \mu} L(\phi', \theta', \lambda, \eta, \mu) - \max_{\phi', \theta', \lambda, \mu} L(\phi', \theta', \lambda, 1, \mu) \right]. \tag{11}$$

Under H_0, η lies on the boundary of the parameter space, and given the difficulties discussed above, it seems likely that the distribution of the test statistic will be non-standard. In light of the problems with the existing theory, computational methods may be preferred.

As a reasonable approach, one could use the estimated model under H_0 to form a parametric bootstrap. Data of the desired length could be simulated based on sampling with replacement from the residuals of the null model, and the critical values of the distribution could be formed based on the test statistic in (11). This procedure was applied to the unemployment rate data described in Section 2, using the residuals from an estimated ARFIMA(2,1) model. Results based on 5000 simulations are presented in Table 7.[7] Here, we see that the likelihood ratio test statistic associated with the hypothesis $\eta = 1$ takes on a value of 6.31, exceeding simulated critical values at the 5%, but not at the 1% levels. As discussed above, these findings are likely to be useful to applied researchers in economics, since the null model is non-stationary, but strong stationary cycles result under the alternative.

The results provide support favoring $|\eta| < 1$ and also highlight the difficulties that could be encountered for researchers employing CSS theory. More specifically, theoretical confidence bands constructed under the null hypothesis $|\eta| = 1$ suffer from such severe size distortion they are likely uninformative. For the example here, a more careful testing analysis shows a marginal rejection of $\eta = 1$, whereas strong rejection occurs for virtually any size when using the theory of Chung (1996a,b) based on the assumed T^2 rate of convergence.

5. Conclusions

In this manuscript, we analyzed the performance of two parametric estimators of the GARMA model parameters, including the CSS and frequency-based Whittle estimators. These methods, which are based on approximations to the Gaussian log-likelihood function, have at least partial distributional results that have been proposed by Chung (1996a,b) and Giraitis et al. (2001). Here, for the CSS method, we were specifically interested in studying the proposed inference of Chung (1996a,b) for the parameter, η, governing the length of long memory cycles when using an algorithm that searches over all possible values.

An extensive Monte Carlo analysis revealed that both estimators are highly robust in estimating model parameters. As the CSS estimator admits a continuous parameter space, it was found to be superior in estimation of η compared to the Whittle method. The empirical results also showed support for the proposed distribution theory of Chung (1996a,b) for stationary parameterizations in a neighborhood of a unit root process, especially for larger samples. Additionally, in most cases, the distribution theory was shown to be quite useful in constructing confidence bands, where empirical coverage areas are generally close to theoretical asymptotic quantities.

The findings are not without important caveats. For the CSS estimator, in smaller samples, empirical distributions can have larger kurtosis than implied by theory. Although this may be a small sample problem, the simulation results showed that constructed wider confidence bands are likely to be uninformative under very weak long memory and in samples less than 500 observations. From a practical perspective, we would not recommend the use of the proposed asymptotic theory for very small samples when weak persistence is suspected, perhaps as evidenced by a memory parameter, $\hat{\lambda}$, in the neighborhood of zero.

From a testing perspective, we found that there are some concerns in using the existing asymptotic results. In particular, we found that when $|\eta| = 1$, tests based on the theory of Chung (1996a) can yield severe size distortion and may cause confusion in determining whether a series is a stationary GARMA process or a non-stationary ARFIMA/unit root process. With rejection rates under the null approaching 15% for a 5% test size in samples of about 2000, this does not appear to be a minor problem.

The paper concluded with an application to monthly unemployment rates, where evidence using a parametric bootstrap provided support for the existence of stationary, long memory cycles in the labor market.

[7] Given the non-stationarity of the null model, the data are simulated recursively using the coefficients of the expansion of $(1 - L)^{-0.6602}$ and the standard assumption that pre-sample observations are 0.

The overall conclusions suggest that GARMA parameters can be well estimated by existing likelihood-based techniques. For researchers interested in obtaining confidence bands for finite cycle lengths, the existing theory also appears to be largely applicable. Nonetheless, the results also showed there are still several gaps in the existing theory that merit additional exploration, specifically as it relates to testing for values of $|\eta|$ in the neighborhood of 1. In lieu of additional theoretical distribution results, we propose that computational methods such as our bootstrap approach are likely to be the most informative for researchers interested in statistical inference for these cases.

References

References

Arteche, J. and Robinson, P. M. (2000). Semiparametric inference in seasonal and cyclical long memory processes. *Journal of Time Series Analysis*, 21(1):1–25.

Artiach, M. and Arteche, J. (2012). Doubly fractional models for dynamic heteroscedastic cycles. *Computational Statistics & Data Analysis*, 56(6):2139 – 2158.

Asai, M., Peiris, S., McAleer, M., and Allen, D. (2018). Cointegrated dynamics for a generalized long memory process: An application to interest rates. Technical Report EI2018-32, Econometric Institute Research Papers, University of Madrid.

Beaumont, P. M. and Smallwood, A. D. (2019). Constrained sum of squares estimation of multiple frequency long memory models. Working Paper, Florida State University.

Caporale, G. M. and Gil-Alana, L. (2014). Long-run and cyclical dynamics in the us stock market. *Journal of Forecasting*, 33(2):147–161.

Caporale, G. M. and Gil-Alana, L. A. (2011). Multi-factor gegenbauer processes and european inflation rates. *Journal of Economic Integration*, 26(2):386–409.

Chung, C.-F. (1996a). Estimating a generalized long memory process. *Journal of Econometrics*, 73(1):237 – 259.

Chung, C.-F. (1996b). A generalized fractionally integrated autoregressive movingaverage process. *Journal of Time Series Analysis*, 17:111–140.

Diongue, A. K. and Ndongo, M. (2016). The k-factor garma process with infinite variance innovations. *Communications in Statistics-Simulation and Computation*, 45(2):420–437.

Dissanayake, G. S., Peiris, M. S., and Proietti, T. (2018). Fractionally differenced gegenbauer processes with long memory: A review. *Statistical Science*, 33(3):413–426.

Ferrara, L. and Guégan, D. (2001). Forecasting with k-factor gegenbauer processes: Theory and applications. *Journal of Forecasting*, 20(8):581–601.

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

Gil-Alanaa, L. A. (2007). Testing the existence of multiple cycles in financial and economic time series. *Annals of Economics & Statistics*, 8(1):1–20.

Giraitis, L., Hidalgo, J., and Robinson, P. M. (2001). Gaussian estimation of parametric spectral density with unknown pole. *The Annals of Statistics*, 29(4):987–1023.

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(1):15–24.

Gray, H. L., Zhang, N., and Woodward, W. A. (1989). On generalized fractional processes. *Journal of Time Series Analysis*, 10:233–257.

Hidalgo, J. (2005). Semiparametric estimation for stationary processes whose spectra have an unknown pole. *The Annals of Statistics*, 33(4):1843–1889.

Hidalgo, J. and Soulier, P. (2004). Estimation of the location and exponent of the spectral singularity of a long memory process. *Journal of Time Series Analysis*, 25(1):55–81.

Hosking, J. R. M. (1981). Fractional differencing. *Biometrika*, 68:165–76.

Leschinski, C. and Sibbertsen, P. (2019). Model order selection in periodic long memory models. *Econometrics and Statistics*, 9:78 – 94.

Lu, Z. and Guegan, D. (2011). Estimation of time-varying long memory parameter using wavelet method. *Communications in Statistics—Simulation and Computation®*, 40(4):596–613.

McElroy, T. S. and Holan, S. H. (2012). On the computation of autocovariances for generalized gegenbauer processes. *Statistica Sinica*, 22(4):1661–1687.

Nielsen, M. Ø. and Frederiksen, P. H. (2005). Finite sample comparison of parametric, semiparametric, and wavelet estimators of fractional integration. *Econometric Reviews*, 24(4):405–443.

Peiris, M. and Asai, M. (2016). Generalized fractional processes with long memory and time dependent volatility revisited. *Econometrics*, 4(3):1–21.

Ramachandran, R. and Beaumont, P. (2001). Robust estimation of garma model parameters with an application to cointegration among interest rates of industrialized countries. *Computational Economics*, 17(2/3):179–201.

Reisen, V. A., Zamprogno, B., Palma, W., and Arteche, J. (2014). A semiparametric approach to estimate two seasonal fractional parameters in the sarfima model. *Mathematics and Computers in Simulation*, 98:1 – 17.

Robinson, P. M. (1995). Log-periodogram regression of time series with long range dependence. *The Annals of Statistics*, 23(3):1048–1072.

Woodward, W. A., Cheng, Q. C., and Gray, H. L. (1998). A k-factor garma long-memory model. *Journal of Time Series Analysis*, 19:485–504.

Table 1: Bias and RMSE for Whittle and CSS estimates of GARMA(0,0) model parameters

T	CSS η Bias	GHR η Bias	CSS λ Bias	GHR λ Bias	CSS η RMSE	GHR η RMSE	CSS λ RMSE	GHR λ RMSE	Chung Asymp
Model #1: GARMA(0,0); $\eta = 1$, $\lambda = 0.20$									
100	**-0.00644**	-0.00659	**0.00564**	-0.03109	0.0262	**0.0237**	**0.0444**	0.0591	0.0390
300	**-0.00089**	-0.00093	**0.00311**	-0.01106	0.0058	**0.0043**	**0.0251**	0.0278	0.0225
500	**-0.00027**	-0.00028	**0.00228**	-0.00630	0.0017	**0.0012**	**0.0194**	0.0199	0.0174
1000	**-0.00006**	-0.00008	**0.00173**	-0.00320	**0.0003**	0.0004	**0.0131**	0.0133	0.0123
2000	**-0.00002**	-0.00002	0.00210	**-0.00158**	**0.0001**	0.0002	0.0095	**0.0092**	0.0087
Model #2: GARMA(0,0); $\eta = -1$, $\lambda = 0.20$									
100	0.00724	**0.00624**	**0.00498**	-0.03160	0.0277	**0.0248**	**0.0448**	0.0593	0.0390
300	**0.00090**	0.00091	**0.00306**	-0.01025	0.0052	**0.0039**	**0.0252**	0.0276	0.0225
500	0.00031	**0.00027**	**0.00249**	-0.00670	0.0019	**0.0012**	**0.0191**	0.0205	0.0174
1000	**0.00007**	0.00009	**0.00183**	-0.00342	**0.0004**	0.0004	**0.0134**	0.0138	0.0123
2000	0.00002	**0.00002**	**0.00167**	-0.00176	**0.0002**	0.0002	0.00935	**0.0092**	0.0087
Model #3: GARMA(0,0); $\eta = 0.9995$, $\lambda = 0.40$									
100	**-0.00070**	-0.00118	0.01696	**0.00227**	**0.0029**	0.0044	**0.0470**	0.0546	0.0396
300	**-0.00008**	-0.00018	0.01057	**0.00999**	**0.0005**	0.0008	**0.0274**	0.0316	0.0229
500	**-0.00002**	-0.00006	**0.00809**	0.00918	**0.0002**	0.0004	**0.0213**	0.0245	0.0177
1000	**1.18E-06**	-1.99E-06	0.00562	**-0.00107**	**0.0001**	0.0001	0.0149	**0.0133**	0.0125
2000	**1.59E-06**	2.22E-06	0.00495	**-0.00007**	**0.0001**	0.0001	0.0108	**0.0092**	0.0089
Model #4: GARMA(0,0); $\eta = -0.9995$, $\lambda = 0.40$									
100	**0.00089**	0.00120	0.01566	**0.00180**	**0.0039**	0.0042	**0.0453**	0.0458	0.0390
300	**0.00009**	0.00014	0.01049	**0.00929**	**0.0006**	0.0007	**0.0276**	0.0309	0.0229
500	**0.00002**	0.00006	**0.00823**	0.00839	**0.0003**	0.0004	**0.0214**	0.0243	0.0177
1000	1.63E-06	**1.43E-06**	0.00553	**-0.00123**	**0.0001**	0.0001	0.0149	**0.0131**	0.0125
2000	6.80E-07	**-3.76E-06**	0.00452	**-0.00022**	**0.0001**	0.0001	0.0106	**0.0093**	0.0089
Model #5: GARMA(0,0); $\eta = 0.50$, $\lambda = 0.40$									
100	**0.00102**	-0.00587	0.02535	**0.00602**	**0.0298**	0.0437	**0.0805**	0.1050	0.0675
300	**0.00029**	0.00051	**0.01620**	-0.01930	**0.0102**	0.0114	**0.0458**	0.0480	0.0390
500	**0.00001**	0.00152	0.01245	**0.00467**	**0.0060**	0.0079	**0.0361**	0.0387	0.0302
1000	**-0.00002**	-0.00080	0.00754	**0.00472**	**0.0030**	0.0038	**0.0248**	0.0268	0.0214
2000	**-0.00002**	0.00036	0.00521	**0.00397**	**0.0016**	0.0020	**0.0170**	0.0630	0.0151
Model #6: GARMA(0,0); $\eta = -0.50$, $\lambda = 0.40$									
100	**-0.00002**	0.00621	0.02950	**-0.02176**	**0.0313**	0.0429	**0.0812**	0.1039	0.0675
300	**-0.00006**	-0.00016	**0.01535**	-0.01945	**0.0102**	0.0113	**0.0447**	0.0490	0.0390
500	**0.00002**	-0.00151	0.01112	**0.00245**	**0.0063**	0.0080	**0.0346**	0.0374	0.0302
1000	**-0.00005**	0.00063	0.00736	**0.00456**	**0.0031**	0.0039	**0.0239**	0.0260	0.0214
2000	**-0.00001**	-0.00035	0.00460	**0.00328**	**0.0016**	0.0020	0.0166	**0.0020**	0.0151

Notes: Results are based on 2500 simulations for various sample sizes. The asymptotic standard error for λ using the theory of Chung (1996a) is presented in the last column. Quantities appearing in bold font indicate the smaller bias/RMSE between the CSS and GHR estimators.

Yajima, Y. (1996). Estimation of the frequency of unbounded spectral densities. ASA Proceedings Business and Economic Statistics Section 4-7, American Statistical Association.

Table 2: Bias and RMSE for Whittle and CSS estimates of GARMA(1,0) model parameters

T	CSS: η	GHR: η	CSS: λ	GHR: λ	CSS: ϕ	GHR: ϕ	Chung
	Model #7: GARMA(1,0); $\eta = 0.50$, $\lambda = 0.40$, $\phi = 0.80$						(Asymp λ)
	Parameter Bias						
100	**-0.00058**	-0.00901	**0.01937**	-0.05005	-0.02735	**-0.02068**	0.06779
300	-0.00036	**-0.00005**	**0.01320**	-0.02692	-0.01035	**-0.00357**	0.03914
500	**-0.00016**	0.00119	0.00999	**-0.00106**	**-0.00683**	-0.00724	0.03032
1000	**-0.00001**	-0.00086	0.00710	**0.00249**	-0.00348	**-0.00319**	0.02144
2000	**-0.00001**	0.00037	0.00655	**0.00488**	**-0.00146**	-0.00150	0.01516
	Model #7: GARMA(1,0); $\eta = 0.50$, $\lambda = 0.40$, $\phi = 0.80$						(Asymp ϕ)
	Parameter RMSE						
100	**0.03235**	0.04947	**0.07735**	0.11326	**0.07617**	0.07828	0.06024
300	**0.01075**	0.01152	**0.04481**	0.05360	0.03854	**0.03769**	0.03478
500	**0.00641**	0.00784	**0.03435**	0.03796	**0.02918**	0.03093	0.02694
1000	**0.00334**	0.00408	**0.02422**	0.02584	**0.01951**	0.02004	0.01905
2000	**0.00161**	0.00201	**0.01736**	0.01867	**0.01369**	0.01394	0.01347
T	CSS: η	GHR: η	CSS: λ	GHR: λ	CSS: ϕ	GHR: ϕ	Chung
	Model #8: GARMA(1,0); $\eta = 0.9995$, $\lambda = 0.40$, $\phi = 0.80$						(Asymp λ)
	Parameter Bias						
100	**-0.01416**	-0.02782	**-0.20227**	-0.30174	**0.09923**	0.10871	0.12635
300	**-0.00223**	-0.00796	-0.15377	**-0.11110**	**0.09992**	-0.11546	0.07295
500	**-0.00065**	-0.00337	-0.12132	**-0.01833**	**0.07958**	-0.24442	0.05651
1000	**-0.00009**	-0.00066	-0.08719	**0.00541**	**0.06242**	-0.13533	0.03996
2000	**0.000002**	-0.00004	**0.01609**	0.01929	**-0.02224**	-0.10542	0.02825
	Model #8: GARMA(1,0); $\eta = 0.9995$, $\lambda = 0.40$, $\phi = 0.80$						(Asymp ϕ)
	Parameter RMSE						
100	**0.03787**	0.05880	**0.24362**	0.33197	0.22420	**0.20178**	0.06024
300	**0.00981**	0.01753	**0.19453**	0.24230	**0.15666**	0.41871	0.03478
500	**0.00466**	0.00822	**0.16658**	0.20506	**0.14172**	0.48856	0.02694
1000	**0.00162**	0.00255	**0.12891**	0.14197	**0.12002**	0.33196	0.01905
2000	**0.00006**	0.00037	**0.04347**	0.10770	**0.07160**	0.27240	0.01347

Notes: Results are based on 2500 simulations for various sample sizes. The asymptotic standard error for λ using the theory of Chung (1996a) is presented in the first panel for both models and in the last column. The analogous quantity for ϕ is presented in the second panel, and again in the last column. Quantities appearing in bold font indicate a smaller bias/RMSE between the CSS and GHR estimators.

Table 3: Distribution of $T(\hat{\eta} - \eta)$ at selected percentiles

Model: GARMA(0,0) with $\eta = 0.9995$, $\lambda = 0.40$										
	0.001	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990	0.995
Chung	**-0.335**	**-0.284**	**-0.220**	**-0.172**	**-0.124**	**0.124**	**0.172**	**0.220**	**0.284**	**0.335**
CSS/100	-1.639	-1.268	-0.689	-0.386	-0.169	0.047	0.067	0.092	0.120	0.140
GHR/100	-3.092	-1.721	-0.739	-0.739	-0.147	0.050	0.050	0.050	0.050	0.050
CSS/300	-1.002	-0.628	-0.319	-0.194	-0.096	0.066	0.115	0.149	0.166	0.182
GHR/300	-1.493	-0.902	-0.442	-0.442	-0.113	0.084	0.150	0.150	0.150	0.150
CSS/500	-0.722	-0.431	-0.275	-0.155	-0.075	0.071	0.137	0.202	0.247	0.260
GHR/500	-0.737	-0.737	-0.382	-0.382	-0.105	0.092	0.211	0.211	0.250	0.250
CSS/1000	-0.460	-0.368	-0.209	-0.122	-0.072	0.079	0.152	0.229	0.345	0.494
GHR/1000	-0.763	-0.467	-0.211	-0.211	0.007	0.007	0.184	0.184	0.322	0.421
CSS/2000	-0.369	-0.276	-0.189	-0.118	-0.064	0.064	0.126	0.236	0.343	0.461
GHR/2000	-0.421	-0.421	-0.194	-0.194	-0.194	0.013	0.201	0.201	0.368	0.516
Model: GARMA(1,0) with $\eta = 0.50$, $\lambda = 0.40$, $\phi = 0.80$										
	0.001	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990	0.995
Chung	**-9.176**	**-7.786**	**-6.012**	**-4.700**	**-3.399**	**3.399**	**4.700**	**6.012**	**7.786**	**9.176**
CSS/100	-14.111	-10.745	-6.949	-4.284	-2.348	2.226	4.476	7.491	10.681	12.771
GHR/100	-19.098	-13.188	-7.422	-7.422	-1.825	3.583	3.583	8.779	13.7421	22.897
CSS/300	-13.187	-11.189	-7.236	-4.840	-2.388	2.048	4.125	6.488	10.263	12.863
GHR/300	-11.011	-11.011	-5.474	-5.474	0.000	0.000	5.408	5.408	10.748	16.018
CSS/500	-14.513	-10.683	-6.761	-4.417	-2.164	1.975	3.966	6.564	10.153	13.467
GHR/500	-14.648	-14.648	-9.123	-3.636	-3.636	1.812	7.220	7.220	12.587	12.587
CSS/1000	-12.619	-9.842	-6.535	-3.972	-2.139	2.107	4.117	6.774	9.783	13.259
GHR/1000	-12.750	-12.750	-7.273	-7.273	-1.815	3.623	3.623	9.041	9.041	14.440
CSS/2000	-15.543	-11.177	-6.657	-3.872	-2.046	2.043	3.752	6.869	10.212	13.488
GHR/2000	-14.545	-9.083	-9.083	-3.630	-3.630	1.813	7.246	7.246	12.670	12.670

Notes: Results are based on 2500 simulations. In bold font, we present the values at theoretical percentiles for the test statistic using equation (5) and the associated simulated quantities for Y_0 from Chung (1996a). The remaining elements yield the associated values at a given percentile for $\hat{\eta}$ using the CSS and GHR estimators.

Table 4: Bias in confidence intervals for η for GARMA(0,0) models with $\eta = 0.50$

	68%L	68%U	90%L	90%U	95%L	95%U	99%L	99%U
Theory: 500	**0.4799**	**0.5201**	**0.4623**	**0.5376**	**0.4517**	**0.5482**	**0.4267**	**0.5733**
				BIAS WITH $\lambda = 0.10$				
CSS/100	0.2504	-0.1563	0.4764	-0.2420	0.6080	-0.2297	0.8701	-0.1318
CSS/300	0.0518	-0.0627	0.1154	-0.1123	0.1753	-0.1429	0.4368	-0.2568
CSS/500	0.0328	-0.0361	-0.0641	-0.0657	0.0830	-0.0766	0.1623	-0.1394
CSS/1000	0.0140	-0.0178	-0.0290	-0.0317	0.0354	-0.0378	0.0469	-0.0499
CSS/2000	0.0077	-0.0090	-0.0151	-0.0158	0.0175	-0.0194	0.0231	-0.0273
Theory: 500	**0.4900**	**0.5101**	**0.4812**	**0.5188**	**0.4759**	**0.5241**	**0.4634**	**0.5366**
				BIAS WITH $\lambda = 0.20$				
CSS/100	0.0110	-0.0251	0.0858	-0.1001	0.1279	-0.1321	0.2952	-0.2209
CSS/300	0.0045	-0.0092	-0.0303	-0.0336	0.0435	-0.0435	0.0655	-0.0674
CSS/500	0.0032	-0.0056	0.0166	-0.0201	0.0239	-0.0248	0.0348	-0.0396
CSS/1000	0.0026	-0.0029	0.0098	-0.0093	0.0122	-0.0122	0.0198	-0.0174
CSS/2000	0.0016	-0.0018	0.0055	-0.0053	0.0068	-0.0067	0.0105	-0.0085
Theory: 500	**0.4933**	**0.5067**	**0.4874**	**0.5125**	**0.4839**	**0.5161**	**0.4756**	**0.5244**
				BIAS WITH $\lambda = 0.30$				
CSS/100	-0.0161	0.0162	0.0044	-0.0095	0.0200	-0.0326	0.0553	-0.0731
CSS/300	-0.0042	0.0041	0.0050	-0.0046	-0.0115	-0.0118	0.0256	-0.0214
CSS/500	-0.0019	0.0017	0.0019	-0.0040	0.0064	-0.0080	0.0157	-0.0168
CSS/1000	-0.0006	0.0005	0.0022	-0.0020	0.0040	-0.0040	0.0081	-0.0077
CSS/2000	0.0001	-0.0001	0.0018	-0.0018	0.0027	-0.0028	0.0043	-0.0042
Theory: 500	**0.4950**	**0.5050**	**0.4906**	**0.5094**	**0.4879**	**0.5120**	**0.4817**	**0.5183**
				BIAS WITH $\lambda = 0.40$				
CSS/100	-0.0181	0.0195	-0.0301	0.0305	-0.0266	0.0269	-0.0151	-0.0064
CSS/300	-0.0054	-0.0057	-0.0076	0.0075	-0.0053	0.0040	0.0037	-0.0034
CSS/500	-0.0030	0.0031	-0.0036	0.0028	-0.0016	0.0013	0.0022	-0.0022
CSS/1000	-0.0012	0.0012	-0.0006	0.0007	0.0004	-0.0001	0.0027	-0.0011
CSS/2000	-0.0003	0.0002	0.0004	-0.0006	0.0008	-0.0010	0.0022	-0.0023

Notes: The table reports the difference between the value of η associated with theoretical confidence bands, which have been constructed using equation (5) along with simulated values for Y_0, and the estimated value of η associated with a given percentile. 68%L and 68%U refer to 68% lower and upper confidence bands, with similar meaning for other quantities. Values appearing in bold font in the rows labeled "Theory" are the actual values of η that are generated from equation (5) for a sample size of 500 observations. Remaining rows show the bias (theoretical minus estimated) in the lower and upper confidence intervals of η for various sample sizes. To obtain more precision, results are based on 5000 simulations as compared to 2500 in other tables.

Table 5: Bias in confidence intervals for η for GARMA(0,0) models with $\eta = 0.98$

	68%L	68%U	90%L	90%U	95%L	95%U	99%L	99%U
Theory: 500	**0.9754**	**0.9846**	**0.9713**	**0.9866**	**0.9689**	**0.9911**	**0.9632**	**0.9968**
BIAS WITH $\lambda = 0.10$								
CSS/100	0.1104	0.0063	0.9097	0.0145	0.9732	-0.0211	1.1108	0.0267
CSS/300	0.0151	-0.0059	0.0336	-0.0052	0.0478	-0.0015	0.0987	0.0078
CSS/500	0.0079	-0.0066	0.0172	-0.0089	0.0218	-0.0079	0.0358	-0.0031
CSS/1000	0.0036	-0.0036	0.0073	-0.0062	0.0099	-0.0068	-0.0158	-0.0071
CSS/2000	0.0018	-0.0018	-0.0036	-0.0032	0.0043	-0.0037	0.0061	-0.0051
Theory: 500	**0.9777**	**0.9823**	**0.9757**	**0.9843**	**0.9745**	**0.9855**	**0.9716**	**0.9884**
BIAS WITH $\lambda = 0.20$								
CSS/100	0.0105	0.001	0.0047	0.0036	0.0760	0.0081	0.2348	0.0206
CSS/300	0.0023	-0.0014	0.0090	-0.0049	0.0129	-0.0057	0.0235	-0.0047
CSS/500	0.0013	-0.0010	0.0055	-0.0037	0.0067	-0.0045	0.0120	-0.0051
CSS/1000	0.0007	-0.0006	0.0025	-0.0020	0.0035	-0.0027	0.0056	-0.0035
CSS/2000	0.0004	-0.0004	0.0013	-0.0012	0.0017	-0.0016	0.0027	-0.0021
Theory: 500	**0.9785**	**0.9815**	**0.9771**	**0.9829**	**0.9763**	**0.9837**	**0.9744**	**0.9856**
BIAS WITH $\lambda = 0.30$								
CSS/100	-0.0030	0.0041	0.0047	0.0028	0.0107	0.0020	0.0387	0.0081
CSS/300	-0.0009	0.0009	0.0015	-0.0007	0.0036	-0.0020	0.0071	-0.0028
CSS/500	-0.0004	0.0005	0.0012	-0.0008	0.0024	-0.0015	0.0047	-0.0030
CSS/1000	-0.0001	0.0001	0.0007	-0.0005	0.0011	-0.0011	0.0021	-0.0019
CSS/2000	0.00002	-0.00002	0.0005	-0.0004	0.0007	-0.0006	0.0013	-0.0012
Theory: 500	**0.9788**	**0.9812**	**0.9778**	**0.9822**	**0.9772**	**0.9828**	**0.9758**	**0.9842**
BIAS WITH $\lambda = 0.40$								
CSS/100	-0.0040	0.0043	-0.0068	0.0072	-0.0064	0.0070	0.0034	0.0059
CSS/300	-0.0012	0.0013	-0.0016	0.0018	-0.0007	0.0010	0.0005	-0.0005
CSS/500	-0.0007	0.0007	-0.0006	0.0009	-0.0001	0.0005	0.0004	-0.0002
CSS/1000	-0.0003	0.0003	-0.0001	0.0003	0.0001	0.0001	0.0009	-0.0005
CSS/2000	-0.0001	0.0001	0.0001	-0.0001	0.0002	-0.0002	0.0007	-0.0004
GARMA(0,0) model $\eta = 0.9995$, $\lambda = 0.40$								
Theory: 500	**0.9900**	**0.9997**	**0.9992**	**0.9998**	**0.9991**	**0.9999**	**0.9998**	**1.0002**
BIAS FOR MODEL NEAR UNIT ROOT BOUNDARY								
CSS/100	-0.0006	0.0004	-0.0008	0.0009	0.0001	0.0011	0.0066	0.0017
CSS/300	-0.0002	0.0002	-0.0003	0.0002	0.00002	0.0003	0.0013	0.0005
CSS/500	-0.0001	0.0001	-0.0001	0.0001	0.00004	0.0001	0.0002	0.0002
CSS/1000	-0.00005	0.00003	-0.00004	-0.00003	0.00001	-0.00003	0.0001	-0.00004
CSS/2000	-0.00002	0.000001	0.00001	-0.00003	0.00003	-0.00004	0.00009	-0.00006

Notes: The table provides the bias in estimated confidence intervals for CSS estimates of η for GARMA(0,0) models with $\eta = 0.98$, in the first four panels, and $\eta = 0.9995$ in the bottom panel. Values appearing in bold font in the rows labeled "Theory" are the actual values of η that are generated from equation (5) for a sample size of 500 observations. Remaining rows show the bias (theoretical minus estimated) in the estimated lower and upper confidence intervals for η for various sample sizes, again based on 5000 simulations. 68%L and 68%U, for example, refer to 68% lower and upper confidence bands.

Table 6: Results for $|\eta| = 1$

(a) The empirical distributions of $T(\hat{\eta}-1)$, $T^2(\hat{\eta}-1)$

MODEL: GARMA(0,0): $\eta = 1.00$, $\lambda = 0.20$										
Percentile	0.001	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990	0.995
Chung	**-244.44**	**-185.29**	**-121.23**	**-81.00**	**-75.13**	**10.14**	**15.44**	**21.53**	**31.18**	**39.85**
CSS:T=100										
$T(\hat{\eta}-1)$	-17.791	-13.262	-7.194	-3.825	-1.238	0.074	0.101	0.133	0.167	0.211
$T^2(\hat{\eta}-1)$	-1779.1	-1326.2	-719.39	-382.49	-123.84	7.365	10.106	13.297	16.654	21.114
CSS:T=300										
$T(\hat{\eta}-1)$	-9.974	-6.843	-2.328	-0.846	-0.297	0.024	0.032	0.042	0.053	0.060
$T^2(\hat{\eta}-1)$	-2992.3	-2052.8	-698.42	-253.80	-89.105	7.306	9.543	12.469	15.824	18.050
CSS:T=500										
$T(\hat{\eta}-1)$	-5.020	-3.332	-1.116	-0.494	-0.1774	0.014	0.019	0.024	0.031	0.037
$T^2(\hat{\eta}-1)$	-2510.1	-1666.2	-557.87	-246.83	-88.696	7.041	9.350	11.796	15.253	18.704
CSS:T=1000										
$T(\hat{\eta}-1)$	-2.072	-1.4230	-0.593	-0.286	-0.087	0.007	0.009	0.012	0.015	0.017
$T^2(\hat{\eta}-1)$	-2072.1	-1429.8	-592.49	-286.06	-86.90	6.839	9.208	11.978	14.823	17.289
CSS:T=2000										
$T(\hat{\eta}-1)$	-0.911	-0.644	-0.247	-0.010	-0.034	0.003	0.005	0.006	0.007	0.009
$T^2(\hat{\eta}-1)$	-1821.5	-1288.0	-493.68	-199.01	-68.111	6.656	8.930	11.097	14.235	18.194

Notes: The results in 6a provide the empirical distribution values at selected percentiles when $\eta = 1$ for a GARMA(0,0) model with $\lambda = 0.20$, based on both an assumed convergence rate of T and T^2. The row labeled "Chung" in bold font shows the theoretical values for the test statistic $T^2(\hat{\eta}-1)$ using equation (6) and the associated simulated quantities for Y_1 from Chung (1996a). The remaining rows show the computed distribution values using the CSS estimator for various sample sizes and 2500 simulations.

(b) Rejection rates of the hypothesis $\eta = 1$ (left panel) and $\eta = -1$ (right panel)

	$\eta = 1,\ \lambda = 0.20$		$\eta = -1,\ \lambda = 0.20$	
Sample Size	99%	95%	99%	95%
100	0.1344	0.1960	0.1432	0.1980
300	0.1112	0.1584	0.1236	0.1896
500	0.1108	0.1608	0.1296	0.1888
1000	0.1100	0.1572	0.1156	0.1632
2000	0.0960	0.1480	0.1072	0.1620

Notes: The data are generated with $\eta = 1$ or $\eta = -1$. The results in 6b are from one-sided tests based on the test statistic given by $T^2(|\hat{\eta}|-1)$. Associated critical values are calculated using the percentiles of Y_1 from Chung (1996a) based on equation (6) and the estimated values of λ, using the CSS estimator. Quantities here indicate the proportion of occurrences based on 2500 simulations where the associated test statistics are less than the calculated 99% and 95% critical values.

Table 7: Bootstrapped likelihood ratio test for $H_0 : \eta = 1$ vs. $H_A : \eta < 1$

	η	λ	ϕ_1	ϕ_2	θ	SSR	CSS
ARFIMA:	1	0.3301	0.9234	0.0292	-0.5481	11.613	207.421
GARMA:	0.9986	0.3263	0.9341	0.0263	-0.5674	11.459	210.574
Bootstrapped test statistic for $H_0 : \eta = 1$							
Test stat: $2(CSS_u - CSS_r) = 6.3067(0.0224)$							
Bootstrapped Critical Values							
		1%	5%	10%			
		3.7671	5.0200	7.8154			

Notes: For the unemployment rate, u_t, the ARFIMA model has been obtained with the restriction $\eta = 1$ imposed. Given the functional relationship $d = 2\lambda$, the model can be written as:

$$(1 - L)^{0.6602}(1 - 0.923L - 0.029L^2)(u_t - 6.20) = (1 - 0.548L)\varepsilon_t.$$

"CSS" denotes the value of the likelihood function in (4), with and without the restriction $\eta = 1$ imposed. The test statistic has been calculated as in equation (11). Critical values have been calculated on the basis of a parametric bootstrap, where the residuals from the ARFIMA model above have been randomly sampled with replacement to generate 5000 samples of 475 observations using the parameters of the ARFIMA model. For computational purposes, the set of starting values for η range from 0.925 to 1, where no constraints have been imposed on the parameter space. Using the estimated values of the CSS function with and without the constraints imposed, the test statistic in (11) has been calculated, and the associated critical values are presented in the final panel.

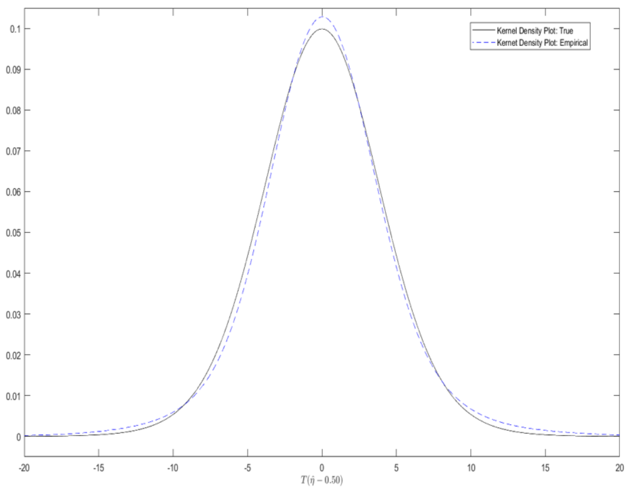

Figure 1: Theoretical and Empirical Kernel Density Plot of $T(\hat{\eta} - \eta)$ for $T = 2000$ and $\eta = 0.50$.

Notes: We use a Gaussian smoothing window and a bandwidth parameter of 3. To calculate the theoretical density, one needs the associated percentiles of Y_0 from (5). These values have been simulated using equation 25 in Chung (1996a) using MATLAB code available upon request.