Multiple Frequency Long Memory Models

Aaron D. Smallwood[*]
Department of Economics
University of Oklahoma

Paul M. Beaumont[†]
Department of Economics
Florida State University.

August, 2004.

Abstract

There has been considerable interest recently in long memory models able to capture complex features in both the spectral density and autocorrelation functions of time series data. We present a multiple frequency Gegenbauer autoregressive moving average, or k-factor GARMA, model and we derive the asymptotic distributions for a conditional sum of squares estimator of the parameters of this model. The parameters that determine the cycle lengths of the process are asymptotically independent, but the differencing parameters associated with these cycle lengths are not. Furthermore, the parameters of the model may converge at rates of $T^{0.5}$, T or T^2 depending upon the cycle lengths in the model. Our estimator exploits these theoretical properties and is able to produce unbiased and efficient estimators even in fairly small samples. We present Monte Carlo simulation results to establish the small sample properties of our estimator. A particularly interesting case of the k-factor GARMA model is when one frequency is consistent with a fractionally integrated process and another frequency is consistent with a GARMA process. This produces a process whose autocorrelation function has a sinusoidal pattern oscillating about a monotonically decaying trend, both of which are decaying at hyperbolic rates. This pattern is common in physical and financial data and is very difficult to reproduce with any other parsimonious model. We show that IBM trading volume follows such a process and we present a 2-factor GARMA model for this time series.

Keywords: GARMA; k-factor GARMA; Gegenbauer process; ARFIMA; Fractional Integration; conditional sum of squares estimation; asymptotic distribution.

JEL classification: C22

---

[*] Norman, OK 73019. *E-mail:* Aaron.Smallwood-1@ou.edu. *Phone:* (405) 325-2643. *Fax:* (405) 325-5842.

[†] Tallahassee, FL 32306. *E-mail:* Paul.Beaumont@fsu.edu. *Phone:* (850) 644-7085. *Fax:* (850) 644-4535.

1 Introduction

The seminal work in long memory models by Granger and Joyeux (1980) and Hosking (1981) has invigorated interest in parsimonious time series models that are capable of reproducing complicated stochastic properties of data in both the time and frequency domains. Recently, Gray, Zhang and Woodward (1989), Chung (1996b), Bierens (2001), Ferrara and Guegan (2001), Ramachandran and Beaumont (2001), and Caporale and Gil-Alana (2004), among others, discuss cyclical long memory models with one or more sources of long memory. These models are quite general and are potentially very useful for modelling many physical, economic and financial time series that exhibit complex long memory features that can be otherwise very difficult to model. To date, however, there are very few results on the properties of consistent estimators for the parameters of these models. In this paper we investigate the multiple frequency, or k-factor, Gegenbauer autoregressive moving average model (GARMA) that includes ARIMA, fractionally integrated ARMA (ARFIMA), and single frequency GARMA models as special cases and may include features of all of these models simultaneously. The k-factor GARMA models are especially intriguing in that they can capture very convoluted patterns in both the spectral density and autocorrelation functions of a stochastic process using only a few parameters. We describe a conditional sum of squares estimator for the model parameters that is asymptotically equivalent to a maximum likelihood estimator and we derive the asymptotic distributions for this estimator.

ARMA models are generally adequate for capturing the short run dynamics (the first several terms of the autocorrelation function) of most stationary stochastic processes. An ARMA process is characterized by a bounded spectrum at all frequencies and an autocorrelation function that eventually decays at a geometric rate toward zero. ARFIMA models (Granger and Joyeux 1980; Hosking 1981) are able to capture some types of long run dynamics commonly observed in physical data and particularly in economic and financial data and are characterized by a monotonic, hyperbolic rate of decay in autocorrelation functions and a bounded spectrum at all frequencies except, for positive difference parameters, at the origin. GARMA models (Gray, Zhang and Woodward 1989; Chung 1996a,b; Ramachandran and Beaumont 2001) are also long memory models but introduce an additional parameter that allows for a damped sinusoidal pattern in the autocorrelation function that decays toward zero at a hyperbolic rate. The spectrum of a GARMA process is bounded at all frequencies except, if the difference parameter is positive, at the frequency that determines the cycle in the autocorrelation function. Another important feature of a GARMA process is that the sinusoidal pattern in the autocorrelation function is symmetric about zero so there must be negative autocorrelations at some lags.

None of these models is able to reproduce the autocorrelation function displayed in Figure 1, which decays non-monotonically at a hyperbolic rate and is asymmetric about zero. This stochastic process was generated from a model that contains both an ARFIMA component and a GARMA component and is an example of a two frequency GARMA process. Many other patterns that are impossible to capture for the models previously described may be generated by multiple frequency GARMA processes, including one long memory sinusoidal pattern imposed upon another with differing amplitudes. Woodward, Cheng and Gray (1998) demonstrate that atmospheric CO_2 data follow a two frequency GARMA process and we have found that several other physical and financial processes contain multiple long memory frequencies.

The key feature of multiple frequency GARMA models is that they allow for more than one source of long memory and, since they contain ARIMA, GARMA and ARFIMA models as nested special cases, multiple frequency GARMA models provide a convenient framework for tests of the properties of a very wide range of stochastic processes. In the next section we present the details of the multiple frequency GARMA model. These properties were alluded to by Chung (1996b) and more fully developed by Woodward et al. (1998), but they do not derive the asymptotic properties of a consistent estimator that can be used for statistical inference. We derive the properties of a consistent estimator in Section 3, and in Section 4 we present Monte Carlo evidence for the finite sample precision of the iterative CSS estimation method that we propose. An interesting property of k-factor GARMA models that makes them difficult to estimate is that the various parameters may converge at the rates $T^{0.5}$, T, or T^2. In Section 5 we show that the weekly trading volume of IBM stocks is best modelled with a two frequency GARMA process. We summarize and draw conclusions in Section 6.

2 Multiple Frequency Long Memory Processes

The multiple frequency GARMA model was originally discussed by Chung (1996b) and presented in greater detail by Woodward et al. (1998) who refer to the model as a k-factor GARMA model. Although Chung's term is perhaps more literally descriptive, the k-factor usage is probably more commonly employed. Either way, the model is written as

$$\phi(L)\prod_{i=1}^{k}\left(1-2\eta_i L+L^2\right)^{\lambda_i}(x_t-\mu)=\theta(L)\,\varepsilon_t, \tag{1}$$

where $\phi(L)$ and $\theta(L)$ are polynomials in the lag operator L such that $\phi(z)=0$ and $\theta(z)=0$ have roots outside the unit circle, $\{\varepsilon_t\}$ is a white noise disturbance sequence, the λ_i are the difference parameters, and the η_i are the parameters describing the periodic features of the process. The namesake of the GARMA models are the Gegenbauer polynomials $(1-2\eta_i L+L^2)^{\lambda_i}$ which have a pair of complex roots with modulus one and expand to an infinite order polynomial in L. When $k=1$ we get the single frequency GARMA model (Gray et al. 1989), and when, in addition, $\eta=1$ the model further reduces to an ARFIMA(p,d,q) model (Granger and Joyeux 1980; Hosking 1981) where, in this context, $d=2\lambda$. Finally, we get an ARIMA model when $\eta=1$ and $\lambda=0.5$ and an ARMA process when $\lambda=0$.

Assuming that each η_i is distinct, the k-factor GARMA model is stationary if $\lambda_i<0.5$ whenever $|\eta_i|<1$ and when $\lambda_i<0.25$ when $|\eta_i|=1$. The model is invertible if $\lambda_i>-0.5$ whenever $|\eta_i|<1$ and when $\lambda_i>-0.25$ when $|\eta_i|=1$. The proofs of these claims are straightforward (Woodward, Cheng and Gray 1998).

For the stationary cases, the moving average representation of the k-factor GARMA model is

$$(x_t-\mu)=\frac{\theta(L)}{\phi(L)}\prod_{i=1}^{k}\left(1-2\eta_i L+L^2\right)^{-\lambda_i}\varepsilon_t \tag{2}$$

from which the spectral density function is obtained as

$$f(\omega)=\left|\frac{\theta(e^{-i\omega})}{\phi(e^{-i\omega})}\right|^2\prod_{j=1}^{k}\left\{2\,|\cos(\omega)-\cos(\upsilon_j)|\right\}^{-2\lambda_j},\ \omega\in[0,\pi] \tag{3}$$

where $\upsilon_j=\cos^{-1}(\eta_j)$ are the unique Gegenbauer frequencies. The spectral density function is unbounded at all frequencies υ_i where $\lambda_i>0$ and vanishes at all frequencies where $\lambda_i<0$.

Closed form expressions for the autocovariances for a k-factor GARMA process are not available so they must be computed from the sprectrum as

$$\gamma_j = 2\int_0^{\pi} f(\omega)\cos(\omega j)\, d\omega, \tag{4}$$

where care must be taken with the singularities in $f(\omega)$. The partial autocorrelations can then be computed by applying the Durbin-Levinson recursion (Woodward et al., 1998).

Granger and Joyeux (1980) show that the autocorrelations, ρ_j, for a fractional process (when $\eta = 1$) for large j and $\lambda < 0.25$ decay monotonically at a hyperbolic rate and are approximately $\rho_j \approx K^* j^{2\lambda-1}$ where K^* is independent of j. Chung (1996a) shows that for large j the autocorrelation function for a GARMA(0,0) model with $|\eta| < 1$ and $0 < \lambda < 0.5$ can be approximated as $\rho_j \approx J^* \cos(j\, \upsilon)\, j^{2\lambda-1}$ where J^* does not depend upon j. This expression makes clear the hyperbolically damped sinusoidal pattern of the autocorrelation function of a stationary GARMA model with $|\eta| < 1$.

A particularly interesting case for economic and financial applications is when one of the frequencies is zero, corresponding to an $\eta_i = 1$, so that we get a combination of the ARFIMA and GARMA models. This is the case illustrated by the autocorrelation function in Figure 1, which was generated from a two frequency GARMA model with parameters $(\eta_1, \lambda_1) = (1, 0.15)$ and $(\eta_2, \lambda_2) = (0.992, 0.25)$. Note that the first frequency corresponds to an unbounded spike at the origin of the spectrum, and the second frequency corresponds to an unbounded spike at the frequency $\upsilon_2 = \cos^{-1}(0.992) = 0.1266$ radians or 0.0201Hz, which is very close to the origin, so it would be difficult to detect from inspection of the smoothed periodogram. However, the autocorrelation function clearly displays the influence of the second GARMA frequency.

3 Estimation

Our approach is to estimate the k-factor GARMA model using a time domain parametric estimator that is asymptotically equivalent to maximum likelihood estimation. Specifically, we generalize the conditional sum of squares (CSS) estimator described by Chung and Baillie (1993) for fractional models and by Chung (1996a,b) for single frequency GARMA models. Other estimators, both time domain (Woodward et al. 1998), and frequency domain estimators (Giriatis et al., 2001) exist for the k-factor GARMA model. Unfortunately, neither approach has yet to yield an estimator with an established asymptotic theory. Our procedure simultaneously estimates all parameters, including the ARMA components. Furthermore, we provide an exact analytic distribution for all of the parameters of the model. This can be important, for example, because it allows one to test the null hypothesis that the data generating process has an ARFIMA component.

As noted by Chung (1996a,b), if the assumption is made that the initializing disturbances are zero, then maximization of the CSS function is asymptotically equivalent to maximum likelihood estimation. Under the assumption that the disturbances, ε_t, are iid normal with variance σ^2, then the $p+q+2+2k$ parameters $\Psi = \phi_{1,}....., \phi_p,\theta_1, ..., \theta_q, \mu, \sigma^2, \eta_1, \lambda_1, ..., \eta_k, \lambda_k$ can be estimated via maximization of the CSS function

$$\mathcal{L}^*(\Psi) = -\frac{T}{2}\log(2\pi) - \frac{T}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{t=1}^{T}\varepsilon_t^2. \tag{5}$$

Note that the normality assumption is made here only to justify the construction of the CSS function and that the asymptotic theory derived below does not depend upon this assumption. All that is needed for the asymptotic results to hold is that $\{\varepsilon_t\}$ are martingale differences with respect to an increasing sequence of sigma-fields F_t such that, for some $\beta > 0$, $\sup_t E(|\varepsilon_t|^{2+\beta} \| F_{t-1}) < \infty$, almost surely, and $E(\varepsilon_t^2 \| F_{t-1}) = \sigma^2$, almost surely. It is also worth noting that an additional advantage of the CSS estimator is that it is easily generalized to handle GARCH residuals and other types of non-normal distributional assumptions that often arise in financial and economic applications.

To derive the standard errors for the CSS estimator we must consider four cases. The first case is for those models for which $|\eta_i| < 1$ for all $i = 1, \ldots, k$. The second case is for those models for which there exists a single η_i such that $\eta_i = 1$ and $|\eta_j| < 1$ for all other frequencies. The third case is for those models for which there exists a value $\eta_i = -1$ and $|\eta_j| < 1$ otherwise. The fourth case is for those models for which there exists two values η_i and η_j such that $\eta_i = 1$ and $\eta_j = -1$ with $|\eta_m| < 1$ otherwise. Note that if two or more η_i's are equal then the individual λ_i's for those components of the polynomial will not be identified, but we may simply combine those polynomials and estimate $\lambda_1 + \lambda_2$.

The next theorem establishes that the asymptotic information matrix for the k-factor GARMA model is block diagonal.

Theorem 1 *Let $\delta = (\lambda_1, ..., \lambda_k, \phi', \theta')'$ and $\eta^* = (\eta_1, ..., \eta_k)'$ be the parameters associated with the CSS function for the k-factor GARMA model. The asymptotic distribution of δ is independent of η^*.*

The proof of this theorem is given in the Appendix. The essential idea is to establish the different rates of stochastic convergence for the elements of δ and η^*. Note that there is no condition placed on the value of η_i relative to η_j, $i \neq j$, so this theorem holds for all four cases described above. As a result of this theorem, the asymptotic distribution of δ can be considered independently of η^*. The following theorem establishes the asymptotic distribution of the estimators of δ and μ.

Theorem 2 *Let $\hat{\delta}$ be the CSS estimator of δ for the stationary and invertible k-factor GARMA model. (If the mean of the process is unknown we must add the restriction that $\eta_i \neq 1$ for all $i = 1, \ldots, k$.) Then*

$$\sqrt{T}(\hat{\delta} - \delta) \rightsquigarrow N(0, I_{\delta}^{-1}), \tag{6}$$

where $\rightsquigarrow$ denotes the weak convergence of the random vectors $\hat{\delta}$, and where

$$\underset{(k+p+q)\times(k+p+q)}{I_\delta} = \begin{bmatrix} I_{\lambda_1} & \cdots & I_{\lambda_1\lambda_k} & I_{\lambda_1,\phi} & I_{\lambda_1,\theta} \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ I_{\lambda_1\lambda_k} & \cdots & I_{\lambda_k} & I_{\lambda_k,\phi} & I_{\lambda_k,\theta} \\ I_{\lambda_1,\phi} & \cdots & I_{\lambda_k,\phi} & I_{\phi} & I_{\phi,\theta} \\ I_{\lambda_1,\theta} & \cdots & I_{\lambda_k,\theta} & I_{\phi,\theta} & I_{\theta} \end{bmatrix}. \tag{7}$$

The elements of I_δ are defined as follows,

$$I_{\lambda_i} = 2[\frac{\pi^2}{3} - \pi\upsilon_i + \upsilon_i^2],\ i = 1, \ldots, k \tag{8}$$

$$I_{\lambda_i\lambda_j} = 2[\frac{\pi^2}{3} - \pi\upsilon_i + \frac{\upsilon_i^2 + \upsilon_j^2}{2}],\ \upsilon_i > \upsilon_j, \tag{9}$$

$$I_{\lambda_i\phi_j} = 2\sum_{l=0}^{\infty}\phi_l^* \frac{\cos[(l+j)\upsilon_i]}{(l+j)},\ i = 1, \ldots, k,\ j = 1, \ldots, p \tag{10}$$

$$I_{\lambda_i\theta_m} = 2\sum_{l=0}^{\infty}\theta_l^* \frac{\cos[(l+m)\upsilon_i]}{(l+m)}, i = 1, \ldots, k,\ m = 1, \ldots, q \tag{11}$$

where ϕ_l^* and θ_l^* denote the l^{th} coefficients in the infinite order expansions of $\phi^{-1}(L)$ and $\theta^{-1}(L)$, respectively. Finally, the I_ϕ, $I_{\phi,\theta}$ and I_θ submatrices consists of elements that are the same as the corresponding submatrices of the usual information matrix of an ARMA model. Finally, let $\eta_i < 1$, i=1,....,k. For the CSS estimator of the mean, $\hat{\mu}$, we have the following

$$\sqrt{T}(\hat{\mu} - \mu) \rightsquigarrow N(0, 2\pi f(0)), \tag{12}$$

where f(0) denotes the spectral density function evaluated at $\omega = 0$. Further, the distribution of $\hat{\mu}$ is equivalent to the sample mean $\bar{x}$.

The proof of this theorem is given in the Appendix and extends the result of Chung (1996b) to handle multiple frequencies. The following theorem is the central result and provides the asymptotic distribution of the η's for all four cases described above.

Theorem 3 *Let $\hat{\eta}_1, \ldots, \hat{\eta}_k$ be the CSS estimators of $\eta_1, \ldots, \eta_k$, for a stationary and invertible k-factor GARMA model based on a sample $\{X_t\}$, $t = 1, \ldots, T$, with $\eta_i \neq \eta_j, i \neq j$. Without loss of generality, order the elements of η^* from smallest to largest. Then let I_{η_1} denote the indicator function, which takes on the value 1 if $\eta_1 = -1$ and 0 otherwise, and let I_{η_k} denote the indicator function which takes on the value 1 if $\eta_k = 1$ and 0 otherwise. If $\lambda_i \neq 0$, $i = 1, \ldots, k$, then,*

$$T(\hat{\eta}_i - \eta_i) \rightsquigarrow \frac{\sin(\upsilon_i)}{\lambda_i} \frac{\left[\int_0^1 W_{2i-1-I_{\eta_1}}\, dW_{2i-I_{\eta_1}} - \int_0^1 W_{2i-I_{\eta_1}}\, dW_{2i-1-I_{\eta_1}}\right]}{\left[\int_0^1 W^2_{2i-1-I_{\eta_1}}(r)\, dr + \int_0^1 W^2_{2i-I_{\eta_1}}(r)\, dr\right]} \tag{13}$$

with $|\eta_i| < 1$, where $i = 1 + I_{\eta_1}, \ldots, k - I_{\eta_k}$ and,

$$T^2(\widehat{\eta}_1 + 1) \rightsquigarrow -\frac{1}{2\lambda_1} \frac{\int_0^1 \left[\int_0^r W_1(s) ds\right] dW_1(r)}{\int_0^1 \left[\int_0^r W_1(s)\, ds\right]^2 dr}, \ \textit{if } \widehat{\eta}_1 = -1 \tag{14}$$

$$T^2(\widehat{\eta}_k - 1) \rightsquigarrow \frac{1}{2\lambda_k} \frac{\int_0^1 \left[\int_0^r W_{2k-1-I_{\eta_1}}(s) ds\right] dW_{2k-1-I_{\eta_1}}(r)}{\int_0^1 \left[\int_0^r W_{2k-1-I_{\eta_1}}(s) ds\right]^2 dr}, \textit{if } \widehat{\eta}_k = 1 \tag{15}$$

where $W_1, W_2, \ldots., W_{2k-I_{\eta_1}-I_{\eta_k}}$, are $2k - I_{\eta_1} - I_{\eta_k}$ independent Brownian motions.

The proof of this theorem is given in the Appendix. An important result of this theorem is the asymptotic independence of the values in the vector η^*. In addition, for each $\hat{\eta}_i$, the values of λ_i and υ_i enter the equation for the asymptotic distribution proportionally so one only needs the values of the stochastic integrals depicted in theorem 3 to calculate asymptotic confidence intervals. The values for these integrals are reported in Chung (1996a).

These theorems provide important practical information for designing an efficient estimator for the parameters of the k-factor GARMA model. We know that the asymptotic distributions of the $\hat{\lambda}_i$'s are not independent of the $\hat{\phi}$'s and the $\hat{\theta}$'s. Also, the asymptotic distributions of $\hat{\delta}$ and $\hat{\eta}^*$ are independent, but the elements of $\hat{\delta}$ are $O_p(T^{-1/2})$, whereas $\hat{\eta}_i$ is $O_p(T^{-1})$ if $|\eta_i| < 1$ and $O_p(T^{-2})$ if $|\eta_i| = 1$. This suggests that the procedure advocated by Woodward et al. (1998) of estimating the ARMA parameters conditional upon the (η_i, λ_i) pairs in a grid will not produce consistent and efficient estimators. It would be more appropriate to use an extension of Chung's (1996a,b) method by conducting a grid search over η^* combined with a gradient method over δ. However, Monte Carlo simulations indicate that the grid over the η's must be very fine since the likelihood function has many local minima. This makes such an approach computationally burdensome when there are more than two frequencies in the model.

The computational burden of the grid search could be reduced if one had good starting values for the long memory parameters (η_i, λ_i). One possibility is the exploitation of the information contained in the autocorrelations of the data. Tieslau, Schmidt and Baillie (1996) suggest fitting sample autocorrelations to theoretical counterparts to estimate the parameters of an ARFIMA model. But as Chong (2000) notes for ARFIMA models, such a procedure must be used with caution since the parameterization is not necessarily unique, and despite the long memory properties of the data, most of the information in the sample autocorrelations is still found in the first few terms. Chong (2000) reports that relying on the long lags in the autocorrelation function produces a downward bias in the differencing parameter d for the ARFIMA model.

In addition, one can exploit the asymptotic independence of the values of the vector η^* to obtain starting values. The computational burden for the k-factor GARMA model is derived from the fact that a grid from -1 to 1 must be set up for each of the k frequencies. Thus, for a step size of .1, there are a total of 21^k grid points. Since η_i is stochastically independent of η_j for $i \neq j$, one should obtain reasonable estimates for these values by estimating a single frequency GARMA model. The residuals can be refit with another single frequency GARMA model. This procedure continues until the residuals in the last step resemble white noise. Simulation results reported below indicate that this procedure results in a remarkably small bias in the estimates of the values of the vector η^*. It should be pointed out that this algorithm does not produce a consistent estimator but should only be used to limit the number of grid points for η^*. Clearly, the elements of δ are not stochastically independent, and the use of this procedure will result in biased estimates for this vector. The simulation results below confirm this.

The computational complexity of this approach can be better appreciated when we consider the step of recursively computing the residuals for the CSS estimator. The i^{th} Gegenbauer polynomial in the k-factor GARMA model can be expanded as

$$(1 - 2\eta_i z + z^2)^{-\lambda_i} = \sum_{j=0}^{\infty} C_j^{(\lambda_i)}(\eta_i)\, z^j \tag{16}$$

where (Gray et. al. 1989),

$$C_j^{(\lambda_i)}\left(\eta_i\right) = \sum_{l=0}^{[j/2]} \frac{(-1)^l \left(2\eta_i\right)^{j-2l}\,\Gamma\left(\lambda_i - l + j\right)}{l!\,(j-2l)!\,\Gamma\left(\lambda_i\right)} \tag{17}$$

and $[j/2]$ is the integer part of $j/2$. As Chung (1996a) notes, the easiest way to calculate the coefficients $C_j^{(\lambda_i)}$ is via the following recursion,

$$C_j^{(\lambda_i)}(\eta_i) = 2\eta_i\left(\frac{\lambda_i - 1}{j} + 1\right) C_{j-1}^{(\lambda_i)}(\eta_i) - \left(2\frac{\lambda_i - 1}{j} + 1\right) C_{j-2}^{(\lambda_i)}(\eta_i) \tag{18}$$

where $C_0^{(\lambda_i)}(\eta_i) = 1$ and $C_1^{(\lambda_i)}(\eta_i) = 2\,\lambda_i\,\eta_i$. Under the assumption that $\varepsilon_0 = \varepsilon_{-1} = \ldots = 0$, one can calculate the residuals of the k-factor GARMA model recursively from the expression

$$\phi(L)(x_t - \mu) = \prod_{i=1}^{k}\left[\sum_{j=0}^{\infty} C_j^{(\lambda_i)}(\eta_i)\,L^j\right]\theta(L)\,\varepsilon_t. \tag{19}$$

It is the combination of the k-dimensional product over the infinite dimensional sums that create the computational burden, because the sums cannot be truncated at too low a value without introducing significant errors.

Our Monte Carlo simulations, which we report in detail in the following section, indicate that a combination of the methods described above work quite well. We begin by fitting a polynomial in η^* and δ to the sample autocorrelation function. Like Chong (2000) and Tieslau et al. (1996), we find that 20 lags are generally sufficient to capture the movement in the multiple frequency GARMA process. It is possible that weighting the first several lags of the autocorrelations more heavily and making use of the partial autocorrelation function may improve the fit, but we leave these issues for future research. We combine the procedure of autocorrelation fitting with the sequence of single GARMA estimations as described above. Next a grid is created for each element of η^* about the initial estimates. Fixing the values $\eta_2, \ldots, \eta_k$, a gradient method is used to update the value of η_1. Using this updated η^* vector, a gradient method is then used to update the δ vector. Then we fix $\eta_1, \eta_3, \ldots, \eta_k$ and use a gradient method to update η_2, etc. The final model results from the set of parameters that produce the smallest sum of squared errors. The procedure is computationally burdensome but necessary given the properties of the estimators and the behavior of the CSS function.

4 Finite Sample Performance

In the previous section we showed that the asymptotic information matrix is block diagonal in the parameter groups η^* and δ, and the convergence rates of the elements of η^* are faster than those of δ. Our two-stage, iterative CSS estimator of the k-factor GARMA model parameters is designed to acknowledge these two facts. In this section we report simulation results that examine the finite sample properties of the CSS estimation. We are interested in examining the finite sample bias in the parameter estimates and in comparing the finite sample standard errors of the parameter estimates with the asymptotic standard errors.

Since these are *long memory* models, we are not particularly interested in examining the estimator's properties for very small sample sizes. When $\eta = 0.93$, for instance, the cycle in the data is $0.06\,Hz$ or 16.7 periods per cycle, so it would be quite possible to detect this cycle with about 100 observations. However, when $\eta = .998$ the cycle becomes 100 periods long, so clearly estimation with 100 observations is difficult. Chung (1996a,b) and Ramachandran and Beaumont (2001) have done extensive simulations for the single frequency GARMA model with the latter paying particular attention to the parametric region where η is close to one and λ is close to one-half. Based upon those results, we will use a sample size of 300 observations, and we will concentrate on two frequency models with parameter ranges that we believe are most relevant for economic and financial applications. We pay particular attention to the mixed ARFIMA/GARMA case.

The simulation results are presented in Tables 1–4. The columns of each table list the parameters of the simulated two frequency GARMA model and each block in the tables gives the results for a specific parameterization. The first row in each block gives the true parameter values (TRUE); the second row presents the mean bias (MEAN BIAS); the third row gives the median bias (MED. BIAS); the fourth row reports the root mean square error (RMSE); the fifth row reports the mean of the numerical standard errors (MNSE) that are computed from the gradient vector in the last iteration of the estimation; and the sixth row is the mean absolute deviation (MAD). In each case, we carry out 1000 replications with 300 observations per replication. We report both the mean bias and median bias as well as the RMSE and the MAD in the event that outliers, that may occur in any simulation exercise, distort the mean calculations. For computational purposes, we use an iterative procedure to generate data. We generate a large amount of observations, and then discard all but the last 300 observations. On extremely rare occasions, the generated data do not take on the properties of a multiple frequency GARMA model. For the 17000 generated series below, it was necessary to discard three.

Table 1 presents the results for six different two frequency GARMA(0,0) models with η values of $-\frac{1}{2}, 0, \frac{1}{2}$ and λ values of 0.2 and 0.4. In a sense, these are the *easy* cases, because the η's are well away from unity and there are no nuisance ARMA parameters to deal with. The estimation biases are all quite small especially for the η's that converge at a faster rate than the λ's. For the λ's, there is not much difference between the mean bias and the median bias or the RMSE and the MAD indicating that the distribution of these parameters is quite robust. Generally speaking, a larger value of λ mitigates the already small bias in the η's. The η's appear to be marginally more sensitive to estimation outliers. This is probably due to the fact that an estimate of λ_i near zero can lead to wildly wrong estimates of the corresponding η_i since that Gegenbauer polynomial will have very little impact on the likelihood function no matter what the value of η_i is. In these cases, the mean, μ, is estimated with the sample mean, which is again asymptotically equivalent to the CSS estimator of μ provided $\eta_i < 1,\ i = 1, ..., k$. As noted above, the estimator for the mean is $O_p(T^{-1/2})$ just as the other parameters in δ, so its bias is also quite small.

To further validate our estimator, we compare the mean numerical standard errors calculated from the estimated Hessian matrix in the last iteration with the true asymptotic standard errors calculated with the aid of Theorem 2. In particular, for cases 1-6 of Table 1 the true asymptotic standard errors of the corresponding values of λ_1 are 0.0394, 0.0450, 0.0394, 0.0455, 0.0450, and 0.0455, respectively. These values are quite comparable to the MNSE and RMSE of the corresponding numbers in Table 1. The true asymptotic standard errors for λ_2 are 0.0455, 0.0450, 0.0455, 0.0394, 0.0450, and 0.0394, which again are very close to their numerical counterparts. The true asymptotic standard errors for the values of μ are 0.0438, 0.0372, 0.0503, 0.0324, 0.0463, and 0.0351. Recall that the mean is estimated with the sample mean, and thus the mean numerical standard errors are not reported. However, the RMSE is comparable to the true asymptotic standard error, although it is interesting to note that the RMSE slightly underestimates the standard deviation of the mean in small samples. Finally, in light of the results of Theorem 3, it is not surprising to see that the MNSE and RMSE for the η's are quite different, since the RMSE assumes convergence at the rate $\sqrt{T}$.

To examine the influence of the nuisance parameters ϕ and θ we choose a particular parameterization (case 2 from Table 1) and estimate various two frequency GARMA(p, q) models with p and q being either zero or one. The results are reported in Table 2 and are similar to those in Table 1. Interestingly, the main consequence of the inclusion of ARMA parameters is a relatively wide distribution for the sample mean when a positive autoregressive parameter exists. Again, for all of the cases considered in Table 2, the median and mean biases are quite small, and the RMSE compares favorably with the MNSE.

Table 3 examines the particularly interesting case where one of the η's equals one so that we get a combination ARFIMA and GARMA model. Compared to η_2, the estimator for $\eta_1 = 1$ has very little bias and extremely small RMSE and MNSE reflecting the fact that this parameter is $O_p(T^{-2})$, an extremely fast rate of convergence. The results for $|\eta_2| < 1$ are similar to those in Tables 1 and 2, as are the results for the λ's. Note, however, that the RMSE's for the η's in Table 2 are much greater, as expected, than the RMSE's for the η_1's in Table 3. In addition, the corresponding true asymptotic standard errors are of the same magnitude as the corresponding MNSE and RMSE. For example, for case 1 of Table 3, the true asymptotic standard errors for λ_1 and λ_2 are 0.0274 and 0.0400, which compare favorably to the estimated standard errors (0.0288 and 0.0410 respectively). When $\eta_i = 1$, however, the sample mean and CSS estimate of μ are no longer asymptotically equivalent. Thus we use the CSS estimator for the mean in these cases. The computational difficulties of time domain estimators for ARFIMA models when the mean is unknown has been well documented (Adenstedt 1974; Yajima 1991; Chung and Baillie 1993; Cheung and Diebold 1994; Hauser, Potscher and Reshenhofer 1999; and Ooms and Doornik 1999). For example, when λ is positive the estimator for the mean of ARFIMA models is $O_p(T^{-1/2+\lambda})$. From inspection of Theorem 2, it is obvious that this result carries over to the mixed ARFIMA/GARMA case. In spite of these difficulties, the mean is fairly unbiased albeit with a wide distribution. Again, the remaining parameters suffer from very little distortion.

As noted above, the computational burden of the CSS estimator grows rapidly with the number of frequencies due to the grid search over the η's. Thus, if we could narrow the range of the grid search over the η's we could greatly improve the efficiency of the algorithm. Since the η's are asymptotically independent of one another and of the δ parameters, it may be possible to first estimate the η's sequentially to get good starting values and then re-estimate the process using fairly tight grids over the η's. In Table 4, we investigate this possibility. First we estimate a single frequency GARMA model and then filter the data with the resulting Gegenbauer polynomial before estimating the second frequency using a single freqency model on this filtered data. This process should produce good starting values for the η's as long as the biases in the η's are not too large.

The first two cases in Table 4 are cases from the previous simulations and the third case represents a mixed ARFIMA/GARMA model in which the ARFIMA component is short memory ($\lambda < 0$). The latter process may result from differencing a non-stationary ARFIMA process (Smith Jr., Sowell and Zin 1997). For each of the cases considered in Table 4, the sample mean is used to estimate μ. We find that the single frequency estimator generally first selects the frequency with the largest corresponding value of λ, thus capturing the most dominate feature of the autocorrelation function. The results in Table 4 indicate that the small sample biases in η_1 and η_2 are quite reasonable, suggesting that the possibility of choosing a tight grid around these point estimates may work quite well. The relatively large biases in the values of the vector δ, however, confirm the results of Theorem 2 that a consistent estimator is obtained only through joint estimation of all parameters.

5 Application

Many financial time series are good candidates for k-factor GARMA models. For example, the trading volume, v, of IBM equities from June 3, 1977 through April 12, 2002 is well modeled as the 2-factor GARMA(0,1) process

$$(1 - 2L + L^2)^{0.269}(1 - 1.152L + L^2)^{0.152}(v_t - \hat{\mu}) = (1 - 0.324L)\widehat{\varepsilon}_t \tag{20}$$

which is a mixed ARFIMA/GARMA process with $(\eta_1, \lambda_1) = (1.00, 0.269)$ and $(\eta_2, \lambda_2) = (0.576, 0.152)$ and mean $\hat{\mu} = 3.175\,10^7$. The numerical standard errors for λ_1 (associated with $\eta_1 = 1$), λ_2 (associated with $\eta_2 = 0.576$) and θ are: 0.0255, 0.0233 and 0.0706, respectively, indicating high levels of significance. The asymptotic 95% confidence intervals for the η's are: [0.9999, 1.0000] and [0.5646, 0.5876]. Note that, since $\hat{\lambda}_1 > 0.25$, the estimated process is nonstationary. The second frequency associated with $\hat{\eta}_2 = 0.576$ suggests a cycle of 6.6 weeks in the data. If the volume data are estimated with an ARFIMA model, the residuals still exhibit long memory features of the GARMA type and the autocorrelations of the residuals show strong serial correlations exist even as far out as 100 lags. However it is not valid to estimate a single frequency GARMA model to these residuals since the estimators for λ_1 from the ARFIMA process and the λ_2 from the GARMA process are not independent. Indeed, λ_1 would tend to be overestimated and λ_2 would tend to be underestimated in this case.

6 Conclusions

In this paper we investigate the properties of a model that captures very diverse patterns in the autocorrelation functions of data. The k-factor GARMA model generalizes existing long memory models and has the particular advantage that the autocorrelations can decay at a non-monotonic rate that is not necessarily symmetric about zero. In addition, the k-factor GARMA model can accommodate multiple singularities in the spectral density function.

The multiple frequency GARMA model has not been studied in much detail. Woodward et al. (1998) do investigate some preliminary aspects of the model, but they do not present an estimator that extends readily to statistical inference. We propose a conditional sum of squares estimator and present the asymptotic properties in Theorems 1–3 whose proofs are presented in the Appendix. The key feature of these theorems is that the asymptotic distribution of η_i is independent of η_j whenever $i \neq j$. Furthermore, η_i is independent of the remaining parameters of the model denoted by the vector δ. These results hold for all potential values of η_i and η_j. It is very important to note, however, that the values in the vector δ are not asymptotically independent of each other and therefore estimators that sequentially estimate these values will likely suffer from severe bias. Finally, the model parameters are shown to converge at differing rates. Our proposed CSS estimator of the model parameters exploits the particular convergence rates of the parameters found in Theorems 1–3.

The simulation results in Section 4 demonstrate that the estimator performs quite well and that the finite sample standard errors are very close to the asymptotic calculations. In addition, the simulation results suggest that the computational complexity associated with a k-dimensional grid search can be greatly reduced via repeated estimation of a single frequency GARMA model to obtain starting values. The application in Section 5 demonstrates the practical value of the k-factor GARMA model. The trading volume of IBM is shown to be a two frequency ARFIMA/GARMA mixture.

Appendix

Proof of Theorem 1:

Consider the first order Taylor series expansion of the CSS estimators of the invertible and stationary k-factor GARMA model of the process $\{x_t\}_{t=1}^{T}$ about the true parameter values $\delta = (\lambda_1, \ldots, \lambda_k, \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q)$ and $\eta = (\eta_1, \ldots, \eta_k):$

$$\begin{bmatrix} \frac{1}{\sqrt{T}} \frac{\partial \mathcal{L}^*}{\partial \delta} \\ \frac{1}{f_T} \odot \frac{\partial \mathcal{L}^*}{\partial \eta} \end{bmatrix} + \begin{bmatrix} \frac{1}{T} \frac{\partial^2 \mathcal{L}^*}{\partial \delta\, \partial \delta'} & \frac{1}{\sqrt{T}} \frac{1}{F_T} \odot \frac{\partial^2 \mathcal{L}^*}{\partial \delta\, \partial \eta'} \\ \frac{1}{\sqrt{T}} \frac{1}{F_T} \odot \frac{\partial^2 \mathcal{L}^*}{\partial \delta\, \partial \eta'} & \left(\frac{1}{f_T}\right) \left(\frac{1}{f_T}\right)' \odot \frac{\partial^2 \mathcal{L}^*}{\partial \eta\, \partial \eta'} \end{bmatrix}$$

$$\times \begin{bmatrix} \sqrt{T} \left(\hat{\delta} - \delta\right) \\ f_T \odot (\hat{\eta} - \eta) \end{bmatrix} = o_p\left(1\right) \quad \text{(A.1)}$$

where $\odot$ denotes element by element multiplication, f_T and $\frac{1}{f_T}$ denotes k×1 vectors whose j^{th} elements are T and $\frac{1}{T}$ when $|\eta_j| < 1$ and T^2 and $\frac{1}{T^2}$ when $|\eta_j| = 1$,. and $\frac{1}{F_T}$ denotes the matrix formed by stacking the row vector, $\frac{1}{f_T}'$, on top of itself k times.

We will show that $\frac{1}{T} \frac{\partial^2 \mathcal{L}^*}{\partial \delta\, \partial \delta'}$ and $\left(\frac{1}{f_T}\right) \left(\frac{1}{f_T}\right)' \odot \frac{\partial^2 \mathcal{L}^*}{\partial \eta\, \partial \eta'}$, are O$_p(1)$, while $\frac{1}{\sqrt{T}} \frac{1}{F_T} \odot \frac{\partial^2 \mathcal{L}^*}{\partial \delta\, \partial \eta'}$ possesses elements that are all $o_p(1)$. Below, we show that the remaining elements are bounded and find the exact distribution of these elements. Initially we assume that the remaining elements are $O_p(1)$. For large T, we get

$$\begin{bmatrix} \frac{1}{\sqrt{T}\, \sigma^2} \sum_{t=1}^{T} \varepsilon_t \frac{\partial \varepsilon_t}{\partial \delta} \\ \frac{1}{\sigma^2} \frac{1}{f_T} \odot \sum_{t=1}^{T} \varepsilon_t \frac{\partial \varepsilon_t}{\partial \eta} \end{bmatrix} + \begin{bmatrix} I_\delta & \sqrt{T} \frac{1}{F_T} \odot I_{\delta\eta} \\ \sqrt{T} \frac{1}{F_T} \odot I_{\delta\eta} & \frac{1}{\sigma^2} \left(\frac{1}{f_T}\right) \left(\frac{1}{f_T}\right)' \odot \sum_{t=1}^{T} \frac{\partial^2 \varepsilon_t}{\partial \eta\, \partial \eta'} \end{bmatrix}$$

$$\times \begin{bmatrix} \sqrt{T} \left(\hat{\delta} - \delta\right) \\ f_T \odot (\hat{\eta} - \eta) \end{bmatrix} = o_p\left(1\right). \quad \text{(A.2)}$$

The cases for $I_{\phi_i,\eta_j}, I_{\theta_i,\eta_j}$, and I_{λ_i,η_j} when $|\eta_j| = 1$ follow from Chung (1996a,b). We thus concentrate on the I_{λ_i,η_j} expressions when $|\eta_j| < 1$. From the equation for the innovations, and using Gradshteyn and Ryzhik (1980) equations 1.514 and 8.937.1 we find the I_{λ_i,η_j} information matrix elements to be

$$-E\left[\frac{1}{T} \frac{\partial^2 \mathcal{L}^*}{\partial \lambda_i\, \partial \eta_j}\right] = E\left[\frac{4\lambda_j}{\sigma^2 T} \sum_{t=1}^{T} \left(\sum_{l=1}^{\infty} \frac{\cos(l\upsilon_i)}{l} \varepsilon_{t-l}\right) \left(\sum_{l=1}^{\infty} \frac{\sin(l\upsilon_j)}{\sin(\upsilon_j)} \varepsilon_{t-l}\right)\right]. \quad \text{(A.3)}$$

Under the assumptions governing ε_t, if $\upsilon_j > \upsilon_i$, and $\upsilon_i \neq \upsilon_j$, Gradshteyn and Ryzhik (1980) equation 1.441.1 yields

$$I_{\lambda_i \eta_j} = \frac{2\lambda_j}{\sin(\upsilon_j)}(\sum_{l=1}^{\infty}\frac{\sin[l(\upsilon_i+\upsilon_j)]}{l} + \sum_{l=1}^{\infty}\frac{\sin[l(\upsilon_j-\upsilon_i)]}{l}) = \frac{2\lambda_j(\pi-\upsilon_j)}{\sin(\upsilon_j)}. \tag{A.4}$$

Thus, $I_{\lambda_i \eta_j} < \infty$. If $\upsilon_j < \upsilon_i$, then the infinite sums in (A.4) are equal to $\sum_{l=1}^{\infty}\frac{\sin[l(\upsilon_j-\upsilon_i)+2\pi l]}{l}$. From Gradshteyn and Ryzhik (1980) equation 1.444.1, we see that the infinite sum converges. The same is true if $\upsilon_i = \upsilon_j$.

If the remaining terms of all of the elements in (A.2) are $O_p(1)$ as shown below, then the matrix in (A.2) is asymptotically block diagonal, and the distribution of $\sqrt{T}(\hat{\delta}-\delta)$ can be considered independently of $f_T\ (\hat{\eta}-\eta)$ as claimed.

Proof of Theorem 2:.

From (A.2), the assumption that the remaining elements involving η^* in (A.2) are bounded, and Chan and Wei's (1988) central limit theorem,

$$\sqrt{T}(\hat{\delta}-\delta) = -I_{\delta}^{-1}\left[\frac{1}{\sqrt{T}\sigma^2}\sum_{t=1}^{T}\varepsilon_t\frac{\partial\varepsilon_t}{\partial\delta}\right] + o_p(1) \rightsquigarrow N(0, I_{\delta}^{-1}). \tag{A.5}$$

Information numbers for the diagonal terms of I_δ are given in Chung (1996a, page 251). The off diagonal terms $I_{\lambda_i \lambda_j}$, which for large T and $i \neq j$ are,

$$-E\left(T^{-1}\frac{\partial^2\mathcal{L}^*}{\partial\lambda_i\partial\lambda_j}\right) = E\frac{1}{T\sigma^2}\sum_{t=1}^{T}\left[\log(1-2\eta_i L+L^2)\varepsilon_t \, \log(1-2\eta_j L+L^2)\varepsilon_t\right]. \tag{A.6}$$

Using Gradshteyn and Ryzhik (1980) equations 1.514 and 1.443.3 yields,

$$I_{\lambda_i\lambda_j} = 2\sum_{l=1}^{\infty}\frac{\cos(l(\upsilon_i+\upsilon_j))+\cos(l(\upsilon_i-\upsilon_j))}{l^2} = 2(\frac{\pi^2}{3}-\pi\upsilon_i+\frac{\upsilon_i^2+\upsilon_j^2}{2}), \tag{A.7}$$

For the CSS estimator for $\hat{\mu}$, the information number is:

$$I_\mu = -E\left(T^{-1}\frac{\partial^2\mathcal{L}^*}{\partial d_\mu^2}\right) = \frac{1}{\sigma^2}\left|\frac{\phi(1)}{\theta(1)}\right|^2\prod_{i=1}^{k}(2-2\eta_i)^{2\lambda_i} = \frac{1}{2\pi}f(0)^{-1}, \tag{A.8}$$

where $f(0)$ denotes the spectral density function evaluated at $\omega = 0$. Now consider the variance for $\bar{x}$. We have:

$$var\left[\sqrt{T}(\bar{x}-\mu)\right] = \frac{\sigma^2}{T}\left[T\left\{\left|\frac{\theta(1)}{\phi(1)}\right|^2\prod_{i=1}^{k}(2-2\eta_i)^{-\lambda_i}\right\}\right] = 2\pi f(0). \tag{A.9}$$

It is clear that the variance of $\bar{x}$ is equivalent to that of the CSS estimator of μ. Thus, by the central limit theorem of Chan and Wei, we also have $\sqrt{T}(\bar{x}-\mu) \rightsquigarrow N(0, 2\pi f(0))$. The proof of the results for $I_{\lambda_1\phi_j}, I_{\lambda_1\theta_m}, \ldots, I_{\lambda_k\phi_j}$, and $I_{\lambda_k\theta_m}$ follows directly from Chung (1996b) in the single frequency case.

Proof of theorem 3:

Before proving Theorem 3 we state and prove the following useful lemma.

Lemma 4 *Let $\hat{\eta}_1, \ldots, \hat{\eta}_k$ be the CSS estimators for $\eta^* = (\eta_1, \ldots, \eta_k)$ in a stationary and invertible k-factor GARMA model. Then, with $i \neq j$,*

$$\frac{1}{T^\alpha}\frac{\partial^2 \mathcal{L}^*}{\partial \eta_i \partial \eta_j} = o_p(1),$$

where $\alpha = 2$ if $|\eta_i|, |\eta_j| < 1, \{i, j \in [1, k] : i \neq j\}$ (case 1), $\alpha = 3$ if $\eta_i = \pm 1$ and $|\eta_j| < 1$ (cases 2 and 3), and $\alpha = 4$ if $\eta_i = -1$ and $\eta_j = 1$ (case 4).

Proof of the Lemma: Consider case 1 where $|\eta_i|, |\eta_j| < 1, \{i, j \in [1, k] : i \neq j\}$. Without loss of generality, and for ease of notation, rearrange the polynomials in η such that $\eta_i = \eta_1, \eta_j = \eta_2$. Let,

$$Z_{at} = -\frac{1}{2\lambda_a}\frac{\partial \varepsilon_{t+1}}{\partial \eta_a} = \frac{\varepsilon_t}{(1 - 2\eta_a L + L^2)}, \ a = 1, 2. \tag{A.10}$$

Applying Gradshteyn and Ryzhik (1980) equation 8.937.1,

$$Z_{at} = \frac{1}{\sin(\upsilon_a)}\sum_{j=1}^{t} \sin[(t+1)\upsilon_a - j\upsilon_a]\varepsilon_j, \ a = 1, 2. \tag{A.11}$$

which follows if $\varepsilon_0 = \varepsilon_{-1} = ... = 0$. Now, define the random elements

$$S_T(\upsilon_a, r) = \frac{\sqrt{2}}{\sqrt{T\sigma^2}}\sum_{j=1}^{[Tr]} \cos(j\upsilon_a)\varepsilon_j, \ a = 1, 2 \tag{A.12a}$$

$$T_T(\upsilon_a, r) = \frac{\sqrt{2}}{\sqrt{T\sigma^2}}\sum_{j=1}^{[Tr]} \sin(j\upsilon_a)\varepsilon_j, \ a = 1, 2 \tag{A.12b}$$

where $r \in [0, 1]$ and $[Tr]$ is the integer part. Finally, from the expressions in (A.12) and using $\omega_1 = \upsilon_1 + \upsilon_2$,

$\omega_2 = \upsilon_1 - \upsilon_2$, $\omega_3 = \upsilon_2 - \upsilon_1$ along with a few rules of trigonometry, we get the following expression,

$$\begin{aligned}
&\frac{4\sin(\upsilon_1)\sin(\upsilon_2)}{\sigma^2}\frac{1}{T^2}\sum_{t=1}^{T-1} Z_{1t}\, Z_{2t} \\
&= \frac{1}{T}\sum_{t=1}^{T-1}(\cos[(t+1)\,\omega_2] - \cos[(t+1)\omega_1])S_T(\upsilon_1, t/T)S_T(\upsilon_2, t/T) \\
&- \frac{1}{T}\sum_{t=1}^{T-1}(\sin[(t+1)\omega_1] + \sin[(t+1)\omega_3])S_T(\upsilon_2, t/T)\, T_T(\upsilon_1, t/T) \\
&- \frac{1}{T}\sum_{t=1}^{T-1}(\sin[(t+1)\omega_1] + \sin[(t+1)\omega_2])S_T(\upsilon_1, t/T)T_T(\upsilon_2, t/T) \\
&+ \frac{1}{T}\sum_{t=1}^{T-1}(\cos[(t+1)\omega_1] + \cos[(t+1)\omega_2])T_T(\upsilon_1, t/T)T_T(\upsilon_2, t/T). \qquad \text{(A.13)}
\end{aligned}$$

Consider the random elements

$$S_n^*(\upsilon_1) = \sum_{j=1}^{n}\cos(j\upsilon_1)\,\varepsilon_j \quad \text{and} \quad T_n^*(\upsilon_1) = \sum_{j=1}^{n}\sin(j\upsilon_1)\,\varepsilon_j, \qquad \text{(A.14)}$$

and similarly for $S_n^*(\upsilon_2)$ and $T_n^*(\upsilon_2)$. Let $\{X_n\} = \{S_n^*(\upsilon_1)\, S_n^*(\upsilon_2)\}$, and consider the first term in the expansion of (A.13). It is clear from the definition of $S_T(\upsilon_1, \frac{t}{T})$ and $S_T(\upsilon_2, \frac{t}{T})$ that $\frac{1}{T}\sum_{t=1}^{T-1}\cos[(t+1)\omega_2]S_T(\upsilon_1, t/T)S_T(\upsilon_2, t/T) = o_p(1)$ if

$$\underset{1\le j\le T}{sup}\left|\sum_{n=1}^{j} e^{in\theta}X_n\right| = o_p(T^2). \qquad \text{(A.15)}$$

First, observe that

$$E|S_n^*(\upsilon_1)S_n^*(\upsilon_2)| \le \{ES_n^{*2}(\upsilon_1)\}^{1/2}\{ES_n^{*2}(\upsilon_2)\}^{1/2} \le \sigma^2 n \qquad \text{(A.16)}$$

so that $E|S_n^*(\upsilon_1)S_n^*(\upsilon_2)| = O(n)$. Now let $n \ge m$ and consider

$$|X_n - X_m| \le |S_n^*(\upsilon_1)||S_n^*(\upsilon_2) - S_m^*(\upsilon_2)| + |S_m^*(\upsilon_2)||S_n^*(\upsilon_1) - S_m^*(\upsilon_1)|. \qquad \text{(A.17)}$$

Noting that

$$E|S_n^*(\upsilon_1)|^2 = E(S_n^*(\upsilon_1))^2 = \sigma^2\left(\sum_{j=1}^{n}\cos(j\upsilon_1)\right)^2 \le \sigma^2 n \qquad \text{(A.18)}$$

yields $E|S_n^*(\upsilon_1)|^2 = O(n)$. Given $m \le n$, this also implies $E|S_m^*(\upsilon_2)|^2 \le \sigma^2 n$.

Next consider the expression

$$E|S_n^*(\upsilon_2) - S_m^*(\upsilon_2)|^2 = \sigma^2\left(\sum_{j=m+1}^{n}\cos^2(j\upsilon_2)\right) \le \sigma^2(n-m). \qquad \text{(A.19)}$$

Thus, $E|S_n^*(\upsilon_2) - S_m^*(\upsilon_2)|^2 = O(n-m)$. Similar reasoning implies that $E|S_n^*(\upsilon_1) - S_m^*(\upsilon_1)|^2 = O(n-m)$. If $\upsilon_1 \neq \upsilon_2$, by Chan and Wei's (1988) Theorem 2.1, we see that the first term in (A.13) is $o_p(1)$. By similar reasoning, the remaining terms in (A.13) are also seen to be $o_p(1)$. Thus, we have established that

$$\frac{4\sin(\upsilon_1)\sin(\upsilon_2)}{\sigma^2}\frac{1}{T^2}\sum_{t=1}^{T-1} Z_{1t}Z_{2t} = o_p(1). \tag{A.20}$$

This expression is asymptotically equivalent to

$$-\frac{4\lambda_1\lambda_2}{4\sin(\upsilon_1)\sin(\upsilon_2)}\frac{4\sin(\upsilon_1)\sin(\upsilon_2)}{\sigma^2}\frac{1}{T^2}\sum_{t=1}^{T-1} Z_{1t}Z_{2t}, \tag{A.21}$$

which is $o_p(1)$. So this completes the proof of Case 1 in the Lemma.

Case 2: Without loss of generality let $\eta_k = 1,\ |\eta_j| < 1,$ and $j \neq k$. Rearrange the polynomials in η such that $\eta_j = \eta_1$, and define the following elements,

$$Z_{1t} = -\frac{1}{2\lambda_1}\frac{\partial\varepsilon_{t+1}}{\partial\eta_1} = \frac{\varepsilon_t}{(1-2\eta_1 L + L^2)}\ ,\ Z_{2t} = -\frac{1}{2\lambda_k}\frac{\partial\varepsilon_{t+1}}{\partial\eta_k} = \frac{\varepsilon_t}{(1-L)^2}. \tag{A.22}$$

Define the auxiliary process and its associated truncation.

$$Y_t = (1-L)Z_{2t} = \sum_{j=1}^{t}\varepsilon_j. \tag{A.23}$$

This gives the following truncated series forZ_{2t},

$$Z_{2t} = \sum_{j=1}^{t} Y_t = \sum_{j=1}^{t} j\,\varepsilon_{t-j+1}. \tag{A.24}$$

For ease of exposition, define the random process

$$X_T(r) = \frac{1}{T}\frac{1}{\sqrt{T}\sigma}\sum_{j=1}^{[Tr]} Y_j, \tag{A.25}$$

and define $S_T(\upsilon_1, t/T)$ and $T_T(\upsilon_1, t/T)$ precisely as in (A.12). We then get,

$$\begin{aligned}\frac{\sqrt{2}\sin(\upsilon_1)}{\sigma^2}\frac{1}{T^3}\sum_{t=1}^{T-1} Z_{1t}Z_{2t} = \frac{1}{T}\sum_{t=1}^{T-1}\sin[(t+1)\upsilon_1]\,S_T(\upsilon_1, t/T)\,X_T(t/T)\\ -\frac{1}{T}\sum_{t=1}^{T-1}\cos[(t+1)\upsilon_1]\,T_T(\upsilon_1, t/T)\,X_T(t/T).\end{aligned} \tag{A.26}$$

Note that the expression

$$-\frac{4\lambda_1\lambda_k}{\sqrt{2}\sin(\upsilon_1)}\frac{\sqrt{2}\sin(\upsilon_1)}{\sigma^2}\frac{1}{T^3}\sum_{t=1}^{T-1} Z_{1t}Z_{2t}, \tag{A.27}$$

is asymptotically equivalent to $\frac{1}{T^3}\frac{\partial^2 \mathcal{L}^*}{\partial \eta_1 \partial \eta_k}$. Define the processes

$$S_n^*(\upsilon_1) = \sum_{j=1}^{n} \cos(j\upsilon_1)\varepsilon_j,\ T_n^*(\upsilon_1) = \sum_{j=1}^{n} \sin(j\upsilon_1)\varepsilon_j, \text{ and } \ X_n^* = \sum_{j=1}^{n} Y_j, \tag{A.28}$$

to facilitate the analysis. It is easy to verify that

$$\frac{1}{T}\sum_{t=1}^{T-1} \sin[(t+1)\upsilon_1]\, S_T(\upsilon_1, t/T)\, X_T(t/T) = o_p(1) \tag{A.29}$$

if

$$\sum_{n=1}^{T-1} \sin[(n+1)\upsilon_1]\, S_n^*(\upsilon_1) X_n^* = o_p(n^3). \tag{A.30}$$

The same is true for the second term in (A.26). From (A.18) $ES_n^*(\upsilon_1)^2 \leq \sigma^2 n$. From Gradshteyn and Ryzhik (1980) equation 0.121.2, we have

$$EX_n^{*2} = E\left[\sum_{j=1}^{n} j\,\varepsilon_{t-j+1}\right]^2 = \sigma^2 \sum_{j=1}^{n} j^2 = \sigma^2 \frac{2n^3+3n^2+n}{6} \leq \sigma^2 n^3. \tag{A.31}$$

Given, $E|S_n^*(\upsilon_1)X_n^*| \leq \{ES_n^*(\upsilon_1)^2\}^{1/2}\{EX_n^{*2}\}^{1/2}$, we see that $E|S_n^*(\upsilon_1)X_n^*|$ is $O(n^2)$. Now let $n \geq m$ and consider

$$|S_n^*(\upsilon_1)X_n^* - S_m^*(\upsilon_1)X_m^*| \leq |S_n^*(\upsilon_1)||X_n^* - X_m^*| + |X_m^*||S_n^*(\upsilon_1) - S_m^*(\upsilon_1)|. \tag{A.32}$$

Clearly, $E|S_n^*(\upsilon_1)|^2 \leq \sigma^2 n$, and from (A.19), $E|S_n^*(\upsilon_1) - S_m^*(\upsilon_1)|^2 \leq \sigma^2(n-m)$. From (A.31) we have, $E|X_m^*|^2 \leq \sigma^3 m^3 \leq \sigma^2 n^3$. Finally, given Y_j from (A.23), yields,

$$E|X_n^* - X_m^*|^2 = E\left(\sum_{j=m+1}^{n} Y_j\right)^2 \tag{A.33}$$

$$= (n-m)^2 \sum_{j=1}^{m} \sigma^2 + \sigma^2 \sum_{j=1}^{n-m} j^2 \leq \sigma^2(n^3 - 2n^2 m + n^2 m) = \sigma^2\{n^2(n-m)\}. \tag{A.34}$$

Thus, from Chan and Wei's Theorem 2.1,

$$\underset{1\leq j \leq n}{sup}\left|\sum_{t=1}^{j} e^{it\theta} S_t^* X_t^*\right| = o_p(n^3) \tag{A.35}$$

which implies that the first term in (A.26) is $o_p(1)$. Following the same reasoning, the second term in (A.26) is also $o_p(1)$ and this proves Case 2 of the Lemma.

Case 3: Without loss of generality $\eta_1 = -1$ and $|\eta_j| < 1$, $j \neq 1$. Rearrange the polynomials in η^* such that $\eta_2 = \eta_j$. Then,

$$Z_{1t} = -\frac{1}{2\lambda_1}\frac{\partial \varepsilon_{t+1}}{\partial \eta_1} = \frac{\varepsilon_t}{(1+L)^2} = \sum_{j=1}^{t} (-1)^{j+1} j\, \varepsilon_{t-j+1}. \tag{A.36}$$

Define the process

$$X_T(t/T) = \begin{cases} \frac{1}{T}\frac{1}{\sqrt{T}\sigma}\sum_{j=1}^{t}(-1)^{j+1} j\,\varepsilon_{t-j+1} \text{ if } t \text{ is odd} \\ \frac{1}{T}\frac{1}{\sqrt{T}\sigma}\sum_{j=1}^{t}(-1)^{j} j\,\varepsilon_{t-j+1} \text{ if } t \text{ is even} \end{cases}. \tag{A.37}$$

Let $\omega_1 = (\upsilon_2 + \pi)$ and $\omega_2 = (\upsilon_2 - \pi)$. Noting that $T^{3/2}\sigma\, X_T(t/T)\, cos[(t+1)\pi] = Z_{1t}$, and defining Z_{2t} as in (A.10), we get

$$\begin{aligned} &\frac{2\sqrt{2}\sin(\upsilon_2)}{\sigma^2}\frac{1}{T^3}\sum_{t=1}^{T-1} Z_{1t}Z_{2t} \\ &= \frac{1}{T}\sum_{t=1}^{T-1}(\sin[(t+1)\omega_1] + \sin[(t+1)\omega_2])S_T(\upsilon_2, t/T)X_T(t/T) \\ &- \frac{1}{T}\sum_{t=1}^{T-1}(\cos[(t+1)\omega_1] + \sin[(t+1)\omega_2])S_T(\upsilon_2, t/T)\, X_T(t/T) \end{aligned} \tag{A.38}$$

Construct the variable $S_n^*(\upsilon_2)$ as above and the auxiliary variable X_n^* as

$$X_n^* = \begin{cases} \sum\limits_{j=1}^{n}(-1)^{j+1} j\varepsilon_{n-j+1} \text{ if } n \text{ is odd} \\ \sum\limits_{j=1}^{n}(-1)^{j} j\varepsilon_{n-j+1} \text{ if } n \text{ is even.} \end{cases}. \tag{A.39}$$

Using these definitions we get

$$\frac{1}{T}\sum_{t=1}^{T-1}\sin[(t+1)\omega_1]\, S_T(\upsilon_2, t/T)\, X_T(t/T) = o_p(1) \tag{A.40}$$

if $\sum_{n=1}^{T-1}\sin[(n+1)\omega_1]\, S_n^*(\upsilon_2)\, X_n^* = o_p(n^3)$. Again, $\{ES_n^*(\upsilon_2)^2\}^{1/2} \le \sigma\sqrt{n}$ and $E|S_n^*(\upsilon_2)X_n^*| \le \{ES_n^*(\upsilon_2)^2\}^{1/2}\{EX_n^{*2}\}^{1/2}$. Now, if n is odd, we have

$$E\left(X_n^*\right)^2 = E\left(\sum_{j=1}^{n}(-1)^{j+1} j\,\varepsilon_{t-j+1}\right)^2 = \sigma^2\sum_{j=1}^{n} j^2 \le \sigma^2 n^3, \tag{A.41}$$

and precisely the same reasoning holds if n is even. This implies that $E|S_n^*(\upsilon_2)X_n^*|$ is $O(n^2)$. We know that

$$|S_n^*(\upsilon_2)X_n^* - S_m^*(\upsilon_2)X_m^*| \le |S_n^*(\upsilon_2)||X_n^* - X_m^*| + |X_m^*||S_n^*(\upsilon_2) - S_m^*(\upsilon_2)| \tag{A.42}$$

where the bounds on $|S_n^*(\upsilon_2)|$ and $|S_n^*(\upsilon_2) - S_m^*(\upsilon_2)|$ were established in (A.18) and (A.19), respectively, and the bound on $|X_m^*|$ was established in the discussion above (A.33). Now, choosing $n \ge m$ for n odd and m even, gives

$$E|X_n^* - X_m^*|^2 = E[(n-m)\sum_{j=1}^{m}(-1)^{j+1}\varepsilon_j + \sum_{j=1}^{n-m}(-1)^{j+1} j\varepsilon_{n-j+1}]^2 \le \sigma^2[n^2(n-m)] \tag{A.43}$$

The result holds for any permutations of n and m. By Theorem 2.1 of Chan and Wei (1988), the first term in (A.38) is $o_p(1)$ and, by exactly the same reasoning, the remaining terms are also $o_p(1)$. This completes the proof of Case 3.

Case 4: Without loss of generality, let $\eta_1 = 1, \eta_k = 1$, with $|\eta_j| < 1$, for $j \neq 1, k$. Define the following elements:

$$Z_{1t} = -\frac{1}{2\lambda_1}\frac{\partial \varepsilon_{t+1}}{\partial \eta_1} = \frac{\varepsilon_t}{(1+L)^2} = \sum_{j=1}^{t}(-1)^{j+1} j\ \varepsilon_{t-j+1} \tag{A.44}$$

$$Z_{kt} = -\frac{1}{2\lambda_k}\frac{\partial \varepsilon_{t+1}}{\partial \eta_k} = \frac{\varepsilon_t}{(1-L)^2} = \sum_{j=1}^{t} j\ \varepsilon_{t-j+1} \tag{A.45}$$

$$X_{1t}^* = \begin{cases} \sum_{j=1}^{t}(-1)^{j+1} j \varepsilon_{t-j+1} \text{ if } t \text{ is odd} \\ \sum_{j=1}^{t}(-1)^{j} j \varepsilon_{t-j+1} \text{ if } t \text{ is even} \end{cases}. \tag{A.46}$$

Then,

$$\frac{4\lambda_1\lambda_k}{T^4}\sum_{t=1}^{T-1} Z_{1t}\, Z_{kt} = \frac{4\lambda_1\lambda_k}{T^4}\sum_{t=1}^{T-1}\cos[(t+1)\pi]\, X_{1t}^*\, X_{kt}^* \tag{A.47}$$

where X_{kt}^* is defined similarly to X_n^* in (A.28). This form allows us to directly apply Chan and Wei's Theorem 2.1 to show that the last expression is $o_p(1)$ if

$$\underset{1\leq j\leq n}{sup}\left|\sum_{t=1}^{j} e^{it\theta} X_{1t}^* X_{kt}^*\right| = o_p(n^4). \tag{A.48}$$

Now let X_{1n}^* and X_{kn}^* be defined equivalently to X_{1t}^* and X_{kt}^* with the sequence of partial sums running to n rather than t. From the definition of X_{1n}^* we have

$$EX_{1n}^{*2} = \sigma^2 \sum_{j=1}^{n} j^2 \leq \sigma^2 n^3. \tag{A.49}$$

From (A.31)

$$E|X_{1n}^* X_{kn}^*| \leq \{EX_{1n}^{*2}\}^{1/2}\{EX_{kn}^{*2}\}^{1/2} \leq \sigma^2\{n^{3/2}\}\{n^{3/2}\}. \tag{A.50}$$

Choose n and m as integers greater than 0 with $n \geq m$. Then,

$$|X_{1n}^* X_{kn}^* - X_{1m}^* X_{km}^*| \leq |X_{1n}^*||X_{kn}^* - X_{km}^*| + |X_{km}^*||X_{1n}^* - X_{1m}^*|. \tag{A.51}$$

From (A.31), and (A.49) we know that $E|X_{km}^*|^2$ and $E|X_{1n}^*|^2$ are both $O(n^3)$, while from (A.34) and (A.43), $E|X_{kn}^* - X_{km}^*|^2$ and $E|X_{1n}^* - X_{1m}^*|^2$ are $O(n^2(n-m))$. By Chan and Wei's Theorem 2.1, the sequence in

(A.48) is $o_p(n^4)$ and thus the sequence of partial sums in (A.47) is $o_p(1)$. This completes the proof of the Lemma.

Theorem 3 follows from the lemma, Chan and Wei's Theroem 2.2, and Chung's (1996a) Theorem 1. Note that for the jth element of η^*, we get

$$T^a(\hat{\eta}_j - \eta_j) = -[\frac{1}{T^{2a}}\sum_{t=1}^{T}(\frac{\partial \varepsilon_t}{\partial \eta_j})^2]^{-1}[\frac{1}{T^a}\sum_{t=1}^{T}\varepsilon_t\frac{\partial \varepsilon_t}{\partial \eta_j}] + o_p(1), \tag{A.52}$$

where $a = 1$ if $|\hat{\eta}_j| < 1$, and $a = 2$ if $|\hat{\eta}_j| = 1$. The theorem is complete as this is precisely the k-factor version of equation (A.5) in Chung (1996a).

References

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

Bierens, H.(2001), "Complex Unit Roots and Business Cycles: Are They Real?," *Econometric Theory*, 17, 962-983.

Caporale, M., and Gil-Alana, L. (2004), "Long-run and Cyclical Dynamics in the US Stock Market," Institute for Advanced Studies, Viena, Working Paper, Economic Series No. 155, 1-37.

Chan, N. H., and Wei, C. Z. (1988), "Limiting distributions of least squares estimates of unstable autoregressive processes," *The Annals of Statistics*, 16, 367-401.

Cheung, Y-W, and Diebold, F. X. (1994), "On maximum likelihood estimation of the differencing parameter of fractionally-integrated noise with unknown mean," *Journal of Econometrics*, 62, 301-16.

Chong, T.T.L (2000), "Estimating the differencing parameter via the partial autocorrelation function," *Journal of Econometrics*, 97, 365-81.

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

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

Chung, C.-F., and Baillie, R. T. (1993), "Small sample bias in conditional sum of squares estimators of fractionally integrated ARMA model," *Empirical Economics*, 18, 791-806.

Ferrara, L., and Guegan, D. (2001), "Forecasting with k-factor Gegenbauer processes: Theory and Applications," *Journal of Forecasting*, 20, 581-601.

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

Gradshyteyn, I.S. and Ryzhik, I. M. (1980), *Tables of integrals, series, and products* (4th edition), New York, NY: Academic Press.

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.

Gray, H.L., Zhang, N-F,, and Woodward, W. (1989), "On generalized fractional processes," *Journal of Time Series Analysis*, 10, 233-57.

Hauser, M.A, B.A. Potscher, and Reshenhofer, E. (1999), "Measuring persistence in aggregate output: ARMA Models, fractionally integrated ARMA models and nonparametric procedures," *Empirical Economics*, 24, 243-69.

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

Ooms, M., and Doornik, J. (1999), "Inference and forecasting for fractional autoregressive integrated moving average models, with an application to US and UK inflation," *Econometric Institute Report*, 9947/A, 1-33.

Ramachandran, R., and Beaumont, P. M. (2001), "Robust estimation of GARMA model parameters with an application to cointegration among interest rates of industrialized countries," *Computational Economics*,17, 179-201.

Smith, Jr A., Sowell, F. B., and Zin, S. E. (1997), "Fractional integration with drift: Estimation in small samples," *Empirical Economics*, 22, 103-116.

Tieslau, M.A., Schmidt, P., and Baillie, R. T. (1996), “A minimum distance estimator for long-memory processes,” *Journal of Econometrics*, 71, 249-64.

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.

Yajima, Y. (1991), “Asymptotic properties of the LSE in a regression model with long-memory stationary errors,” *The Annals of Statistics*, 19, 158-77.

Table 1: Simulation on the estimation of 2-factor GARMA processes

	η_1	η_2	λ_1	λ_2	μ
True	0.5	0	0.2	0.4	0
Mean Bias	0.0017	-0.0011	-0.0033	0.0029	0.0006
Med. Bias	0.0003	-0.0004	-0.0035	0.0041	0.0004
RMSE	0.0473	0.0155	0.0507	0.0463	0.0356
MNSE	0.0080	0.0133	0.0461	0.0396	N/A
MAD	0.0296	0.0091	0.0329	0.0395	0.0289
True	0.5	-0.5	0.2	0.4	0
Mean Bias	0.0007	-0.0000	-0.0002	0.0061	-0.0008
Med. Bias	0.0004	-0.0000	0.0004	0.0076	-0.0004
RMSE	0.0428	0.0136	0.0448	0.0484	0.0314
MNSE	0.0134	0.0069	0.0454	0.0457	N/A
MAD	0.0267	0.0081	0.0363	0.0388	0.0253
True	0	0.5	0.2	0.4	0
Mean Bias	-0.0028	0.0001	-0.0016	0.0025	-0.0008
Me. Bias	-0.0014	-0.0002	-0.0037	0.0033	0.0007
RMSE	0.0495	0.0146	0.0457	0.0442	0.0422
MNSE	0.0154	0.0070	0.0459	0.0399	N/A
MAD	0.0317	0.0086	0.0365	0.0348	0.0340
True	0	-0.5	0.2	0.4	0
Mean Bias	0.0025	0.0003	-0.0009	0.0046	-0.0006
Med. Bias	0.0005	0.0002	-0.0032	0.0073	-0.0015
RMSE	0.0500	0.0135	0.0458	0.0448	0.0274
MNSE	0.0155	0.0067	0.0460	0.0399	N/A
MAD	0.0318	0.0080	0.0366	0.0360	0.0219
True	-0.5	0.5	0.2	0.4	0
Mean Bias	-0.0018	0.0004	-0.0011	0.0025	0.0002
Med. Bias	-0.0004	0.0001	-0.0007	0.0023	0.0008
RMSE	0.0446	0.0140	0.0442	0.0487	0.0383
MNSE	0.0134	0.0070	0.0454	0.0457	N/A
MAD	0.0274	0.0080	0.0352	0.0388	0.0308
True	-0.5	0	0.2	0.4	0
Mean Bias	-0.0010	0.0012	0.0004	0.0039	0.0002
Med. Bias	0.0001	0.0003	0.0019	0.0055	0.0003
RMSE	0.0447	0.0160	0.0381	0.0490	0.0285
MNSE	0.0131	0.0080	0.0397	0.0461	N/A
MAD	0.0283	0.0098	0.0299	0.0395	0.0228

Table 2: Simulation on the estimation of 2-factor GARMA processes with nuisance parameters

	η_1	η_2	λ_1	λ_2	ϕ	θ	μ
True	0.5	-0.5	0.2	0.4	0.8	N/A	0
Mean Bias	-0.0010	-0.0003	-0.0015	-0.0025	-0.0114	-	-0.0101
Med. Bias	0.0003	-0.0002	-0.0016	0.0003	-0.0099	-	-0.0107
RMSE	0.0444	0.0138	0.0456	0.0547	0.0429	-	0.1507
MNSE	0.0134	0.0070	0.0462	0.0519	0.0409	-	N/A
MAD	0.0275	0.0082	0.0363	0.0441	0.0335	-	0.1212
True	0.5	-0.5	0.2	0.4	N/A	0.8	0
Mean Bias	0.0014	0.0003	-0.0011	0.0054	-	-0.0099	-0.0007
Med. Bias	0.0005	-0.0001	-0.0011	0.0062	-	-0.0086	-0.0023
RMSE	0.0488	0.0125	0.0480	0.0463	-	0.0459	0.0567
MNSE	0.0137	0.0068	0.0518	0.0463	-	0.0407	N/A
MAD	0.0295	0.0077	0.0389	0.0368	-	0.0362	0.0454
True	0.5	-0.5	0.2	0.4	-0.8	N/A	0
Mean Bias	0.0020	0.0016	0.0013	0.0078	0.0042	-	0.0004
Med. Bias	0.0009	0.0001	0.0011	0.0109	-0.0003	-	-0.0000
RMSE	0.0461	0.0344	0.0516	0.0513	0.0431	-	0.0177
MNSE	0.0137	0.0068	0.0517	0.0464	0.0401	-	N/A
MAD	0.0283	0.0089	0.0398	0.0381	0.0323	-	0.0141
True	0.5	-0.5	0.2	0.4	N/A	-0.8	0
Mean Bias	-0.0009	0.0009	-0.0026	0.0031	-	0.0167	-0.0001
Med. Bias	0.0000	0.0003	-0.0018	0.0058	-	0.0142	0.0001
RMSE	0.0526	0.0128	0.0457	0.0538	-	0.0539	0.0080
MNSE	0.0137	0.0070	0.0461	0.0521	-	0.0413	N/A
MAD	0.0299	0.0078	0.0364	0.0432	-	0.0413	0.0064
True	0.5	-0.5	0.2	0.4	0.8	0.8	0
Mean Bias	0.0002	0.0015	-0.0011	-0.0013	-0.0112	-0.0380	0.0045
Med. Bias	0.0003	0.0002	-0.0002	0.0016	-0.0041	-0.0281	0.0009
RMSE	0.0458	0.0221	0.0524	0.0578	0.0488	0.0739	0.2663
MNSE	0.0137	0.0069	0.0557	0.0548	0.0438	0.0467	N/A
MAD	0.0292	0.0083	0.0407	0.0442	0.0371	0.0546	0.2112

Table 3: Estimation of simulated ARFIMA/GARMA processes

	η_1	η_2	λ_1	λ_2	μ
True	1	0.75	0.2	0.3	0
Mean Bias	-0.0016	0.0026	-0.0127	0.0059	0.0417
Med. Bias	-0.0000	0.0010	-0.0121	0.0080	0.0612
RMSE	0.0062	0.0195	0.0345	0.0420	0.4345
MNSE	0.0005	0.0069	0.0288	0.0410	0.3679
MAD	0.0016	0.0109	0.0275	0.0338	0.3543
True	1	0.5	0.2	0.3	0
Mean Bias	-0.0023	0.0048	-0.0100	0.0041	0.0399
Med Bias	-0.0001	0.0014	-0.0085	0.0049	0.0539
RMSE	0.0085	0.0274	0.0283	0.0437	0.3486
MNSE	0.0006	0.0092	0.0245	0.0413	0.2812
MAD	0.0023	0.0150	0.0222	0.0348	0.2756
True	1	0	0.2	0.3	0
Mean Bias	-0.0040	0.0070	-0.0097	-0.0034	0.0023
Med. Bias	-0.0005	0.0019	-0.0096	-0.0016	0.0002
RMSE	0.0116	0.0393	0.0269	0.0498	0.2574
MNSE	0.0010	0.0108	0.0242	0.470	0.1859
MAD	0.0040	0.0176	0.0214	0.0387	0.2022
True	1	-0.5	0.2	0.3	0
Mean Bias	-0.0015	0.0041	-0.0079	-0.0032	-0.0074
Med. Bias	0.0000	0.0010	-0.0077	0.0002	-0.0203
RMSE	0.0085	0.0232	0.0316	0.0486	0.2797
MNSE	0.0004	0.0091	0.0287	0.0486	0.2365
MAD	0.0015	0.0132	0.0251	0.0386	0.2267

Table 4: Estimation of simulated ARFIMA/GARMA processes with single frequency models

	η_1	η_2	λ_1	λ_2	ϕ	θ	μ
True	0.5	0.0	0.2	0.4	N/A	N/A	0
Mean Bias	0.0099	0.0265	-0.0274	0.0487	-	-	0.0006
Med. Bias	0.0023	0.0087	-0.0276	0.0535	-	-	0.0007
RMSE	0.0603	0.0486	0.0468	0.0656	-	-	0.0356
MNSE	0.0141	0.0473	0.0257	0.0103	-	-	N/A
MAD	0.0319	0.0279	0.0379	0.0566	-	-	0.0285
True	0.5	-0.5	0.2	0.4	0.8	0.8	0
Mean Bias	-0.0103	0.0027	-0.0975	-0.1278	-0.0796	0.0466	0.0045
Med. Bias	0.0159	0.0007	0.0971	-0.1256	-0.0752	0.0514	0.0009
RMSE	0.0787	0.0155	0.1010	0.1376	0.0955	0.0620	0.2663
MNSE	0.0202	0.0074	0.0374	0.0429	0.0481	0.0316	N/A
MAD	0.0406	0.0084	0.0975	0.1278	0.0814	0.0542	0.2112
True	1.0	0.75	-0.2	0.3	N/A	N/A	0
Mean Bias	-0.0019	-0.0151	0.0631	-0.1266	-	-	-0.0006
Med. Bias	0.0000	-0.0024	0.0611	-0.1254	-	-	-0.0008
RMSE	0.0336	0.0463	0.0689	0.1325	-	-	0.0130
MNSE	0.0007	0.0085	0.0243	0.0302	-	-	N/A
MAD	0.0019	0.0143	0.0631	0.1275	-	-	0.0105